Search for jet-jet resonances in association with a leptonic $W$ decay at the ATLAS experiment

Candidato
Francesco Nuti

Relatrice
Dott. Chiara Roda
A Francesca e ai miei genitori
Contents

Introduction 1

1 Theoretical Overview 3
  1.1 Standard Model .......................................................... 3
  1.2 Diboson studies ............................................................ 8
    1.2.1 Diboson cross-section ............................................. 8
    1.2.2 Triple gauge couplings ......................................... 12
  1.3 An excess in jet-jet invariant mass ................................. 13

2 The Large Hadron Collider and the ATLAS experiment 17
  2.1 The Large Hadron Collider ........................................... 17
  2.2 The ATLAS detector .................................................... 19
    2.2.1 Inner Detector ..................................................... 20
  2.3 The Calorimeter System ............................................... 23
  2.4 Muon Spectrometer .................................................... 26
  2.5 Trigger ................................................................. 29

3 Reconstruction and identification of physics objects 31
  3.1 Electrons ................................................................. 31
    3.1.1 Electron reconstruction ......................................... 31
    3.1.2 Electron identification ........................................ 32
    3.1.3 Electron performances ......................................... 33
  3.2 Muons ................................................................... 35
    3.2.1 Muon reconstruction ............................................. 35
    3.2.2 Muon performances ............................................... 35
  3.3 Jets ...................................................................... 36
  3.4 Missing Transverse Energy ($E_T^{miss}$) ............................ 38

4 Event samples and selection 39
  4.1 Event samples .......................................................... 39
    4.1.1 Monte Carlo samples ............................................. 40
    4.1.2 Data sample ....................................................... 42
  4.2 Event selection ........................................................ 44
    4.2.1 Electron selection ............................................... 45
    4.2.2 Muon selection ................................................... 46
    4.2.3 $W \rightarrow l\nu$ reconstruction ............................... 47
Acknowledgements

This thesis is the fruit of the work that I developed in 2011 as member of the AT-LAS collaboration. ATLAS is about the most cutting edge experiment with the highest discovery potential nowadays, therefore I am pleased for having had the opportunity to take part in it. Beyond, I found a very nice environment that allowed me to do my work with passion and enthusiasm. I had also the pleasure to know a lot of people involved in the ATLAS experiment and many of them have influenced on this work and some on my mind. Thank you to all them and in a special way to the following people.

The biggest thank goes to Chiara Roda which is simply the ideal supervisor for me. You assisted me constantly and this helped me to work at best. At the same time, you let me develop my ideas about the analysis. It is evident the joy and the devotion with which you do your work and you transmit them to the people around you. Thanks also for not being only a supervisor but also a great friend whose opinion, about a lot of different topics, has been and will be useful to me to take my decisions.

I would like to thank Federico Bertolucci because he is a very good friend and I worked very well and enjoyed with him. I learnt a lot from him especially as regards programming and he has provided me the tools to perform my analysis. He is a very good guy nevertheless his only bad flaw: he want always pay the bill when you have lunch with him.

Many thanks go to all the members of the ATLAS Pisa group. I am grateful to Prof. Vincenzo Cavasinni and Prof. Tarcisio Del Prete that welcomed me heartily and that have been very willing during this year. I want to thank Paola Giannetti and Alberto Annoi that offered me useful suggestions about how to develop my analysis. I greet and thank Paolo Francavilla, Zinonos Zinonas, Michele Cascella, Francesco Crescioli, Vincent Giangiobbe, Daniele Puddu and Francesco Cervigni for the help received in using analysis and computing tools and above all for the nice time spent in Pisa and in Geneva with them.

Thanks a lot to the WW → lnujj group in which I am proud to be engaged and in a special way thank you to Giovanni Zevi Della Porta, Pierluigi Catastini and Alexander Solodkov who gave me many useful advices. I used some of the Giovanni’s results in this thesis, so thank you again Giovanni. Thanks to the Milano Grid technical staff to have been so efficient in solving computing problems.

Greetings also to the other people that I knew during this experience like Matteo Volpi, Andrea Dotti, Iacopo Vivarelli, Giulio Usai, Elisabetta Barberio, Peter Loch.

Finally, an immense thank you to Francesca and to my relatives who, although I have never explained to them why I spend my time in studying physics, support and love me. Thank you also to all my friends because there is at least one reason for which they deserve it.
Introduction

The Standard Model of particle physics is the theory which describes the strong and electroweak interactions of the elementary particles. Many experimental results have confirmed extensively the prediction made by this theory but some aspects of the theory like the origin of the masses have not found an experimental confirmation yet.

The Large Hadron Collider (LHC) offers the possibility to answer to the open questions and to probe the Standard Model at new kinematic regions unreachable before. The LHC is a proton-proton collider designed to reach the center-of-mass energy ($\sqrt{s}$) of 14 TeV and an instantaneous luminosity of $10^{34}$ cm$^{-2}$s$^{-1}$. The ATLAS experiment is one of the main experiments at the LHC and it has been built with very high performance detectors in order to sustain the huge collision rate and to collect and precisely reconstruct the interesting events for physics analysis. In the 2011 the ATLAS experiment has collected about 5 $fb^{-1}$ of proton-proton collisions at $\sqrt{s} = 7$ TeV.

The concurrent production of two vector bosons ($\gamma$, $W$, $Z$) is termed diboson production. The measurement of the diboson cross-section constitutes an interesting test of the Standard Model. In fact, the cross-section of these processes is sensitive to the coupling constants of the interaction vertexes among three vector bosons (Triple Gauge Couplings) which are expected from the gauge symmetry of the Standard Model. A measurement of diboson cross-section not compatible with the Standard Model predictions would imply the presence of new physics. The ATLAS collaboration is investigating all the diboson processes. This thesis describes the studies of the $WW$ and $WZ$ productions in the final state where one $W$ decays leptonically and the second $W$ or $Z$ boson decays hadronically (diboson semileptonic decay). These studies are not only important to measure the triple gauge couplings but also for the search of the Higgs decay in two $W$s where the $WW$ production constitutes a background.

There is an additional interest in studying the semileptonic $WW/WZ$ channel since the CDF collaboration, at the beginning of the 2011, observed an excess in the dijet invariant mass distribution in the mass range 120 – 160 GeV/$c^2$. The excess is measured in events with a jet pair produced in association with a $W$ which decays leptonically. This measurement constitutes one of the most significant deviations from the Standard Model predictions. Therefore it is important to verify this result with other studies. One has been done with the ATLAS experiment and it is described here in this thesis.

This thesis has two main topics: the studies of the semileptonic $WW/WZ$ production and the investigation of the excess observed by the CDF experiment. It is divided in 6 chapters.

Chapter 1 introduces the Standard Model and the mechanism used to justify the mass of the particles. Then it provides an overview of the diboson productions and of the origin
of the triple gauge couplings. The experimental results on diboson cross-sections and on the triple gauge couplings are also reported as well as those on the excess measured by the CDF experiment.

Chapter 2 describes briefly the LHC and the ATLAS experiment highlighting their performances.

The physics object used to reconstruct the $WW/WZ$ semileptonic decay and to study the CDF excess are electrons, muons, jets and missing transverse energy. Chapter 3 explains how these objects are reconstructed and identified combining the information of the ATLAS sub-detectors.

Chapter 4 is divided in two parts. The first one describes the data and Monte Carlo samples used in the two analysis; the second one explains the selection criteria common to the two analysis and the method used to extract the multijet QCD background from data.

Chapter 5 describes the analysis of the dijet mass excess produced in association with a $W$ leptonic decay. It gives details of the selection and of the statistical method used in the analysis and also of the systematics which affects the analysis. In this part of the analysis I have mainly contributed to the checks carried out on the cut-flow. I have also contributed to estimate the effect given by the jet energy scale uncertainty on the dijet mass distribution.

Chapter 6 collects the most relevant studies done by the author on the $WW/WZ$ production in the semileptonic channel. These studies are the first ones done in ATLAS for this type of measurement. Although the analysis has not reached the necessary sensitivity to reliably extract the signal the two described here are at the moment the best results obtained in term of statistical significance and signal to background ratio.

The analysis on the excess in the dijet mass distribution has been published in July 2011; the analysis on the $WW/WZ$ is still ongoing and it aims at measuring the diboson cross-section in the 2012.
Chapter 1

Theoretical Overview

Elementary particle physics investigates the nature of fundamental particles and their interactions. Four types of fundamental forces are known: strong, electromagnetic, weak and gravitational. At elementary level the gravitational interaction is negligible compared to the other forces. Indeed, the strength of the force between particles is represented by coupling constants which are ordered in magnitude as strong ($\alpha_s = 1$), electromagnetic ($\alpha = 1/137$), weak ($\alpha_W = 10^{-6}$) and gravitational ($\alpha_G = 10^{-39}$).

1.1 Standard Model

Physicists have developed many theories to describe the origin and properties of the four interactions. At present, the theory of reference is the Standard Model (SM) [1, 2, 3], a relativistic quantum field theory which describes all the known elementary particles and their strong, weak and electromagnetic interactions. It has been corroborated by predictions of new particles later discovered like the quarks bottom ($b$) and top ($t$) quarks and the neutrino $\tau$. However, some fundamental questions are left unanswered. One of the main problems in the SM theory is that it cannot justify the mass of particles, and in particular that of the gauge bosons. This problem is elegantly solved by the Higgs mechanism [4, 5, 6, 7], which hypothesizes the existence of a new particle, the Higgs boson ($H$).

The Higgs mechanism introduces in the Standard Model the mass terms in a way that preserves the structure of the gauge theory on which the SM is based. At present, the Higgs boson has not been identified and in case this will be confirmed by the analysis of the data collected at the LHC collider [8] this will lead to a failure of the Standard Model theory requiring to elaborate new ideas to solve the mass puzzle. The ATLAS [9] and CMS [10] experiments will cover the whole possible mass range therefore an answer on this problem is expected soon.

In terms of particle content, the Standard Model has 12 spin-1 bosons, mediators of strong, weak and electromagnetic forces, and 12 spin-1/2 fermions plus their corresponding antiparticles. Eight bosons are the mediators of the strong interaction (gluons $g$), three mediate the weak force ($W^-, W^+, Z$) and one the electromagnetic force ($\gamma$). Among the fermions there are the leptons: the neutrinos ($\nu_e, \nu_\mu, \nu_\tau$) which interact only weakly and the charged leptons ($e^-, \mu^-, \tau^-$) which can interact weakly and electromagnetically.
CHAPTER 1. THEORETICAL OVERVIEW

The remaining fermions are called quarks \((u, d, s, c, t, b)\) and are subject to all the three interactions. The properties of the SM particles are summarized in Fig. 1.1. In addition to charge, spin and mass, quarks and gluons are identified by an additional characteristic number called color. Because of confinement, in nature only uncolored particles are observed, so quarks combine together in hadrons where their color charge is neutralized.

![Fig. 1.1: Elementary particles in the Standard Model.](image)

In the SM the particles are described by fields and their free propagation by a Lagrangian terms of the form in Eq. 1.1, 1.2, 1.3 depending on their spin:

\[ \partial_\mu \Phi(x) \partial^\mu \Phi(x) \]  

(1.1)

where \(\Phi(x)\) is a spin-0 boson field;

\[ \bar{\psi}(x)i\gamma^\mu \partial_\mu \psi(x) \]  

(1.2)

where \(\psi(x)\) is a spin-1/2 boson field;

\[ -\frac{1}{4} F^a_{\mu\nu}(x) F^{a\mu\nu} \]  

(1.3)

where \(F^a_{\mu\nu}(x) = \partial_\mu W^a_\nu(x) - \partial_\nu W^a_\mu(x) - g \epsilon^{abc} W^b_\mu(x) W^c_\nu(x) \) and \(W^a_\mu(x)\) is a spin-1 boson field.

The interactions between fermions are derived from the invariance of the SM Lagrangian under local transformations of the gauge group:

\[ G_{SM} = SU(3)_C \otimes SU(2)_L \otimes U(1)_Y \]
The $SU(3)_C$ group invariance describes the strong interactions among quarks which is mediated by eight boson fields corresponding to the gluons. The local symmetry under $SU(2)_L \otimes U(1)_Y$ transformations generates four gauge fields responsible for the electroweak interaction. The physics fields corresponding to the $\gamma$, $Z$, $W^\pm$ bosons are linear combination of these gauge fields. Defining $W^\mu_a$ with $a = 1, 2, 3$ the gauge fields of the $SU(2)_L$ group, $B^\mu$ that of the $U(1)_Y$ group, $W^\pm_\mu$ and $Z_\mu$ the fields associated to the weak physics bosons and $A_\mu$ that representing the photon($\gamma$), the relationships between gauge fields and physics fields are reported in Eq. 1.4 and 1.5:

$$\begin{pmatrix}
  Z_\mu \\
  A_\mu 
\end{pmatrix} = \begin{pmatrix}
  \cos \theta_W & -\sin \theta_W \\
  \sin \theta_W & \cos \theta_W 
\end{pmatrix} \begin{pmatrix}
  W^3_\mu \\
  B_\mu 
\end{pmatrix}$$

$$W^\pm_\mu = \frac{1}{\sqrt{2}} (W^1_\mu \mp iW^2_\mu)$$

where $\theta_W = 0.23116 \pm 6.5 \times 10^{-5}$ [11] is the weak mixing angle or Weinberg angle.

To preserve the Lagrangian gauge invariance the fermions and the bosons acquire mass through the Higgs mechanism [4, 5, 6, 7]. In this mechanism the Lagrangian is modified introducing a complex scalar boson field $\Phi(x)$ subjected to the potential:

$$\frac{k^2}{4} \left( \Phi^\dagger(x)\Phi(x) - v^2 \right)^2$$

where $v$ is the non zero vacuum expectation value of the fields and $k$ is an adimensional coupling constant and they are connected with the Higgs boson mass ($M^2_H = k^2 v^2$). The requirement to preserve the local gauge invariance for this modified Lagrangian, implies the presence of interaction terms between the Higgs boson and the massive vector bosons ($W^\pm$, $Z$) and in addition the masses of the last ones. Furthermore it is possible to built with the left and right components of the fermion fields and with the $\Phi(x)$ field, the Lagrangian terms which describe the interactions between the Higgs boson and the fermions and the ones which assign them the mass.

The SM does not predict the value of Higgs mass ($M_H$), but relates it to the coupling constant of Higgs with fermions and bosons. The decay width in fermions is proportional to $M_H$ and to the mass of the fermion while the width in bosons to $M^3_H$ and to the squared mass of the boson. As a result, the fermion channels are favourite for low Higgs masses ($M_H < 135 \text{ GeV}/c^2$) while the boson channels for high $M_H$. Fig. 1.2 shows the Higgs branching ratios (BR) as a function of the Higgs mass. The dominant decay processes involve the couplings to the $W^\pm$ and $Z$ bosons or to the third generation of quarks and leptons. The decay topology depends on the mass of the Higgs, therefore the search for this particle requires a detector capable to cover all these signatures.

The first limits on the intervals of the allowed Higgs masses have been set by the Large Electron Positron (LEP) experiments [12] and subsequently by Tevatron experiments [13]. At LHC the protons are collided at the highest center-of-mass energy ($\sqrt{s}$) ever reached at an hadron collider ($\sqrt{s} = 7 \text{ TeV}$ at LHC against $\sqrt{s} = 1.96 \text{ TeV}$ at Tevatron). This raises the Higgs production cross-section of about two order of magnitude with respect to the one at Tevatron [14, 15]. This along with an instant luminosity of about $10^{34} \text{ cm}^{-2}\text{s}^{-1}$ by project, allows the ATLAS and CMS experiments at LHC to cover the remaining mass regions giving an answer to the Higgs existence.
CHAPTER 1. THEORETICAL OVERVIEW

At the time of writing, combining the studies on different Higgs decay channels in 4.7 fb$^{-1}$, the CMS experiment excludes at 95% of confidence level or higher the Higgs boson in the mass range $127 - 600$ GeV/c$^2$ [16]. Furthermore, CMS observes an excess compatible with a SM Higgs hypothesis in the vicinity of $124$ GeV/c$^2$, but with a global\(^1\) significance of less than 2 standard deviations (2σ) from the known backgrounds [16]. ATLAS in 4.9 fb$^{-1}$ of integrated luminosity has observed that data are compatible with the background only hypothesis and that the SM Higgs boson is excluded at 95% of confidence level or higher in the mass ranges $112.7 - 115.5$ GeV/c$^2$, $131 - 237$ GeV/c$^2$ and $251 - 453$ GeV/c$^2$ [17]. An excess of events is observed for a Higgs boson mass hypothesis close to $M_H = 126$ GeV/c$^2$ with global significance of 2.3 σ [17]. Fig. 1.3(a) shows the regions excluded experimentally at 95% of confidence level by CMS, LEP and Tevatron, while Fig. 1.3(b) shows the ATLAS results in the mass region $110 - 150$ GeV/c$^2$ where the most significant excess is observed by ATLAS. The results of ATLAS and CMS seem to show hint of a particle with mass around $125$ GeV/c$^2$. However further studies need to be done to confirm this possibility.

---

\(^1\)Global indicates that the significance has been obtained taking into account the “look-elsewhere” effect.
1.1. STANDARD MODEL

Fig. 1.3: Experimental limits from the CMS, LEP and Tevatron on Standard Model Higgs production in the mass range $100 - 600 \text{ GeV/c}^2$ (a). The hatched regions show the exclusions from the searches at the different colliders [16]. Experimental limits from ATLAS on Standard Model Higgs production in the mass range $110 - 150 \text{ GeV/c}^2$ (b). The most significant deviation is found at $M_H = 126 \text{ GeV/c}^2$ and has a global significance of $2.3 \sigma$ [17]. In both plots, the solid curve reflects the observed experimental limits for the production of Higgs of each possible mass value (horizontal axis). The region for which the solid curve dips below the horizontal line at the value of 1 is excluded with a 95% confidence level (CL). The dashed curve shows the expected limit in the absence of the Higgs boson, based on simulations. The green and yellow bands correspond (respectively) to 68%, and 95% confidence level regions from the expected limits.
1.2 Diboson studies

The study of the diboson production at the LHC provides an important test of the high energy behaviour of electroweak interactions. The vector boson self-couplings, resulting from the non-Abelian nature of the electroweak interaction, are fundamental predictions of the SM theory and their measurement provides useful information on the SM. Any theory predicting physics beyond the Standard Model while maintaining the Standard Model as a low-energy limit may introduce deviations in the gauge couplings at some high energy scale. Precise measurements of the couplings will not only provide stringent tests of the Standard Model, but will also probe for new physics in the bosonic sector. One of the aim of the study presented in this thesis is to measure the production of the $WW/WZ$ bosons in the semileptonic channel. A simple introduction to the $WW/WZ$ boson production and decay is given in section 1.2.1. In section 1.2.2 the theoretical ingredients of the Triple Gauge Coupling (TGC) are briefly discussed. The measurement of the TGCs is one of the most important information that can be obtained from the measurement of the $WW/WZ$ production.

1.2.1 Diboson cross-section

$WW$ and $WZ$ processes are a fundamental benchmarks of TGCs but their understanding is also important for the Higgs search in $WW$ final state where the diboson productions are irreducible backgrounds.

The leading order Feynman diagrams for the dominant $WW/WZ$ production mechanism at the LHC are shown in Fig. 1.4(a), for the $WW$ production, and in Fig. 1.5, for the $WZ$ production. These productions are characterized by a quark-antiquark initial state. Another mechanism of diboson production at the LHC is the gluon-gluon fusion as shown in Fig. 1.4(b). This process contributes few percent ($\sim 3\%$) to the total $WW/WZ$ production at $\sqrt{s} = 7$ TeV [18].

The NLO $W^+W^-$ and $W^\pm Z$ cross-sections at LHC ($\sqrt{s} = 7$ TeV) and at Tevatron ($\sqrt{s} = 1.96$ TeV) are summarized in Table 1.1. The production cross-section depends strongly by the center-of-mass energy of the $pp$ system. When two protons collides, the process which originates the $WW$ and $WZ$ is the interaction between two partons (hard-scattering), one for each proton. Kinematically, the $WW$ and $WZ$ can be produced if the center-of-mass energy of the two partons is at least the sum of the masses of the two bosons. The partons which compose the proton can be classified in valence quarks, gluons and sea quarks. The fraction of proton momentum carried by these partons is described by the parton distribution functions (PDF) of the protons. The PDF represent the probability densities to find a parton carrying a momentum fraction $X$ of the proton. If the

<table>
<thead>
<tr>
<th>Condition</th>
<th>$\sigma$ (pb)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$pp$ at $\sqrt{s} = 7$ TeV</td>
<td>$46 \pm 3$</td>
</tr>
<tr>
<td>$p\bar{p}$ at $\sqrt{s} = 1.96$ TeV</td>
<td>$11.7 \pm 0.7$</td>
</tr>
</tbody>
</table>

Table 1.1: NLO production cross-sections of the diboson processes $W^+W^-$ and $W^\pm Z$ for $pp$ collisions at $\sqrt{s} = 7$ TeV and $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV [19, 20].
1.2. DIBOSON STUDIES

Fig. 1.4: Feynman diagrams of $W^+W^-$ production in hadron collision: (a) the two leading order diagrams, one with a TGC vertex (right); (b) next to leading order diagrams, the one on the right with a TGC vertex.

Fig. 1.5: Feynman diagrams of $W^\pm Z$ production in hadronic collision at the leading order. The diagram on the right has a TGC vertex.
Fig. 1.6: Diboson cross-sections as a function of the center-of-mass energy ($\sqrt{s} = 1.96 \text{ TeV}$ for Tevatron and $\sqrt{s} = 7 \text{ TeV}$ for LHC) [21] for $pp$ and $p\bar{p}$ collisions.

$\sqrt{s}$ raises, partons with lower momentum fraction can give origin to the $WW/WZ$, so the probability of the two interacting partons to produce the diboson increases. In addition, at higher $\sqrt{s}$ the final state can be produced in more kinematic configurations, i.e. the phase space of the interaction products expands. This effect contributes to increase the cross-section too. The dependence of the diboson cross-section by $\sqrt{s}$ and by the colliding particles (proton-proton and proton-antiproton) is shown in Fig. 1.6. At low $\sqrt{s}$ the cross-section from $pp$ collision is lower than that from $p\bar{p}$ while at large $\sqrt{s}$ the two tend to the same value. This is due to the fact that only in $p\bar{p}$ collisions, valence quarks may produce the diboson and that those interactions between valence quarks are significant at low $\sqrt{s}$ while weigh less at high $\sqrt{s}^2$. The final state $WW/WZ$ has total charge equal to 0, ±1. In $pp$ collisions there is no combination of valence quarks which has charges 0, ±1 while in $p\bar{p}$ ones there are ($u\bar{u}$, $d\bar{d}$ with charge 0 and $ud$, $d\bar{u}$ with charge ±1). Therefore, due to the charge conservation, in $pp$ collisions the quarks interaction that produces the diboson is necessarily originated by at least a sea quark. This explains why the diboson cross-section in $pp$ interactions is smaller than that in $p\bar{p}$ ones.

The $W$ and $Z$ bosons decay with mean lifetime of $10^{-25} \text{ s}$, so they are not detected directly but through their products. The $W$ boson may decay in a lepton ($l$) and its flavor conjugate neutrino ($\nu_l$) or in two quarks $qq$; the $Z$ boson may decay in a couple fermion anti-fermion as two leptons ($ll$), two neutrinos ($\nu_l\bar{\nu}_l$) or two quarks ($qq$). The $W$ and $Z$ branching ratios (BR) are summarized in Table 1.2.

The decay products of $WW$ and $WZ$ are the combination of the two single decays.

\footnote{The valence quarks carry a momentum fraction $X$ on average larger than those of sea quarks and gluons. Therefore, at low $\sqrt{s}$ the valence quarks have more chances to reach the kinematic condition needed to produce the diboson than the sea quarks.}
1.2. DIBOSON STUDIES

Table 1.2: Branching ratios for $W$ and $Z$ bosons. The channels in leptons or quarks are summed on all flavors [11].

<table>
<thead>
<tr>
<th>Boson</th>
<th>$l\nu_l$</th>
<th>$qq'$</th>
<th>$ll$</th>
<th>$\nu_l\nu_l$</th>
<th>$qq$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$W$</td>
<td>32.4%</td>
<td>67.6%</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>$Z$</td>
<td>-</td>
<td>-</td>
<td>10.1%</td>
<td>20%</td>
<td>69.9%</td>
</tr>
</tbody>
</table>

Table 1.3: Production cross-sections of the $WW$ measured by the CDF [24], D0 [25], ATLAS [18] and CMS [26] experiments in the pure leptonic channel. The integrated luminosity used in the measurements is specified.

<table>
<thead>
<tr>
<th>Experiment</th>
<th>$\sigma_{WW}$ (pb)</th>
<th>Int. Lumi. ($fb^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CDF</td>
<td>$12.1 \pm 0.9(stat.)^{+1.4}_{-1.6}(syst.)$</td>
<td>3.6</td>
</tr>
<tr>
<td>D0</td>
<td>$11.5 \pm 2.1(stat + syst) \pm 0.7(lumi)$</td>
<td>1</td>
</tr>
<tr>
<td>ATLAS</td>
<td>$51.0 \pm 4.5(stat.) \pm 6.4(syst.) \pm 1.9(lumi.)$</td>
<td>0.83</td>
</tr>
<tr>
<td>CMS</td>
<td>$55.3 \pm 3.3(stat.) \pm 6.9(syst.) \pm 3.3(lumi.)$</td>
<td>1.1</td>
</tr>
</tbody>
</table>

The final states can be classified in three categories:

the pure hadronic decay channels these channels are characterized by two bosons decaying in four quarks whose hadronization produces four jets. It is the final state with the highest branching ratio (45.7% in $WW$ and 47.3% in $WZ$) but it is difficult to recognize at an hadron collider where the multijet QCD has a much higher production rate.

the pure leptonic decay channels in these cases the two bosons decay in four particles among charged leptons and neutrinos. These final states have the smallest branching ratios (10.5% in $WW$ and 9.8% in $WZ$) but clear signatures in particular in channels with at least two leptons that allow to highly reject the jet background. The pure leptonic decay is the first channel where Tevatron and LHC have measured the $WW$ and $WZ$ cross-sections. The cross-sections measured in this channel by the CDF [22], D0 [23], ATLAS and CMS experiments are summarized in Tables 1.3, 1.4.

the semileptonic channels these channels are those in which a boson decays in hadrons and the other one in leptons and neutrinos. The decay with a $W \rightarrow l\nu_l$ and the other $W/Z \rightarrow qq'$ has a branching ratio equal to 43.8% in $WW$ and 22.6% in $WZ$. The signature given by the leptonic $W$ decay (a lepton and missing transverse energy produced by neutrino) allows to remove most of multijet QCD background. Nevertheless a large irreducible background remains: the $W$ production in association with jets.

This thesis work is part of an analysis whose aim is the measurement of $WW/WZ$ cross-section through the isolation of the $WW/WZ$ signal in the semileptonic decay channel $WW/WZ \rightarrow l\nu_lqq'$ with $l = e, \mu$. The jet energy resolution does not allow to separate the $W$ and $Z$ resonances in the jet-jet invariant mass distribution. Therefore, in the semileptonic channel it is possible to measure only the sum of $WW$ and $WZ$ cross-sections.

At Tevatron the cross-section measurement in the semileptonic channel has been done by CDF [30] and D0 [31]. CDF measures $\sigma(WW/WZ) = 18.1 \pm 3.3(stat.) \pm 2.5(syst.)$ pb
with a significance of $5.24$ standard deviations while D0 measures $\sigma(WW/WZ) = 19.6^{+3.2}_{-3.0} \text{ pb}$ and rejects the background-only hypothesis at a level of $7.9$ standard deviations. Both measurements are consistent with the SM predictions. Fig. 1.7 shows the dijet invariant mass distributions measured by CDF in the electron (left) and muon (right) channels. The $WW/WZ$ cross-section has been obtained by fitting the data distribution with the background+signal mass shape as obtained from Monte Carlo (MC) and data-driven measurements.

At LHC the measurement of the $WW/WZ$ production in the semileptonic channel is much more challenging than at Tevatron. In fact the production cross-section for $W+(n \geq 2)\text{jets}$ times the BR($W \rightarrow l\nu$) where $l = e, \mu$ grows by approximately a factor of 20, from 11 pb measured by CDF [32] to 220 pb measured by ATLAS [33], while the theoretical cross-section for the $WW/WZ \rightarrow lvqq'$ production only grows by approximately a factor of 4, from 2 pb [31] at the Tevatron to 8 pb [18, 29] at the LHC$^3$. Chapter 6 presents a study of the selections that I have developed to optimize the signal to noise ratio and/or the signal significance in this channel.

### 1.2.2 Triple gauge couplings

An important test of SM is the measurement of the triple gauge boson couplings (TGC) which are the coupling constants associated with the interaction vertex among three vector bosons. The term that describes the boson propagation and the interaction in the SM Lagrangian is:

$$\mathcal{L}_{TGC} = g_{WWV} (i \left(1 + \frac{\Delta g_1^V}{1}ight) (W^+_{\mu} W^{\gamma \mu} V^\nu - W^+_{\mu} W^{\gamma \mu} V^\nu) + i \left(1 + \frac{\Delta k^V}{1}ight) W^+_{\mu} W^{\gamma \mu} V^\nu)$$

where $\Delta g_1^V = \frac{g_{WWV}^V}{g_{WWV}^V}$ and $\Delta k^V = \frac{g_{WWV}^V}{g_{WWV}^V}$ are the new gauge couplings. This is responsible of the interactions between three bosons fields is written in Eq. 1.7:

$$\mathcal{L}_{TGC} = g_{WWV} (i \left(1 + \frac{\Delta g_1^V}{1}ight) (W^+_{\mu} W^{\gamma \mu} V^\nu - W^+_{\mu} W^{\gamma \mu} V^\nu) + i \left(1 + \frac{\Delta k^V}{1}ight) W^+_{\mu} W^{\gamma \mu} V^\nu)$$

where $V$ stands for the $Z$ and the $A$ fields, $X_{\mu\nu} = \partial_{\mu} B_{\nu} - \partial_{\nu} B_{\mu}$ where $X = W, V$ and $g_{WWV}$ are the triple gauge boson couplings. The SM allows interactions only among $W^+ W^- \gamma$

$^3$The values reported refer to the single lepton channel (electron or muon).
1.3. AN EXCESS IN JET-JET INVARIANT MASS

The CDF experiment published in April 2011 an analysis which asserts the presence of an excess in the invariant mass distribution of jet pairs produced in association with a $W$ boson. The excess is observed in the region around $M_{jj} = 145 \text{ GeV/c}^2$ with a significance of 3 standard deviations [37]. The signature of the excess is the same as the one of the $WW/WZ$ analysis but the selection adopted is optimized to measure larger dijet invariant masses.

Some theories [34] predict anomalous TGCs as large as $10^{-3}$ to $10^{-4}$. Limits on TCGs has been set by LEP II [11] through $W^+W^-$ and $ZZ$ production and by Tevatron [35] using $W^\pm W^\mp$, $W^\pm Z$ and $W^\pm \gamma$. These results are fully compatible with the standard model. The sensitivity on the anomalous TGCs increases with the center-of-mass energy, therefore at LHC the limits on anomalous TGC could be enhanced. The ATLAS experiment had measured the TGCs from $WZ$ [29] and $ZZ$ [36] leptonic decays and no anomalies has been noticed.

1.3 An excess in jet-jet invariant mass

The excess is observed in the region around $M_{jj} = 145 \text{ GeV/c}^2$ with a significance of 3.2 standard deviations [37].
Fig. 1.8: Dijet invariant mass ($M_{jj}$) distribution measured by the CDF collaboration in events containing one leptonically decayed $W$ boson (a). The distribution is shown for the sum of electron and muon events. In the plot the contributions of each known process plus an additional hypothetical Gaussian component are shown as colored histograms. The data are shown as full circles. In plot (b), by subtraction, only the diboson ($WW$, $WZ$) and the hypothetical Gaussian contributions are shown. The band in the subtracted plot represents the sum of all background shape systematic uncertainties. The muon sample has $158 \pm 45$ excess events and the electron sample $240 \pm 55$. The peak of the Gaussian excess is at $147 \pm 4$ GeV/$c^2$ with an RMS of $14$ GeV/$c^2$. The $\chi^2$ is quoted for the fit region of $28 < M_{jj} < 200$ GeV/$c^2$ [38].

masses. This excess could be the evidence of a new resonance not predicted by the SM. In fact, in SM the only process which has a similar signature is the $WH$ but is predicted with a cross-section $\times \text{BR}(H \rightarrow bb)$ of about $12$ fb incompatible with that estimated for this resonance ($4$ pb).

Recently, CDF has updated the measurement increasing the statistics of the analysed data sample from $4.3$ fb$^{-1}$ to $7.3$ fb$^{-1}$ and the significance of the excess is increased to $4.1$ standard deviations [38]. The significance is obtained by modeling the excess as a Gaussian with a width compatible with the dijet invariant mass resolution, and performing a $\Delta \chi^2$ test for the presence of this additional component. The test return a $p$-value$^4$ of $1.9 \times 10^{-5}$ corresponding to a significance of $4.1$ standard deviations and a cross-section of the order of $4$ pb [38]. The peak of the Gaussian excess is at $147 \pm 4$ GeV/$c^2$ and the RMS is $14$ GeV/$c^2$. In Fig. 1.8 is reported the dijet invariant mass where the excess is measured.

The D0 experiment has performed the same study with a $4.3$ fb$^{-1}$ data sample and they find no evidence for a resonant excess in the dijet invariant mass distribution. They also state that the probability that the D0 data are consistent with the presence of a dijet

$^4$Given a null hypothesis and a set of data it is performed a statistical test to verify the agreement between data and the null hypothesis. The $p$-value is the probability of obtaining a statistic test at least as extreme as the one that was actually observed, assuming that the null hypothesis is true.
1.3. AN EXCESS IN JET-JET INVARIANT MASS

![Dijet invariant mass distribution measured by the D0 collaboration in events containing one leptonically decayed W boson (a). The distribution is shown for the sum of electron and muon events. In the plot the contributions of each known process are shown as colored histograms. The data are shown as full circles. In the right plot (b), by subtraction, only the diboson (WW, WZ) contribution is shown, along with the ±1 s.d. systematic uncertainty on all SM predictions. The χ² fit probability, P(χ²), is based on the residuals using data and MC statistical uncertainties. Also shown is the relative size and shape for a model with a Gaussian resonance with a production cross-section of 4 pb at M_{jj} = 145 GeV/c² [39].](image)

The ATLAS collaboration has recently published the result of a search of jet pairs produced in association of a leptonically decaying W boson [40]. In this analysis the same CDF selection used to observe the excess is applied to a data sample of about 1 fb⁻¹. I have taken part to this analysis and the results of this study are discussed in the chapter 5.

resonance with a 4 pb production cross-section at 145 GeV/c² is 8 × 10⁻⁶ [39]. The dijet invariant mass distribution obtained by D0 collaboration is shown in Fig. 1.9. The results of D0 has placed many doubts on the physical nature of the CDF excess, however the question is still open.
Chapter 2

The Large Hadron Collider and the ATLAS experiment

2.1 The Large Hadron Collider

The Large Hadron Collider (LHC) is designed to produce proton-proton collision at a center-of-mass energy of 14 TeV and at a luminosity up to $10^{34} \text{cm}^{-2}\text{s}^{-1}$. The accelerator can also collide heavy ions. LHC is installed in a underground ring, originally built for Large Electron-Positron collider at the CERN laboratory. It has a circumference of 27 km along which 1232 superconducting dipole magnets bend the proton beams. There are four interaction points where the experiments ATLAS, CMS, LHCb \cite{41} and ALICE \cite{42} are located as shown in Fig. 2.1.

The LHC has successfully accelerated the first beam the 10th September of 2008. Unfortunately, nine days after the first beam circulation, a superconducting magnets had a fault which damaged an extended area of the collider. After the reparation, at the beginning of 2010, LHC started to collide protons reaching in a few days stable beam conditions at a center-of-mass energy of 7 TeV. The collision rate has undergone in 2010/2011 data taking a steady rise thanks to the work done on the beam conditions.

The collision rate is measured using the instantaneous luminosity $\mathcal{L}$ which is a measurement of the density of interaction centers in the time unit and it has the dimension of $\left[ L \right]^{-2}\left[ T \right]^{-1}$. In a collider $\mathcal{L}$ is a function of the number of bunches ($b$), their dimension ($\sigma_X$ and $\sigma_Y$) transverse to the beam direction, the number of particles in each of them ($n$) and the revolution frequency ($\nu$):

$$\mathcal{L} = \frac{n^2 b \nu}{4\pi \sigma_X \sigma_Y} \quad (2.1)$$

The instantaneous luminosity times the cross-section of a process provides the process production rate. Consequently, a necessary condition to measure rare processes is the high instantaneous luminosity. However a large instantaneous luminosity comes at the price of a high pile-up. The pile-up is normally divided in in-time pile-up and out-of-time pile-up. The in-time pile-up is the superposition in the detector of signal produced from multiple

\footnote{Conseil Européen pour la Recherche Nucléaire.}
interactions\(^2\) between protons belonging to the bunch crossing where the triggering interaction occurred. The *in-time* pile-up grows in a collider when the transversal dimension of bunches decreases or the number of particle in the bunch increases. The *out-of-time* pile-up is caused by the proton collisions that occurred in bunch crossings before or after the one that caused the trigger. The effect caused by the Pile-up is directly connected to the instantaneous luminosity; for the *out-of-time* pile-up the integration time (or memory) of the detectors is also important. The distortions generated by the pile-up on the physics are corrected at various level of the analysis using mostly data-driven techniques.

The integrated luminosity\(^3\) collected by the ATLAS experiment during the 2011 and the day by day peak luminosity are shown in Fig. 2.2. The data used in this thesis has been collected from March to July 2011. For the analysis of the resonance observed by CDF the data sample amounts to 1.02 \(fb^{-1}\) while for the diboson analysis to 1.33 \(fb^{-1}\).

During these periods an instantaneous luminosity of about \(10^{32} - 10^{33} \text{cm}^{-2}\text{s}^{-1}\) (depending on the exact period of data taking) was reached colliding \(10^2 - 10^3\) bunches per beam, assembled in consecutive bunches with a time distance of 50 ns (\(\simeq 15\) m). Since the distance between bunches was steady with the increase of the number of bunches, the *Out of time* pile-up did not changed significantly. Each bunch had a length of \(\simeq 6\) cm (\(\simeq 0.2\) ns) and radial dimensions of \(\sigma_X \simeq \sigma_Y \simeq 2.7\) \(\mu m\), and consisted of \(10^{11}\) protons [44].

---

\(^2\)In the first half of the 2011 at LHC the mean number \(< \mu >\) of interactions in one bunch crossing is \(\sim 6\) [43].

\(^3\)The integrated luminosity is the integral of the instantaneous luminosity with respect to the data acquisition time.
2.2. THE ATLAS DETECTOR

ATLAS\(^4\) is a multi purpose high energy physics detector designed to measure the largest possible variety of physics processes that might contain indication of new physics. One key point that the ATLAS experiment aims at clarifying is the existence of the Higgs boson, a basic ingredient of the successful Standard Model Theory [1]. For this goal, ATLAS is equipped with high-resolution complementary detectors that allow the identification and reconstruction of electrons, muons, jets and missing transverse energy.

The ATLAS detector has a barrel shape with a diameter of 25 m and a length of 44 m (Fig. 2.3). The beam direction is the axis of the barrel and defines the Z-axis of a right-handed coordinate system whose origin is the collision point. The X-axis is directed from the origin to the center of the ring, and Y-axis points upwards as shown in Fig. 2.1. It is also useful to define the coordinates \( \phi \) and \( \eta \) which are respectively the azimuthal angle in \( X \times Y \) plane and the pseudorapidity \( \eta = -\ln [\tan (\theta/2)] \) where \( \theta \) is the polar angle. Furthermore, in \( \eta \times \phi \) space, \( \Delta R \) is defined as the angular distance between two points \( (\eta_1, \phi_1), (\eta_2, \phi_2) \):

\[
\Delta R = \sqrt{\Delta \eta^2 + \Delta \phi^2}
\]

where \( \Delta \eta = \eta_2 - \eta_1 \) and \( \Delta \phi = \phi_2 - \phi_1 \in (-\pi, \pi) \). Along the Z-axis the detector is divided in five parts: a central part called barrel region, the two lateral parts called end-cap and the two parts close to the beam axis called forward regions.

\(^4\)ATLAS is the acronym for A Toroidal LHC ApparatuS.
CHAPTER 2. THE LARGE HADRON COLLIDER AND THE ATLAS EXPERIMENT

Fig. 2.3: Section of the ATLAS detector.

The ATLAS detector [9] consists of several sub-detector systems and magnets disposed around the interaction point. From inside to outside, there are the tracking detectors, the superconducting solenoid, the calorimeters and the muon chambers surrounded by the toroid magnets. The analysis described in this thesis is based on signatures from electrons, muons, jets and missing transverse momentum therefore needs information from all the sub-systems. A brief description of each sub-detector system is provided in the following sections.

2.2.1 Inner Detector

The Inner Detector (ID) consists of three tracking devices placed in the most internal part of the ATLAS: the Pixel Detector (PIXEL), the Silicon Microstrip Tracker (SCT), the Transition Radiation Tracker (TRT). The disposition of the PIXEL, the SCT and the TRT detectors is shown in Fig. 2.4 (a) along with a view of the radial disposition of these detectors in the barrel region Fig. 2.4 (b). The PIXEL is the most internal detector followed by the SCT and then by the TRT. The Inner Detector acceptance in $\eta$ is limited to $|\eta| < 2.5$ while the $\phi$ coverage is complete.

The Inner Detector is placed inside a 2 T magnetic field provided by a superconducting solenoid which bends the trajectory of the charged particles inside the ID. The main purpose of the Inner Detector is to provide the space point measurements to reconstruct, with a very high precision, the particle trajectories from which the transverse momentum of the particle is extrapolated. For this purpose, the alignment of all the sub-detectors has to be known accurately too. Fig. 2.5 summarizes the resolutions and alignment uncertainties of the PIXEL, SCT, TRT detectors. Combining the information of all hits, the ID measures the transverse momentum ($p_T$) of charged particles with a relative uncertainty $\frac{\sigma_{p_T}}{p_T} = 0.05\% p_T \pm 1\%$.

The ID is important in particular for the reconstruction of muons and electrons. In fact, the electron identification is based on ID and calorimeter parameters while the muon...
2.2. THE ATLAS DETECTOR

Fig. 2.4: Cutaway view of the Inner Detector (a). Zoomed view of the barrel region of the ID (b). The distances of the PIXEL, the SCT and the TRT from the beams are provided.

<table>
<thead>
<tr>
<th>Item</th>
<th>Intrinsic accuracy (μm)</th>
<th>Alignment tolerances (μm)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Radial (R)</td>
</tr>
<tr>
<td><strong>Pixel</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Layer-0</td>
<td>10 (R-φ) 115 (z)</td>
<td>10</td>
</tr>
<tr>
<td>Layer-1 and -2</td>
<td>10 (R-φ) 115 (z)</td>
<td>20</td>
</tr>
<tr>
<td>Disks</td>
<td>10 (R-φ) 115 (R)</td>
<td>20</td>
</tr>
<tr>
<td><strong>SCT</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Barrel</td>
<td>17 (R-φ) 580 (z)</td>
<td>100</td>
</tr>
<tr>
<td>Disks</td>
<td>17 (R-φ) 580 (R)</td>
<td>50</td>
</tr>
<tr>
<td><strong>TRT</strong></td>
<td>130</td>
<td></td>
</tr>
</tbody>
</table>

1 Arises from the 40 mrad stereo angle between back-to-back sensors on the SCT modules with axial (barrel) or radial (end-cap) alignment of one side of the structure. The result is pitch-dependent for end-cap SCT modules.

2 The quoted alignment accuracy is related to the TRT drift-time accuracy.

Fig. 2.5: Intrinsic measurement accuracies and mechanical alignment tolerances for the inner detector sub-systems, as defined by the performance requirements of the ATLAS experiment. The numbers in the table correspond to the single-module accuracy for the pixels, to the effective single-module accuracy for the SCT and to the drift-time accuracy of a single straw for the TRT [9].
CHAPTER 2. THE LARGE HADRON COLLIDER AND THE ATLAS EXPERIMENT

Identification combines information from the ID and muon chambers. Due to the closeness to interaction point, the ID is crossed by the highest radiation. Furthermore the energy lost by particles in its percolation affects the measurement made by the outer detectors. Therefore the three sub-detectors are designed to have:

- high resistance to radiation;
- high granularity and low time occupancies;
- low thickness to reduce energy loss and multiple scattering.

PIXEL

The PIXEL consists of barrel layers and discs containing pixel detectors. It is composed of three cylindrical layers in the barrel region and three disks in each end-cap region. Since its layers are close up to 50.5 mm to the interaction point, it is subject to a large radiation density. Therefore in order to distinguish tracks of different particles it has the highest granularity \((\sigma_{R-\phi} \times \sigma_Z = 10 \times 115 \text{ nm}^2)\) given by 80 million pixels with a size of 400 \(\mu\text{m}\) in \(Z\) direction and 50 \(\mu\text{m}\) in \(R-\phi\) direction. One hit per layer/disk is expected for a charged particle with \(|\eta| < 2.5\). The hits are used to reconstruct the particle tracks. The first layer of PIXEL (\(b\)-layer) is placed just outside the beam pipe and due to its closeness to interaction point, provides a high resolution measurement of the vertex positions and of the track impact parameters.

SCT

The Silicon Microstrip Tracker (SCT) is built of four cylindrical layers and nine disks at each end-cap. Each layer is made of two sub-layers of silicon strips: one with strips along \(Z\) direction; the strips of the other one form an angle of 40 mrad with the \(Z\) direction. The two layers form a grid that allows the detection of particle position in \((R-\phi) \times (Z)\) space with a resolution of \(17 \times 580 \text{ nm}^2\). In total there are 6.3 million of 6.4 cm-long silicon strips with a pitch of 80 nm. The SCT has the role to supply four additional hits used to reconstruct the tracks of charged particles.

TRT

The Transition Radiation Tracker (TRT) uses 351000 Gas-filled (70% \(Xe\), 27% \(CO_2\) and 3% \(O_2\)) straws to detect particles which leave in them a wake of ionized gas. In the barrel region, there are 73 planes of straws parallel to the beam axis and in the end-cap region, straws constitute the radii of 160 wheels. Each straw has a diameter of 4 \(mm\) and contains a tungsten goldplated wire which works as anode. When a charged particle across the straw, ionizes the gas whose electrons are collected by the wire. The time of collection (drift time) is related to the distance of the particle track from the wire. With this method the resolution on the track position in the TRT is 130 \(\mu\text{m}\). Typically, a particle hits in 36 straws which provide the \(R\) and \(\phi\) coordinates of the track. No information on the \(Z\) direction is given, however the hits contributes to increase the resolution on the \(p_T\) of the particle.
2.3. THE CALORIMETER SYSTEM

The TRT is important in the electron identification. In fact, the passage of the electron through the straws and the gas inside them produces the so-called transition radiation. The photons produced in the transition interact with the Xe producing additional charge. As a result, the signal collected during the transit of an electron is larger than that produced by other particles allowing the identification of the electron.

2.3 The Calorimeter System

The solenoid is surrounded by the ATLAS calorimeter system that extends over the range $|\eta| < 4.9$. Fig. 2.6 shows a cutaway view of the ATLAS calorimeter. The electromagnetic calorimeter placed around the solenoid measures the energy and the direction of electrons and photons. The electromagnetic calorimeter is enclosed in the hadronic calorimeter composed by the Tile Calorimeter, the Hadronic Endcap Calorimeter and the Forward Calorimeter which are used to detect hadrons.

Primarily, calorimeters have to measure the particle energies. This is achieved by stopping the particles through interactions with matter and counting the detectable energy emerging from these interactions.

Energetic photons and electrons in matter form electromagnetic showers through chains of bremsstrahlung ($e \rightarrow e\gamma$) and $e^+e^-$ pair production ($\gamma \rightarrow e^+e^-$) processes. The loss of energy of the electron due to bremsstrahlung interactions per length unit and per material density unit follows a decreasing exponential function with a characteristic length called radiation length $X_0$. The mean free path covered by a photon in matter before the production of a pair $e^+e^-$, multiplied for the matter density, is $\frac{7}{5}X_0$. Thus $X_0$ governs both the interactions and determines the longitudinal development of the electromagnetic shower. $X_0$ depends on the atomic number $Z$ and mass number of the material and is usually measured in $g/cm^2$ to be independent from the material density.

As well as the energetic electrons, hadrons lose energy in matter generating cascades of particles with lower energy. However, since hadrons are subject also to the strong interaction the characteristic length for a hadron shower is governed by the interaction length ($\lambda$) which is the mean distance travelled by a hadron before undergoing an inelastic nuclear interaction.

The calorimeters need to be enough thick to contain all the shower of the incident particles, otherwise some of the energy of the particles is undetectable. The depth of the ATLAS calorimeters expressed in $X_0$ (a) and $\lambda$ units (b) is shown in Fig. 2.7.

The ATLAS calorimeters are sampling devices, i.e., they are composed of alternated layers of active material, where the energy from the interactions is measured, and passive material, where the particles interact and slow down more but the energy is not detected.

The signal measured with the calorimeter depends on the particle which originates the shower. Hadrons in matter can lose their energy through nondetectable processes like nuclear excitation, nuclear breakup or neutron, $\mu$, $\nu$ leaks. Therefore, usually the signal produced by a hadron in the calorimeter is lower than that one of an electron or a photon of the same energy. For this reason, the calorimeter signal needs to be calibrated differently for hadrons and electrons to obtain the energy of the interacting particle. By default, the calorimeter is calibrated to measure the energy of electrons and photons, and an offline calibration...
correction is then applied on the measured hadron energy to rescale it to the appropriate value.

**LAr Electromagnetic calorimeter**

The EM calorimeter is a sampling calorimeter which uses liquid argon (LAr) as active medium and lead as passive material. The use of the Liquid Argon offers the high radiation resistance needed in this region. Accordion-shaped kapton electrodes collect free charges in LAr. The EM calorimeter is composed by a barrel covering $|\eta| < 1.475$ and two coaxial wheels at each *end-cap* covering $1.375 < |\eta| < 3.2$. All these parts are segmented in depth. In particular for $|\eta| < 2.5$, the same region covered by the ID, the EM calorimeter is divided in three longitudinal layers. The inner layer has high $\eta$ granularity and is used to distinguish isolated photons from couples of photons coming from $\pi^0$ decays. The following layer is 16 radiation length long and collects the largest fraction of the energy of the electromagnetic showers. The outer layer collects the tail of the shower and is therefore less segmented. In $|\eta| < 1.8$ an additional 1.1 cm-thick layer, called “presampler”, is placed between solenoid and calorimeter. It is designed to recover the energy loss in the passive material inside calorimeter that varies from 2 up to 6 $X_0$. Details on the calorimeter size and granularity are given in Fig. 2.8. The overall thickness of EM calorimeter is 24 $X_0$ at $\eta = 0$ which allows to stop the most of the electrons and photons. The energy resolution for electrons is $\frac{\sigma_E}{E} = 10\% + 0.7\%$.

**Hadronic calorimeter**

The *barrel* part of the hadronic calorimeter is a sampling calorimeter made of steel and plastic scintillators (TileCal). The light produced in each scintillator tile is coupled by wavelength shifting fibers to read-out photomultipliers. TileCal covers $|\eta| < 1.7$ and...
2.3. THE CALORIMETER SYSTEM

Fig. 2.7: (a): For the barrel (left) and end-cap (right) regions, cumulative amount of material, in units of $X_0$, as a function of $|\eta|$ in front of the electromagnetic calorimeters (yellow), and in the three layers of the electromagnetic calorimeters themselves (green, purple, cyan). (b): Cumulative amount of material, in units of $\lambda$, as a function of $|\eta|$, in front of the electromagnetic calorimeters (light grey), in the electromagnetic calorimeters themselves (EM calo), in each hadronic layer (Tile0, HEC0, FCal0), and the total amount at the end of the active calorimetry. Also shown for completeness is the total amount of material in front of the first active layer of the muon spectrometer (cyan).
Fig. 2.8: Sketch of a barrel module of the EM calorimeter segmented in depth in three layers whose length in $X_0$ is reported. The granularity in $\eta$ and $\phi$ of the cells of each of the three layers and of the trigger towers is shown. It is also depicted the geometry of electrodes in the $\phi \times R$ plane.

radially, is segmented in three layers corresponding to $9.7 \lambda$. The Hadronic Endcap Calorimeter (HEC) covers $1.5 < |\eta| < 3.2$ and is a LAr-copper-sampling calorimeter. The design energy resolution for hadrons is $\frac{\sigma_E}{E} = 50\% \oplus 3\%$. The choice of a different technique for the end-cap and for the barrel regions is based on the different amount of radiation which the two detectors have to sustain. The TileCal technology is cheaper than the HEC one but provides similar performances in term of energy resolution. Jets, emerging from the fragmentations of the partons, are measured assembling information of the electromagnetic calorimeter and hadronic calorimeter systems.

Forward calorimeter

The Forward Calorimeter (FCal) is made of LAr as active medium interposed to an internal layer of copper and two layers of tungsten. The choice of material is made to be heat dissipating and hard against radiation which is higher in region $3.1 < |\eta| < 4.9$ covered by FCal. FCal measures the energy of forward jets and is very important for the measurement of transverse missing energy ($E_T^{miss}$). In fact, the larger is the calorimeters coverage, the better is the resolution on $E_T^{miss}$. The resolution on jet energy is $\frac{\sigma_E}{E} = 100\% \oplus 10\%$.

2.4 Muon Spectrometer

Muons with energy lower than 100 GeV lose mainly energy by ionization in matter. The energy loss per length unit due to this process is lower than that due to chains of Bremsstrahlung and pair production interactions (radiative losses). However in high-Z materials like those of the ATLAS calorimeter the radiative effects can be already signif-
2.4. MUON SPECTROMETER

The energy lost by a 10 GeV muon through the ID and in the calorimeters is about 3 GeV \cite{46}. Therefore, while electrons with tens of GeV of energy, like those used in this thesis, are stopped in the ATLAS calorimeter system, muons of the same energy lose only a fraction of energy and emerge from the calorimeter reaching the muon spectrometer where the muon momentum is measured. The momentum and direction of these muons is then measured by a system of magnets and detectors placed outside the calorimeters. The momentum is corrected for the energy lost in the ID and the calorimeter which is estimated from parametrizations or from a measurement of the energy deposited in the calorimeters.

As shown in Fig. 2.9 there are three toroidal superconducting magnets: one has a barrel shape and covers $|\eta| < 1.4$; other two have a wheel shape and cover $1.6 < |\eta| < 2.7$. The magnetic toroidal field bends muon tracks which are detected by Monitored Drift Tubes (MDT) in the \textit{barrel} region and by Cathode Strip Chambers (CSC) in the \textit{end-cap} region.

The MDTs are tubes filled with gas which collect avalanche amplified charges produced in gas ionization committed by muons. There are 3 radial MDT chambers in the \textit{barrel} region, 4 wheels chambers in the \textit{end-cap} region. Each chamber consists of three to eight layers of drift tubes that measure the coordinates in the bending plane ($R - Z$). The other coordinate ($\phi$) is provided by other detectors (RPC and TGC) used also as trigger for muons.

The CSC chambers are multi-wire proportional chambers and have segmented cathodes which provides the \( \phi \) coordinate. The signal induced on the wires by charged particles gives information about their $R - Z$ coordinates. There are two disks of CSC chambers in each \textit{end-cap} made by four layers of wires. The CSC technology has higher rate capability and time resolution than the MDT one in order to cope with the particle rate in the \textit{end-cap} region. Both MDT and CSC have a total momentum resolution of $\frac{\Delta p}{p} < 10\%$ at $p = 1 \text{ TeV}$. To provide with good accuracy the muon momentum, the position of the chambers is controlled by 12000 precision-mounted alignment sensors and the magnetic field is measured by 1800 Hall sensors.

The detection in $|\eta| < 2.4$ region is completed by faster but lower resolution detectors. These are the Resistive Plate Chambers (RPC) and the Thin Gap Chambers (TGC) designed to provide a fast trigger when crossed by an energetic muon. Details of the four muon detectors like the spatial and time resolutions are listed in Fig. 2.10.
Fig. 2.9: Section of the ATLAS muon spectrometer.

<table>
<thead>
<tr>
<th>Type</th>
<th>Function</th>
<th>$z/R$</th>
<th>$\phi$</th>
<th>$t$</th>
<th>barrel</th>
<th>end-cap</th>
<th>chambers</th>
<th>channels</th>
</tr>
</thead>
<tbody>
<tr>
<td>MDT</td>
<td>tracking</td>
<td>35 $\mu$m (c)</td>
<td>—</td>
<td>—</td>
<td>20</td>
<td>20</td>
<td>1088 (1150)</td>
<td>339k (354k)</td>
</tr>
<tr>
<td>CSC</td>
<td>tracking</td>
<td>40 $\mu$m (c)</td>
<td>5 mm</td>
<td>7 ns</td>
<td>—</td>
<td>4</td>
<td>32</td>
<td>30.7k</td>
</tr>
<tr>
<td>RPC</td>
<td>trigger</td>
<td>10 mm (c)</td>
<td>10 mm</td>
<td>1.5 ns</td>
<td>6</td>
<td>—</td>
<td>544 (606)</td>
<td>359k (373k)</td>
</tr>
<tr>
<td>TGC</td>
<td>trigger</td>
<td>2-6 mm (c)</td>
<td>3-7 mm</td>
<td>4 ns</td>
<td>—</td>
<td>9</td>
<td>3588</td>
<td>318k</td>
</tr>
</tbody>
</table>

Fig. 2.10: Parameters of the four sub-systems of the muon detector. The quoted spatial resolution (columns 3, 4) does not include chamber-alignment uncertainties. Column 5 lists the intrinsic time resolution of each chamber type, to which contributions from signal-propagation and electronics contributions need to be added. Numbers in brackets refer to the complete detector configuration as planned for 2009 [9].
2.5 Trigger

The ATLAS trigger and data acquisition (TDAQ) is the system built to select and store the physically interesting events. The decision to save or discard events is taken by three consecutive level of triggers: the Level 1 (L1), the Level 2 (L2), and the event filter (EF). Each trigger level refines the decision of the previous one, accessing to more detailed information from the detector. The time employed to process the event increases with the amount of information used and the three level structure is designed to cope with the collision rate. Only the events which passes all the three levels are recorded.

The triggers used in this thesis search for an electron with transverse energy $E_T > 20 \text{ GeV}$ (EF\_e20\_medium) or a muon with transverse momentum $p_T > 18 \text{ GeV}/c$ (EF\_mu18\_MG OR EF\_mu40\_MSonly\_barrel).

The L1 trigger has to take fast decision in order to process all the events at a production rate of 40 $\text{MHz}$. It employs less than 2.5 $\mu$s to take a decision and meanwhile data are buffered in memories located within the detector front-end electronics. The rate of events passing the L1 trigger is 75 $\text{KHz}$. The L1 trigger algorithms analyse only a limited number of information from the detectors to find high transverse-momentum muons, electrons, photons, jets, and $\tau$ as well as large missing and total transverse energy in the event.

The L1 electron triggers are based on the signal information from about 7000 analogue EM trigger towers of reduced granularity ($\Delta \eta \times \Delta \phi = 0.1 \times 0.1$). A Cluster Processor (CP) identifies and counts squares 2×2 clusters of trigger towers in which at least one of the four possible two-tower sums exceeds predefined thresholds (eight programmable thresholds for electron). For this analysis the threshold is set to 14 GeV and once a cluster exceeds it, the information about the geometric location of trigger object is retained and passed to the L2 trigger as Region of Interest (RoI).

The L1 muon triggers are based on signals from the muon trigger chambers: RPCs and TGCs. These detectors are grouped in stations at different distance from the interaction point. The L1 trigger searches for any coincidence of hits in the different trigger stations and includes the hits within a rectilinear road, which tracks the path of a muon from the interaction point through the detector. Due to the magnetic field which bend the muon trajectory, the smaller is the road which contains the hits, the higher is the $p_T$ of the particle. Once the road width is smaller than one of the six programmable width thresholds, the event is triggered and the RoI sent to L2 trigger. In this thesis the minimum $p_T$ required at L1 is 10 GeV/c.

The L2 analyses the events triggered by L1 looking at the RoIs with a complete detector information and performing stricter selection cuts. The triggering object is reconstructed at this level, however the reconstruction algorithms used are faster with respect to the offline ones in order to reduce the processing time. The L2 decision time is about 40 ms and the L2 selection reduces the event rate from 75 $\text{KHz}$ to 3.5 $\text{KHz}$. For electron triggers, L2 has access to the full calorimeter granularity in the RoI and for this analysis the $E_T$ of the cluster obtained using the refined information must be at least 20 GeV. Some electron identification cuts are also applied. The electron has to pass the online medium criterion which is softer than the offline one (see section 3.1). For muon triggers, the L2 uses also the information from the MDTs and CSCs in the RoI to take a decision. The momentum and track parameter resolution of the muon candidate are refined by fast fitting algorithms using the information provided by the muon spectrometer and the ID. A muon momentum
$p_T > 18 \text{ GeV/c}$ is required.

Event filter look at full detector information for events passing L2 trigger and reduces further the event rate to approximately 200 Hz, with an average event processing time of order four seconds. At the EF level, the muon reconstruction is more accurate than at L2 but the requirement on the $p_T$ remains 18 GeV/c. For the electrons, $E_T > 20 \text{ GeV}$ is required and the electron have to pass the offline medium criterion.

Once event filter has fired, all raw data coming from million of readout channels are stored in Tier One grid computing centers with a rate of 10/100 MB/s and from there they are accessible to the ATLAS community.
Chapter 3

Reconstruction and identification of physics objects

The raw data collected by the ATLAS detector consists of the ensemble of the signals recorded from about 100 millions electronic channels. This information is elaborated from several algorithms to extract the nature and four-momentum of the particles that generated these signals. The algorithms used to identify and reconstruct electrons, muons, jets and neutrinos are discussed in this chapter. In particular the presence of neutrinos is inferred using the missing transverse energy ($E_T^{\text{miss}}$) which is discussed in the last paragraph.

3.1 Electrons

The electrons in $|\eta| < 2.47$ are reconstructed with an algorithm that combines the information from the EM calorimeter with those form the ID. Other algorithms reconstruct electrons up to $|\eta| < 4.9$ using only the calorimetric information. This analysis uses only electron within $|\eta| < 2.47$ and the following is a description of the reconstruction algorithm adopted in this geometrical region.

3.1.1 Electron reconstruction

Electrons are reconstructed as energy deposits (clusters) in the EM calorimeter matched with a track\(^1\) in the ID. The sliding-window algorithm [48] searches for seed clusters in the second layer of the EM calorimeter with a transverse energy $E_T > 2.5$ GeV. The seed cluster is a window of size $3 \times 5$ cell units (each cell of the second layer of the EM calorimeter has dimensions $0.025 \times 0.025$) in the $\eta \times \phi$ plane. In case of superimposed windows it is retained the one with highest $E_T$.

It is required that at least one track from the ID matches with the seed cluster. The match is done requiring that the $\eta$ distance between the track and the cluster is $\Delta \eta < 0.05$. In the $\phi$ direction an asymmetric cut is used to take into account the difference between

---

\(^1\)Tracks are charged particles trajectories reconstructed using a pattern recognition algorithm that starts with the PIXEL and SCT information and then adds hits in the TRT. One further pattern recognition step starts from the TRT and works inwards adding silicon hits looking at hits not previously used. This allows the reconstruction of tracks from secondary interactions, such as photon conversions and long-lived hadron decays. A detailed description of the track reconstruction is presented in [47].
the impact point and the cluster position due to bremsstrahlung losses. Therefore in the $\phi$ direction toward the bending it is required a match $\Delta \phi < 0.05$ and away from the bending $\Delta \phi < 0.1$. In case of multiple matches, the track with silicon hits and the smallest $\Delta R$ is chosen.

Finally, the cluster is rebuilt using a $3 \times 7$ ($5 \times 5$) window which is the typical size of the electron deposit in the barrel (end-cap). The energy of the cluster is the sum of the estimated energy deposit in the material in front of the EM calorimeter, the measured energy deposit in the cluster, the estimated external energy deposit outside the cluster and the estimated energy deposit beyond the EM calorimeter.

The reconstructed electron four-momentum has energy given by the cluster energy and the direction given by the $\eta$ and $\phi$ coordinate of the track at the interaction vertex.

The basic definition of electrons is the starting point of a more refined identification used to increase the purity of the reconstructed electron sample.

### 3.1.2 Electron identification

There are three different levels of identification provided by ATLAS experts of electrons and photons (egamma group): loose, medium, tight. Each identification add to the previous some additional requirements. Therefore the tighter is the identification the larger is the jet rejection and at the same time the lower is the efficiency. In the analysis described in this thesis the $W$ decay in electron is reconstructed requiring an electron passing the tight criteria which has an expected rejection of fake electrons from jets of 50000.

The loose selection uses EM shower shape information from the second layer of the EM calorimeter and energy leakage into the hadronic calorimeters to discriminate between electrons and jets. The leakage is determined from ratio of the $E_T$ loss in the first and second layers of the hadronic calorimeter with respect to the $E_T$ of the cluster in the EM calorimeter. The lateral shape and the lateral leakage of energy is studied in a $7 \times 7$ cell unit window around the EM cluster. Hadrons are expected to have larger longitudinal leakage and wider transverse shape than electrons.

The medium selection improves rejection against hadrons by evaluating the energy deposit patterns in the first layer of the EM calorimeter, track quality variables and the cluster-track matching variables. Jets with single or multiple energetic $\pi^0$ and $\eta$ may produce an EM cluster similar to that of the electron because $\pi^0$ and $\eta$ decay in two photons $\gamma \gamma$ which form two close EM showers indistinguishable in the second EM calorimeter layer. Thanks to the high granularity of the first layer of EM calorimeter, the deposit due to a $\pi^0$ has often two maxima corresponding to the two photons. A cut on the difference between energy of second maximum and the minimal energy deposit between the two maxima is applied along with shower shape and width selection to remove signals originating from $\pi^0$ and $\eta$ hadrons.

The medium selection imposes to the track to have a stricter match with the EM cluster by cutting on $\eta$ difference between cluster and extrapolated track in the first EM layer ($\Delta \eta < 0.01$). The track needs also to have at least seven precision hits in PIXEL+SCT, at least one hits in PIXEL and a transverse distance of closest approach to the primary vertex $|d_0| < 5$ mm.

In addition to the cuts used in the medium selection, the tight selection applies further track and matching requirements. Tracks are required to release at least one hit in
3.1. ELECTRONS

the first layer of PIXEL to reject electrons from conversions and a minimum number of hits (depending on the \( \eta \)) in the TRT. The impact parameter to the primary vertex of the track \( d_0 \) has to be at most 1 \( mm \) and a constraint is put on the ratio of the track \( p_T \) to the EM cluster \( E_T \). Finally, the number of TRT straws with an high charge deposit is used as further discriminating cut.

3.1.3 Electron performances

The efficiencies (\( \epsilon \)) for reconstruction and identification are measured from data using the tag-and-probe method. The tag-and-probe method allows to measure the efficiency of any cut applied in an electron selection and is used to calculate muon selection efficiencies too. On data is applied a selection for the \( Z \rightarrow l^+l^- \) analysis which searches two leptons of the same flavor. One lepton (tag lepton) is reconstructed with best identification criteria, then another lepton (probe lepton) is selected by requiring basic acceptance and that the invariant mass with the tag lepton is included within 10 \( GeV/c^2 \) from the \( Z \) mass. The efficiency of a specific selection is the ratio of the number of events with the probe lepton which passes the selection divided by the number of events with the probe lepton. The efficiency is measured in Monte Carlo too, and a scale factor (SF) is defined as the ratio of the efficiencies (\( \epsilon_{\text{data}}/\epsilon_{\text{MC}} \)). The SFs are important because allow to reweigh Monte Carlo events in order to correct for small discrepancy in the MC.

The reconstruction efficiency for an electron is about 95% in the barrel region while in the end-cap regions varies between 90 and 95%. On average, the SF are about 1% [48]. The identification efficiency measured in data and the expected efficiency from MC are shown as a function of the electron \( \eta \) and \( p_T \) in Fig. 3.1. On average, the measured efficiency is about 80% and the identification SFs, excluded those in the transition region \( (1.37 < |\eta| < 1.52) \), differ from one at most of 3%.

In this analysis electrons are selected using the lowest available unprescaled\(^2\) single electron trigger, which is for the data analysed the \( EF_\text{e20}_\text{medium} \) described in section 2.5. Electrons candidate are selected with \( E_T > 25 \ GeV \) to ensure to be on the trigger efficiency plateau that is about 98% [49]. The electron trigger efficiency as a function of the \( E_T \) is shown if Fig. 3.2(a). The trigger scale factors varies from 0.97 to 1 depending on the \( \eta \).

\(^2\)Every event which satisfies this trigger is acquired.
CHAPTER 3. RECONSTRUCTION AND IDENTIFICATION OF PHYSICS OBJECTS

Fig. 3.1: Efficiencies measured from $Z \rightarrow ee$ events and predicted by MC for tight identification as a function of the electron $\eta$ and integrated over $20 < E_T < 50$ GeV (a) and $E_T$ and integrated over $|\eta| < 2.47$ excluding the transition region $1.37 < |\eta| < 1.52$ (b).

Fig. 3.2: Electron (a) and muon (b) trigger efficiencies as a function of the lepton $p_T$. 
3.2 Muons

Muons used in this analysis are reconstructed using information from the outer muon spectrometer (MS), the inner tracking detectors and the calorimeters. The reconstruction of these muons called combined muons (in contrast with stand-alone muons obtained from the MS information only) is obtained with a tracking algorithm that associates a track found in the muon spectrometer with an inner detector track, after the former is corrected for the energy loss in the calorimeter.

3.2.1 Muon reconstruction

The track reconstruction in the muon spectrometer starts from the search for straight track segments in each single muon station (MDT or CSC) in regions of size $\Delta \eta \times \Delta \phi = 0.4 \times 0.4$ where activity in the trigger systems (RPC or TGC) is detected. The segments have to loosely point to the center of interaction and to satisfy some quality factors in order to suppress random hits and background. In case of segments in the MDT detector, the $\phi$ coordinate of this segment is picked up from trigger detectors. A track is then obtained as a combination of two or more segments in different muon chambers using a least-square fitting method taking into account the magnetic field displacement.

The MS tracks are then extrapolated to the perigee considering the effect of multiple scattering and energy losses in the calorimeters and dead material, obtaining the candidate muon coordinates ($p_T$, $\eta$, $\phi$ and the impact parameters) at the interaction point.

The muon ID tracks are reconstructed with the same procedure used for the electron tracks. However, the quality criteria for the muon track are slightly different than that required for tight electron tracks. The track is required to have one hit in the inner layer of PIXEL (b-layer), at least two in the PIXEL and at least six in the SCT. Dead sensors crossed by tracks are counted as hits but they can be at most two. Finally, five hits in TRT are required with a constraint on the number of TRT associated as outliers which depends from $\eta$. The candidate ID muon coordinates are taken as the track parameters extrapolated at the perigee.

A $\chi^2$ test, defined from the difference between the respective track extrapolated coordinates weighted by their combined covariance matrices, is used to match tracks in ID with tracks in MS. The momentum of the combined muon is then calculated by the weighted average of the ID and the MS momentum measurements. The ID dominates the transverse momentum measurement up to $p_T \sim 80$ GeV/c in the barrel and $p_T \sim 20$ GeV/c in the end-caps. For $p_T$ close to 100 GeV/c the ID and MS measurements have similar weights while the MS dominates at $p_T > 100$ GeV/c.

3.2.2 Muon performances

The muon reconstruction efficiency have been measured from experimental data and estimated with MC simulations using the tag-and-probe methods with $Z \rightarrow \mu^+ \mu^-$ decays [50]. Fig. 3.3 shows the muon reconstruction efficiency as a function of $\eta$ and $p_T$ evaluated on 2010 data with the tag-and-probe method. The average reconstruction efficiency for $p_T^{\mu} > 20$ GeV/c is larger than 95% for all the $\eta$ range with the exception of the transition regions. The lower plots on Fig. 3.3 show the efficiency scale factors, defined as the ratio between the efficiency obtained on data over the efficiency obtained from MC. The mean
value of the $\eta$ dependent scale factor curve is $0.989 \pm 0.003$ [50]. The deviation from 1 is due to an efficiency drop in the transition region ($|\eta| \sim 1$) attributed to the limited accuracy of the magnetic field map used in the reconstruction of the ATLAS data in this region.

For this analysis the muon trigger employed is the logical OR of two trigger: the $\mu_{18\ MG}$ and the $\mu_{40\ MS\ only\ barrel}$. The first one is the unprescaled single muon trigger with the lowest $p_T$ threshold. It has a flat efficiency for muons with $p_T > 20 \ GeV/c$ (those used in this analysis) but in Monte Carlo has an efficiency loss at high $p_T$ due to a wrongly configured back extrapolator. The loss is recovered by the $\mu_{40\ MS\ only\ barrel}$ trigger which is fully efficient at those $p_T$ values. The $\mu_{18\ MG}$ trigger efficiency as a function of the muon $p_T$ is shown if Fig. 3.2(b). The overall data trigger efficiency is $0.8125 \pm 0.0015$ while the MC one is $0.7902 \pm 0.0002$, statistical errors only. The SFs correct the MC efficiencies of about 2% [51].

3.3 Jets

Once quarks and gluons are produced in an hard interaction, they are subject to gluons emissions ($q \rightarrow qg$, $g \rightarrow gg$) and decays ($g \rightarrow q\bar{q}$) as a result of the growth of QCD potential with the distance from the other partons. The overall effect is the formation of a shower of low energy partons. Once the energy is as low as QCD is no more perturbative, colored partons combine together to create colorless hadrons [11]. The detectable result of a gluon or quark production is a spray of hadrons called jet from which the parton properties (four-momentum and interaction point) are reconstructed.

Jets are reconstructed in the ATLAS detector as ensemble of energy deposits in the calorimeter system. Reconstructed jets use as input objects, three dimensional clus-
ters (topological clusters [52]) built associating calorimeter cells on the basis of the signal-to-noise ratio. Clusters are constructed around cells with a high signal-to-noise ratio and the discrimination to start and expand a cluster is based on the absolute value of the signal-to-noise ratio. This algorithm obtains a large noise suppression while introducing a small bias on the cluster energy. Ideally, topological clusters allow the association of calorimeter signals produced by the same particle shower. Topological clusters are the input objects for the jet reconstruction algorithm that aims at grouping the topoclusters generated from a single parton shower fragmentation.

In this thesis jets are reconstructed with the anti-$kt$ algorithm with distance parameter $R = 0.4$ [53]. The anti-$kt$ algorithm uses the following definition of distance between two objects:

$$d_{ij} = \begin{cases} \min \left( \frac{E_{T_i}^{-2}}{E_{T_j}^{-2}}, \frac{\Delta R_{ij}^2}{R^2} \right) & \text{if } i \neq j \\ E_{T_i}^{-2} & \text{if } i = j \end{cases}$$

(3.1)

where $\Delta R_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$, $y_i (y_j)$ and $E_{T_i} (E_{T_j})$ are respectively the rapidity and the transverse energy of the $i$-th ($j$-th) object. The algorithm searches for the combination of indices $i, j$ which minimize the $d_{ij}$. If $i = j$ the $i$-th object becomes a jet and is no more considered by the algorithm. If $i \neq j$ the $i$-th object is combined with the $j$-th one and they are substituted by a new object whose four-momentum is the sum of the four-momenta of the two original objects. Topoclusters are assumed to be massless objects. The result of applying the algorithm recursively is a list of composite objects far at least $R$ from each other which are the reconstructed jets.

The baseline calibration of the topological clusters is the electromagnetic scale, defined with electrons and muons during test-beams. At this energy scale the effect of calorimeter non-compensation or energy losses in uninstrumented material are not corrected. Therefore, after the jets are identified, a calibration scheme must be applied to correct for these effects and, in general, for any effects that make the reconstructed jet energy different from the reference true jet energy. The reference truth jets are obtained by running the anti-$kt$ algorithm on the ideal final state of a proton-proton collision where all particles with a lifetime longer than 10 ps are considered stable. This definition does not include muons and neutrinos from hadronic decays. The calibration scheme used in this thesis is obtained by applying $\eta$ and $p_T$ dependent scale factors that bring the reconstructed energy to the average truth energy [54, 55].

Pile-up affects the jet energy measurements because particles from pile-up interactions may overlap to jets and release energy in the same calorimetric cells. In this case the jet energy would be overestimated. At present a specific jet energy correction for this pile-up effect has not been studied yet. However simulation and collision data have been used to estimate the average energy shift that pile-up would give and this is used to evaluate the pile-up uncertainty on the jet energy scale (JES). In 2011 data this is evaluated to be $5 - 7\%$ for $20 < p_T < 50 \text{ GeV/c}$ and $2 - 3\%$ for $50 < p_T < 100 \text{ GeV/c}$ and negligible for $p_T > 100 \text{ GeV/c}$.

Pile-up not only affects the energy of jets produced in the primary collision but can also produce additional jets. This effect is minimized using the Jet Vertex Fraction (JVF) variable. This variable is obtained matching the tracks to the reconstructed vertexes. JVF is then evaluated by calculating the amount of transverse momentum carried by tracks matched to the jet that are also generated from the primary vertex. The more is the
pile-up contamination in the jet $E_T$, the lower is the JVF value. The JVF does not take into account the neutral component of the jet energy. Nevertheless it can help to remove jets highly affected by pile-up.

### 3.4 Missing Transverse Energy ($E_T^{\text{miss}}$)

Neutrinos and other weakly interactive particles escape from detector without interacting. However the presence of these particles in the event may be inferred using a quantity called missing transverse energy ($E_T^{\text{miss}}$). The $E_T^{\text{miss}}$ is defined as the vectorial sum of transverse energies of all detected particles changed in sign. Since the total transverse momentum is conserved, in absence of neutrinos (or other exotic non interacting particles), the $p_T$ of the detectable particles has to balance, therefore the $E_T^{\text{miss}}$, within the resolution, should be 0. In events with one neutrino, the $E_T^{\text{miss}}$ tends to point in the neutrino direction in the transverse plane with a magnitude similar to the neutrino’s $p_T$.

In ATLAS there are various algorithms to reconstruct the $E_T^{\text{miss}}$. The one used in this analysis determines the $E_T^{\text{miss}}$ from the energy collected by the electromagnetic and hadronic calorimeters and from the muons momentum measured by the muon spectrometer and inner detector. The contribute to the $E_T^{\text{miss}}$ of electrons, photons and hadrons is obtained as the sum of the $E_T$ of all topological clusters in the ATLAS calorimeter system, appropriately calibrated, changed in sign. The $E_T^{\text{miss}}$ obtained from the topoclusters is termed MET_LocHadTopo. Topological clustering extends to $|\eta| < 4.5$ allowing the usage of the $E_T$ of very forward particles. The region from $4.5 < |\eta| < 4.9$ is excluded because of energy calibration problems in the FCal.

The transverse momentum carried by muons which escape from calorimeters is accounted by the term MET_MuonBoy. For isolated muons ($\Delta R(\mu, j) > 0.3$) in $|\eta| < 2.4$ the $p_T$ is a combination of the inner detector and muon spectrometer measurement corrected for energy loss in calorimeters. For non-isolated muons or muons outside of the inner detector acceptance ($2.4 < |\eta| < 2.7$) the $p_T$ is obtained from the muon spectrometer track only. The transverse energy released by isolated muons in calorimeters is accounted in MET_RefMuonTrack and is used to avoid energy double counting. The overall $E_T^{\text{miss}}$ is:

\[
E_X^{\text{miss}} = \text{MET}_{\text{LocHadTopo}} - \text{MET}_{\text{MuonBoy}} - \text{MET}_{\text{RefMuonTrack}}
\]

\[
E_Y^{\text{miss}} = \text{MET}_{\text{LocHadTopo}} + \text{MET}_{\text{MuonBoy}} - \text{MET}_{\text{RefMuonTrack}}
\]

\[
E_T^{\text{miss}} = \sqrt{(E_X^{\text{miss}})^2 + (E_Y^{\text{miss}})^2}
\]
Chapter 4

Event samples and selection

My work has mainly concentrated on two analysis: the search for the CDF resonance at LHC and a study to evaluate the sensitivity to measure the diboson $WW/WZ$ signal at LHC. Since these two analysis have the same final state topology they share most of the background sources and event selection. In this chapter all the information that is common to the two analysis is described specifying, when needed, the differences between them.

4.1 Event samples

The searches of the two dijet resonances (CDF excess and $W/Z \to jj$ boson) are made by comparing, with a statistical method, the jet-jet invariant mass obtained from data to the Standard Model expectation. The aim in the $WW/WZ$ analysis is to measure the diboson cross-section through a fit of the fractions of each SM process in data. The Standard Model expectation distribution is obtained from a data-driven method for what concern the multijet QCD component and from samples of Monte Carlo simulated data for the other SM processes. These latter data are produced using event generators (examples of these are HERWIG [56], PYTHIA [57], ALPGEN [58]) which generate the requested processes from the proton-proton collisions. The generators are based both on theory and on models for a number of physics aspects which might include hard and soft interactions, parton distributions, initial and final state parton showers, multiple interactions, fragmentations and decays. The output of the generators is a list of the particles, including their 4-vectors and types, that form the ideal final state that is “seen” from the detector. The effect of the detector on the generated final state is than simulated using the program GEANT [59].

This program contains a very detailed description of the detector material and geometry, simulates the interactions of particles with matters and gives on output the exact same information that is contained in the real detector data. At this point the reconstruction program ATHENA [60] is used to obtain from the detector information, simulated or real, the information about the physics objects. The distributions obtained from the simulated data give the SM expectation and can therefore be compared to the ones obtained from the real detector data.

In these analysis the dijet invariant mass is investigated. A disagreement between data and SM expectation, incompatible with the estimated uncertainty, would be considered as
CHAPTER 4. EVENT SAMPLES AND SELECTION

non-Standard Model signal as it could be the case of the CDF resonance. In order to state that a significant difference is a new physics signal, it is important to include all possible SM processes that can contribute to the dijet invariant mass distribution after the event selection. This aspect is not less important in the WW/WZ analysis where the fit of data to extract the SM fractions has meaning only if all the contribution are considered.

The two analysis have to select events with at least two jets, one neutrino and one lepton in the final state. Therefore, the Monte Carlo samples considered for the SM prediction are those in which the final state corresponds to exactly the same topology as the signal (irreducible background) or those in which one of the particles may be misidentified making the final state looks like the signal topology (reducible backgrounds).

In the following section the characteristics of the Monte Carlo and data samples are described.

4.1.1 Monte Carlo samples

The Monte Carlo simulated samples used for these analysis include \( W/Z + \text{jets} \), \( t \bar{t} \), single top and WW/WZ events. The \( W + \text{jets} \) and \( Z + \text{jets} \) samples are generated with ALPGEN 2.13 [58] interfaced to HERWIG [56] to simulate the parton shower and hadron fragmentation. MC@NLO [61] interfaced to HERWIG is used to generate \( t \bar{t} \) and single top events in the \( t \) and \( s \) channels. Diboson (WW, WZ) productions are simulated by HERWIG.

As it will be discussed in the following sections the \( W + \text{jets} \) processes give the largest contribution to the expected SM jet-jet invariant mass. The \( Z + \text{jets} \) may also contribute when one of the two leptons from the leptonic \( Z \) decay is not detected or is misidentified giving a large missing transverse energy. The \( W + \text{jets} \) and \( Z + \text{jets} \) processes are simulated considering separately at the diagram level, the emission of a \( W/Z \) with a number of partons varying from 0 to 5 (\( W + n\text{partons} \), \( Z + n\text{partons} \) samples where \( n = 0, 1, 2, 3, 4, 5 \)). It should be noticed however that, for various reasons, the number of jets detected in the final state does not always corresponds to the number of simulated partons. Firstly, when two protons collide, the partons of the protons that do not undergo the hard-scattering process may interact among them generating softer processes that are color connected to the hard-scattering process. This effect, indicated with the name of Underlying Event, may also generate jets that contribute to the total event multiplicity. Secondly, it may happen that the jet algorithm splits the hadrons generated from a single parton fragmentation, causing a single parton being detected as two jets. Similarly two parton fragmentations may be merged in a single jet. As a result, the \( W \) or the \( Z \) samples produced with 0 or 1 parton may pass the selection and contribute to the SM prediction of the dijet invariant mass.

The top quark \( t \) decays almost completely in \( bW \), hence the \( t \bar{t} \) very often produces a couple of \( W \) bosons and a couple of \( b \)-jets. The two \( W \) bosons can than decay semileptonically producing the final state that is searched with one lepton, \( E_T^{\text{miss}} \) from the neutrino and a couple of jets. The only difference is the presence of two more \( b \)-jets. In the CDF resonance analysis the \( b \)-jet identification has not been used therefore jets produced by the fragmentation of the \( b \)-quarks are treated exactly as any other jets. In the WW/WZ analysis it is made an attempt to identify the jets produced from a \( b \)-quark to remove \( t \bar{t} \) events. However this method has a limited efficiency and \( t \bar{t} \) background can not be
neglected. The single top process is also considered a background process since the $W$ may decay leptonically and the $b$-jet together with another jet may give the searched final topology.

In all the Monte Carlo samples, the $W$ and $Z$ leptonic decays are simulated in all three lepton flavors. $\tau$ leptons whose mean lifetime is about $10^{-15}$ s are not identified directly but they can decay in $e$ or $\mu$ and neutrinos ($\text{BR}(\tau^{-} \rightarrow l^{-} \bar{\nu}_{l} \nu_{\tau}) \sim 17.5\%$ with $l = e, \mu$ [11]). For this reason, events with $W$ and $Z$ decays in $\tau$ leptons also enter the selection.

In the diboson analysis I used additional $W + n$partons samples. These are the $W$ produced in association with heavy flavor quarks (charm and bottom quarks) samples. The standard $W + n$partons samples contain heavy flavor quarks but only produced in the parton shower process, for example through the process $g \rightarrow b \bar{b}$. In order to correctly take into account the kinematic characteristics of event samples with heavy flavour quarks, the samples in which the heavy flavour quarks are generated also through the matrix element process are considered (Table 4.2). When these samples are used the heavy flavor contribution is removed from the inclusive samples to avoid the double counting.

The $WW$ and $WZ$ samples are obviously used in the diboson analysis but also the CDF resonance analysis includes these productions. Their semileptonic decay constitutes an irreducible background which contributes to the dijet invariant mass mostly in the region $80 - 90 \text{ GeV}/c^2$. They affect peripherally the region at $147 \text{ GeV}/c^2$ where the CDF resonance is seen.

In each Monte Carlo sample, the pile-up contribution is simulated overlaying to the hard-scattered event a number of minimum bias events similar to the number of multiple proton-proton interactions occurring in real bunch collisions. Since the number of pile-up events changes with the beam conditions the simulation is generated with a nominal distribution of overlaid minimum bias events and than Monte Carlo events are reweighed scaling the distribution of the number of simultaneous collision in Monte Carlo to that one observed in data. The plot (a) in Fig. 4.1 shows the distribution of the number of interactions per event in Monte Carlo and it gives on average 8 interactions per event. In data, the number of interactions is measured event by event using the luminosity detectors [62]. At analysis level the number of interactions is averaged over a period of data corresponding to about one minute (luminosity block). The measured average number of interactions per luminosity block is shown on the plot (b) in Fig. 4.1. In the data samples used in the two analysis the average number of interactions per event is about 6.

As briefly discussed above the detector response to the generated events is simulated through the GEANT4 detector simulation [63]. This is a very important step since any detector response mismodeling could cause discrepancies between Monte Carlo and data. For example, some of these mismodelings are at the basis of the different reconstruction efficiency and $p_T$ resolution noticed in MC reconstructed leptons and in data ones. Corrections obtained from collision data (data-driven corrections) are thus applied to Monte Carlo events in order to remove these discrepancies.

A last contribution to the SM prediction is given by the multijet QCD production. These type of events contribute to the SM prediction since jets faking the lepton identification or leptons from heavy flavor decays can be selected as the $W$ decay products. These type of events have low probability to be selected but the very large multijet QCD cross-section (order of few mb) makes the number of events passing the selection not negligible. The difficulty to reproduce the jet misidentification in Monte Carlo requires to
estimate the multijet QCD contribution with a data-driven method. The one adopted in this thesis is described in section 4.3.

Details of the Monte Carlo samples used in this thesis for both the dijet resonances and WW/WZ studies are reported in Table 4.1. Table 4.2 collects information on the samples used exclusively in the diboson analysis. The Tables report the production cross-section $\sigma$, the efficiency of the filter applied at production level to remove non interesting events and the $k$-factor used to correct the leading order (LO) estimated cross-section to the next to leading order (NLO) value. These numbers are used to obtain the correct normalizations of the SM predictions. $t\bar{t}$ and single top samples are produced at the NLO, while all the other processes at the leading order of the perturbative theory. The sample sizes generally correspond to those expected in 1 $fb^{-1}$ of integrated luminosity; some samples like WW, $t\bar{t}$ have an higher statistics.

4.1.2 Data sample

The data sample used corresponds to $pp$ collisions at $\sqrt{s} = 7$ TeV acquired in 2011. The data are divided in two sets indicated with egamma and muon streams. The egamma stream is a subsample of events collected using electron and photon trigger types. These events are used to search for the electron decay channel described in section 4.2.1. The muon stream data are triggered by muon trigger types and are used to search for the muon decay channel (section 4.2.2).

A first preliminary selection is applied to keep only those events acquired with stable proton-proton collisions and with all the detector well functioning. The integrated luminosity is calculated after this preliminary selection which discards a few percent of the total events. In the search of the dijet resonance the integrated luminosity used amounts to 1.02 $fb^{-1}$, while for the WW/WZ studies it is upgraded to 1.33 $fb^{-1}$. The luminosity

Fig. 4.1: Distribution of the number of interactions per event in Monte Carlo (a) and data (b). For data, the black line shows the distribution for the data used in the analysis of the CDF bump while the red one the distribution for the sample used in the WW/WZ cross-section studies.
### 4.1. EVENT SAMPLES

<table>
<thead>
<tr>
<th>Process Sample</th>
<th>$\sigma$(nb)</th>
<th>$\epsilon$</th>
<th>$k$-factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>$W + jets$: Alpgen.JimmyWenuNp0</td>
<td>$6.92E + 00$</td>
<td>1.2</td>
<td>1.2</td>
</tr>
<tr>
<td>Alpgen.JimmyWenuNp1</td>
<td>$1.30E + 00$</td>
<td>1.2</td>
<td>1.2</td>
</tr>
<tr>
<td>Alpgen.JimmyWenuNp2</td>
<td>$3.78E - 01$</td>
<td>1.2</td>
<td>1.2</td>
</tr>
<tr>
<td>Alpgen.JimmyWenuNp3</td>
<td>$1.02E - 01$</td>
<td>1.2</td>
<td>1.2</td>
</tr>
<tr>
<td>Alpgen.JimmyWenuNp4</td>
<td>$2.57E - 02$</td>
<td>1.2</td>
<td>1.2</td>
</tr>
<tr>
<td>Alpgen.JimmyWenuNp5</td>
<td>$7.02E - 03$</td>
<td>1.2</td>
<td>1.2</td>
</tr>
<tr>
<td>$W + jets$: Alpgen.JimmyWmunuNp0</td>
<td>$6.92E + 00$</td>
<td>1.2</td>
<td>1.2</td>
</tr>
<tr>
<td>Alpgen.JimmyWmunuNp1</td>
<td>$1.30E + 00$</td>
<td>1.2</td>
<td>1.2</td>
</tr>
<tr>
<td>Alpgen.JimmyWmunuNp2</td>
<td>$3.78E - 01$</td>
<td>1.2</td>
<td>1.2</td>
</tr>
<tr>
<td>Alpgen.JimmyWmunuNp3</td>
<td>$1.02E - 01$</td>
<td>1.2</td>
<td>1.2</td>
</tr>
<tr>
<td>Alpgen.JimmyWmunuNp4</td>
<td>$2.56E - 02$</td>
<td>1.2</td>
<td>1.2</td>
</tr>
<tr>
<td>Alpgen.JimmyWmunuNp5</td>
<td>$6.99E - 03$</td>
<td>1.2</td>
<td>1.2</td>
</tr>
<tr>
<td>$W + jets$: Alpgen.JimmyWtaunuNp0</td>
<td>$6.92E + 00$</td>
<td>1.2</td>
<td>1.2</td>
</tr>
<tr>
<td>Alpgen.JimmyWtaunuNp1</td>
<td>$1.30E + 00$</td>
<td>1.2</td>
<td>1.2</td>
</tr>
<tr>
<td>Alpgen.JimmyWtaunuNp2</td>
<td>$3.78E - 01$</td>
<td>1.2</td>
<td>1.2</td>
</tr>
<tr>
<td>Alpgen.JimmyWtaunuNp3</td>
<td>$1.02E - 01$</td>
<td>1.2</td>
<td>1.2</td>
</tr>
<tr>
<td>Alpgen.JimmyWtaunuNp4</td>
<td>$2.57E - 02$</td>
<td>1.2</td>
<td>1.2</td>
</tr>
<tr>
<td>Alpgen.JimmyWtaunuNp5</td>
<td>$6.99E - 03$</td>
<td>1.2</td>
<td>1.2</td>
</tr>
<tr>
<td>$single\ top$: st_schanemu_McAtNlo_Jimmy</td>
<td>$4.69E - 04$</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>st_tchanemu_McAtNlo_Jimmy</td>
<td>$7.12E - 03$</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>$single\ top$: st_schannumu_McAtNlo_Jimmy</td>
<td>$4.69E - 04$</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>st_tchannumu_McAtNlo_Jimmy</td>
<td>$7.11E - 03$</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>$single\ top$: st_schanmutau_McAtNlo_Jimmy</td>
<td>$4.69E - 04$</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>st_tchannmutau_McAtNlo_Jimmy</td>
<td>$7.10E - 03$</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 4.1: Types of Monte Carlo samples, theoretical cross-sections $\sigma$, filter efficiencies $\epsilon$ and $k$-factors used in the analysis.
CHAPTER 4. EVENT SAMPLES AND SELECTION

<table>
<thead>
<tr>
<th>Process</th>
<th>Sample</th>
<th>$\sigma$(nb)</th>
<th>$\epsilon$</th>
<th>k-factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>$W + HF$:</td>
<td>AlpgenWcNp0</td>
<td>$6.44E-01$</td>
<td>1</td>
<td>1.2</td>
</tr>
<tr>
<td></td>
<td>AlpgenWcNp1</td>
<td>$2.05E-01$</td>
<td>1</td>
<td>1.2</td>
</tr>
<tr>
<td></td>
<td>AlpgenWcNp2</td>
<td>$5.08E-02$</td>
<td>1</td>
<td>1.2</td>
</tr>
<tr>
<td></td>
<td>AlpgenWcNp3</td>
<td>$1.14E-02$</td>
<td>1</td>
<td>1.2</td>
</tr>
<tr>
<td></td>
<td>AlpgenWcNp4</td>
<td>$2.8E-03$</td>
<td>1</td>
<td>1.2</td>
</tr>
<tr>
<td></td>
<td>AlpgenWccFullNp0</td>
<td>$1.28E-01$</td>
<td>1</td>
<td>1.2</td>
</tr>
<tr>
<td></td>
<td>AlpgenWccFullNp1</td>
<td>$1.05E-01$</td>
<td>1</td>
<td>1.2</td>
</tr>
<tr>
<td></td>
<td>AlpgenWccFullNp2</td>
<td>$5.20E-02$</td>
<td>1</td>
<td>1.2</td>
</tr>
<tr>
<td></td>
<td>AlpgenWccFullNp3</td>
<td>$1.70E-02$</td>
<td>1</td>
<td>1.2</td>
</tr>
<tr>
<td></td>
<td>Alpgen.JimmyWbbFullNp0</td>
<td>$4.73E-02$</td>
<td>1</td>
<td>1.2</td>
</tr>
<tr>
<td></td>
<td>Alpgen.JimmyWbbFullNp1</td>
<td>$3.58E-02$</td>
<td>1</td>
<td>1.2</td>
</tr>
<tr>
<td></td>
<td>Alpgen.JimmyWbbFullNp2</td>
<td>$1.73E-02$</td>
<td>1</td>
<td>1.2</td>
</tr>
<tr>
<td></td>
<td>Alpgen.JimmyWbbFullNp3</td>
<td>$6.63E-03$</td>
<td>1</td>
<td>1.2</td>
</tr>
</tbody>
</table>

Table 4.2: Monte Carlo heavy flavor samples, theoretical cross-sections $\sigma$, filter efficiencies $\epsilon$ and $k$-factors used in the diboson analysis.

is known with an uncertainty of about 3.7% [64].

In all plots where the data distribution are compared with the Standard Model expectation the color convention which identifies the samples used is: red and purple for $WW/WZ$, blue for $t\bar{t}$, blue checked for single top, green for $Z + jets$, grey for multijet QCD, orange checked for $W \rightarrow \tau\nu + jets$, a scale of yellows for $W \rightarrow e\nu + jets$ and a scale of cyans for $W \rightarrow \mu\nu + jets$. In the diboson analysis the heavy flavor component of the $W + jets$ background is colored with dark purple. Black dots represents data.

4.2 Event selection

What follows is a description of the standard selection which is adopted in ATLAS to reconstruct a $W$ decay in the leptonic final state. It is made of a set of cleaning cuts and prescriptions to reconstruct the lepton, the $E_T^{miss}$ and the $W$. The requirements used to select jets for the analysis explained in this thesis are presented in the last part of this section.

The first part of the selection aims at selecting well reconstructed interactions and at eliminating residual detector problems not treated at the central level. Events are required to have a reconstructed primary vertex with at least three associated tracks with $p_T > 0.5$ GeV/c. The primary vertex is defined as the one with the highest sum of track $p_T$ squared. This condition is used to select events with at least one hard-scattering.

Events with large noise bursts and data integrity errors in the Liquid Argon calorimeter are discarded. In a large part of the events the Liquid Argon calorimeter has a non sensitive area ($\eta \times \phi \in [0.0, 1.45] \times [-0.78847, -0.59213]$) due to an hardware failure. The method to deal with this problem is described in the following sections and the area interested is indicated as the $LAr$ Hole region.

The next step of the selection searches for events with one lepton with high transverse momentum and with high missing transverse energy. These conditions along with a large
transverse mass of the system lepton plus $E_T^{\text{miss}}$ are the standard selections used to identify a leptonically decaying $W$. In these type of events the trigger is easily obtained from the lepton, electron or muon, with relatively high transverse momentum.

The selection criteria to identify the leptonically decaying $W$, described in sections 4.2.1, 4.2.2, 4.2.3, are common to the CDF dijet resonance and to the $WW/WZ$ searches. The same is true also for most of the jet selection (section 4.2.4). The two analysis differ in a set of kinematic selections in the first case optimized for invariant dijet masses above 80 $GeV/c^2$, in the latter studied to remove the background in the region of the $W$ and the $Z$ masses. These exclusive selections are described in chapter 5 for the dijet resonance search and in chapter 6 for the $WW/WZ$ analysis.

### 4.2.1 Electron selection

Events for the electron channel analysis are selected by requiring that the unprescaled electron trigger with the lowest trigger $p_T$ threshold ($EF_e20\text{medium}$) is satisfied. This trigger selects electrons with transverse energy, at online level, $E_T > 20$ GeV. The trigger reaches the efficiency plateau for electrons with $p_T > 25$ GeV/c. As a consequence the electrons are required to have $p_T > 25$ GeV/c. Since the electron energy resolution in MC is slightly smaller than in data (they differ of about 2.5%) the electron energy in MC is smeared to adjust the resolution to that one measured in data. This is done before any selection is applied.

Candidate electrons are accepted only in the region covered by the ID and by the EM calorimeter ($|\eta| < 2.47$) excluding the transition region ($1.37 < |\eta| < 1.52$) between the barrel and end-cap electromagnetic calorimeters and also the ones pointing to the LAr Hole region. Whatever is the electron direction, a quality control is made on the single calorimeter cells excluding electrons measured by cells with problems.

The candidate electrons are requested to be generated from the primary interaction by requiring that their tracks are within 1 cm from the primary vertex in the $Z$ direction and within 10 standard deviations (the distance at most ten times larger than the error on the distance $\sigma_{d0} \sim 0.05$ mm) in plane transverse to the beams.

Candidate electrons must satisfy the tight identification criteria described in section 3.1 and have to be isolated. The isolation criteria is satisfied when the calorimetric transverse energy deposits in a cone of radius $R = 0.3$ around the electron cluster, corrected for the pile-up contribution, is lower than 4 GeV. The pile-up correction is applied on the basis of the number of primary vertexes with at least two tracks in the event. The larger is the number of primary vertexes the larger is the deposit in calorimeters from secondary interactions. The correction is made with a data-driven method and depends on the electron spatial coordinates $\eta$ and $\phi$. The isolation condition and the tight identification selections are studied to have the largest rejection against jets ($\sim 10^5$) while keeping a high efficiency on electrons ($\sim 75\%$). Fig. 4.2 shows the reduction of multijet QCD events given by this set of cuts with respect to using a set of less stringent (medium) identification criteria and an uncorrected isolation.

Finally, events containing a second electron are removed. The veto on the second electron is applied to electrons passing all previous requirements but with medium identification criterion and $p_T > 20$ GeV/c. This cut aims at reducing the contribution from $Z + jets$ event with a $Z \to ee$. 
Fig. 4.2: Jet-jet invariant mass distributions obtained for events with medium electrons isolated without pile-up correction (a) and for events with tight electrons isolated accounting the pile-up (b). This is a preliminary study made with limited integrated luminosity and background components. Data are represented by red dots with bars corresponding to the statistical uncertainties. The filled area is the total Standard Model prediction and each colored zone represents a specific process (color convention also explained in section 4.1.2). The grey component is the multijet QCD. The jet selection is similar to that used in the WW/WZ analysis.

In summary the electron cut-flow is:

1. trigger $EF_e20_{\text{medium}}$;
2. electron quality selections;
3. transverse impact parameter $d_0 < 10$, longitudinal impact parameter $z_0 < 1 \text{ cm}$;
4. electron $|\eta| < 2.47$ and not in the transition region;
5. electron identification: tight;
6. electron $p_T > 25 \text{ GeV/c}$;
7. electron isolation $\Sigma E_{\text{Tcorr}} < 4 \text{ GeV}$ in a cone $R = 0.3$;
8. veto on events with a second medium electron with $p_T > 20 \text{ GeV/c}$.

4.2.2 Muon selection

Muon events are collected using two triggers: the $EF_{\text{mu}18_{\text{MG}}}$ and the $EF_{\text{mu}40_{\text{MSonly_barrel}}}$. They are both unprescaled and allow an uniform efficiency for muons with $p_T > 20 \text{ GeV/c}$ which is the threshold used for reconstructed muons. As for the electron channel, in Monte Carlo events the muon momentum is smeared to reach the resolution observed in data (the difference of resolution is about 2%).


4.2. EVENT SELECTION

A muon candidate must be a combined muon reconstructed using the information from the inner detector and the muon spectrometer as explained in section 3.2. Muon candidate tracks must have distance of closest approach to the primary vertex smaller than 1 cm in the Z direction and within 10 standard deviations ($\sigma_{d0} \sim 0.02 \text{ mm}$) in the transverse plane. These cuts aim at removing muons originated from cosmic ray interactions and muons generated from heavy flavor decays in jets which tends to have larger impact parameters. Muons originated from decays in jets are also removed by requiring isolated muons. The muon isolation requirement is applied to the sum of the transverse momentum of the tracks in a cone of radius $R = 0.2$ around the muon track, this sum is required to be less than the 10% of the muon transverse momentum. As further requirement the muon candidate must be within $|\eta| < 2.4$ and a final veto is applied to events with a second muon in $|\eta| < 2.5$ and satisfying all the other selection criteria. This veto allows to reject a part of the events with $Z \rightarrow \mu\mu$ decays. The muon reconstruction is not affected by the problem in the LAr calorimeter, hence no further condition is applied to muons in the region of the hole.

In summary the muon selected satisfies the following requirements:

1. trigger $EF\_mu18\_MG$ or $EF\_mu40\_MOnly\_barrel$;
2. muon quality selections;
3. inner detector track quality cuts;
4. transverse impact parameter $d_0 < 10$, longitudinal impact parameter $z_0 < 1$ cm;
5. muon $|\eta| < 2.4$;
6. muon reconstruction: combined;
7. muon $p_T > 20$ GeV/c;
8. muon isolation $\Sigma p_T < 0.1 p_T$, in a cone $R = 0.2$;
9. veto on events with a second isolated combined muon with $p_T > 20$ GeV/c and $|\eta| < 2.5$.

4.2.3 W → lν reconstruction

Whatever is the lepton flavor selected, the event selection continues with further requirements equal for the two channels. In the following to distinguish the two flavors they are indicated as the electron and the muon selection.

The presence of one neutrino in the leptonic $W$ decay is accounted by introducing a lower threshold on the amplitude of the $E_T^{miss}$. Since detector problems may easily produce fake $E_T^{miss}$, before the $E_T^{miss}$ cut is applied, events are inspected to reject problematic events. The quality criteria applied to muons and electrons have already been described above. Looking at the jet properties as electromagnetic fraction, jet charged fraction, jet quality in the LAr and in the hadronic end-cap calorimeters three levels of jet quality are defined (loose, medium, tight). Jets that fall in the worst quality criteria (loose) are rejected since they are most probably produced from electronic noise. Moreover jets which corresponds to energy deposition in a region where the energy measurement is not accurate due to dead cells or transition among parts of the calorimeter are also rejected.
Events are then selected if the $E_T^{\text{miss}} > 25$ GeV. In this way, a large percentage of multijet QCD events is removed because they tend to have balanced jets and not energetic neutrinos. Some of the multijet QCD events can still pass this selection cuts because of the limited $E_T^{\text{miss}}$ resolution.

The final requirement applied to identify the candidate leptonically decayed $W$ is a cut on the transverse mass ($M_T$) which is computed with the Eq. 4.1:

$$M_T = \sqrt{(E_X^{\text{miss}} + p_{Tlep})^2 - (E_X^{\text{miss}} + p_{Xlep})^2 - (E_Y^{\text{miss}} + p_{Ylep})^2} = \sqrt{2(E_T^{\text{miss}})(p_{Tlep})(1 - \cos(\phi_{Tlep} - \phi_{lep}))}$$

where $p_{Tlep}$, $p_{Xlep}$, $p_{Ylep}$ are the transverse, X and Y momentum components of the lepton; $E_X^{\text{miss}}$, $E_Y^{\text{miss}}$ are the missing energies along the X and Y axis. The cut is set to $M_T > 40$ GeV/c$^2$ and removes the multijet QCD tail at low $M_T$. Fig. 4.3 shows the transverse mass distribution before the $M_T$ cut is applied for the electron and muon channels for data and for the SM predictions. The agreement between Monte Carlo and data at this step of the selection is not optimal but it improves considerably when only events with at least two good jets (see section 4.2.4) are selected (Fig. A.5). A plausible explanation for this disagreement could be a light mismodeling of the contribution given by $W$ produced with 0 partons. In fact while the $W+0\text{partons}$ is the dominant process after the $E_T^{\text{miss}}$ cut once two good jets are required, this contribution is highly decreased (see Fig. 5.1) and the agreement improves. The same effect is noticed also in the distribution of the $E_T^{\text{miss}}$. The $E_T^{\text{miss}}$ distribution before and after the requirement at least of two good jets, can be found in Fig. A.6(a,b) (A.7(a,b)) for the electron (muon) selection. After the selection of at least two jets the $E_T^{\text{miss}}$ distribution presents a better agreement between data and Monte Carlo than before.

In brief, the selection used to reconstruct the leptonic $W$ requires:

1. no jets with quality lower than loose;
2. missing transverse energy $E_T^{\text{miss}} > 25 \text{ GeV}$;
3. transverse mass $M_T > 40 \text{ GeV}/c^2$.

4.2.4 Jet selection

The following step of the analysis is the selection of the jets. I used jets reconstructed with the anti-$k_t$ algorithm with parameter $R = 0.4$ as explained in section 3.3. Since the jets selected are used to calculate the invariant mass, the quality required for them is better than the one used to remove events with fake jets. The medium quality is adopted and it consists in stricter requirements on electromagnetic fraction, jet charged fraction, jet quality in LAr and in hadronic end-cap calorimeters.

Jet transverse momentum has to be at least $30 \text{ GeV}/c$. This requirement is a compromise between the need to have a large statistical sample while containing the effect of the Jet Energy Scale uncertainty. Jets with lower transverse momenta have larger systematic uncertainties. The jet pseudorapidity must be within $2.8$ the region where jets are measured with the best resolution and lower systematic uncertainty.

The jets position cannot overlap with the lepton selected in the $\eta \times \phi$ plane ($\Delta R(j, l) < 0.5$) to avoid any double identification of the same object as a jet and a lepton. It should be noted that the discrimination between electrons and jets is done at this level in fact the list of jets by default contains also electrons which must be excluded.

The jets which pass all the selection explained above are termed good jets and, unless differently specified, the jets in the rest of the analysis are of good type. The events which are interesting for the two analysis are those which have at least two good jets ($N_{\text{jets}} \geq 2$).

A further condition is applied to remove those events with energetic jets pointing to the LAr Hole. Since the energy of a jet fallen in the LAr hole is not well measured, if the jet has $p_T > 30 \text{ GeV}/c$ then the event is removed. The $p_T$ is extrapolated from the information of the neighbour cells to the LAr Hole.

The following list summarizes the selections which define a good jet:

1. medium quality cuts;
2. jet $p_T > 30 \text{ GeV}/c$;
3. jet $|\eta| < 2.8$;
4. distance from the lepton selected $\Delta R(j, \text{lep}) < 0.5$.

The events selected for the analysis have:

- no jet with $p_T > 30 \text{ GeV}/c$ pointing to the LAr Hole;
- at least two good jets $N_{\text{jets}} \geq 2$.

4.3 Estimate of the multijet QCD contribution

As explained in section 4.2.3, most of the multijet QCD background is removed by requiring $E_T^{\text{miss}} > 25 \text{ GeV}$ and by selecting events with well identified and isolated muons.
or electrons. Nevertheless the multijet QCD background is not negligible due to its high
cross-section and even after the selection mentioned above it constitutes about 10% of
the SM predicted events. In this case the Monte Carlo is not reliable to evaluate this
contribution since it is very difficult to correctly model the rate of jets faking the lepton
identification. Therefore the multijet QCD is estimated using a data-driven method.

The data-driven method exploits a suitably modified lepton selection to define a control
sample dominated by multijet QCD background and with kinematic distributions as close
as possible to those of the standard selection. This sample is used to define the shape of
the distributions for the multijet QCD background at various level of the cut-flow.

For the electron channel this multijet QCD enriched sample is obtained by those data
where the electron identified as medium do not pass two of the tight requirements. These
electrons are called anti-electrons and fails exactly two of the five requirements on these
variables:

1. number of hits in the b-layer;
2. ratio of track momentum and cluster energy;
3. impact parameters of the track;
4. number of hits in the TRT;
5. identification criteria in TRT.

The larger is the number of requirements failed, the higher is the purity because less true
electrons enters in the sample. On the other hand, the less is the number of failed cuts,
the more is the statistics and the compatibility. The choice to invert two out of the five
requirements is a compromise solution between statistics and purity of the sample and
gives the best overall performance [65]. With the assumption made, the distributions
obtained from data inverting two of the tight criteria has the shape which should assume
the multijet QCD samples after the standard cut-flow.

For the muon channel the multijet QCD control sample is defined by applying the muon
selections described above, but inverting the cut on the transverse impact parameter with
respect to the primary vertex. The track must still be within 10 mm of the vertex in Z. In
this way the sample is composed of muons that do not originate from the primary vertex,
as expected for muons produced from heavy-flavor decays in jets. Fig. 4.4 [66] shows the
$E_T^{\text{miss}}$ distribution for data after the muon selection with the $d_0$ cut inverted (anti-$d_0$) is
applied and the same distribution obtained from a Monte Carlo sample of multijet QCD
by applying the standard muon selection. It is also shown the $E_T^{\text{miss}}$ distribution obtained
from data inverting the isolation criterion. The sample which passes this selection is
enriched of multijet QCD events since the muons produced in decays within the jets are
not isolated. Therefore the inversion of the isolation requirement is another candidate
method to extrapolate the multijet QCD distribution. The overall agreement between the
data distribution obtained with the $d_0$ requirement inverted and the Monte Carlo multijet
QCD prediction is better than that between the data distribution with non isolated muons
and the Monte Carlo multijet QCD one. This supports the use of the data-driven method
which uses the muon selection with the $d_0$ cut inverted to extract the shapes of the
distributions for multijet QCD events.
4.3. ESTIMATE OF THE MULTIJET QCD CONTRIBUTION

Fig. 4.4: Missing transverse energy distribution for data with the \( d_0 \) cut inverted (red), data with the isolation cut inverted (blue) and Monte Carlo multijet QCD (black), all normalized to the same area. The Monte Carlo of multijet QCD agrees better with data selected with the anti-\( d_0 \) cut than with those selected with the anti-isolation cut [66].

The inversion methods determine the shapes of multijet QCD distributions but cannot be used to obtain their normalizations. The normalization is extrapolated from a fit to the \( E_T^{\text{miss}} \) distribution which also determines the \( W + \text{jets} \) normalization. It is a likelihood fit which takes into account both the data and Monte Carlo statistical uncertainties [67]. The choice to obtain the normalization using the \( E_T^{\text{miss}} \) distribution is motivated by the very different shapes of the \( E_T^{\text{miss}} \) for the multijet QCD and \( W + \text{jets} \) samples. For example these two samples would not be separated using the dijet invariant mass (Fig. A.10). The multijet QCD covers the low \( E_T^{\text{miss}} \) region and peaks near to 20 GeV while the \( W + \text{jets} \) peaks at 40 GeV as can be seen in Fig. 4.5.

The MC \( E_T^{\text{miss}} \) distribution used to fit the data distribution is obtained by fixing the normalization and the shape of the \( t\bar{t}, \text{single top}, Z + \text{jets}, WW \) and \( WZ \) samples to the values predicted by the Monte Carlo simulations. The template for the \( W + \text{jets} \) contribution is also taken from the Monte Carlo, and the multijet QCD background shape from the control samples but their normalizations are obtained from the fit. The \( W + \text{jets} \) normalization is also fitted because \( W + \text{jets} \) is the Monte Carlo sample with the highest theoretical uncertainty on its normalization. Typically the resulting normalization is compatible with the theoretical one within few percent.

The fit is done in the range 15 GeV < \( E_T^{\text{miss}} \) < 100 GeV. In these analysis it is not possible to extend the interval to 0 GeV because the data used have a preselection which rejects most of the events with \( E_T^{\text{miss}} \) < 15 GeV. The preselection is not applied in the Monte Carlo samples, thus a comparison of data with Monte Carlo is meaningless in the interval 0 GeV < \( E_T^{\text{miss}} \) < 15 GeV.

The scale factors from the fit to the \( E_T^{\text{miss}} \) distribution are then used to normalize the multijet QCD and \( W + \text{jets} \) samples in the dijet invariant mass distribution after requiring
$E_T^{\text{miss}} > 25 \text{ GeV}$. The error on the fraction of multijet QCD events resulting from the fit is of the order of 5%. Fig. 4.6 shows the $E_T^{\text{miss}}$ distribution once the fit is performed for both the electron and muon channel at the step of the selection with at least two jets in the event. Data distributions agree with Monte Carlo ones within 5% in the interval [15, 100] GeV. For larger $E_T^{\text{miss}}$ values however the distributions are compatibles within the statistical error. Section A.4 contains the $E_T^{\text{miss}}$ distributions for each step of the CDF resonance selection for both the lepton channels, while section B.5 collects those for the WW/WZ selection.
Fig. 4.5: $E_{T}^{miss}$ distributions for $W + jets$ (green) and multijet QCD (grey) samples at the end of the CDF resonance selection (top, see section 5.1 for details) and at the end of the WW/WZ Sel1 (bottom, see chapter 6 for details). Plots (a,c) are obtained from the electron selection and plots (b,d) from the muon one.
Fig. 4.6: $E_{\text{miss}}^{T}$ distributions after the $M_T$ cut and requiring at least two jets in the event for the electron selection (a) and for the muon one (b). The distribution for data (dots) is superimposed to the expected SM predictions (filled area). The statistical error is reported for data (black crosses) and for SM expectations (dashed violet lines). The panel underneath displays the bin-by-bin ratio of data to Monte Carlo with error bars corresponding to the statistical error in data and MC.
Chapter 5

Search for a jet-jet resonance

The resonance measured by the CDF experiment at dijet mass $147 \pm 4 \, \text{GeV}/c^2$ could represent the evidence of physics not predicted by Standard Model, so it is important to investigate the presence of this signal in other experiments. The same resonance should also be produced in proton-proton collisions at LHC therefore the ATLAS collaboration has searched for this signal firstly in $33 \, \text{pb}^{-1}$ of integrated luminosity [68] and recently in $1.02 \, \text{fb}^{-1}$ [40]. The analysis described in this chapter is the search for the dijet resonance in $1.02 \, \text{fb}^{-1}$ of data applying a selection as close as possible to that one used by CDF. The kinematic conditions at LHC differ from those at the Tevatron both because of the higher center-of-mass energy and of the different colliding particle nature. For these reasons the production cross-sections of all processes and therefore the contributions to the jet-jet invariant mass are also different with respect to what is observed at Tevatron.

In this first analysis phase no model has been hypothesized for the resonance production therefore no expectation can be calculated for the fraction of the resonant events produced at LHC. As a result, the sensitivity of the ATLAS experiment to the dijet resonance is unknown.

This chapter contains the peculiar selections used to control the SM background in the region of the CDF resonance. It is given a description of the main systematic uncertainties which affect the measurement and of the statistical method adopted to assess the presence of the resonance. The data and MC samples, the selection of the leptonic $W$ and of the jets and the method to estimate the multijet QCD background are described in chapter 4.

5.1 CDF selection

At present, there is not a unique interpretation of the dijet resonance observed by CDF therefore to keep the analysis as much as possible model-independent the kinematic selection adopted here closely follows the one applied by CDF.

Fig. 5.1 shows the jet multiplicity\(^1\) distributions for events that pass all the selections described in chapter 4 up to the $M_T$ cut. The $W + 0\text{partons}$ events have the lowest multiplicity and many of them are removed when a jet multiplicity larger than 2 is required. The $t\bar{t}$ sample has, as expected, the highest multiplicity. The agreement between data and SM expectation is within 15% with the exception of the bins with low statistics. The Jet

\(^1\)In the jet multiplicity are considered only those jets which pass the requirements in section 4.2.4.
Energy Scale uncertainty, described in section 5.2, has a large effect on these distributions since the number of jets which pass the \( p_T > 30 \text{ GeV}/c \) cut changes significantly with the Jet Energy Scale value. Considering the systematic error due to the Jet Energy Scale uncertainty, the data and MC jet multiplicity distributions are compatible (Fig. B.27).

The jet-jet invariant mass distributions are computed with the two jets in the event which have the larger \( p_T \). The dijet mass distributions are made for all events with two or more jets (allowed by the increased jet activity at LHC energies compared to the Tevatron), and also with the CDF requirement of exactly two jets. A set of further requirements have been used to better control the SM prediction especially in the region of the resonant signal as it was done in the CDF selection:

- transverse momentum of the dijet system \( p_{Tjj} > 40 \text{ GeV}/c \);
- jet separation in \( \eta \): \( |\Delta \eta(j_{1st}, j_{2nd})| < 2.5 \);
- azimuthal angular separation between the leading jet and the missing transverse energy direction \( \Delta \phi(j_{1st}, E_T^{miss}) > 0.4 \).

The transverse momentum cut on the dijet system aims at selecting only boosted jet-jet systems. This requirement models the shape of the SM prediction so that the shape in the region of the signal is smoothly decreasing. The \( \Delta \eta(j_{1st}, j_{2nd}) \) cut rejects back to back events. Both selections are justified by the idea that the dijet system should balance the leptonic \( W \) decay products so the two jets are expected to be boosted with small opening angle. The \( \Delta \phi(j_{1st}, E_T^{miss}) > 0.4 \) cut rejects some of the multijet QCD background because in these events the \( E_T^{miss} \) is mainly generated from a mismeasurement of the energy of a jet, thus it tends to point in the direction of the leading jet.
5.1. CDF SELECTION

<table>
<thead>
<tr>
<th>Sample</th>
<th>Electron ch.</th>
<th>Muon ch.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$N_{jets} \geq 2$</td>
<td>$N_{jets} = 2$</td>
</tr>
<tr>
<td>WW/WZ</td>
<td>1.2</td>
<td>1.3</td>
</tr>
<tr>
<td>$t\bar{t}$</td>
<td>13.2</td>
<td>4.8</td>
</tr>
<tr>
<td>Single top</td>
<td>1.2</td>
<td>1.3</td>
</tr>
<tr>
<td>$Z + jets$</td>
<td>5.6</td>
<td>5.7</td>
</tr>
<tr>
<td>$W + jets$</td>
<td>72.1</td>
<td>79.9</td>
</tr>
<tr>
<td>Multijet QCD</td>
<td>6.6</td>
<td>7.1</td>
</tr>
</tbody>
</table>

Table 5.1: Fraction of expected events in data for each Standard Model process at the end of the selection. The cases with $N_{jets} = 2$ and $N_{jets} \geq 2$ are displayed for both the electron and muon channels.

No further selection is applied but the events with at least two jets in the final state and exclusively two jets are both investigated. Fig. 5.2 and 5.3 show respectively for the electron selection and for the muon one the distributions of $p_{Tjj}$, $\Delta\eta(j_{1st}, j_{2nd})$, $\Delta\phi(j_{1st}, E_T^{miss})$ and the jet multiplicity each one displayed just before cutting on the variable. Data agree with SM predictions within 10% except bin with low statistics. A light mismodeling is noticed in the $p_{Tjj}$ distributions for $p_{Tjj} < 60 \text{ GeV/c}$ and in the $\Delta\phi(j_{1st}, E_T^{miss})$ ones for $\Delta\phi(j_{1st}, E_T^{miss}) < 1.5$. The invariant mass distributions of the two leading jets at the end of the cut-flow are shown in Fig. 5.4 (Fig. 5.5) for the electron channel (muon channel). For each channel two distribution are obtained: one is made with all the events with at least two jets $N_{jets} \geq 2$; the other one with events with exactly two jets ($N_{jets} = 2$). A complete evolution of the jet-jet invariant mass through each step of the selection is provided in Fig. A.8 and A.9 of section A.5. Table 5.1 shows the expected fractions of Standard Model processes in the jet-jet invariant mass distributions of Fig. 5.4 and 5.5.
Fig. 5.2: $p_T^{jj}$ (a), $\Delta \eta(j_{1st}, j_{2nd})$ (b), $\Delta \phi(j_{1st}, E_T^{miss})$ (c) and jet multiplicity (d) distributions obtained for the electron selection just before a cut is applied on these variables. Plot (d) has logarithmic vertical scale. For each plot, the distribution for data (dots) is superimposed to the expected SM predictions (filled area). The statistical error is reported for data (black crosses) and for SM expectations (dashed violet lines). The panel underneath displays the bin-by-bin ratio of data to Monte Carlo with error bars corresponding to the statistical error in data and MC.
Fig. 5.3: $p_{T;jj}$ (a), $\Delta \eta(j_{1st},j_{2nd})$ (b), $\Delta \phi(j_{1st},E_T^{miss})$ (c) and jet multiplicity (d) distributions obtained for the muon selection just before a cut is applied on these variables. Plot (d) has logarithmic vertical scale. For each plot, the distribution for data (dots) is superimposed to the expected SM predictions (filled area). The statistical error is reported for data (black crosses) and for SM expectations (dashed violet lines). The panel underneath displays the bin-by-bin ratio of data to Monte Carlo with error bars corresponding to the statistical error in data and MC.
Fig. 5.4: Jet-Jet invariant mass distributions in the electron channel at the end of the CDF selection for events with at least two jets (a) and only two jets (b). For each plot, the distribution for data (dots) is superimposed to the expected SM predictions (filled area). The statistical error is reported for data (black crosses) and for SM expectations (dashed violet lines). The panel underneath displays the bin-by-bin ratio of data to Monte Carlo with error bars corresponding to the statistical error in data and MC.
Fig. 5.5: Jet-Jet invariant mass distributions in the muon channel at the end of the CDF selection for events with at least two jets (a) and only two jets (b). For each plot, the distribution for data (dots) is superimposed to the expected SM predictions (filled area). The statistical error is reported for data (black crosses) and for SM expectations (dashed violet lines). The panel underneath displays the bin-by-bin ratio of data to Monte Carlo with error bars corresponding to the statistical error in data and MC.
5.2 Systematic uncertainties

The analysis takes into account the most important systematics which affect the dijet mass distribution. The main uncertainties come from:

1. the Jet Energy Scale (JES);
2. the Jet Energy Resolution (JER);
3. the reconstruction and identification scale factors of the selected lepton;
4. the energy or transverse momentum resolution of the lepton selected;
5. the normalization of the multijet QCD sample due to the chosen interval in which the fit is performed.

The jet systematic uncertainties considered are those on the JER and on the JES. The JES uncertainty is the overall systematic error resulting from uncertainties on calorimeter scale, dead material description, cluster reconstruction, fragmentation and Underlying Event modeling as well as pile-up [54, 55]. The effect of this systematic on the dijet mass distribution has been evaluated looking at how this distribution changes shifting the JES up and down of one standard deviation from the central value. The JES uncertainty is the main systematic error which affects the analysis and its effect on the dijet mass distribution is reported in Fig. 5.6. The uncertainty on the $M_{jj}$ distribution is estimated as the bin-by-bin maximum difference of the shifted distributions with respect to the nominal one.

The JER uncertainty is the measurement of how different is the resolution on the jet energy in data and in Monte Carlo, and it is evaluated using dijet QCD events. The energy resolution of all the MC reconstructed jets is worsened of one uncertainty. The $M_{jj}$ distribution is then compared with the one obtained with the nominal energy resolution and their difference is taken as the systematic error on the $M_{jj}$ distribution.

Scale factors have been used to reweigh the MC events in order to recover the lepton reconstruction and identification efficiencies observed in data. Two $M_{jj}$ distribution have been obtained shifting the scale factors up and down of their uncertainty.

In Monte Carlo samples the electron energy is smeared to reproduce the resolution observed in data. The uncertainty on the smearing is a systematic which has been treated. The smearing parameters has been first shifted up and then down of their uncertainties obtaining new energy values for each electron. The muon is treated as well as the electron through the shift of the momentum smearing parameters. The selections have been applied again and the $M_{jj}$ distributions have been recomputed. The systematic error on the $M_{jj}$ distribution due to the smearing and scale factor uncertainties is estimated as the bin-by-bin maximum difference of the shifted distributions with respect to the nominal one.

The multijet QCD normalization has been extrapolated from the fit to the $E_T^{miss}$ distribution as explained in section 4.3. The interval where the fit is performed has been changed and data without any $E_T^{miss}$ preselection has been chosen to extend the fit up to 0 GeV. The aim is to understand how the choice of the fit range changes the multijet QCD normalization and the $M_{jj}$ distribution. The fit is done in three intervals: [0, 75], [0, 100] and [0, 125] GeV. As systematic error is taken the bin-by-bin maximum difference of the $M_{jj}$ distributions obtained from the three fits.
The CDF resonance is searched on the $M_{jj}$ distribution which is the sum of two obtained from the electron selection and the muon one (see section 5.3). Therefore, the lepton uncertainties have been propagated to the dijet mass plot of the sum of the electron and muon channels considering the uncertainties as uncorrelated. The jet uncertainties are instead completely correlated for the two channels so it has been obtained as the linear sum of uncertainties in the single channel distributions. The overall systematic uncertainty in each bin is obtained as the sum in quadrature of all the uncertainty coming from the different sources and is shown for $N_{jets} \geq 2$ ($N_{jets} = 2$) in the lower panel in Fig. 5.7(a) (Fig. 5.7(b)).
Fig. 5.6: Jet-jet invariant mass distributions for the electron selection (left) and muon selection (right) for the case of $N_{jets} \geq 2$ (top) and $N_{jets} = 2$ (bottom). For all plots, the distributions obtained by shifting the JES up (red) and down (blue) by its uncertainty are compared to the distribution obtained with nominal JES (black). The black dots correspond to data. The errors bars are the statistical errors. The panel underneath displays the bin-by-bin ratio of data to Monte Carlo (red crosses) with error bars corresponding to the statistical error in data and MC and the cyan bands represent the relative systematic uncertainty due to the JES uncertainty estimated as the maximum difference between the shifted bin values and the unshifted one.
5.3. ANALYSIS OF THE RESULTS

5.3 Analysis of the results

The invariant mass distribution obtained as combination of the two channels is processed with a statistical analysis whose role is the identification of the most significant discrepancy between data and Monte Carlo backgrounds. The mass distributions analysed are those obtained after the whole selection with \( N_{jets} = 2 \) and with \( N_{jets} \geq 2 \) and are shown respectively in Fig. 5.7(a) and Fig. 5.7(b).

The first step of the analysis consists in fitting the data with the background distribution which is left free to vary within the systematic errors in order to best describe the data. To each of the five sources of systematic uncertainty (section 5.2) is associated a nuisance parameter whose variation, shifts the background mass distribution within the uncertainty of that source. It is assumed that the uncertainties are distributed as Gaussian distributions and are fully correlated between bins. The set of parameter values which maximizes the likelihood of data to be the result of the background-only hypothesis is considered as the best fit.

The region considered for the fit covers the dijet invariant mass interval \( 100 < M_{jj} < 300 \, \text{GeV}/c^2 \). This is similar to the range explored by CDF, and avoids the low mass region affected by the diboson contribution. The upper plots in Fig. 5.8 show the comparison of data and background description as obtained from the fit. Both cases of 2 jets in the event and at least 2 jets in the event are shown. This procedure provides a conservative upper limit on the existence of a possible resonance because the systematics are set to compensate part of an eventual discrepancy due to the resonance. The lower plots in Fig. 5.8 show the difference between the data and the background prediction in each bin in term of standard deviations. The significance is determined by calculating in each bin the probability to obtain a deviation as extreme as that one observed (p-value). In the case of an excess of data in a given bin, the p-value represents the probability to see an excess of that size or larger, while for a deficit of data, the p-value gives the probability of seeing a deficit so large or larger. The p-value in each bin is then converted in standard deviations (\( \sigma \)) by integrating a Gaussian distribution with width equal to the statistical error. The significance is plotted as positive for an excess of data, negative otherwise.

A global likelihood test is used as goodness-of-fit statistical test to determine the agreement between the data and the conditioned background. It has been calculated as the product of the likelihood in each single bin. A set of pseudo-experiments obtained by the variation of background within the statistical error is used to determine the probability to obtain a lower likelihood than that measured.

The global p-value for the dijet mass distribution is 88\% for a number of jets \( N_{jets} = 2 \) and 65\% for \( N_{jets} \geq 2 \), indicating overall agreement between the data and the background prediction. This test has not optimal sensitivity to localized excesses at mass interval with lower statistics because high statistics bins have an high weight. Furthermore, this method do not consider whether bin-by-bin fluctuations go in the same direction, or randomly swing up and down as expected from statistical fluctuation.

The BumpHunter algorithm [69], [70] is a more sensitive test to search for the potential contributions from a resonance in an interval. It searches for a window of at least 4 bins with the most significant excess in data above background. The window is shifted in position and increased until it exceeds half the search region. Four bins correspond to the approximate resolution (20 GeV/c\(^2\)) expected for a resonance with mass 145 GeV/c\(^2\).
This resolution is obtained considering that the width of the $W$ boson decay in jets, as measured in $t\bar{t}$ events, is about 15% of $M_W$. Assuming the same relative resolution, the width of the resonance should be $22 \text{ GeV/c}^2$.

The most significant departure from background is defined by the set of consecutive bins that have the smallest probability of arising from the background-only hypothesis assuming Poisson statistics. The probability is corrected for the look-elsewhere effect which depends on the dimension of the interval where the resonance is searched. The larger is the interval, the more is the probability of having a fluctuation of at least 4 bin as significant as that searched by BumpHunter. This probability is obtained on the basis of a series of pseudo-experiments.

The most significant discrepancy identified by the BumpHunter algorithm, in the $M_{jj}$ distribution for $N_{jets} = 2$, is a 9-bin excess in the dijet mass interval $M_{jj} = 225 - 270 \text{ GeV/c}^2$, as shown in Fig. 5.8. The p-value of observing an excess at least as large as this is 37.1% which corresponds to 0.33 $\sigma$. For $N_{jets} \geq 2$ the most significant window ($M_{jj} = 160 - 265 \text{ GeV/c}^2$) exceeds half of the mass interval and has a corrected p-value of 17.1% (0.95 $\sigma$). Considering the region of the CDF excess, $M_{jj} = 120 - 170 \text{ GeV/c}^2$, and neglecting the look-elsewhere effect, the minimum p-value is 29.8% (0.53 $\sigma$) in the $N_{jets} = 2$ analysis, and in that with $N_{jets} \geq 2$ it is 42.4% (0.19 $\sigma$).

As final test, the BumpHunter algorithm is used in the interval $M_{jj} < 100 \text{ GeV/c}^2$ to determine the sensitivity of this analysis to the $WW/WZ$ signal. First, from the background template is subtracted the $WW/WZ$ contribution, then is searched a interval with the maximum discrepancy data/backgrounds. No significant excess is identified as expected. In fact, the fraction of $WW/WZ$ is about 1% in this analysis and it is compensated by the statistical and systematic uncertainties.

To summarize, the data sample analysed do not present significant excess but it is not possible to predict how much is the signal expected. In fact, without any hypothesis on the production mechanism, the cross-section at LHC cannot be estimated and compared to that measured at CDF.
5.3. ANALYSIS OF THE RESULTS

Fig. 5.7: Dijet mass distribution for electron and muon events combined, with $N_{jets} \geq 2$ (a) and $N_{jets} = 2$ (b). The data points are plotted (full circles) with the statistical uncertainties. The blue solid line is the total Standard Model prediction and the yellow hatched band indicates the systematic uncertainties calculated as described in the text. The colored histograms indicate the expected sample composition, and are stacked in order of increasing number of events, except for the $WW/WZ$ contribution, which is plotted last for clarity. Lower panel: data over total Standard Model prediction. The error bars and the yellow hatched band indicate the statistical and systematic uncertainties respectively [40].
Fig. 5.8: Finely binned comparison of the dijet invariant mass obtained in data with $N_{jets} \geq 2$ (a) and $N_{jets} = 2$ (b), to a background hypothesis setting nuisance parameters to their maximum likelihood estimate (top), and the resulting statistical bin-to-bin significances (bottom). The mass interval with the most significant excess found by the BumpHunter algorithm is shown by blue vertical lines [40].
Chapter 6

Sensitivity studies on $WW/WZ$ resonance

As introduced in chapter 1, studies of $WW/WZ$ processes constitute a test of Standard Model as well a contribution to the understanding of background for other analysis. The analysis I took part to has the aim to measure the cross-section of $WW$ and $WZ$ decays in the semileptonic channel:

$$WW/WZ \rightarrow l \nu jj$$

where $l$ stands for electron or muon and $j$ for jet. The inclusive $W$ production in association with jets ($W + jets$) is the dominant source of background in events containing a lepton, missing transverse energy and at least two jets. Other contributions are multijet QCD events, electroweak processes including $Z \rightarrow ll + jets$, $t\bar{t}$ and single top decays. A description of each background considered is given in section 4.1.1.

The cross-section is obtained by the fraction of $WW/WZ$ events evaluated fitting the jet-jet invariant mass obtained from collision data with the expected signal and background, estimated with simulated events and data-driven techniques. The $WW$ and $WZ$ signals can not be distinguished in the jet-jet invariant mass distribution because the expected resolution of the mass of a reconstructed $W$ from jets is about $11 \text{ GeV}/c^2$ ($FWHM \sim 25 \text{ GeV}/c^2$) and similar for the $Z$ and the difference in mass between $W$ and $Z$ is $\sim 11 \text{ GeV}/c^2$. Therefore the fitting procedure extracts the sum of the number of events from the $WW$ and $WZ$ processes. The $WW/WZ$ cross-section in the semileptonic channel has been measured at Tevatron by CDF [30] and D0 [31]. These analysis found a production cross-section for the signal of about $18 - 20 \text{ pb}$ compatible with the theoretical prediction of $15.2 \text{ pb}$ [31]. At LHC it is expected to be $63 \text{ pb}$ [18, 29] due to the larger $\sqrt{s}$ than Tevatron. Despite the signal is 4 times bigger, at LHC, it will be more challenging the measurement of the cross-section since the main background, the $W + n \text{jets}$ ($n = 2, 3, 4, 5$), grows about 20 times [32, 33]. This has brought to the study of an optimized selection which could allow to measure the $WW/WZ$ cross-section even if the initial ratio of signal ($S$) to background ($B$) is 5 times worse than at Tevatron.

This chapter describes the selection that I developed as possible solution to measure the $WW/WZ$ signal. It was studied with the intention to maximize the $S/\sqrt{B}$ ratio and to raise the $S/B$ one. To obtain this, I studied the kinematic of the signal and background events and their difference, and I applied selections to reduce the dominant backgrounds.
which are $W + \text{jets}$, multijet QCD and $t\bar{t}$. An alternative selection has been developed to
the improve more the $S/B$ ratio paying some points in $S/\sqrt{B}$ ratio. I have also developed
specific selections to isolate regions enriched of a particular background process, indicated
in the following as control regions. The data to Monte Carlo agreement in the control
region is used as check of the goodness in simulating the background selected. At last, the
effect of the most relevant systematic uncertainties on the measurement is estimated.

6.1 Data samples and selection

This analysis uses the samples and the selection described in chapter 4. Further cuts
shown in this chapter will be added to the selection. The multijet QCD is derived by the
data-driven method discussed in section 4.3. The amount of data analysed these $WW/WZ$
studies is 1.33 $fb^{-1}$.

6.1.1 Selection of hadronic $W$ decay candidates

The hadronic $W$ decay products are selected as the two leading jets having $p_T > 30$ GeV/c
and $|\eta| < 2.8$. The jet selection for this analysis has been carried out with the most recent
dataset available and this has allowed to use a new variable, not used in the previous
chapter, to identify the jets produced by the pile-up interactions. This new variable
is termed Jet Vertex Fraction (JVF) [71] and is obtained using the information of the
reconstructed tracks in the Inner Detector. This information is therefore available only
for jets having $|\eta| < 2.1$. The JVF is defined as the percentage of transverse momentum
carried by the tracks pointing to the jet and originating from the primary vertex with
respect to the total transverse momentum carried by tracks pointing to the jet. The
JVF is required to be larger of 75% for all the considered jets. This condition allows to
largely suppress jets originating from pile-up interaction vertexes. This does not correct
completely for the effect of pile-up in fact in events where a jet is originating from the
primary vertex its energy can be modified for the presence of some activity due to pile-up.
This effect is addressed by part of the systematic uncertainty on the Jet Energy Scale. For
the rest of the analysis, the jets used are those which pass the selection in section 4.2.4
and in addition the JVF cut unless differently specified.

Fig. 6.1 shows the multiplicity of jets passing the selection in data and MC for the
electron and the muon channels. The data and MC jet multiplicity distributions are com-
patible within the systematic error due to the Jet Energy Scale uncertainty (see Fig. B.27).
For the MC the fractions of signal events and $t\bar{t}$ ones are also shown as a function of the
number of jets. The $t\bar{t}$ background has the largest multiplicity and most of these events
have 3 – 4 jets. Indeed, a large part of them decay into one lepton, one neutrino, two
jets from $W$ and two jets from $b$-quarks. About 70% of the times, the signal has less
than two jets. Some of these losses are due to events where the $WW$ and $WZ$ decays
fully leptonically but more frequently one or both jets from the hadronic $W/Z$ decay are
rejected by the jet selection. A possible way to reduce these losses is to decrease the $p_T$
threshold. However in this case both the signal and backgrounds would increase. Moreover
the systematic uncertainty due to the Jet Energy Scale would raise. In section 6.3 the
effects of a decreased $p_T$ threshold are discussed.
6.1. DATA SAMPLES AND SELECTION

Fig. 6.1: Jet multiplicity distributions after the $M_T$ cut for electron (left) and muon (right) selections. Only the jets which pass the jet selection (section 4.2.4) and the JVF cut are counted. On the top: the distribution for data (dots) is superimposed to the expected SM predictions (filled area). The statistical error is reported for data (black crosses) and for SM expectations (dashed violet lines). The vertical scale is logarithmic. The panel underneath displays the bin-by-bin ratio of data to Monte Carlo with error bars corresponding to the statistical error in data and MC. On the bottom: distributions for signal events (red) and $t\bar{t}$ background ones (blue). The two distributions are normalized to unitary area.
Fig. 6.2: Invariant mass distributions of the two leading jets (blue) and of the two with lowest 
$p_T$ (red) for $WW$ events. The events considered satisfy the electron selection and have $N_{jets} \geq 3$. 

When an event has more than two jets, it presents the problem of the choice of the 
couple of jets used for the computation of the jet-jet invariant mass. Fig. 6.2 shows for 
$WW$ events with at least three jets the $M_{jj}$ distribution choosing the two leading jets in 
the event and the two with lowest $p_T$. The choice of the non leading jets seems find the 
right combination of jets from $W$ more times, however the overall gain of events in the peak 
of the signal considering all the events with at least two jets is of the order of 1%. Despite 
the small loss of events in the signal peak, for this analysis I have used the two leading 
jets in the events because they have the lowest relative systematic uncertainties. Further 
studies should be done to find a selection criterion which improves more the number of 
good combinations. 

The mass distributions obtained selecting the two leading jets are shown in Fig. 6.3, 6.4. 
The plots in Fig. 6.3 display the distributions for data and for the SM predictions and 
the ratio of the two. The agreement between data and SM expectation is within 5%. The 
invariant mass distributions are dominated by the $W + jets$ background which peaks in the 
region from 60 to 150 $GeV/c^2$. Therefore, the signal is located right in the region where 
the background has the maximum magnitude and maximum slope variation (Fig. 6.4). 
The $W + jets$ constitutes the 71 − 73% of the events while the signal only the 1%. The 
second background in order of magnitude is the multijet QCD which is larger in muon 
channel (14.3%) than in the electron one (9.6%). Jets in the multijet QCD sample fakes
6.1. DATA SAMPLES AND SELECTION

<table>
<thead>
<tr>
<th>Sample</th>
<th>Fraction of SM predictions in data (%)</th>
<th>Electron ch.</th>
<th>Muon ch.</th>
</tr>
</thead>
<tbody>
<tr>
<td>WW/WZ</td>
<td>1.05</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>$t\bar{t}$</td>
<td>10.05</td>
<td>9.31</td>
<td></td>
</tr>
<tr>
<td>single top</td>
<td>1.19</td>
<td>1.11</td>
<td></td>
</tr>
<tr>
<td>$Z + jets$</td>
<td>4.77</td>
<td>3.23</td>
<td></td>
</tr>
<tr>
<td>$W + jets$</td>
<td>73.3</td>
<td>71.0</td>
<td></td>
</tr>
<tr>
<td>multijet QCD</td>
<td>9.62</td>
<td>14.3</td>
<td></td>
</tr>
</tbody>
</table>

Table 6.1: Expected fractions of the SM processes in data for events with $N_{jets} \geq 2$.

...more easily the muon selection mainly because of the lower muon $p_T$ threshold with respect to the electron one. In the interval from 20 to 25 GeV/c in the muon $p_T$ distribution, the multijet QCD constitutes half of the background as shown in Fig. 6.31(a). The $t\bar{t}$ is the third source of background contributing about 10% in both channels. A summary of the contribution of the various background samples for the electron and the muon channels is shown in Table 6.1.
Fig. 6.3: Invariant mass distributions of the two leading jets for the electron (a) and muon (b) selections. For each plot, the distribution for data (dots) is superimposed to the expected SM predictions (filled area). The statistical error is reported for data (black crosses) and for SM expectations (dashed violet lines). The panel underneath displays the bin-by-bin ratio of data to Monte Carlo with error bars corresponding to the statistical error in data and MC.
6.2 Optimization of the selection

The aim of the study which I developed is the definition of a selection which improves the measurement of the $WW/WZ$ signal through the fitting procedure. I mainly tried to improve the following two factors which contribute to the success and feasibility of the fit:

- the signal yield in data in order to be large enough with respect to the background fluctuation. I have concentrated my studies on improving the signal significance\(^1\) and possibly also the signal to noise ratio. However a large statistical significance it is not enough to detect the signal in fact the effect of the systematic uncertainty should also be considered. The selection I developed does not study in details the relation between each single cut and the systematic uncertainty. But in general, the increase of the $S/B$ ratio also improve the signal magnitude with respect to the systematic error. In this sense the selection also tries to improve the signal sensitivity with respect to the systematic uncertainty. I have also developed another selection aimed at improving efficiently the $S/B$ ratio but which leads to a lower $S/\sqrt{B}$ ratio (section 6.8).

- the smoothness of the background distribution in the signal region. If the background has a smooth shape in the signal region the fit procedure, used for the extraction of

\(^1\)In this thesis with statistical significance of the signal in a interval it is meant the improbability of the background to fluctuate in such a way, that it appears at least as large as the signal in the same interval. Since the study here presented proposes candidate selections that could allow the measurement of the $WW/WZ$ signal but do not measure it, the signal significance is not calculated but an approximation of it is used. The ratio of the signal to the square root of the background provides an estimate of the significance expressed in standard deviations. A significant improvement of the $S/\sqrt{B}$ ratio entails the improvement of the significance too. Therefore, it can be said that if the selections improve the $S/\sqrt{B}$ ratio they improve the significance too.
CHAPTER 6. SENSITIVITY STUDIES ON WW/WZ RESONANCE

the fraction of signal and background, is less affected by uncertainties. In fact the variation of the background distribution is less distorted by the systematic uncertainty and is less sensitive to the statistical fluctuations. Moreover a different shape between signal and background helps in discriminating one from the other.

6.2.1 Selection strategy

The selection is improved through a careful study of the signal and background kinematic differences. The jet, lepton and $E_T^{\text{miss}}$ distribution are investigated in order to find those variables in which there are significant differences in the shape of the signal and background distributions. Once one of these distribution is identified, a cut is applied to raise the $S/B$ and $S/\sqrt{B}$ ratios.

This method of selection is heavily based on Monte Carlo simulations and data plays a role only in determining the multijet QCD samples and the $W + \text{jets}$ normalization. Therefore, for each variable investigated the distribution obtained summing the Monte Carlo samples reweighed for the integrated luminosity is compared with data to verify the overall agreement of the two. In the plots which show data superimposed to Monte Carlo prediction the color conventions are those specified in section 4.1.2. As further check, for some distributions the ratio of data to predictions or the residual\(^2\) in each bin of the distribution is plotted. If data and MC do not present significant disagreements, the signal and backgrounds distributions are considered well modeled by Monte Carlo and a cut on them is applied. This check prevents the introduction of further systematic error.

The choice of the variables on which the cut should be applied is based on the following parameters:

- the maximum global $S/\sqrt{B}$ ratio achievable cutting on the distribution and the corresponding $S/B$ ratio;
- the maximum $S/\sqrt{B}$ ratio achievable cutting on the distribution considering the events in the $M_{jj}$ signal region\(^3\) and the corresponding $S/B$ ratio in the same region;
- the change of the $M_{jj}$ background distribution as a function of the cut looking in particular at how the distribution becomes smooth in the region of signal as consequence of the cut.

Given a distribution of a variable, the cut that maximizes the statistical significance is found by an algorithm. The algorithm explores the interval where the distribution is plotted and searches for those consecutive bins where the $S/\sqrt{B}$ ratio is maximum. The algorithm returns the cut which maximize the $S/\sqrt{B}$, the maximum $S/\sqrt{B}$ ratio and the corresponding $S/B$ ratio. The error on the number of signal events $S$ and background one $B$ in the interval is the square root of the sum in quadrature of the errors in each bin. The bin error is obtained as the square root of the sum of the weight, used to fill the bin, squared. The errors on the $S/\sqrt{B}$ and on the $S/B$ ratios come from the propagation of the errors on $S$ and $B$.

\(^2\)The residual is the bin-by-bin difference between the data and the MC prediction divided by the standard deviation.

\(^3\)In this thesis for signal or peak region it is meant the interval in the $M_{jj}$ distribution where the $S/\sqrt{B}$ ratio is maximum.
The method used to measure the maximum $S/\sqrt{B}$ and $S/B$ ratios in the $M_{jj}$ signal region is similar to that explained above but it considers 2D histograms of the $M_{jj}$ versus the variable to cut for the signal and the background. An algorithm on these 2D distributions finds the rectangle of bin which has the maximum $S/\sqrt{B}$ ratio and returns the value of the cut to apply, the interval in $M_{jj}$ which identifies the peak region, the maximum $S/\sqrt{B}$ ratio and the $S/B$ ratio. Whatever is the variable where the algorithm is applied, the interval of mass with the maximum $S/\sqrt{B}$ ratio obtained slightly changes: the lower limit varies from 60 to 65 GeV/$c^2$; the upper one from 105 to 110 GeV/$c^2$. Therefore for each cut the peak region is practically steady.

The 2D distributions offer also the possibility to understand which is the $M_{jj}$ region in signal and in background affected by the cut and how the cut shapes the two distributions. Thank to this is possible to avoid selections that improve the $S/\sqrt{B}$ ratio but constraint the kinematic of the background to resemble that of the signal. As a consequence, doing these cuts, the $M_{jj}$ distributions of the signal and background become very similar decreasing the fitting power.

The selection applied in the next section has the role to adjust the background shape in the $M_{jj}$ region of the signal. After it will be applied a series of requirements which aim mainly at maximize the $S/\sqrt{B}$ ratio improving also the $S/B$ ratio in the signal region without worsening the background shape.

The $S/\sqrt{B}$ and $S/B$ ratios for the full jet-jet invariant mass distributions are respectively $3.70 \pm 0.03$ and $(1.06 \pm 0.01)$% for the electron channel and $4.26 \pm 0.03$ and $(1.03 \pm 0.01)$% for the muon one. The $M_{jj}$ distributions in Fig. 6.3 have a significance $S/\sqrt{B} = 4.04 \pm 0.04$ in the interval $[60, 110]$ GeV/$c^2$ for the electron selection and $S/\sqrt{B} = 4.59 \pm 0.04$ in the interval $[60, 110]$ GeV/$c^2$ for the muon one. The significance indicates that the signal peak has very low probability ($\sim 10^{-5}$) to be the result of the statistical fluctuation of the background in that region, nevertheless it seems that data are not sensitive to the signal. This is due to the systematic uncertainty which is not taken into account in the significance calculation. The $S/B$ ratio in the peak region is $(2.16 \pm 0.03)$% for the electron channel and $(2.02 \pm 0.02)$% for the muon one. For a systematic error of the order of the 2%, the signal is compatible with the background error hardening the possibility to measure it in data. In section 6.4 the systematic uncertainty on the jet-jet invariant mass distribution due to JES uncertainty will be shown and its magnitude will be compared to that of the signal at this step of the selection (Fig. B.26) and at the end of it (Fig. 6.23).

6.2.2 Background shaping

The dijet invariant mass distributions in Fig. 6.4 are characterized by a background which has maximum magnitude in the region of the signal. In addition, in that region the background has the maximum shape variation, hence any systematics or mismodeling has an amplified effect there. Furthermore, as mentioned at the beginning of this chapter, the fit works better when the signal and the background have different shapes. The mentioned features have called for a selection which shapes the background making it more regular in the signal region. The selection which seems to have the best capacity to smooth the background shape, is a cut on the $p_T$ of the selected dijet system ($p_{T_{jj}}$). The dijet $p_T$ distribution is plotted in Fig. 6.5 for the electron channel. This section
contains only distributions obtained for the electron selection but the comments made on them remain valid also for the distributions obtained for the muon selection. These latter are presented in section B.1 (Fig. B.1). Fig. 6.5(a) displays the distribution for data and for SM processes. The distributions are compatible and the agreement is improved with respect to the same distribution obtained without the JVF requirement and the $W + HF$ samples (Fig 5.2(a)). For the muon channel there is less agreement between data and Monte Carlo predictions, with data distribution that exceeds MC one in the region $20 \text{ GeV}/c^2 < M_{jj} < 40 \text{ GeV}/c^2$ (Fig. B.1(a)). However, once the multijet QCD is removed as shown in section 6.8, the discrepancies between data and SM expectations disappear (Fig B.11). The plot (b), in Fig. 6.5, compares the $p_{Tjj}$ distribution for signal and background normalized to the same area. This plot shows that the dijet system in signal events tends to be slightly more boosted than in background events. In the range $[60, 200] \text{ GeV}/c$ the distribution decreases monotonically, while immediately below $60 \text{ GeV}/c$ it has a relative minimum. This shape is due to the jet $p_T$ cut which is set to $30 \text{ GeV}/c$. If no cut is applied on the jet $p_T$, the dijet $p_T$ distribution should decrease monotonically on the whole interval $[0, 200] \text{ GeV}/c$. Instead when jets have $p_T > 30 \text{ GeV}/c$, the interval $p_{Tjj} < 60 \text{ GeV}/c$ loses the events which have close jets. The lower is the dijet $p_T$, the larger is the minimum $\phi$ angle allowed between the selected jets ($\Delta \phi(j_{1st}, j_{2nd})$) as shown in Fig. 6.6. This loss generates the relative minimum in the $p_{Tjj}$ distribution at about $60 \text{ GeV}/c$.

The removal of events with close jets in the region with dijet $p_T < 60 \text{ GeV}/c$ has consequences on the dijet invariant mass (Fig. 6.7). Since the jets cannot be close, the jet-jet invariant mass tends to assume high values with the result that at the low masses the $M_{jj}$ distribution has fewer events than in the signal region. This explains the background maximum in the signal region. On the other side if $p_{Tjj} > 60 \text{ GeV}/c$ jets are free to be close as well as back to back, therefore the region at low mass is populated by these events. A cut on dijet $p_{Tjj}$ equal to two times the jet $p_T$ has the role to remove a part of those events that peak in the signal region. Therefore the cut $p_{Tjj} > 60 \text{ GeV}/c$ is applied and its effect on the dijets mass distribution is displayed in Fig. 6.8.

The background shoulder in the signal region has been flattened as shown in Fig. 6.8 (c) and (d). Now the backgrounds and in particular the $W + jets$ has a shape which is monotonically decreasing from 30 to 200 $\text{ GeV}/c^2$. The residual irregularity of the background is mainly given by the $t\bar{t}$ background shape which shows two maxima.

The $t\bar{t}$ decay chain that is selected with this selection is:

$$t\bar{t} \rightarrow bW^+bW^- \rightarrow b \ b \ l \ \nu \ j \ j$$

When the two jets chosen in the selection are those from the $W$, the event falls in the region around 80 $\text{ GeV}/c^2$ forming one peak. When one or two jets selected are $b$-jets, the events form a wide peak centered at $\sim 120 \text{ GeV}/c^2$. With jets from $b$, the $M_{jj}$ distribution has higher value than using the jets from $W$ because the $b$-jets have an average larger $p_T$ since they are a $t$ quark decay product. There is a second possible way in which the $t\bar{t}$ could enter in the selection but it occurs with lower probability. It is when

$$t\bar{t} \rightarrow bW^+bW^- \rightarrow bbl\nu\nu'$$
6.2. OPTIMIZATION OF THE SELECTION

Fig. 6.5: Transverse momentum distribution of the system of two leading jets for the electron selection. (a): distribution for data (dots) superimposed to the expected SM predictions (filled area). The statistical error is reported for data (black crosses) and for SM expectations (dashed violet lines). The panel underneath displays the bin-by-bin ratio of data to Monte Carlo with error bars corresponding to the statistical error in data and MC. (b): distribution for overall background events (cyan) superimposed to the distribution for signal events (red). The two distributions are normalized to unitary area.
Fig. 6.6: Two dimensional distributions of the angle $\Delta \phi(j_{1st}, j_{2nd})$ between the two leading jets as a function of the $p_{Tjj}$ for signal events (left) and for background events (right) for the electron selection. The bin color indicates the fraction of events in that bin. Bins filled with warmer colors contain a larger statistics. For $p_{Tjj} < 60$ GeV/c events with small $\Delta \phi(j_{1st}, j_{2nd})$ are not allowed.

with a lepton which is not identified. In this case the resulting invariant mass contributes to the wide peak at high $M_{jj}$ since is made with the two $b$-jets.

The cut on $p_{Tjj}$ diminishes the statistical significance, but increases the $S/B$ ratio and refine the background shape. The new values of $S/\sqrt{B}$ and $S/B$ ratios for the electron and muon channels are summarized in Table 6.2. The amount of signal is halved and this lead to the loss of a part of the significance. However, about three standard deviations remain in both channels. The cut on $p_{Tjj}$ is necessary since now the background peaks no more in the signal region and its shape is regular which helps the fit to separate the signal from the background. The next steps of the selection will be applied with the aim of restoring a part of the significance lost and at the same time of continuing to improve the $S/B$ ratio. Furthermore the following cuts do not worsen the background shape.

6.2.3 Improving the signal significance

The selection which allows the improvement of the background shape has the drawback of decreasing the signal significance. For this reason is important to exploit at best kinematic criteria to "restore" as much as possible the $S/\sqrt{B}$ ratio. Since the statistical error on the significance is about 1% only cuts that increases the significance by at least 1% are considered. It will be shown that the selection will improve the significance a few percent which is the maximum obtainable from cutting on the variables studied. However the $S/B$ ratio will be appreciably improved. All the distributions that have been considered in this study are shown in section B.6. Among them, the distribution of the azimuthal angle between the two leading jets ($\Delta \phi(j_{1st}, j_{2nd})$) shows the largest difference between signal
6.2. OPTIMIZATION OF THE SELECTION

Fig. 6.7: On the top: two dimensional distributions of the angle $p_{Tjj}$ as a function of the $M_{jj}$ for signal events (left) and for background events (right) for the electron selection. The bin color indicates the fraction of events in that bin. Bins filled with warmer colors contain a larger statistics. On the bottom: $M_{jj}$ distribution for events with $p_{Tjj} > 60$ GeV/c (b) and for those events with $p_{Tjj} < 60$ GeV/c (c). The electron selection is applied. The distribution for data (dots) is superimposed to the expected SM predictions (filled area). The statistical error is reported for data (black crosses) and for SM expectations (dashed violet lines). The panel underneath displays the bin-by-bin ratio of data to Monte Carlo with error bars corresponding to the statistical error in data and MC. The background for $p_{Tjj} < 60$ GeV/c peaks on the mass region of the signal and few events have $M_{jj} < 60$ GeV/c². For $p_{Tjj} > 60$ GeV/c the maximum of the background $M_{jj}$ distribution is at 30 GeV/c² and the distribution is monotonically decreasing.
Fig. 6.8: Invariant mass distributions of the two leading jets after the $p_{Tjj} > 60$ GeV/c cut for the electron (left) and muon (right) selections. (a), (b): the distribution for data (dots) is superimposed to the expected SM predictions (filled area). The statistical error is reported for data (black crosses) and for SM expectations (dashed violet lines). The panel underneath displays the bin-by-bin ratio of data to Monte Carlo with error bars corresponding to the statistical error in data and MC. (c), (d): the distribution for overall background events (cyan) is superimposed to the distribution for signal events (red). The two distributions are normalized to unitary area.
6.2. OPTIMIZATION OF THE SELECTION

| Electron channel \((p_{Tjj} \text{ cut})\) | S/\sqrt{B} \quad S/B(\%) \quad \text{signal (\%)} |
|---------------------------------|----------------|----------------|----------------|
| total                           | peak           | total           | peak           |
| 2.67 ± 0.03                     | 1.10 ± 0.01    | 2.50 ± 0.04     | 50             |

| Muon channel \((p_{Tjj} \text{ cut})\) | S/\sqrt{B} \quad S/B(\%) \quad \text{signal (\%)} |
|---------------------------------|----------------|----------------|----------------|
| total                           | peak           | total           | peak           |
| 3.03 ± 0.03                     | 1.10 ± 0.01    | 2.48 ± 0.04     | 47             |

Table 6.2: Significance and signal to background ratio after the \(p_{Tjj}\) cut considering both the whole sample which passes the cut and the subsample of events that belongs to the interval \([60,110] \text{ GeV}/c^2\) of dijet mass for the two channels. The last column provides the expected percentage of signal which passes the cut.

and background. The distribution of \(\Delta \phi(j_{1\text{st}},j_{2\text{nd}})\) for all events passing the previous cuts is shown on Fig. 6.9 for the electron selection. The distribution obtained with the muon selection are exposed in Appendix B. We choose to require \(\Delta \phi(j_{1\text{st}},j_{2\text{nd}}) < 2.5\) which reduces all types of background but mainly the \(W+\text{jets}\). This cut is efficient because the jets from the \(W/Z\) decays are unlikely to be back to back in the laboratory reference system. In fact, the system of two jets is boosted with at least 60 \(\text{GeV}/c\) of transverse momentum (as required by the cut \(p_{Tjj} > 60 \text{ GeV}/c\)). If the two jets are decay products of the \(W/Z\) their momenta, in the center-of-mass reference system, is 40/45 \(\text{GeV}/c\). Then, in the laboratory frame they tend to have a small angle between them. This cut increases the significance from 2.67 ± 0.03 to 2.72 ± 0.03 in the electron channel and from 3.03 ± 0.03 to 3.09 ± 0.03 in the muon channel. This cut has a small impact on the background events in the signal peak region but improves the shape of the \(M_{jj}\) distribution at higher masses acting mainly on the \(t\bar{t}\) and \(W+\text{jets}\) contributions (Fig. 6.10). The next section discusses the selection cuts used to decrease the multijet QCD background whose percentage at this level of the selection is 7.4% in the electron channel and 10.7% in the muon one.

6.2.4 Rejection of the multijet QCD background

The multijet QCD background is the most difficult background to model and predict. For this reason, contrarily to the other backgrounds, the multijet QCD contribution is estimated from data using an ad-hoc data selection as described in section 4.3. It is difficult to estimate the precise reliability of this technique. For this reason, the multijet QCD background could be the main cause of the discrepancies between data and Monte Carlo. In Fig. 6.10 data tend to be slightly below the prediction in the signal region where the multijet QCD fraction is higher, and the effect is more evident in the muon channel which has a higher multijet QCD background than the electron channel (Fig. 6.10).

I studied four cuts to reduce the multijet QCD background imposing constraints on the kinematic of the system. The effect of these four cuts is compared and the one that provides both the best \(S/\sqrt{B}\) ratio and multijet QCD suppression is chosen.

One aspect emerged from the analysis is that in multijet QCD events the \(E_T^{\text{miss}}\) tends to point in the same direction of the leading jet. The distribution of the angle between \(E_T^{\text{miss}}\) and the leading jet \((\Delta \phi(j_{1\text{st}},E_T^{\text{miss}}))\) for the multijet QCD sample in Fig. 6.11(e) (and
CHAPTER 6. SENSITIVITY STUDIES ON WW/WZ RESONANCE

Fig. 6.9: Distributions of the azimuthal angle between the two leading jets ($\Delta \phi(j_{1st}, j_{2nd})$) for the electron channel. (a): the distribution for data (dots) is superimposed to the expected SM predictions (filled area). The statistical error is reported for data (black crosses) and for SM expectations (dashed violet lines). The panel underneath displays the bin-by-bin ratio of data to Monte Carlo with error bars corresponding to the statistical error in data and MC. (b): the distribution for overall background events (cyan) is superimposed to the distribution for signal events (red). The two distributions are normalized to unitary area.
Fig. 6.10: Invariant mass distributions of the two leading jets for the electron selection after the $\Delta \phi (j_{1\text{st}}, j_{2\text{nd}}) < 2.5$ cut. (a): the distribution for data (dots) is superimposed to the expected SM predictions (filled area). The statistical error is reported for data (black crosses) and for SM expectations (dashed violet lines). The panel underneath displays the bin-by-bin ratio of data to Monte Carlo with error bars corresponding to the statistical error in data and MC. (b): the distribution for overall background events (cyan) is superimposed to the distribution for signal events (red). The two distributions are normalized to unitary area.
in Fig. B.5(e) for the muon channel) shows this behaviour. Fig. 6.11 offers a deepened overview of the $\Delta\phi(j_{1st}, E_T^{\text{miss}})$ distributions constituted by:

(a): distribution for data (dots) superimposed to the expected SM predictions (filled area). The statistical error is reported for data (black crosses) and for SM expectations (dashed violet lines). The panel underneath displays the bin-by-bin difference between data and Monte Carlo measured in standard deviations (residuals). The standard deviation is evaluated considering only the statistical error in data and in MC samples.

(b): distribution for background events (cyan) superimposed to the distribution for signal events (red). The two distributions are normalized to unitary area to highlight the differences between the shapes of the signal and background.

(c): two dimensional distributions of the $\Delta\phi(j_{1st}, E_T^{\text{miss}})$ as a function of the invariant mass $M_{jj}$ for signal events (left) and for background events (right). The bin color indicates the fraction of events in that bin. Bins filled with warmer colors contain a larger statistics.

(d): $S/\sqrt{B}$ ratio as a function of the lower $\Delta\phi(j_{1st}, E_T^{\text{miss}})$ threshold.

(e): distribution for multijet QCD background events (grey) superimposed to the distribution for signal events (red). The two distributions are normalized to unitary area.

(f): distribution for $t\bar{t}$ background events (blue) superimposed to the distribution for signal events (red). The two distributions are normalized to unitary area.

For each cut here treated a figure structured like Fig. 6.11 is shown and the description given above still holds (changing the $\Delta\phi(j_{1st}, E_T^{\text{miss}})$ with another variable). The figures presented in this section are obtained by events which pass the electron selection, while the same plots made with the muon selection are listed in Appendix B. The consideration made are valid for the two channels unless differently specified.

For the $\Delta\phi(j_{1st}, E_T^{\text{miss}})$ variable, the multijet QCD distribution fills mainly the small angles while the signal and the other backgrounds occupy mainly the larger ones. There is a slight mismodeling, especially in the electron channel, with data higher than MC below $\Delta\phi(j_{1st}, E_T^{\text{miss}}) \sim 1.5$, and data lower than MC in the other interval. A lower cut on this variables removes a large percentage of multijet QCD but the algorithm applied to this variable do not find any cut which increase the significance in the $M_{jj}$ peak region nor in the whole $M_{jj}$ interval. Therefore, a cut is applied at $\Delta\phi(j_{1st}, E_T^{\text{miss}}) > 0.8$ which brings to a small loss of significance (see Fig. 6.11(d)) while significantly reducing the multijet QCD background. The effects of the selection on the jet-jet invariant mass distribution and on $S/\sqrt{B}$ and $S/B$ ratios are shown in Fig. 6.15 and in the summary Tables 6.4, 6.5.

It should be noticed that in the multijet QCD background passing all the selection criteria there are at least three jets: the two jets candidates for the $W$ hadronic decay and a third jet that has faked the lepton identification. The leading jet is the one measured with the largest absolute uncertainty therefore its error has the highest impact on the missing transverse energy. When the energy of the leading jet is underestimated the $E_T^{\text{miss}}$ points in the leading jet direction. On the contrary the $E_T^{\text{miss}}$ in signal events points basically
Fig. 6.11: Wide view of the distribution of the azimuthal angle between the $E_T^{\text{miss}}$ and the leading jet ($\Delta \phi(j_{\text{1st}}, E_T^{\text{miss}})$). For the description of each single plot refer to the beginning of section 6.2.4. The electron selection is applied until the $\Delta \phi(j_{\text{1st}}, j_{\text{2nd}}) < 2.5$ cut.
on the opposite side of the leading jet. The analysis of the event kinematics help to understand this behaviour. For the principle of conservation of the transverse momentum, the \( p_T \) of the two bosons in the signal events balance each other (assuming no other jet is produced). The \( W/Z \) bosons which decay hadronically are boosted by at least 60 \( \text{GeV}/c \) of transverse momentum. As a consequence, the system lepton \( E_T^{\text{miss}} \) should also have a transverse momentum higher than 60 \( \text{GeV}/c \). Therefore in the final state of signal events the lepton and \( E_T^{\text{miss}} \) are close in azimuthal angle opposite to the dijet system. For this reason the \( E_T^{\text{miss}} \) tends to be opposite to the leading jet.

As mentioned above, if no additional activity is produced in the hard interaction the two bosons are produced back-to-back to each other. If additional activity emerges either because of multi-parton interactions or because of pile-up, or if some of the physics objects are not well reconstructed, the \( E_T^{\text{miss}} \) can be modified so that the two bosons are no more \( p_T \) balanced. In general, however, the multi-parton interactions and the pile-up produce low \( p_T \) jets therefore we can expect that in well reconstructed \( WW/WZ \) events the \( p_T \) balance remains valid. On the contrary, in backgrounds events where there are more than two jets at medium or high \( p_T \) produced in the hard interaction such as in \( t\bar{t}, W+3\text{partons} \) and multijet QCD the system of two jets does not counterbalance the system lepton \( E_T^{\text{miss}} \). A cut sensitive to the conservation of the momentum should therefore removes background, in particular the multijet QCD one. The variables which I studied are:

\[
\Delta \phi(W_{\text{lep}},W_{\text{had}}) : \text{the azimuthal angle } \phi \text{ between the system lepton } E_T^{\text{miss}} (W_{\text{lep}}) \text{ and the dijet system } (W_{\text{had}}); \\
\text{balance}(W_{\text{lep}},W_{\text{had}}) : \text{the ratio (or balance) of the } p_T \text{ of the leptonically decaying } W \text{ candidate } (W_{\text{lep}}) \text{ to the } p_T \text{ of the hadronically decaying } W \text{ (W}_{\text{had}}); \\
p_{TWW} : \text{the vectorial sum of the } p_T \text{ of the } W_{\text{lep}} \text{ and the one of the } W_{\text{had}}.
\]

As discussed above, in the signal the \( \Delta \phi(W_{\text{lep}},W_{\text{had}}) \) angle tends to be around \( \pi \), the \( \text{balance}(W_{\text{lep}},W_{\text{had}}) \) around 1 and the transverse momentum \( p_{TWW} \) close to 0.

The plots in Fig. 6.12 (Fig. B.6 for the muon channel), show the \( \Delta \phi(W_{\text{lep}},W_{\text{had}}) \) distribution for data, for the signal and for all the most significant backgrounds. I have used the algorithm to find the rectangle in the two dimensional distribution \( \Delta \phi(W_{\text{lep}},W_{\text{had}}) \) vs. \( M_{jj} \) which maximizes the \( S/\sqrt{B} \) ratio in both the electron and the muon channel. The result is a region between 60 and 110 \( \text{GeV}/c^2 \) of \( M_{jj} \) and over 2.3 – 2.5 of \( \Delta \phi(W_{\text{lep}},W_{\text{had}}) \) depending on the channel. I choose to cut in \( \Delta \phi(W_{\text{lep}},W_{\text{had}}) \) using the value 2.3 on both channels.

Fig. 6.13 shows the distributions of the balance of the \( W_{\text{had}} \) and the \( W_{\text{lep}} \) systems for the electron channel (Fig. B.7 for the muon channel). The background components tends to have a transverse momentum of \( W_{\text{lep}} \) systems lower than that of the jet-jet system. This is also true for the signal with an average balance closer to 1 than the backgrounds. This is probably due the cut at dijet \( p_{Tjj} > 60 \text{ GeV}/c \) which sharply cuts the \( p_{Tjj} \) distribution.

The balance for the multijet QCD background shows the lowest average value. This is more evident in the muon channel. Also the \( t\bar{t} \) has a large number of unbalanced events. In fact \( t\bar{t} \) events have six objects in the final state and when the jets selected as \( W \) decay products are the ones originating from the \( t \rightarrow Wb \) decay they have a quite large \( p_T \) therefore the dijet \( p_{Tjj} \) tends to be higher than that of the \( W_{\text{lep}} \) system.
6.2. OPTIMIZATION OF THE SELECTION

Fig. 6.12: Wide view of the distribution of the azimuthal angle between the $W_{\text{had}}$ and the $W_{\text{lep}}$ ($\Delta \phi(W_{\text{lep}}, W_{\text{had}})$). For the description of each single plot refer to the beginning of section 6.2.4. The electron selection is applied until the $\Delta \phi(j_{1\text{st}}, j_{2\text{nd}}) < 2.5$ cut.
On the basis of the results obtained from the algorithm which maximizes the significance in the signal region, it has been chosen to cut the events with balance $< 0.35$. The vectorial sum of the momenta of the two systems $p_{TWW}$ is a variable which combines the discriminant power of the cut in angle with that in balance. The Fig. 6.14 (Fig. B.8) shows this distribution for the electron (muon) channel. The $p_{TWW}$ distribution for Monte Carlo events is in close agreement with the experimental data. The backgrounds, in particular the multijet QCD and $t\bar{t}$ components, tend to have a higher $p_{TWW}$ than the signal. The cut which maximize the $S/\sqrt{B}$ ratio in the peak region is $p_{TWW} < 65$ GeV/c.

In summary we studied four cuts to reduce the multijet QCD background and try to increase the significance. These cuts are:

- the angle between the leading jet and the missing transverse energy $\Delta \phi(j_{1st}, E^{miss}_{T}) > 0.8$;
- the angle between the dijet and the lepton+$E^{miss}_{T}$ system $\Delta \phi(W_{lep}, W_{had}) > 2.3$;
- the ratio of the lepton+$E^{miss}_{T}$ transverse momentum to the dijet balance $W_{lep}, W_{had}) > 0.35$;
- the vectorial sum of the lepton+$E^{miss}_{T}$ and dijet transverse momenta $p_{TWW} < 65$ GeV/c.

Fig. 6.15 and 6.16 give an overview of how the jet-jet invariant mass is modified by the cuts which are applied in the electron and muon channel. All the four cuts work well in reducing the multijet QCD but the $p_{TWW}$ cut acts better in reducing also the $t\bar{t}$ background (see Fig. 6.14(f)).

Tables 6.4 (for the electron ch.) and 6.5 (for the muon ch.) report the overall $S/\sqrt{B}$ and $S/B$ values calculated for the whole $M_{jj}$ range and for the signal region only. The residual percentage of the multijet QCD background after the various cut is also specified. The $\Delta \phi(j_{1st}, E^{miss}_{T})$ is more efficient in removing the multijet QCD background in the electron channel than in the muon one but in both cases the $S/\sqrt{B}$ ratio decreases when applying this cut. This variable is not very well described from the MC (Fig. 6.11(a)). The Monte Carlo exceeds data for large angles while the situation is opposite for small angles. The $\Delta \phi(W_{lep}, W_{had})$, balance($W_{lep}, W_{had}$) and $p_{TWW}$ cuts have all the effect of raising the $S/\sqrt{B}$ and the $S/B$ ratios and while reducing the multijet QCD background. However looking at the values in the tables, the cut which has the most significant impact in removing the multijet QCD background and the largest significance increase is the $p_{TWW}$ cut. Furthermore, this cut brings the percentage of multijet QCD background in the two channels to a similar value giving to the $M_{jj}$ distribution in the two channels a similar shape and contribution of backgrounds. For these reasons the requirement added to the selection is $p_{TWW} < 65$ GeV/c. The percentage of the various background and signal sources in data after this cut are reported in Table 6.3 for the electron and muon channels.

### 6.2.5 Reduction of the $t\bar{t}$ background

The plots of the dijet invariant mass after the $p_{TWW}$ cut are represented enlarged in Fig. 6.17(a,b). The background with the most irregular shape in the signal region remains
6.2. OPTIMIZATION OF THE SELECTION

Fig. 6.13: Wide view of the distribution of the momentum balance of the $W_{\text{had}}$ and the $W_{\text{lep}}$ ($\text{balance}(W_{\text{lep}}, W_{\text{had}})$). For the description of each single plot refer to the beginning of section 6.2.4. The electron selection is applied until the $\Delta \phi(j_{\text{1st}}, j_{\text{2nd}}) < 2.5$ cut.
Fig. 6.14: Wide view of the distribution of the transverse momentum resulting from the vectorial sum of the \(W_{\text{had}}\) \(p_T\) and the \(W_{\text{lep}}\) \(p_T\) (\(p_{TWW}\)). For the description of each single plot refer to the beginning of section 6.2.4. In plot (d), the \(S/\sqrt{B}\) ratio is shown as a function of the upper \(p_{TWW}\) threshold. The electron selection is applied until the \(\Delta \phi(j_{1\text{st}}, j_{2\text{nd}}) < 2.5\) cut.
### 6.2. Optimization of the Selection

<table>
<thead>
<tr>
<th>Sample</th>
<th>Fraction of SM predictions in data (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Electron ch.</td>
</tr>
<tr>
<td>$WW/WZ$</td>
<td>1.32</td>
</tr>
<tr>
<td>$tt$</td>
<td>10.29</td>
</tr>
<tr>
<td>single top</td>
<td>1.28</td>
</tr>
<tr>
<td>$Z + jets$</td>
<td>6.00</td>
</tr>
<tr>
<td>$W + jets$</td>
<td>74.5</td>
</tr>
<tr>
<td>multijet QCD</td>
<td>6.6</td>
</tr>
</tbody>
</table>

Table 6.3: Expected fractions of the SM processes in data after the $p_{TW} < 65 \text{ GeV}/c$ cut.

<table>
<thead>
<tr>
<th>Cut</th>
<th>$S/\sqrt{B}$</th>
<th>$S/B(%)$</th>
<th>multijet QCD(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>total</td>
<td>peak</td>
<td>total</td>
</tr>
<tr>
<td>$\Delta\phi(j_{1st},j_{2nd})$</td>
<td>2.72 ± 0.03</td>
<td>2.90 ± 0.04</td>
<td>1.23 ± 0.01</td>
</tr>
<tr>
<td>$\Delta\phi(j_{1st},E_{T}^{miss})$</td>
<td>2.67 ± 0.03</td>
<td>2.89 ± 0.04</td>
<td>1.32 ± 0.02</td>
</tr>
<tr>
<td>balance($W_{lep},W_{had}$)</td>
<td>2.71 ± 0.03</td>
<td>2.92 ± 0.04</td>
<td>1.27 ± 0.02</td>
</tr>
<tr>
<td>$\Delta\phi(W_{lep},W_{had})$</td>
<td>2.70 ± 0.03</td>
<td>2.92 ± 0.04</td>
<td>1.34 ± 0.02</td>
</tr>
<tr>
<td>$p_{TW}$</td>
<td>2.70 ± 0.03</td>
<td>2.95 ± 0.04</td>
<td>1.34 ± 0.02</td>
</tr>
</tbody>
</table>

Table 6.4: Significance and signal to background ratio as a function of the cut considering both the whole sample which passes the cut and the subsample of events that belongs to the interval $[60,110] \text{ GeV}/c^2$ of dijet mass for the electron channel. The last column provides the expected percentage of multijet QCD background in data after each cut. The $\Delta\phi(j_{1st},E_{T}^{miss})$, $\text{balance}(W_{lep},W_{had})$, $\Delta\phi(W_{lep},W_{had})$, $p_{TW}$ requirements are applied independently after the $\Delta\phi(j_{1st},j_{2nd})$ cut.

<table>
<thead>
<tr>
<th>Cut</th>
<th>$S/\sqrt{B}$</th>
<th>$S/B(%)$</th>
<th>multijet QCD(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>total</td>
<td>peak</td>
<td>total</td>
</tr>
<tr>
<td>$\Delta\phi(j_{1st},j_{2nd})$</td>
<td>3.09 ± 0.03</td>
<td>3.27 ± 0.04</td>
<td>1.23 ± 0.01</td>
</tr>
<tr>
<td>$\Delta\phi(j_{1st},E_{T}^{miss})$</td>
<td>3.01 ± 0.03</td>
<td>3.19 ± 0.04</td>
<td>1.30 ± 0.01</td>
</tr>
<tr>
<td>balance($W_{lep},W_{had}$)</td>
<td>3.14 ± 0.03</td>
<td>3.33 ± 0.04</td>
<td>1.32 ± 0.01</td>
</tr>
<tr>
<td>$\Delta\phi(W_{lep},W_{had})$</td>
<td>3.13 ± 0.03</td>
<td>3.34 ± 0.04</td>
<td>1.33 ± 0.01</td>
</tr>
<tr>
<td>$p_{TW}$</td>
<td>3.13 ± 0.03</td>
<td>3.35 ± 0.04</td>
<td>1.40 ± 0.02</td>
</tr>
</tbody>
</table>

Table 6.5: Significance and signal to background ratio as a function of the cut considering both the whole sample which passes the cut and the subsample of events that belongs to the interval $[60,110] \text{ GeV}/c^2$ of dijet mass for the muon channel. The last column provides the expected percentage of multijet QCD background in data after each cut. The $\Delta\phi(j_{1st},E_{T}^{miss})$, $\text{balance}(W_{lep},W_{had})$, $\Delta\phi(W_{lep},W_{had})$, $p_{TW}$ requirements are applied independently after the $\Delta\phi(j_{1st},j_{2nd})$ cut.
CHAPTER 6. SENSITIVITY STUDIES ON WW/WZ RESONANCE

Fig. 6.15: Dijet invariant mass distributions for the electron channel after the selections: $(\Delta \phi(j_{1st}, E_T^{miss}) > 0.8)$ (a), $(\Delta \phi(W_{lep}, W_{had}) > 2.3)$ (b), $(balance(W_{lep}, W_{had}) > 0.35)$ (c), $(pt_{WW} < 65 \text{ GeV/c})$ (d). For each plot, the distribution for data (dots) is superimposed to the expected SM predictions (filled area). The statistical error is reported for data (black crosses) and for SM expectations (dashed violet lines). The panel underneath displays the bin-by-bin ratio of data to Monte Carlo with error bars corresponding to the statistical error in data and MC.
6.2. OPTIMIZATION OF THE SELECTION

Fig. 6.16: Dijet invariant mass distributions for the muon channel after the selections: $(\Delta \phi(j_{1st}, E_{T}^{miss}) > 0.8)$ (a), $(\Delta \phi(W_{lep}, W_{had}) > 2.3)$ (b), $(balance(W_{lep}, W_{had}) > 0.35)$ (c), $(p_{T}^{WW} < 65 \text{ GeV/c})$ (d). For each plot, the distribution for data (dots) is superimposed to the expected SM predictions (filled area). The statistical error is reported for data (black crosses) and for SM expectations (dashed violet lines). The panel underneath displays the bin-by-bin ratio of data to Monte Carlo with error bars corresponding to the statistical error in data and MC.
Fig. 6.17: Dijet mass distributions for the electron (a) and muon (b) channels after the $p_T^{WW} < 65$ GeV/c cut. For each plot, the distribution for data (dots) is superimposed to the expected SM predictions (filled area). The statistical error is reported for data (black crosses) and for SM expectations (dashed violet lines). The panel underneath displays the bin-by-bin ratio of data to Monte Carlo with error bars corresponding to the statistical error in data and MC.
the $t\bar{t}$ while the multijet QCD, the $W+jets$ and the $Z+jets$ have been modeled to a shape monotonically decreasing. The data do not show significant disagreement with respect to the Monte Carlo even if they still tend to be lower than the MC in the signal region.

The goal of this section is to find a variable which allows a further reduction of the $t\bar{t}$ along with an improvement of the significance. Since the $t\bar{t}$ events are usually characterized by four jets in the final state, they have often a jet multiplicity larger than two. For this reason I have studied the kinematic of the non leading jets to understand whether the application of a cut on them could remove the $t\bar{t}$ while preserving the signal.

The Fig. 6.18 displays the jet multiplicities $N_{jets}$ at this level of the selection for the electron channel (distributions for muon channel in Fig. B.9). About the 70% of the $t\bar{t}$ background can be removed if the jet multiplicity is required to be exactly two. However this cut removes about 20% of the signal. As a result, the significance is reduced even if most of $t\bar{t}$ is removed. The $S/\sqrt{B}$, the $S/B$ ratios and the fraction of $t\bar{t}$ after applying this selection are given in Tables 6.6 (electron channel), 6.7 (muon channel). Fig. 6.18(c) is the dijet mass distribution for events with $N_{jets} = 2$.

Further studies have been done on the distributions of the following variables:

- $p_T$, energy $E$, $\eta$ of the 3rd and 4th jet ($p_T$ ordered);
- scalar sum of the $E$ of all jets with $p_T > 20$ GeV/c not considered as candidate for the $W$ hadronic decay;
- scalar sum of the $p_T$ of all jets with $p_T > 20$ GeV/c not considered as candidate for the $W$ hadronic decay ($\Sigma p_{T_{j_{\text{3rd},4th}}}$).

This last variable is the one that showed the best performance. In the calculation of $\Sigma p_{T_{j_{\text{3rd},4th}}}$ enters only jets passing all the selection criteria used for the leading jets (see section 4.2.4) but with a $p_T$ which could be at minimum 20 GeV/c. The calculation has been done also summing the $p_T$ of jets with at least $p_T > 30$ GeV/c but fewer events have at least three of these jets, and due to this the variable has less power to discriminate signal from $t\bar{t}$. Fig. 6.19 and B.10 shows the $\Sigma p_{T_{j_{\text{3rd},4th}}}$ distributions for the electron and muon channels respectively. The data are well described by the Monte Carlo prediction (plot (a)) and this give us confidence on cutting on this variable. Only 10% of the events in the $t\bar{t}$ background do not have at least a third jet with $p_T > 20$ GeV/c and the $\Sigma p_{T_{j_{\text{3rd},4th}}}$ distribution extends to higher values than the signal one. This is in agreement with the statement that $b$-jets from the $t$ decay have high transverse momentum. In the signal sample instead there are rarely jets with high $p_T$ except the two leading jets.

The cut which maximizes the significance in the signal region is found to be $\Sigma p_{T_{j_{\text{3rd},4th}}} < 70$ GeV/c. The invariant mass distribution obtained after the application of this requirement is shown in Fig. 6.19(c) (electron channel) and B.10(c) (muon channel).

The values of the significance and of the $S/B$ ratio are reported in Tables 6.6 (electron) and 6.7 (muon). The $\Sigma p_{T_{j_{\text{3rd},4th}}}$ cut raises both $S/B$ and $S/\sqrt{B}$ values in the peak region and reduces $t\bar{t}$ fraction from 10.3% to 6.9%.

### 6.2.6 Summary

I applied a selection on $p_{T_{jj}}$ to shift the signal from the top of the background distribution to a region where the background decreases monotonically. Then a cut on $\Delta \phi(j_{1\text{st}},j_{2\text{nd}})$
CHAPTER 6. SENSITIVITY STUDIES ON WW/WZ RESONANCE

Fig. 6.18: Wide view of the distribution of the jet multiplicity ($N_{jets}$). For the description of each single plot (except (c)) refer to the beginning of section 6.2.4. In plot (d), the $S/\sqrt{B}$ ratio is shown as a function of the upper $N_{jets}$ threshold. The electron selection is applied until the $p_{TW} < 65$ GeV/c cut. (c): jet-jet invariant mass distribution obtained for $N_{jets} = 2$. The distribution for data (dots) is superimposed to the expected SM predictions (filled area). The statistical error is reported for data (black crosses) and for SM expectations (dashed violet lines). The panel underneath displays the bin-by-bin ratio of data to Monte Carlo with error bars corresponding to the statistical error in data and MC.
6.2. OPTIMIZATION OF THE SELECTION

![Graphs and data]

Fig. 6.19: Wide view of the $\Sigma p_{T_{j, \text{th}}}$ distribution. For the description of each single plot (except (c)) refer to the beginning of section 6.2.4. In plot (d), the $S/\sqrt{B}$ ratio is shown as a function of the upper $\Sigma p_{T_{j, \text{th}}}$ threshold. The electron selection is applied until the $p_{T_{WW}} < 65$ GeV/c cut. (c): jet-jet invariant mass distribution obtained after the $\Sigma p_{T_{j, \text{th}}} < 70$ GeV/c cut. The distribution for data (dots) is superimposed to the expected SM predictions (filled area). The statistical error is reported for data (black crosses) and for SM expectations (dashed violet lines). The panel underneath displays the bin-by-bin ratio of data to Monte Carlo with error bars corresponding to the statistical error in data and MC.
### Electron channel

<table>
<thead>
<tr>
<th>Cut</th>
<th>$S/\sqrt{B}$</th>
<th>$S/B(%)$</th>
<th>$t\bar{t}(%)$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>total</td>
<td>peak</td>
<td>total</td>
</tr>
<tr>
<td>$p_{TWW}$</td>
<td>2.70 ± 0.03</td>
<td>2.95 ± 0.04</td>
<td>1.34 ± 0.02</td>
</tr>
<tr>
<td>$N_{jets} = 2$</td>
<td>2.45 ± 0.03</td>
<td>2.90 ± 0.05</td>
<td>1.40 ± 0.02</td>
</tr>
<tr>
<td>$\Sigma p_{T_{j_{1}} &lt; 70 \text{ GeV/c}}$</td>
<td>2.69 ± 0.03</td>
<td>2.98 ± 0.05</td>
<td>1.39 ± 0.02</td>
</tr>
</tbody>
</table>

Table 6.6: Significance and signal to background ratio as a function of the cut considering both the whole sample which passes the cut and the subsample of events that belongs to the interval $[60, 110] \text{ GeV/c}^2$ of dijet mass for the electron channel. The last column provides the expected percentage of $t\bar{t}$ background in data after the cut. The $N_{jets} = 2$ and $\Sigma p_{T_{j_{1}} < 70 \text{ GeV/c}}$ requirements are applied independently after the $p_{TWW}$ cut.

### Muon channel

<table>
<thead>
<tr>
<th>Cut</th>
<th>$S/\sqrt{B}$</th>
<th>$S/B(%)$</th>
<th>$t\bar{t}(%)$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>total</td>
<td>peak</td>
<td>total</td>
</tr>
<tr>
<td>$p_{TWW}$</td>
<td>3.13 ± 0.03</td>
<td>3.35 ± 0.04</td>
<td>1.40 ± 0.02</td>
</tr>
<tr>
<td>$N_{jets} = 2$</td>
<td>2.84 ± 0.03</td>
<td>3.26 ± 0.05</td>
<td>1.46 ± 0.02</td>
</tr>
<tr>
<td>$\Sigma p_{T_{j_{1}} &lt; 70 \text{ GeV/c}}$</td>
<td>3.13 ± 0.03</td>
<td>3.40 ± 0.05</td>
<td>1.47 ± 0.02</td>
</tr>
</tbody>
</table>

Table 6.7: Significance and signal to background ratio as a function of the cut considering both the whole sample which passes the cut and the subsample of events that belongs to the interval $[60, 110] \text{ GeV/c}^2$ of dijet mass for the muon channel. The last column provides the expected percentage of $t\bar{t}$ background in data after the cut. The $N_{jets} = 2$ and $\Sigma p_{T_{j_{1}} < 70 \text{ GeV/c}}$ requirements are applied independently after the $p_{TWW}$ cut.
6.2. Optimization of the Selection

<table>
<thead>
<tr>
<th>Cut</th>
<th>$S/\sqrt{B}$ (S/B)</th>
<th>$S/B$ (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N_{jets} \geq 2$</td>
<td>3.70 ± 0.03</td>
<td>1.06 ± 0.01</td>
</tr>
<tr>
<td>$p_{T_{jj}}$</td>
<td>2.67 ± 0.03</td>
<td>1.10 ± 0.01</td>
</tr>
<tr>
<td>$\Delta \phi(j_{1st}, j_{2nd})$</td>
<td>2.72 ± 0.03</td>
<td>1.23 ± 0.01</td>
</tr>
<tr>
<td>$p_{TW}$</td>
<td>2.70 ± 0.03</td>
<td>1.34 ± 0.02</td>
</tr>
<tr>
<td>$\Sigma p_{T_{j_{1th}}}$</td>
<td>2.69 ± 0.03</td>
<td>1.39 ± 0.02</td>
</tr>
</tbody>
</table>

Table 6.8: Significance and signal to background ratio after the application of each requirement for the electron channel.

<table>
<thead>
<tr>
<th>Cut</th>
<th>$S/\sqrt{B}$ (S/B)</th>
<th>$S/B$ (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N_{jets} \geq 2$</td>
<td>4.23 ± 0.03</td>
<td>1.03 ± 0.01</td>
</tr>
<tr>
<td>$p_{T_{jj}}$</td>
<td>3.03 ± 0.03</td>
<td>1.10 ± 0.01</td>
</tr>
<tr>
<td>$\Delta \phi(j_{1st}, j_{2nd})$</td>
<td>3.09 ± 0.03</td>
<td>1.23 ± 0.01</td>
</tr>
<tr>
<td>$p_{TW}$</td>
<td>3.13 ± 0.03</td>
<td>1.40 ± 0.02</td>
</tr>
<tr>
<td>$\Sigma p_{T_{j_{1th}}}$</td>
<td>3.13 ± 0.03</td>
<td>1.47 ± 0.02</td>
</tr>
</tbody>
</table>

Table 6.9: Significance and signal to background ratio after the application of each requirement for the muon channel.

has had the effect to reduce a part of all background components and has refined the background shape. The $p_{TW}$ cut has been used to remove a part of the multijet QCD and of the $t\bar{t}$ fractions and, finally, the $\Sigma p_{T_{j_{1th}}}$ cut gave a further reduction of the $t\bar{t}$ background. Tables 6.8, 6.9 summarize significance and signal to background ratio after each requirement of the selection for the electron and muon channels respectively. Table 6.10 contains the percentage of Standard Model samples in data. The jet-jet invariant mass distributions at the end of the selection are shown in Fig. 6.20. The data distribution, the MC distribution and their bin-by-bin ratio, the statistical error and the systematic error band are shown. The data and the SM expectations are in good agreement. The largest discrepancies are of the order of 10% and are located in the region $M_{jj} < 40$ GeV/c$^2$ for both electron and muon selections. The final $S/\sqrt{B}$ ($S/B$) ratios are 2.98 ± 0.05 (2.88 ± 0.05) for the electron channel and 3.40 ± 0.05 (2.99 ± 0.05) for the muon channel. Respect to the starting point the $S/B$ ratio is increased and the background shape is improved. The statistical significance is decreased but is at least 3 in both channels. Fig. B.28 shows the evolution of the shape and magnitude of the $M_{jj}$ distributions as a function of the cut applied for the signal and for the dominant backgrounds. The section 6.4 estimates the magnitude of the systematic uncertainty in the $M_{jj}$ distribution obtained at the end of the selection in order to understand if the $S/B$ ratio obtained can be enough to measure the signal cross-section.
Fig. 6.20: Dijet mass distributions of the two leading jets for the electron (a) and muon (b) selections after the $\Sigma p_T^{j_{1,2}} < 70 \text{ GeV}/c$ cut. The distribution for data (dots) is superimposed to the expected SM predictions (filled area). The statistical error is reported for data (black crosses) and for SM expectations (dashed violet lines). The panel underneath displays the bin-by-bin ratio of data to Monte Carlo (red crosses) with error bars corresponding to the statistical error in data and MC, the relative systematic error due to the JES uncertainty (woven cyan bands) and the sum in quadrature of the two errors (black bands).
6.3 Shifting the jet $p_T$ threshold

I repeated the analysis for different values of jet $p_T$ threshold. The $p_{T_{jj}}$ threshold has been changed too according to the statements made in section 6.2.2 ($p_{T_{jj}}$ threshold > $2p_T$ threshold). I selected events with at least two jets with a transverse momentum of at least 20 or 25 GeV/c and respectively with $p_{T_{jj}}$ more than 40 and 50 GeV/c. The other cuts remain unchanged. The $M_{jj}$ distributions obtained are compared with the one made with the $p_T > 30$ GeV/c cut in Fig. 6.21. The Tables 6.11 and 6.12 report the $S/B$ and $S/\sqrt{B}$ ratios for the three configurations. A weaker cut on the jet $p_T$ brings to a larger $S/\sqrt{B}$ ratio and a smoother background shape. On the other hand the $S/B$ ratio falls therefore the ratio signal to the systematic error is smaller. The systematic uncertainty is estimated in section 6.4 and is about 2% in the signal region therefore at the moment the better choice is that with the larger $S/B$ ratio which is $p_T > 30$ GeV/c and $p_{T_{jj}} > 60$ GeV/c.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Electron ch.</th>
<th>Muon ch.</th>
</tr>
</thead>
<tbody>
<tr>
<td>$WW/WZ$</td>
<td>1.38</td>
<td>1.43</td>
</tr>
<tr>
<td>$t\bar{t}$</td>
<td>6.96</td>
<td>6.86</td>
</tr>
<tr>
<td>single top</td>
<td>1.34</td>
<td>1.41</td>
</tr>
<tr>
<td>$Z + jets$</td>
<td>6.17</td>
<td>3.54</td>
</tr>
<tr>
<td>$W + jets$</td>
<td>77.3</td>
<td>80.7</td>
</tr>
<tr>
<td>multijet QCD</td>
<td>6.9</td>
<td>6.1</td>
</tr>
</tbody>
</table>

Table 6.10: Expected fractions of the SM processes in data at the end of the selection.

<table>
<thead>
<tr>
<th></th>
<th>$S/\sqrt{B}$ (total)</th>
<th>$S/\sqrt{B}$ (peak)</th>
<th>$S/B$ (%) (total)</th>
<th>$S/B$ (%) (peak)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_T &gt; 20, p_{T_{jj}} &gt; 40$</td>
<td>3.32 ± 0.03</td>
<td>3.57 ± 0.04</td>
<td>0.98 ± 0.01</td>
<td>1.92 ± 0.03</td>
</tr>
<tr>
<td>$p_T &gt; 25, p_{T_{jj}} &gt; 50$</td>
<td>3.09 ± 0.03</td>
<td>3.37 ± 0.04</td>
<td>1.24 ± 0.01</td>
<td>2.45 ± 0.04</td>
</tr>
<tr>
<td>$p_T &gt; 30, p_{T_{jj}} &gt; 60$</td>
<td>2.69 ± 0.03</td>
<td>2.98 ± 0.05</td>
<td>1.39 ± 0.02</td>
<td>2.88 ± 0.05</td>
</tr>
</tbody>
</table>

Table 6.11: Significance and signal to background ratio as a function of jet $p_T$ and $p_{T_{jj}}$ thresholds for the electron channel.
Fig. 6.21: $M_{jj}$ distributions in the electron (left) and muon (right) channels at the end selection. The jet $p_T$ threshold is increased from 20 GeV/c (top) to 25 GeV/c (center) to 30 GeV/c (bottom). In each plot, the distribution for data (dots) is superimposed to the expected SM predictions (filled area). The statistical error is reported for data (black crosses) and for SM expectations (dashed violet lines). The panel underneath displays the bin-by-bin ratio of data to Monte Carlo with error bars corresponding to the statistical error in data and MC.
6.4 SYSTEMATIC UNCERTAINTIES

The main systematic uncertainties which affect the diboson measurement are the same studied in the search of the CDF resonance (section 5.2). Another important source of systematic error is the uncertainty on the shape of the $M_{jj}$ distribution due to the model used to generate the MC samples. Particularly important in this case is the shape of the $W+\text{jets}$ background, which is the dominant background process. One example of this uncertainty is the dependence of the $W+\text{jets}$ shape on the renormalization scale. The MC samples are generated at the leading order of the perturbation theory (except $t\bar{t}$ and single top produced at NLO). This means that the physics generated is only an approximation of what is predicted by the theory. The prediction made with a theory broken at a certain perturbative order depends on an unphysical renormalization scale factor which is arbitrarily set to some value. In the $W+\text{jets}$ case is equal to the center-of-mass energy $Q$ of the hard-scattering. The physics results should be independent from the renormalization scale. A look at the $M_{jj}$ distributions obtained changing the renormalization scale to $2\cdot Q$ and $Q/2$ allows to evaluate the error due to the approximation made. The systematic error on the $M_{jj}$ distribution is estimated as the bin-by-bin maximum difference of the distributions with renormalization scale set to $2\cdot Q$ and $Q/2$ from the one with renormalization scale equal to $Q$. A preliminary study done at truth level on the $W+\text{jets}$ sample with a standard selection for the muon channel plus the $p_{Tjj}>60$ GeV/c cut is shown in Fig. 6.22 [72]. The differences due to the renormalization scale in the signal peak region are of the order of 5 $\sim$ 10%. Since the $W+\text{jets}$ is the dominant background 5 $\sim$ 10% is the systematic error which we expect for in the $M_{jj}$ distribution made with all SM processes. This error is large and studies are on-going to have a more precise evaluation and to find a way to decrease it.

The study I made on systematic uncertainties, focused on the effect of the Jet Energy Scale (JES) uncertainty on the dijet mass distribution. The JES uncertainty is one of the most important systematic uncertainties and is the overall systematic error resulting from uncertainties on calorimeter scale, dead material description, cluster reconstruction, fragmentation and Underlying Event modeling as well as pile-up [54, 55]. To evaluate the impact on the $M_{jj}$ distribution, the energy of each jet is shifted up and down by a percentage value corresponding to the JES uncertainty, the selection is applied again and a new $M_{jj}$ distribution is obtained. The systematic error on the $M_{jj}$ distribution is estimated as the bin-by-bin maximum difference of the distributions with shifted JES with respect to the one at nominal JES. Fig. 6.23 shows the $M_{jj}$ distributions for MC with shifted JES and with nominal JES and the distribution for data. The lower panel shows for both channels

<table>
<thead>
<tr>
<th>$p_T &gt;, p_{Tjj} &gt;$</th>
<th>$S/\sqrt{B}$</th>
<th>$S/B(%)$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>total</td>
<td>peak</td>
</tr>
<tr>
<td>20, 40</td>
<td>3.81 ± 0.04</td>
<td>4.04 ± 0.04</td>
</tr>
<tr>
<td>25, 50</td>
<td>3.58 ± 0.03</td>
<td>3.80 ± 0.04</td>
</tr>
<tr>
<td>30, 60</td>
<td>3.13 ± 0.03</td>
<td>3.40 ± 0.05</td>
</tr>
</tbody>
</table>

Table 6.12: Significance and signal to background ratio as a function of jet $p_T$ and $p_{Tjj}$ thresholds for the muon channel.
CHAPTER 6. SENSITIVITY STUDIES ON WW/WZ RESONANCE

Fig. 6.22: Jet-jet invariant mass distribution obtained with the renormalization scale set to $Q/2$ (red), $2Q$ (blue) and to the standard value $Q$ (black) for $W + jets$ samples. The lower pad displays the bin-by-bin ratio of the distribution obtained with $Q/2$ and $2Q$ to the standard one. [72].

the bin-by-bin ratio data to MC with the statistical error (red markers). The systematic uncertainty is also shown as a cyan band. In the signal region ($[60, 110]$ GeV/$c^2$) the JES systematic uncertainty is about 2.1%.

The $S/B$ ratio obtained with the selection described in this chapter is comparable to the systematic uncertainty given by the JES uncertainty only. As a result the sensitivity of the fit to the WW/WZ signal is not enough to measure its cross-section unless the $S/B$ ratio is increased or the systematic uncertainty is reduced. A careful study of the JES and in particular of the effect due to the pile-up on the jet energy is on-going and this could result in a reduction of the systematic error. The $S/B$ ratio could also be improved by using a more stringent selection. However this has the price to reduce the $S/\sqrt{B}$ ratio. To simultaneously keep the statistical significance at about 3 and improve the $S/B$ ratio a larger statistics is needed. In section 6.8 it is described a selection suitable for a data sample of about $5 $ fb$^{-1}$ where the $S/B$ ratio can be improved while the signal significance can still reach at least 3 in both channels.
Fig. 6.23: Jet-jet invariant mass distributions for electron (a) and muon (b) selections. For each plot, the distributions obtained by shifting the JES up (red) and down (blue) by its uncertainty are compared to the distribution obtained with nominal JES (black). The black dots correspond to data. The errors bars are the statistical errors. The panel underneath displays the bin-by-bin ratio of data to Monte Carlo (red crosses) with error bars corresponding to the statistical error in data and MC and the cyan bands represent the relative systematic uncertainty due to the JES uncertainty estimated as the maximum difference between the shifted bin values and the unshifted one.
6.5 \( Z \) control region

The expected fraction of \( Z + \text{jets} \) events in data at the end of the selection studied in this chapter is 6.2\% for the electron channel and 3.6\% for the muon one. This background is not negligible and it is important to test if the simulation correctly reproduces it. The normalization and the shape of the \( Z + \text{jets} \) distribution can be checked using a sample created to enrich the percentage of \( Z + \text{jets} \) events. The selection is obtained by substituting the cuts used to select the leptonic \( W \) decay with those that select the leptonic \( Z \) decay. The jet selection is not subjected to changes.

The cuts replaced in the electron selection are:
- \( \text{electron identification: } \textit{tight} \); 
- \( \text{electron } p_T > 25 \text{ GeV}/c \); 
- \( \text{veto on events with a second medium electron with } p_T > 20 \text{ GeV}/c \).

The cuts added to the selection are:
- \( \text{electron identification: } \textit{medium} \); 
- \( \text{electron } p_T > 20 \text{ GeV}/c \); 
- \( \text{two reconstructed electron with opposite charge in the event} \).

In the muon selection I substituted the requirement:
- \( \text{veto on events with a second isolated combined muon with } p_T > 20 \text{ GeV}/c \) and \( |\eta| < 2.5 \); 

with:
- \( \text{two reconstructed muons with opposite charge in the event} \).

The cuts used to reconstruct the leptonic \( W \) decay:
- \( \text{missing transverse energy } E_T^{\text{miss}} > 25 \text{ GeV} \); 
- \( \text{transverse mass } M_T > 40 \text{ GeV}/c^2 \).

are replaced by those for the leptonic \( Z \) decay selection:
- \( \text{missing transverse energy } E_T^{\text{miss}} > 15 \text{ GeV} \); 
- \( \text{lepton-lepton invariant mass } 66 < M_{ll} < 116 \text{ GeV}/c^2 \).

The other steps of the selection remain unchanged.

The \( M_{jj} \) distributions obtained with these specific selections are displayed in Fig. 6.24. There is an overall agreement between data and Monte Carlo even if data are slightly below the MC expectations in the signal region. This could in part justify why the data are below the MC expectations in the signal region for the standard selection too. At the moment this discrepancy has not yet been understood.
Fig. 6.24: Invariant mass distributions of the two leading jets obtained with the selection of the $Z$ control sample for the electron (a) and muon (b) channels. For each plot, the distribution for data (dots) is superimposed to the expected SM predictions (filled area). The statistical error is reported for data (black crosses) and for SM expectations (dashed violet lines). The panel underneath displays the bin-by-bin ratio of data to Monte Carlo with error bars corresponding to the statistical error in data and MC.
6.6 Multijet QCD control region

The multijet QCD background is obtained through the data-driven method described in section 4.3. The selections with the reverse electron identification criteria and with the non-pointing muons are used to find the shape of the multijet QCD background. Since these selections are applied on data it is not guaranteed that all the events that pass the selection are effectively multijet QCD events. I applied the selection to the MC samples to evaluate the fraction of non multijet QCD events which passes the selection. Fig. 6.25 shows the data distributions and the non multijet QCD events expected for these selections. The blank gap between data and MC is constituted by multijet QCD events and it is about the 90% of data.

In the analysis I used the shape of the data distribution after this selection for the determination of the multijet QCD shape in the standard selection. However the shape of the distribution is in part affected by the shape of the non multijet QCD distributions. An improvement of the method could be done using as multijet QCD shape the shape of the data distribution removing the MC expectation from it.

![Invariant mass distributions of the two leading jets obtained from the electron selection with the inverted tight cut (a) and from the muon selection with the inverted $d_0$ cut (b).](image)

Fig. 6.25: Invariant mass distributions of the two leading jets obtained from the electron selection with the inverted tight cut (a) and from the muon selection with the inverted $d_0$ cut (b). For each plot, the distribution for data (dots) is superimposed to the expected electroweak predictions (filled area). The statistical error is reported for data (black crosses) and for the electroweak expectations (dashed violet lines).

6.7 $t\bar{t}$ control region

The $t\bar{t}$ background constitutes about the 6.9% of the data and is the one whose $M_{jj}$ distribution has the most irregular shape in this analysis. In fact it has multiple structures depending on the jets that are chosen to reconstruct the mass. The goodness with which the Monte Carlo reproduces the $t\bar{t}$ shape should be verified. This is done in this section.
using four selections which improves the quantity of $t\bar{t}$ events in data. The comparison between data and MC for these $t\bar{t}$ enriched selections is used to evaluate how the MC correctly reproduces the $t\bar{t}$ in collision data.

Since the $t\bar{t}$ sample has the highest jet multiplicity, the invariant mass distribution of the two leading jets for events with at least 4 jets is calculated. The distribution for the electron and the muon channel are shown in Fig. 6.26. These distributions are obtained by applying the standard selection but excluding the $\Sigma p_T^{j_{ith}}$ cut which would drastically reduce the four jet sample. It can be seen that this selection highly enriches the sample of $t\bar{t}$ events that constitute more than 50% of the whole sample.

Another characteristic of the $t\bar{t}$ events is that they have two jets which originate from $b$-quarks. An algorithm of $b$-tagging is used to test whether the jet is originated from a $b$-quark or not. A neural network uses the information of the tracks that match with the jet and on the basis of them (it mainly evaluates if tracks form secondary vertices or have high impact parameter) the network assigns a value to the jet. This value is termed COMBNN (or IP3D+JetFitter [73]) and the larger the value the higher is the probability that the jet originates from a $b$-quark. The Fig. 6.27 shows the COMBNN distribution for jets of signal events and for those of $t\bar{t}$ events, the latter have an higher average COMBNN. The efficiency of this technique in identifying $b$-jets depends on the jet’s $p_T$ but it has mean value of 70% for the cut COMBNN $> 2$. For this cut the mean rejection of light jets is of the order of 200 while for $c$-jets is about 9. The invariant mass distribution of the two leading jets in events with at least one $b$-tagged jet (Fig. 6.28) is a sample rich of $t\bar{t}$ events therefore I used it to control of the $t\bar{t}$ shape. The $\Sigma p_T^{j_{ith}}$ cut is not applied to obtain this
Fig. 6.27: COMBNN distributions for jets which pass the jet selection (section 4.2.4) and the JVF cut. The jets belong the events selected with the standard selection excluded the $\Sigma p_T$ cut. Plot (a) is obtained for the electron selection; (b) for the muon one. The distribution for $t\bar{t}$ background jets (cyan) is superimposed to the distribution for signal jets (red). The two distributions are normalized to unitary area. In this analysis the jets considered as $b$-jets are those with $\text{COMBNN} \geq 2$.

Furthermore I selected events with at least three jets requiring at least one $b$-tagged jets and at least two non $b$-tagged jets. The data which pass this selection are almost completely $t\bar{t}$ events. The invariant mass distribution obtained with the two leading jets non $b$-tagged is shown in Fig. 6.29 and nicely shows the peak of the hadronic $W$ produced by the top decay. The $\Sigma p_T$ cut is not applied to obtain this plots. The $t\bar{t}$ MC reproduces pretty well shape of data even if the MC magnitude is slightly less than expected.

A final check is made reversing the $\Sigma p_T$ cut. The $M_{jj}$ distributions for events with $\Sigma p_T > 70$ GeV/$c$ are displayed in Fig. 6.30

The plots obtained in this section from the control regions present a good agreement between data and MC which certifies the goodness of the $t\bar{t}$ background in reproducing true $t\bar{t}$ events.
6.7. $T\bar{T}$ CONTROL REGION

Fig. 6.28: Invariant mass distributions of the two leading jets obtained in events with at least one $b$-jets for the electron (a) and muon (b) channels. For each plot, the distribution for data (dots) is superimposed to the expected SM predictions (filled area). The statistical error is reported for data (black crosses) and for SM expectations (dashed violet lines). The panel underneath displays the bin-by-bin ratio of data to Monte Carlo with error bars corresponding to the statistical error in data and MC.

Fig. 6.29: Invariant mass distributions of the two leading jets not $b$-tagged obtained in events with at least one $b$-jets for the electron (a) and muon (b) channels. For each plot, the distribution for data (dots) is superimposed to the expected SM predictions (filled area). The statistical error is reported for data (black crosses) and for SM expectations (dashed violet lines). The panel underneath displays the bin-by-bin ratio of data to Monte Carlo with error bars corresponding to the statistical error in data and MC.
6.8 Alternative Selection

The first selection developed for the WW/WZ semileptonic analysis, described in previous sections of this chapter, aims at reaching a signal significance of at least 3 in a sample of about $1.33 \text{fb}^{-1}$. However as soon as some of the largest systematic uncertainties have been studied it was realized that the weakness of this selection is represented by the $S/B$ ratio which is not large enough compared to the size of the systematic uncertainty. For this reason I studied another selection which aims at improving the signal to background ratio at the price of reducing the signal significance. In the following I will compare the results obtained with this new selection with the one described in the previous section that I will refer to as “Sel1” and that I summarize here for clarity:

- $p_{Tjj} > 60 \text{GeV/c}$;
- $\Delta \phi(j_{1st}, j_{2nd}) < 2.5$;
- $p_{TW} < 65 \text{GeV/c}$;
- $\Sigma p_{T_{j_{1st}}} < 70 \text{GeV/c}$.

In this thesis I analysed $1.33 \text{fb}^{-1}$ of data, but the total integrated luminosity collected with the ATLAS detector in the 2011 amounts to about $5 \text{fb}^{-1}$ and a larger amount will be collected in 2012. Assuming that the increase pile-up contribution will not give a significant change in the efficiency of the analysis it can be assumed that a $S/\sqrt{B}$
ratio equal to 1.5 in 1.33 fb$^{-1}$ corresponds to 3 in 5 fb$^{-1}$. Setting as minimum limit for $S/\sqrt{B}$ ratio with 1.33 fb$^{-1}$ the value of 1.5 I studied a selection able to raise significantly the $S/B$ ratio. The choice of the particular cuts to apply does not uses an optimization strategy but is made by looking at the signal and background distributions and selecting those cuts which suppress the background by a larger percentage than the signal.

This selection shares with the previous one all the cuts until the $p_T^{\mu}$ cut is kept unchanged since it improves both the $S/B$ ratio and the background shape. The $E_T^{miss}$ and the muon $p_T$ cuts have the role to improve the $S/B$ ratio along with reducing the multijet QCD background which has a large uncertainty. The multijet QCD background in the muon channel is larger than in the electron channel since the muon $p_T$ threshold (20 GeV/c) is lower than that one required for the electrons (25 GeV/c). Fig. 6.31(a), (c), (e) show the muon transverse momentum distributions for events with $N_{jets} \geq 2$. The multijet QCD background covers mainly the region at low $p_T$ (Fig. 6.31(e)) and below 25 GeV/c it is about half of the whole background (Fig. 6.31(a)). In addition, the muon $p_T$ distribution for the signal is harder than for the background (Fig. 6.31(e)), therefore a cut $p_T > 25$ GeV/c reduces the multijet QCD and raises the $S/B$ ratio.

One of the variables where the multijet QCD is differently distributed with respect to the signal is the $E_T^{miss}$ (Fig. 6.31(b), (d), (f)). The multijet QCD covers mainly the region of $E_T^{miss}$ below 40 GeV while more than a half of the signal lies in the region $E_T^{miss} > 40$ GeV (Fig. 6.31(d)). For this reason, I selected events with $E_T^{miss} > 40$ GeV. The multijet QCD background percentage decreases from 9.6% (14.4%) to 3.2% (2.9%) for the electron (muon) channel. The cut also reduces the $Z + jets$ background which is the other process with a low average $E_T^{miss}$. The distributions of the $M_{jj}$ made with the $E_T^{miss}$ and muon $p_T$ cuts set respectively to 40 GeV and 25 GeV/c are compared to those ones obtained with the older thresholds in Fig. 6.32. It can be seen that the multijet QCD background (shown in grey in the plots) is highly reduced with the new selection. The overall agreement between the data and the SM predictions (lower panels in Fig. 6.32) is not significantly changed by the new selection. Table 6.13 shows how the $S/B$ and $S/\sqrt{B}$ ratios change with the new $E_T^{miss}$ and muon $p_T$ cuts. The loss of significance in the signal region is about 0.3 – 0.5 while the $S/B$ ratio is improved by about 0.3%. The sample compositions with the new and the old cuts are shown in Table 6.14. With the new cuts the multijet QCD background contribution is decreased to about 3% in both channels.

The $t\bar{t}$ background can be reduced by hardening the $p_{T_{WW}}$ cut. The quality of the description of the $p_{T_{WW}}$ is shown on Fig. 6.33(a) and the difference between the signal and $t\bar{t}$ background distributions for this variable is shown in Fig. 6.33(c). I chose to select events with $p_{T_{WW}} > 35$ GeV/c because in that range the fraction of diboson events is lower than the fraction of background events (Fig. 6.33(b)). The $t\bar{t}$ is reduced from 17 – 18% to 8 – 9%. The $M_{jj}$ distribution after the $p_{T_{WW}}$ cut is shown in Fig. 6.33(d) for the electron selection. The values of the $S/B$ and $S/\sqrt{B}$ ratios after the $p_{T_{WW}}$ cut are reported in the summary Table 6.15.

The twist ($T$) between two jets is defined as:

$$ T = \arctan \left( \frac{\Delta \phi(j_{1st},j_{2nd})}{\Delta \eta(j_{1st},j_{2nd})} \right) $$
Fig. 6.31: Muon $p_T$ distributions (left) and $E_{\text{miss}}$ distributions (right) in events with $N_{\text{jets}} \geq 2$ for the muon channel. Top: the distributions for data (dots) are superimposed to the Standard Model predictions. Center: distributions for signal (red) and multijet QCD background (grey) normalized to unitary area. Bottom: distributions for signal (red) and total background (cyan) normalized to unitary area.
Fig. 6.32: $M_{jj}$ distributions of the two leading jets for the muon (top) and electron (bottom) channels for events with $E_T^{miss} > 25$ GeV (left) and with $E_T^{miss} > 40$ GeV (right). For the muon channel the muon $p_T$ is also varied from 20 GeV/c (left) to $p_T > 25$ GeV/c (right). The distributions are obtained after the $p_T^{jj} > 60$ GeV/c cut. For each plot, the distribution for data (dots) is superimposed to the expected SM predictions (filled area). The statistical error is reported for data (black crosses) and for SM expectations (dashed violet lines). The panel underneath displays the bin-by-bin ratio of data to Monte Carlo with error bars corresponding to the statistical error in data and MC.
CHAPTER 6. SENSITIVITY STUDIES ON WW/WZ RESONANCE

Electron channel (pTjj cut)

<table>
<thead>
<tr>
<th>Cuts (GeV), (GeV/c)</th>
<th>S/√B</th>
<th>S/B(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>total</td>
<td>peak</td>
</tr>
<tr>
<td>E^{miss}_T &gt; 25, p_T &gt; 25</td>
<td>2.67 ± 0.03</td>
<td>2.90 ± 0.04</td>
</tr>
<tr>
<td>E^{miss}_T &gt; 40, p_T &gt; 25</td>
<td>2.35 ± 0.02</td>
<td>2.58 ± 0.04</td>
</tr>
</tbody>
</table>

Muon channel (pTjj cut)

<table>
<thead>
<tr>
<th>Cuts (GeV), (GeV/c)</th>
<th>S/√B</th>
<th>S/B(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>total</td>
<td>peak</td>
</tr>
<tr>
<td>E^{miss}_T &gt; 25, p_T &gt; 20</td>
<td>3.03 ± 0.03</td>
<td>3.26 ± 0.04</td>
</tr>
<tr>
<td>E^{miss}_T &gt; 40, p_T &gt; 25</td>
<td>2.57 ± 0.02</td>
<td>2.75 ± 0.03</td>
</tr>
</tbody>
</table>

Table 6.13: Significance and signal to background ratio after the pTjj cut considering both the whole sample which passes the cuts and the subsample of events that belongs to the interval [60, 110] GeV/c^2 of dijet mass for the two channels. The values are given for the two thresholds of E^{miss}_T and muon p_T.

Cuts (GeV), (GeV/c) | Percentage in the electron channel (pTjj cut) | Percentage in the muon channel (pTjj cut)
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>WW/WZ</td>
<td>tt</td>
</tr>
<tr>
<td>E^{miss}_T &gt; 25, p_T &gt; 25</td>
<td>1.09</td>
<td>14.15</td>
</tr>
<tr>
<td>E^{miss}_T &gt; 40, p_T &gt; 25</td>
<td>1.21</td>
<td>17.8</td>
</tr>
</tbody>
</table>

Table 6.14: Expected composition of the SM sample for events with pTjj > 60 GeV/c as a function of the E^{miss}_T and muon p_T thresholds.
Fig. 6.33: $p_{TWW}$ distributions for the electron channel (a,b,c). (a): distribution for data (dots) superimposed to the expected SM predictions (filled area). The statistical error is reported for data (black crosses) and for SM expectations (dashed violet lines). The panel underneath displays the bin-by-bin difference between data and Monte Carlo measured in standard deviations (residuals). The standard deviation is evaluated considering only the statistical error in data and in the MC samples. (b): distribution for signal (red) and background (cyan) events normalized to the same area. (c): distribution for signal (red) and $t\bar{t}$ background (blue) events normalized to the same area. $M_{jj}$ distribution after the $p_{TWW} < 35$ GeV/c cut for the electron channel (d). The distribution for data (dots) is superimposed to the expected SM predictions (filled area). The statistical error is reported for data (black crosses) and for SM expectations (dashed violet lines). The panel underneath displays the bin-by-bin ratio of data to Monte Carlo with error bars corresponding to the statistical error in data and MC.
CHAPTER 6. SENSITIVITY STUDIES ON WW/WZ RESONANCE

It is a variable which quantifies if the two jets tend to be separated by a larger angle in the azimuthal plane than in the longitudinal one. In the following I use the absolute value of the twist. The distribution of $|T|$ in data is well described from the MC predictions (Fig. 6.34(a)). The diboson signal lies at larger $|T|$ values than the background (Fig. 6.34(b)). In Fig. 6.34(c) are shown the distributions of $M_{jj}$ vs $|T|$ for signal and background. The plots show that the percentage of signal events in a given $|T|$ interval smoothly increases as $|T|$ increases. On the contrary the percentage of background events in a given $|T|$ interval is approximately constant as the $|T|$ increase up to $|T| = 1$. Moreover Fig. 6.34(c) shows that the background distribution in the region $|T| < 1$ is mostly concentrated in the low $M_{jj}$ region therefore a selection that rejects event at low $|T|$ values is also expected to modify the global $M_{jj}$ shape. I decide to select events with $|T| > 0.8$. The selection removes a larger fraction of background events than the signal ones in the $M_{jj}$ region of the diboson peak. The $M_{jj}$ distribution after this cut is shown in Fig. 6.35. The $W + \text{jets}$ background has now a plateau from 30 to 150 GeV/c$^2$ due to the twist cut. The values of the $S/B$ and $S/\sqrt{B}$ ratios after the twist cut are reported in the summary Table 6.15.

In this last selection I also exploited the $b$-tagging to veto events containing jets produced by the fragmentation of a $b$-quark. This has allowed to remove most of the remaining $t\bar{t}$ background. The $b$-jet veto is based on the presence of secondary vertexes and high impact parameter tracks as described in section 6.7. For a $b$-jet identification efficiency equal to 70% the probability for a light jet to fake a $b$-jet is 1/200 while for a $c$-jet is 1/9. These figures have to be taken as indicative values because both the efficiency and fake probability depend on the jet $p_T$. Fig. 6.36(a) shows the multiplicity of $b$-tagged jets for the data overlaid to the SM expectation. The $b$-jet multiplicities for signal and $t\bar{t}$ events are compared in Fig. 6.36(b). The loss of signal due to the veto is less than 10% while only about the 25% of the $t\bar{t}$ survives to this selection.

The jet-jet invariant mass distribution after the $b$-jet veto is shown for the electron and the muon channel in Fig. 6.37 and the values of the $S/B$ and $S/\sqrt{B}$ ratios are listed in Table 6.15 for the whole $M_{jj}$ range and for the peak signal region only. The definition of the range for the peak signal region, obtained by finding the $M_{jj}$ range with the highest $S/\sqrt{B}$ ratio, is [70, 100] GeV/c$^2$. This range is slightly narrower than the definition used in previous selection ([60, 110] GeV/c$^2$) in fact the tighter selection used here result in narrowing the $W \rightarrow jj$ peak.

In summary the new selection, that I will refer to as “Sel2”, corresponds to the following list of cuts:

- muon $p_T > 25$ GeV/c (only for muon channel);
- $E_{T}^{\text{miss}} > 40$ GeV;
- $p_{Tjj} > 60$ GeV/c;
- $p_{TWW} < 35$ GeV/c;
- $|T| > 0.8$;
- $N_{\text{btags}} = 0$. 

Fig. 6.34: Wide view of the $|T|$ distribution. For the description of each single plot refer to the beginning of section 6.2.4. The electron selection is applied until the $p_{TW} < 35 \text{ GeV/c}$ cut.
Fig. 6.35: Invariant mass distributions of the two leading jets after the $|T| > 0.8$ cut for the electron channel. (a): distribution for data (dots) superimposed to the expected SM predictions (filled area). The statistical error is reported for data (black crosses) and for SM expectations (dashed violet lines). The panel underneath displays the bin-by-bin ratio of data to Monte Carlo with error bars corresponding to the statistical error in data and MC. (b): distribution for overall background events (cyan) superimposed to the distribution for signal events (red). The two distributions are normalized to unitary area.

<table>
<thead>
<tr>
<th>Cut</th>
<th>$S/\sqrt{B}$ (total)</th>
<th>$S/\sqrt{B}$ (peak)</th>
<th>$S/B(%)$ (total)</th>
<th>$S/B(%)$ (peak)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_{Tjj}$</td>
<td>2.35 ± 0.02</td>
<td>2.56 ± 0.04</td>
<td>1.22 ± 0.01</td>
<td>3.62 ± 0.06</td>
</tr>
<tr>
<td>$p_{TWW}$</td>
<td>1.96 ± 0.02</td>
<td>2.38 ± 0.05</td>
<td>1.40 ± 0.02</td>
<td>4.47 ± 0.10</td>
</tr>
<tr>
<td>$</td>
<td>T</td>
<td>$</td>
<td>1.62 ± 0.03</td>
<td>2.11 ± 0.05</td>
</tr>
<tr>
<td>$N_{blags} = 0$</td>
<td>1.61 ± 0.03</td>
<td>2.09 ± 0.05</td>
<td>1.54 ± 0.03</td>
<td>5.42 ± 0.15</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Cut</th>
<th>$S/\sqrt{B}$ (total)</th>
<th>$S/\sqrt{B}$ (peak)</th>
<th>$S/B(%)$ (total)</th>
<th>$S/B(%)$ (peak)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_{Tjj}$</td>
<td>2.57 ± 0.02</td>
<td>2.70 ± 0.04</td>
<td>1.21 ± 0.01</td>
<td>3.46 ± 0.06</td>
</tr>
<tr>
<td>$p_{TWW}$</td>
<td>2.11 ± 0.02</td>
<td>2.49 ± 0.04</td>
<td>1.37 ± 0.02</td>
<td>4.32 ± 0.09</td>
</tr>
<tr>
<td>$</td>
<td>T</td>
<td>$</td>
<td>1.74 ± 0.02</td>
<td>2.18 ± 0.05</td>
</tr>
<tr>
<td>$N_{blags} = 0$</td>
<td>1.73 ± 0.02</td>
<td>2.14 ± 0.05</td>
<td>1.52 ± 0.02</td>
<td>5.06 ± 0.13</td>
</tr>
</tbody>
</table>

Table 6.15: Significance and signal to background ratio after the application of each cut of Sel2 considering both the whole sample which passes the cut and the subsample of events that belongs to the interval [70, 100] GeV/c² of dijet mass.
Fig. 6.36: Distributions of the number of jets b-tagged with the COMBNN cut ($N_{b\text{tags}}$). The electron selection is applied until the $|T| > 0.8$ cut. (a): distribution for data (dots) superimposed to the expected SM predictions (filled area). The statistical error is reported for data (black crosses) and for SM expectations (dashed violet lines). The panel underneath displays the bin-by-bin difference between data and Monte Carlo measured in standard deviations (residuals). The standard deviation is evaluated considering only the statistical error in data and in MC samples. (b): distribution for $t\bar{t}$ background events (blue) superimposed to the distribution for signal events (red). The two distributions are normalized to unitary area.
Fig. 6.37: Invariant mass distributions of the two leading jets for the electron (a) and muon (b) selections after the $N_{b\text{tags}} = 0$ cut. The distribution for data (dots) is superimposed to the expected SM predictions (filled area). The statistical error is reported for data (black crosses) and for SM expectations (dashed violet lines). The panel underneath displays the bin-by-bin ratio of data to Monte Carlo (red crosses) with error bars corresponding to the statistical error in data and MC, the relative systematic error due to the JES uncertainty (woven cyan bands) and the sum in quadrature of the two errors (black bands).
In Sel2, each of the cut applied raises the $S/B$ ratio significantly (Table 6.15) until it reaches a final value of 5.4% (5.1%) in the signal region for the electron (muon) channel. On the other hand, in the same region, the $S/\sqrt{B}$ ratio decreases significantly reaching the value 2.09 (2.14) for the electron (muon) channel. However the $S/\sqrt{B}$ ratio is well above 1.5 in both channels and it is expected to be about 4 in 5 $fb^{-1}$. The effect of each single cut of the Sel2 selection on the $M_{jj}$ distributions for the signal, $W + jets$, $t\bar{t}$ and multijet QCD is shown in Fig. B.29 of the Appendix B.

The final $M_{jj}$ distributions obtained for Sel2 selection are shown in Fig. 6.37 for the electron (top) and muon (bottom) channels. The lower panels show the bin-by-bin data over MC ratio. In these plots both the statistical and systematic uncertainty are shown. The systematic uncertainty on the $M_{jj}$ distribution due to JES uncertainty (including the contribution due to the pile-up uncertainty) has been estimated using the method described in section 6.4. In the signal region the relative systematic error is about 3.2%, therefore although Sel2 has a larger $S/B$ ratio than Sel1 it also shows a larger systematic uncertainty (in Sel1 the systematic uncertainty in the signal peak region is about 2.1%). However only a careful study of all the main systematics can establish which of the two selections leads the smaller systematic uncertainty on the $M_{jj}$ distribution. A deeper study on this has to be done as well as a search for cuts which improves more the $S/B$ ratio without raising the effect of the JES systematic uncertainty.

### 6.9 Final remarks and further steps

The final number of events and efficiencies expected for the different processes for Sel1 and Sel2 are compared in Table 6.16. As expected the efficiencies of Sel1 are larger both for the signal and the background than those ones obtained for Sel2. The larger background reduction is obtained for the $t\bar{t}$ and the multijet QCD processes. Further studies should be focused on the reduction of the $W + jets$ background which is still giving a large contribution. One idea in development is the use of a tagger to distinguish jets emerging from quark or gluon fragmentation. A first version of the quark-gluon tagger has just been developed in ATLAS and it is based on the information of the tracks associated to the jet. This tool could be used to remove a part of the $W + jets$ background requiring that the two leading jets in the event originate from quarks as it is expected for the signal. In $W + jets$ events jets can be originated by a quark or a gluon. A MC based study has shown that, in the signal region, in only about 20% of the $W + jets$ events both jets selected as $W$ decay candidates originate from quarks. If this technique will demonstrate to work we can expect a significant improvement on the $S/B$ ratio. However a deeper study of the quark-gluon tagger has to be done to also assess the systematic uncertainty introduced by this type of selection.

A further improvement in the selection technique could be done using a selection based on likelihood ratios instead than on successive cuts. This is also under study.

Finally it should be noted that any new development must be confronted with the impact given on the systematic uncertainty. It is however expected that the JES uncertainty, one of the main sources of systematic uncertainty, will be significantly decreased using data-driven techniques. Hopefully this will greatly help the extraction of the signal in this analysis.
Table 6.16: Expected number of events in 1.33 \( fb^{-1} \) for the SM processes considered in this analysis. The numbers are provided for the events with \( E_T^{miss} > 25 \) GeV and \( N_{jets} \geq 2 \), for the selection used to maximize the \( S/\sqrt{B} \) ratio (Sel1) and that used to improve more the \( S/B \) ratio (Sel2). It is also provided the efficiency of the Sel1 and Sel2 on events with \( N_{jets} \geq 2 \) for each physics process.
Conclusions

In the course of this thesis two main topics have been treated: the investigation of the jet-jet resonances produced in association with a leptonically decaying \( W \); the study of the \( WW/WZ \) semileptonic decay. These studies are based on the data collected by the ATLAS experiment in the 2011.

The search for the dijet resonances produced in association with a \( W \) leptonic decay has been done using about 1.02 \( fb^{-1} \). Since there is not yet an accredited hypothesis on the mechanism which produces the excess at CDF, there are no predicted cross-sections for this hypothetical process at LHC. For this reason the analysis in ATLAS has been done using a selection as close as possible to the one with which the CDF collaboration has observed the excess. The effect of the main systematic uncertainties has been evaluated. The best fit to data has then been obtained varying the Monte Carlo predictions within the systematic uncertainties. The agreement between data and the SM predictions in the \( M_{jj} \) region \([100, 300]\) GeV/c\(^2\) has been determined with a statistical algorithm (BumpHunter algorithm). No significant excess over the Standard Model expectation has been found in the mass range considered. Since no hypothesis has been done on the signal cross-section nothing can be stated about the compatibility of the ATLAS and CDF measurements but only that there is no significant disagreement between the 1 \( fb^{-1} \) of data collected by ATLAS and the SM predictions. This analysis on the search for the CDF excess in the dijet mass distribution has been published in July 2011.

The ATLAS analysis of the diboson decay in the semileptonic channel has begun in 2011. Since the beginning the ATLAS Pisa group has played a leading role in the analysis and I contributed to it with the studies shown in this thesis. The selection I developed are, at present, the most efficient and most promising selections to arrive at the measurement of the \( WW/WZ \) cross-section. Since the signal to background ratio at LHC is 5 times less than at Tevatron the signal observation is much more difficult. Nevertheless the possibility to measure the \( WW/WZ \) signal with the ATLAS experiment is not precluded. In fact both the statistical and the systematic errors are going to decrease with the use of the full 2011 statistics (5 \( fb^{-1} \)) and with a better understanding of the jet energy scale.

The studies of the \( WW/WZ \) decay in the semileptonic channel that I developed have been performed on 1.33 \( fb^{-1} \). Data have been used to extract the multijet QCD distributions and to check the goodness of the MC simulations. The SM expectations have been carefully investigated in search of selections which increases the probability to measure the signal. Two studies have been developed and they have to be considered as test of the feasibility of the diboson measurement in the semileptonic channel with a cut based method.

One selection has been thought to maximize the \( S/\sqrt{B} \) once an initial cut is applied
to suitably shape the background $M_{jj}$ distribution and to improve the $S/B$ ratio. The $S/\sqrt{B}$ and $S/B$ ratios in the $M_{jj}$ region [60, 110] GeV/$c^2$ are respectively $2.98 \pm 0.05$ and $(2.88 \pm 0.05)\%$ for the electron channel while $3.40 \pm 0.05$ and $(2.99 \pm 0.05)\%$ for the muon one. The systematic uncertainty on the $M_{jj}$ distribution due to JES uncertainty is about $2.1\%$.

The second selection focused on the increase of the $S/B$ ratio. The $S/\sqrt{B}$ and $S/B$ ratios obtained at the end of the selection in the $M_{jj}$ region [70, 100] GeV/$c^2$ are respectively $2.09 \pm 0.05$ and $(5.42 \pm 0.15)\%$ for the electron channel while $2.14 \pm 0.05$ and $(5.06 \pm 0.13)\%$ for the muon one. The systematic uncertainty on the $M_{jj}$ distribution due to JES uncertainty is about $3.2\%$.

From these results emerges that the main problem of this analysis is constituted by the smallness of the $S/B$ ratio with respect to the systematic error. The expected reduction of the statistical and systematic errors will favour the diboson measurement, however it will be fundamental also the development of new analysis techniques and tools which could increase more the $S/B$ ratio. For this purpose, the most promising improvement, at the moment, seems to be the introduction of quark-gluon tagger in the analysis to select jets from quarks. The tests made by the ATLAS Pisa group show that the $S/B$ ratio could increase up to reach $20\%$. 

128
Appendix A

Additional plots for the CDF bump analysis

This appendix gathers additional distributions obtained with the selection described in chapter 5. It is divided in five parts that collect the lepton distributions (section A.1) and the jet ones (section A.2). One section is dedicated to the $M_T$ distributions (section A.3), another to the $E_T^{\text{miss}}$ distributions (section A.4) which have been used to determine the multijet QCD distributions, and the last to the $M_{jj}$ distributions (section A.5). The data used for this section correspond to an integrated luminosity of 1.02 fb$^{-1}$.

A.1 Lepton’s variables

The distribution of the kinematic variables $p_T$, $E$, $\eta$ and $\phi$ of the leptons selected with the electron and the muon channels are shown in Fig. A.1 and A.2. Fig. A.1 collects the electron’s kinematic variable distributions as they appear at the end of the CDF selection (section 5.1). Fig. A.2 shows the same distributions for the muon.

A.2 Jet’s variables

The jet-jet invariant mass distribution is calculated with the two leading jets in $p_T$ among those that are selected. For these jets, the energy and transverse momentum distributions are shown in Fig. A.3 and A.4 for the electron and muon selections respectively.
Fig. A.1: Transverse momentum $p_T$(a), energy $E$ (b), $\eta$ (c) and $\phi$ (d) distributions of the electron selected at the end of the CDF cut-flow. For each plot, the distribution for data (dots) is superimposed to the expected SM predictions (filled area). The statistical error is reported for data (black crosses) and for SM expectations (dashed violet lines). The panel underneath displays the bin-by-bin ratio of data to Monte Carlo with error bars corresponding to the statistical error in data and MC.
A.2. JET’S VARIABLES

Fig. A.2: Transverse momentum $p_T$(a), energy $E$ (b), $\eta$ (c) and $\phi$ (d) distributions of the muon selected at the end of the CDF cut-flow. For each plot, the distribution for data (dots) is superimposed to the expected SM predictions (filled area). The statistical error is reported for data (black crosses) and for SM expectations (dashed violet lines). The panel underneath displays the bin-by-bin ratio of data to Monte Carlo with error bars corresponding to the statistical error in data and MC.
Fig. A.3: Transverse momentum $p_T$ (left) and energy $E$ (right) distributions of the leading jet (top) and of the sub-leading one (bottom). The distributions are obtained at the end of the CDF resonance selection with one electron identified. For each plot, the distribution for data (dots) is superimposed to the expected SM predictions (filled area). The statistical error is reported for data (black crosses) and for SM expectations (dashed violet lines). The panel underneath displays the bin-by-bin ratio of data to Monte Carlo with error bars corresponding to the statistical error in data and MC.
Fig. A.4: Transverse momentum $p_T$ (left) and energy $E$ (right) distributions of the leading jet (top) and of the sub-leading one (bottom). The distributions are obtained at the end of the CDF resonance selection with one muon identified. For each plot, the distribution for data (dots) is superimposed to the expected SM predictions (filled area). The statistical error is reported for data (black crosses) and for SM expectations (dashed violet lines). The panel underneath displays the bin-by-bin ratio of data to Monte Carlo with error bars corresponding to the statistical error in data and MC.


A.3 \( M_T \) distributions

The transverse mass distributions for the electron selection and the muon one after requiring at least two jets in the event is shown in Fig. A.5

![Transverse mass distributions](image)

Fig. A.5: Transverse mass (\( M_T \)) distributions after the \( M_T \) cut for events with at least two jets and which pass the electron selection (a) or the muon selection (b). The distribution for data (dots) is superimposed to the expected SM predictions (filled area). The statistical error is reported for data (black crosses) and for SM expectations (dashed violet lines).

A.4 \( E_T^{\text{miss}} \) distributions

This section shows the \( E_T^{\text{miss}} \) distributions at each step of the selection for the CDF bump search. The fit to these distributions has been used to find the normalization of the multijet QCD distributions and the one of the \( W + \text{jets} \) distributions. The distributions in Fig. A.6 are obtained using the electron selection and those in Fig. A.7 using the muon one. The \( E_T^{\text{miss}} \) distributions are shown after the cuts: \( M_T > 40 \text{ GeV}/c^2, N_{jets} \geq 2, p_{Tjj} > 40 \text{ GeV}/c, \Delta \eta(j_{1st}, j_{2nd}) < 2.5, \Delta \phi(j_{1st}, E_T^{\text{miss}}) > 0.4 \) and \( N_{jets} = 2 \).

A.5 \( M_{jj} \) distributions

The \( M_{jj} \) distributions after the \( N_{jets} \geq 2, p_{Tjj} > 40 \text{ GeV}/c, \Delta \eta(j_{1st}, j_{2nd}) < 2.5, \Delta \phi(j_{1st}, E_T^{\text{miss}}) > 0.4 \) requirements are shown in Fig. A.8 and Fig. A.9 for the electron and the muon channels respectively. For each cut, the data distribution and the SM one are shown and the ratio of the two is provided along with the relative statistical and systematic errors. Fig. A.10 compares the \( M_{jj} \) distributions for \( W + \text{jets} \) events and multijet QCD ones and underlines the likeness of the shape of the two distributions.
A.5. $M_{jj}$ DISTRIBUTIONS

Fig. A.6: $E_T^{\text{miss}}$ distributions after the $M_T > 40 \text{ GeV}/c^2$ (a), $N_{\text{jets}} \geq 2$ (b), $p_{T,ij} > 40 \text{ GeV}/c$ (c), $\Delta \eta(j_{1,\text{st}}, j_{2,\text{nd}}) < 2.5$ (d), $\Delta \phi(j_{1,\text{st}}, E_T^{\text{miss}}) > 0.4$ (e), $N_{\text{jets}} = 2$ (f) cuts for the electron selection. For each plot, the distribution for data (dots) is superimposed to the expected SM predictions (filled area). The statistical error is reported for data (black crosses) and for SM expectations (dashed violet lines). The panel underneath displays the bin-by-bin ratio of data to Monte Carlo with error bars corresponding to the statistical error in data and MC.
Fig. A.7: $E_{\text{miss}}$ distributions after the $M_T > 40 \text{ GeV}/c^2$ (a), $N_{\text{jets}} \geq 2$ (b), $p_{T,j} > 40 \text{ GeV}/c$ (c), $\Delta\eta(j_{1st}, j_{2nd}) < 2.5$ (d), $\Delta\phi(j_{1st}, E_T^{\text{miss}}) > 0.4$ (e), $N_{\text{jets}} = 2$ (f) cuts for the muon selection. For each plot, the distribution for data (dots) is superimposed to the expected SM predictions (filled area). The statistical error is reported for data (black crosses) and for SM expectations (dashed violet lines). The panel underneath displays the bin-by-bin ratio of data to Monte Carlo with error bars corresponding to the statistical error in data and MC.
A.5. $M_{jj}$ DISTRIBUTIONS

Fig. A.8: Jet-Jet invariant mass distributions as a function of the cuts applied for the electron channel: $N_{jets} \geq 2$ (a), $p_{T jj} > 40$ GeV/c (b), $\Delta \eta(j_{1st}, j_{2nd}) < 2.5$ (c), $\Delta \phi(j_{1st}, E_{miss}) > 0.4$ (d). For each plot, the distribution for data (dots) is superimposed to the expected SM predictions (filled area). The statistical error is reported for data (black crosses) and for SM expectations (dashed violet lines). The panel underneath displays the bin-by-bin ratio of data to Monte Carlo (red crosses) with error bars corresponding to the statistical error in data and MC, the relative systematic error due to the JES uncertainty (woven cyan bands) and the sum in quadrature of the two errors (black bands).
Fig. A.9: Jet-Jet invariant mass distributions as a function of the cuts applied for the muon channel: $N_{jets} \geq 2$ (a), $p_{Tjj} > 40 \text{ GeV/c}$ (b), $\Delta y(j_{1st},j_{2nd}) < 2.5$ (c), $\Delta \phi(j_{1st}, E_{T^{miss}}) > 0.4$ (d). For each plot, the distribution for data (dots) is superimposed to the expected SM predictions (filled area). The statistical error is reported for data (black crosses) and for SM expectations (dashed violet lines). The panel underneath displays the bin-by-bin ratio of data to Monte Carlo (red crosses) with error bars corresponding to the statistical error in data and MC, the relative systematic error due to the JES uncertainty (woven cyan bands) and the sum in quadrature of the two errors (black bands).
Fig. A.10: $M_{jj}$ distributions for $W + jets$ (green) and multijet QCD (grey) samples at the end of the CDF resonance selection for the electron selection (a) and the muon one (b). The two distributions are normalized to unitary area.
APPENDIX A. ADDITIONAL PLOTS FOR THE CDF BUMP ANALYSIS
Appendix B

Additional plots for the diboson analysis

This appendix gets together additional distributions which refer to the $WW/WZ$ studies described in chapter 6. The first two sections contain those distributions for the muon channel which are shown in chapter 6 only for the electron channel. The distributions for the Sel1 are shown in section B.1 while those for the Sel2 in section B.2. The lepton distributions and those of the jets are collected in section B.3 and section B.4 respectively. One section is dedicated to the $E_T^{miss}$ distributions (section B.5), an other to additional distributions studied (section B.6).

B.1 Sel1 distributions (muon selection)

This section collects the distribution obtained for the muon channel of the variables used in the Sel1. The distributions shown are those of the:

- $p_{Tjj}$ (Fig. B.1);
- $p_{Tjj}$ vs $M_{jj}$ (B.2(a));
- $\Delta \phi (j_{1st}, j_{2nd})$ (Fig. B.3);
- $\Delta \phi (j_{1st}, E_T^{miss})$ (Fig. B.5);
- $\Delta \phi (W_{lep}, W_{had})$ (Fig. B.6);
- $balance(W_{lep}, W_{had})$ (Fig. B.7);
- $p_{TW}$ (Fig. B.8);
- $N_{jets}$ (Fig. B.9);
- $\Sigma p_{Tj_{jth}}$ (Fig. B.10);
- $M_{jj}$ (Fig. B.2(b,c), Fig. B.4, Fig. B.9(c), Fig. B.10(c)).
Fig. B.1: Transverse momentum distribution of the system of two leading jets for the muon selection. (a): distribution for data (dots) superimposed to the expected SM predictions (filled area). The statistical error is reported for data (black crosses) and for SM expectations (dashed violet lines). The panel underneath displays the bin-by-bin ratio of data to Monte Carlo with error bars corresponding to the statistical error in data and MC. (b): distribution for overall background events (cyan) superimposed to the distribution for signal events (red). The two distributions are normalized to unitary area.
B.1. SEL1 DISTRIBUTIONS (MUON SELECTION)

Fig. B.2: On the top: two dimensional distributions of the angle $p_{Tjj}$ as a function of the $M_{jj}$ for signal events (left) and for background events (right) for the muon selection. The bin color indicates the fraction of events in that bin. Bins filled with warmer colors contain a larger statistics. On the bottom: $M_{jj}$ distribution for events with $p_{Tjj} > 60$ GeV/c (b) and for those events with $p_{Tjj} < 60$ GeV/c (c). The muon selection is applied. The distribution for data (dots) is superimposed to the expected SM predictions (filled area). The statistical error is reported for data (black crosses) and for SM expectations (dashed violet lines). The panel underneath displays the bin-by-bin ratio of data to Monte Carlo with error bars corresponding to the statistical error in data and MC. The background for $p_{Tjj} < 60$ GeV/c peaks on the mass region of the signal and few events have $M_{jj} < 60$ GeV/c^2. For $p_{Tjj} > 60$ GeV/c the maximum of the background $M_{jj}$ distribution is at 30 GeV/c^2 and the distribution is monotonically decreasing.
Fig. B.3: Distributions of the azimuthal angle between the two leading jets ($\Delta \phi (j_{1\text{st}}, j_{2\text{nd}})$) for the muon channel. (a): the distribution for data (dots) is superimposed to the expected SM predictions (filled area). The statistical error is reported for data (black crosses) and for SM expectations (dashed violet lines). The panel underneath displays the bin-by-bin ratio of data to Monte Carlo with error bars corresponding to the statistical error in data and MC. (b): the distribution for overall background events (cyan) is superimposed to the distribution for signal events (red). The two distributions are normalized to unitary area.
B.1. SEL1 DISTRIBUTIONS (MUON SELECTION)

Fig. B.4: Invariant mass distributions of the two leading jets for the muon selection after the $\Delta \phi (j_{1ST}, j_{2ND}) < 2.5$ cut. (a): the distribution for data (dots) is superimposed to the expected SM predictions (filled area). The statistical error is reported for data (black crosses) and for SM expectations (dashed violet lines). The panel underneath displays the bin-by-bin ratio of data to Monte Carlo with error bars corresponding to the statistical error in data and MC. (b): the distribution for overall background events (cyan) is superimposed to the distribution for signal events (red). The two distributions are normalized to unitary area.
Fig. B.5: Wide view of the distribution of the azimuthal angle between the $E_T^{miss}$ and the leading jet ($\Delta \phi(j_{1st}, E_T^{miss})$). For the description of each single plot refer to the beginning of section 6.2.4. The muon selection is applied until the $\Delta \phi(j_{1st}, j_{2nd}) < 2.5$ cut.
Fig. B.6: Wide view of the distribution of the azimuthal angle between the $W_{had}$ and the $W_{lep}$ ($\Delta \phi(W_{lep}, W_{had})$). For the description of each single plot refer to the beginning of section 6.2.4. The muon selection is applied until the $\Delta \phi(j_{1st}, j_{2nd}) < 2.5$ cut.
Fig. B.7: Wide view of the distribution of the momentum balance of the $W_{had}$ and the $W_{lep}$ ($balance(W_{lep}, W_{had})$). For the description of each single plot refer to the beginning of section 6.2.4. The muon selection is applied until the $\Delta \phi(j_{1st}, j_{2nd}) < 2.5$ cut.
Fig. B.8: Wide view of the distribution of the transverse momentum resulting from the vectorial sum of the $W_{\text{had}}$ $p_T$ and the $W_{\text{lep}}$ $p_T$ ($p_{TWW}$). For the description of each single plot refer to the beginning of section 6.2.4. In plot (d), the $S/\sqrt{B}$ ratio is shown as a function of the upper $p_{TWW}$ threshold. The muon selection is applied until the $\Delta\phi(j_{1\text{st}},j_{2\text{nd}}) < 2.5$ cut.
APPENDIX B. ADDITIONAL PLOTS FOR THE DIBOSON ANALYSIS

Fig. B.9: Wide view of the distribution of the jet multiplicity ($N_{\text{jets}}$). For the description of each single plot (except (c)) refer to the beginning of section 6.2.4. In plot (d), the $S/\sqrt{B}$ ratio is shown as a function of the upper $N_{\text{jets}}$ threshold. The muon selection is applied until the $p_{\text{TWW}} < 65$ GeV/c cut. (c): jet-jet invariant mass distribution obtained for $N_{\text{jets}} = 2$. The distribution for data (dots) is superimposed to the expected SM predictions (filled area). The statistical error is reported for data (black crosses) and for SM expectations (dashed violet lines). The panel underneath displays the bin-by-bin ratio of data to Monte Carlo with error bars corresponding to the statistical error in data and MC.
Fig. B.10: Wide view of the $\Sigma_{pT_{j1h}}$ distribution. For the description of each single plot (except (c)) refer to the beginning of section 6.2.4. In plot (d), the $S/\sqrt{B}$ ratio is shown as a function of the upper $\Sigma_{pT_{j1h}}$ threshold. The muon selection is applied until the $p_{TWW} < 65$ GeV/c cut. (c): jet-jet invariant mass distribution obtained after the $\Sigma_{pT_{j1h}} < 70$ GeV/c cut. The distribution for data (dots) is superimposed to the expected SM predictions (filled area). The statistical error is reported for data (black crosses) and for SM expectations (dashed violet lines). The panel underneath displays the bin-by-bin ratio of data to Monte Carlo with error bars corresponding to the statistical error in data and MC.
B.2 Sel2 distributions (muon selection)

This section collects the distribution obtained for the muon channel of the variables used in the Sel2. The distributions shown are those of the:

- $p_{Tjj}$ (Fig. B.11);
- $p_{TW}$ (Fig. B.12);
- $|T|$ (Fig. B.13);
- $N_{b\text{tags}}$ (Fig. B.15);
- $M_{jj}$ (Fig. B.14).
Fig. B.11: Transverse momentum distribution of the system of two leading jets for the muon selection. (a): distribution for data (dots) superimposed to the expected SM predictions (filled area). The statistical error is reported for data (black crosses) and for SM expectations (dashed violet lines). The panel underneath displays the bin-by-bin ratio of data to Monte Carlo with error bars corresponding to the statistical error in data and MC. (b): distribution for overall background events (cyan) superimposed to the distribution for signal events (red). The two distributions are normalized to unitary area.
Fig. B.12: $p_{TWW}$ distributions for the muon channel (a,b,c). (a): distribution for data (dots) superimposed to the expected SM predictions (filled area). The statistical error is reported for data (black crosses) and for SM expectations (dashed violet lines). The panel underneath displays the bin-by-bin difference between data and Monte Carlo measured in standard deviations (residuals). The standard deviation is evaluated considering only the statistical error in data and in the MC samples. (b): distribution for signal (red) and background (cyan) events normalized to the same area. (c): distribution for signal (red) and $t\bar{t}$ background (blue) events normalized to the same area. $M_{jj}$ distribution after the $p_{TWW} < 35 \text{ GeV/c}$ cut for the muon channel (d). The distribution for data (dots) is superimposed to the expected SM predictions (filled area). The statistical error is reported for data (black crosses) and for SM expectations (dashed violet lines). The panel underneath displays the bin-by-bin ratio of data to Monte Carlo with error bars corresponding to the statistical error in data and MC.
Fig. B.13: Wide view of the $|T|$ distribution. For the description of each single plot refer to the beginning of section 6.2.4. The muon selection is applied until the $p_{TWW} < 35$ GeV/c cut.
Fig. B.14: Invariant mass distributions of the two leading jets after the $|T| > 0.8$ cut for the muon channel. (a): distribution for data (dots) superimposed to the expected SM predictions (filled area). The statistical error is reported for data (black crosses) and for SM expectations (dashed violet lines). The panel underneath displays the bin-by-bin ratio of data to Monte Carlo with error bars corresponding to the statistical error in data and MC. (b): distribution for overall background events (cyan) superimposed to the distribution for signal events (red). The two distributions are normalized to unitary area.
### B.2. SEL2 DISTRIBUTIONS (MUON SELECTION)

The muon selection is applied until the \(|T| > 0.8\) cut. (a): distribution for data (dots) superimposed to the expected SM predictions (filled area). The statistical error is reported for data (black crosses) and for SM expectations (dashed violet lines). The panel underneath displays the bin-by-bin difference between data and Monte Carlo measured in standard deviations (residuals). The standard deviation is evaluated considering only the statistical error in data and in MC samples. (b): distribution for \(t\bar{t}\) background events (blue) superimposed to the distribution for signal events (red). The two distributions are normalized to unitary area.

---

**Table:**

<table>
<thead>
<tr>
<th>btag</th>
<th>N_{btag}</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
</tr>
</tbody>
</table>

**Figure B.15:** Distributions of the number of jets b-tagged with the COMBNN cut \(N_{btags}\). The muon selection is applied until the \(|T| > 0.8\) cut.
B.3 Lepton’s variables

The distribution of the kinematic variables $p_T$, $E$, $\eta$, and $\phi$ of the leptons selected with the electron and the muon channels are shown in Fig. B.16 and B.17. Fig. B.16 collects the electron’s kinematic variable distributions as they appear at the end of the Sel1. Fig. A.2 shows the same distributions for the muon.

B.4 Jet’s variables

The jet-jet invariant mass distribution is calculated with the two leading jets in $p_T$ among those that are selected. For these jets, the energy and transverse momentum distributions are shown in Fig B.18 and B.19 for the electron and muon selections respectively. These results are obtained at the end of the $WW/WZ$ Sel1.

B.5 $E_T^{miss}$ distributions

This section shows the $E_T^{miss}$ distributions at each step of the Sel1 and Sel2 developed in the $WW/WZ$ studies. The fit to these distributions has been used to find the normalization of the multijet QCD distributions and the one of the $W+jets$ distributions. The distributions in Fig. B.20 are obtained using the electron Sel1 and those in Fig. B.21 using the muon Sel1. The $E_T^{miss}$ distributions are shown after the cuts: $M_T > 40\text{ GeV}/c^2$, $N_{jets} \geq 2$, $p_{Tjj} > 60\text{ GeV}/c$, $\Delta \phi(j_{1st}, j_{2nd}) < 2.5$, $p_{TWW} < 65\text{ GeV}/c$, $\Sigma p_{T_{j_{th}}} < 70\text{ GeV}/c$. The $E_T^{miss}$ distributions for the cuts of the Sel2 ($p_{TWW} < 35\text{ GeV}/c$, $|T| \geq 0.8$ and $N_{btags} = 0$) are displayed in Fig. B.22 for both the channels.
Fig. B.16: Transverse momentum $p_T$ (a), energy $E$ (b), $\eta$ (c) and $\phi$ (d) distributions of the electron selected at the end of the WW/WZ Sel1. For each plot, the distribution for data (dots) is superimposed to the expected SM predictions (filled area). The statistical error is reported for data (black crosses) and for SM expectations (dashed violet lines). The panel underneath displays the bin-by-bin ratio of data to Monte Carlo with error bars corresponding to the statistical error in data and MC.
Fig. B.17: Transverse momentum $p_T$ (a), energy $E$ (b), $\eta$ (c) and $\phi$ (d) distributions of the muon selected at the end of the WW/WZ Sel1. For each plot, the distribution for data (dots) is superimposed to the expected SM predictions (filled area). The statistical error is reported for data (black crosses) and for SM expectations (dashed violet lines). The panel underneath displays the bin-by-bin ratio of data to Monte Carlo with error bars corresponding to the statistical error in data and MC.
Fig. B.18: Transverse momentum $p_T$ (left) and energy $E$ (right) distributions of the leading jet (top) and of the sub-leading one (bottom). The distributions are obtained at the end of the $WW/WZ$ Sel1 with one electron identified. For each plot, the distribution for data (dots) is superimposed to the expected SM predictions (filled area). The statistical error is reported for data (black crosses) and for SM expectations (dashed violet lines). The panel underneath displays the bin-by-bin ratio of data to Monte Carlo (red crosses) with error bars corresponding to the statistical error in data and MC.
APPENDIX B. ADDITIONAL PLOTS FOR THE DIBOSON ANALYSIS

Fig. B.19: Tansverse momentum $p_T$ (left) and energy $E$ (right) distributions of the leading jet (top) and of the sub-leading one (bottom). The distributions are obtained at the end of the $WW/WZ$ Sel1 with one muon identified. For each plot, the distribution for data (dots) is superimposed to the expected SM predictions (filled area). The statistical error is reported for data (black crosses) and for SM expectations (dashed violet lines). The panel underneath displays the bin-by-bin ratio of data to Monte Carlo (red crosses) with error bars corresponding to the statistical error in data and MC.
B.5. $E_T^{\text{miss}}$ DISTRIBUTIONS

Fig. B.20: $E_T^{\text{miss}}$ distributions after the $M_T > 40 \text{ GeV}/c^2$ (a), $N_{\text{jets}} \geq 2$ (b), $p_{T,jj} > 60 \text{ GeV}/c$ (c), $\Delta \phi(j_{\text{jet}},j_{\text{lead}}) < 2.5$ (d), $p_{T\text{WW}} < 65 \text{ GeV}/c$ (e), $\Sigma p_{T,j_{\text{lead}}} < 70 \text{ GeV}/c$ (f) cuts for the electron selection of the WW/WZ analysis. For each plot, the distribution for data (dots) is superimposed to the expected SM predictions (filled area). The statistical error is reported for data (black crosses) and for SM expectations (dashed violet lines). The panel underneath displays the bin-by-bin ratio of data to Monte Carlo with error bars corresponding to the statistical error in data and MC.
APPENDIX B. ADDITIONAL PLOTS FOR THE DIBOSON ANALYSIS

Fig. B.21: $E_T^{miss}$ distributions after the $M_T > 40$ GeV/c$^2$ (a), $N_{jets} \geq 2$ (b), $p_T^{jet} > 60$ GeV/c (c), $\Delta\phi(j_{lead},j_2) < 2.5$ (d), $p_T^{WW} < 65$ GeV/c (e), $\Sigma p_T^{jets} < 70$ GeV/c (f) cuts for the muon selection of the WW/WZ analysis. For each plot, the distribution for data (dots) is superimposed to the expected SM predictions (filled area). The statistical error is reported for data (black crosses) and for SM expectations (dashed violet lines). The panel underneath displays the bin-by-bin ratio of data to Monte Carlo with error bars corresponding to the statistical error in data and MC.
DISTRIBUTIONS

Fig. B.22: $E_T^{\text{miss}}$ distributions after the $p_T^{\ell\ell} < 35$ GeV/c (a,b), $|T| \geq 0.8$ (c,d) and $N_{\text{btags}} = 0$ (e,f) cuts for the electron (left) and muon (right) selections. For each plot, the distribution for data (dots) is superimposed to the expected SM predictions (filled area). The statistical error is reported for data (black crosses) and for SM expectations (dashed violet lines). The panel underneath displays the bin-by-bin ratio of data to Monte Carlo with error bars corresponding to the statistical error in data and MC.
B.6 Further distributions

In this section are reported the distributions of all the variables investigated in search of those which could improve more the $S/\sqrt{B}$ and the $S/B$ ratios. Each plot in this section is obtained for the electron selection and shows the distribution for background events (cyan) superimposed to the distribution for signal events (red). The two distributions are normalized to unitary area to highlight the differences between the shapes of the signal and background. The variables investigated are combinations of the kinematic properties of the leading jet ($j_{1st}$), the subleading jets ($j_{2nd}$), the electron ($e$) and the $E_T^{miss}$.

Fig B.23 shows the distributions of:

(a): the distance $\Delta R$ between the two leading jets;
(b): the angle between the momenta of the two leading jets;
(c): the vectorial difference of the jet momenta;
(d): the scalar difference of the jet momenta;
(e): the scalar sum of the jet momenta;
(f): the twist between the two jets;
(g): the distance $\Delta \eta$ between the two leading jets;
(h): the distance $\Delta \phi$ between the two leading jets;
(i): the energy of the dijet system;
(j): the $\eta$ of the dijet system;
(k): the $p_T$ of the dijet system;
(l): the $p_T$ of the $e$-$E_T^{miss}$ system.

Fig B.24 shows the distributions of:

(a): the distance $\Delta \phi$ between $j_{1st}$ and $e$;
(b): the distance $\Delta \phi$ between $j_{2nd}$ and $e$;
(c): the vectorial difference of the momenta of $j_{1st}$ and $e$;
(d): the vectorial difference of the momenta of $j_{2nd}$ and $e$;
(e): the scalar difference of the momenta of $j_{1st}$ and $e$;
(f): the scalar difference of the momenta of $j_{2nd}$ and $e$;
(g): the $p_T$ of the $j_{1st}$-$e$ system;
(h): the $p_T$ of the $j_{2nd}$-$e$ system;
(i): the scalar sum of the momenta of $j_{1st}$ and $e$;
Fig. B.23: Distribution for background events (cyan) superimposed to the distribution for signal events (red). Refer to text for a detailed description.
APPENDIX B. ADDITIONAL PLOTS FOR THE DIBOSON ANALYSIS

(j): the scalar sum of the momenta of $j_{2nd}$ and $e$;
(k): the distance $\Delta \phi$ between $j_{1st}$ and $E_T^{miss}$;
(l): the distance $\Delta \phi$ between $j_{2nd}$ and $E_T^{miss}$;

Fig B.25 shows the distributions of:

(a): the vectorial difference of the momenta of $j_{1st}$ and $E_T^{miss}$;
(b): the vectorial difference of the momenta of $j_{2nd}$ and $E_T^{miss}$;
(c): the scalar difference of the momenta of $j_{1st}$ and $E_T^{miss}$;
(d): the scalar difference of the momenta of $j_{2nd}$ and $E_T^{miss}$;
(e): the $p_T$ of the $j_{1st}$-$E_T^{miss}$ system;
(f): the $p_T$ of the $j_{2nd}$-$E_T^{miss}$ system;
(g): the scalar sum of the momenta of $j_{1st}$ and $E_T^{miss}$;
(h): the scalar sum of the momenta of $j_{2nd}$ and $E_T^{miss}$;
(i): the vectorial difference of the momenta of $e$ and $E_T^{miss}$;
(j): the scalar difference of the momenta of $e$ and $E_T^{miss}$;
(k): the scalar sum of the momenta of $e$ and $E_T^{miss}$;
(l): the distance $\Delta \phi$ between $e$ and $E_T^{miss}$;
Fig. B.24: Distribution for background events (cyan) superimposed to the distribution for signal events (red). Refer to text for a detailed description.
Fig. B.25: Distribution for background events (cyan) superimposed to the distribution for signal events (red). Refer to text for a detailed description.
The systematic error has been evaluated also at the beginning of the selection when it is required \( N_{jets} \geq 2 \). The systematic error in the signal region amounts to 1.7%. Fig. B.26 compare the shifted mass distributions to the standard one at this step of the analysis. The systematics are evaluated also for the jet multiplicity distribution at the beginning of the selection (Fig. B.27). Data agrees with SM expectations within the systematic error.

Fig. B.26: Jet-jet invariant mass distributions for electron (a) and muon (b) selections after the \( M_T > 40 \text{ GeV}/c^2 + N_{jets} \geq 2 \) cuts. For each plot, the distributions obtained by shifting the JES up (red) and down (blue) by its uncertainty are compared to the distribution obtained with nominal JES (black). The black dots correspond to data. The errors bars are the statistical errors. The panel underneath displays the bin-by-bin ratio of data to Monte Carlo (red crosses) with error bars corresponding to the statistical error in data and MC and the cyan bands represent the relative systematic uncertainty due to the JES uncertainty estimated as the maximum difference between the shifted bin values and the unshifted one.
Fig. B.27: Jet multiplicity distributions after the $M_T$ cut for electron (a) and muon (b) selections. Only the jets which pass the jet selection (section 4.2.4) and the JVF cut are counted. For each plot, the distribution for data (dots) is superimposed to the expected SM predictions (filled area). The statistical error is reported for data (black crosses) and for SM expectations (dashed violet lines). The vertical scale is logarithmic. The panel underneath displays the bin-by-bin ratio of data to Monte Carlo (red crosses) with error bars corresponding to the statistical error in data and MC, the relative systematic error due to the JES uncertainty (woven cyan bands) and the sum in quadrature of the two errors (black bands).
B.8. OTHER PLOTS

B.8 Other plots

Fig. B.28: Dijet invariant mass distributions for the signal (a), the $W + \text{partons}$ (b), the $t\bar{t}$ (c) and multijet QCD (d). For each sample, the $M_{jj}$ distributions after the $N_{jets} \geq 2$ cut (black), the $p_T^{jj} > 60$ GeV/c (blue), the $\Delta \phi(j_{1st}, j_{2nd}) < 2.5$ cut (green), the $p_{TWW} < 65$ GeV/c cut (purple) and the $\Sigma p_T^{j_{ith}} < 70$ GeV/c cut (red) is provided. In these plots the electron selection is applied.
Fig. B.29: Dijet invariant mass distributions for the signal (a), the $W + \text{partons}$ (b), the $t\bar{t}$ (c) and multijet QCD (d). For each sample, the $M_{jj}$ distributions after the $N_{jets} \geq 2$ cut (black), the $p_T^{jj} > 60$ GeV/c (blue), the $p_T^{TWW} < 35$ GeV/c cut (purple), the $|T| > 0.8$ cut (green) and the $N_{btags} = 0$ cut (red) is provided. In these plots the electron selection is applied.
Bibliography


[66] Internal results obtained by G. Zevi della Porta.


[72] Internal results obtained by G. Zevi della Porta.