



**UNIVERSITÀ  
DEGLI STUDI  
DI PADOVA**

Sede Amministrativa: UNIVERSITÀ DEGLI STUDI DI PADOVA  
Dipartimento di Fisica "G. Galilei"

SCUOLA DI DOTTORATO DI RICERCA IN FISICA - CICLO XXV

# **Study of the ZZ diboson production at CDF II**

**Direttore della Scuola:** Prof. Andrea Vitturi

**Supervisore:** Prof.ssa Donatella Lucchesi

**Dottorando:** Dott. Matteo Bauce



*To those who never give up*

*If you're in pitch blackness,  
all you can do is sit tight  
until your eyes get used to the dark.  
- H. M.*



---

# Study of the ZZ diboson production at CDF II

## Abstract

The subject of this Thesis is the production of a pair of massive  $Z$  vector bosons in the proton antiproton collisions at the Tevatron, at the center-of-mass energy  $\sqrt{s} = 1.96$  TeV. We measure the  $ZZ$  production cross section in two different leptonic decay modes: into four charged leptons ( $e$  or  $\mu$ ) and into two charged leptons plus two neutrinos. The results are based on the whole dataset collected by the Collider Detector at Fermilab (CDF), corresponding to  $9.7 \text{ fb}^{-1}$  of data. The combination of the two cross section measurements gives  $\sigma(p\bar{p} \rightarrow ZZ) = 1.38_{-0.27}^{+0.28}$  pb, and is the most precise  $ZZ$  cross section measurement at the Tevatron to date.

We further investigate the four lepton final state searching for the production of the scalar Higgs particle in the decay  $H \rightarrow ZZ^{(*)} \rightarrow \ell\ell\ell'\ell'$ . No evidence of its production has been seen in the data, hence was set a 95% Confidence Level upper limit on its production cross section as a function of the Higgs particle mass,  $m_H$ , in the range from 120 to 300 GeV/ $c^2$ .

---



---

# Studio della produzione di dibosoni ZZ a CDF II

## Prefazione

In questa Tesi è studiata la produzione di coppie di bosoni vettori massivi  $Z$  nelle collisioni protone antiprotone a Tevatron, ad un'energia nel centro di massa di  $\sqrt{s} = 1.96$  TeV. La sezione d'urto di produzione di  $ZZ$  è misurata in due diversi modi di decadimento leptonici: in quattro leptoni carichi ( $e$  o  $\mu$ ) e in due leptoni carichi e due neutrini. I risultati si basano su tutto il campione di dati raccolto dal Collider Detector at Fermilab (CDF), corrispondente a  $9.7 \text{ fb}^{-1}$  di dati. La combinazione delle due misure di sezione d'urto dà come risultato  $\sigma(p\bar{p} \rightarrow ZZ) = 1.38^{+0.28}_{-0.27} \text{ pb}$ , ed è al momento la misura più precisa di produzione di  $ZZ$  al Tevatron.

Abbiamo analizzato ulteriormente lo stato finale con quattro leptoni carichi ricercando la produzione della particella scalare di Higgs nel decadimento  $H \rightarrow ZZ \rightarrow \ell\ell'\ell''$ . Non è stata osservata nessuna evidenza della produzione di questa particella nei dati quindi si è ricavato un limite superiore ad un Livello di Confidenza del 95% sulla sezione d'urto di produzione in funzione della massa del bosone di Higgs  $m_H$ , nell'intervallo compreso tra 120 e 300  $\text{GeV}/c^2$ .

---

# Contents

<b>Introduction</b>	<b>iii</b>
<b>1 The Standard Model</b>	<b>1</b>
1.1 Fundamental Particles and Interactions	2
1.1.1 Quantum Chromo Dynamics	3
1.1.2 Electroweak Interactions	4
1.2 Electroweak Symmetry Breaking and the Higgs Boson	7
<b>2 Diboson Production and Higgs Phenomenology</b>	<b>11</b>
2.1 Massive Boson Mediators: $Z$ and $W^\pm$	11
2.1.1 $Z$ properties from $e^+e^-$ colliders	12
2.2 Physics Measurements of Diboson Final States	14
2.2.1 LEP II	15
2.2.2 Tevatron	16
2.2.3 LHC	18
2.3 Higgs phenomenology	18
2.3.1 Theoretical constraints	19
2.3.2 Higgs boson production and decay	20
2.3.3 Experimental Situation	22
<b>3 The Tevatron Collider and CDF experiment</b>	<b>25</b>
3.1 Tevatron Accelerator Complex	25
3.1.1 Proton Source and Acceleration	26
3.1.2 Main Injector	27
3.1.3 Antiproton Production and Storage	27
3.1.4 The Tevatron Collider	28
3.2 The CDF Detector	28
3.2.1 Tracking System	30
3.2.2 Calorimeter System	35
3.2.3 Muon Detector System	37
3.2.4 Cherenkov luminosity counters	39
3.2.5 Online Data Acquisition	40
<b>4 Physical Object Reconstruction</b>	<b>47</b>
4.1 Track Reconstruction	48
4.1.1 Tracking Algorithms	48
4.2 Primary Vertex Identification	50
4.3 Electron Identification	50
4.4 Muon Identification	52

---



4.4.1	Tracks Identification of unknown lepton flavor . . . . .	54
4.5	Jet Identification . . . . .	55
4.5.1	CDF Cone Algorithm . . . . .	57
4.5.2	Jet Correction . . . . .	58
4.6	Neutrino Identification: Energy Unbalance . . . . .	58
4.7	Trigger paths and efficiencies . . . . .	59
4.7.1	Trigger Efficiencies . . . . .	59
<b>5</b>	<b>ZZ reconstruction in leptonic final states</b>	<b>61</b>
5.1	$ZZ \rightarrow \ell\ell'\ell'$ . . . . .	63
5.1.1	Data Collection and Offline Selection . . . . .	63
5.1.2	Monte Carlo Simulations . . . . .	64
5.1.3	Data Driven Background Estimation . . . . .	68
5.1.4	Event Selection . . . . .	73
5.1.5	Systematic Uncertainties . . . . .	77
5.1.6	Cross section measurement . . . . .	78
5.2	$ZZ \rightarrow \ell\nu\nu$ . . . . .	80
5.2.1	Initial Sample Composition . . . . .	81
5.2.2	Kinematic Properties and Signal Selection . . . . .	83
5.2.3	Test of the Background Modeling in a Control Region . . . . .	92
5.2.4	Artificial Neural Network for Signal Extraction . . . . .	98
5.2.5	Systematic Uncertainties . . . . .	100
5.2.6	Cross Section Measurement . . . . .	107
<b>6</b>	<b>Standard Model Higgs Search in four lepton final state</b>	<b>109</b>
6.0.7	Event Reconstruction and Sample Composition . . . . .	110
6.0.8	Background Determination . . . . .	115
6.0.9	Event kinematics . . . . .	117
6.0.10	Systematic Uncertainties . . . . .	121
6.0.11	Limit Calculation on the Higgs production cross section . . . . .	122
6.0.12	Results . . . . .	124
<b>7</b>	<b>Discussion of the Results</b>	<b>127</b>
7.1	Diboson measurements . . . . .	127
7.1.1	CDF ZZ cross section measurement combination . . . . .	127
7.1.2	Dibosons at hadron colliders . . . . .	129
7.2	Higgs search summary . . . . .	132
7.2.1	Tevatron Higgs searches sensitivity . . . . .	132
7.2.2	LHC Higgs searches sensitivity . . . . .	132
<b>8</b>	<b>Conclusions</b>	<b>137</b>
<b>A</b>	<b>Lepton Reconstruction Categories</b>	<b>139</b>
A.1	Reconstructed Lepton Categories . . . . .	139
A.1.1	Tight Central Electrons . . . . .	139
A.1.2	Likelihood Based Electrons . . . . .	141
A.1.3	Forward Electrons . . . . .	143
A.1.4	Central Muons . . . . .	144

---

---

A.1.5	Intermediate Muons . . . . .	145
A.1.6	Forward Muons . . . . .	146
A.1.7	Stubless Muons . . . . .	147
A.1.8	Isolated Tracks . . . . .	147
A.2	Lepton Probes and Denominator Objects . . . . .	148
A.3	Lepton reconstruction efficiencies and scale factors . . . . .	152
<b>B</b>	<b>Single Object Triggers</b>	<b>157</b>
B.1	Trigger efficiencies . . . . .	158
<b>C</b>	<b>Lepton momentum corrections</b>	<b>163</b>

---

# Introduction

The past decades of high-energy physics research have been dedicated to investigate the Standard Model of particle physics, as a theory explaining the strong and electroweak interactions between the fundamental particles. Several experiments focused on the precise measurement of the physics observables theoretically predicted by this model at particle colliders, considering processes in a wide range of production rates.

An interesting sector of the Standard Model is the production of a pair of massive gauge bosons ( $Z, W^\pm$ ) in high-energy particle collisions. The coupling between the gauge bosons are particularly sensitive to the underlying theoretical model that describes particle interactions. Many exotic models of particle physics predict deviation from the Standard Model which can be significantly observed in processes involving the production of a pair of massive vector bosons.

Diboson production is sensitive to anomalous trilinear Gauge couplings [1] and large extra dimension models [2] that can modify the rate of boson pair production. Additional contribution to diboson production can be due to the presence of a high-mass resonant state ( $X$ ) which kinematically can decay to a boson pair. The simplest example is the case of the Higgs boson, the scalar particle introduced by the electroweak symmetry breaking to justify the mass of the fundamental particles. Other than the Higgs boson, several Standard Model extensions predict the existence of new resonances that can decay to dibosons (e.g. SUSY,  $Z'$ , Randall-Sundrum gravitons).

For this reason, the experimental measurement of diboson production is a fundamental step to proceed in further tests of the Standard Model and search for new physics phenomena.

In this Thesis we focus on the associated production of two  $Z$  bosons in  $p\bar{p}$  collisions at the Tevatron, at  $\sqrt{s} = 1.96$  TeV. Among all the dibosons this is the one expected to be produced less frequently, according to the Standard Model predictions.

We studied  $ZZ^{(*)1}$  production in two different decay modes:

$$p\bar{p} \rightarrow ZZ^{(*)} \rightarrow \ell^+ \ell^- \ell'^+ \ell'^- \quad p\bar{p} \rightarrow ZZ^{(*)} \rightarrow \ell^+ \ell^- \nu \bar{\nu} ,$$

exploiting all the data collected by the CDF II experiment. The available experimental sample, corresponding to  $9.7 \text{ fb}^{-1}$  of integrated luminosity, and improved analysis techniques allowed to measure the production cross section for this process with a precision of the order of 20%.

The measurement of the  $ZZ^{(*)}$  production cross section at CDF sets the basis for further investigation of similar processes. Among the two decay modes considered for the cross section measurement we focused on the four charged lepton final state to study the production of the Higgs boson in  $p\bar{p}$  collisions. The main contribution expected of Higgs boson production in the aforementioned final state according to the Standard Model is due to the Higgs boson decay to  $ZZ$ :

$$p\bar{p} \rightarrow H \rightarrow ZZ^{(*)} \rightarrow \ell^+ \ell^- \ell'^+ \ell'^-$$

---

<sup>1</sup> The superscript  $(*)$  represents a possible virtual  $Z$  boson.

but minor contributions are expected from different Higgs boson production mechanisms. This process is considered the *golden channel* for the Higgs boson search at the LHC, where this particle is expected to be produced with higher cross sections than at the Tevatron. The analysis of the full available CDF dataset would set the achievable sensitivity at CDF for the Higgs search in this final state, sustained by the precise  $ZZ$  cross section measurement in the same decay mode. Moreover, this sensitivity study will contribute to the overall searches of the Higgs boson in all its possible decay modes.

In **Chapter 1** of this Thesis is presented the Standard Model of particle physics, which is the reference frame for the processes studied in this work.

The state of the Art of the experimental knowledge of massive gauge bosons is summarized in **Chapter 2**, with particular reference to the phenomenology of the diboson production. Here is also reported an overview of the experimental measurements of  $VV$  production at electron and hadron colliders, as well as the current status on the search for the Higgs boson.

The experimental environment where the measurements took place is described in **Chapter 3**: the  $p\bar{p}$  Tevatron collider at Fermilab and the CDF experiment.

The information from the CDF detector are analyzed with appropriate algorithms and tools described in **Chapter 4** to reconstruct physical objects (e.g. particle tracks, energy deposits) in the particle collisions.

We analyze the whole sample of data collected by CDF (corresponding to  $9.7 \text{ fb}^{-1}$ ) to measure the  $ZZ$  production cross section as described in **Chapter 5**, considering two leptonic decay modes:  $ZZ \rightarrow \ell^+ \ell^- \ell'^+ \ell'^-$  and  $ZZ \rightarrow \ell^+ \ell^- \nu \bar{\nu}$ .

In **Chapter 6** we study the sensitivity of CDF to the search for the Standard Model Higgs boson in four charged lepton events, setting an upper limit on the production cross section as a function of its mass  $m_H$ .

The results obtained from the analyses presented in this Thesis are combined and discussed in **Chapter 7**. We compare the result obtained at CDF within the context of the Tevatron and LHC most recent results.

In **Chapter 8** we conclude the Thesis with a brief summary of the work presented and its impact on the current knowledge of the Standard Model.

---

# The Standard Model

## Contents

1.1	Fundamental Particles and Interactions . . . . .	2
1.1.1	Quantum Chromo Dynamics . . . . .	3
1.1.2	Electroweak Interactions . . . . .	4
1.2	Electroweak Symmetry Breaking and the Higgs Boson . . . . .	7

The Standard Model (SM) of particles physics is a gauge field theory which incorporates both quantum mechanics and Einstein's theory of special relativity in the attempt to describe the fundamental particles and their interactions. The Standard Model describes three of the four known fundamental forces of nature: strong interactions, electromagnetic interactions and weak interactions. The fourth force, gravity, is far weaker<sup>1</sup> and is not expected to contribute significantly to the physical processes which are of current interest in high energy particle physics. The Standard Model is described by the gauge group

$$SU(3)_C \times SU(2)_L \times U(1)_Y \quad (1.1)$$

and is invariant under local gauge transformations. The  $C$  is a reminder that  $SU(3)$  represents the symmetry group of the *colored* strong interactions of quantum chromodynamics (QCD). The  $L$  indicates that the  $SU(2)$  group contains left-handed weak isospin doublets and the  $Y$  is a reminder that the  $U(1)$  group contains the right-handed weak hypercharge singlets. Together, the  $SU(2)_L \times U(1)_Y$  groups govern the unified electroweak force.

There are two main classifications of particles within the SM: the spin- $\frac{1}{2}$  *fermions* that are the constituents of normal matter and the integer spin *bosons* which are also the mediators of the strong and electroweak forces. The mass of the particles in the SM is ascribed to the interaction with the so-called Higgs field which is the result of a spontaneously broken symmetry arising in the  $SU(2)_L \times U(1)_Y$  electroweak sector. The predicted Higgs boson resulting from this broken symmetry is the particle in the SM which has been investigated in the more recent years at lepton and hadron colliders. An evidence of a resonant state compatible with being this one has been observed from the experiments located at the Large Hadron Collider (LHC) at CERN (Geneva). This chapter describes the fundamental particles, their interactions, and a short overview of the Higgs mechanism.

<sup>1</sup> Roughly 40 orders of magnitude smaller than the strong nuclear force.

## 1.1 Fundamental Particles and Interactions

### Fermions

Fermions are spin- $\frac{1}{2}$  fundamental particles that appear as two different types: quarks and leptons. Quarks are the constituents of familiar composite particles such as the proton and neutron, but they can combine to form other hadrons which are bound state of quark-antiquark pairs or 3 quarks called mesons and baryons respectively. Quarks interact via both strong and electroweak forces. Leptons, such as the electron, have only electroweak interactions.

**Quarks** There are six types of quarks plus their anti-quarks: up, down, charm, strange, top and bottom. Quarks have fractional electric charge ( $Q$ ) and a quantum property called *color charge* similar to the electric charge of electromagnetism, but in three types  $r$ ,  $g$  and  $b$ . The charges ( $Q$ ) and the masses of the quarks are listed in Table 1.1 along with their flavor.

	Flavor	Charge ( $Q$ )	Mass (GeV)
1st generation	$u$	$+2/3$	$(1.5 - 3.3) \times 10^{-3}$
	$d$	$-1/3$	$(3.5 - 6.0) \times 10^{-3}$
2nd generation	$c$	$+2/3$	$1.27^{+0.07}_{-0.11}$
	$s$	$-1/3$	$104^{+26}_{-34} \times 10^{-3}$
3rd generation	$t$	$+2/3$	$171.2 \pm 2.1$
	$b$	$-1/3$	$4.2^{+0.17}_{-0.07}$

Table 1.1: Flavor, charge and mass [3] of the quarks.

Quarks are subject to both strong interactions as well as electroweak interactions which are discussed in section 1.1.1 and 1.1.2 respectively.

**Leptons** There are six types of leptons (plus their anti-particles) in the SM. These are the electron, muon, tau ( $e$ ,  $\mu$ ,  $\tau$ ) and their respective neutrinos ( $\nu_e$ ,  $\nu_\mu$ ,  $\nu_\tau$ ). They are classified by their charge and lepton quantum numbers ( $L_e$ ,  $L_\mu$  and  $L_\tau$ ) as shown in Table 1.2.

Flavor	Charge ( $Q$ )	Lepton Numbers	Mass (GeV)
$e$	-1	$L_e = 1$	$0.511 \times 10^{-3}$
$\nu_e$	0	$L_e = 1$	$< 225 \times 10^{-9}$
$\mu$	-1	$L_\mu = 1$	$105.7 \times 10^{-3}$
$\nu_\mu$	0	$L_\mu = 1$	$< 0.19 \times 10^{-3}$
$\tau$	-1	$L_\tau = 1$	1.777
$\nu_\tau$	0	$L_\tau = 1$	$< 18.2 \times 10^{-3}$

Table 1.2: Charge, lepton number and mass [3] for the leptons.

Leptons are subject to the electroweak force. They are colorless and thus do not participate in strong interactions.

Although the neutrino masses are listed in Table 1.2 only with an upper limit, there is now strong evidence from neutrino mixing measurements that they are in fact non-zero [4]. Also, it is worth to note that the  $\tau$  lepton is the only lepton with enough mass to decay to final states containing hadrons, and it does it with a branching fraction of 65% [3].

## Bosons - Force Mediators

SM interactions are mediated by spin-1 bosons. The gluon ( $g$ ) is the mass-less mediator of the strong force. The photon ( $\gamma$ ),  $W^\pm$  and  $Z$  are the force carriers of the electroweak interactions. The photon is mass-less while the  $W^\pm$  and  $Z$  are massive particles. The gauge bosons and their properties are summarized in Table 1.3. The role of force carriers in particle interactions is described in Sections 1.1.1 and 1.1.2. Since this Thesis focuses on the study of dibosons production, more details about the gauge massive bosons will be given in Chapter 2.

	Interactions	Charge	Mass [GeV/c <sup>2</sup> ]	Width [GeV/c <sup>2</sup> ]
$g$	strong	0	0	
$\gamma$	electromagnetic	0	0	
$Z$	electro-weak	0	$91.1876 \pm 0.0021$	$2.4952 \pm 0.0023$
$W^\pm$	(electro-)weak	$\pm 1$	$80.398 \pm 0.025$	$2.141 \pm 0.041$

Table 1.3: Summary of the force carriers and their masses [3]

### 1.1.1 Quantum Chromo Dynamics

Quantum chromodynamics (QCD) is the theory of the strong interactions between quarks and gluons. Quarks carry a single color charge while a gluon is a color-anticolor state and is the mediator of color flow. In  $SU(3)$  the  $3 \times \bar{3}$  colors ( $rgb$ ) give nine total color states for the gluon: an octet

$$\frac{(r\bar{r} + g\bar{g} - 2b\bar{b})}{\sqrt{6}} \quad \frac{(r\bar{r} - g\bar{g})}{\sqrt{2}} \quad c\bar{c}' , \quad (1.2)$$

where in the last expressions  $c$  and  $c'$  can be any different colors, and a color singlet

$$\frac{(r\bar{r} + b\bar{b} + g\bar{g})}{\sqrt{3}} . \quad (1.3)$$

However, the singlet is colorless and so in nature there are only 8 possible colored gluons. Systems of quarks and quark-antiquark only exist in colorless bound states with integer charge. For instance the combination of three quarks (e.g. the proton), can form the colorless state:

$$\frac{(rgb - grb + gbr - bgr + brg - rbg)}{\sqrt{6}} . \quad (1.4)$$

Also quark-antiquark combinations can form colorless bound states (e.g. the pion  $\pi^+$ ) of the form:

$$\frac{(r\bar{r} + b\bar{b} + g\bar{g})}{\sqrt{3}} . \quad (1.5)$$

The feature of the strong force is that the coupling becomes increasingly large with separation distance. This indicates that colored partons will be confined in objects which are as a whole colorless. The coupling constant of QCD ( $\alpha_s$ ) is a running function of the 4-momentum squared transferred in the interaction, given by the equation

$$\alpha_s(q^2) = \frac{12\pi}{(33 - 2n_f) \log(q^2/\Lambda^2)} . \quad (1.6)$$

Here  $\Lambda \sim 0.1$  GeV indicates the lower energy limit at which SM should be considered valid (as approximation of more extended theories) and  $n_f$  is the number of quark flavors whose mass is greater than the  $q^2$  of interest [5]. At very large  $q^2$  (corresponding to very short approach distances)  $\alpha_s$  becomes small. This phenomenon is known as asymptotic freedom. This property allows, for high- $q^2$  interaction, perturbative expansion of QCD processes which remain finite. Gluons couple to quarks and other gluons via the  $qqg, ggg$  and  $gggg$  vertexes shown in Figure 1.1. Color, charge and flavor are always conserved in strong interactions.

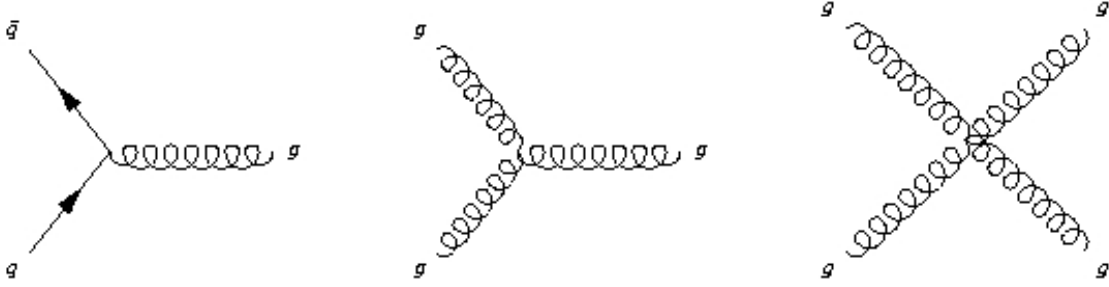


Figure 1.1: The primary vertexes of QCD.

A more detailed description of QCD theory can be found in [6, 7]

### 1.1.2 Electroweak Interactions

The electroweak interaction of quarks and leptons is described by the  $SU(2)_L \times U(1)_Y$  gauge group. Weak isospin ( $T_L$ ) and hypercharge ( $Y$ ) are the respective generators of symmetry transformations. They satisfy the equation

$$Q = T_3 + \frac{1}{2}Y \quad (1.7)$$

where  $T_3$  is the projection of the third component of the weak isospin vector.

The electroweak Lagrangian is given by

$$\mathcal{L} = -\frac{1}{4}W^{\mu\nu}W_{\mu\nu} - B^{\mu\nu}B_{\mu\nu} + \bar{\psi}i\gamma^\mu D_\mu\psi \quad (1.8)$$

where the covariant derivative is given by

$$D_\mu = \partial_\mu + igW_\mu T + \frac{1}{2}ig'B_\mu Y. \quad (1.9)$$

$T$  is the weak isospin operator,  $g$  and  $g'$  are two different electroweak coupling constants and  $B_{\mu\nu}$ , similar to the electromagnetic field tensor, is given by

$$B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu \quad (1.10)$$

where  $B_\nu$  is the massless gauge field representing the singlet of  $U(1)_Y$ .  $W_\mu$  are the gauge fields of  $SU(2)$  and  $W_{\mu\nu}$  the field tensor which is defined as

$$W_{\mu\nu} = \partial_\mu W_\nu - \partial_\nu W_\mu - gW_\mu \times W_\nu \quad (1.11)$$



The electroweak gauge fields  $W_\mu^1$ ,  $W_\mu^2$ ,  $W_\mu^3$  and  $B_\mu$  used to write the electroweak Lagrangian have basically a theoretical meaning. Instead, in particle physics we can express  $\mathcal{L}$  as function of four physical fields:  $A_\mu$ ,  $Z_\mu$ ,  $W_\mu^+$  and  $W_\mu^-$ .  $A_\mu$  is the (neutral) electromagnetic field,  $Z_\mu$  is the field corresponding to electroweak neutral current while  $W_\mu^\pm$  correspond to the electroweak charged currents. By requiring the electromagnetic and weak forces to be unified and to describe the gauge bosons observed experimentally it is required that there be two neutral and two charged bosons. Thus the electromagnetic field  $A$  and neutral current  $Z$  must be some linear combination of the unified electroweak fields. This can be written in terms of the electroweak mixing angle  $\theta_W$  as

$$\begin{pmatrix} Z \\ A \end{pmatrix} = \begin{pmatrix} \cos\theta_W & -\sin\theta_W \\ \sin\theta_W & \cos\theta_W \end{pmatrix} \cdot \begin{pmatrix} W^3 \\ B \end{pmatrix} \quad (1.12)$$

from which can be shown that the parameters  $g$  and  $g'$  have the relation  $g' = g \tan\theta_W$  and are also related to the charge of the electron  $e$  by the relation  $e = g \sin\theta_W = g' \cos\theta_W$ . The remaining two components of the  $W_\mu$  are then related to the observables  $W^+$  and  $W^-$ . The real fields are then given by

$$W^\pm = \frac{1}{\sqrt{2}}(W_\mu^1 \mp iW_\mu^2) \quad (1.13)$$

$$Z_\mu = \frac{-g'B_\mu + gW_\mu^3}{\sqrt{g^2 + g'^2}} \quad (1.14)$$

$$A_\mu = \frac{gB_\mu + g'W_\mu^3}{\sqrt{g^2 + g'^2}}. \quad (1.15)$$

The term in the electroweak Lagrangian given in Equation 1.8 responsible for the interaction of quarks and leptons with the gauge bosons is  $\bar{\psi}i\gamma^\mu D_\mu\psi$  which can be rewritten as

$$eJ_{EM}^\mu A_\mu + \frac{g}{\sqrt{2}}(J_L^{+\mu}W_L^+ + J_L^{-\mu}W_L^-) + \frac{gg'}{e}J_Z^\mu Z_\mu \quad (1.16)$$

where

$$J_L^{\pm\mu} = \sqrt{2}\bar{\psi}\gamma^\mu T_L^\pm\psi \quad (1.17)$$

$$J_Z^\mu = \bar{\psi}\gamma^\mu(T_{3L} - \sin^2\theta_W Q)\psi \quad (1.18)$$

$$J_{EM}^\mu = \bar{\psi}\gamma^\mu Q\psi \quad (1.19)$$

are the charged and neutral current interaction terms describing the interaction of the gauge bosons with the fermion  $\psi$  fields.

Singlet state  $\psi_R$  do not survive operations by  $T$  and  $T_{3L}$ . Since neutrinos do not carry charge it can be seen that there are no right-handed neutrino state  $\nu_R$ .

Electroweak interactions involve both leptons and quarks and the relevant quantum numbers for the allowed fermion states are shown in Table 1.4.

In electroweak interactions the leptons numbers  $L_e$ ,  $L_\mu$  and  $L_\tau$  are always conserved. Electroweak interactions among the leptons only occur within a single family and there is no inter-generational mixing. This is not the case for quarks in charged current interactions. Apparently the quark mass eigenstates are not exactly the same as the electroweak eigenstates. The quark

Family			$T$	$T_3$	$Y$	$Q$
$\begin{pmatrix} \nu_e \\ e \end{pmatrix}_L$	$\begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix}_L$	$\begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix}_L$	$1/2$	$+1/2$	$-1$	$0$
$e_R$	$\mu_R$	$\tau_R$	$0$	$0$	$-2$	$-1$
$\begin{pmatrix} u \\ d \end{pmatrix}_L$	$\begin{pmatrix} c \\ s \end{pmatrix}_L$	$\begin{pmatrix} t \\ b \end{pmatrix}_L$	$1/2$	$+1/2$	$+1/3$	$+2/3$
$u_R$	$c_R$	$t_R$	$0$	$0$	$+4/3$	$+2/3$
$d_R$	$s_R$	$b_R$	$0$	$0$	$-2/3$	$-1/3$

Table 1.4: Weak isospin, hypercharge and electric charge for the quarks and leptons.

eigenstates of electroweak charged current interactions are given by

$$\begin{pmatrix} u \\ d' \end{pmatrix}, \begin{pmatrix} c \\ s' \end{pmatrix}, \begin{pmatrix} t \\ b' \end{pmatrix} \quad (1.20)$$

where the mixing is described by the unitary Cabibbo-Kobayashi-Maskawa (CKM) matrix

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \cdot \begin{pmatrix} d \\ s \\ b \end{pmatrix} \quad (1.21)$$

Here the different  $V_{ij}$  are constants to consider in calculation of Feynman vertexes connecting the quark  $i$  with the quark  $j$  and are fundamental in electroweak processes amplitude and cross section calculation. Although there are 9 elements in the CKM matrix, imposing unitarity, the CKM matrix is function of only 4 free parameters, which can be expressed as 3 angles and one CP (*charge-parity*) violating phase. This corresponds to the Cabibbo-Kobayashi-Maskawa form

$$V = \begin{pmatrix} c_1 & -s_1 c_3 & -s_1 s_3 \\ s_1 c_2 & c_1 c_2 c_3 - s_2 s_3 e^{i\delta} & c_1 c_2 s_3 + s_2 c_3 e^{i\delta} \\ s_1 s_2 & c_1 s_2 c_3 + c_2 s_3 e^{i\delta} & c_1 s_2 s_3 + c_2 c_3 e^{i\delta} \end{pmatrix} \quad (1.22)$$

where  $s$  and  $c$  refer to  $\sin$  and  $\cos$  and their subscript to the angle  $\theta_i$ .  $\delta$  is then the CP violating phase. Thus there are four parameters which are  $\theta_1$ ,  $\theta_2$ ,  $\theta_3$  and  $\delta$ . CP violation, though very small, is now well established with CPT (charge, parity and time operations) believed to be the preserved underlying symmetry.

In a similar way it is possible to parametrize the CKM matrix in terms of  $\lambda$ ,  $A$ ,  $\rho$  and  $\eta$  according to the Wolfenstein parametrization, that maintain the unitarity to  $\mathcal{O}(\lambda^4)$

$$V = \begin{pmatrix} 1 - \lambda^2/2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \lambda^2/2 & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4) \quad (1.23)$$

Before introducing the Higgs mechanism (described in Section 1.2) the unified theory of the electromagnetic and weak interactions describes physical reality rather well at high energies where  $q^2$  is much greater than the mass of the partons involved.

## 1.2 Electroweak Symmetry Breaking and the Higgs Boson

The Higgs mechanism is the mechanism by which particles in the SM acquire mass. In 1964 Peter W. Higgs published a very short paper on “Broken Symmetries and the Masses of Gauge Bosons” [8] describing a mechanism by which gauge bosons can acquire mass. The Higgs mechanism has been fully incorporated into the Standard Model and it is thought to be responsible for the gauge boson and fermion masses. This mechanism gives rise to particle masses, at the price of introducing a new spin-0 scalar particle, called the Higgs boson.

It should be noted that other theories exist which attempt to explain spontaneous symmetry breaking and particle masses but are not considered in this brief overview.

In order to give mass to the gauge boson, a scalar field  $\Phi$  and potential term  $V(\Phi) = \mu^2|\Phi|^2 + \lambda|\Phi|^4$  is introduced in the electroweak Lagrangian given in Equation 1.8 as

$$\mathcal{L}_\Phi = |D_\mu\Phi|^2 - \mu^2|\Phi|^2 - \lambda|\Phi|^4. \quad (1.24)$$

If  $\mu^2$  is positive then the potential  $V(\Phi)$  is symmetric about its minimum which is 0. However, in the case where  $\mu^2 < 0$  the potential has a minimum at

$$|\Phi| = \sqrt{\frac{-\mu^2}{2\lambda}} \quad (1.25)$$

as can be seen in Figure 1.2.

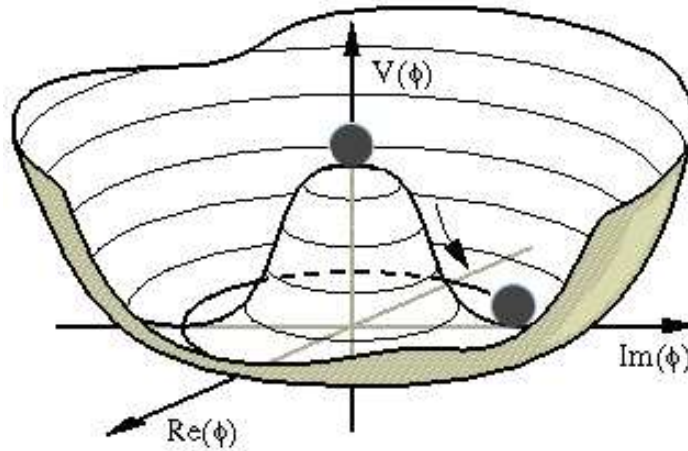


Figure 1.2: Higgs potential  $V(\phi)$  for  $\mu < 0$ .

The ground state is said to have spontaneously picked a direction, which has given rise to a non-zero vacuum expectation value and a broken symmetry. The complex doublet  $\Phi$  can be written in terms of the vacuum expectation value and two real fields with zero vacuum expectation value  $\xi$  and  $H$  as

$$\Phi(x) = \exp\left(\frac{i\xi(x) \cdot \tau}{2v}\right) \begin{pmatrix} 0 \\ (v + H(x))/\sqrt{2} \end{pmatrix} \quad (1.26)$$

where  $v = \sqrt{-\mu^2/\lambda}$ . Here,  $H$  will be the Higgs field and  $\xi(x)$  are non-physical fields known as Goldstone bosons. Since the SM theory must be invariant under local gauge transformations, with an appropriate gauge transformation is possible to cancel the dependence from the  $\xi$  fields. The

choice of the gauge is called the unitary gauge where

$$\Phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + H(x) \end{pmatrix} \quad (1.27)$$

and the Lagrangian  $\mathcal{L}_\Phi$  becomes

$$\mathcal{L}_\Phi = \frac{1}{2}(\partial H)^2 + \frac{1}{4}g^2 W^+ W^- (v + H)^2 + \frac{1}{8} \left( \frac{gg'}{e} \right)^2 Z Z (v + H)^2 - V \left( \frac{1}{2}(v + H)^2 \right). \quad (1.28)$$

What has happened here is that the Goldstone bosons  $\xi$  have vanished and as a result, the gauge bosons have acquired mass terms. For the  $W$  and  $Z$  bosons they can essentially be read off the Equation 1.28 and are given by

$$M_W = \frac{1}{2}g v, \quad M_Z = \frac{1}{2} \frac{gg'}{e} v. \quad (1.29)$$

The Higgs itself has an associated mass term coming from the potential term  $V$

$$M_H = \sqrt{-2\mu^2}. \quad (1.30)$$

Recall that  $g$  and  $g'$  are related by the electroweak mixing angle  $\theta_W$ . It follows that  $M_W$  and  $M_Z$  are related by

$$M_W = M_Z \cos \theta_W. \quad (1.31)$$

Experimentally one can measure both  $M_W$  and  $M_Z$  and given the relationship between  $g$  and  $g'$  show that  $v = 246$  GeV [9]. Then the only undetermined parameter is  $\mu$  which implies that the mass of the Higgs is undetermined.

Fermion masses are generated in the Higgs mechanism by what is called the Yukawa coupling. The Yukawa interaction term in the Lagrangian for a lepton is given by

$$\mathcal{L}_l = -G_l [\bar{l}_R (\Phi^\dagger l_L) + (\bar{l}_L \Phi) l_R] \quad (1.32)$$

for the singlet  $l_R$  and doublet  $l_L$  where  $G_l$  is a coupling constant. In the unitary gauge this becomes

$$\mathcal{L}_l = -\frac{1}{\sqrt{2}} G_l v \bar{l} l - \frac{1}{\sqrt{2}} G_l H \bar{l} l \quad (1.33)$$

from which the lepton mass can be read off as

$$M_l = \frac{1}{\sqrt{2}} G_l v.$$

The direct coupling of the leptons to the Higgs is evident in the  $H \bar{l} l$  term in Equation 1.33. Although the Higgs existence and its mass has only recently been seen (details in Section 2.3), its couplings to all particles were already well defined by the SM and depend on the particle masses, which have been measured experimentally. Similarly for the quarks, a Yukawa coupling can be added of the form

$$\mathcal{L}_q = - \sum_{i=1}^3 \sum_{j=1}^3 [\tilde{G}_{ij} \bar{u}_{iR} (\Phi^\dagger D_{jL}) + G_{ij} \bar{d}_{iR} (\Phi^\dagger D_{jL})] + \text{h.c.} \quad (1.34)$$

where  $u_i$  and  $d_i$  refer to the up and down-type quarks. Here  $G$  is related to the up-type quark

mass matrix and  $\tilde{G}$  to the down-type quark mass matrix by

$$M_{ij}^u = \frac{v}{2} \tilde{G}_{ij}, \quad M_{ij}^d = \frac{v}{2} G_{ij}.$$

In this spontaneous symmetry breaking model the mass of the Higgs boson is a free parameter that has to be determined experimentally. Some indirect constraints on this observable can be inferred from the precise measurement of SM parameters, that will be discussed in Section 2.3, as well as direct constraints come from Higgs boson searches at colliders.



# Diboson Production and Higgs Phenomenology

## Contents

2.1	Massive Boson Mediators: $Z$ and $W^\pm$ . . . . .	<b>11</b>
2.1.1	$Z$ properties from $e^+e^-$ colliders . . . . .	12
2.2	Physics Measurements of Diboson Final States . . . . .	<b>14</b>
2.2.1	LEP II . . . . .	15
2.2.2	Tevatron . . . . .	16
2.2.3	LHC . . . . .	18
2.3	Higgs phenomenology . . . . .	<b>18</b>
2.3.1	Theoretical constraints . . . . .	19
2.3.2	Higgs boson production and decay . . . . .	20
2.3.3	Experimental Situation . . . . .	22

## 2.1 Massive Boson Mediators: $Z$ and $W^\pm$

The interactions mediated by the  $Z$  boson were not noticed until about half a century ago, when observation of non-leptonic decay modes of strange mesons first hinted the need to complement the earlier  $V - A$  theory with neutral intermediate boson. By 1967, however, it was a key unobserved piece of the near-modern electroweak theory. Thereafter, the Gargamelle experiment (1973) saw the evidence for it in neutrino-nucleon collision where neutral current collisions accompanied the charged current  $\nu + \text{nucleon} \rightarrow l^- + \text{hadrons}$  events predicted by the  $V - A$  theory. With its direct observation at UA1 and UA2 experiment at CERN in 1983 as a few resonant events in the dielectron and dimuon mass spectrum [10], detailed studies of the  $Z$ 's properties began. Nowadays the  $Z$  boson properties are known to a precision below the per-mil level [3], as shown by the measured branching ratios reported in Tables 2.1.

In the same year (1983) the UA2 experiment observed the existence of the  $W$  boson [11], that completed the set of massive bosons of the unified electroweak theory. The properties of the  $W$  boson have been investigated in the following years, with particular attention to the value of its mass  $m_W$ , which is a crucial parameter of the SM <sup>1</sup>. Given the subject of the analyses presented

<sup>1</sup> The mass of the  $W$  boson is strongly correlated with the Higgs boson mass, as will be discussed in Section 2.3.1. A precise measurement of the former [12] will give indirect constraints on the latter.

Z Decay modes	Fraction $\Gamma_i/\Gamma$
$e^+e^-$	$(3.363 \pm 0.004) \%$
$\mu^+\mu^-$	$(3.366 \pm 0.007) \%$
$\tau^+\tau^-$	$(3.370 \pm 0.008) \%$
invisible	$(20.00 \pm 0.06) \%$
hadrons	$(69.91 \pm 0.06) \%$
$(u\bar{u} + c\bar{c})/2$	$(11.6 \pm 0.6) \%$
$(d\bar{d} + s\bar{s} + b\bar{b})/3$	$(15.6 \pm 0.4) \%$
$c\bar{c}$	$(12.03 \pm 0.21) \%$
$b\bar{b}$	$(15.12 \pm 0.05) \%$
$b\bar{b}b\bar{b}$	$(0.036 \pm 0.013) \%$

Table 2.1: Main  $Z$  decay modes.

in this Thesis we will not focus on the details of the  $W$  boson but more information can be found in [3].

### 2.1.1 $Z$ properties from $e^+e^-$ colliders

In the years after the discovery of  $Z$  boson, different experiments worked a lot to study its properties, especially its mass, width and branching ratios. Here are described the main results obtained with the study of  $Z$  properties at the two main  $e^+e^-$  collider that investigated it, LEP at CERN and SLC at SLAC.

The properties of the  $Z$  boson and the underlying electroweak theory were verified at LEP and SLC with several measurements, among these the overall  $e^+e^-$  production cross-section, the  $Z$  line-shape and width, the number of neutrino families, and the forward-backward asymmetries of the leptons and quarks. The experimental analyses of the  $Z$  line-shape (see Figure 2.1) of the decay branching ratios and the asymmetries were performed with very high precision. The fit of the LEP and SLC data gave as results

$$M_Z = 91187.5 \pm 2.1 \text{ MeV} \quad (2.1)$$

$$\Gamma_Z = 2495.2 \pm 2.3 \text{ MeV} \quad (2.2)$$

$$\sin^2 \theta_{\text{eff}} = 0.23147 \pm 0.00016 . \quad (2.3)$$

where  $\theta_{\text{eff}}$  is obtained from the electroweak mixing angle  $\theta_W$  including the effect of radiative corrections.

Thus, the electroweak sector of the Standard Model successfully passed the examination at the per-mille level, as highlighted by global analysis of the electroweak mixing parameter  $\sin^2 \theta_W$ . Figure 2.2 shows the observables that were precisely measured at LEP and SLC, found in a good agreement with the Standard Model expectations.

However, beyond this most stringent test of the electroweak theory itself,  $Z$  physics at  $e^+e^-$  colliders allowed important conclusions to be drawn on several other aspects of the Standard Model and potential physics beyond it. The first of these concerned the three families of leptons in the Standard Model. The width  $\Gamma$  of the  $Z \rightarrow \ell\ell$  depends on the number of lepton families. Comparing  $\Gamma_Z$  and  $\Gamma_{Z \rightarrow \ell\ell}$  could be determined the number of light neutrinos<sup>2</sup>. The ensuing

<sup>2</sup> Light neutrinos are those with  $m_\nu < m_Z/2$ , hence kinematically accessible for the  $Z$  decay.



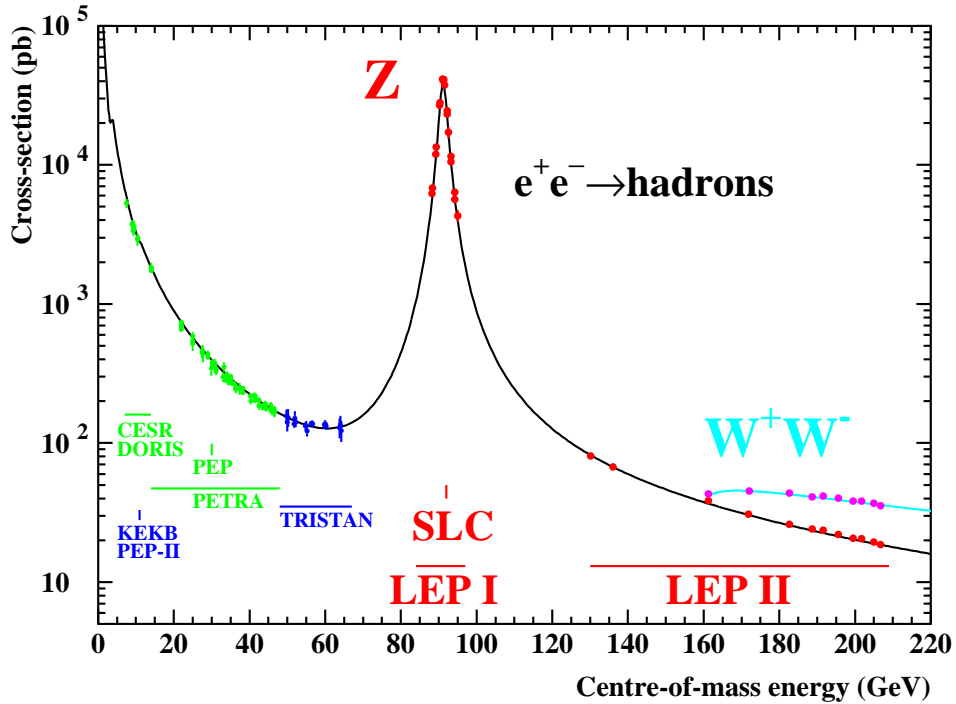


Figure 2.1: The  $e^+e^-$  annihilation cross-section to hadrons, from initial low energies in early colliders to the maximum energy at LEP.

difference determines the number of light neutrino species to be:

$$N_\nu = 2.985 \pm 0.008 . \quad (2.4)$$

Thus, LEP put the lid on the Standard Model with three families of matter particles.

In the same period LEP started to study  $Z$  properties, the SLAC (Stanford Linear Accelerator Collider) improved the pre-existing accelerator to produce 50 GeV  $e^\pm$  beams for  $e^+e^-$  collisions. The Stanford Large Detector (SLD) was designed and built to study the  $Z$  properties especially in the heavy flavor decay channel. The main feature and difference of SLAC was the use of longitudinal polarized  $e$  beam. This permitted to study with high precision the angular asymmetries of polarized  $Z$  produced. SLD measured the parity-violating parameter  $A_b$  by analyzing the left-right (back-forward) asymmetry of  $b$  quarks in  $e^+e^- \rightarrow Z \rightarrow b\bar{b}$  with different analysis techniques. Similarly was studied the asymmetry parameter  $A_c$  from  $Z \rightarrow c\bar{c}$  decay. From data collected from 1993 to 1995, in a sample of  $\sim 150000$   $Z$  the asymmetry measurement gave

$$A_b = 0.910 \pm 0.068(\text{stat.}) \pm 0.037(\text{syst.}) \quad (2.5)$$

$$A_c = 0.642 \pm 0.110(\text{stat.}) \pm 0.063(\text{syst.}) . \quad (2.6)$$

Combined with the results found at LEP those two experiments gave a precise measurement of the most important  $Z$  properties [13]. These measurements were found to be in a really nice agreement with the SM prediction, testing not only the tree-level theory, but also its higher order corrections.

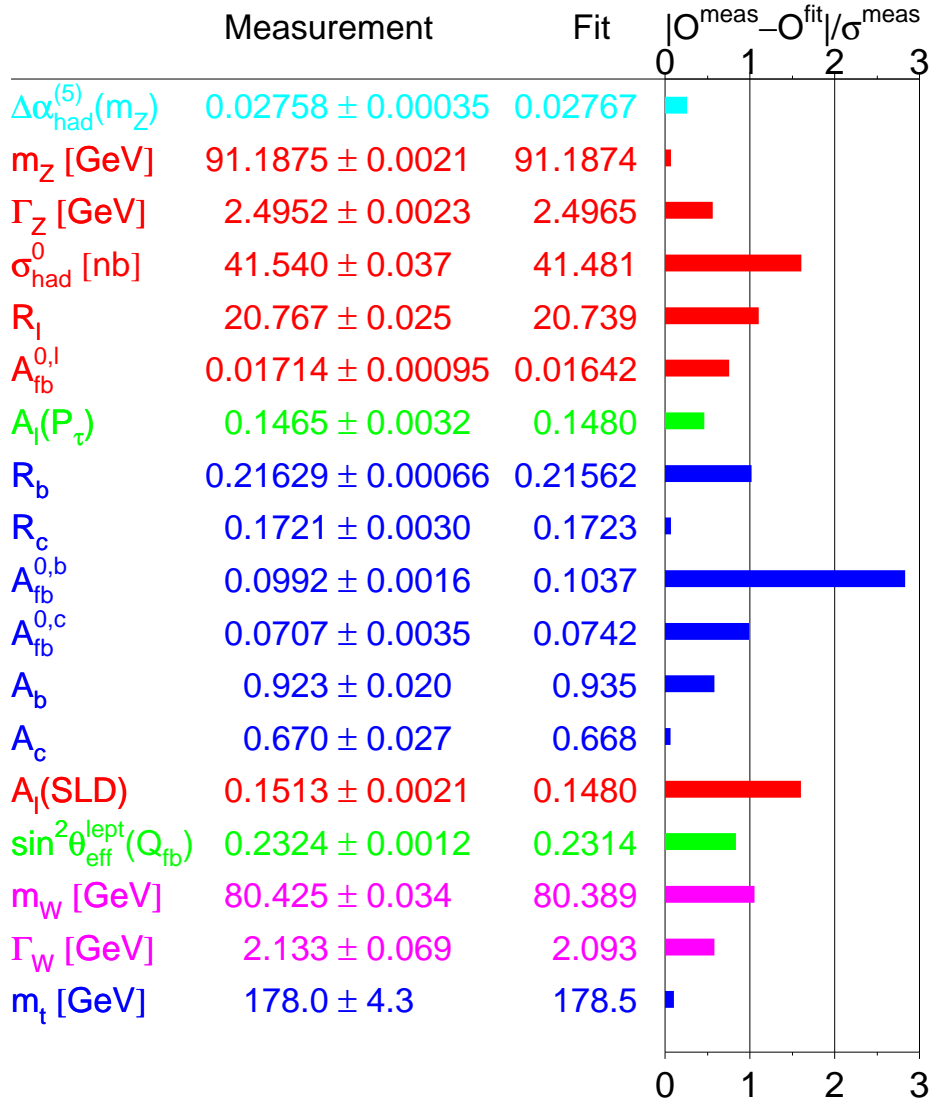


Figure 2.2: Precision observables of the electroweak part of the Standard Model, measured at LEP (when not specified) and SLC (by the SLD experiment).

## 2.2 Physics Measurements of Diboson Final States

After the discovery of the  $Z$  and  $W$  boson, with all their features, many physical studies have been done to measure the cross-sections of processes involving the production of two of such massive vector bosons. These kind of processes are less frequent than those involving just one

vector boson and needed more years and different experimental conditions to be studied. The study of dibosons production is very important to test the Standard Model predictions at a second order precision, and search for any hint of deviations from these as effects of new physics.

The diboson production has been studied both in electron-positron colliders (LEP II,  $e^+e^-$  collisions at  $\sqrt{s} = 183 - 207$  GeV) and in hadronic collider (Tevatron,  $p\bar{p}$  collisions at  $\sqrt{s} = 1.96$  TeV and LHC,  $pp$  collisions at  $\sqrt{s} = 7(8)$  TeV). In the two different environments dibosons are produced through different processes that will be discussed in the following two sections. First measurements of diboson cross-sections have been done at LEP collider; since LEP was an  $e^+e^-$  collider, the only boson pairs that could have been produced were  $ZZ$  and  $W^+W^-$  while at Tevatron and LHC, several measurements have been done in  $p\bar{p}$  collisions of  $ZZ$ ,  $W^+W^-$  and  $ZW^\pm$  production cross section.

In the following we propose a short review of the fundamental analysis done during the past years at different colliders (leptonic and hadronic) involving diboson production cross sections.

### 2.2.1 LEP II

At LEP II collider the  $ZZ$  diboson events were produced by  $e^+e^-$  collisions by mean of the processes illustrated by Feynman diagrams in Figure 2.3

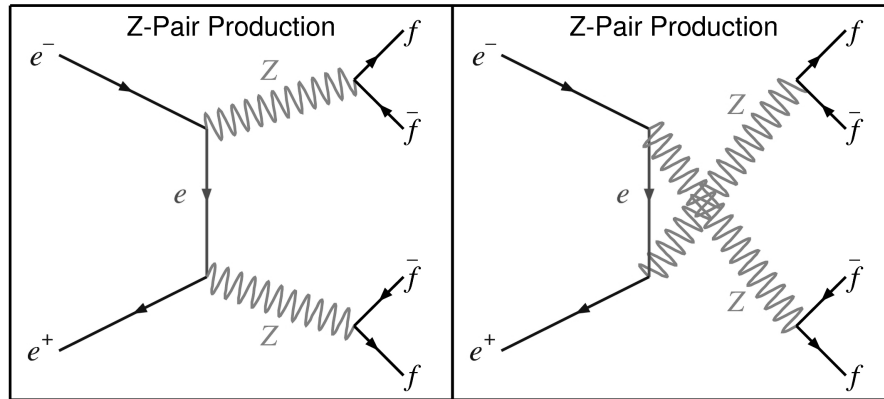


Figure 2.3: Tree level Feynman diagrams for  $ZZ$  production at the LEP collider.

The first evidence of the  $ZZ$  production was studied during 1997 by the DELPHI experiment [14], when LEP was operating at a center of mass energy of  $\sqrt{s} = 182.6$  GeV, corresponding to the threshold for this channel. In the following years the analysis was extended with the growth of the center-of-mass energy, from  $\sqrt{s} = 182.6$  GeV to  $\sqrt{s} \simeq 207$  GeV, collecting totally  $665.3 \text{ pb}^{-1}$  of data. The choice of the DELPHI experiment was to select fermion pairs in a narrow range around the  $Z$  mass value, from 81 GeV to 101 GeV, for both fermion pairs reconstructed and then to scale the calculated cross section with the ratio of the theoretical total cross section and the one in the four fermion chosen window.

All the results obtained in the different final states studied,  $q\bar{q}q\bar{q}$ ,  $\ell^+\ell^-q\bar{q}$ ,  $\nu\bar{\nu}q\bar{q}$ ,  $\nu\bar{\nu}\ell^+\ell^-$  and  $\ell^+\ell^-\ell'^+\ell'^-$  have been combined to obtain the final cross sections at the different energies, reported in Table 2.2 and shown in Figure 2.4: all the results are in really good agreement with the SM prediction.

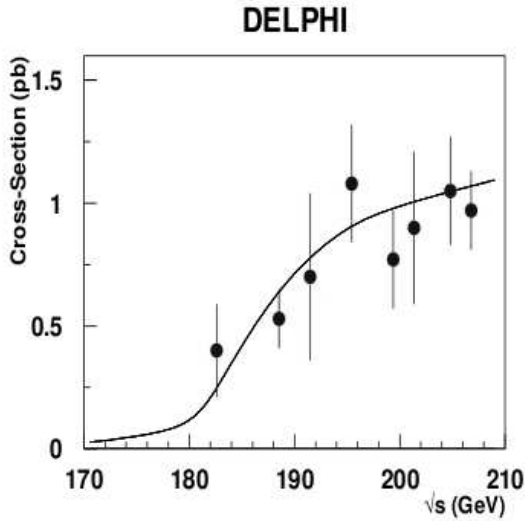


Figure 2.4: Combined cross-sections measured from data collected in 1997-2000. The errors shown are sums in quadrature of the statistical and systematic errors.

$\sqrt{s}$ [GeV]	$\sigma(ZZ)$ [pb]	SM exp. [pb]
183	$0.40^{+0.21}_{-0.16} \pm 0.02$	0.25
189	$0.53^{+0.12}_{-0.11} \pm 0.02$	0.65
192	$0.70^{+0.37}_{-0.31} \pm 0.02$	0.78
196	$1.08^{+0.25}_{-0.22} \pm 0.02$	0.90
200	$0.77^{+0.21}_{-0.18} \pm 0.02$	0.99
202	$0.90^{+0.33}_{-0.29} \pm 0.02$	1.00
205	$1.05^{+0.23}_{-0.20} \pm 0.02$	1.05
207	$0.97^{+0.16}_{-0.15} \pm 0.02$	1.07

Table 2.2: Cross section obtained by DELPHI measurements. The first errors are statistical while the seconds are systematic.

Similar analyses have been done by the other LEP experiments, ALEPH, L3 and OPAL but we will not discuss them here.

For a center-of-mass energy above  $\sim 2 \cdot m_W$  at LEP, it was possible to produce pairs of  $W$  boson. The  $W^+W^-$  production cross section has been measured by the four LEP experiments in the energy range  $\sqrt{s} = 183 \div 207$  GeV and compared with the theoretical predictions [15, 16], which are shown in Figure 2.5(a). The combination of the results lead to a measurement with an uncertainty of 1%, in agreement with the predictions, as shown in Figure 2.5(b). Further details about the  $WW$  cross section measurement at LEP can be found in [17].

## 2.2.2 Tevatron

At the Tevatron, the production of a pair of massive boson is dominated by  $q\bar{q}$  annihilation ( $s$ -channel) and scattering ( $t$ -channel), as shown in Figure 2.6. At a hadron collider it is favorable to search for these processes in their leptonic or semi-leptonic decay modes. Even if the fully hadronic final states have larger branching ratios [3], the small diboson signals is overwhelmed by the background from hadronic activity produced by higher-rate QCD and single boson production.

The diboson production at Tevatron has been investigated by the CDF and D0 experiment, two multi-purpose experiment analysing the products of the high energy  $p\bar{p}$  collisions. The  $WW$  production cross section is expected to be  $12.4 \pm 0.8$  pb at Next-to-Leading Order (NLO) [18]. The first observation of that has been reported by D0, that found 25 candidate events of which  $8.1 \pm 1.0$  were expected to be background events in approximately  $250 \text{ pb}^{-1}$  of data; the measurement extended to  $1 \text{ fb}^{-1}$  reported a cross section  $\sigma(WW) = 13.8^{+4.3}_{-3.8}(\text{stat.})^{+1.2}_{-0.9}(\text{syst.}) \pm 0.9(\text{lum.})$  pb, with  $5.2 \sigma$  of significance [19], consistent with the SM prediction. CDF made a more precise analysis, on  $3.6 \text{ fb}^{-1}$  of data, based on a likelihood ratio method, that measured a cross section  $\sigma(WW) = 12.1 \pm 0.9(\text{stat.})^{+1.6}_{-1.4}(\text{syst.})$  pb [20], again consistent with SM predictions.

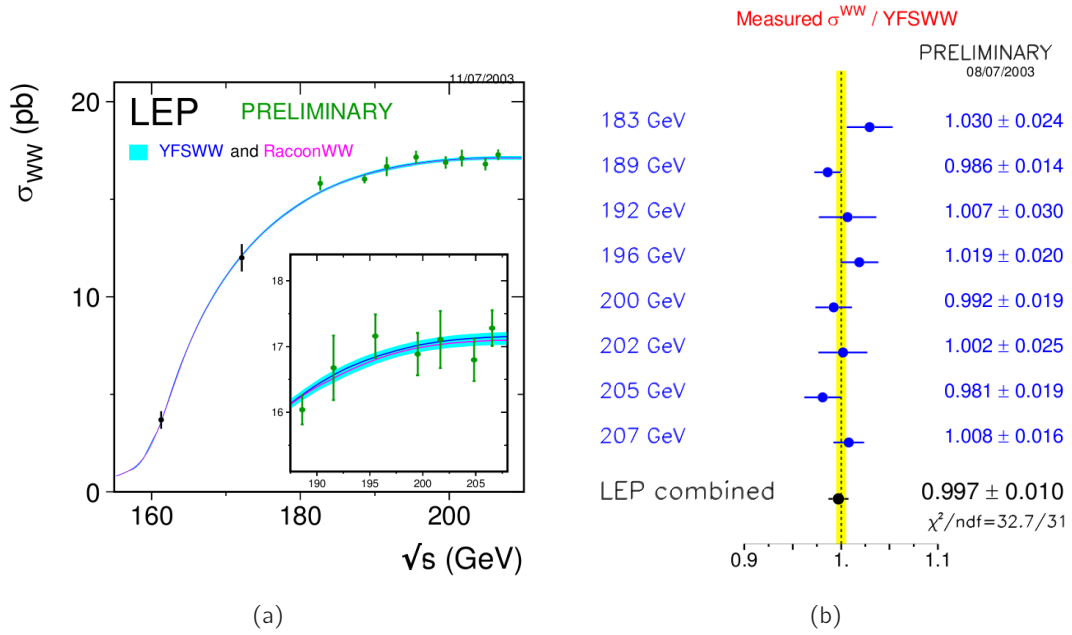


Figure 2.5: (a)  $W$  pair production production cross section as a function of the center-of-mass energy  $s$  at LEP. (b) Ratio of the measured cross section to the prediction.

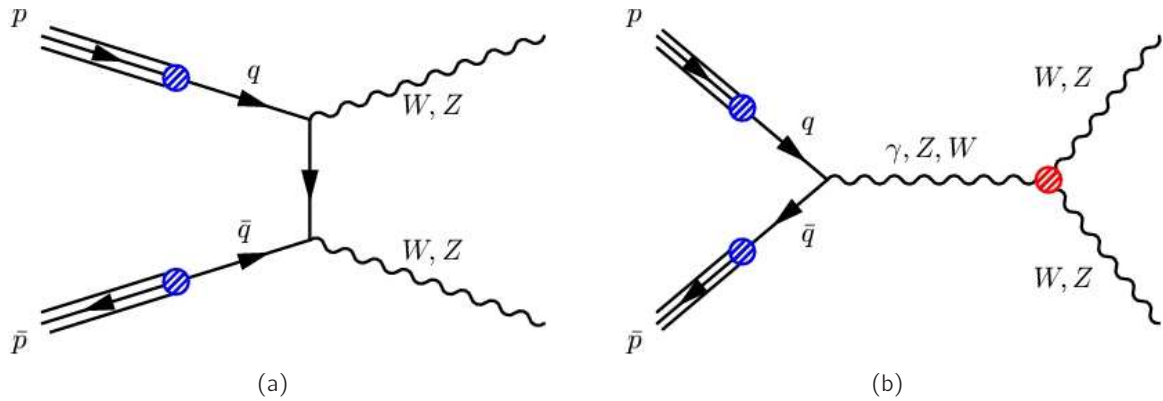


Figure 2.6: Tree-level  $t$ -channel(a) and  $s$ -channel(b) Feynman diagram for  $ZZ$  diboson production at the Tevatron.

**WZ** production at Tevatron, expected to have a cross section  $\sigma(WZ) = 3.47 \pm 0.21$  pb at NLO, was first observed by CDF with a data sample of approximately  $1.1 \text{ fb}^{-1}$  in the  $WZ \rightarrow \ell\nu\ell'\ell'$  channel ( $\ell=e, \mu$ ) [21]. More accurate measurements was carried out later by CDF, considering  $7.1 \text{ fb}^{-1}$ , which measured a cross section  $\sigma(WZ) = 3.93^{+0.60}_{-0.53}(\text{stat.})^{+0.59}_{-0.46}(\text{syst.})$  pb [22], and by D0, considering  $8.6 \text{ fb}^{-1}$ , that obtained as result  $\sigma(WZ) = 4.5^{+0.6}_{-0.7}$  pb [23].

CDF also reported the evidence of  $WW$  and  $WZ$  production considering their semi-leptonic decay mode  $WW/WZ \rightarrow \ell\nu jj^3$ . The branching ratio for this process is higher than the leptonic decay channel but this final state faces difficulties in the separation of the diboson signal from the large background due to  $W + jets$  events. The first analysis in this particular channel, using  $2.7 \text{ fb}^{-1}$  of data give a cross section of  $\sigma(WW/WZ) = 17.7 \pm 3.9$  pb with a significance of 5.4

<sup>3</sup> It is not possible with current jets resolution (see Section 4.5) to distinguish between a  $W$  jet pair and a  $Z$  jet pair so the analysis has been done combining the two processes.

$\sigma$ . Both CDF and D0 carried out analyses on the  $\ell\nu jj$  final state, reported in [24, 25, 26], using approximately half of the full CDF available dataset ( $4.3 - 4.6 \text{ fb}^{-1}$ ).

**ZZ** has been studied by CDF in a small sample of data, corresponding to  $1.9 \text{ fb}^{-1}$ , considering the  $\ell\ell'\ell'$  and  $\ell\ell\nu\nu$  decay mode [27]. D0 performed similar analyses considering the  $\ell\ell'\ell'$  decay mode in a sample of  $1.7 \text{ fb}^{-1}$  and the  $\ell\ell\nu\nu$  decay mode using approximately  $2.7 \text{ fb}^{-1}$  [28]. Each experiment combined the measurement in the two final states. The result for the CDF analysis is a cross section of  $\sigma(ZZ) = 1.4^{+0.7}_{-0.6}(\text{stat.} + \text{syst.}) \text{ pb}$  [27] while D0 measured a cross section of  $\sigma(ZZ) = 1.60 \pm 0.63(\text{stat})^{+0.16}_{-0.17}(\text{syst}) \text{ pb}$  [29] with a total significance of  $5.7 \sigma$ ; both measurements are consistent with the SM prediction. The cross section measurements presented in this Thesis exploit the full CDF available data sample and has been published (in a reduced version, exploiting  $6 \text{ fb}^{-1}$  of collected data) [30].

All the aforementioned results, as well as other relevant diboson measurement performed at the Tevatron, are reported in Table 2.3.

### 2.2.3 LHC

The measurement of diboson production has been studied by the CMS and Atlas experiments at the CERN LHC  $pp$  collider, using data collected during 2010-'11 at  $\sqrt{s} = 7 \text{ TeV}$  and during 2012 at  $\sqrt{s} = 8 \text{ TeV}$ .

The Atlas experiment measured the  $WW$  [31],  $WZ$  [32], and  $ZZ$  [33] production cross section in their leptonic decays at  $\sqrt{s} = 7 \text{ TeV}$  and the  $ZZ$  production cross section in the  $\ell\ell'\ell'$  final state at  $\sqrt{s} = 8 \text{ TeV}$  [34].

Similarly CMS measured the  $WW$  [35],  $WZ$  [36], and  $ZZ$  [37] at  $\sqrt{s} = 7 \text{ TeV}$ , and  $WW \rightarrow \ell\nu\ell'\nu'$  [38] and  $ZZ \rightarrow \ell\ell'\ell'$  [39] at  $\sqrt{s} = 8 \text{ TeV}$ .

Using  $5.0 \text{ fb}^{-1}$  of the data collected at  $\sqrt{s} = 7 \text{ TeV}$  the CMS experiment reported a preliminary measurement of  $WW/WZ$  production in the semileptonic decay mode [40] which is in agreement with the SM NLO expectation.

All the most relevant measurement of diboson production cross section carried out at the LHC are reported in Table 2.3.

## 2.3 Higgs phenomenology

According to the Higgs mechanism described in Section 1.2, a Higgs boson particle can be produced in high energy collisions, and several efforts have been done to detect its signature both at lepton and hadron colliders. The production of this unstable particle can be observed through the detection of the products of its decay. The Higgs boson favorable decays are those to a pair of massive particles, given its Yukawa couplings outlined in Equation 1.33, when kinematically accessible depending on the Higgs boson mass  $m_H$ . This is an unknown parameter of the SM spontaneous symmetry breaking mechanism, which determines afterward the Higgs boson couplings, branching ratios, production mechanisms and rates<sup>4</sup>. Constraints on the Higgs boson mass exist from theory, phenomenology and experimental measurements, which we briefly report in the following.

<sup>4</sup> This properties are determined once measured the Higgs boson mass if we assume the SM Higgs mechanism described in Section 1.2. Different theoretical models can give rise to a similar scalar boson, with different features. All the Higgs properties must be tested experimentally to understand precisely the theoretical model responsible for its existence.

Process	Experiment	$\mathcal{L}$ (fb $^{-1}$ )	Measured $\sigma$ (pb)	Theory $\sigma$ (pb)
$WW \rightarrow \ell\nu\nu$	CDF[20]	3.6	$12.1 \pm 0.9^{+1.6}_{-1.4}$	$11.3 \pm 0.7$
	D0[19]	1.0	$11.5 \pm 2.2$	$11.3 \pm 0.7$
	CMS[35]	4.9	$52.4 \pm 2.0 \pm 4.7$	$47.0 \pm 2.0$
	Atlas[31]	4.6	$51.9 \pm 2.0 \pm 4.4$	$44.7^{+2.1}_{-1.9}$
	CMS†[38]	3.5	$69.9 \pm 2.8 \pm 6.4$	$53.7^{+2.4}_{-1.6}$
$ZZ \rightarrow \ell\ell + (\ell\ell/\nu\nu)$	CDF[30]	6.1	$1.64^{+0.44}_{-0.38}$	$1.4 \pm 0.1$
	D0[23]	8.6	$1.44^{+0.35}_{-0.34}$	$1.4 \pm 0.1$
	Atlas[33]	4.6	$6.7 \pm 0.7^{+0.5}_{-0.4}$	$5.89^{+0.22}_{-0.18}$
$ZZ \rightarrow \ell\ell\ell'\ell'$	CMS[37]	5.0	$6.24^{+0.86+0.43}_{-0.80-0.35}$	$6.5^{+0.3}_{-0.2}$
	CMS†[39]	5.3	$8.4 \pm 1.0 \pm 0.8$	$7.7 \pm 0.4$
	Atlas†[34]	5.8	$9.3^{+1.1+0.5}_{-1.0-0.4}$	$7.4 \pm 0.4$
$WZ \rightarrow \ell\ell\nu$	CDF[22]	7.1	$3.9 \pm 0.8$	$3.47 \pm 0.21$
	D0[23]	8.6	$4.5^{+0.6}_{-0.7}$	$3.47 \pm 0.21$
	CMS[36]	1.1	$17.0 \pm 2.4 \pm 1.5$	$17.3^{+1.3}_{-0.8}$
	Atlas[32]	4.6	$19.0^{+1.4}_{-1.3} \pm 1.0$	$17.6^{+1.1}_{-1.0}$
$WW/WZ \rightarrow \ell\nu + \text{jets}$	CDF[24]	4.3	$18.1 \pm 3.3 \pm 2.5$	$16.8 \pm 0.5$
	CDF[25]	4.6	$16.5^{+3.3}_{-3.0}$	$16.8 \pm 0.5$
	D0[26]	4.3	$19.6^{+3.2}_{-3.0}$	$16.8 \pm 0.5$
	CMS[40]	5.0	$68.9 \pm 8.7 \pm 9.8$	$65.6 \pm 2.2$
$WZ/ZZ \rightarrow \ell\ell/\ell\nu/\nu\nu + \text{H.F.}$	CDF[41]	9.45	$4.08^{+1.38}_{-1.26}$	$4.4 \pm 0.2$
	D0[42]	8.4	$5.0 \pm 1.0^{+1.3}_{-1.2}$	$4.4 \pm 0.2$
	D0, CDF[43]	7.5 - 9.5	$4.47 \pm 0.67^{+0.73}_{-0.72}$	$4.4 \pm 0.3$
$VV' \rightarrow \nu\nu + \text{jets}$	CDF[44]	3.5	$18.0 \pm 2.8 \pm 2.6$	$16.8 \pm 0.5$

Table 2.3: Summary of the recent measurements of the diboson production cross section in leptonic and semi-leptonic final states. The reference to the individual measurements are reported in the table as well as the used integrated luminosity; the statistical uncertainty appears before the systematic uncertainties (when both are available); the theoretical predictions are calculated at NLO [18, 45]. The results marked with † are obtained at the LHC working at  $\sqrt{s} = 8$  TeV; for all the other results the measurement from CDF and D0 are assumed to be obtained at  $\sqrt{s} = 1.96$  TeV and from CMS and Atlas at  $\sqrt{s} = 7$  TeV.

### 2.3.1 Theoretical constraints

At high energy, the elastic scattering amplitude of massive weakly interacting bosons  $W$ ,  $WW \rightarrow WW$ , increases indefinitely with the center-of-mass energy for longitudinally polarized particles, due to the linear dependence of the longitudinal wave function on particle energy [46]. This constrains the Higgs boson mass to  $m_H < \sqrt{8\pi\sqrt{2}}/(3G_F) \simeq 1$  TeV.

An additional upper limit on  $m_H$  is provided by an analysis on the behavior of the Higgs potential [47]. Large values of  $m_H$  would lead to an infinite value of the parameter  $\lambda$  in Equation 1.24 at some energy scale, with implications in lattice calculations [48]. Reference [47] sets  $m_H \lesssim 180 \pm 6$  GeV/ $c^2$ .

A lower bound on the Higgs mass can instead be obtained requiring the minimum of the Higgs potential to be an absolute minimum, since possible instabilities are generated by the quantum

loop corrections (also referred to as stability bound) [49] suggesting  $m_H \gtrsim 135 \text{ GeV}/c^2$ . These theoretical indirect constraints, based on a-priori assumptions on the model relevant energy scales, are not used to set tight constraint on the Higgs boson mass, but as an indication of the preferable kinematic range.

### 2.3.2 Higgs boson production and decay

The Higgs boson couplings determine the dominant production processes at hadron and  $e^+e^-$  colliders, for which the Leading Order Feynman diagrams are shown in Figure 2.7.

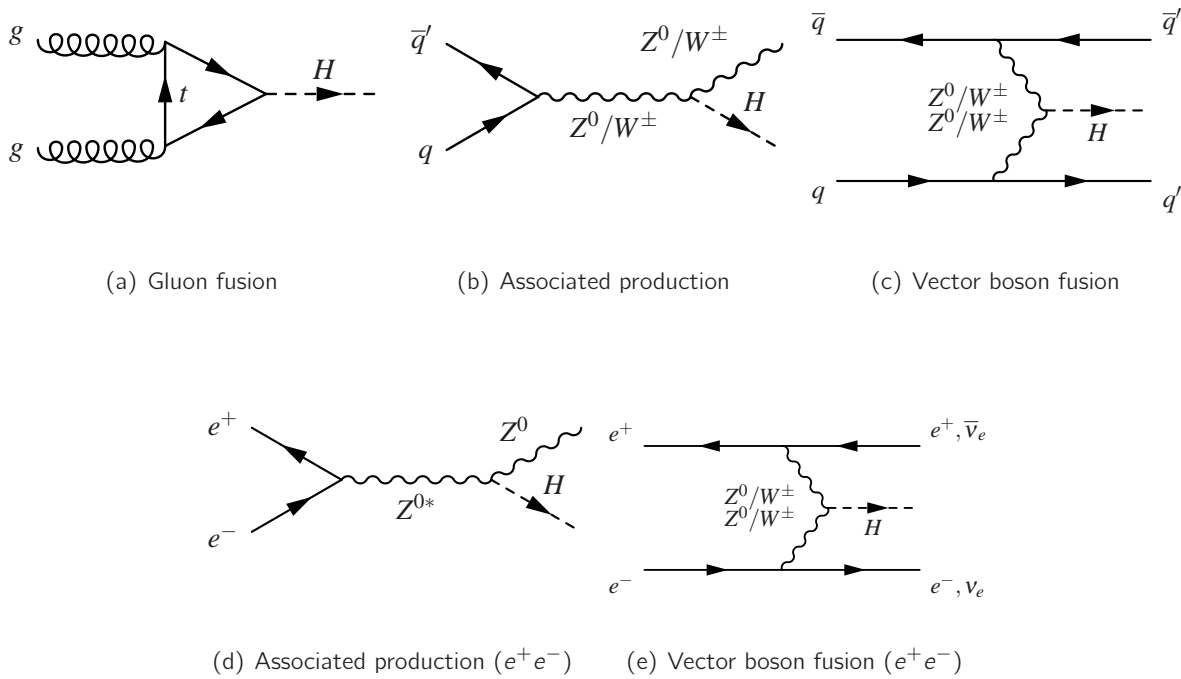


Figure 2.7: Leading order Feynman diagrams of the dominant Higgs production mechanisms at colliders: (a) gluon fusion, (b) associated production, and (c) vector boson fusion at hadron colliders, and (d) associated production and (e) vector boson fusion at  $e^+e^-$  colliders.

The dominant production mechanism is the **direct production** ( $gg \rightarrow H$ ) via gluon fusion and the top quark loop. The latest available calculations [50, 51, 52], implement a full QCD Next-to-Next-to-Leading Order (NNLO) approximation, with a total uncertainty of the order of 10% [50].

An alternative Higgs boson production mechanism is the **associated production**  $q\bar{q}(e^+e^-) \rightarrow Z^0 H$  and  $q\bar{q}' \rightarrow W^\pm H$ , where the Higgs boson is radiated by a  $W$  or  $Z$  boson. These are electroweak processes which receive small ( $\sim 1\%$ ) contributions by radiative corrections. Calculation at Next-to-Next-to-Leading Order (NNLO) [53] are available with an accuracy of less than 5%.

In addition to those, the **vector boson fusion** takes place when each of the incoming quarks interacts through charged or neutral current processes and the two  $W$ s or  $Z$ s bosons couple to the Higgs:  $q\bar{q} \rightarrow Hq\bar{q}$ . This is a pure electroweak process that is known at NLO with an estimate accuracy of  $\sim 10\%$  [53].

Figure 2.8(a) shows the expected production cross section at the Tevatron with the four processes shown separately. The gluon fusion contributes for 78% to the inclusive Higgs production



cross section while the associated production and vector boson fusion contribute with  $\sim 15\%$  and  $\sim 7\%$  respectively.

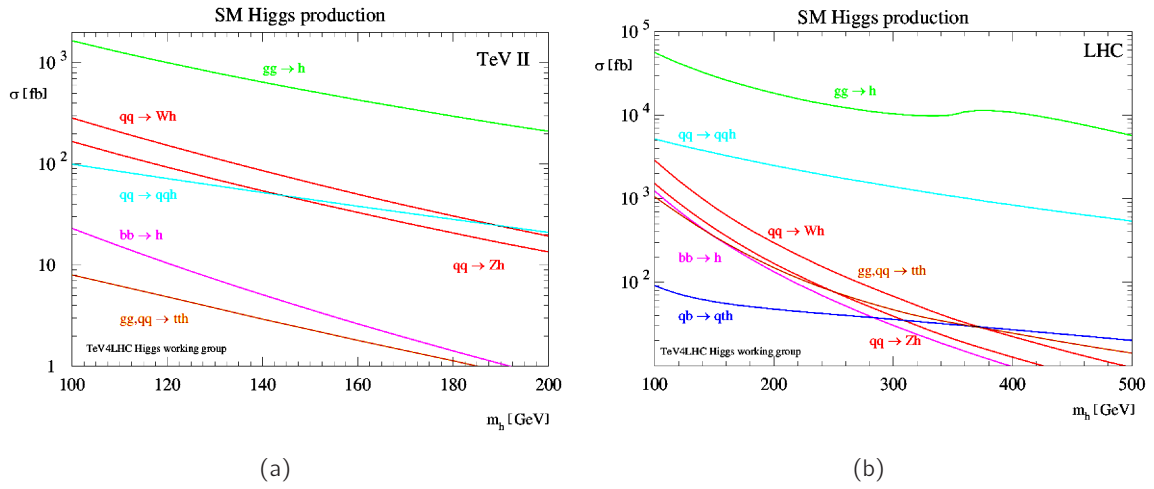


Figure 2.8: Production cross sections at the (a)  $p\bar{p}$  collider Tevatron ( $\sqrt{s} = 1.96$  TeV) and at the (b)  $pp$  collider LHC ( $\sqrt{s} = 7$  TeV), as function of the Higgs boson mass.

At the Large Hadron Collider (LHC) the Higgs production cross sections are about two order of magnitude larger as shown in Figure 2.8(b), with the expectations for the four production mechanisms considered.

The strong coupling of the Higgs boson to heavy Standard Model particles implies a preferential decay to top quarks and vector bosons, if kinematically allowed, and to  $b$ -quarks. The Higgs boson Branching fractions, calculated at Next-to-Leading order, are shown in Figure 2.9 as functions of the Higgs boson mass.

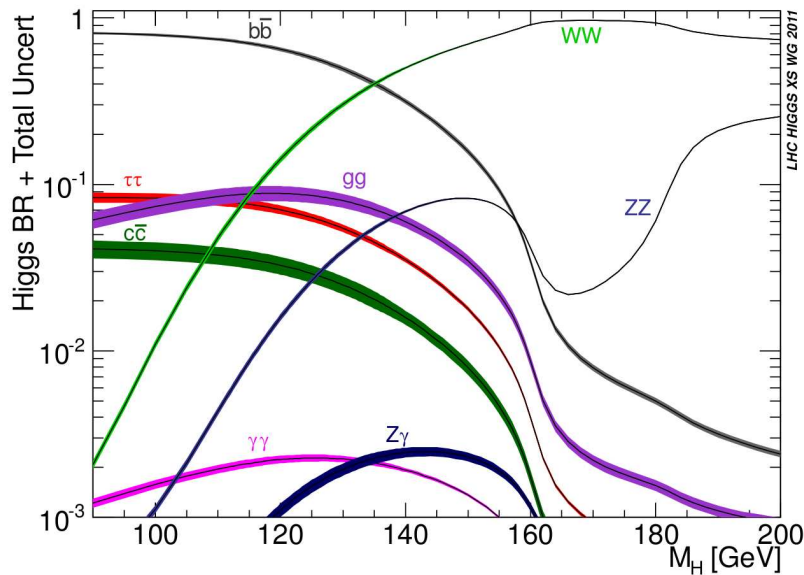


Figure 2.9: Higgs boson decay branching ratios [54].

For  $m_H < 2 m_W$  the decay width to two vector bosons is suppressed for the presence of at least one off-shell  $W$  or  $Z$ . Since the decay to a top pair is not kinematically allowed the main

decay channel is into  $b\bar{b}$  pair. For  $2 m_W < m_H < 2 m_Z$  the decay to a pair of  $W$  bosons is available while the one to a pair of on-shell  $Z$  boson is not; in this range dominates the  $H \rightarrow WW$  decay, resulting in a large suppression of the  $H \rightarrow ZZ$  decay branching fraction. In the mass range above  $m_H > 2 m_Z$  both the decay channels to  $WW$  and  $ZZ$  produce on-shell vector bosons and  $ZZ$  decay branching ratio rise to  $\sim 1/3$  of the branching ratio to  $WW$ . For very high Higgs boson masses,  $m_H > 2 m_{top}$ , also the decay into a top pair becomes important. Higgs decays to leptons less often than to heavy quarks, given the lower mass. However the most favored leptonic decay mode  $H \rightarrow \tau^+\tau^-$  reaches branching fractions of  $\sim 7-8\%$  for  $m_H = 100-140 \text{ GeV}/c^2$ . The Higgs boson does not couple directly to massless particles, but it can also decay through loop-mediated processes, as it is for the  $H \rightarrow \gamma\gamma$  decay mode, which, despite a branching fractions  $\sim 0.1\%$  for  $m_H \lesssim 160 \text{ GeV}/c^2$ , is a favorable final state to its search.

For  $m_H \gtrsim 135 \text{ GeV}/c^2$  the most favorable Higgs boson searches rely on diboson reconstruction in the final state, hence the accurate knowledge of these latter processes increase the experimental sensitivity to the Higgs boson production.

### 2.3.3 Experimental Situation

Direct searches for the production of the Higgs boson have been performed at the four experiments (ALEPH, DELPHI, L3 and OPAL) of the  $e^+e^-$  LEP collider. The main production mode of the Higgs boson at the LEP collider were the Higgs-strahlung from a virtual  $Z$  boson (see Figure 2.7(d)) and, for a smaller fraction, the vector boson fusion (see Figure 2.7(e)). The LEP experiment performed detailed analyses on the  $\sim 2461 \text{ pb}^{-1}$  of total integrated luminosity collected by the four detectors (at  $189 \leq \sqrt{s} \leq 209 \text{ GeV}$ ) by considering various  $Z$  and  $H$  decay modes. No evidence of the Standard Model Higgs boson production has been found and the combination of these searches [55] establishes a lower bound of  $114.4 \text{ GeV}/c^2$  at 95% Confidence Level (C.L.) on its mass  $m_H$ .

In addition to direct searches, indirect constraints on the Higgs mass come from electroweak measurements of the SM observables. The radiative corrections to the relative strength  $\rho$  of the charged and the neutral current depends in particular on the  $W$  mass  $m_W$ , the top mass  $m_t$ , and the Higgs boson mass  $m_H$  [56]. The parameter  $\rho = 0$  at tree level, but deviation from this value are due to one-loop correction that might involve the Higgs boson. It is clear that precise measurements of the top and  $W$  mass are crucial in constraining the Higgs boson mass. Other SM electro-weak observables depend on the Higgs mass through radiative corrections, although their dependence is in general milder than for the  $\rho$  parameter.

Precision measurement results on electroweak observables from the four LEP experiments, CDF, D0, NuTeV, and SLD are combined by the LEP electroweak working group [57] and is evaluated the dependence of the SM observables on the Higgs boson mass. A global SM fit as a function of the Higgs mass is then performed using as input all the available observables [58]. Figure 2.10 shows the result of the fit to the electroweak constraints (as of 2008) as a function of the Higgs mass.

In the latest years the experiments located at the Tevatron and LHC colliders carried out several searches for the SM Higgs boson, providing direct constraints on its mass. The SM Higgs boson search strategy at hadron colliders is driven by the most dominant decay modes, since these channels provide the greatest predicted sensitivity for observing a Higgs boson signal in the data. Formally, the sensitivity of a particular search channel is determined from the calculation of an expected limit on its production rate (as will be discussed in Section 6.0.11). At first the Tevatron excluded at 95% C.L. the Higgs mass range  $100-103 \text{ GeV}/c^2$  and  $147-180 \text{ GeV}/c^2$  [59]. The CMS and Atlas experiments at LHC extended the exclusion to the regions  $110.0-117.5 \text{ GeV}/c^2$ ,

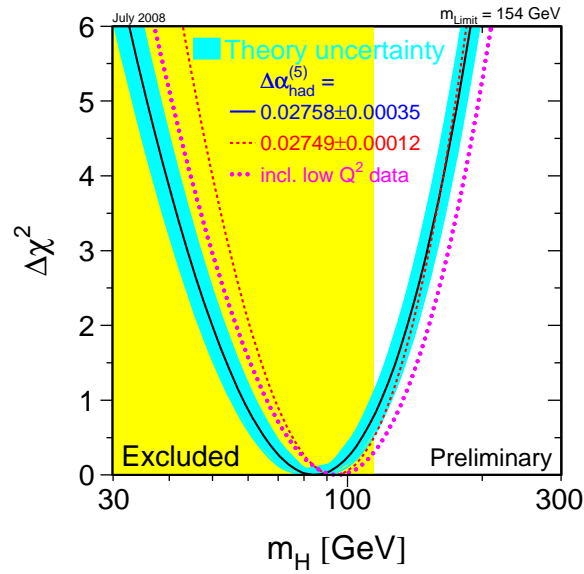


Figure 2.10:  $\Delta\chi^2 = \chi^2 - \chi_{min}^2$  of the fit to electroweak precision measurements performed by the LEP-EWWG.

118.5-122.5 GeV/ $c^2$ , and 127-600 GeV/ $c^2$  [60, 61].

Figure 2.11 shows a summary of the current sensitivity of Lep, Tevatron, and LHC experiments, as of June 2012, highlighting the excluded mass range and the excess observed in Tevatron data.

### Higgs Boson Discovery

On Summer 2012 both the Tevatron and LHC experiments reported an excess of data in the Higgs boson mass range  $115 < m_H < 135$  GeV/ $c^2$ , which motivated the scientific community to further improve Higgs search analyses in this mass region. In December 2012, analyzing data collected at  $\sqrt{s} = 7$  and 8 TeV, the ATLAS and CMS experiment reported the observation of a new particle compatible with being a SM Higgs boson with  $m_H \simeq 125$  GeV/ $c^2$  [60, 61]. The discovery of this new particle has been a great achievement in high-energy physics and set the starting point for further investigations of the properties of this particle, to understand and eventually confirm that this particle is the SM Higgs boson.

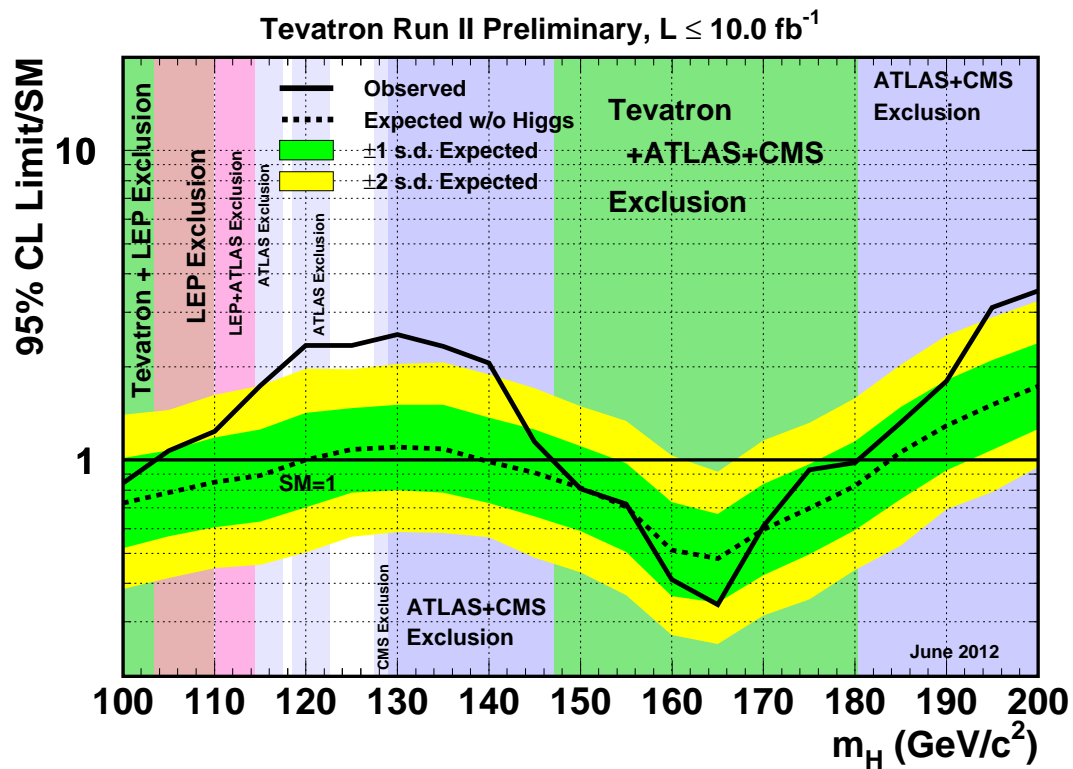


Figure 2.11: Mass exclusion limits (95% C.L.) obtained from the combination of all the Tevatron searches for a SM Higgs boson [59]. Exclusion limits obtained from the CMS [61] and Atlas [60] Collaborations are overlaid.

# The Tevatron Collider and CDF experiment

## Contents

---

3.1	Tevatron Accelerator Complex . . . . .	<b>25</b>
3.1.1	Proton Source and Acceleration . . . . .	26
3.1.2	Main Injector . . . . .	27
3.1.3	Antiproton Production and Storage . . . . .	27
3.1.4	The Tevatron Collider . . . . .	28
3.2	The CDF Detector . . . . .	<b>28</b>
3.2.1	Tracking System . . . . .	30
3.2.2	Calorimeter System . . . . .	35
3.2.3	Muon Detector System . . . . .	37
3.2.4	Cherenkov luminosity counters . . . . .	39
3.2.5	Online Data Acquisition . . . . .	40

---

The measurements presented in this thesis are based on the complete sample of data collected by the CDF II (Collider Detector at Fermilab for Run II) experiment, from February 2002 to September 30<sup>th</sup> September 2011. The CDF multi-purpose detector is located at one of the two instrumented interaction points along the Tevatron  $p\bar{p}$  collider, producing collisions at  $\sqrt{s}=1.96$  TeV.

In this Chapter are described the accelerator facility (Section 3.1) and the CDF detector apparatus (Section 3.2)

## 3.1 Tevatron Accelerator Complex

The Tevatron in Batavia, Illinois, USA is the first large-scale superconducting synchrotron in the world. Originally named the *Energy Doubler* since as a proton-synchrotron it was reaching twice the energy of the original Fermilab facility (the “Main Ring”), it began operation in 1983 in fixed target mode and in 1985 as a proton-antiproton collider.

From 1985 to 1996 various periods of collider or fixed target operations alternate with each other, during what is called the Tevatron Run I. From January 2001 to September 2011 the Tevatron fully operated as a  $p\bar{p}$  collider in the so-called Run II, producing particles interactions at an energy in the center of mass frame of 1.96 TeV. Along the Tevatron ring there are two multi-purpose collider detectors, CDF and D0, that have undergone extensive upgrades during the 6 years long (1996 to 2001) preparations for Run II.

---

A schematic layout of the Fermilab accelerator complex is shown in figure 3.1. In section 3.1.1 this complex (accelerator feeding the Tevatron, and Tevatron Collider) will be briefly described.

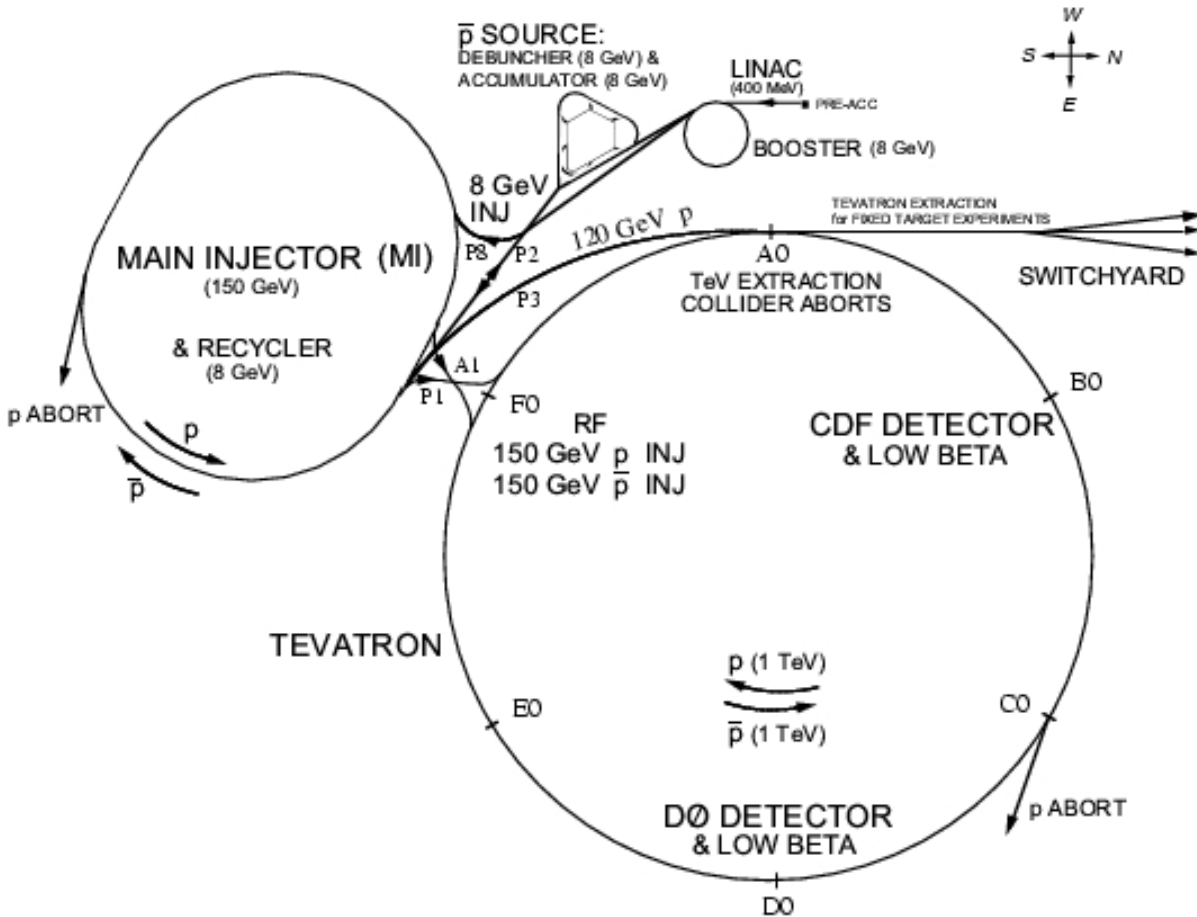


Figure 3.1: The accelerator complex of the Fermi National Accelerator Laboratory.

### 3.1.1 Proton Source and Acceleration

Acceleration begins with a **Cockroft-Walton electrostatic Preaccelerator**. Here  $H^-$  ions are accelerated from a ion source to 750 KeV.

Ion bunches at 750 KeV are fed into the **Linear Accelerator (Linac)**. The Linac is approximately 140 m long and comprises two sections; in the first one five accelerating cavities with a drift tube design fed by a single RF generator accelerate ions to approximately 166 MeV. The second one, comprising 7 RF cavities fed by a more modern set of Klystron amplifiers, ramps ions to 400 MeV<sup>1</sup>. At the Linac exit the ion beam strikes a thin carbon target and turns into a proton beam by electron stripping.

Stripped 400 MeV protons enter the **Booster**, a 8 GeV synchrotron whose diameter is about 150 m. To maintain a constant circular orbit the dipole magnetic field in the Booster increases

<sup>1</sup> The 400 MeV final energy is the result of a Linac upgrade, that took place in 1993 and increased the boost in the second Linac sector from 200 MeV to 400 MeV. This effort allowed to double the number of protons per bunch and to increase by about 50% the production rate of antiprotons.

from 0.74 Tesla to 7 Tesla during acceleration.

Both Linac and Booster provide pulses up to  $5 \cdot 10^{12}$  protons at a rate of about 5 Hz for antiproton production every 1.5 seconds, or  $6 \cdot 10^{10}$  protons per bunch in series of 5 to 7 bunches, repeated 9 times per second (in average).

### 3.1.2 Main Injector

From the Booster the proton beam is fed into the Main Injector whose role is either to accelerate protons as needed for injection in the Tevatron and to deliver beam to the antiproton production target.

The original Tevatron injector was the Main Ring, built to provide primarily 400 GeV protons to fixed target experiments. The Main Ring limited aperture was a limit to the whole accelerator performances. The *Main Injector* was designed to solve this problem and located in separated tunnel for an easier operation of the complex.

The **Main Injector** is a synchrotron with a circumference of about 3 km. It is able to accelerate protons of 8 GeV energy up to 150 GeV. It operates at 120 GeV for antiproton production, while 150 GeV protons are delivered to the Tevatron.

The Main Injector is also used to give beam to number of fixed target experiments, noticeably on secondary neutrino beams.

### 3.1.3 Antiproton Production and Storage

A pulse of  $5 \cdot 10^{12}$  protons at 120 GeV is extracted every  $\sim 2.2$  seconds from the Main Injector and directed to the antiproton station, a rotating 7 cm-thick target made of nickel alloys containing chromium, iron and other metals. The resulting particles spray contains some antiprotons with a broad momentum and wide-spread spatial distribution.

A cylindrical lithium lens ( $760 T/m$ ) focuses the particles produced around the forward direction. Negative particles in  $35 mrad$  cone about the forward direction are selected by using a 1.5 T pulsed dipole magnet and injected in the *Debuncher Storage Ring*. Typically,  $\sim 21$  antiproton per  $10^6$  protons on target are collected.

In the **Debuncher ring**, a rounded triangular-shaped synchrotron with mean radius of 90 meters stochastic cooling and bunch phase rotation<sup>2</sup> reduce by about 10 times the momentum spread in the bunches.

After each beam pulse the Debuncher is emptied. The antiproton bunches (with an energy of  $8 \pm 0.018$  GeV) are transferred with a 60%-70% efficiency to the **Antiproton Accumulator**, a 75 m mean radius storage ring of larger acceptance housed in the same tunnel as the Debuncher. In the accumulator multiple beam pulses are stacked and  $\bar{p}$  are further cooled to increase the antiproton phase space density.

For the time being problems in antiproton collection, cooling and stacking are among the main causes limiting the final Tevatron luminosity since a very small fraction of the proton incident on target produces antiprotons and only a part of these can be stored.

A further improvement of the antiproton source is the **Recycler**, a post-accumulator storage ring of constant 8 GeV energy, located in the Main Injector enclosure and composed of permanent

<sup>2</sup> Stochastic technique is a way of narrowing the particle distribution in transverse and longitudinal momentum around the average value. There is not any accompanying beam-loss. This goal is achieved by applying iteratively a mechanism which recognizes deviation from spatial orbit of a 8-GeV antiproton in upstream sensors and makes appropriate correction downstream.

magnets. Because of its larger acceptance it can store an antiproton current twice as large as the Accumulator.

When a new store is ready for collider operation, previously used antiprotons are transferred to the Recycler while protons are thrown away. Then new antiprotons are transferred from either Accumulator or the Recycler to the Main Injector in order to increase their energy up to 150 GeV. Antiprotons are finally transferred to the *Tevatron*, where an opposite proton beam of the same energy was previously stored.

### 3.1.4 The Tevatron Collider

The **Tevatron** is a 1 km radius synchrotron that comprises about 1000 superconducting magnets including 772 dipoles. Each dipole is approximately 6 m in length and 4 tons in weight. The superconducting coils are made up of niobium-titanium wires embedded in copper. A 4400 A current in the dipoles provides a 4.2 T magnetic field. All superconducting materials are kept at 4 K temperature.

As written in section 3.1.2 the Tevatron receives protons and antiprotons from the Main Injector at 150 GeV. At injection 36 bunches composed typically of  $300 \times 10^9$  p are transferred at 150 GeV with a timing separation of 360 ns from each other.

Both protons and antiprotons orbit are in the same vacuum pipe. Electrostatic separators reduce to a negligible amount the unwanted interactions, by keeping the beams away from each other at all points in the orbit helix<sup>3</sup>.

Protons and antiprotons are accelerated to 980 GeV. A tour of the Tevatron takes about 21  $\mu$ s. About one minute is needed to reach the final beam energy.

High-power focusing quadrupole magnets minimize the beam section at the interaction regions to maximize the collision rate.

The resulting transverse beam distributions may be approximated by 2D Gaussian functions, with  $\sigma_T \approx 30 \mu$ m. The typical longitudinal dimension of a bunch is 60-70 cm. The event source is roughly distributed longitudinally as a Gaussian with  $\sigma_z = 28$  cm<sup>4</sup>.

Tevatron bunches are organized in three trains; within a train the inter-bunch time is 396 ns while inter-train time is 2.6  $\mu$ s. The intra-train empty sectors are needed for the fast *kicker magnets* to abort the beam into a dump before the arrival of the next train in case of emergency.

In Figures 3.2 and 3.3 are respectively shown the integrated and peak luminosity of Tevatron through different runs.

The record peak luminosity reached by the Tevatron is  $\sim 3.65 \cdot 10^{32} \text{ cm}^{-2}\text{s}^{-1}$  corresponding to about 5 interactions per bunch-crossing on average.

## 3.2 The CDF Detector

The Collider Detector at Fermilab (CDF) was designed to study  $p\bar{p}$  collision at the Tevatron. Commissioned in 1987 it was upgraded in 2001 in order to be adapted to the higher collision rate coming from the increased instantaneous luminosity delivered by the accelerator.

The CDF coordinate system is a right handed coordinate system defined such that the z axis lay along the beam line, in the direction of the proton beam at the nominal ( $z_0$ ) collision point.

<sup>3</sup> Intrabeam distance is typically 5 times the sum of the beam widths (in a Gaussian approximation).

<sup>4</sup> The about 28 cm length of the interaction region is determined by the overlap of the two approximately longitudinally Gaussian bunches.



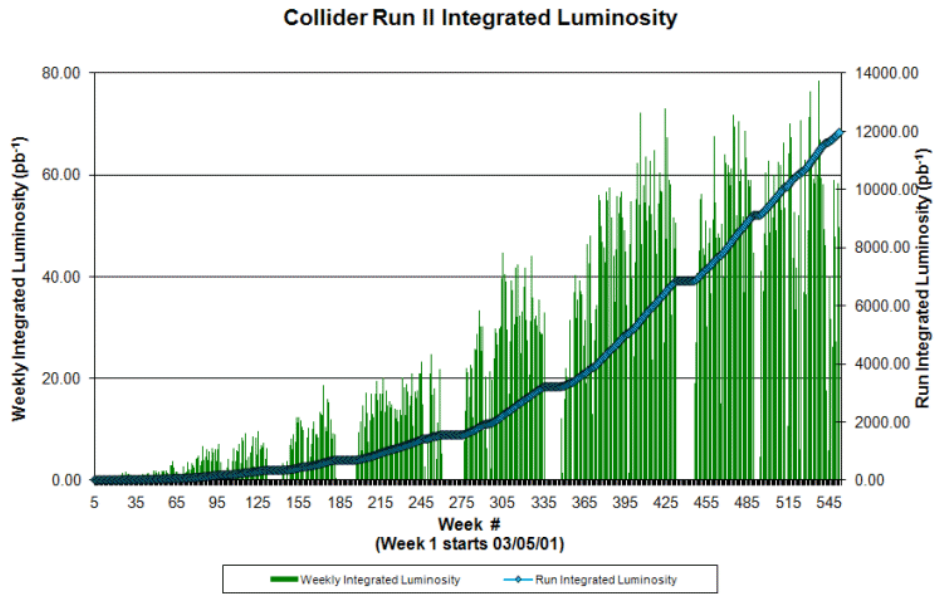


Figure 3.2: Run 2 Integrated luminosity as a function of time.

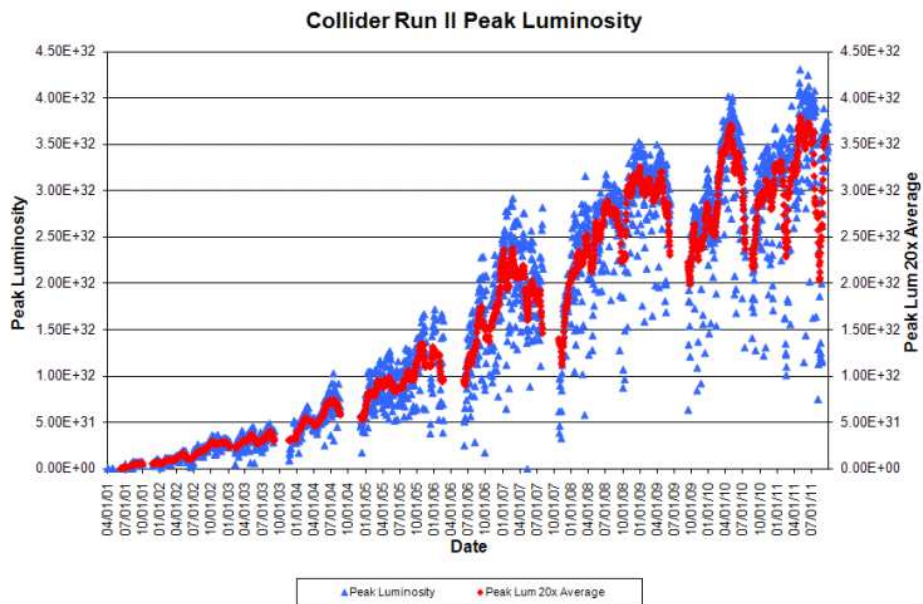


Figure 3.3: Run 2 Peak luminosity in the stores and its average, as a function of time.

The positive  $y$  direction is defined to point vertically upward. This leaves the  $x$  direction pointing outward (roughly northwest) where unit vectors satisfy  $\hat{z} = \hat{x} \times \hat{y}$ .

It is useful to describe CDF detector geometry also using a cylindrical  $(r, \phi, z)$  coordinate-system, where  $r$  is the radial distance from the beam line and  $\phi$  is the polar angle in the plane perpendicular to the beam line. The origin is the geometric center of the detector.

It is often convenient to use a polar variable invariant under boost along  $\hat{z}$ . This variable is

the *rapidity* defined as

$$y = \frac{1}{2} \ln \left( \frac{E + p \cdot \cos\theta}{E - p \cdot \cos\theta} \right) \quad (3.1)$$

where  $E$ ,  $p$ ,  $\theta$  are respectively the energy, momentum and polar angle of the considered particle. At high energies and away from very forward angles  $y \approx \eta = -\ln[\tan(\theta/2)]$  called *pseudo-rapidity*<sup>5</sup>.

The Run II Detector (Fig. 3.4 and 3.5) is composed of several subdetectors, each one optimized for a specific task.

Starting from the interaction point and following the path of an outgoing particle within acceptance there are:

1. a tracking system enclosed by a superconducting solenoid (1.5 m in radius and 4.8 m in length), which generates 1.4 T magnetic field parallel to the beam axis. The magnetic field is uniform in the tracking region at the level of 1% or better.
2. finely segmented calorimeters.
3. planar muon drift chambers backed by scintillation counters.

In the next sections these sub-systems will be discussed.

Some of the components (the time-of-flight detector, the small angle spectrometers on beam pipe, etc.) of the full CDF II detectors have been neglected since are not used in this thesis.

A detailed description of the upgraded detector can be found in [62].

### 3.2.1 Tracking System

Charged particle within the tracking system acceptance encounter an inner silicon tracking system and outer gas drift-chamber as shown in Figure 3.6.

Within the solenoid field they follow helical trajectories which are measured by the system in order to estimate their momentum.

#### Inner Tracker

The **Inner Tracker** is composed of eight layers (seven at  $\theta = 90^\circ$ ) of silicon sensors arranged in approximately cylindrical sub-systems coaxial with the beam-pipe: Layer 00 (L00), the Silicon Vertex Detector (SVXII), and the Intermediate Silicon Layers (ISL). Figure 3.7 zooms on the Inner Tracker structure.

All silicon microstrip sensors have a space resolution of  $\sim 12 \mu\text{m}$  in the direction transverse to the beam. They also provide  $z$ -measurements with reduced accuracy.

**L00** is 90 cm long, radiation hardened single side micro strip detector, and it is mounted directly on the beam pipe. It is located at radial distance of 1.35 to 1.62 cm from the beam axis. L00 purpose is to improve the track impact parameter (the measured distance of minimum approach to the beam axis) resolution ( $\sim 40 \mu\text{m}$  for tracks with  $p_T \geq 2 \text{ GeV}/c$ ) and compensate for the multiple scattering degradation for particles that travel across bulkhead.

<sup>5</sup> In CDF literature are usually distinguished  $\eta_{det}$ , which is relative to the geometrical center of the detector, and  $\eta$ , which is measured with respect to the interaction point  $z_0$  where particles originated. Usually the former symbol is used for describing the detector geometry while the latter for outgoing particles. For simplicity the same symbol will be used in both cases.

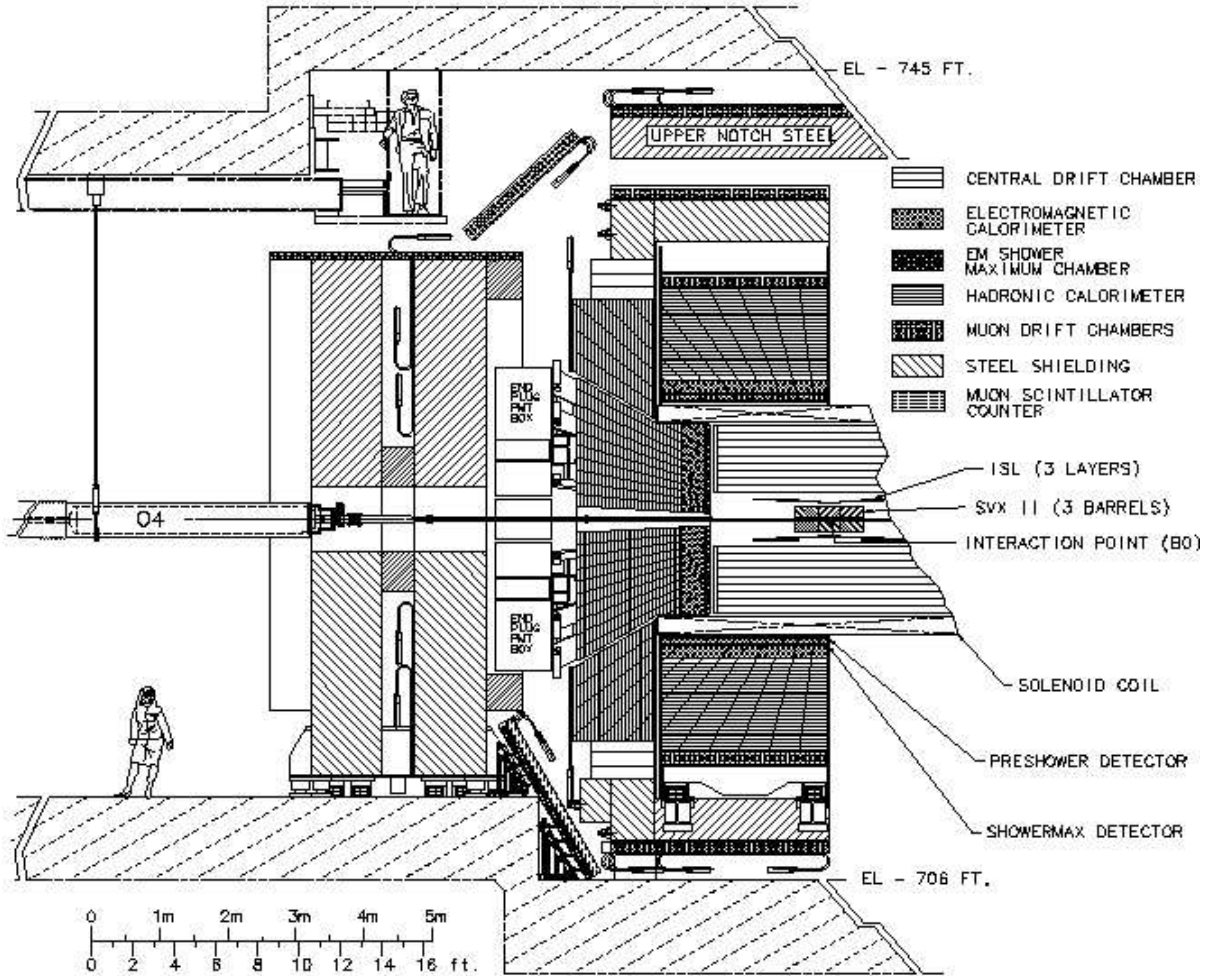


Figure 3.4: Elevation view of the CDF II detector.

**SVXII**, shown in fig 3.8, extends radially from 2.5 cm to 10.7 cm. It is segmented into three 29 cm barrels along the  $z$  axis: this allows for a  $|\eta| < 2.0$  coverage. Each barrel carries 5 layers of double-sided microstrip wafers. Four silicon wafers are mounted on light support structures called ladders. Twelve ladders placed along a circumference make a layer. The double-sided imprint of the wafers allow for 3D position measurements: one side of the wafer has strips along the beam axis, the other one has either  $90^\circ$  or  $1.2^\circ$  stereo strips. This sub-detector provides also some information on the Energy deposit profile as a function of the path length traveled by a particle ( $dE/dx$ ) and has a  $12 \mu m$  resolution in the direction transverse to the beam on the single hit.

The main **ISL** purpose is to compensate for the incomplete coverage of other sub-detectors in the region  $|\eta| > 1$  by providing precision tracking at  $1 < |\eta| < 2$ . It consist of 5 layers of double sided silicon wafers (same wafers as for SVX II). Four layers are at  $1 < |\eta| < 2$  (at radii of 20 and 28 cm, as shown in Fig. 3.7), one layer is at  $|\eta| < 1$ .

The combined resolution of the CDF inner trackers for high momentum tracks is  $\sim 40 \mu m$  in impact parameter and  $\sim 70 \mu m$  along  $z$  direction.

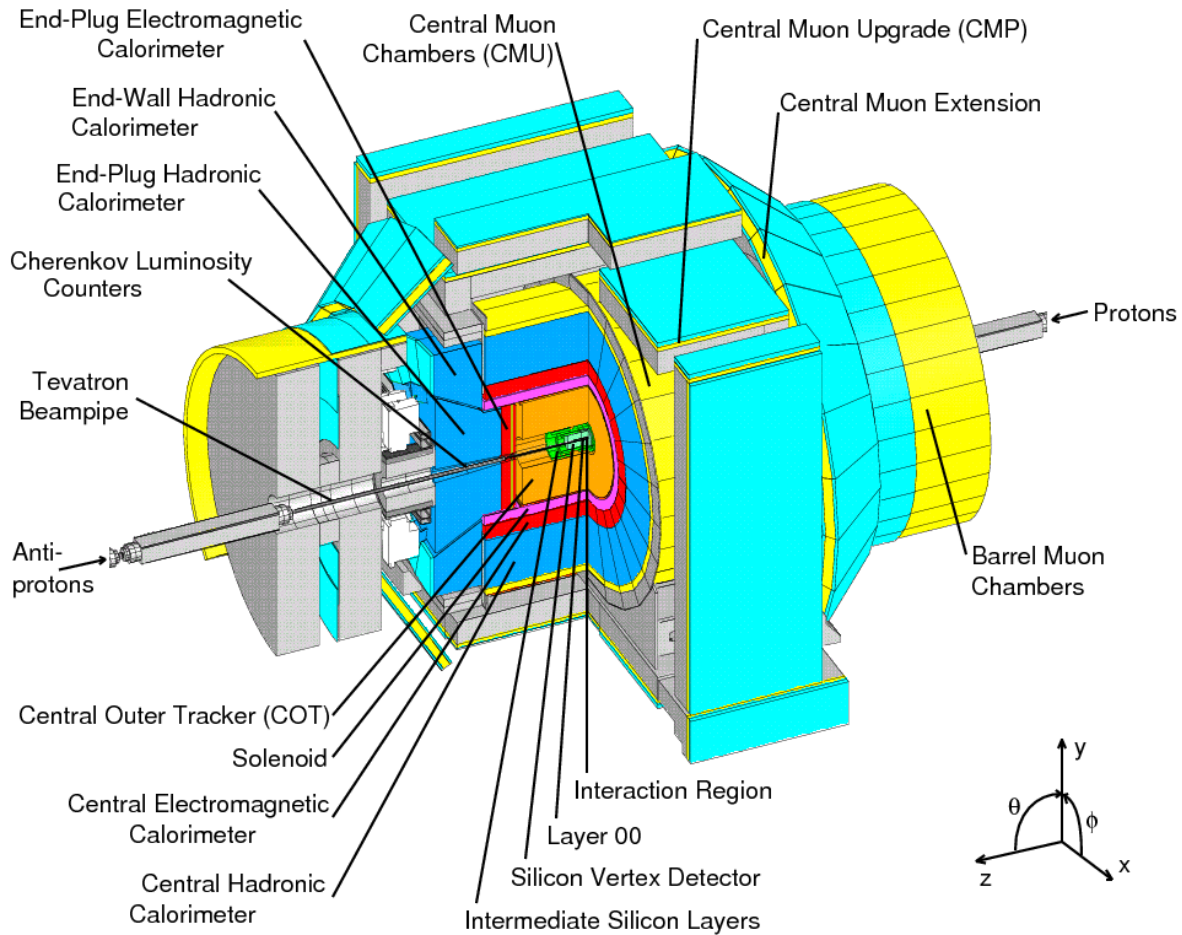


Figure 3.5: Isometric view of the CDF II detector.

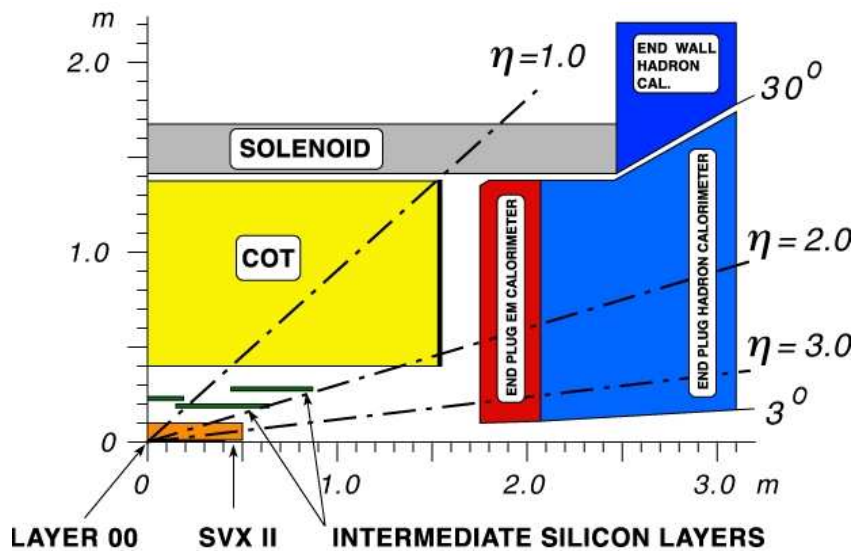


Figure 3.6: Longitudinal view of the Tracking System of the CDF II detector.

### Central Outer Tracker

The **Central Outer Tracker** (COT) is an open-cell wire drift chamber filled with a gas mixture of Argon, Ethane and  $CF_4$  in proportion 50%, 35% and 15%; it has a cylindrical shape and is

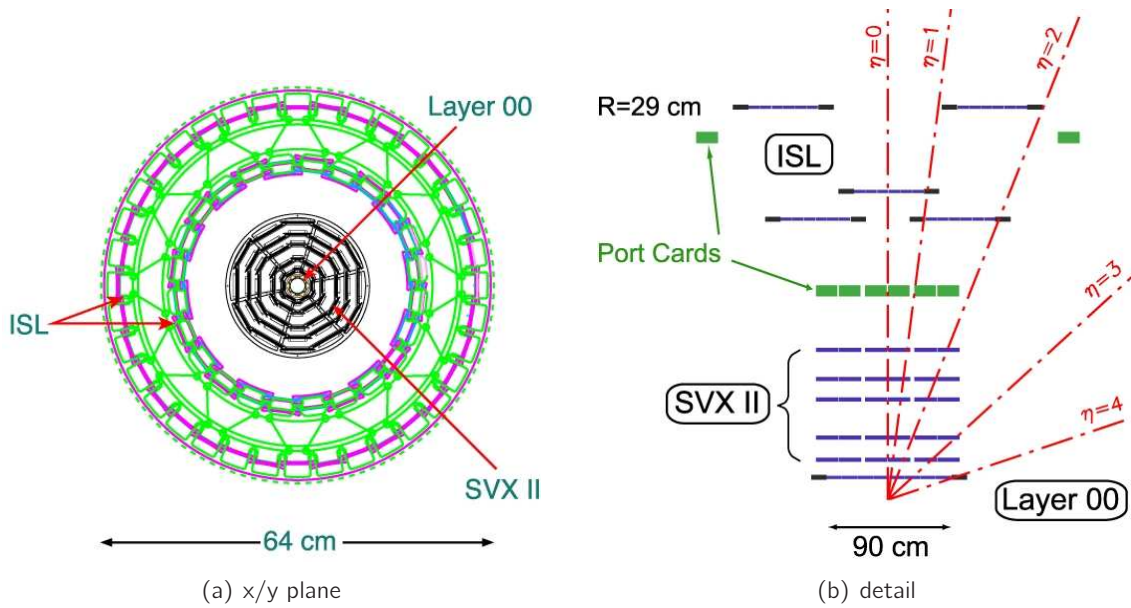


Figure 3.7: (a) Sketch of the silicon detector in a  $x/y$  projection. (b) Cutaway transverse to the beam of the three inner tracker sub-systems.

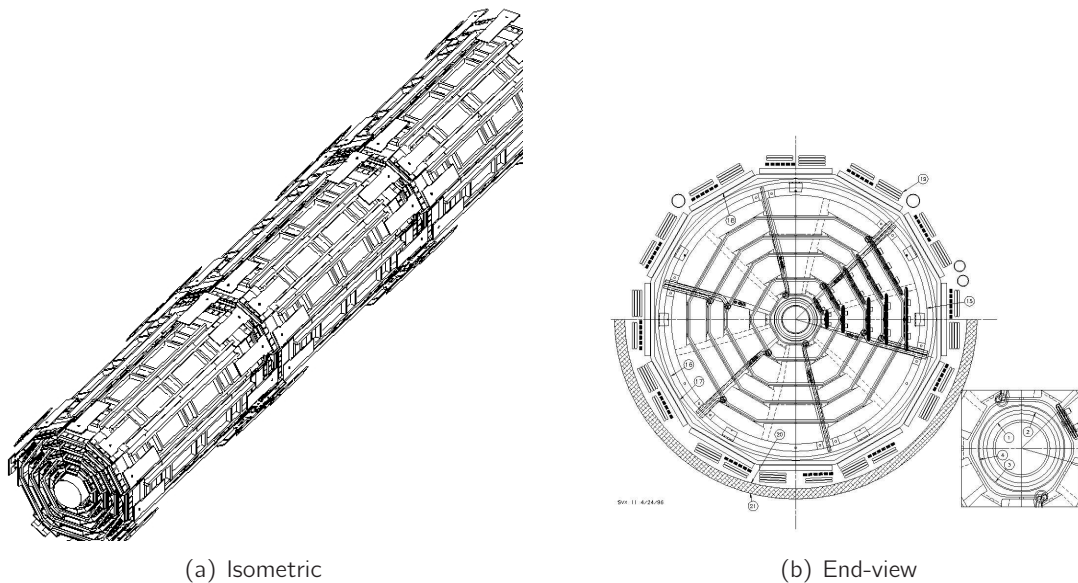


Figure 3.8: Isometric (a) and end-view (b) of the CDF Silicon Vertex Detector.

radially right outside the ISL. Its internal and external radii are 43 cm and 137 cm respectively. The COT provides full tracking in the central region ( $|\eta| < 1$ ), even if its maximum geometrical acceptance reaches up to  $|\eta| < 2$  (see Fig. 3.6), where tracking performances are reduced. The COT is composed of 4 axial and 4 stereo<sup>6</sup> superlayers of azimuthal cells. Each cell has alternated sense and field shaping wires (Fig. 3.9). Within the cell width, the trajectory of a particle is sampled 12 times (by sense wires spaced 0.583 mm apart). Figures 3.9 show a portion of the COT endplate.

<sup>6</sup> Stereo superlayers are tilted at  $\pm 2^\circ$  width respect to the  $z$  direction. Axial layers provide tracking information in  $r - \phi$  plane, stereo layers are also sensitive to the  $z$  direction.

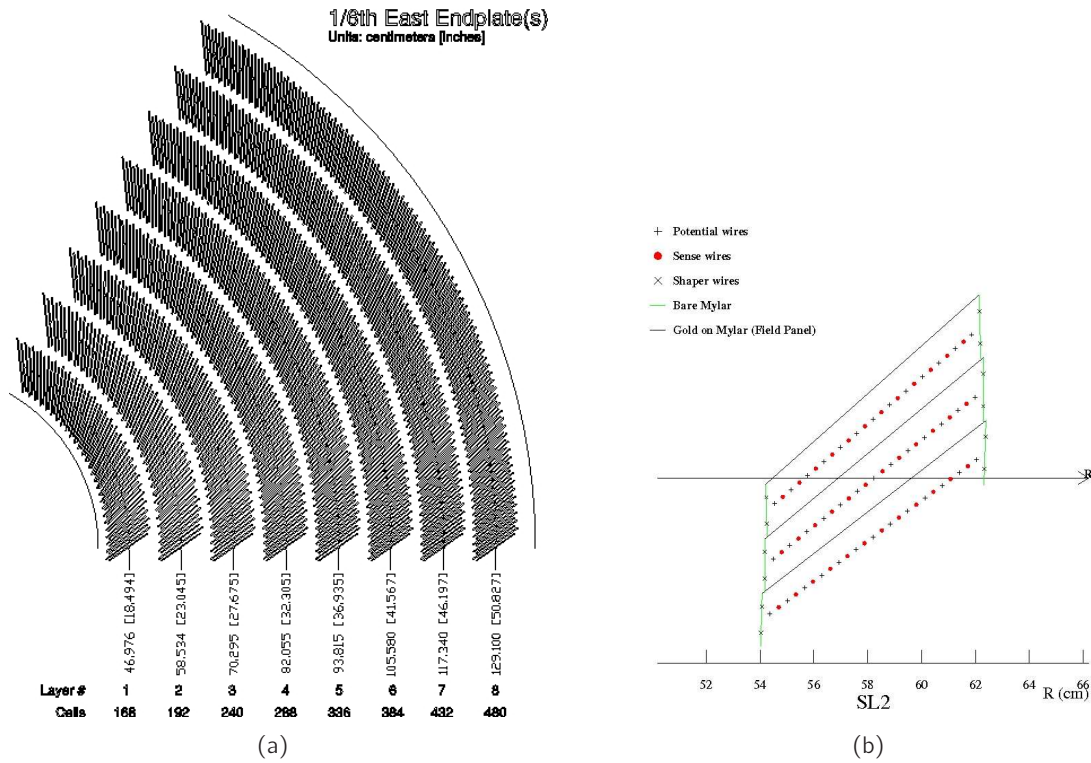


Figure 3.9: Slots housing the wire holding fixtures of one quadrant of a COT endplate (a). Drift and field wires in three cells (b). The horizontal arrow shows the radial direction.

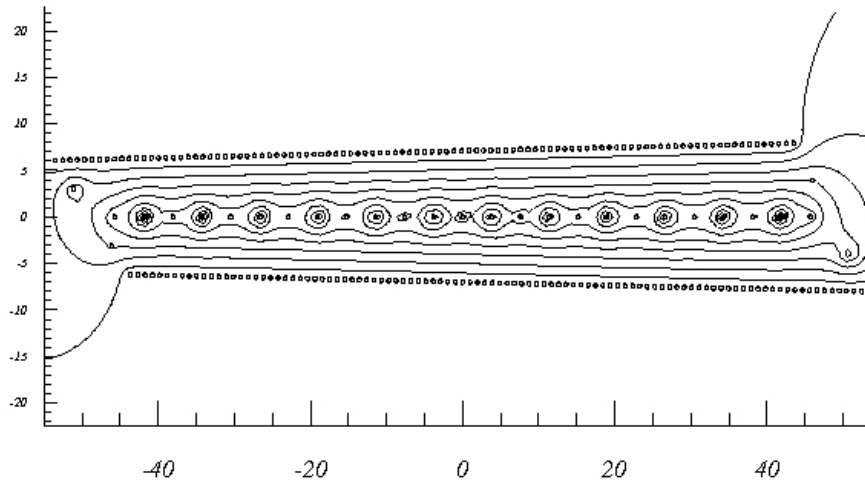


Figure 3.10: Equipotential lines in a COT cell.

Inside the solenoid magnetic field, the drifting electrons experience a Lorentz force which rotates their path. The cells are tilted by  $35^\circ$  with respect to the radial direction in order to make the electrons drifting perpendicularly to the radius for optimal momentum resolution<sup>7</sup>.

<sup>7</sup> For best momentum resolution, the optimal correlation between drift time and hit distance from wire is for

The single hit position is measured with an uncertainty of  $\sim 140 \mu\text{m}$  which translates into an overall transverse resolution

$$\sigma(P_T)/P_T = 0.15\% \cdot P_T[\text{GeV}/c] \quad (3.2)$$

$P_T$  being the transverse momentum of the particle.

### 3.2.2 Calorimeter System

The CDF calorimeter performs a destructive measurement of the particle energy, providing a signal proportional to it. CDF uses scintillator sampling calorimeters divided into an inner electromagnetic and an outer hadronic compartment. Both calorimeters are segmented into projective towers. Each tower consists of alternating layers of passive absorber material ( $Pb$  in the electromagnetic and  $Fe$  in the hadronic compartment) and plastic scintillator for shower sampling. The light from the scintillator plates is read out through wavelength-shifting bars or plates and (WLS) light guides by photo-multiplier tubes (PMT)(see figure 3.11).

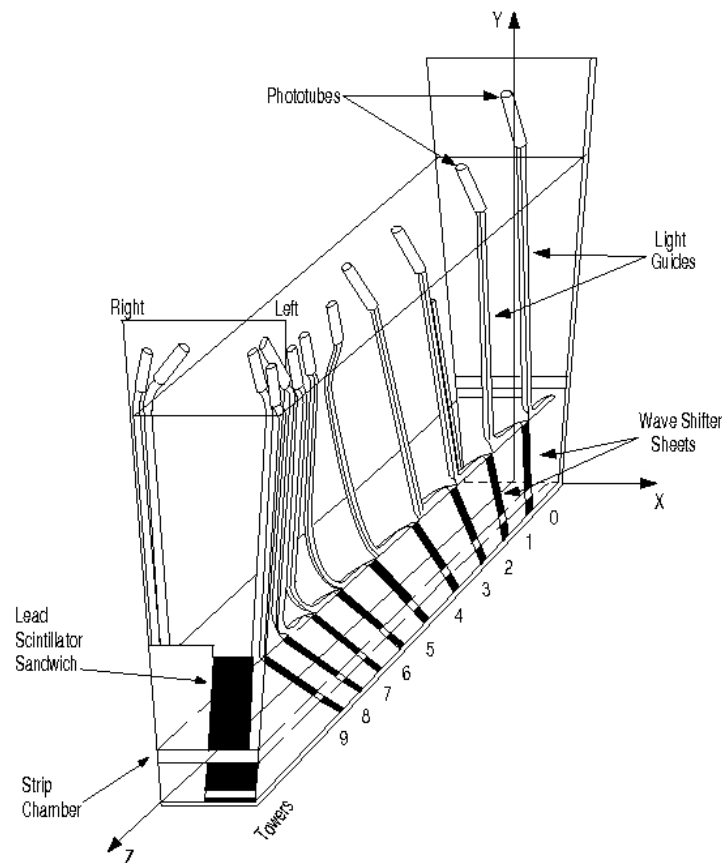


Figure 3.11: Light-shifter plates connected to light guides and to photomultipliers of the front electromagnetic sector of a central calorimeter wedge.

High energy electrons and photons generate an electromagnetic shower which is mostly absorbed in the front calorimeter compartment<sup>8</sup>. For charged particles heavier than the electron, electrons drifting in the direction transverse to the radius.

<sup>8</sup> A shower is a cascade of particle. In the case of photons and electrons the showers are composed mainly of electrons, positrons and photons.

radiative energy losses are negligible to a first approximation. High energy hadrons interact with the detector matter mostly through inelastic collisions with nuclei of the absorbing medium. Particles produced in the nuclear interactions can lose their energy by ionisation and secondary nuclear interactions. Mixed electromagnetic and hadron showers that originate are absorbed in the entire (em+had) calorimeter.

CDF calorimeters provide full azimuthal coverage and up to 3.6 in  $|\eta|$ . The calorimeter system includes the Central Electromagnetic Calorimeter (CEM) and the Central Hadronic Calorimeter (CHA) in the  $|\eta| < 0.9$  region, the Endwall Hadronic Calorimeter (WHA) in  $0.9 < |\eta| < 1.3$  and the electromagnetic and hadronic plug calorimeters (PEM,PHA) in  $1.1 < |\eta| < 3.6$  (see Figures 3.5 and 3.6).

### Central Calorimeter

The central calorimeters, CEM, CHA and WHA are composed of two parts joining at the geometrical centre of CDF<sup>9</sup>. Central calorimeters are azimuthally divided into 24 wedges covering  $\sim 15^\circ$  in  $\phi$  each. Each wedge is divided into projective towers of size  $\delta\eta = 0.1$  in pseudorapidity. The **CEM calorimeter** is made of 31 alternate layers of 0.5 cm plastic scintillator and 0.32 cm thick lead absorbers: the total amount of material is  $18 \cdot X_0$  ( $X_0$  is the electron radiation length). The CEM energy resolution is:

$$\sigma_{E_t}/E_T = 13.5\%/\sqrt{E_T[\text{GeV}]} \oplus 2\% \quad (3.3)$$

$E_T$  being the energy of an electron or a photon hitting the calorimeter perpendicularly to its front. In this case the  $\oplus 2\%$  term represent the contribution to the resolution due to the lateral or longitudinal shower leakage, approximately constant.

CEM also includes two additional specialized detector: the Central Electron Strip Chambers (CES) and the Central Preshower (CPR). CES is a combined strip/wire gas proportional chamber embedded in CEM at about  $6 \cdot X_0$  since there is expected the maximum longitudinal development of the electromagnetic shower. The CES purpose is to measure the position and the shape of electro-magnetic showers in both transverse plane and longitudinal direction. CES resolution is about 1 cm in  $z$  and 1 mm in  $r - \phi$ . CPR is a set of scintillator tiles located in front of the calorimeter wedges which help distinguishing electrons from charged hadrons by gauging their probability of showering in the detector material prior to entering the calorimeter.

The **CHA calorimeter**, located behind CEM, is composed of 32 alternate layers of 1 cm plastic scintillators and 2.5 cm thick steel. The **WHA calorimeter** employs the same technology as CHA except for the smaller number of layers (15) and the larger thickness of the radiator medium (5 cm per layer). The total calorimeter thickness is  $\sim 4.7 \lambda_0$  ( $\lambda_0$  is the absorption length) for both CHA and WHA.

Resolutions of CHA and WHA for perpendicular particle entrance are:

$$\sigma_{E_t}/E_T = 50\%/\sqrt{E_T[\text{GeV}]} \oplus 3\% \quad (3.4)$$

$$\sigma_{E_t}/E_T = 75\%/\sqrt{E_T[\text{GeV}]} \oplus 4\% \quad (3.5)$$

### Plug Calorimeters

The **PEM calorimeters** (Fig. 3.12) have the same structure as the CEM: same tower segmentation in  $\eta$ , but finer in  $\phi$  (a  $\phi$  coverage of  $7.5^\circ$ ) for  $|\eta| < 2.11$ , 22 layers of 4.5 mm thick

<sup>9</sup> In this zone,  $\eta = 0$ , there is an instrumented area (*crack*).



lead alternate with 22 layers of 4 mm thick scintillator.  
The PEM transverse energy resolution is:

$$\sigma_{E_T}/E_T = 16\%/\sqrt{E_T[\text{GeV}]} \oplus 1\% \quad (3.6)$$

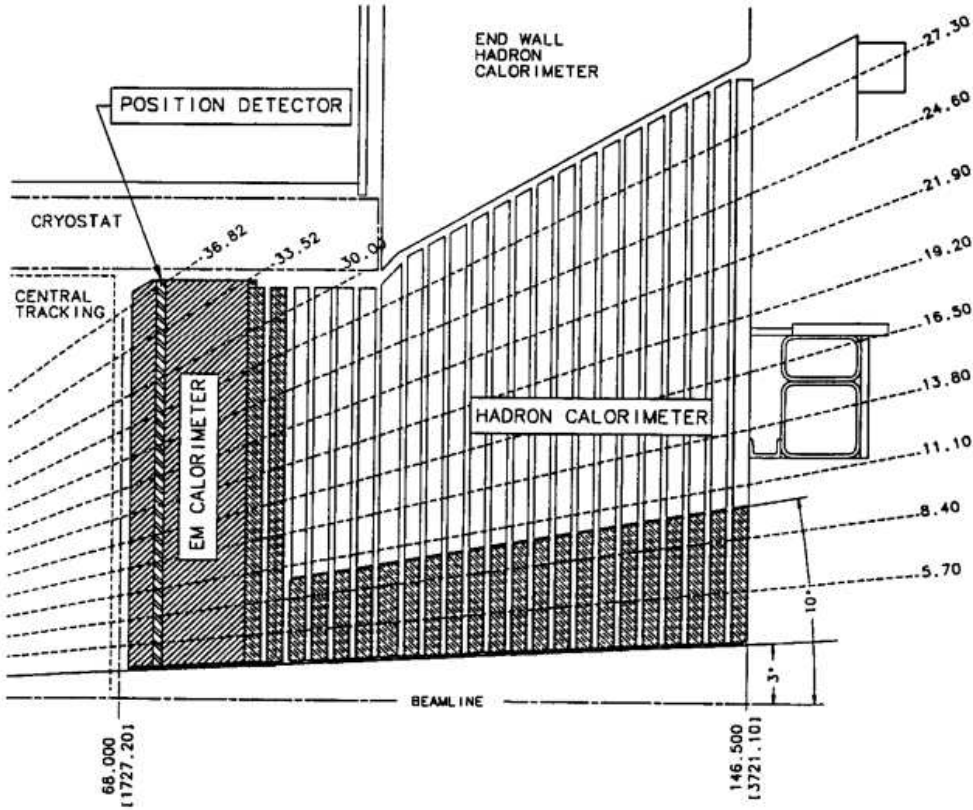


Figure 3.12: Longitudinal view of Plug Calorimeters.

As for CEM, PEM is equipped with a shower maximum detector (PES). More details can be found in [63]. PHA, located behind PEM, has the same tower segmentation. The technology is the same as for CHA, with 23 layers of 2cm thick steel absorber alternating with 6 mm thick scintillator. The total amount of material corresponds to  $\sim 4.7 \lambda_0$ . PHA resolution is:

$$\sigma_{E_t}/E_T = 80\%/\sqrt{E_T[\text{GeV}]} \oplus 5\% \quad (3.7)$$

### 3.2.3 Muon Detector System

Muons interact electromagnetically but, since they have a higher mass than electrons, when they cross electromagnetic calorimeter they do not shower in it. At the typical energies they are produced in  $p\bar{p}$  collision at the Tevatron they interact in calorimeters as minimum ionizing particles (MIP) so they lose just a minimal fraction of their energy crossing it. For this reason systems dedicated to detect muons are located in the outermost shell of the detector. Muon momenta are measured in the tracker.

Four independent systems are used to detect muons in the  $|\eta| < 1.5$  region: the Central Muon Detector (CMU), the Central Muon Upgrade Detectors (CMP), the Central Muon Extension

(CMX) and the Intermediate Muon detector (IMU). The  $\eta - \phi$  coverage of the Run II muon detectors is shown in figure 3.13.

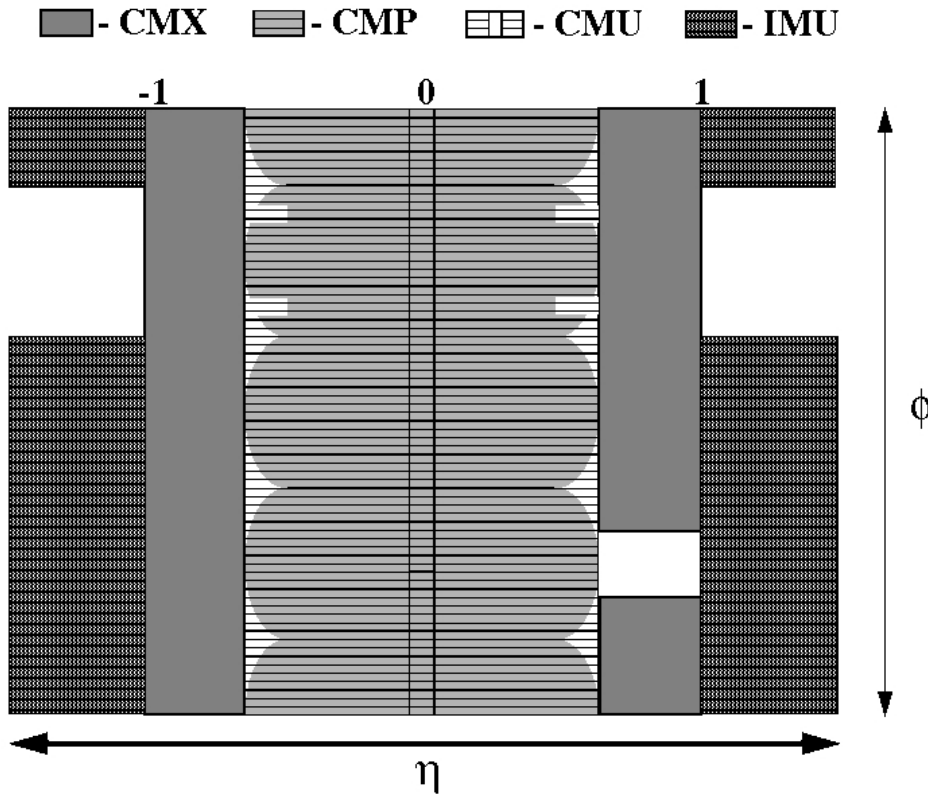


Figure 3.13: Coverage of muon detectors in the  $(\eta, \phi)$  space.

Muon detectors share common features. They consist of stack of rectangular drift chamber modules<sup>10</sup> composed of single-wire cells. Stacks are four layers deep with laterally displaced cells from layer to layer to compensate for cell inefficiencies. The difference in drift-electrons arrival-times between neighbor cells provides a typical resolution of  $250 \mu m$  for the hit position in the transverse plane. Charge division at the wire ends measures the  $z$  coordinate with a  $1.2 mm$  resolution.

Chambers are coupled with scintillator counters in order to suppress backgrounds due to secondary interactions in the detector and cosmic rays.

A muon candidate is reconstructed when a short track segment (*stub*) in the muon chambers match the extrapolation of a COT track.

The **CMU detector** is behind CHA at a radius of 347 cm from the beam axis and covers the  $|\eta| < 0.6$  region. CMU consist of 144 modules with 16 cells each. The **CMP detector** is arranged to enclose the  $|\eta| < 0.6$  region in an approximately central box (figure 3.13). Scintillator layers (CSP) on the outermost side of the CMP chambers allow identifying bunch crossing. CMU/CMP system is called CMUP. It detects muon with a minimum energy of  $\sim 3 GeV$ .

The **CMX detector** extends the muon identification in  $0.6 < |\eta| < 1$  region and is composed from different section of chambers. Two main *arches* were placed, since Run I, in both sides of the detector and cover the  $\phi$  region  $-45^\circ < \phi < 75^\circ$  and  $105^\circ < \phi < 225^\circ$ . The top *gap* in the west

<sup>10</sup>Chambers are filled with a mixture of Argon and ethane (50% each).

side ( $\eta < 0$ ) of the detector<sup>11</sup>, for  $75^\circ < \phi < 105^\circ$ , has been covered from Run II with a muon chamber system called *KeyStone*. As for CMP, cells of the two *arches* and of the *KeyStone* are sandwiched to scintillators (CSX). The bottom  $90^\circ$  gap of CMX penetrates the nominal floor of the collision hall but was filled in Run II with a slightly different muon chambers system<sup>12</sup>, called *MiniSkirt*, for  $-90^\circ < \phi < -45^\circ$  and  $225^\circ < \phi < 270^\circ$ . One additional layer of scintillator, MSX, is installed in the inner side of the *MiniSkirt*, and is read by photomultipliers at both ends. The forward region of muon system is the **IMU detector**, composed by the **BMU** muon chamber for  $1 < |\eta| < 1.5$  and  $-45^\circ < \phi < 225^\circ$  and the associated scintillator counters sub-systems called **BSU**(Barrel Scintillator Upgrade) and **TSU**(Toroid Scintillator Upgrade).

### 3.2.4 Cherenkov luminosity counters

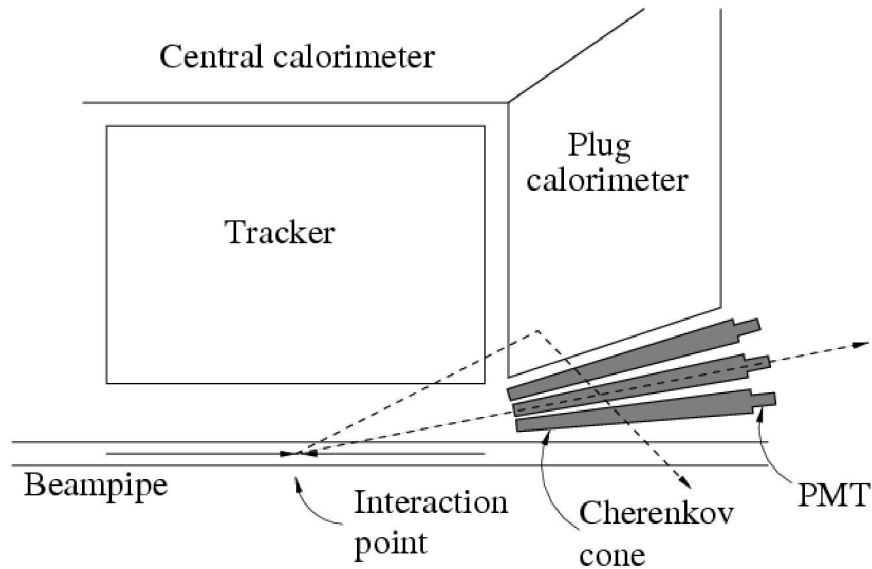


Figure 3.14: Schematic view of the luminosity monitor inside a quadrant of CDF.

CDF measures the collider luminosity with a coincidence between two arrays of Cherenkov counters, the CLC, placed around the beam pipes on the two detector sides [64], as shown in Figure 3.14. They are located inside the endplug calorimeters, in the forward and backward regions ( $3.7 < |\eta| < 4.7$ ). Each module consist of 48 thin, long, conical, gas filled Cherenkov counters. These counters are arranged around the beam pipe in three concentric layers with 16 counters each and pointing to the center of the interaction region.

The counters measure the average number of interactions per bunch crossing  $\mu$ , which is used to provide a measurement of the instantaneous luminosity  $\mathcal{L}$ :

$$\mu \cdot f_{bc} = \sigma_{p\bar{p}} \cdot \mathcal{L}, \quad (3.8)$$

where  $\sigma_{p\bar{p}}$  is the total  $p\bar{p}$  cross section at  $\sqrt{s} = 1.96$  TeV ( $\sigma_{p\bar{p}} = 60.7 \pm 2.4$  mb) and  $f_{bc}$  is the bunch crossing rate in the Tevatron. This system measures the luminosity with about the 6%

<sup>11</sup> The same coverage was not possible in the east side ( $\eta > 0$ ) because of the presence of cryogenic utilities servicing the solenoid.

<sup>12</sup> The design and chamber geometry for this section is similar but different from the upper part of the detector. The geometry is a plane of chambers arranged as a sector of a flattened cone.

systematic uncertainty.

### 3.2.5 Online Data Acquisition

At hadron collider experiments the collision rate is much higher than the rate at which data can be stored on tape. At CDF the predicted inelastic cross section for  $p\bar{p}$  scattering is  $60.7 \pm 2.4$  mb, which, considering an instantaneous luminosity of the order  $10^{32} \text{ cm}^{-2}\text{s}^{-1}$ , results in a collision rate of about 6 MHz, while the maximum rate at which events can be written on tape is only of the order of 100 Hz. The role of the trigger is to efficiently select the most interesting physics events. Events selected by the trigger system are saved permanently on a mass storage and subsequently fully reconstructed offline.

The CDF trigger system has a three-level architecture providing a rate reduction sufficient to allow more sophisticated event processing one level after another with a minimal deadtime (see Figure 3.15). The front-end electronics of all detectors is interfaced to a synchronous pipeline where up to 42 subsequent events can be stored for  $5.5 \mu\text{s}$  while the hardware is taking decision. Level 1 (L1) always occurs at a fixed time  $<4 \mu\text{s}$  so that it doesn't cause any dead time. Using a custom designed hardware, L1 makes a raw reconstruction of physical objects and takes a decision after counting them. Events passing the L1 trigger requirements are then moved to one of four on-board Level 2 (L2) buffers. Each separate L2 buffer is connected to a two-step pipeline, each step having a latency time of  $10 \mu\text{s}$ : in step one, single detector signals are analyzed, while in step two the combination of the outcome of step one are merged and trigger decisions are made. The data acquisition system allows a L2 trigger accept rate of  $\sim 1$  kHz and a L1 + L2 rejection factor of about 2500. Events satisfying both L1 and L2 requirements are transferred to the Level 3 (L3) trigger processor farm where they are reconstructed and filtered using the complete event information, with an accept rate  $< 150$  Hz and a rejection factor  $> 6$ , and then finally written to permanent storage.

According to the prerogatives of the events one would like to select, specific sets of requirements are established by exploiting the physics objects (*primitives*, such as raw tracks, calorimeter energy deposits, etc.) available for each trigger level. Then links across different levels are established by defining *trigger paths*, identified by a unique combination of a L1, a L2, and a L3 trigger; datasets (or data streams) are then finally formed by merging the data samples collected via different trigger paths.

Some trigger paths have output rates that exceed the maximum allowed value. To avoid the introduction of further selections that would bias the collected data sample such trigger paths are *prescaled* by a factor  $N$ : namely one event out of  $N$  that pass the selection is actually written on tape. When the prescale factor is fixed the prescale is called *static*, while it is called *dynamic* when it is varied during the data taking, according to the instantaneous luminosity, to exploit at the most the available bandwidth.

#### Level 1

##### Tracks

The most significant tool for L1 trigger is the possibility of track finding by means of a hardwired algorithm named eXtremely Fast Tracker (XFT). The XFT has been designed to work with COT signals at high collision rates, returning track  $p_T$  and  $\phi_0$  by means of a fast  $r - \phi$  reconstruction. These tracks are then extrapolated to the central calorimeter wedges and to the muon chambers (CMU and CMX), allowing the first electron or muon identification.

---

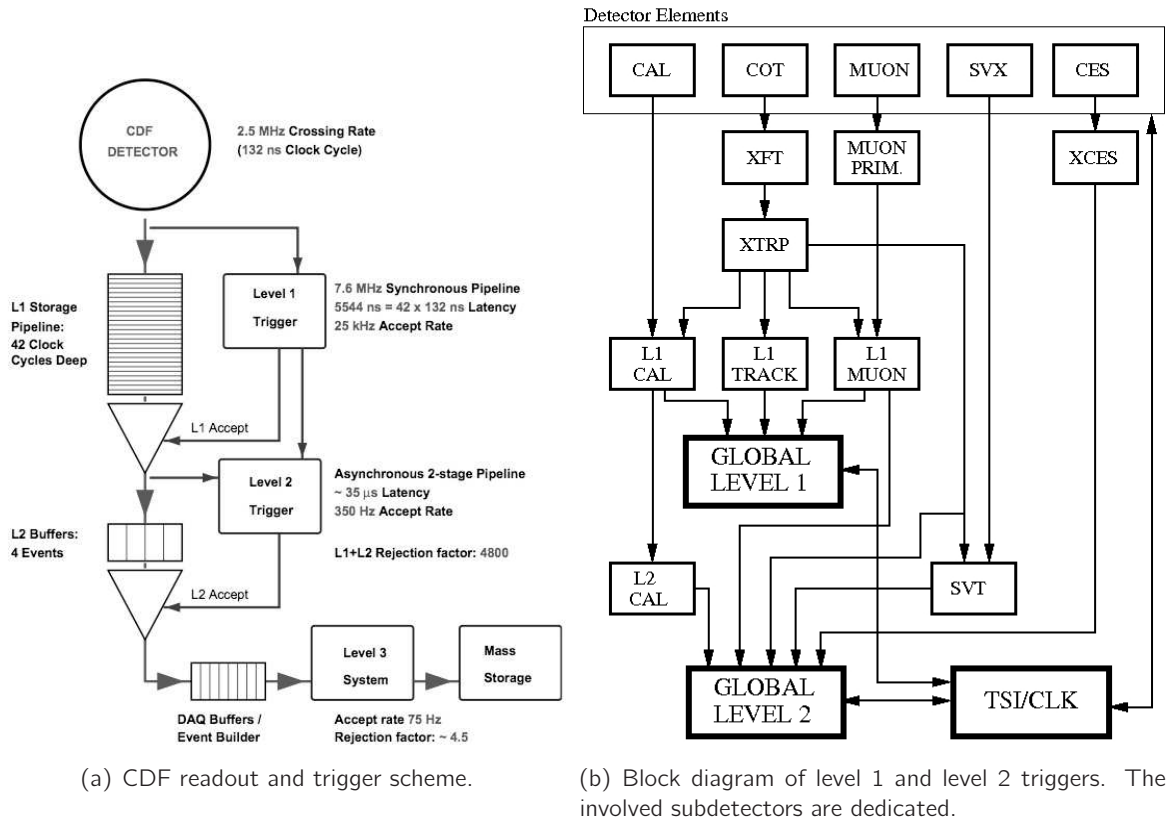


Figure 3.15: CDF trigger system.

### Calorimetric primitives

All L1 calorimetric towers are merged in pairs along  $\eta$  to define *trigger towers*, which are the basis for two types of primitives:

**object primitives:** electromagnetic and hadronic transverse energy contributions are used to define electron/photon and jet primitives respectively;

**global primitives:** transverse energy deposits in all trigger towers above 1 GeV are summed to compute the event total energy deposit in the calorimeter and its distribution in the transverse plane (e.g. the  $\cancel{E}_T$ <sup>13</sup>, described in Section 4.6).

Correspondingly, object and global triggers can be defined by applying a threshold to the respective primitives.

### Leptons

As already mentioned above, L1 muon and electron triggers are obtained by matching a XFT track to a corresponding primitive: for electrons, primitives are essentially the calorimetric trigger towers described above, while for muons they are obtained from clusters of hits in the muon chambers.

<sup>13</sup> The  $\cancel{E}_T$  represents the unbalance in the transverse energy distribution, ascribable to undetected particles.

## Level 2

L2 trigger takes a decision on a partially reconstructed event, exploiting data collected from L1 and from the calorimeter shower maximum detectors. Simultaneously a hardware cluster finder processes data from calorimeters while a track processor finds tracks in the silicon vertex detector.

### Calorimeter clusters

Since jets are expected not to be fully contained into a single calorimeter trigger tower, the energy threshold on L1 jet primitives must be set much lower than the typical jet energy in order to maintain high selection efficiency. As a consequence, jet trigger rates are too high to be fed directly into L3. An effective rate reduction can be obtained at L2 by triggering both on multiplicity and transverse energy of trigger tower clusters. The algorithm for cluster finding is based on a simple iterative algorithm clumping together neighbor towers [62]. L2 clusters can be used to build object triggers by applying a cut on their transverse energy and position (provided from  $\eta - \phi$  address of the seed towers), and global triggers by selecting on the number and  $\sum E_T$  of the clusters.

### SVT tracks

The Silicon Vertex Tracker (SVT) [65] exploits the potential of a high precision silicon vertex detector to trigger on tracks with a large impact parameter.

The architecture of SVT is shown in Figure 3.16. Its inputs are the list of axial COT tracks found by XFT and the data from SVXII. First SVXII hits are found by a Hit Finder algorithm and stored in hit buffers; then association between XFT and SVXII tracks is performed by *Associative Memory* (AM), a massive parallel mechanism based on the search of *roads* among the list of SVXII hits and XFT tracks; a road is a coincidence between hits on four out of five of the silicon layers and XFT tracks. Upon receiving a list of hits and tracks, each AM chip checks if all the components of one of its roads are present in the list of hits and XFT tracks. When AM has determined that a road might contain a track, hits belonging to that road are retrieved from the input buffer and passed to a track fitter to compute track parameters.

### Leptons

L2 muon primitives are almost the same exploited in L1, the only difference consists in an improved  $\phi$ -matching (within  $1.25^\circ$ ) between XFT tracks and track segments (*stubs*) formed with hits in the muon chambers. In the case of electrons, a finer  $\phi$ -matching can be instead performed at L2 thanks to the information from the central and plug shower maximum detectors.

## Level 3

The L3 trigger is a software trigger that runs on a Linux PC farm where all events are almost fully reconstructed using C++ codes and object-oriented techniques. In particular jets, COT tracks and leptons are identified. The algorithms used for the reconstruction are similar to the ones used in the offline analysis. Events coming from L2 are addressed to the *Event Builder* (EVB), which associates information on the same event from different detector parts. The final decision to accept an event is made on the basis of its features of interest (large  $E_T$  leptons, large missing  $E_T$ , large energy jets and a combination of such) for the physics process under study. Events exit L3 at a rate up to about 100 Hz and are permanently stored on tapes for further offline analysis. The size of each stored event is about 250 kB. Further offline processing is then performed on the selected events.

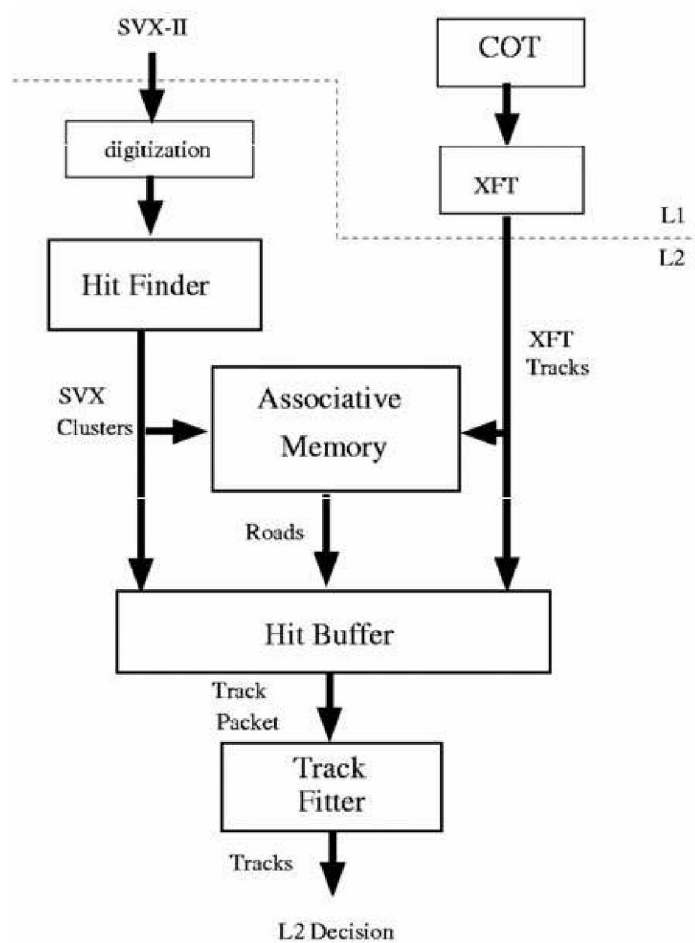


Figure 3.16: The SVT architecture.

### Trigger Upgrades

CDF has undergone two major trigger upgrades during Run II in order to deal with high trigger rates with increasing luminosity and to augment signal acceptance: an XFT upgrade and an upgrade in L2CAL system [66, 67].

XFT upgrades regards both Level 1 (L1) and Level 2 (L2) trigger systems. At L1 it rejects fake axial tracks by requiring the association with *stereo* segments, with a rejection factor of about 7. Moreover XFT segments of finer granularity can be sent to L2 where a 3D-track reconstruction can be performed with a good resolution on  $\cot\theta$  ( $\sigma_{\cot\theta} = 0.12$ ) and  $z_0$  ( $\sigma_{z_0} = 11$  cm).

The upgraded L2CAL system used a *fixed cone* cluster finding algorithm which prevents fake clusters formation and exploits full 10-bit trigger tower energy information for  $\vec{E}_T$  and  $\sum E_T$  calculation<sup>14</sup>. A jet is reconstructed starting from a *seed* tower above a 3 GeV threshold and adding all the towers inside a fixed cone centered at the seed tower and having a radius  $\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2} = 0.7$  units in the azimuth-pseudorapidity space. Jet position is calculated weighting each tower inside the cone according to its transverse energy. This upgrade has reduced L2 trigger rate and has provided at L2 jets energy and position measurements with nearly equivalent resolution to the offline one.

<sup>14</sup> The old system, due to hardware limitations, used only 8-bit tower information

### Offline data processing

The raw data flow from L3 triggers, segmented into streams according to triggers sets tuned to a specific physics process, is then stored on fast-access disks in real time (*online*), as the data are collected. All other manipulations of the collected data are referred to as *offline* data handling. The most important of these operations is the so-called *production*, which consists of the complete reconstruction of the collected data. At this stage the raw data are unpacked and the physics objects suitable for the analysis are reconstructed (such as tracks, vertexes, leptons), applying the most up-to-date detector calibrations available<sup>15</sup>. The output of the production is stored and further categorized into datasets which are used as input to physics analyses.

Homogeneous data are grouped during online acquisition in *run numbers*; during offline processing, events corresponding to several run numbers are grouped in *run periods*, each one with an integrated luminosity of the order of  $\sim 100 \text{ pb}^{-1}$ . Table 3.1 reports the conventional classification of the data acquired during the Tevatron Run II.

---

<sup>15</sup>At this stage, for example, are applied track reconstruction corrections accounting for the *real-time* beam position.

---



Period	Run numbers	Online dates	Luminosity [ $\text{pb}^{-1}$ ]	Cumulative luminosity[ $\text{pb}^{-1}$ ]
38	310472 - 312510	15 Aug 11 - 30 Sep 11	252	10121
37	308570 - 310441	04 Jul 11 - 15 Aug 11	174	9869
36	306791 - 308554	13 May 11 - 04 Jul 11	462	9695
35	304266 - 306762	06 Mar 11 - 13 May 11	364	9233
34	301952 - 303854	06 Jan 11 - 06 Mar 11	359	8869
33	299368 - 301303	01 Nov 10 - 24 Dec 10	357	8510
32	294778 - 299367	21 Aug 10 - 01 Nov 10	435	8153
31	293826 - 294777	20 Jun 10 - 17 Jul 10	172	7718
30	291294 - 293800	13 Apr 10 - 19 Jun 10	460	7546
29	289273 - 291025	26 Feb 10 - 13 Apr 10	360	7086
28	287294 - 289197	06 Jan 10 - 25 Feb 10	333	6726
27	284858 - 287261	25 Oct 09 - 05 Jan 10	422	6393
26	282976 - 284843	15 Sep 09 - 25 Oct 09	189	5971
25	275873 - 277511	05 May 09 - 13 Jun 09	283	5782
24	274123 - 275848	22 Mar 09 - 04 May 09	283	5546
23	272470 - 274055	15 Feb 09 - 21 Mar 09	232	5263
22	271072 - 272214	02 Jan 09 - 10 Feb 09	292	5031
21	268155 - 271047	12 Oct 08 - 01 Jan 09	520	4739
20	266528 - 267718	24 Aug 08 - 04 Oct 08	256	4219
19	264101 - 266513	01 Jul 08 - 24 Aug 08	287	3963
18	261119 - 264071	18 Apr 08 - 01 Jul 08	407	3676
17	258880 - 261005	28 Feb 08 - 16 Apr 08	188	3269
16	256840 - 258787	27 Jan 08 - 27 Feb 08	142	3081
15	254800 - 256824	05 Dec 07 - 27 Jan 08	159	2939
14	252836 - 254683	28 Oct 07 - 03 Dec 07	44.5	2780
13	241665 - 246231	13 May 07 - 04 Aug 07	317	2736
12	237845 - 241664	01 Apr 07 - 13 May 07	185	2419
11	233133 - 237795	31 Jan 07 - 30 Mar 07	264	2234
10	228664 - 233111	24 Nov 06 - 31 Jan 07	280	1970
9	222529 - 228596	01 Sep 06 - 22 Nov 06	180	1690
8	217990 - 222426	09 Jun 06 - 01 Sep 06	210	1510
7	210012 - 212133	14 Jan 06 - 22 Feb 06	50	1300
6	206990 - 210011	10 Nov 05 - 14 Jan 06	110	1250
5	203819 - 206989	05 Sep 05 - 09 Nov 05	135	1140
4	201350 - 203799	20 Jul 05 - 04 Sep 05	95	1005
3	198380 - 201349	21 May 05 - 19 Jul 05	100	910
2	195409 - 198379	19 Mar 05 - 20 May 05	130	810
1	190697 - 195408	07 Dec 04 - 18 Mar 05	130	680
0	138425 - 186598	04 Feb 02 - 22 Aug 04	550	550

Table 3.1: Data acquired by the CDF detector. The table shows conventional separation into run periods, and for each the online range information and luminosity acquired.



# Physical Object Reconstruction

## Contents

---

4.1	Track Reconstruction . . . . .	<b>48</b>
4.1.1	Tracking Algorithms . . . . .	48
4.2	Primary Vertex Identification . . . . .	<b>50</b>
4.3	Electron Identification . . . . .	<b>50</b>
4.4	Muon Identification . . . . .	<b>52</b>
4.4.1	Tracks Identification of unknown lepton flavor . . . . .	54
4.5	Jet Identification . . . . .	<b>55</b>
4.5.1	CDF Cone Algorithm . . . . .	57
4.5.2	Jet Correction . . . . .	58
4.6	Neutrino Identification: Energy Unbalance . . . . .	<b>58</b>
4.7	Trigger paths and efficiencies . . . . .	<b>59</b>
4.7.1	Trigger Efficiencies . . . . .	59

---

The offline data processing exploits several algorithms to reconstruct high- $p_T$  physics objects and observables quantities to be used in any data analysis. Basic algorithms are strongly connected with detector subsystems, like tracking algorithms that process information from the inner part of the detector to identify charged particles tracks and their fundamental kinematic parameters (described in Section 4.1). Parallel algorithms process data from the calorimeters to identify energy deposits from particles produced in the interaction (leptons or hadrons). Combining information from these two detector components it is possible to form higher-level objects, like hadronic jets (see Section 4.5) or identified particles (e.g. leptons) and evaluate event global properties (see Section 4.6).

The data analyses described in this thesis focus on reconstruction of diboson production in leptonic final states. This mainly relies on the identification of leptons (electrons and muons) in the detector. We therefore dedicate particular care to the optimization and increase of the detector acceptance in the lepton reconstruction. This is obtained exploiting the best from any part of the detector and optimizing the reconstruction according to the resolution of the detector subsystem involved. For simplicity we define several types of leptons reconstructed in the detector (*categories*), each one characterized by slightly different identification features. The identification procedures are described in Sections 4.3 and 4.4, while further details are reported in Appendix A.1.

In the following sections are described the techniques combined to reconstruct physical objects in the detector.

---

## 4.1 Track Reconstruction

The ability to detect and reconstruct charged particle trajectories is essential for particle identification and momentum reconstruction. Precise, high efficient tracking plays a central role for particle identification and separation.

At CDF the following five parameters are used to describe the helix trajectory of a charged particle in the magnetic field  $\vec{B}$  (see Figure 4.1):

- $C$ : the half-curvature ( $C = 1/2r$ , where  $r$  is the helix radius) of the trajectory, it has the same sign of the particle charge and it is related to the transverse momentum of the track:

$$p_T = \frac{B \cdot q}{2|C|} \quad (4.1)$$

being  $q$  the charge of the particle and  $B$  the magnetic field.

- $d_0$ : the impact parameter, i.e. the distance of the closest approach in the transverse plane between the helix and the beam line ( $z$ -axis), defined as:

$$|d_0| = \sqrt{x_0^2 + y_0^2} - r \quad (4.2)$$

where  $x_0$  and  $y_0$  are the coordinate of the track, in the transverse plane, in the point of closest approach to the beam line. With the *impact parameter significance*, defined as  $|d_0/\sigma_{d_0}|$ , is possible to estimate if the particle comes from primary vertex,  $d_0 \approx 0$ , or from a secondary one.

- $\lambda$ : the helix pitch, i.e. the cotangent of the polar angle between the track and the  $z$ -axis ( $\cot\theta_0$ ). The longitudinal component of the momentum is given by:

$$p_z = p_T \cdot \cot \theta_0 . \quad (4.3)$$

- $z_0$ : the position of the track vertex in  $z$ , defined as the interception between the track and the  $z$  axis in the transverse plane.
- $\phi_0$ : the azimuthal angle of the track at its vertex.

The Helix is completely described by these five parameters, in fact every point along the trajectory satisfy the following equation [68]:

$$x = r \sin \phi - (r - d_0) \sin \phi, \quad (4.4)$$

$$y = r \cos \phi - (r - d_0) \cos \phi, \quad (4.5)$$

$$z = z_0 + s\lambda. \quad (4.6)$$

where  $s$  is the track length projected in a transverse plane, and  $\phi = 2Cs + \phi_0$ .

### 4.1.1 Tracking Algorithms

A track pattern recognition algorithm searches among the several signals (*hits*) in the tracking system those that can be associated with the same track. Then a track fitting algorithm use those hits to reconstruct a track with its parameters. The experiment exploits several tracking algorithms [69], each optimized for the information available in different detector regions. In the

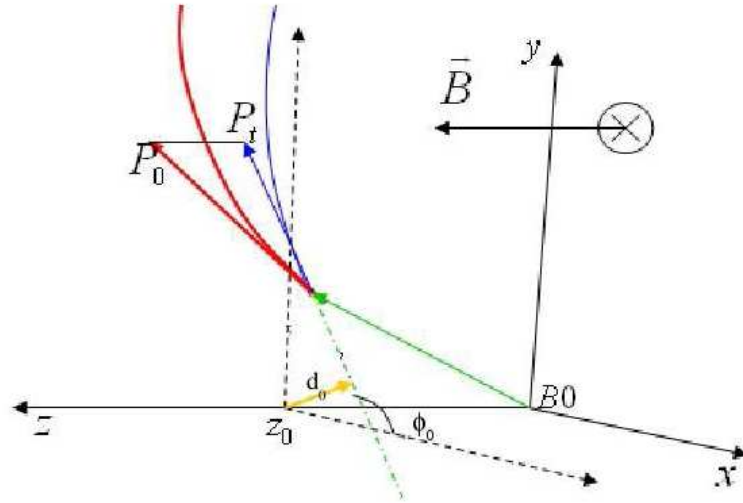


Figure 4.1: CDF track parameters and coordinate system.

following paragraphs we describe the main features of the three tracking algorithms most used: the Outside-In algorithm (OI), the Silicon-Stand-Alone (SiSA) algorithm and the Inside-Out (IO) algorithm.

### Outside-In Algorithm

The Outside-In is the most used CDF tracking algorithms and it is based on COT hits. Track pattern recognition starts in the COT outer layers, where the hit density is smaller, and proceeds through four steps: first each superlayer is searched for groups of three aligned hits and they are fit to a straight line with the least squares method. Then the tracks are reconstructed using information of the *axial* superlayers that are linked by two different algorithms (*segment linking* and *histogram linking* algorithms [69]). During the third step, the information of the stereo layers are added and the algorithm searches for the vertex of the track. As final step a global refit of the track is performed taking into account corrections for the non-uniformity of the magnetic field and for the modeling of electron drift.

At second stage of reconstruction, the track found in the COT is propagated into the silicon system. A road around a track is defined using the errors on the COT track parameters and silicon hits are added if they lie inside this predefined road. When a hit is added, the track parameters are recalculated and the search is performed again. The impact parameter resolution of COT+SVX tracks is found to be  $\sigma_{d_0} \simeq 20 \mu m$ .

### Silicon-Stand-Alone Algorithm

The hits in silicon subdetectors not used by OI tracking are available to the Silicon-Stand-Alone algorithm [69] to search for tracks in the region  $|\eta| < 2$  with few residual capability up to  $|\eta| \simeq 2.8$ .

The SiSA algorithm starts from a collection of at least four hits in the SVXII detector in the  $r - \phi$  plane and fits the  $C$ ,  $d_0$  and  $\lambda$  parameters to obtain a projection of the helix on the transverse plane. Then the algorithm creates a 3-D seed track adding small angle hits and the primary vertex information. At this point the  $90^\circ$  stereo hits are added and a global refit is performed.

SiSA tracks reconstructed only with SVXII have a poor resolution for high  $p_T$  tracks so hits are

searched in L00 and ISL with the SVXII track as seed. The track is refit if other layers can be added. However, the performances on momentum and impact parameter resolution are limited and indeed SiSA tracks are not used for secondary vertexing.

### Inside-Out Algorithm

The third tracking algorithm, the Inside-Out [70], tries to recover efficiency and  $p_T$  resolution in the region  $1.2 < |\eta| < 1.8$  where the COT coverage is limited. Practically SiSA tracks are used as seeds which are extrapolated to the COT inner cylinder. Matching hits in the COT are added, track is refitted and all duplicates are removed.

## 4.2 Primary Vertex Identification

The primary vertex is the position of the interaction point of a given event. The algorithm uses tracks information (`PrimVtx`): a seed vertex is calculated as the average  $z$  position of all tracks passing predefined quality requirements and is provided as input. Then all tracks with  $|z_0 - z_{vtx}| < 1 \text{ cm}$ ,  $|d_0| < 1 \text{ cm}$  and  $|d_0/\sigma_{d_0}| < 3$  are selected and ordered in decreasing  $p_T$ . They are fitted to a new vertex and the tracks with  $\chi^2 > 10$  are removed. The procedure is iterated until all accepted tracks have  $\chi^2 < 10$ . A quality index is assigned to the primary vertex depending on parameters like the number of final tracks.

## 4.3 Electron Identification

Electrons are reconstructed in the detector combining calorimetric and tracking information: the presence of a track pointing to a clustered energy deposit in the calorimeter is a clear signature of such candidate lepton. From the fit to the track is reconstructed the three-dimensional direction of the electron  $\vec{p}/|\vec{p}|$ , while the magnitude  $E \equiv |\vec{p}|$  is obtained from the calorimetric energy measurement. In this identification, the electron is assumed to be massless, which is a good approximation for leptons with momentum of the order considered. The  $E/p$  distribution is peaked at 1 but it has large radiative tails because the electron can radiate bremsstrahlung collinear photons in the passage through the tracking volume. The EM energy measurement is not much influenced by this effect since the photons generally deposits their energy in the same EM cluster, but the momentum measurement underestimate the original electron momentum.

The CDF EM clustering algorithm works in a simple but efficient way. The physical space corresponding to the calorimeter towers is mapped in an  $\eta - \phi$  plane, the algorithm creates two lists of the calorimeter towers ordered by decreasing energy revealed on them: the *usable list* (working towers with energy  $> 100 \text{ MeV}$ ) and the *seed list* (towers with energy  $> 2 \text{ GeV}$ ). It then takes the first seed tower and create an  $\eta - \phi$  cluster by adding the neighboring towers to form a  $2 \times 2$  or  $3 \times 3$   $\eta - \phi$  area.

As final step the  $\eta - \phi$  centroid of the cluster is calculated and the used towers are removed from the lists. The algorithm selects the next seed tower and iterate the process until all the seed towers have been used.

Usually  $3 \times 3$  clustering is used in the CEM region while  $2 \times 2$  clusters are used in the PEM region, this reduces the probability to overlap the clusters of two different electrons. A cluster is not allowed to cross the boundary between different subdetectors. Several corrections are applied to reconstruct the initial energy of the EM object. The clusters are corrected for lateral leakage, location inside the physical tower, on-line calibration and response curve drawn by test beam data.

Also the energy measured in the shower max (PES) and pre-shower (PPR) detectors is added to the final reconstructed energy. PES is also used to compare the shower profile of electrons or photons and it is used to measure the spatial position of the EM shower centroid.

Beyond the raw EM energy measurement, the calorimeter information can be further exploited for a better particle identification. The  $E_{Had}/E_{EM}$  ratio is used to identify electrons, in fact studies performed with candidate  $Z^0 \rightarrow e^+e^-$  events [71] show that electrons detected in the central or in the plug region have a little deposit in the hadronic part of the calorimeter (see Fig. 4.2) and cutting  $E_{Had}/E_{EM} < 0.125$  clean the  $e^\pm$  signal.

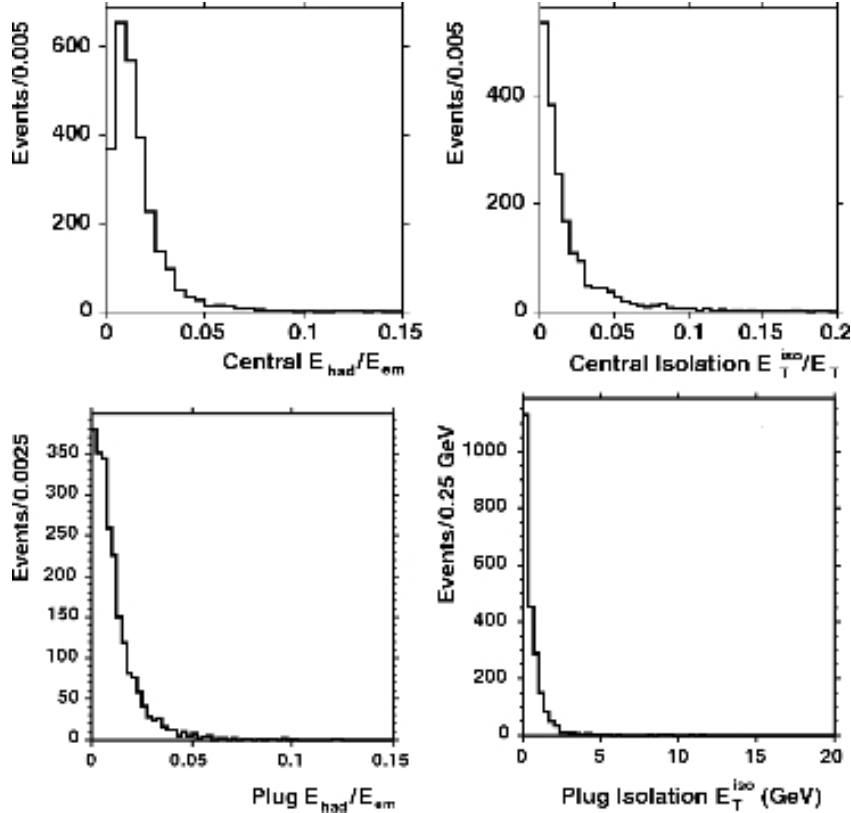


Figure 4.2:  $E_{Had}/E_{EM}$  (left) and isolation (right) distribution of central (top) and plug (bottom) calorimeter electron selection from unbiased, second legs of  $Z^0 \rightarrow e^+e^-$  candidate events in Data [71].

The correct identification of an electron is improved by considering the calorimetric isolation of the candidate energy cluster ( $E_T^C$ ), defined as:

$$\text{Cal Iso} \equiv \frac{E_T^{\Delta R=0.4} - E_T^{\text{cluster}}}{E_T^{\text{cluster}}}, \quad (4.7)$$

where  $E_T^{\Delta R=0.4}$  is the energy collected by the calorimeter within a radius  $\Delta R = 0.4$  from the centroid of the EM cluster. The calorimetric isolation is small for candidate electrons, with different resolution in the central and plug calorimeter (see Figure 4.2); electrons from boson decay are expected to be not correlated with jet directions.

Together with the electromagnetic cluster reconstruction we expect to find a matching track in the detector to complete electron reconstruction. Based on these two main elements is then possible to optimize peculiar selections to improve the purity of reconstructed electron sample,

exploiting calorimetric and tracking information from the different detector subsystems.

In this analyses we define four categories of reconstructed electrons, two relying on the central calorimeter ( $|\eta| \leq 1$ ) and two on the endplug calorimeters ( $1 \leq |\eta| \leq 2$ ). The details about these definitions are reported in Appendix A. In the central region of the detector we identify Tight Central Electrons (TCE) by applying strict requirements on the track quality and the *shape* of the electromagnetic energy cluster (details in Table A.2). We increase the central detector region acceptance by loosening the requirements on these calorimetric and tracking variables, using them to built a likelihood function for electron identification. We define Likelihood Based Electrons (LBE) (details in Section A.1.2) the objects with a high likelihood function output value ( $\mathcal{L} \geq 0.9$ ) that doesn't pass the requirements for TCE identification.

In the forward (and backward) region of the detector we define the Phoenix electrons (PHX) based on a silicon detector reconstructed track<sup>1</sup> matched to an EM cluster in the endplug calorimeters. The tight requirements applied to tracking and calorimetric information are summarized in Table A.4. Similarly to what is done in the central region, we reconstruct Plug Likelihood Based Electrons (PLBE) to increase the overall electron reconstruction efficiency in the forward (and backward) region.

Figure 4.3 shows the detector coverage for electron reconstruction, highlighting the different identified lepton categories.

## 4.4 Muon Identification

While electrons and hadrons lose all of their energy and stop in the calorimeter section of the detector muons at CDF behave like *minimum ionizing particles* (m.i.p.) and, since they leave just a very small amount of energy along their path, are the only particles that reach the outer part of the detector, where muon chambers are located. An algorithm fit the hits produced by the muon to a track segment which is called *stub*. The complete identification of a muon is then composed by three pieces: a charged-particle track reconstructed by the tracking algorithm that points to a detected stub in one of the muon chambers with an additional requirement of calorimetric energy deposit consistent with those of a minimum ionizing particle. A common set of requirements is applied to identify the candidate muons, summarized in Table A.5.

To increase the acceptance for the detector regions not covered by muon chambers, we can reconstruct *stubless* muon based only on a good quality track and m.i.p. requirement in the calorimeter. The purity of stubless muons sample is lower with respect to stubbed ones but it's still high enough for a lepton based analysis.

As described in Section 3.2.3 the muon detection system is composed by sub-detectors with different geometry and intrinsic resolution. We therefore define different muon categories, characterized by geometrical position inside the different muon chambers.

In the central part of the detector ( $|\eta| \leq 0.6$ ) candidate muons can have a stub in either the CMU or CMP chambers, which have different geometries and overlap for their major surface. A central muon that satisfies requirements described in Table A.5 and have corresponding stubs both in CMU and CMP chambers is categorized as CMUP muon. Those that have matching stubs in CMU(CMP) chambers but pointing to spatial region not covered by the CMP(CMU) muon chambers are identified as CMU and CMP only muons respectively (see Tables A.6 and A.7 for details).

In the intermediate region ( $0.6 \leq |\eta| \leq 1.0$ ) we identify as CMX muons those pointing to the arches of the intermediate muon detector, which are required to entirely cross the tracking system

<sup>1</sup> The COT tracking system does not cover this  $\eta$  region.



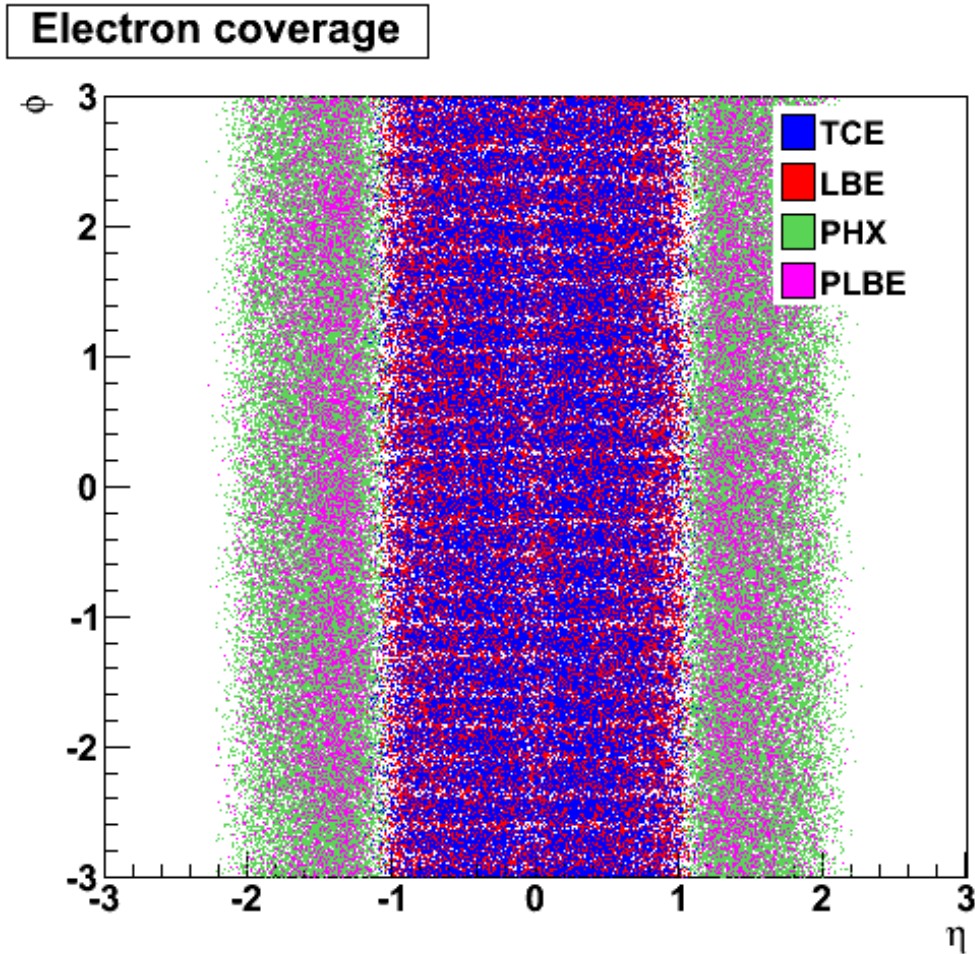


Figure 4.3: Distribution of different electron types in  $\eta - \phi$  plane.

(both silicon and COT tracker), with a stub in that muon chambers, and small energy deposit in the central calorimeter (details in Table A.9). Similarly is defined the CMX-MsKs category, for muons pointing to the CMS Miniskirts and Key-Stone muon chambers.

The forward (and backward) muons, defined as BMU, rely on a good-quality track matched to a stub in the IMU detector, satisfying a m.i.p. requirement in the endplug calorimeters (details in Table A.10).

Besides the six stubbed muon categories defined above the muon identification efficiency is extended defining two categories of *stubless* muons, based only on tracks and calorimeter information: CMIOCES muons, fiducial to the central calorimeter and CMIOPEs muons, fiducial to the endplug calorimeter. The tight requirements applied to identify these two categories are summarized in Tables A.11 and A.12.

Additional *spare* tracks pointing to uninstrumented part of the detector can represent a real muon, passing none of the previously described requirement, and will be promoted to be a lepton, as described in Section 4.4.1.

Figure 4.4 shows the  $\eta - \phi$  distribution of the different muon categories.

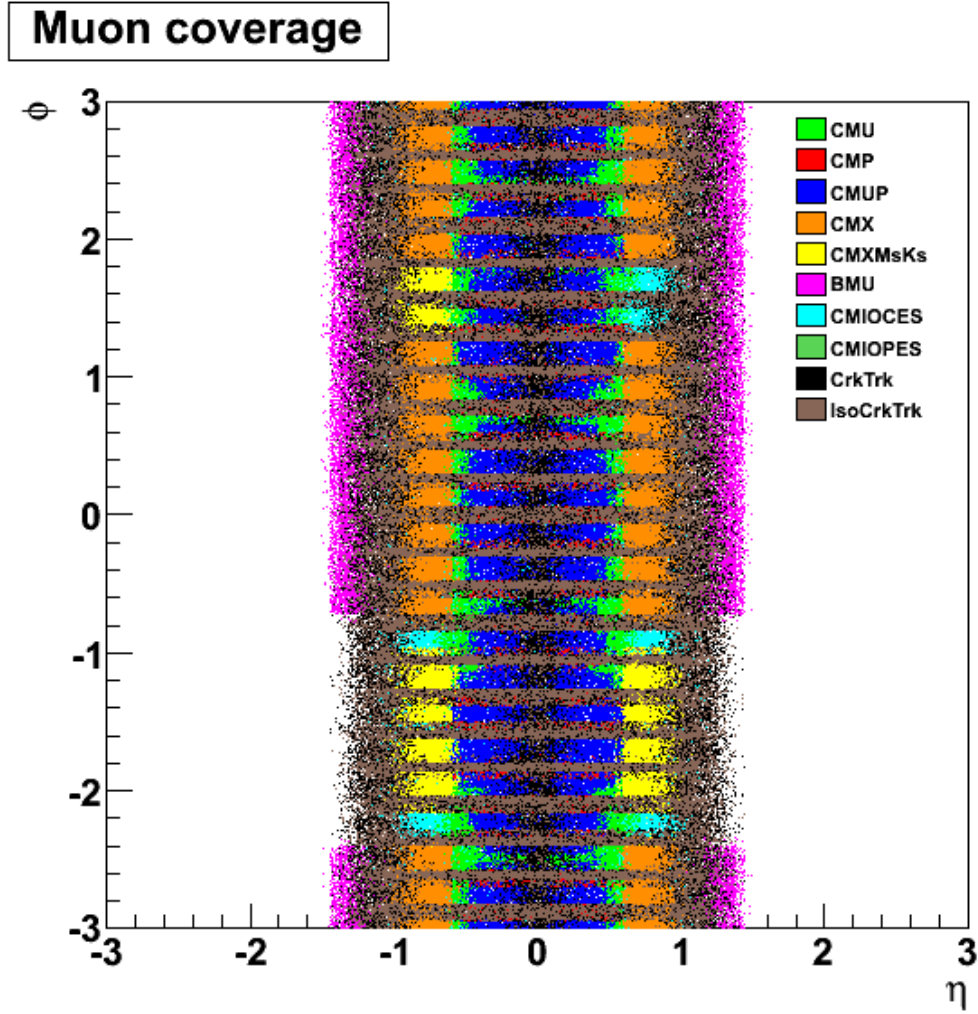


Figure 4.4: Distribution of different muon types in  $\eta - \phi$  plane.

#### 4.4.1 Tracks Identification of unknown lepton flavor

In order to recover some of the acceptance lost due to uninstrumented regions or gaps in the calorimeters, high- $P_T$  tracks which enter such *cracks* are counted in this analysis in a separate category (CrkTrk). These CrkTrk objects are predominantly electrons and muons that were otherwise lost. The definition of CrkTrk requires a well measured track which specifically points to a crack in the calorimeter as well as little calorimeter or track activity near the primary track. The identification requirements for CrkTrk leptons are given in Table A.13. In this case the Isolation requirement is based both on other muons present and on electromagnetic cluster in the closest calorimeter towers. In addition are used more strict selection on Stereo reconstruction of the track and on the  $\chi^2/dof$  of the track fit. Obviously are considered as CrkTrk only tracks that have not passed any other lepton identification selection.

To reduce the rate of spurious track reconstruction as CrkTrk an additional isolation criterion of the track is required, given by

$$\frac{\sum_{i \neq \text{seed}}^N P_T^i}{P_T^{\text{seed}}} < 0.1 \quad (4.8)$$

where N is the number of tracks within a cone of  $\Delta R < 0.4$  around the candidate track. This

requirement is very important because this category does not have reliable calorimeter information for additional lepton identification. Since these tracks are not expected to leave any large fraction of their energy in the calorimeter they are treated as muons for  $\cancel{E}_T$  corrections discussed in Section 4.6.

## 4.5 Jet Identification

QCD tells us that the parton composing the (anti)proton can be treated perturbatively as free particle if they are struck by an external probe<sup>2</sup> with sufficient high energy (*hard scattering*). High- $p_T$  partons resulting from the interaction cannot exist as free particles because at longer distances the *strong potential* can not be treated perturbatively and partons must form colorless hadrons. This process is called *hadronization* or *showering* and produces a collimated cluster of stable particles named *jet*. A jet approximately retains the total momentum and direction of the initial parton (see Fig. 4.5).

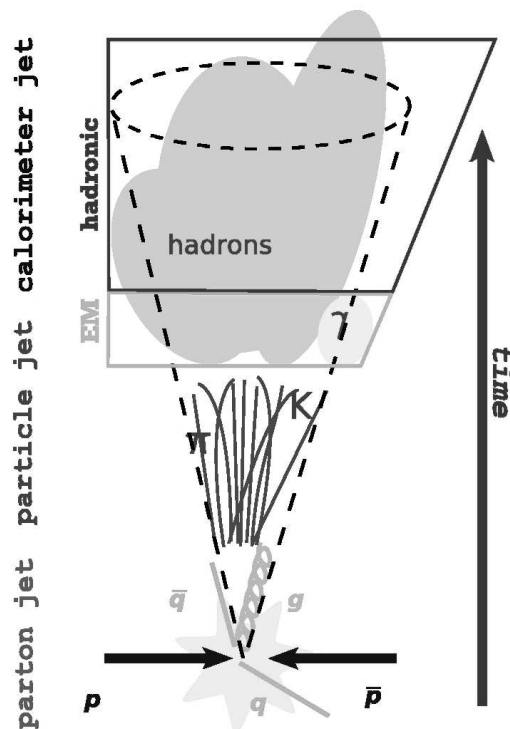


Figure 4.5: A parton originated from a hard scattering hadronizes and gives origin to a collimated spray of particles, a jet.

From the experimental point of view a jet is defined as a large energy deposit in a localized area of the detector (see Fig. 4.6). The challenge of a physics analysis is to recover from the detector information the initial energy, momentum and, possibly, the kind of parton produced in the original interaction. A “jet algorithm” is a tool to reconstruct such information and it must satisfy at best the following requirements [72]:

<sup>2</sup> I.e. a lepton or a parton from another hadron.

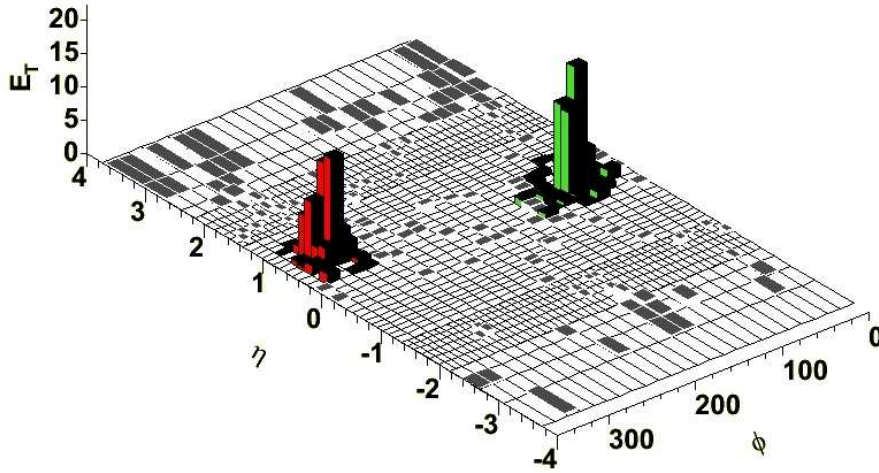


Figure 4.6: Calorimeter deposit in the  $\eta - \phi$  plane as represented in the CDF event display. EM deposits are green (light-colored cluster on the right) while HAD deposits are red (darker-colored cluster on the left).

- *Infrared safety*: the presence of soft radiation between two jets may cause a merging of the two jets. This should not occur to avoid an uncorrected parton attribution.
- *Collinear safety*: the jet reconstruction should be independent from any collinear radiation in the event, i.e. different energy distribution of particles inside calorimetric towers.
- *Invariance under boost*: the same jets should be found independently from boosts in longitudinal direction.
- *Boundary stability*: kinematic variables should be independent from the details of the final state.
- *Order independence*: the same reconstructed quantities should appear looking at parton, particle and detector levels.
- *Straightforward implementation*: algorithm should be easy to implement in perturbative calculations.

Beyond these theoretical aspects a jet algorithm should be characterized by a high reconstruction efficiency, good resolution and stable at different luminosity.

Even though this analysis considers the fully leptonic decay of  $Z$  boson the knowledge of jets reconstruction is relevant since the main background studied is due to the presence of jets misreconstructed as leptons.

In this analysis a jet is defined as a calorimeter cluster of size  $\Delta R < 0.4$  which has a total corrected transverse energy of  $E_T > 15$  GeV and that is within a pseudo-rapidity of  $|\eta| < 2.5$ . Identified

electrons will always satisfy the jet requirements and for this reason an object is not counted as a jet if it is within a  $\Delta R = 0.4$  cone with respect to an identified electron.

### 4.5.1 CDF Cone Algorithm

CDF uses several algorithms, none of them really satisfying all the above requirements. The most common one is JETCLU[73], an iterative fixed cone jet reconstruction algorithm based only on calorimetric information.

The algorithm begins by creating a list of seed towers from all the calorimeter towers with transverse energy above the threshold of 1 GeV. Starting with the highest- $E_T$  seed tower, a precluster is formed by combining together all adjacent seed towers within a cone of given radius  $R$ <sup>3</sup>. This procedure is repeated, starting with the next unused seed tower, until the list is exhausted. The  $E_T$ -weighted centroid is then formed from the towers in the precluster and a new cone of radius  $R$  is formed around this centroid. All towers with energy above the lower threshold of 100 MeV within this new cone are added to the cluster. Then, a new centroid is calculated from the set of towers within the cluster and a new cone drawn. This process is iterated until the centroid of the energy deposition within the cone is aligned with geometric axis of the cone (*stable solution*).

Since each tower can belong to only one jet, in case of jet overlap two clusters are merged if the total energy of the overlapping towers is greater than 75% of the energy of the smaller cluster. If the shared energy is below this cut, the shared towers are assigned to the cluster that is closer in  $\eta - \phi$  space. This process is iterated again until the list of clusters remains fixed.

Massless four-vector momenta are assigned to the towers in the clusters for EM and HAD components with a magnitude equal to the energy deposited in the tower and the direction defined by a unit vector pointing from the event vertex to the center of the calorimeter tower at depth that corresponds to the shower maximum. A cluster four-vector is then defined summing over the towers in the cluster:

$$E = \sum_{i=1}^N (E_i^{EM} + E_i^{HAD}) \quad (4.9)$$

$$p_x = \sum_{i=1}^N (E_i^{EM} \sin \theta_i^{EM} \cos \phi_i^{EM} + E_i^{HAD} \sin \theta_i^{HAD} \cos \phi_i^{HAD}) \quad (4.10)$$

$$p_y = \sum_{i=1}^N (E_i^{EM} \sin \theta_i^{EM} \sin \phi_i^{EM} + E_i^{HAD} \sin \theta_i^{HAD} \sin \phi_i^{HAD}) \quad (4.11)$$

$$p_z = \sum_{i=1}^N (E_i^{EM} \cos \theta_i^{EM} + E_i^{HAD} \cos \theta_i^{HAD}) \quad (4.12)$$

where the index  $i$  runs over the towers in the cluster. Other variables are added to the final jet-object used in the analysis:  $E_T$ ,  $\eta$  and  $\phi$  (calculated from the jet vertex with an energy weighted average over the calorimeter towers associated with the cluster) or other useful information like the number of tracks reconstructed inside the jet cone, the vertex quality or the energy deposited in the *HAD* or *EM* calorimeter.

<sup>3</sup> CDF reconstructs jets using radii 0.4, 0.7 and 1.0.

### 4.5.2 Jet Correction

The ultimate goal of the jet reconstruction algorithm is the best determination of the energy of the outgoing partons coming from the hard interaction. Clearly many factors produce a mismatch between the raw energy produced by the experimental algorithm and the one of the partons before the hadronization.

CDF developed a set of jet energy correction depending of  $\eta$ ,  $E_T^{raw}$  and  $R$  of the jet reconstructed by JETCLU algorithm. The corrections are divided into five levels<sup>4</sup> (“L-levels”) so that can be applied in a standard way to different analysis [74]:  $\eta$ -dependent response of the calorimeter ( $L1$ ), effect of multiple interactions ( $L4$ ), absolute energy scale ( $L5$ ), underlying event ( $L6$ ) and out-of-cone ( $L7$ ) corrections. The correction  $L1$  and  $L5$  are multiplicative factors ( $f_{L1}$  and  $f_{L5}$ ) on the raw  $E_T$  of the jet, the others are additive constants ( $A_{L4}$ ,  $A_{L6}$  and  $A_{L7}$ ). The general equation to apply all corrections is:

$$E_T^{corr}(\eta, E_T^{raw}, R) = (E_T^{raw} f_{L1} - A_{L4}) f_{L5} - A_{L6} + A_{L7}. \quad (4.13)$$

A more detailed description of the different level algorithm can be found in [75] and will not be discussed here.

## 4.6 Neutrino Identification: Energy Unbalance

Momentum conservation is the only way to reveal the presence of neutrinos since they do not interact inside the detector components. Although it is impossible to know the exact momentum of the colliding partons, the transverse component of the system of the two partons,  $p_T$ , is approximately zero in the detector frame. All the detected transverse energies are vectorially summed, and if the sum is greater than zero, we assume the presence of one (or more) candidate neutrino is revealed. The missing transverse energy  $\cancel{E}_T$  gives a measure of the neutrino transverse momentum,

$$\vec{\cancel{E}}_T \equiv - \sum_i \vec{E}_T^i \quad (4.14)$$

where  $\vec{E}_T^i$  is a vector with magnitude equal to the transverse energy collected by the  $i$ -th calorimeter tower and pointing from the interaction vertex to the center of the tower. The sum involves all the calorimetric towers with total energy above 0.1 GeV in the region  $|\eta| < 3.6$ . At offline level, the algorithm corrects for the position of the reconstructed event vertex and for any reconstructed muon (their energy is calculated using track information).

The  $\cancel{E}_T$  used to identify neutrino has to be corrected for several effects. The largest correction is due to muons which are minimum ionizing particles and do not leave much energy in the calorimeter causing an apparent missing energy as most of their energy is carried away as they leave the detector. The  $\cancel{E}_T$  is corrected for muons identified according to section 4.4 by adding back their track momentum measurement and subtracting any small amount of energy which they may have deposited in the calorimeters. High- $P_T$  tracks, supposed to be leptons that fall in an uninstrumented part of the detector, are treated as muons in this calculation since they enter a crack in the calorimeter and did not release their energy in it. The  $\cancel{E}_T$  is also modified to account for the corrections to raw jet energies discussed in section 4.5. The  $\cancel{E}_T$  used at the analysis level

<sup>4</sup> The actual naming skips  $L2$ , because it is absorbed in  $L1$ , and  $L3$ , as it was introduced as a temporary MC calibration in Run II.

is then

$$-\vec{E}_T = \sum_i \vec{E}_T^i + \sum_\mu \vec{P}_T^\mu - \sum_\mu \vec{E}_T^\mu (Em + Had) + \sum_j \vec{E}_T^{j\,corr} - \vec{E}_T^{j\,raw} \quad (4.15)$$

where the  $E_T$  have been corrected for the actual interaction  $z_0$  point since the transverse components are calculated according to  $\sin\theta$  and vertex away from  $z = 0$  would give a different  $\theta$ .

Beyond the neutrinos created in weak interactions there are also several sources of false  $\vec{E}_T$  which are often difficult to control. These sources include the mismeasurement of jet and lepton energies as well as when a lepton or a photon enters a crack in the detector where it would not be possible to reconstruct its energy with any reasonable accuracy.

## 4.7 Trigger paths and efficiencies

Any analysis performed at a hadron collider starts by collecting a sample of interesting events out of the huge amount of collisions. Due to the technological restriction in data acquisition, described in Section 3.2.5, a fundamental element is the online trigger system.

For the analyses described in this thesis we use a three level realtime decision system based on the reconstruction of single high- $p_T$  leptons in the CDF detector. At each level the electron and muon reconstruction parameter resolution increases and a trigger *path* is followed to select events containing electrons or muons. To increase the amount of interesting collected data, we select events using different high- $p_T$  trigger paths. The different electron and muon trigger paths considered are listed in the following:

- ELECTRON\_CENTRAL\_18, selects events with at least one central electron with  $E_T \geq 18$  GeV;
- MUON\_CMUP18, requires one central muon with hits in both the CMU and CMP chambers;
- MUON\_CMX18, requires one central muon with hits in the CMX chambers;
- MUON\_CMP18\_PHIGAP, gathers events with at least one central muon with hits in the CMP detector only, pointing to regions not covered by the CMU chambers;
- MUON\_CMU18\_ETAGAP, gathers events with at least one central muon with hits in the CMU detector only, pointing to regions not covered by the CMP chambers.

In order to study and evaluate specific backgrounds we use data collected by jet-triggers: *JET20*, *JET50*, *JET70*, and *JET100*. These trigger paths require at least one jet to be reconstructed at trigger level with different energy thresholds: 20, 50, 70, and 100 GeV. The detailed requirements at trigger level for all the trigger paths used in this Thesis are described in Appendix B.

### 4.7.1 Trigger Efficiencies

Trigger efficiencies are measured for each lepton type. We define a *triggerable lepton*, i.e. a lepton that could have triggered the event, starting from the lepton categories defined in Appendix A.1. Electrons and stubbed muons are triggerable leptons if  $E_T > 20$  GeV, to avoid significant turn-on effects, while stubless muons and Crack Tracks are not considered triggerable leptons. Each triggerable lepton can be triggered by only one trigger path. Table 4.1 summarizes the triggerable

Lepton category	Trigger path
LBE	ELECTRON_CENTRAL_18
CMUP	MUON_CMUP18
CMP	MUON_CMP18_PHIGAP
CMU	MUON_CMU18_ETAGAP
CMX	MUON_CMX18
CMX-MsKs	MUON_CMX18

Table 4.1: Triggerable lepton types with the associated trigger paths.

lepton types and the associated trigger paths. These objects will constitute the sample used to evaluate the lepton identification efficiencies.

The offline requirements made to select the triggerable leptons are tighter than the trigger requirements. This means that triggerable leptons should pass the trigger requirements, except for the fact that the offline variables have better resolution respect to the online ones.

The trigger efficiencies are evaluated for each run period, to take into account differences in trigger requirements and detector performances. From period 14 on (see Table 3.1), the detector and trigger conditions are stable and we merge the single period measurements in a single efficiency.

Electron trigger efficiencies are evaluated selecting  $W \rightarrow e\nu$  events requiring an identified triggerable electron and significant missing transverse energy. We separately measure efficiencies of calorimetric and tracking requirements using data collected by backup triggers, and multiply them to have the result. To measure muon trigger efficiencies we select  $Z \rightarrow \mu\mu$  data events, requiring the invariant mass of the two identified muons to be  $76 \leq m_{\ell\ell} \leq 106 \text{ GeV}/c^2$ . Both muons are required to be triggerable leptons. One is used as a tag and we require that the corresponding trigger path selected the event online. The number of selected events is the denominator of the trigger efficiency. The numerator is the number of the events for which also the trigger path corresponding to the other muon selected the event.

The statistical uncertainty on trigger efficiencies is found to be smaller than 1%. Appendix B.1 describes the detailed calculations and result for electron and muon trigger efficiencies for each category.

Based on the number and the type of reconstructed triggerable leptons in the events we assign a per-event trigger efficiency, that will be described in detail in Section 5.1.2. Data events are always required to have at least one reconstructed triggerable lepton compatible with the online trigger path that selected the event.



# ZZ reconstruction in leptonic final states

## Contents

5.1	$ZZ \rightarrow \ell\ell'\ell'$	<b>63</b>
5.1.1	Data Collection and Offline Selection	63
5.1.2	Monte Carlo Simulations	64
5.1.3	Data Driven Background Estimation	68
5.1.4	Event Selection	73
5.1.5	Systematic Uncertainties	77
5.1.6	Cross section measurement	78
5.2	$ZZ \rightarrow \ell\nu\nu$	<b>80</b>
5.2.1	Initial Sample Composition	81
5.2.2	Kinematic Properties and Signal Selection	83
5.2.3	Test of the Background Modeling in a Control Region	92
5.2.4	Artificial Neural Network for Signal Extraction	98
5.2.5	Systematic Uncertainties	100
5.2.6	Cross Section Measurement	107

The measurement of the  $ZZ$  production cross section and the comparison with the best available Standard Model (SM) predictions is the main goal of the analysis described in this chapter. This measurement will also allow to test the presence of new physics beyond the SM involving diboson production, that will be discussed in Chapter 6. As we already mentioned in Section 2.2.2, the SM predicts  $ZZ$  production with a cross section of  $\sigma(p\bar{p} \rightarrow ZZ) = 1.4 \pm 0.1$  pb. We will use the full dataset collected from the CDF detector improving the previous measurements done at the Tevatron. We measure this process cross section in the two most favorable decay modes:  $\ell^+\ell^-\ell'^+\ell'^-$  (described in Section 5.1) and  $\ell^+\ell^-\nu\bar{\nu}$  (described in Section 5.2). In Section 5.1 we will describe common elements shared by both channels. In Section 5.1.1 is described how is collected the sample of data used to perform both the analyses (the  $\ell^+\ell^-\ell'^+\ell'^-$  and  $\ell^+\ell^-\nu\bar{\nu}$  final states) while in Section 5.1.2 is reported a description of the Monte Carlo simulation used to model the physics processes investigated. Section 5.1.3 describes the data driven method to extract a prediction of the background contamination to the four charged lepton events, reconstructed as described in Section 5.1.4. Section 5.1.5 summarizes the systematic uncertainties considered for the cross section measurement, obtained with the technique reported in Section 5.1.6.

Section 5.2.1 describe the composition of the data sample used to measure the  $ZZ$  production in the  $\ell^+\ell^-\nu\bar{\nu}$  decay mode, while the peculiar optimization of the kinematic requirement to select the  $ZZ$  signal is described in Section 5.2.2. The test of the background processes modeling is

reported in Section 5.2.3; Section 5.2.4 describes the Multivariate Analysis Technique exploited to extract the investigated signal from the large background contribution. Section 5.2.5 describes the systematic uncertainties considered in the cross section fit performed in Section 5.2.6. The combination of the measurement done in the two decay modes will be reported in Chapter ??.

### Process Properties

According to the branching ratios of the  $Z$  boson [3],  $ZZ$  production can be detected in several leptonic or hadronic final states, summarized with the relative strength in Figure 5.1. The largest decay modes involve hadrons in the final state but the most interesting channels to study this process are the fully leptonic decay modes:  $lll'l'$  and  $ll\nu\nu$ . Given the high efficiency and precision in lepton reconstruction these are the most powerful modes for a cross section measurement.

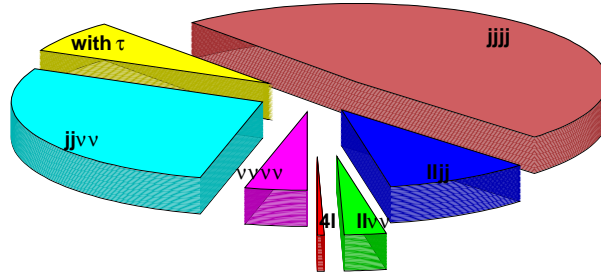


Figure 5.1: Sketch of the relative  $ZZ$  fraction for the different decay modes.

Even if just a small amount of  $ZZ$  decays in the  $lll'l'$  final state ( $\sim 0.5\%$ ) is expected, this process is really clean to extract a diboson signal since a negligible amount of background events can give a similar signature in the detector. No other Leading Order (LO) SM process<sup>1</sup> gives four high- $p_T$  leptons in the final state, similar to the ones produced in the  $ZZ$  decay<sup>2</sup>. The maximum precision achievable in a cross section measurement is limited though by the small statistics of the available sample.

In addition to the four charged lepton decay mode, the  $ll\nu\nu$  decay mode, thanks to a larger branching fraction ( $\sim 3.5\%$ ), gives access to a larger sample of produced  $ZZ$  events. On the other hand, the presence of one  $Z \rightarrow \nu\nu$  prevent from the full reconstruction of the event kinematic, limiting the knowledge of part of the process to global event properties. Several SM processes give two charged leptons in the final state, hence a similar signature in the detector. In particular the single  $Z \rightarrow ll$  production, as well as other diboson processes, lead to an overwhelming background to  $ZZ$  identification in this decay mode.

<sup>1</sup> SM predicts production of four lepton final state through second order correction diagrams, therefore largely suppressed with respect to  $ZZ$  production.

<sup>2</sup> Other SM processes can produce four massive leptons in the final state which are not isolated, resulting for example from the decay of  $b$  quarks. These leptons are usually partially contained within the hadronic jet produced and doesn't satisfy the isolation requirements applied for lepton isolation. These are not the ones we would like to identify as prompt leptons from the boson decay, hence will always be treated as backgrounds.

## 5.1 $ZZ \rightarrow \ell\ell'\ell'$

Since this measurement relies on the identification of four charged leptons the maximization of the lepton detector acceptance (described in Sections 4.3 and 4.4) plays a key role in the sensitivity of this measurement. Reconstruction of electrons and muons produced in the decay gives a fully determined event kinematic, giving information about mass and momentum of each of the two produced  $Z$ .

### 5.1.1 Data Collection and Offline Selection

In this Thesis we analyze data collected during the whole Tevatron Run II (March 2002 - September 2011<sup>3</sup>) by the CDF detector, corresponding to about  $9.7 \text{ fb}^{-1}$  of integrated luminosity. The  $ZZ$  production and decay modes considered in this Thesis involve the production of one or more leptons ( $e, \mu$ ) in the final state, hence we collect experimental data using single high- $p_T$  lepton trigger paths described in Section 4.7 and in detail in Appendix B. We exploit one trigger that reconstructs central electrons with  $E_T \geq 18 \text{ GeV}$  (ELECTRON\_CENTRAL\_18) and four triggers that reconstruct central (MUON\_CMUP18, MUON\_CMP18\_PHIGAP, MUON\_CMU18\_ETAGAP) and intermediate (MUON\_CMX18) muons with  $p_T \geq 18 \text{ GeV}/c$ . All these triggers have an efficiency  $>90\%$ , providing a clean sample of events containing at least one lepton.

The data collected are inserted into different *good run lists* according to the settings of the detector during each data taking period (e.g. sub-detector systems partially deactivated). Each event is required to have a run number in one of these lists, according to the prerogatives of the reconstructed leptons in the event<sup>4</sup>. Table 5.1 reports the integrated luminosities of the samples corresponding to each good run list.

Good run list	$\mathcal{L}(\text{pb}^{-1})$
EM_NOSI	9735.1
EM_CMUP_NOSI	9695.9
EM_MU_NOSI_CMXIGNORED	9653.3
EM_SI	9446.5
EM_CMUP_SI	9409.2
EM_MU_SI_CMXIGNORED	9370.4

Table 5.1: Good run lists considered for different detector conditions with the corresponding integrated luminosity. The first three good run lists doesn't have any requirement on the silicon detector conditions, while the latter three explicitly requires the proper working of the silicon tracking subdetector.

The first good run list (EM\_NOSI) has a minimal set of requirements on the proper operation of the calorimeter and other fundamental parts of the detector (e.g. the COT for track reconstruction) hence collects the largest integrated luminosity and includes all the others. The second and third ones require in addition the proper functioning of the muon chambers, the third ignoring the intermediate muon chamber (CMX) conditions. The last three lists have similar requirements but ask in addition that the Silicon detector works properly. Each event is assigned to the most tight good run list it belongs to, relying on the best detector conditions.  $ZZ$  candidate events are

<sup>3</sup> Details in Table 3.1.

<sup>4</sup> For example, events containing only reconstructed electrons doesn't care about the data-taking conditions of the muon chambers.

reconstructed through their decay to four fully identified charged leptons, i.e. electrons, muons. We only have a marginal acceptance on the decay of one  $Z$  to a pair of  $\tau$  leptons from their subsequent decay to electrons or muons<sup>5</sup>. We consider in the analysis events that contain exactly four isolated leptons in the final state, belonging to one of the lepton categories described in Appendix A.1. Each of these leptons is required to have  $p_T \geq 10$  GeV/ $c$  since this condition guarantees a very good efficiency and purity ( $\sim 100\%$ ) in real lepton identification and reconstruction. We apply offline a requirement that at least one of the four leptons reconstructed in the events is the one that fired the trigger<sup>6</sup>. Even if the considered trigger paths are based on an offline (L3 trigger) requirement of a minimum transverse momentum greater than 18 GeV/ $c$ , the *trigger* lepton in this analysis is required to have  $p_T \geq 20$  GeV/ $c$ . Above this threshold the trigger efficiency reaches its plateau, and can be considered independent from the  $p_T$  of the lepton, ignoring the turn-on profile. The more relevant background in this final state is due to mis-identification of jets or photon as leptons in the detector. Different SM processes (i.e. Drell-Yan, dibosons,  $t\bar{t}$ ) can produce two or three leptons in the final state, in association with hadronic activity or photon production. We expect the dominant contribution to this instrumental background coming from  $Z$  production (Drell-Yan) in association with jets,  $Z\gamma$ +jets, while we expect other processes (e.g.  $t\bar{t}$ ) to give a negligible contribution. In these processes, the misidentification as leptons of additional objects in the detector can mimic the final state of the  $ZZ$  investigated signal. Hence it is fundamental both to have the highest possible efficiency in lepton identification together with the highest purity, to keep the background contribution in this sample as lower as possible.

We optimize this analysis with the help of Monte Carlo simulations of the processes considered, in particular the  $ZZ$  production, to have predictions of the kinematic properties of the signal process we're looking for and background candidate events, as described in the following Section.

### 5.1.2 Monte Carlo Simulations

The production of a pair of  $Z$  bosons and their subsequent decay and interactions with the detector and the production of different processes contributing to the sample selected is reproduced using Monte Carlo simulation. This technique is used also for the modeling of the signal and background processes that contribute to the  $\ell\ell\nu\nu$  decay mode discussed in Section 5.2 and we will discuss here also common aspects that will be exploited later.

The full Monte Carlo simulation of the physics process considered is based on:

- the simulation of the hard parton scattering and the soft remnant interactions;
- the detector simulation, that reproduces the interaction of the produced particles with the CDF detector, correcting for the effect of the online data taking conditions;
- the simulation of the trigger and analysis requirements applied.

**Main Physics Process** The hard scattering physics of the different processes contributing to the sample is simulated using PYTHIA[76] which provides a Leading Order (LO) description of the hard scattering between partons inside the (anti-)protons. PYTHIA is also used to describe

<sup>5</sup> The  $\tau$  has in general different identification and reconstruction algorithm that considers its hadronic decay products. We don't consider it in this analysis and from now on when we mention a *lepton* we mean only electrons and muons. In particular,  $\mathcal{BR}(\tau \rightarrow \ell\bar{\nu}_\ell\nu_\tau)$  is  $\sim 17\%$ ; this will be further reduced by the offline requests on the charged leptons described later on, which are optimized for  $Z \rightarrow e^+e^-$ ,  $Z \rightarrow \mu^+\mu^-$  decay.

<sup>6</sup> This is practically done imposing that the lepton belongs to a category which satisfies the requirement of a given trigger path.

the parton showering, based on a Leading-Logarithm resummation algorithm, for each process, which reproduces initial and final state radiation, and the hadronization of color-charged particles. For a more realistic modeling of the physics process reproduced, the simulations include both the effects due to multiple interaction in the same  $p\bar{p}$  bunch crossing, as well as the interaction of the collision remnants with the detector (underlying event).

Since these simulations require a big amount of computing time, especially for the detector simulation, we often apply preliminary requirements at generator level, i.e. the presence of a minimum of two leptons produced in the collision, and kinematic lower cut-off on their transverse momentum. Efficiencies for these requirements are properly taken into account in the normalization to the data of the simulated sample.

$p\bar{p} \rightarrow ZZ$  production is fully simulated using PYTHIA and includes all the possible decay modes of the  $Z$  boson pair. In this simulation we set the internal PYTHIA parameters according to a configuration which is commonly referred to as Tune A [77], obtained from a fine tuning of the MC simulation parameters on the first inverse femtobarn of CDF collected data. These samples are filtered at generator level for the presence of at least two leptons with  $p_T > 1$  GeV/c.

The simulations are scaled from the generated cross section and branching fractions, to the most recent available calculations, ranging from NLO to Next-to-Next-to-Leading Order (NNLO) accuracy or better.

**Detector Simulation** The event generator produces as output a set of particles with their 4-momenta and the position where these have been created and decayed. We then use GEANT3 [78] to fully simulate the interactions of these particles with the CDF detector.

Data taking conditions are considered into the detector simulation, including time-dependent beam position and the operating conditions of the detector components. Instantaneous luminosity profile is part of the simulation in order to model at best the multiple interactions in the same  $p\bar{p}$  collision. Simulated events are generated to reproduce data collected in a given data-taking period. The samples are then scaled to reproduce the complete dataset considered in the analyses.

**Trigger and Analysis selections** The output of the detector simulation has the same format as the real data so that every selection and reconstruction algorithm can be applied evenly to the data sample and the simulated events.

In the simulated samples we reconstruct as leptons only objects that correspond to a generated lepton. The contribution from particles misidentified as leptons is evaluated separately (see Section 5.1.3). For this reason, in the MC simulation we match each particle identified as lepton to a true particle (electron, muon) at generator level requiring  $\Delta R$  between the direction of the reconstructed lepton momentum and the one of the generated particle to be less than 0.1. The same correspondence between reconstructed and generated particle in the MC simulation is required for photons produced in the interactions.

We assign to simulated event a weight ( $w$ ) so that, for a given process, the weighted sum of the simulated events reproduces the expected number of events in the considered sample of data. This weight takes into account the process cross section, the number of simulated events, as well as different per-event correction factor to the MC simulation.

In the collision simulation the generated  $z$  of the primary vertex is required to be within  $\pm 60$  cm from the center of the detector. The event weight takes into account a correction factor  $\epsilon_{vtx}$  measured in minimum bias events as a function of the run period, which is on average  $\epsilon_{vtx} = 0.9555 \pm 0.0031$  with  $< 2\%$  differences from one period to another.

The *total* trigger efficiency is evaluated for every event as the probability for that event to be

triggered as a function of the number and type of reconstructed triggerable leptons in the event. If only one triggerable lepton is found, then the total trigger efficiency for the event is the one for that lepton. If more than one triggerable lepton is found (e.g. 2), the total trigger efficiency for the event is evaluated based on the different lepton trigger efficiencies. The probability that at least one of the leptons present has fired the trigger is given by  $\epsilon^{trg} = 1 - \sum_{\ell} (1 - \epsilon_{\ell})$ , where  $\epsilon_{\ell}$  are those for each lepton. The prescale that is eventually applied to the corresponding trigger path is taken into account in this calculation. For data-taking periods during which were simultaneously present trigger paths based on the  $r - \phi$  only tracking information (2D) or the complete  $r - \phi - z$  tracking (3D) the trigger efficiency calculation properly handle the overlap between the two<sup>7</sup>.

To consider differences in lepton identification (ID) between collected data and simulated events, we apply to the event weight a scale factor  $s^{lep}$  measured in  $Z \rightarrow \ell\ell$  data samples, described in detail in Section 5.1.2.

All these corrections are taken into account in the weight assigned to each event of the simulated samples. To each event belonging to the  $j$ -th good run lists defined in Table 5.1 is given a weight  $w_j$  defined as following:

$$w_j = \sigma \times \mathcal{B} \times \mathcal{L}_j \times \frac{\epsilon_{\text{filter}} \times \epsilon^{\text{trg}} \times s^{\text{lep}} \times \epsilon_{\text{vtx}}}{N_j^{\text{gen}}(|z_0^{\text{P.V.}}| < 60 \text{ cm})} \quad (5.1)$$

where

$\sigma \times \mathcal{B}$  is the cross section times branching fraction of the physics process simulated

$\epsilon_{\text{filter}}$  is the efficiency of the filter applied to the simulation at generator level

$\epsilon^{\text{trg}}$  is the total *per-event* trigger efficiency

$s^{\text{lep}}$  is the total *per-event* lepton identification scale factor (see Section 5.1.2)

$\epsilon_{\text{vtx}}$  is the efficiency of the z-vertex position requirement ( $|z_0^{\text{P.V.}}| < 60 \text{ cm}$ )

$\mathcal{L}_j$  is the luminosity of the  $j$ -th good run list corresponding to the reconstructed leptons

$N_j^{\text{gen}}(|z_0^{\text{P.V.}}| < 60 \text{ cm})$  is the number of generated events with  $|z_0^{\text{P.V.}}| < 60 \text{ cm}$  for the  $j$ -th good run list corresponding to the reconstructed leptons.

### Lepton Identification Efficiencies and Scale Factors

We measure the lepton identification efficiency for each lepton category using a *Tag and Probe* method applied to  $Z/\gamma^* \rightarrow \ell\ell$  events. In this two-leg events one lepton is required to pass strict selections (*tag*) while the second one is used as a *probe* of the identification efficiency. Real electrons and muons are obtained selecting a sample of  $Z/\gamma^* \rightarrow \ell\ell$  events collected with a the same set of high- $p_T$  (high- $E_T$ ) trigger paths described in Section 5.1.1. We select events with one fully identified lepton (*tag*) and a second object, *probe*, that passes looser and well simulated lepton requirements<sup>8</sup>. The *tag* and the *probe* leptons are required to have opposite charge, an invariant mass within  $\pm 15 \text{ GeV}/c^2$  from the  $Z$  peak ( $76 \leq m(\ell\ell) \leq 106 \text{ GeV}/c^2$ ) and  $\cancel{E}_T < 25 \text{ GeV}$ . Given one tagged *leg* the lepton identification efficiency is evaluated considering the second *leg*

<sup>7</sup> Further details in Appendix B.1

<sup>8</sup> *Probe* definitions are reported in details in Appendix A.2.

in the event, by looking at the full lepton identification of the second one, as depicted in Figure 5.2. The identification efficiency is then defined as:

$$\epsilon' = \frac{N_{\ell\ell}}{N_{\ell p}}. \quad (5.2)$$

where  $N_{\ell\ell}$  is the number of events with two tightly identified leptons and  $N_{\ell p}$  is the number of events with one tight lepton and one probe. The residual background due to non-Drell-Yan leptons is estimated from the invariant mass sidebands  $m(\ell\ell) \in [61,76] \cup [106,121]$  GeV/ $c^2$ , then subtracted from  $N_{\ell\ell}$  and  $N_{\ell p}$ . Figure 5.3 compares, as example, the dilepton invariant mass distribution for real and simulated data with two central muons and shows the definition of the Z signal and sidebands regions.

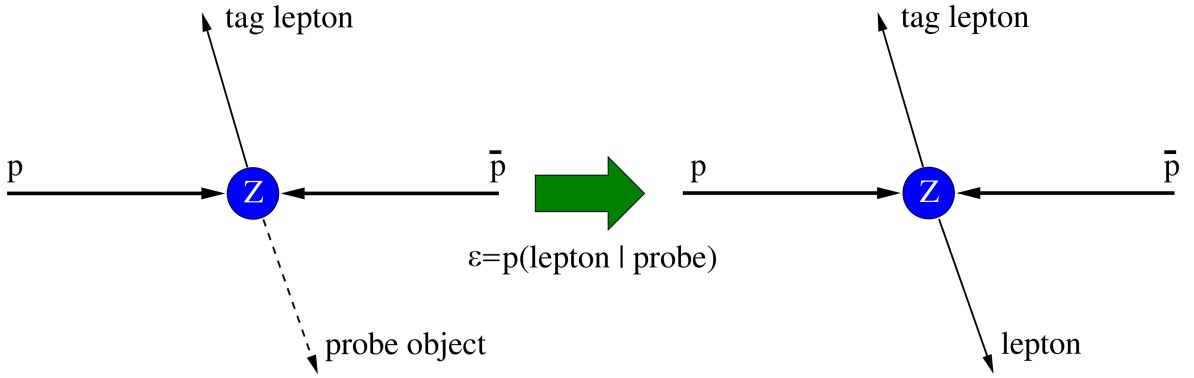


Figure 5.2: Schematic representation of the *tag and probe* method.

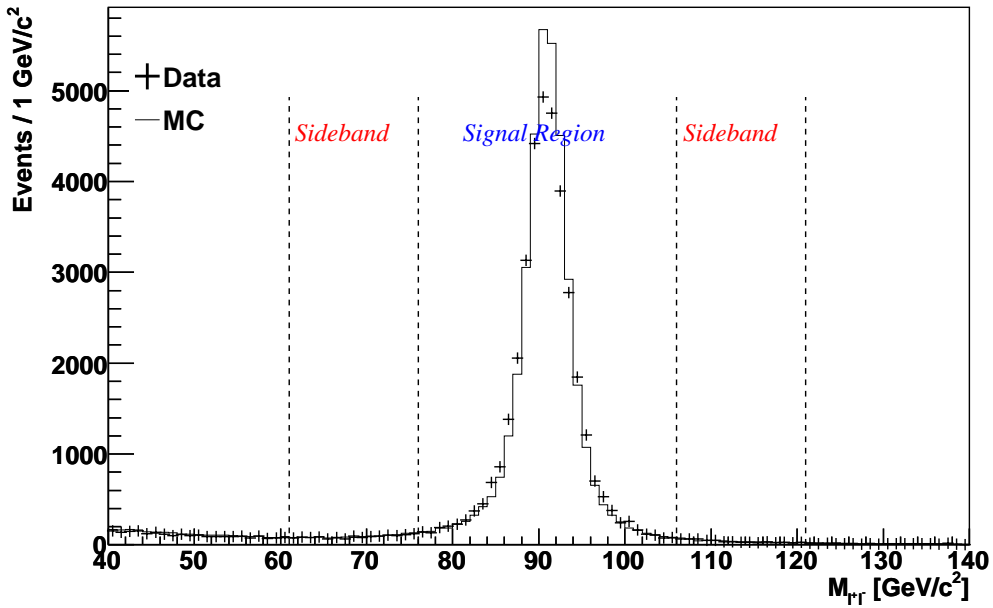


Figure 5.3: Invariant mass distribution of central muon pairs, comparing data and MC simulation.

The probe for each lepton category is based on a common definition of a generic central

or forward electron/muon probe objects, then requiring to pass through a fiducial region of the detector. The selections applied in the definition of the probes are described in Appendix A.2. For forward electrons we separately measure the efficiency for calorimetric (PHXPEM) and tracking (PHXTRK) requirements. The efficiency for identification of Crack Tracks (see Section 4.4.1) is measured separately assuming these are electrons and muons.

The lepton ID efficiencies are evaluated for data and MC simulation separately and for the different run periods. We include in MC simulations event weight (Equation 5.1) a corrective scale factor  $s^{\text{lep}}$  to compensate differences in lepton identification between collected data and simulations. This scale factor is given by the ratio between the ID efficiencies in data ( $\epsilon^{\text{Data}}$ ) and the ones in simulated events ( $\epsilon^{\text{MC}}$ ), namely  $s^{\text{lep}} = \epsilon^{\text{Data}}/\epsilon^{\text{MC}}$ . When we apply the Crack Track correction scale factors to simulated events we use either one or the other depending if the lepton matches and electron or a muon at generator level.

An example of the efficiency measurements and scale factors for run period 14 to 23 is reported in Table 5.2, while those for other run periods are reported in Appendix A.3. The scale factors differ from one run period to another by less than 10%.

Run Period: 22 - 38			
	data	MC	Scale Fac
LBE	$0.122 \pm 0.001$	$0.102 \pm 0.001$	$1.199 \pm 0.012$
TCE	$0.752 \pm 0.001$	$0.787 \pm 0.001$	$0.956 \pm 0.002$
PHXTrk	$0.886 \pm 0.001$	$0.864 \pm 0.001$	$1.026 \pm 0.002$
PHXPEM	$0.769 \pm 0.002$	$0.851 \pm 0.001$	$0.904 \pm 0.002$
PLBEPem	$0.133 \pm 0.001$	$0.095 \pm 0.001$	$1.408 \pm 0.017$
CrkTrk (e)	$0.603 \pm 0.002$	$0.741 \pm 0.001$	$0.814 \pm 0.003$
CMUP	$0.746 \pm 0.003$	$0.877 \pm 0.001$	$0.850 \pm 0.004$
CMU	$0.767 \pm 0.006$	$0.878 \pm 0.003$	$0.874 \pm 0.008$
CMP	$0.752 \pm 0.007$	$0.902 \pm 0.002$	$0.834 \pm 0.008$
CMX	$0.853 \pm 0.004$	$0.898 \pm 0.002$	$0.950 \pm 0.005$
CMXMsKs	$0.772 \pm 0.009$	$0.890 \pm 0.003$	$0.867 \pm 0.010$
BMU	$0.727 \pm 0.007$	$0.695 \pm 0.004$	$1.047 \pm 0.012$
CMIOCES	$0.208 \pm 0.002$	$0.145 \pm 0.001$	$1.434 \pm 0.015$
CMIOPEs	$0.538 \pm 0.005$	$0.575 \pm 0.002$	$0.936 \pm 0.009$
CrkTrk ( $\mu$ )	$0.543 \pm 0.004$	$0.564 \pm 0.002$	$0.962 \pm 0.007$

Table 5.2: Lepton identification efficiencies for data and simulation for periods 22 to 38. Errors are statistical only.

From the  $\epsilon^{\text{Data}}$  and  $\epsilon^{\text{MC}}$  calculation we observe a good agreement between data and MC<sup>9</sup> for central electron identification (less than 2%), while forward electron calorimeter identification efficiency in simulation is overestimated by  $\sim 9\%$ . The central stubbed muons efficiencies tend to be overestimated in our simulation while for central stubless category we observe a deviation in the opposite direction. The isolated track scale factors are similar for electrons and muons and present a slight overestimation of the simulation with respect to data.

### 5.1.3 Data Driven Background Estimation

We expect the four-lepton collected sample to be dominated by ZZ events, but we expect a small contribution from events in which one or more reconstructed leptons are in fact fake lepton

<sup>9</sup> This is indicated by a value of the scale factor  $\sim 1$ .



identifications, which will be considered a background to this measurement. This contribution is evaluated using a data driven method, evaluating the probability of a given object to be identified as a lepton, since we don't expect a MC simulation to properly reproduce the detector particle reconstruction response at this level of precision.

This method relies on few assumptions about the sample used for the jet-to-lepton misidentification rate and the one where those probabilities are applied. One important difference can be due to a different contribution of *heavy flavor* quarks in the samples. In the jet sample where we calculate the misidentification probabilities the contribution from heavy flavor can be estimated comparing the cross section of the processes  $p\bar{p} \rightarrow q\bar{q}$  and  $p\bar{p} \rightarrow b\bar{b}$ . A coarse estimate of this, for jets with  $p_T \sim 50$  GeV/c gives  $\sigma(p\bar{p} \rightarrow b\bar{b})/\sigma(p\bar{p} \rightarrow q\bar{q}) \sim 10^{-2}$ . In the selected sample used for the analysis we expect events with fake leptons coming mainly from single  $Z \rightarrow \ell\ell$  production with additional hadronic activity. For this process the relative contribution of heavy flavor is  $\sigma(Z + b - jet(s))/\sigma(Z + jet(s)) \sim 5 \cdot 10^{-3}$  and the main contribution of  $Z + b$  comes from the process  $ZZ \rightarrow \ell^+ \ell^- b\bar{b}$ . This gives a contribution of  $b$ -quarks in the sample used for background estimation below 1%, hence the data driven estimate will be reliable at the level of precision of  $\sim 1\%$ .

To evaluate the jet-to-lepton misidentification probability we define a *denominator* category for each lepton category by loosening some of the corresponding identification selections and vetoing the full identification of such object as a lepton. This defines a *sideband* in the lepton identification variable space. For likelihood based reconstructed electrons (see Section A.1.2) this is represented by a track with an associated energy deposit in the calorimeter, consistent with being an electron with looser isolation requirement with respect to the standard ones. For muon and isolated track corresponding *denominator* categories different sets of requirements are defined for the central ( $|\eta| \leq 1.0$ ) and forward ( $|\eta| \geq 1.0$ ) region, because of the structure of the calorimeters and the inner tracking system. For the denominator object corresponding to a given muon category, additional requirements are applied according to the involved muon detector region, similar to what is done for trigger efficiency calculation. The details of the selection requirements in the definition of the denominator objects are reported in Appendix A.2. The explicit veto assures that the sample used in this procedure is completely orthogonal to the one that contains only fully identified leptons.

In the jet-enriched sample we define the probability for a denominator object of being reconstructed as a lepton. This is the ratio of the number of fully identified leptons to the number of reconstructed denominator objects, for each lepton category. This probability is referred to as *fake rate* and is evaluated as function of the  $E_T(p_T)^{10}$  of the denominator object. The measurement of this ratio is done separately on data sample collected by JET20, JET50, JET70, JET100 trigger paths<sup>11</sup>, ignoring in the calculation the leading jet in each event; the weighted average in the four is considered as central value while the maximum spread among the various measurements is taken as systematic uncertainty. Figures 5.4 and 5.5 show the *fake rates* for the different lepton categories. We correct this calculation for the contribution of real leptons in the jet dominated sample, expected to be of the order of 1% of the total sample collected by the jet triggers, subtracting the expected real  $Z$  and  $W$  contribution using the prediction from a MC simulation<sup>12</sup>.

From the measurement we can see that the several JET samples covers different regions in

<sup>10</sup> The fake rate is evaluated as function of the  $E_T$  of the denominator object for the electron categories and as a function of the  $p_T$  of the denominator object for the muon categories. For simplicity from now on we will solve this ambiguity simply mentioning the denominator object  $p_T$ .

<sup>11</sup> The different JET-X paths correspond to different jet  $E_T$  threshold requirement of the trigger, details in Section 4.7 and Appendix B.

<sup>12</sup> As we already mentioned in Section 5.1.1 we aim to identify as leptons only those coming from bosons decay, hence leptons produced in heavy flavor meson decay are treated, together with the jet energy deposit, as a misidentified lepton.

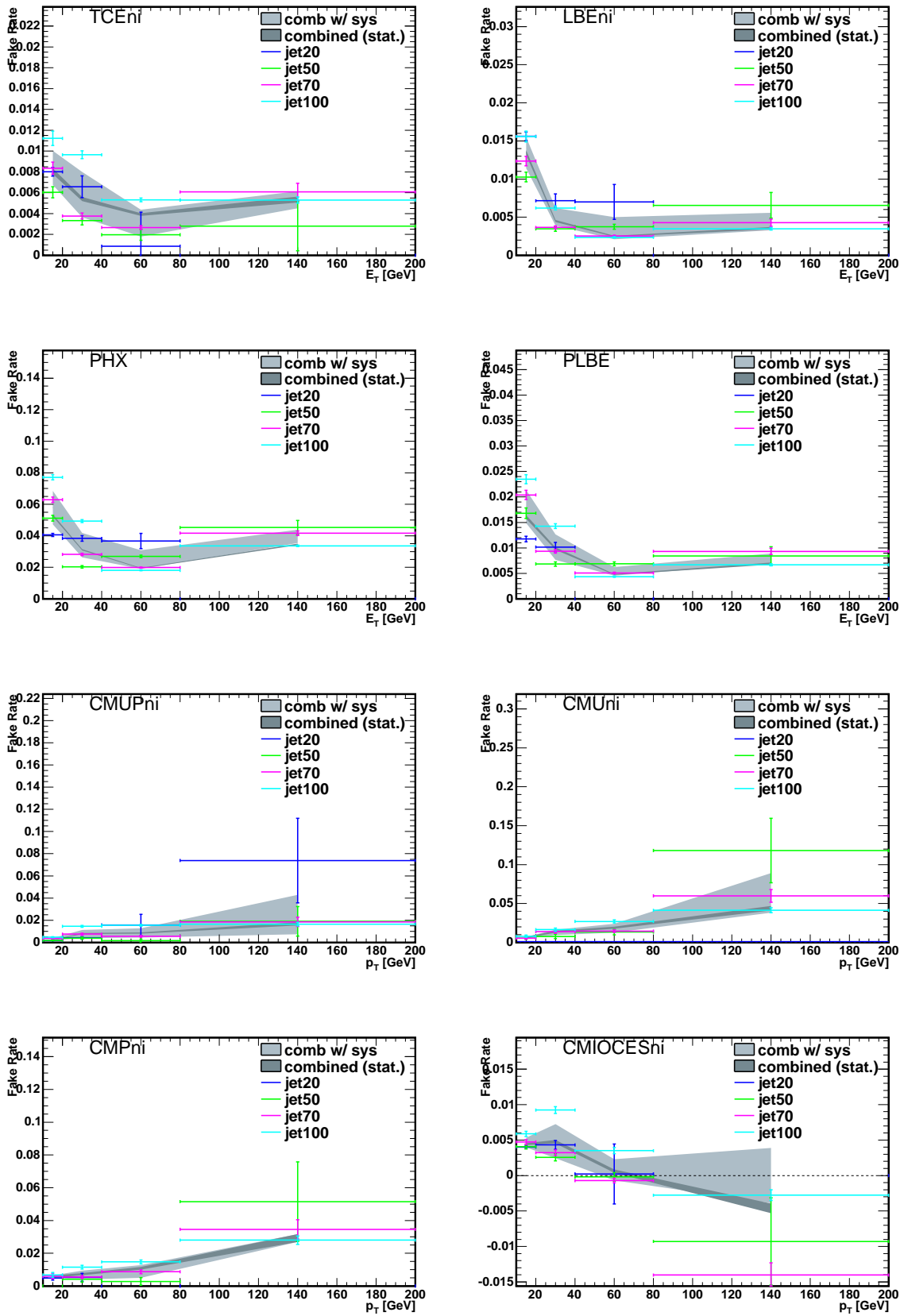


Figure 5.4: Fake rate probabilities as a function of the  $E_T$  ( $p_T$ ) of the denominator object for the several electron(muon) categories.

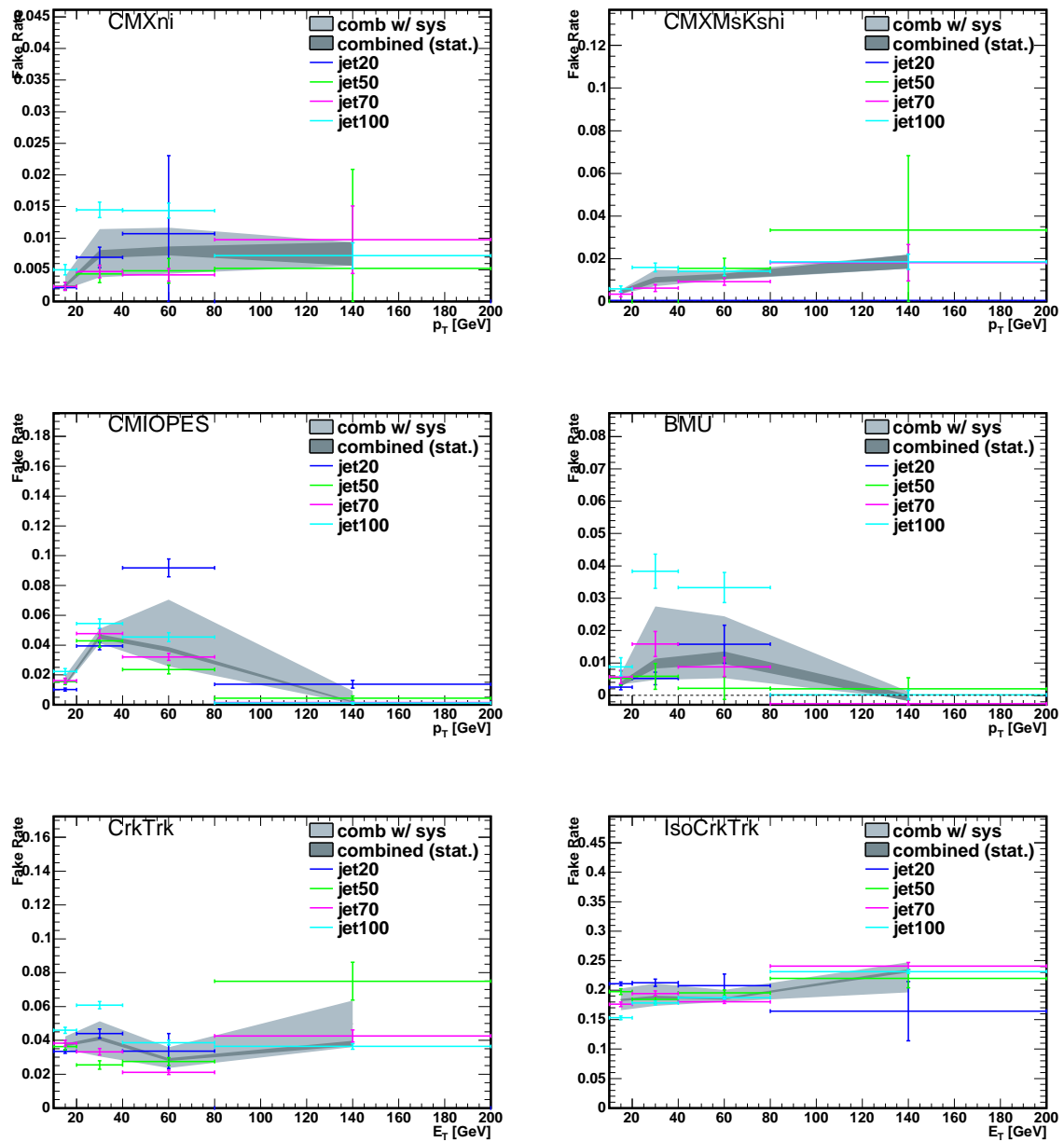


Figure 5.5: Fake rate probabilities as a function of the  $p_T$  of the denominator object for the several muon categories.

transverse energy of the denominator objects. The fake rates are of the order of few percents and, as expected, forward lepton categories has higher mis-reconstruction probabilities, due to the worse resolution in that region of the detector.

Once evaluated the probability for a denominator object to be mis-identified as a lepton we can use such probability to weight events containing one (or more) such denominator objects with the corresponding fake probability. The weighted events are then propagated in the analysis giving an estimate of the expected background events containing one or more *fake* leptons that can contaminate the considered sample. If a denominator object can fake different lepton categories the probability that it mimics one of such leptons is given by the sum of the different misidentification probabilities; for these events the denominator is counted as misidentifiable object of each category.

The procedure used relies on the assumptions that the contribution from real leptons in the denominator sample is negligible and that the identification variables don't bias the kinematic of the physics processes considered. The efficiency of tight identification of real leptons is pretty high, hence the fraction of real leptons which pass the denominator requirement (but not the tight selections) is small.

The background to the four lepton sample is expected to be populated by events with one or more fake leptons (and the remaining real ones). Events with any additional fake lepton are suppressed by a factor of the order of the fake rates ( $\mathcal{O}(10^{-2})$ ). We will consider events with a maximum of two fake leptons since, even if possible, events with more fake leptons are suppressed by 2÷3 order of magnitude with respect to the former case.

### Multiple fake lepton background

To obtain an estimate of the background in the four lepton data sample we consider a sample of collected data containing 2 or 3 fully identified leptons and additional denominator objects for a total of at least 4 objects, these events are weighted with the appropriate fake rate  $f_i(p_T)$  to obtain the estimate of the background contribution to our sample. If an event has more than 2(3) leptons or more than 2(1) denominators we create different candidate events for each possible combination of 3 leptons + 1 fake or 2 leptons + 2 fakes. Since the  $Z$ +jets and  $Z\gamma$ +jets events (where the  $\gamma$  is also misidentified as an electron) are the dominant contributions to the background, the number of expected events with three real leptons and one fake lepton ( $N_{3l+1f}$ ) and the number of events with two real leptons and two fake leptons ( $N_{2l+2f}$ ) are evaluated. The last one,  $N_{2l+2f}$ , is calculated from the sample of events with two leptons and two denominators, weighting the event with the fake rate probability of the two denominators:

$$N_{2l+2f} = \sum_{evts} f_{d1}(p_T^{d1}) \cdot f_{d2}(p_T^{d2}) \equiv f^2 \cdot N_{2l+2d}, \quad (5.3)$$

where  $d_{1,2}$  are the two denominator objects in each event and the last term of the equation is simply the compact definition of the previous one, suppressing the sum symbol.

The number of events with three leptons and one fake is calculated starting from the following relation:

$$N_{3l+1d} = N_{Z\gamma+1d} + N_{2l+1f+1d} = N_{Z\gamma+1d} + 2 \cdot f \cdot N_{2l+2d} \quad (5.4)$$

and then weighting each denominator by the fake probability:

$$f \cdot N_{3l+1d} = f \cdot N_{Z\gamma+1d} + f \cdot N_{Z+1f+1d} = f \cdot N_{Z\gamma+d} + 2 \cdot f^2 \cdot N_{Z+2d}. \quad (5.5)$$

The total number of background events is obtained as the sum of the two contributions:

$$N_{Z\gamma+1fake} + N_{Z+2fake} = f \cdot N_{Z\gamma+1d} + f^2 \cdot N_{Z+2d} \quad (5.6)$$

$$= f \cdot N_{3l+d} - 2f^2 \cdot N_{Z+2d} + f^2 \cdot N_{Z+2d} \quad (5.7)$$

$$= f \cdot N_{3l+1d} - f^2 \cdot N_{Z+2d}. \quad (5.8)$$

With the method described we obtain a data-driven estimate of the background due to fake leptons reconstruction to the four lepton sample, which is expected to be the only relevant background to the  $ZZ$  signal in the four charged lepton decay mode.

### 5.1.4 Event Selection

Once we reconstruct four leptons in the detector we can extract information on the whole  $ZZ$  production to isolate the signal by explicitly applying further requirements. In addition to the  $p_T$  requirement mentioned in Section 5.1.1 we select events where the leptons in the detector are spaced by a minimum  $\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2} = 0.1$ . The four reconstructed leptons are grouped to two pair of same flavor and opposite charge particles, minimizing the function:

$$F(\ell_1, \ell_2, \ell_3, \ell_4) = (M_{\ell_1\ell_2} - M_Z)^2 + (M_{\ell_3\ell_4} - M_Z)^2 \quad (5.9)$$

where  $M_{\ell_1\ell_2}$ ,  $M_{\ell_3\ell_4}$  are the invariant masses of each pair and  $M_Z$  is the nominal  $Z$  mass [3].

Once paired the leptons in such way we reconstruct two  $Z$  bosons in the event by applying specific requirements on the invariant mass of the lepton pairs. We require one of the two reconstructed  $Z$  to be on-shell, hence, naming  $\ell_1$  and  $\ell_2$  the pair with the invariant mass closest to  $m_Z$  [3], we ask that  $76 \leq M_{\ell_1\ell_2} \leq 106$  GeV/ $c^2$ <sup>13</sup>. A looser request is applied on the invariant mass of the second lepton pair,  $40 \leq M_{\ell_3\ell_4} \leq 140$  GeV/ $c^2$  to increase the acceptance on off-shell produced  $Z$ . Figure 5.6 shows the distribution of the two reconstructed  $Z$  invariant masses. The MC simulation shows that the requirements on the dilepton invariant masses are highly efficient on the  $ZZ$  signal ( $\epsilon \sim 80$  % of the events with four charged leptons satisfies the dileptons mass requirements). We expect background contributions from non- $ZZ$  events not to have necessarily one of the two dilepton masses compatible with an on-shell  $Z$ , hence being partly rejected by these selections.

After we apply the requirements discussed before we obtain a clear sample dominated by  $ZZ \rightarrow \ell\ell\ell'\ell'$  candidate events with a tiny expected background from events with fake leptons in them. As summarized in Table 5.3, in the collected sample we observe 7 data events, with respect to the total prediction of  $9.65 \pm 1.55$  events (of which  $0.06 \pm 0.03$  are expected to be background).

With the information from the reconstructed leptons we can compare the kinematic properties of the collected events with the signal and background prediction. Some of the more relevant kinematic variable distributions are reported in Figures 5.7 and 5.8, which show that the observed events are compatible with being  $ZZ \rightarrow \ell\ell\ell'\ell'$  decays.

We can extract information on the  $ZZ$  production cross section from the observed and expected number of events reported in Table 5.3, minimizing a likelihood function that takes into account the possible systematic uncertainties that affects the signal and background prediction. In Section 5.1.5 are illustrated the systematic uncertainties considered in this measurement while in Section 5.1.6 is explained the fitting procedure.

<sup>13</sup>This cut selects the whole  $Z$  peak, see Figure 5.3 for an example of the reconstructed  $Z$  peak width.

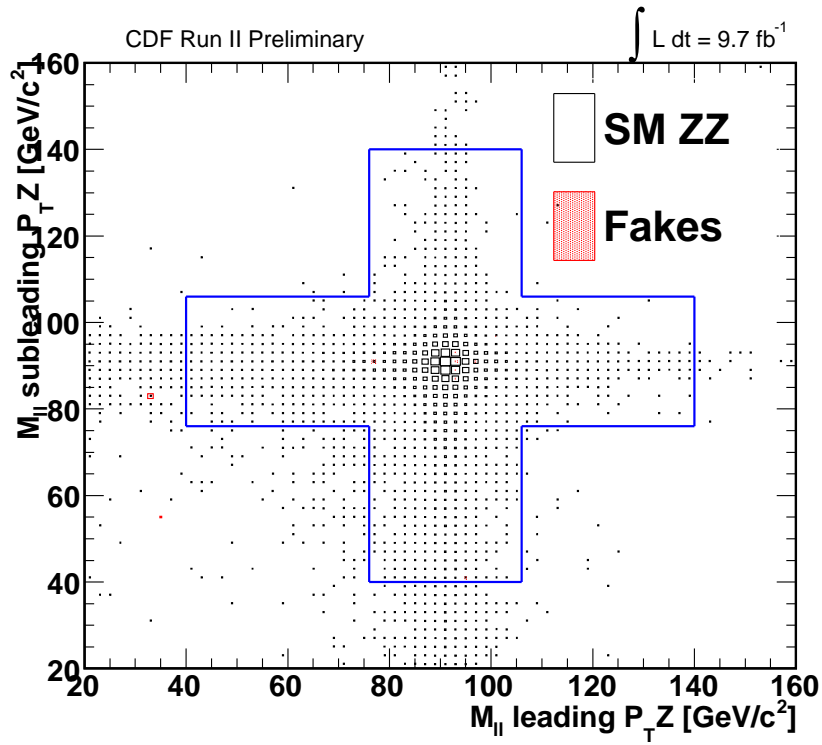


Figure 5.6: Distribution of the sub-leading Z vs. leading Z  $M_{\ell\ell}$  for the predicted signal, fake lepton background contribution.

Process	$\int \mathcal{L} = 9.7 \text{ fb}^{-1}$ candidate events	
ZZ	$9.59 \pm 1.55$	
$Z(\gamma)$ +jets	$0.06 \pm 0.03$	
<b>Total Expected</b>	$9.65 \pm 1.55$	
<b>Data</b>	7	

ZZ  $\rightarrow \ell\ell\ell\ell$  Signal Region

Table 5.3: Expected events in the four-lepton collected sample for the ZZ signal and background from fake lepton events, compared to the observed data.

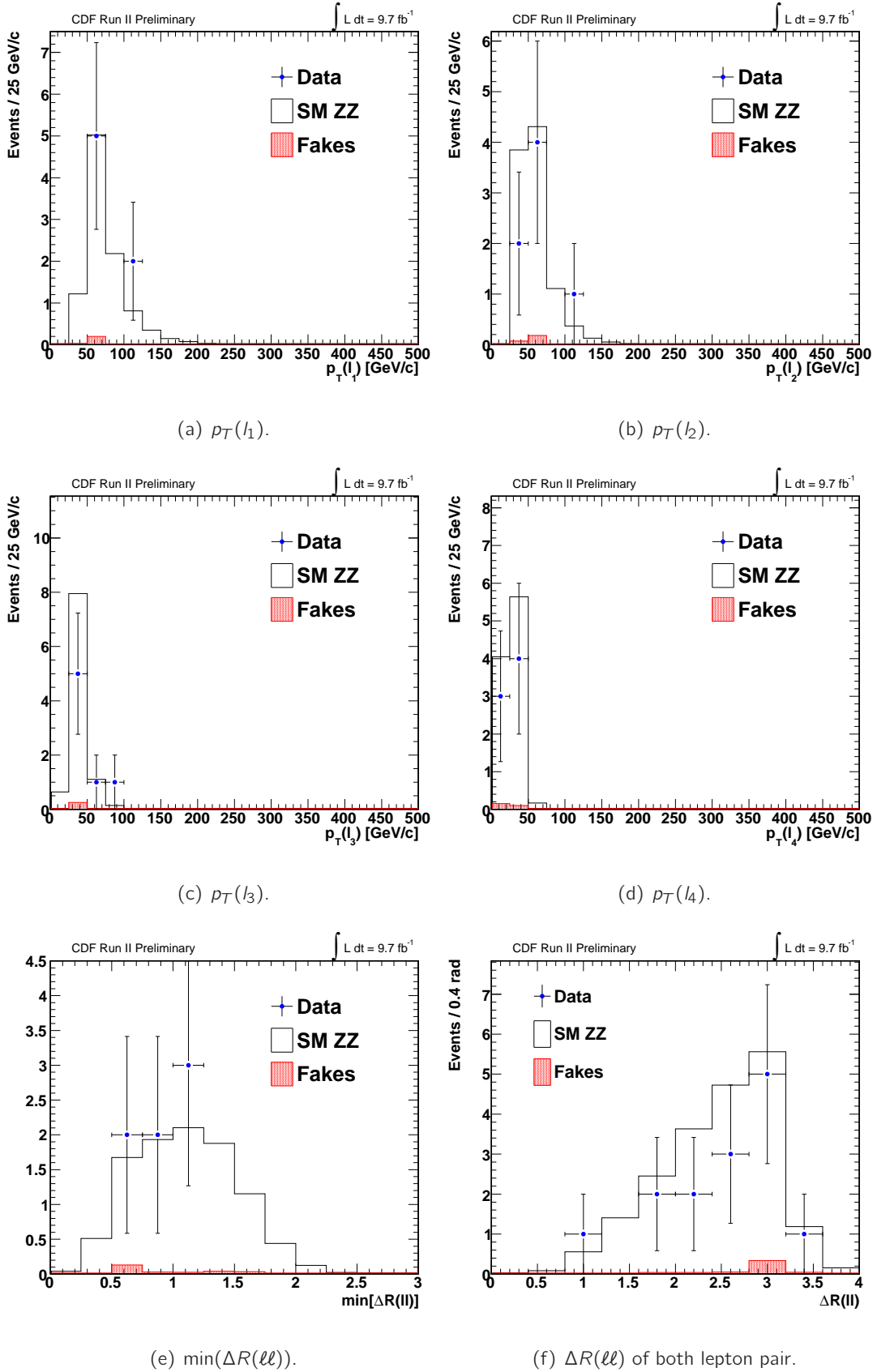


Figure 5.7: Comparison of the MC simulation prediction and observation for some kinematic variables of the events in the collected sample.

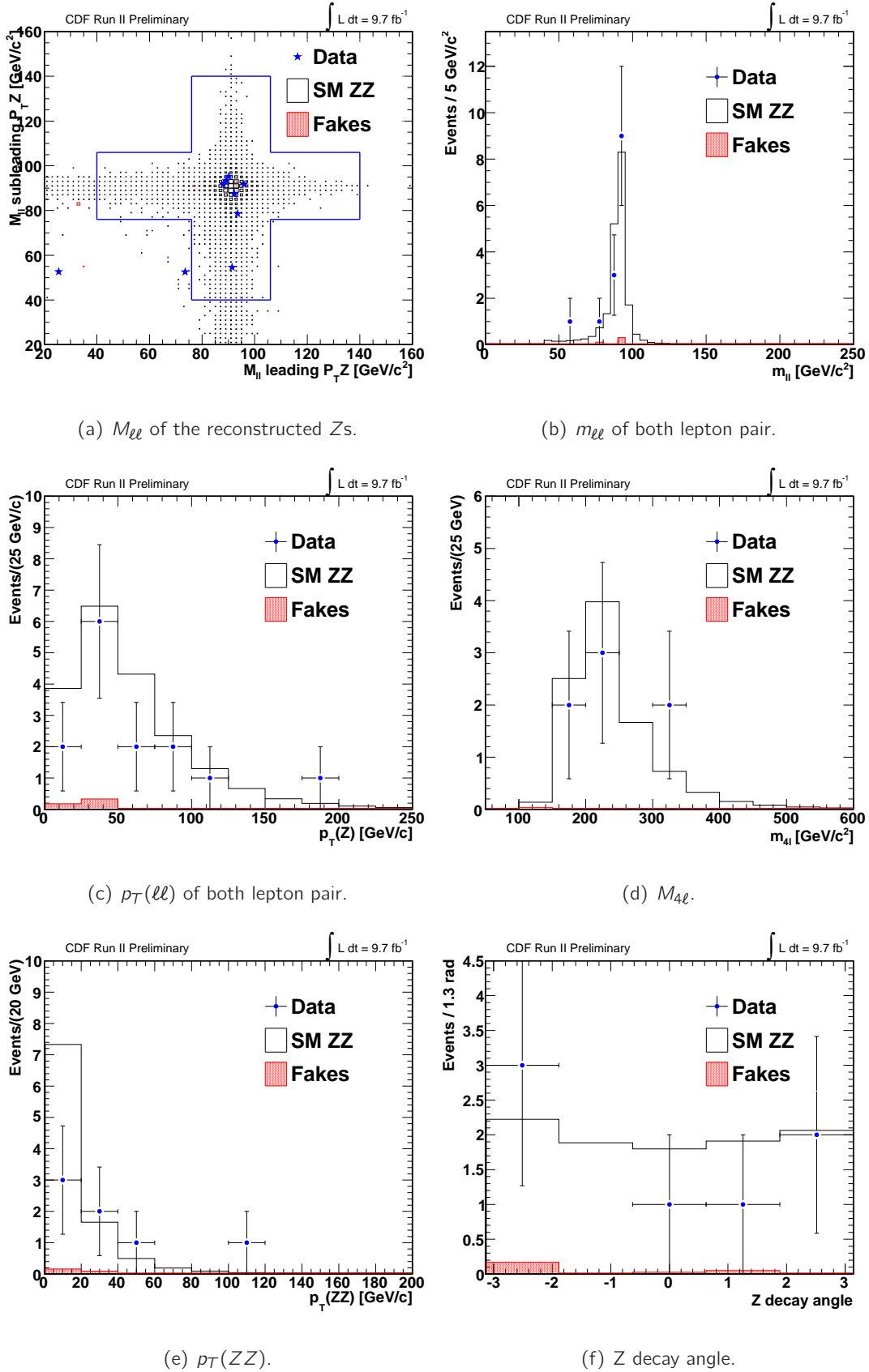


Figure 5.8: Comparison of the MC simulation prediction and observation for some kinematic variables of the events in the collected sample.



### 5.1.5 Systematic Uncertainties

In this section we discuss the sources of systematic uncertainties considered in the  $ZZ$  process investigation. We will focus in particular on the ones that can significantly affect the cross section measurement, which we're interested in. The source of uncertainties can be grouped in two main categories:

**Theoretical uncertainties:** the theoretical model exploited to simulate each process, including the calculation of the cross section and NLO correction factors, reflect its uncertainty on the prediction of the contribution of each process to the sample considered;

**Experimental uncertainties:** the data-taking conditions of the experimental apparatus are not completely reproduced in the detector simulation, hence systematic source of uncertainties can modify the acceptance of the requirements applied in the analysis;

The significant sources of systematics that we consider for the  $ZZ \rightarrow \ell\ell\ell'\ell'$  production cross section are summarized in Table 5.4; each row represents a source of systematic uncertainty while columns express the relative uncertainty on the predicted number of events for each process considered.

Source	ZZ	Z(fl)+jets
PDF Model	2.7 %	
Higher-order Diagrams	2.5 %	
Luminosity	5.9 %	
Jet Fake Rates		50.0 %
Lepton ID efficiencies	3.6 %	
Trigger efficiencies	0.04 %	

Table 5.4: Systematics uncertainties for signal and background that contribute to the final data set.

**PDF** The known uncertainties on the PDF used in MC simulation reflect on the acceptance of each process. Following the prescriptions indicated in [79] the estimate of their effect on the simulation is obtained varying the independent parameters of the PDF function fit according to their respective errors. The resulting modified PDF sets are then used to produce an alternative MC simulation, which is used to estimate the variation on the normalization of the several processes.

**Higher order diagrams** Another uncertainty comes from the difference in acceptance due to the usage of a Next to Leading Order Monte Carlo respect to the PYTHIA Leading Order simulation. This is mainly due to an approximation in the LO simulation of the radiative processes, fragmentation and hadronization. Using a Monte Carlo program MCFM [18] designed to calculate cross sections at the femtobarn level for various processes at hadron colliders, we calculate the difference in the acceptance of this analysis using a LO simulation with respect to a NLO simulation and consider this discrepancy as a systematic uncertainty on the PYTHIA LO simulated process.

**Luminosity** The integrated luminosity of the considered sample is measured using the Cherenkov luminometers (described in Section 3.2.4) with an uncertainty of 5.9% considering two main sources. The inelastic  $p\bar{p}$  cross section has a 4% uncertainty while another 4.2% comes from systematic uncertainties on the geometry description of the Cherenkov detector in the simulation. This error doesn't affect the  $Z(\gamma)+\text{jets}$  background which is estimated from data.

**Jet Fake Rates** The only systematic uncertainty that affects the  $Z(\gamma)+\text{jets}$  background normalization is due to the *fake rate* method used for its determination. We consider the error on the fake rates, taken as the maximum spread among the fake rates measured in the different jet samples (JET\_20/50/70/100), represented by the gray band in Figures 5.4 and 5.5. The positive and negative variation of these probabilities results in a variation of the weight assigned to the lepton-plus-denominator event, and consequently in a difference of the predicted number of background events.

**Lepton ID efficiencies** The scale factor on the lepton reconstruction efficiencies described in Section 5.1.2 are fluctuated up and down by their uncertainties ( $\pm 1\sigma$ ). The difference in acceptance of these variation is considered as the uncertainty on the lepton IDentification efficiency, found to be in the range between 1% and 3.6% for all the processes. To each process is assigned a fixed conservative uncertainty of 3.6% due to this fact, considered completely correlated among the different process samples, which has a small effect on the final result.

**Trigger efficiencies** Similarly to the Lepton ID, trigger efficiencies are varied coherently up and down by their uncertainties and an uniform and correlated uncertainty is assigned to every simulated sample.

### 5.1.6 Cross section measurement

We can extract the ZZ production cross section by minimizing a likelihood function  $\mathcal{L}$  built from the predicted number of events, the observed data, and the information on the systematic uncertainties. We also can obtain the significance of this signal, a measure of whether or not the signal could be due to a background fluctuation.

To extract the cross section we use a Bayesian approach to a maximum likelihood fit [80]. Given the observable  $x$  from which we wish to determine a quantity  $\alpha$ , the posterior probability distribution function (p.d.f)  $p(\alpha|x)$  is determined by the relation:

$$p(\alpha|x) = \frac{\int \mathcal{L}(x|\alpha, \theta) f(\alpha) g(\theta) d\theta}{\int \int \mathcal{L}(x|\alpha') f(\alpha') d\alpha' d\theta} \quad (5.10)$$

where  $\mathcal{L}(x|\alpha, \theta)$  is the likelihood function,  $f(\alpha)$  is the prior p.d.f. for  $\alpha$ , and  $g(\theta)$  is the prior p.d.f for nuisance parameters  $\theta$  representing the systematic uncertainties. The value  $\alpha$  for which is obtained the maximum of the posterior p.d.f. can be quoted as the result of the measurement, and the interval around this value that covers 68% of the total area of the p.d.f. represents the uncertainty on the measurement. The nuisance parameters are integrated over their possible value ranges; this is done with a numerical integration that randomly samples the full space of each nuisance parameter's prior p.d.f., using Markov Chain Monte Carlo integration<sup>14</sup>[81].

<sup>14</sup>This method finds more efficiently the regions of the nuisance parameter space contributing most to the likelihood, hence obtaining a more stable result in the fit.

In this case the likelihood function considered is based on the Poisson probabilities to observe the number of events found in the data, given the predicted signal and background number of events. The prior p.d.f.  $g(\theta)$  is represented in this case by the Gaussian constraints applied to account for the systematic errors,  $\theta_j$ , treated as independent nuisance parameters. The detailed description of the systematic errors can be found in Section 5.1.5 with a summary in Table 5.4. The likelihood function  $\mathcal{L}(n|\alpha, \theta)$  is then given by

$$\mathcal{L}(n|\alpha, \theta) = \frac{\mu^n e^{-\mu}}{n!} \cdot \prod_{j=\text{systr.}} e^{-\frac{\theta_j^2}{2}} \quad (5.11)$$

where the index  $j$  runs over the source of systematic uncertainties,  $\mu$  is the total number of events expected in the considered sample and  $n$  is the number of the observed data events.  $\mu$  takes into account also for the systematic errors and is given by the formula:

$$\mu(\alpha, \theta) = \alpha_{ZZ} \cdot \left[ \prod_j (1 + f_{ZZ}^j \cdot \theta_j) \right] N_{ZZ}^{\text{Exp}} + \left[ \prod_j (1 + f_{Z(\gamma)+jets}^j \cdot \theta_j) \right] N_{Z(\gamma)+jets}^{\text{Exp}} \quad (5.12)$$

In this expression  $f_k^j$  is the fractional uncertainty due to the  $j^{\text{th}}$  systematic source for the physics process  $k(=ZZ, Z(\gamma)+jets)$  and  $N_k^{\text{Exp}}$  is the expected number of events of the physics process  $k$ . This construction of the likelihood function takes into account properly the correlations of the systematic errors. The  $\alpha_{ZZ}$  represents the ratio between the measured cross section and the one used to normalize the  $ZZ$  Monte Carlo, which is left free in the fit, making no assumption on its prior p.d.f. (flat  $f(\alpha)$ ). The  $\alpha_{ZZ}$  returned by the posterior p.d.f. multiplied by the input  $ZZ$  cross section gives the measured value of the  $ZZ$  cross section:

$$\sigma_{ZZ}^{\text{Measured}} = \alpha_{ZZ} \cdot \sigma_{ZZ}^{\text{MC}} \quad (5.13)$$

Integrating the likelihood function over the nuisance parameters we obtain the posterior p.d.f. for the  $ZZ$  cross section shown in Figure 5.9-a, from which we measure a  $ZZ$  production cross section:

$$\frac{\sigma(\rho\bar{\rho} \rightarrow ZZ)}{\sigma_{SM}} = 0.71_{-0.25}^{+0.32}(\text{stat.})_{-0.05}^{+0.08}(\text{syst.}) = 0.71_{-0.25}^{+0.32} \quad (5.14)$$

which is lower than the predicted NLO SM one but compatible within less than  $1\sigma$ .

### Signal Significance

The significance of the considered signal is a measure of how much the data deviates from a model with no  $ZZ$  production, or a background-only hypothesis. This quantity is computed from the  $p$ -value, the probability that the background could fluctuate to create the observed signal<sup>15</sup>. Since we don't have analytical assumption on the  $\alpha_{ZZ}$  prior p.d.f., pseudo-experiments are generated assuming  $ZZ$  production cross section equal zero, sampling the possible values of the nuisance parameters in their existence domain. We determine how *signal-like* a pseudo-experiment (and data) is by evaluating a test statistic based on the likelihood ratio, defined as:

$$T_s = -2 \ln Q = -2 \ln \left( \frac{\mathcal{L}_1}{\mathcal{L}_0} \right) . \quad (5.15)$$

<sup>15</sup> This hypothesis is represented by a likelihood function  $\mathcal{L}_0$  similar to the one in Equation 5.11 but assuming  $\alpha_{ZZ} = 0$ , i.e. no contribution expected from  $ZZ$  production.

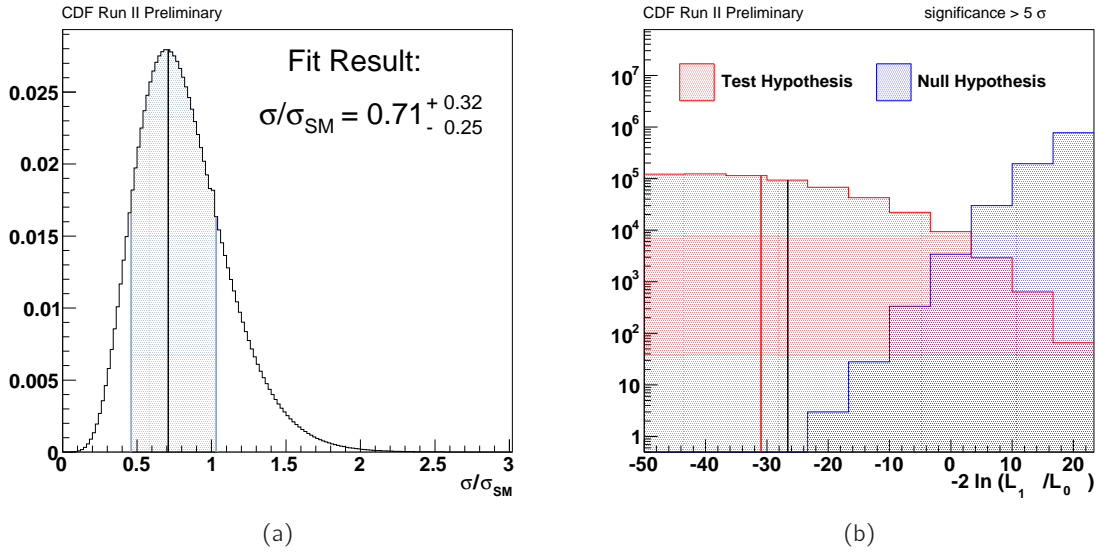


Figure 5.9: (a) Posterior p.d.f. of the parameter  $\alpha_{ZZ}$  after the integration over all the other parameters. The blue band represents the  $\pm 1\sigma$  confidence interval. (b) Test statistic distribution for signal-like (red) and background-only (blue) pseudo-experiments. The black line represents the observed test statistic value for the data, while the red one is the median value of the signal-like pseudo-experiments.

In this expression  $\mathcal{L}_1$  is the likelihood with SM production of  $ZZ$  (Test Hypothesis) whereas  $\mathcal{L}_0$  is the one with  $ZZ$  cross section set to zero, corresponding to the Null Hypothesis (background-only). More signal-like pseudo-experiments would have lower values of the test statistic. The  $p$ -value is the fraction of pseudo-experiments that have a  $T_s$  value lower than the data.

By generating pseudo experiments including SM  $ZZ$  contribution we can evaluate the expected significance of the signal. The  $T_s$  is evaluated for each of these pseudo-experiments, and the expected significance is defined as the probability for a background-only pseudo-experiment to have a  $T_s$  less than the median  $T_s$  of the signal-like pseudo-experiments.

The significance is usually expressed in terms of  $\sigma$ , the standard deviation of a Gaussian distribution. The conversion from the  $p$ -value to the number of  $\sigma$  is done by finding the value  $x$  such that the integral of a Gaussian distribution with mean zero from  $x\sigma$  to  $\infty$  ( $-\infty$  to  $-x\sigma$ ) is equal to the  $p$ -value. For example, a  $p$ -value of  $2.7 \cdot 10^{-3}$  corresponds to a significance of  $3\sigma$ .

Figure 5.9(b) shows the distribution of the  $T_s$  for background only and signal+background pseudo experiments, as well as for the data. Generating 1 billion pseudo-experiments we observe none of them having a smaller  $T_s$  than the one evaluated in the data (and than the median of the signal-like pseudo-experiments). We then calculate as a lower limit on the observed (expected) significance the result in the scenario such that we observe one pseudo-experiment out of the total number of generated ones having a  $T_s$  lower than the one of the data. Hence for this case we conclude that the investigated  $ZZ$  signal has an observed (expected) significance greater than  $6\sigma$ .

## 5.2 $ZZ \rightarrow ll\nu\nu$

In addition to the cross section measurement obtained exploiting the four charged lepton decay mode of the  $ZZ$ , we consider the decay of the same boson pair in a different leptonic

decay mode:  $ZZ \rightarrow \ell^+\ell^-\nu\bar{\nu}$  (<sup>16</sup>). To extract the  $ZZ$  production cross section from this decay mode, we need to detect one  $Z \rightarrow \ell^+\ell^-$  in the final state and obtain indirect information on the undetected  $Z \rightarrow \nu\bar{\nu}$  from energy unbalance in the detector transverse plane. We start from a sample of events containing a  $Z$  boson and apply additional kinematic requirements to reduce background contribution from non- $ZZ$  concurrent processes. We enhance the signal separation from the background processes using an artificial neural network.

### 5.2.1 Initial Sample Composition

We investigate the  $ZZ$  production in the  $\ell\nu\nu$  decay mode starting from the same data sample described in Section 5.1.1, collected using single high- $p_T$ (high- $E_T$ ) lepton trigger paths (described in Appendix B). The collected events are demanded to belong to one of the good run lists reported in Table 5.1, and correspond to an integrated luminosity of about  $9.7 \text{ fb}^{-1}$ . The identification of the  $ZZ \rightarrow \ell\nu\nu$  signal begins with the reconstruction of one  $Z \rightarrow \ell\ell$  in the detector. As a preliminary requirement we select events with exactly two fully identified and isolated high- $p_T$  leptons ( $e, \mu$ ) belonging to one of the categories described in Appendix A.1. At least one of the two leptons is required to satisfy trigger conditions described in Appendix B, in particular having a transverse momentum (energy) greater than 20 GeV/c (GeV). The second lepton must have a transverse momentum (energy) greater than 10 GeV/c (GeV). We reconstruct one  $Z$  in the event by requiring that the two reconstructed leptons have opposite charge and same leptonic flavor ( $e^+e^-, \mu^+\mu^-$  pairs), the invariant mass of the two,  $M_{\ell\ell}$ , being in the range between 76 and 106 GeV/c<sup>2</sup> <sup>17</sup>. The requirements applied on the two reconstructed leptons have only a marginal acceptance on  $Z \rightarrow \tau^+\tau^-$  production:  $\tau \rightarrow \ell\nu\tau\bar{\nu}_\ell$  is  $\sim 17\%$  (for  $\ell = e, \mu$ ) hence only 6% of the times a  $\tau^+\tau^-$  pair decay to  $e^+e^-, \mu^+\mu^-$ . This branching fraction is further reduced by the kinematic requirements applied to the electrons and muons.

This preselection yields a sample of  $\sim 900.000$  events which is expected to be dominated by Drell-Yan production with possible associated jets from initial/final state radiation (ISR/FSR), since this process has a large cross section compared to other processes leading to a similar signature:  $\sigma(p\bar{p} \rightarrow Z/\gamma^* \rightarrow \ell^+\ell^-) \sim 490 \text{ pb}$  for  $m(\ell\ell) \geq 20 \text{ GeV}/c^2$  at NNLO[82].

The main difference between  $ZZ \rightarrow \ell\nu\nu$  signal and the Drell-Yan background are the two additional neutrinos in the final state, whose presence might be inferred from a significant missing transverse energy ( $\cancel{E}_T$ ).

Other processes enter with a smaller fraction in this sample and should be considered. The production of a pair of  $W$  bosons ( $WW$ ) result in the production of two leptons in the final state when both bosons decay to  $\ell\nu$ .  $WZ$  leptonic decay mode ( $WZ \rightarrow \ell\nu\ell'\ell'$ ) can give a similar final state if one of the three leptons is not properly reconstructed or falls outside the detector acceptance. These two processes are predicted to be produced with a cross section of  $12.4 \pm 0.8 \text{ pb}$  and  $3.5 \pm 0.2 \text{ pb}$  respectively, therefore about 8 and 2.5 times more frequently than  $ZZ$  production, with a cross section of  $1.4 \pm 0.1 \text{ pb}$ .

An additional contribution comes from **Wfi** production where only one lepton is real, while the other is due to the misidentification of the photon as an electron, passing the selections applied. Similarly a fake lepton background contribution can result from the production of a single  $W$  boson, in association with a jet mis-identified as a lepton in the detector. The probability for this

<sup>16</sup> This process will be referred to in the following as  $ZZ \rightarrow \ell\nu\nu$ , neglecting in the symbology the charge and the particle anti-particle specification.

<sup>17</sup> Figure 5.3 shows an example of the  $M_{\ell\ell}$  distribution for dilepton events. Despite the different resolution for electron and muon momentum reconstruction, the  $M_{\ell\ell}$  range considered is wide enough so that this selection is not significantly affected by the difference in resolution for the several lepton category combinations.

to happen is smaller than the photon-to-lepton mis-identification rate but given the really high cross section for single boson production, is still comparable with the other concurring processes.

A summary of the physics processes contributing to the collected sample is reported in Table 5.5.

Process	Cross section	Branching Ratio
Drell-Yan	490 pb <sup>a</sup>	
WW	12.4 pb	$\mathcal{BR}(\ell\nu\ell\nu)= 4.5\%$
WZ	3.7 pb	$\mathcal{BR}(\ell\nu\ell\ell)= 1.4\%$
ZZ	1.4 pb	$\mathcal{BR}(\ell\ell\nu\nu)= 2.6\%$
$t\bar{t}$	7.6 pb <sup>b</sup>	$\mathcal{BR}((W \rightarrow \ell\nu)(W \rightarrow \ell\nu))= 4.5\%$
W+jets	103 pb <sup>c</sup>	$\mathcal{BR}(\ell\nu)= 21.2\%$
W $\gamma$	105 pb	$\mathcal{BR}(\ell\nu)= 21.2\%$

<sup>a</sup> Cross section calculated for  $m(\ell\ell) \geq 20$  GeV/ $c^2$ . <sup>b</sup> This assumes  $m_{top}= 172.5$  GeV/ $c^2$ .

<sup>c</sup> Cross section for  $W + \geq 1$ jet production.

Table 5.5: Summary of the properties of the processes concurring to the  $\ell\ell\nu\nu$  signature.

The kinematic properties of the signal, as well as those of the different background processes, are modeled using Monte Carlo (MC) simulations. To simulate the ZZ signal production we exploit the same inclusive MC sample described in Section 5.1.2, produced using the Leading Order (LO) PYTHIA [76] generator, that includes all the possible decay modes of each Z boson<sup>18</sup>. We also simulate WZ,  $t\bar{t}$  and Drell-Yan production using PYTHIA [76]; all the PYTHIA MC simulations use settings referred to as Tune A [77]. Given the availability of a more accurate simulation, WW production is modeled using MC@NLO [83]<sup>19</sup>. In this simulation the MC@NLO generator is interfaced with HERWIG [84] for the shower and hadronization part. To simulate W $\gamma$  production we use the Baur event generator described in [85] for the hard scattering, which provides a better description of the QED radiation. We apply a filter at the generator-level, simulating only events with  $p_T(\gamma) \geq 8$  GeV/ $c$  and  $\Delta R(\ell, \gamma) \geq 0.35$ .

Table 5.6 reports a short summary of the generated MC samples, with the event generator and the cross section used for their normalization.

The contribution from W+jets production with the mis-identified jet giving one of the two leptons is evaluated from a sample of single-lepton experimental data with the *fake rate* method described in Section 5.1.3<sup>20</sup>:

$$N_{W+1fake} = \sum_{evts} \sum_{j=den} f_{d_j} \equiv f \cdot N_{W+1d} . \quad (5.16)$$

We compare the collected data with the full set of simulated samples for events passing the preselections described so far. In Figures 5.10 and 5.11 are shown the compared distribution for the main event kinematic properties, where the simulations of the several processes are stacked over

<sup>18</sup>As we did for the four lepton measurement, we spot the ZZ  $\rightarrow \ell\ell\nu\nu$  decay through offline requirements based on the reconstructed objects (leptons or jets) in the final state. This indirectly allows inclusion in the analysis a tiny fraction of ZZ signal in different decay modes.

<sup>19</sup>We use the most accurate simulation available among the CDF simulated datasets. For WW production we use a NLO simulation, which has been simulated considering higher order correction effects for other analyses.

<sup>20</sup>In this case we consider only event with one real lepton and the possible fake second lepton. The events containing one real lepton and two candidate fake leptons are splitted and two lepton+fake candidate events are considered. Each candidate is weighted with the appropriate misidentification probability and added to the background prediction.

Process	Generator	$\sigma \times \mathcal{BR}$ pb	Accuracy	$\epsilon_{\text{filter}}$
$ZZ$	PYTHIA	1.4	NLO	0.233
$WZ$	PYTHIA	3.46	NLO	0.076
$WW$	MC@NLO	12.4	NLO	1.0
$Z/\gamma^* \rightarrow \ell\ell$	PYTHIA	490	NLO	1.0
$W\gamma \rightarrow \ell\nu\gamma$	BAUR	18.6	NLO	1.0
$t\bar{t} \rightarrow (W \rightarrow \ell\nu)bW(\rightarrow \ell\nu)b$	PYTHIA	0.809 <sup>a</sup>	NLO	1.0

<sup>b</sup>This sample includes all the decays into  $e, \mu, \tau$ . It assumes  $m_{top} = 172.5 \text{ GeV}/c^2$ .

Table 5.6: List of physics processes simulated. We report the event generator, the cross section times branching fraction used in the normalization and the generator-level filter efficiencies ( $\epsilon_{\text{filter}}$ ), if any. Processes with a final state indicating a generic lepton  $\ell$  are generated separately for  $e, \mu$ , and  $\tau$ .

each other and the experimental data are overlaid. In particular we pay attention to the modeling of the  $Z \rightarrow \ell\ell$  reconstruction, i.e. the two lepton transverse momenta, the opening angles between the two leptons ( $\Delta\phi(\ell\ell), \Delta R(\ell\ell)$ ), the reconstructed  $Z$  mass ( $M_{\ell\ell}$ ) and transverse momentum ( $p_T^Z \equiv p_T(\ell\ell)$ ). As a complementary check we control the modeling of the distribution of the number of jets ( $N_{jets}$ ) and the missing transverse energy ( $\cancel{E}_T$ ) in these events. We can see that these properties are well modeled in this Drell-Yan dominated sample; for this reason we can exploit the MC simulation for the optimization of the  $ZZ$  signal extraction.

## 5.2.2 Kinematic Properties and Signal Selection

In order to evaluate the contribution in the preselected sample described above from the  $ZZ \rightarrow \ell\nu\nu$  signal and reject the major part of the Drell-Yan background we exploit several kinematic properties of the signal events. First of all, considering this leptonic decay mode, we don't expect a large hadronic activity, since jets are eventually produced only through initial/final state radiation of a gluon. In particular, in  $p\bar{p} \rightarrow ZZ$  events, the two  $Z$  are approximately balanced in the transverse plane, and no high- $E_T$  jet is expected in the opposite direction with respect to the  $Z \rightarrow \ell\ell$ . On the contrary, the presence of one or more reconstructed jets recoiling against such  $Z$  are typical of the single  $Z$  production (Drell-Yan events) where the jets balance the transverse boost of the produced  $Z$ . Also the  $ZZ$  production in the semileptonic decay mode  $ZZ \rightarrow \ell\ell jj$  usually is characterized by the presence of two jets recoiling against the leptonic  $Z$ . To select the  $ZZ \rightarrow \ell\nu\nu$  signal we apply a veto on the presence of a recoiling jet explicitly requiring that no jet (with  $E_T \geq 15 \text{ GeV}$  and  $|\eta| \leq 2.4$ ) is reconstructed in the region  $\Delta\phi(j, Z) \geq \pi/2$ . Figures 5.12 show the distribution of  $\Delta\phi(j, Z)$  for all the jets reconstructed in the events<sup>21</sup>. For the background events (Drell-Yan,  $W$ +jets) most of the time in which one jet is present<sup>22</sup>, this has an angle with respect to the  $Z \sim \pi$ .

By vetoing the presence of a recoiling jet we select a sample of data events with a small amount of reconstructed jets overall, the remaining ones directed in the same side as the  $Z \rightarrow \ell\ell$ . Figures 5.13 show the distribution in terms of number of jets present in the events (*jet multiplicity*,

<sup>21</sup> From now on we will always count as jets the ones with  $E_T \geq 15 \text{ GeV}$  and  $|\eta| \leq 2.4$ . Lower  $E_T$  and very forward jets are included in the calorimetric energy overall information.

<sup>22</sup> In the preselected sample  $\sim 18\%$  of the events contain at least one reconstructed jet.

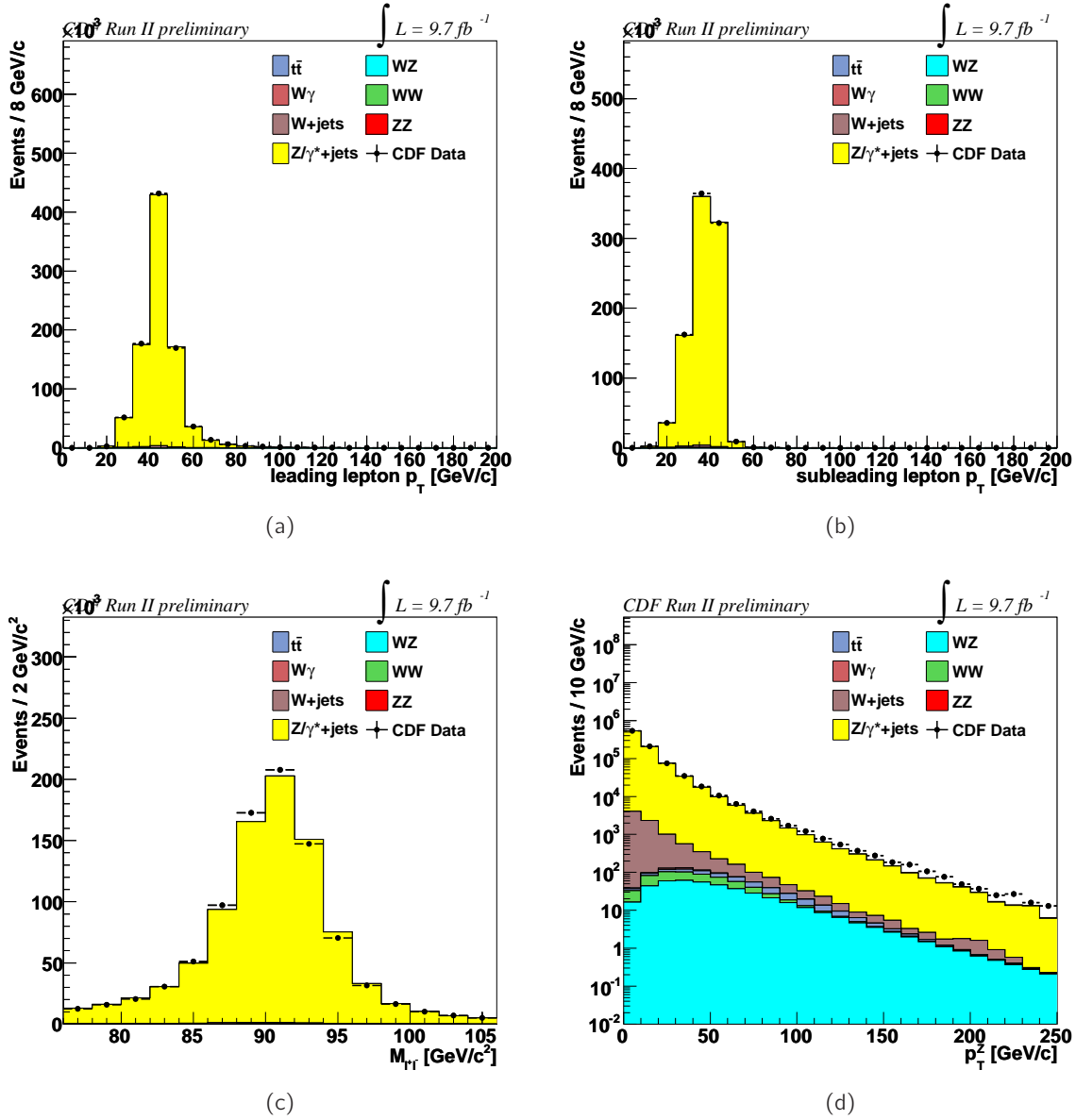


Figure 5.10: Simulation-to-data comparison for events in the preselected sample for (a) leading lepton  $p_T$ , (b) sub-leading lepton  $p_T$ , (c) dilepton invariant mass ( $M_{\ell\ell}$ ), (d) reconstructed Z transverse momentum ( $p_T(Z) \equiv p_T(\ell\ell)$ ).



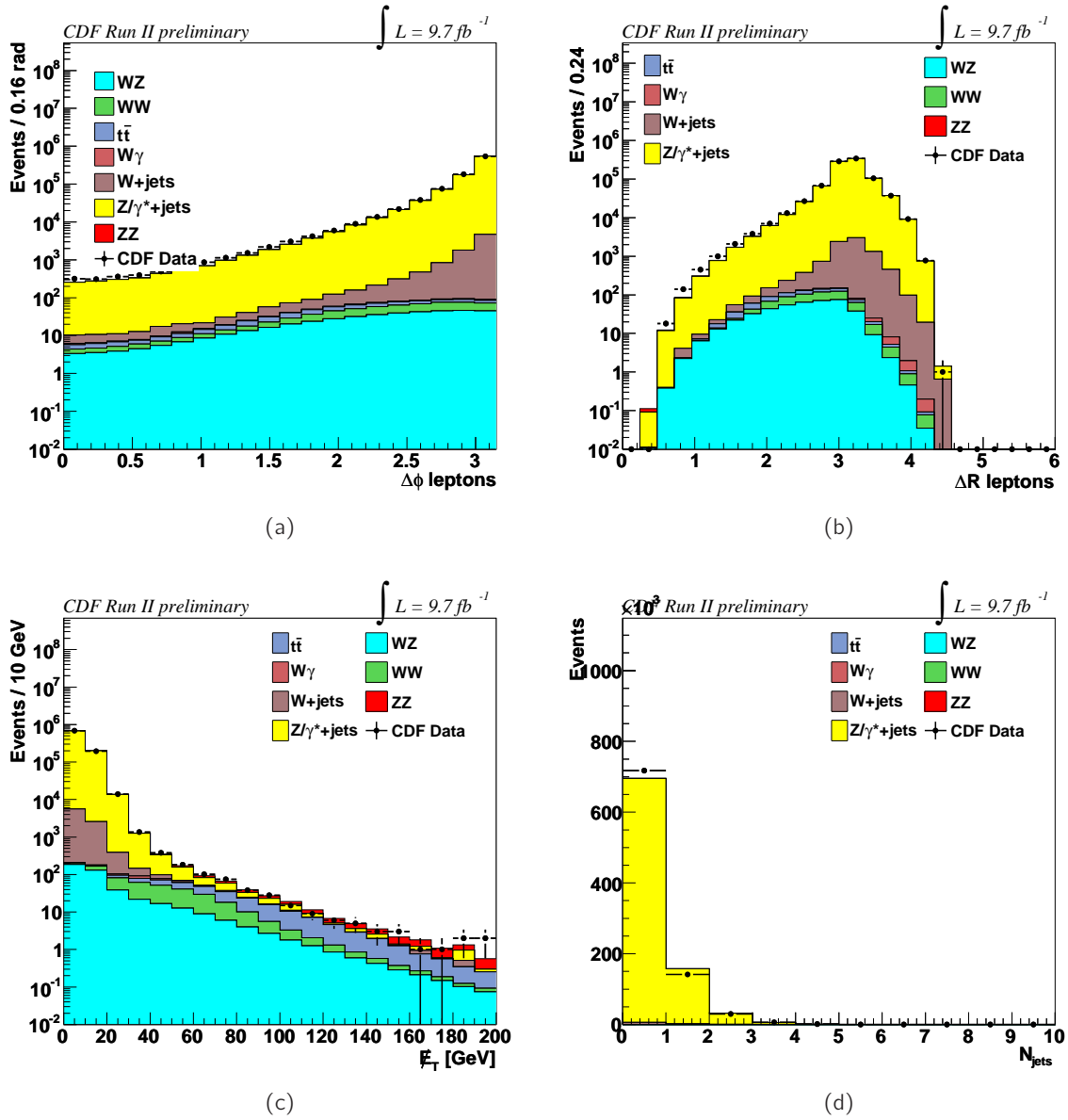


Figure 5.11: Simulation-to-data comparison for events in the preselected sample for (a)  $\Delta\phi(\ell\ell)$ , (b)  $\Delta R(\ell\ell)$ , (c) missing transverse energy ( $E_T$ ), (d) the number of jets.

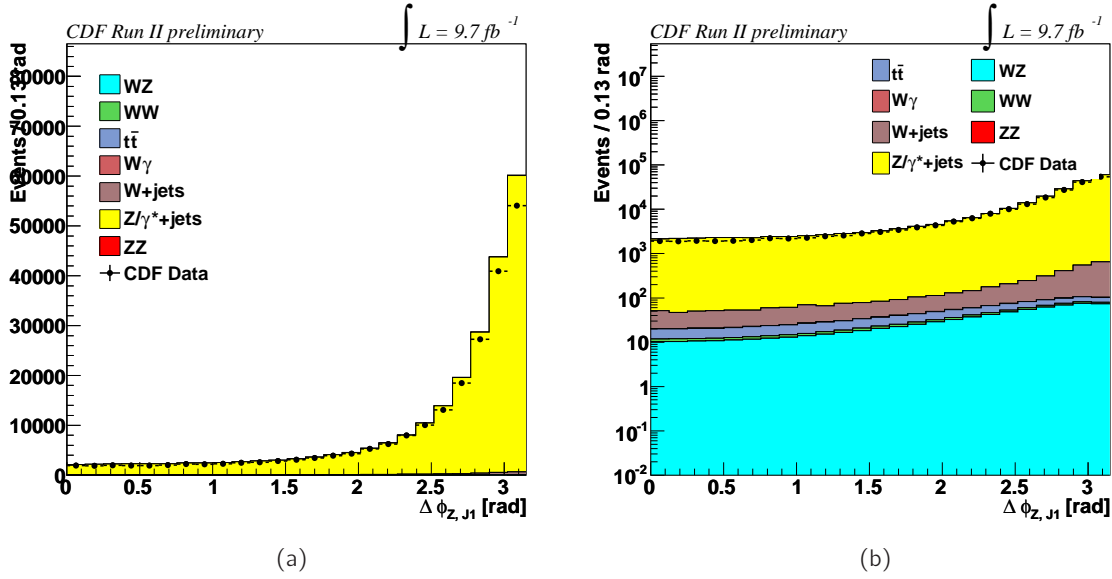


Figure 5.12: Distribution of  $\Delta\phi(j, Z)$  for event in the preselected sample. A comparison of the data distribution with the prediction is shown in (a) linear and (b) logarithmic y-scale.

$N_{jets}$ ) (a) before and (b) after the application of the veto on the recoiling jets. This requirement gives a sample of collected data  $\Delta\phi$  composed for the  $\sim 98\%$  by events with no reconstructed jet at all.

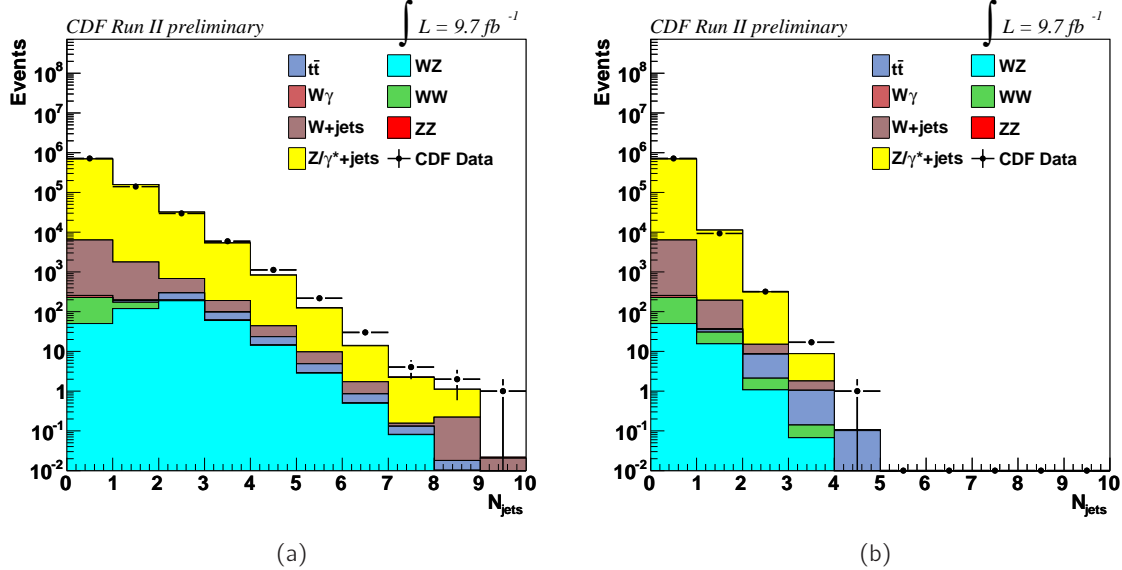


Figure 5.13: The predicted and observed number of jets in the event in the preselected sample (a) and after applying the veto on the presence of a recoiling jet (b).

To further reduce background and select ZZ events we exploit indirect information on the additional Z decaying to a pair of neutrinos, which is not present in the Drell-Yan background events. In the signal investigated process the two neutrinos which are not detected directly in the detector should produce a missing transverse energy in the detector,  $\cancel{E}_T$ , defined in Section 4.6. The presence of two neutrinos from the Z decay allows to extract only indirect information on the

two neutrinos mother particle. The missing transverse energy, in this case, is directly proportional to the second  $Z$  boost in the transverse plane ( $p_T(Z \rightarrow \nu\nu)$ ). For poorly boosted  $Z$  the opening angle between the two neutrinos is  $\sim \pi$ , resulting in a smaller amount of  $\cancel{E}_T$  in the detector, with respect to a single neutrino process. The transverse energy unbalance can also be due to the limited detector resolution and to the presence of jets not fully reconstructed<sup>23</sup>; this is the main source of  $\cancel{E}_T$  for the background processes concurring to the investigated experimental signature (e.g. Drell-Yan). Given the large Drell-Yan production cross section ( $\sim 10^6 \times \sigma(ZZ)$ ) the tail of the  $\cancel{E}_T$  distribution for this process significantly different from zero is still comparable with the  $ZZ$  signal contribution.

For this reason we exploit the peculiar topology of  $\ell\ell\nu\nu$  events to extract the  $ZZ$  signal out of the spurious background. In the plane transverse to the beam we consider the projection of the  $\cancel{E}_T$  along the  $Z$  axis, reconstructed using the two leptons, in the opposite direction, as sketched in Figure 5.14. This axial projection,  $\cancel{E}_T^{Ax}$ , is defined as:

$$\cancel{E}_T^{Ax} = - \cancel{E}_T \cdot \cos\Delta\phi(\hat{\cancel{E}}_T, \hat{p}_T^Z). \quad (5.17)$$

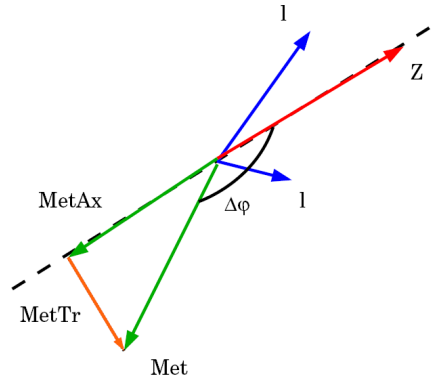


Figure 5.14: Schematic draw of the  $\cancel{E}_T$  projection along the  $Z$  axis.

Figure 5.15 shows the distribution of the  $\cancel{E}_T^{Ax}$  in the sample of preselected events with no recoiling jets. In  $ZZ$  events the  $\cancel{E}_T$  from the two neutrinos is almost aligned with the reconstructed  $Z$  in the opposite direction. For this reason the signal is expected to have a large tail for positive values of  $\cancel{E}_T^{Ax}$ . The dominant background events from Drell-Yan production present  $\cancel{E}_T$  which is uncorrelated to the  $Z$  direction, resulting in a  $\cancel{E}_T^{Ax}$  distribution peaked at low values, falling steeply with increasing values of  $\cancel{E}_T^{Ax}$ . This makes the  $\cancel{E}_T^{Ax}$  a powerful kinematic variables to distinguish  $ZZ$  signal from single  $Z$  production, therefore we select events with  $\cancel{E}_T^{Ax} \geq 30$  GeV, to cut off  $\sim 98\%$  of the Drell-Yan background events.

In summary, we consider for the cross section measurement in this decay mode, events that satisfy the following requirements:

Having two Opposite Sign and Same Flavor<sup>24</sup> leptons in the final state;

$76 \leq M_{\ell\ell} \leq 106$  GeV/ $c^2$  assures that the two reconstructed leptons are consistent with a  $Z$  and reduces non-resonant dilepton background processes;

**No jet in the event with  $\Delta\phi(j, Z) \geq \pi/2$**  to reject events with a high- $E_T$  recoiling jet;

<sup>23</sup> This means even a large number of jets with  $E_T$  below the 15 GeV threshold, used to define the *identified* jets.

<sup>24</sup> Tracks pointing to uninstrumented regions of the detector (defined in Section 4.4.1) are left with undefined flavor so that they can be matched with both electrons and muons to form the  $Z \rightarrow \ell\ell$  candidate.

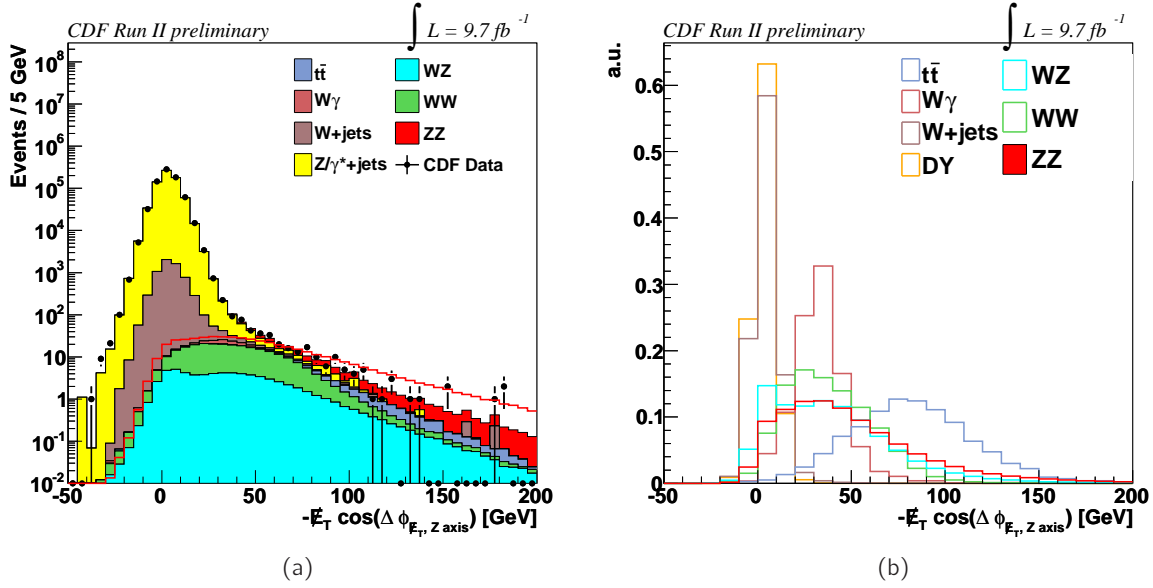


Figure 5.15: (a)  $\cancel{E}_T^{Ax}$  distribution for event in the preselected sample with no recoiling jets, comparing data to simulation. The red lines over-impressed represent the ZZ signal shape, magnified  $\times 5$  with respect to the proper normalization. (b) A shape comparison of the same kinematic variable for the several concurrent processes, normalized to unit area.

$\cancel{E}_T^{Ax} \geq 30 \text{ GeV}$  to reduce background from processes with not physical  $\cancel{E}_T$ , expected to be small and uncorrelated to the  $Z \rightarrow \ell\ell$  direction (see Figure 5.15).

The kinematic region of events passing all these selections will be referred to as Signal Region (SR) and is the one considered to investigate  $ZZ \rightarrow \ell\ell\nu\nu$  production. Table 5.7 summarizes the expected number of events for the ZZ signal and the major background processes in the Signal Region, found to be in agreement with the number of observed event in the data. The sample selected is still dominated by Drell-Yan events (50%), but has also a significant contribution from WW production (18%) and 10% of it is expected to be ZZ signal.

Figures 5.16 and 5.17 show the comparison between data and Monte Carlo for the kinematic variables distributions that characterize the events after all the selection cuts. Of particular interest are first of all the kinematic variables related to the visible  $Z \rightarrow \ell\ell$  decay (i.e.  $p_T(\ell_{1,2})$ ,  $M_{\ell\ell}$ ,  $\Delta\phi(\ell\ell)$ ,  $\Delta R(\ell\ell)$ ,  $p_T(\ell\ell)$ ); in addition to that, we focus on the hadronic activity of the events (represented by  $N_{jets}$ ) which is expected to be small, and the transverse missing energy. We consider the  $\cancel{E}_T$  in relation to the total energy in the event and the relative direction with respect to the other reconstructed objects in the event by considering the following kinematic variables: the angle between the  $\cancel{E}_T$  and the reconstructed Z, the  $\cancel{E}_T^{Ax}$  as defined in Equation 5.17, the angle between the direction of the  $\cancel{E}_T$  and the track momentum imbalance  $\vec{p}_T^{25}$   $\Delta\phi(\cancel{E}_T, \vec{p}_T)$ , and the  $\cancel{E}_T$  significance  $\cancel{E}_T^{Sig} \equiv \cancel{E}_T / \sqrt{\sum_i E_T^i}$ <sup>26,27</sup>. From this comparison we observe that the MC simulations reproduce the experimental data in their fundamental properties. Below each plot

<sup>25</sup> The track momentum imbalance  $\vec{p}_T$  is the result of the vectorial sum of the tracks reconstructed in the tracking system, i.e.  $\vec{p}_T \equiv -\sum_i \vec{p}_T^i$ ,  $p_T \equiv |\vec{p}_T|$ .

<sup>26</sup> In this expression  $\sum_i E_T^i$  is the sum of the energy deposits in all the calorimetric towers.

<sup>27</sup> From the  $\cancel{E}_T$  definition given in Section 4.6, its resolution is  $\sigma(\cancel{E}_T) \simeq \sqrt{\sum_{i=Cal. Tower} \sigma(E_T^i)^2}$ . The calorimeter energy resolution is  $\sigma_E \propto \sqrt{E}$  (see Section 3.2.2) hence  $\sigma(\cancel{E}_T) \simeq \sqrt{\sum_{i=Cal. Tower} E_T^i}$ .

Process	$\int \mathcal{L} = 9.7 \text{ fb}^{-1}$	
	candidate events	
DY	317	$\pm 51.3$
$t\bar{t}$	11.9	$\pm 2.2$
W+jets	69.5	$\pm 18.5$
$W\gamma$	17.3	$\pm 2.2$
WW	114	$\pm 10.6$
WZ	37.5	$\pm 5.3$
<b>Total Background</b>	<b>567</b>	<b><math>\pm 24.4</math></b>
ZZ	63	$\pm 11$
<b>Data</b>	<b>618</b>	

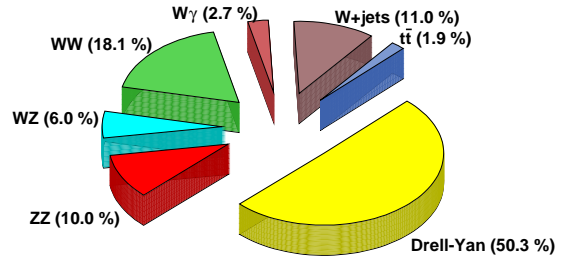
ZZ  $\rightarrow \ell\nu\nu$  Signal Region

Table 5.7: Expected and observed number of events in  $9.7 \text{ fb}^{-1}$  of integrated luminosity for the several process contribution. The uncertainty includes statistical and systematic errors added in quadrature.

is reported the distribution of the residuals, with dashed blue lines representing the systematic uncertainties on the simulated sample (details will be discussed in Section 5.2.5). For all the considered variable the discrepancies between the data and MC simulation are smaller than the uncertainties considered. A not negligible discrepancy is noticeable in the  $\Delta\phi(\ell\ell)$  and  $\Delta R(\ell\ell)$  distributions, that we will take into account with further checks.

In the selected sample the signal-to-background ratio is not so favorable to extract information on the ZZ production cross section from a counting experiment as it has been done for the four charged lepton decay mode (Section 5.1.6). To further isolate the ZZ signal from the other background processes we exploit a Multivariate Analysis Techniques (MVA) that separate signal candidate events from the background based on the event kinematic properties, as we will discuss in Section 5.2.4. This technique relies on the simulated sample for the different background and signal processes, hence it is necessary to have the best modeling of the kinematic properties of the processes involved. The accuracy of the MC simulation of the dominant background processes is tested in data control samples (Control Regions) non-overlapping with the sample of events considered for the cross section measurement, that have nonetheless kinematic properties similar to this one.

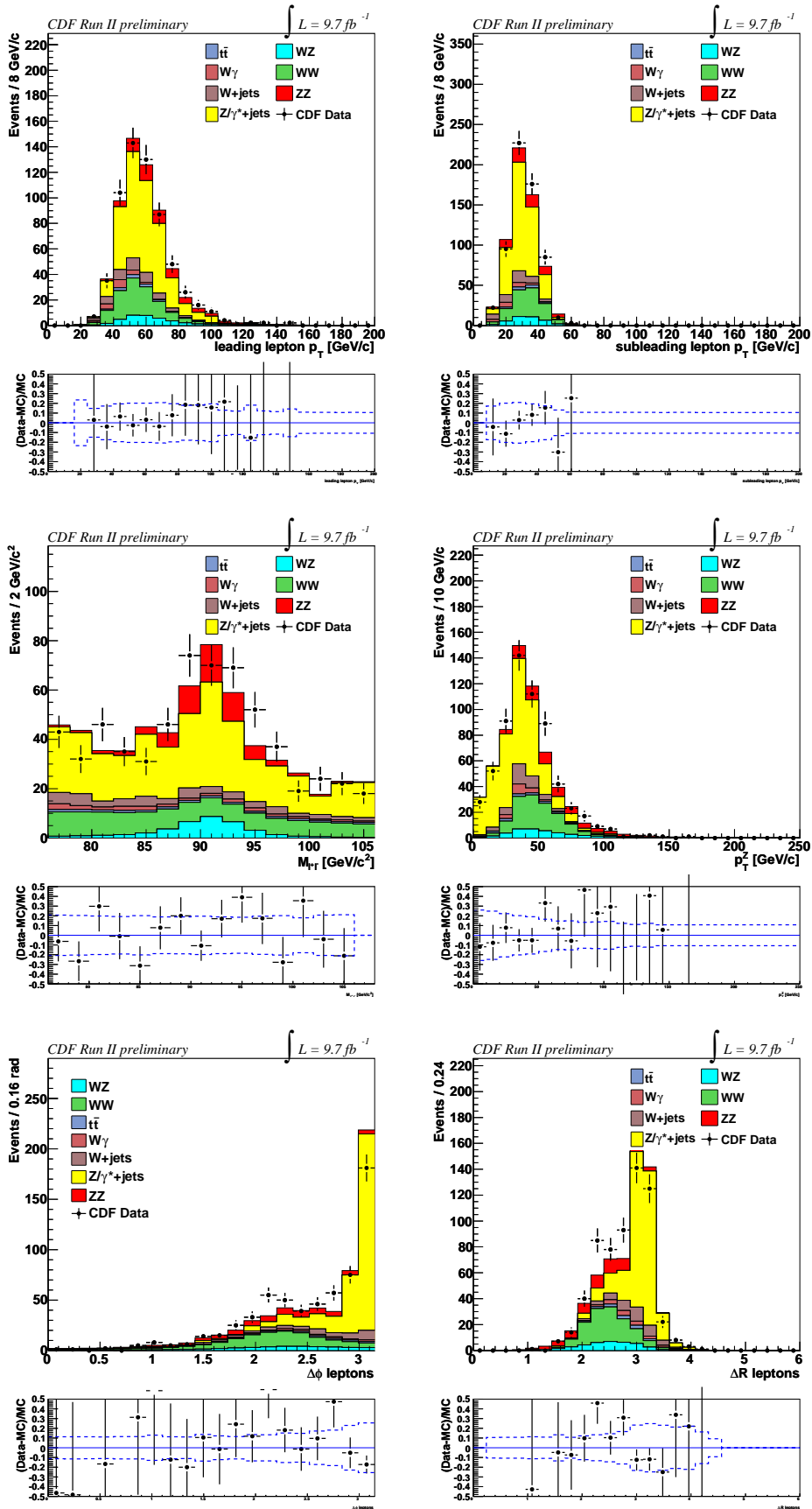


Figure 5.16: Data-to-simulation comparison of some relevant kinematic variable distributions for the events passing the signal region requirements.

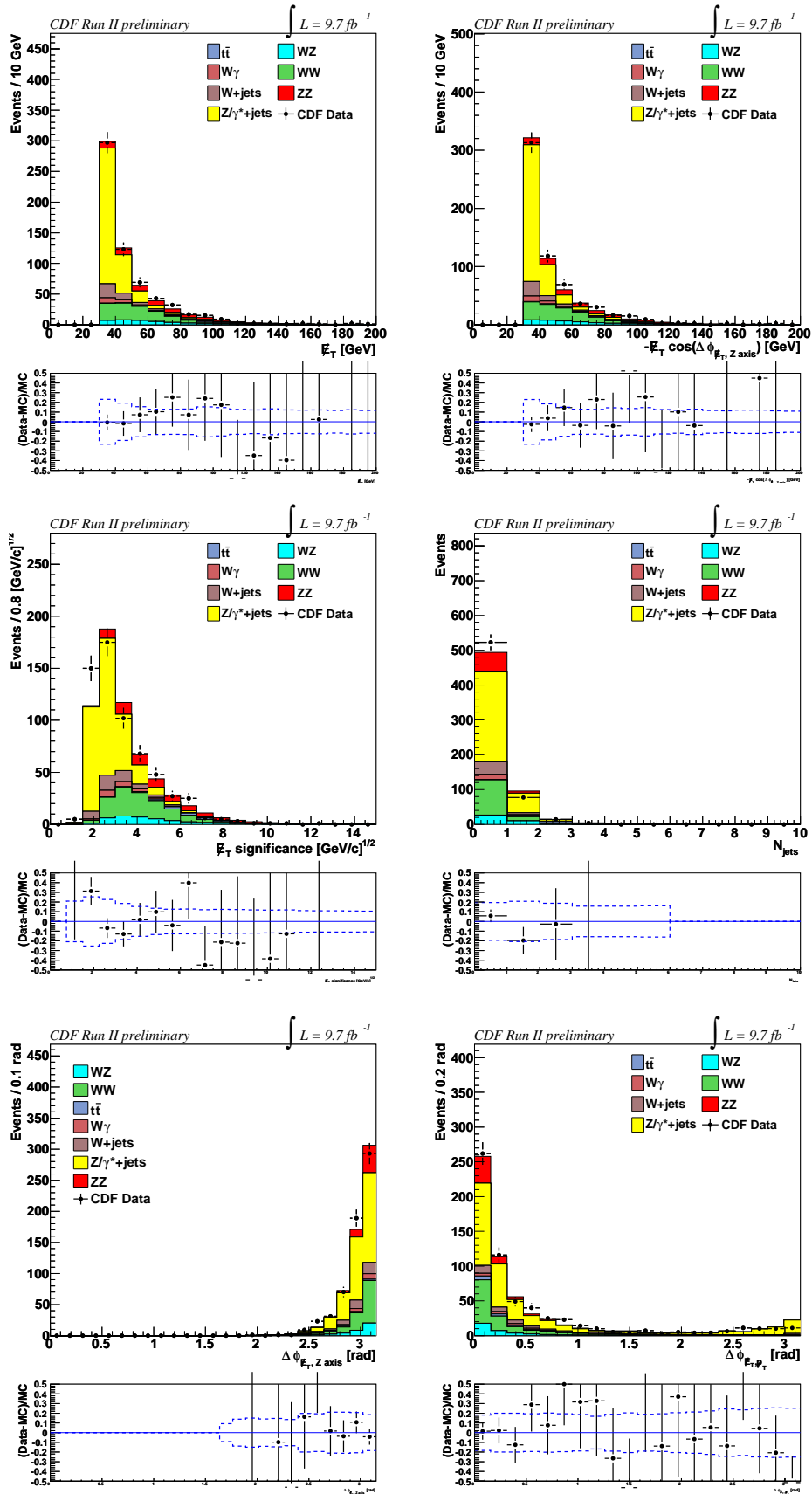


Figure 5.17: Data-to-simulation comparison of some relevant kinematic variable distributions for the events passing the signal region requirements.

### 5.2.3 Test of the Background Modeling in a Control Region

The main background contribution in the previously defined Signal Region is due to Drell-Yan production, which has a similar signature in the final state but a production cross section much higher than that of  $ZZ$ . We know that this process might present a non-negligible amount of missing transverse energy which doesn't reflect the production of undetected particles, but is mostly due to the detector response, which is hard to be modeled in the MC simulation. Secondly, the  $WW$  production constitute also an significant background contribution in the Signal Region. This process is reproduced using a full Next-to-Leading Order simulation and we expect to have a good modeling of the most relevant variable considered (e.g.  $\cancel{E}_T$ ). Drell-Yan simulation will be tested in a non-overlapping data sample of events with a small amount of  $\cancel{E}_T$ , while  $WW$  modeling is tested in a  $e\mu$  sample.

#### Drell-Yan Control Region

The modeling of Drell-Yan events is tested comparing simulation to the data in a control sample of experimental data not overlapping with the Signal Region defined in Section 5.2.2, chosen to have similar kinematic properties as that one and being Drell-Yan dominated. In this way we can extract indirectly information on the modeling of such background process in the Signal Region. The considered Control Region is obtained applying all the Signal Region selections (see Section 5.2.2) but modifying the request on the  $\cancel{E}_T^{Ax}$ , i.e. requiring:

- $\cancel{E}_T^{Ax} \leq 25$  GeV

Table 5.8 summarizes the composition of this sample, with the different contribution from the processes considered. The number of observed events is in agreement, within the systematic uncertainties, with the total predicted number of events from MC simulations.

$\int \mathcal{L} = 9.7 \text{ fb}^{-1}$	
Process	candidate events
DY	699619 $\pm$ 113219
$t\bar{t}$	0.4 $\pm$ 0.1
$W$ +jets	9965 $\pm$ 33.2
$W\gamma$	7.8 $\pm$ 0.5
$WW$	63.5 $\pm$ 1.0
$WZ$	25.3 $\pm$ 0.7
$ZZ$	28.0 $\pm$ 0.7
<b>Total Expected</b>	709709 $\pm$ 113225
<b>Data</b>	725799

Drell-Yan Control Region

Table 5.8: Expected and observed number of events in  $9.7 \text{ fb}^{-1}$  in the Drell-Yan Control Region. The errors include statistical and systematic uncertainties added in quadrature.

The most relevant kinematic variable distributions of the events in the Drell-Yan Control Region for data and Monte Carlo are shown in Figures 5.18 and 5.19. The simulation reproduces the kinematic variables in this control sample with disagreement of the order of 10% or less, which is compatible with the uncertainties that affects the simulations used, discussed in Section



5.2.5. Since the Drell-Yan represents the dominant background to the  $ZZ$  signal extraction, we exploit the data-to-simulation comparison in this control region to evaluate additional systematic uncertainties on the kinematic modeling of this process, as discussed later in this Chapter.

### $WW$ Control Region ( $e\mu$ sample)

To test the  $WW$  modeling in a kinematic region similar to the Signal Region we select events with two isolated leptons in the final state of different flavor, i.e.  $e^\pm\mu^\mp$ , satisfying all the requirements that define the Signal Region, but a different dilepton invariant mass range. Selecting different flavor leptons we drastically reduce contribution from real  $Z$ , i.e. the  $ZZ$  signal contamination and the Drell-Yan production<sup>28</sup>. Since in this case the two leptons are not the products of the  $Z$  decay, we broaden the range of the considered dilepton mass spectrum, to increase the statistic of the control sample:

- $40 \leq M_{e\mu} \leq 140 \text{ GeV}/c^2$

Table 5.9 summarizes the number of events expected from the several processes and the yields in the collected data.

$\int \mathcal{L} = 9.7 \text{ fb}^{-1}$	
Process	candidate events
DY	171.1 $\pm$ 27.7
$t\bar{t}$	23.2 $\pm$ 4.3
$W$ +jets	93.0 $\pm$ 24.7
$W\gamma$	49.3 $\pm$ 6.3
$WW$	265.7 $\pm$ 24.7
$WZ$	6.6 $\pm$ 0.9
$ZZ$	0.7 $\pm$ 0.1
<b>Total Expected</b>	609.6 $\pm$ 45.2
<b>Data</b>	538

$e\mu$  Control Region

Table 5.9: Expected and observed number of events in  $9.7 \text{ fb}^{-1}$  in the  $e\mu$  Control Region.

The most relevant kinematic variable distributions of the events in the  $e\mu$  Control Region for data and Monte Carlo are shown in Figures 5.20 and 5.21. No significant discrepancy is noticeable in the data-to-simulation comparison, with uncertainties dominated by the limited statistics of the sample considered. Once tested in this sample, we assume that the simulation properly models the  $WW$  background also in the Signal Region.

<sup>28</sup>A residual contribution of Drell-Yan events can come from  $Z \rightarrow \tau\tau$  decays with subsequent leptonic decays of the  $\tau$ s that can produce a  $e-\mu$  pair.

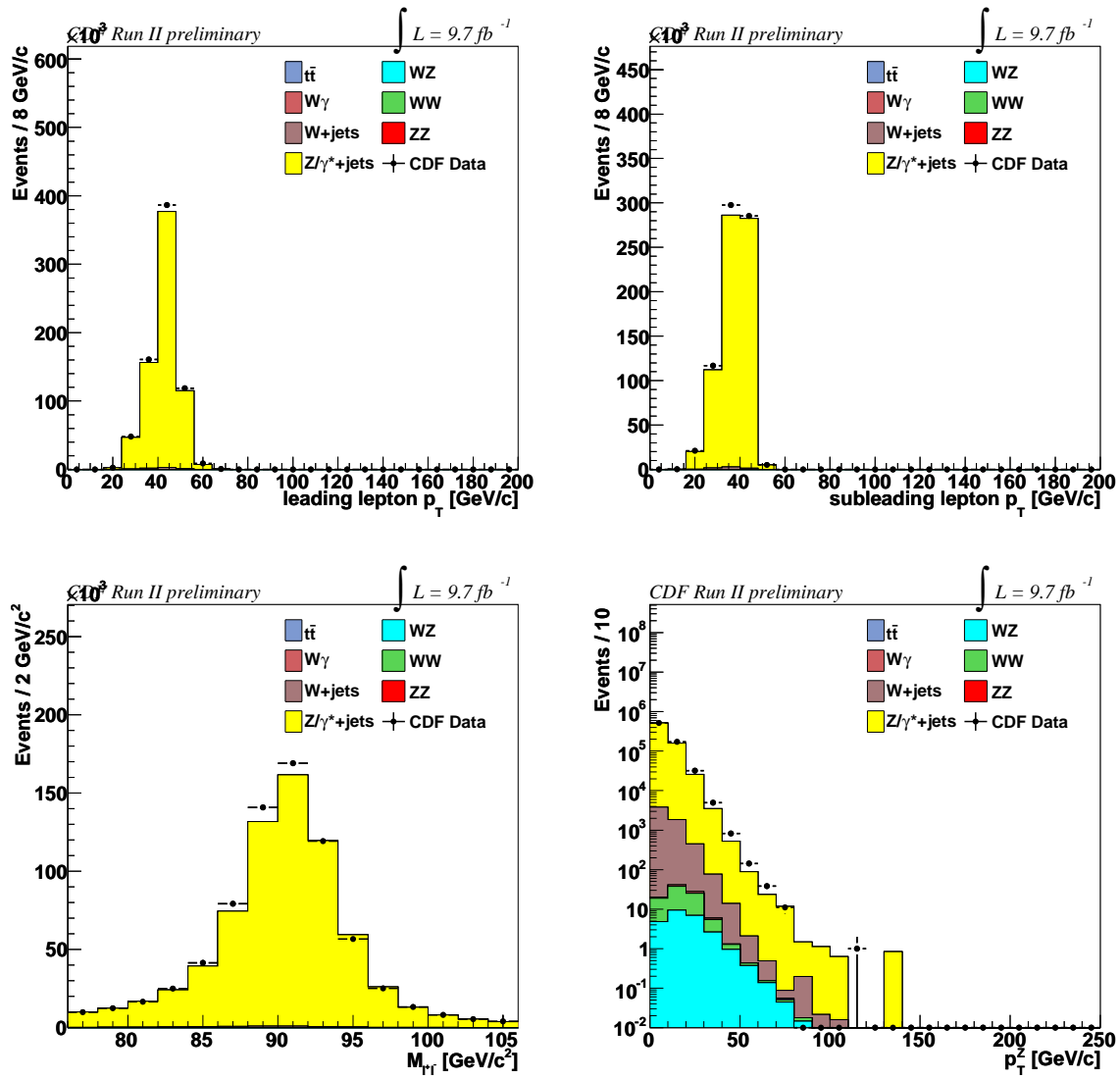


Figure 5.18: Kinematic variables distributions in the Drell-Yan Control Region: simulated predictions compared to data.

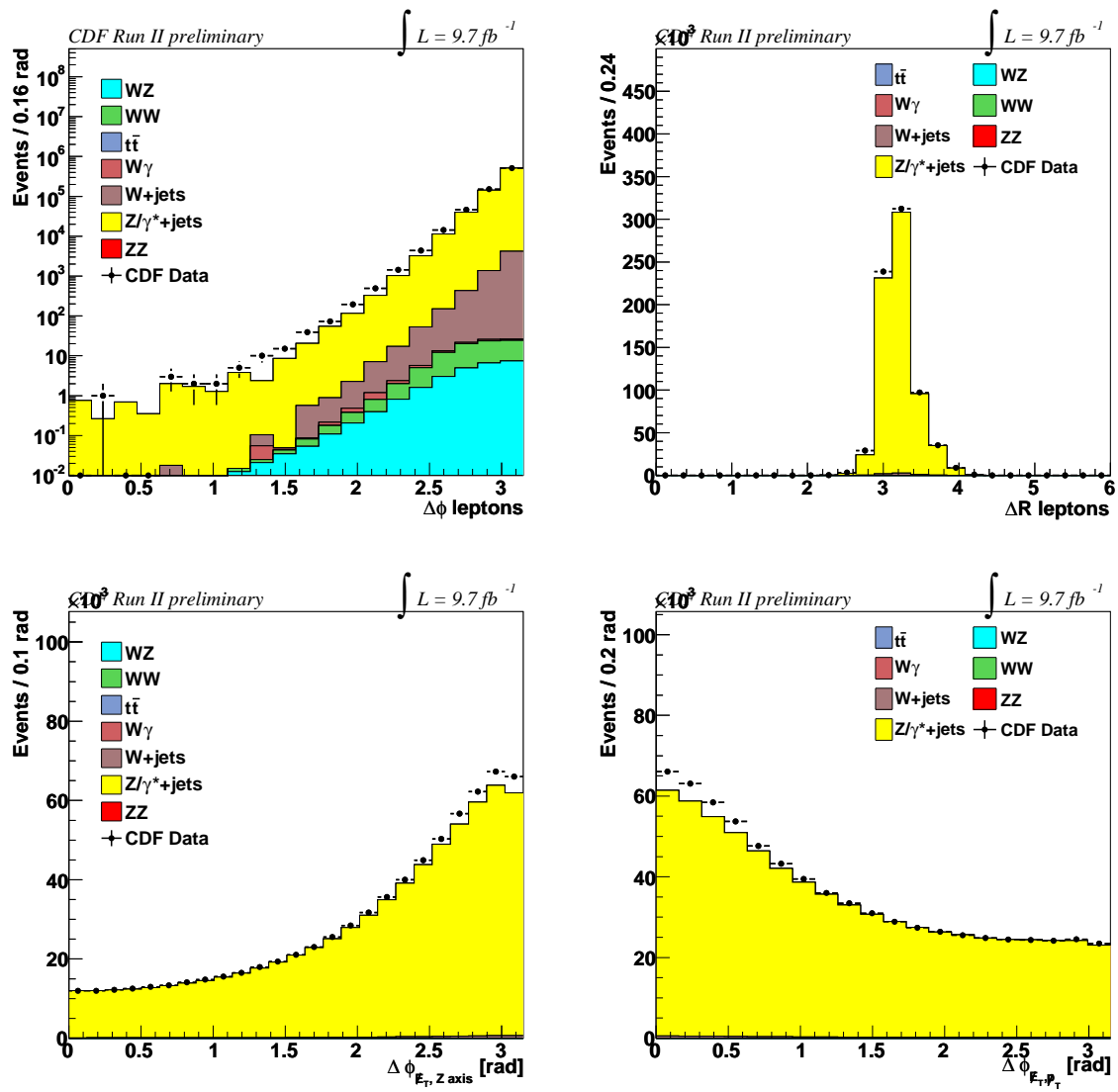


Figure 5.19: Kinematic variables distributions in the Drell-Yan Control Region: simulated predictions compared to data.

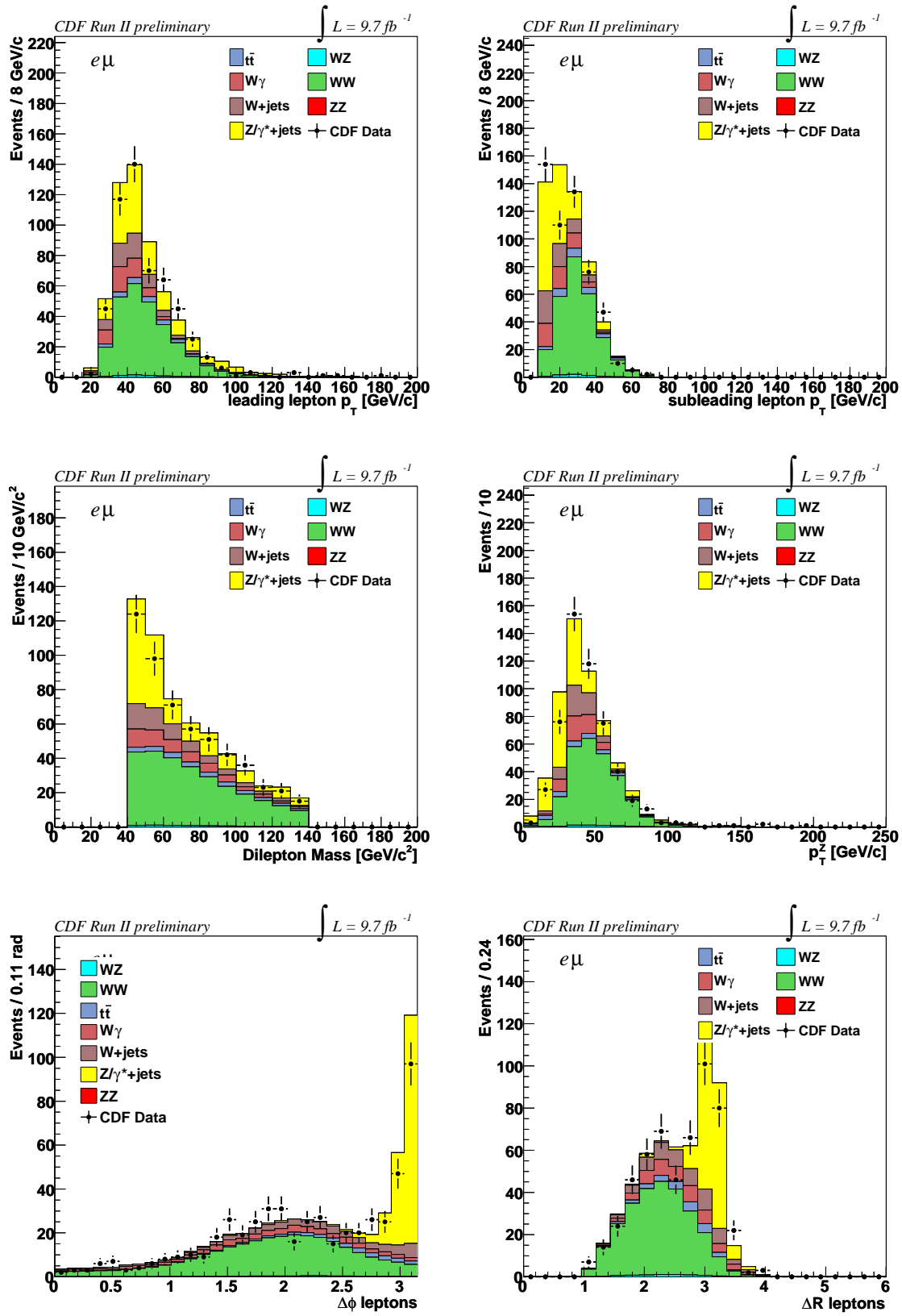


Figure 5.20: Kinematic variables distributions in the  $e\mu$  Control Region: simulated prediction compared to data.

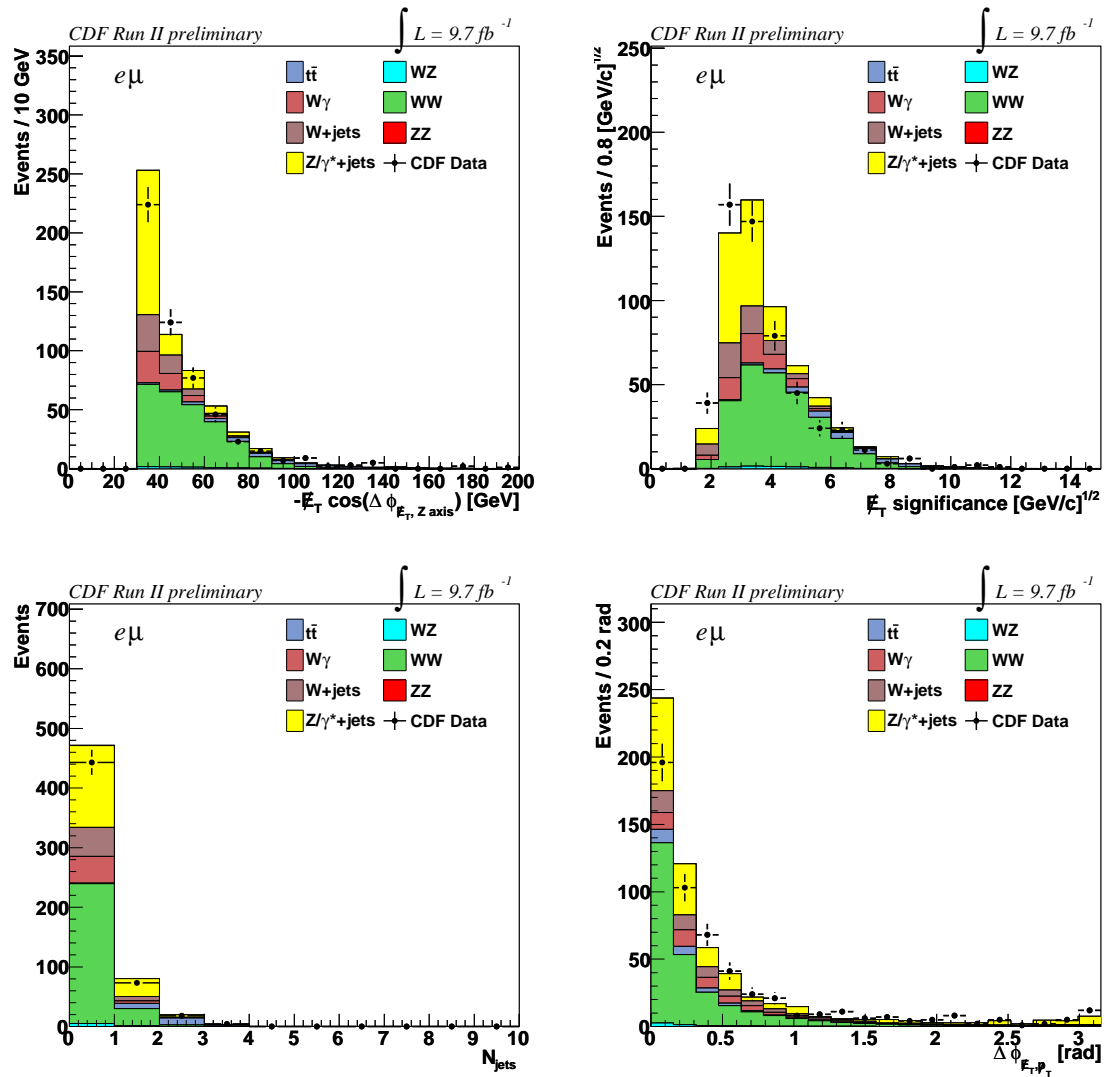


Figure 5.21: Kinematic variables distributions in the  $e\mu$  Control Region: simulated prediction compared to data.

### 5.2.4 Artificial Neural Network for Signal Extraction

After we apply the cuts described in sections 5.2.1 and 5.2.2, the signal over background ratio is still not optimal to measure the cross section simply with a counting experiment. We therefore use a Multivariate Analysis Technique (MVA) to improve the separation of ZZ events from background. This is done using a NeuroBayes<sup>®</sup> [86] Neural Network (NN) [87], a self-learning machine that exploits signal and background kinematic properties to get the best separation between the two samples.

Artificial Neural Networks (NN) are a known method to combine several information in a single discriminant. The principle is to determine a function  $f$  of the input variables of each event that statistically maximize the separation between two event samples, chosen in this case to be the signal and background processes. Once extracted this function, the value it assumes for each event is the NN output. In the best scenario we would like the NN to decide whether an input event is signal-like or background-like, giving an output of +1 or -1 respectively. In its application to real cases a continuous distribution of *scores* in the range  $[-1,+1]$  is obtained, thus expressing the probability for each event of being signal or background. The Neural Network provides a robust approach to determine the  $f$  function.

The NeuroBayes is a *feed-forward* Neural Network, schematically represented in Figure 5.22. The function  $f$  is composed by symmetric sigmoid functions  $f_i : \mathcal{R} \rightarrow \mathcal{R}$ , of the form

$$f_i(x) = \frac{2}{1 + e^{-x}} - 1, \quad (5.18)$$

where  $x$  is the input value of the  $i$ -th node and  $f_i(x)$  is the output. These are smooth step functions that map  $[-\infty, +\infty]$  into  $[-1,+1]$ . In Figure 5.22 each line represents the flow of the information from each node to another: the output information from nodes in one layer are used as input into the nodes of the following layer. Each node considers as input a linear combination of the several inputs represented by different lines *feeding* that node. Each node considers as input value  $x = \sum_i w_i l_i$ , where  $l_i$  are the several input values and  $w_i$  are weights that represent the adjustable parameters describing the behavior of the NN. The network is divided in different *layers* representing the steps through the two samples separation. The nodes in the same layers doesn't interact, therefore there's no exchange of information among them: this is characteristic of the feed-forward NN. The nodes of the first layer have a single input, which is the value of the measured kinematic variables. The second layer is called *hidden* layer since it communicates only with the input and output layers of the network. The third and last layer of the neural network we use has just one node, which gives the NN output value.

The prerogative of the feed-forward neural network is the existence of a relatively simple algorithm to evaluate the best weights  $w_i$  of the function  $f_i$  based on empirical examples. In this case we can exploit the simulated description of signal and background processes to establish which one is more likely an event to come from. This algorithm is called *back-propagation* [87] and uses a gradient-descending algorithm to minimize the squared distance between the NN output values and the target values for these outputs (generally -1 for background-like and +1 for signal-like events). This minimization is called *training* of the NN and relies on MC simulation of the several processes involved.

The NeuroBayes package used in this analysis internally checks the robustness of the training with respect to unphysical values of the input variables. It further assures the training is not influenced by statistical fluctuation of the samples exploited. An schematic example of the output distribution of a NN is shown in Figure 5.23.

Among the many kinematic variables we selected as input for the NN those that provide the

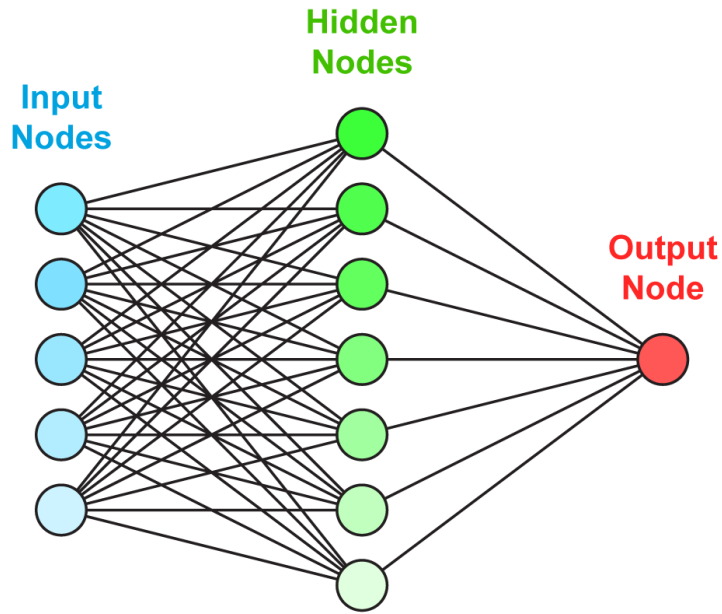


Figure 5.22: The Neural Network structure, consisting of  $N$  input nodes,  $N + 2$  hidden nodes and a single output node.

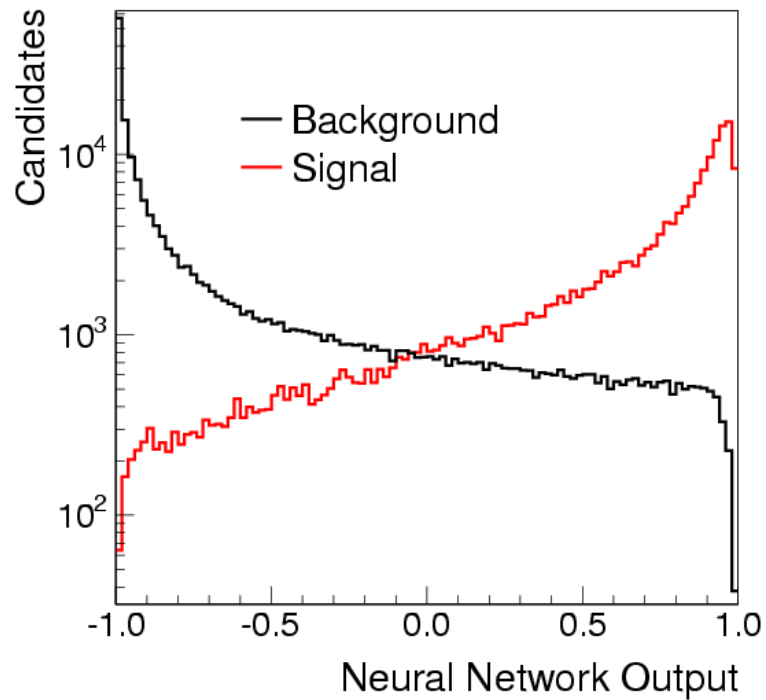


Figure 5.23: Example of the Neural Network output distribution for signal and background sample.

best signal to background separation which are sorted here according to their significance in the neural network training:

- $\Delta\phi(\ell\ell)$ : the opening angle in the transverse plane between the two leptons;

- $p_T(\ell_1)$ : the transverse momentum of the leading lepton;
- $M_{\ell\ell}$ : the dilepton invariant mass;
- $\cancel{E}_T^{Sig}$ : the  $\cancel{E}_T$  significance, defined as  $\cancel{E}_T/\sqrt{\sum_i E_T^i}$ ;
- $N_{jets}$ : the number of reconstructed jets in the event;
- $\Delta\phi(\cancel{E}_T, \not{p}_T)$ : the angle in the transverse plane between the direction of the  $\cancel{E}_T$  and the direction of the track momentum imbalance  $\not{p}_T$ <sup>29</sup>.
- $p_T^Z \equiv p_T(\ell\ell)$ : the dilepton system transverse momentum;

The distributions of these variables are shown in Figures 5.16 and 5.17.

We train the NN to distinguish the ZZ signal from a mixture of background samples, weighted with the relative fractional yields. We split the simulated sample and use 65% of it to train the network and the remaining 35% to test the minimization algorithm, to prevent the dependence on peculiar statistical fluctuation.

The NN output distribution for each simulated process is shown in Figure 5.24 from which we can notice a good separation of the signal from the dominant Drell-Yan background. In particular the neural network pushes the Drell-Yan,  $W$ +jets,  $W\gamma$ ,  $t\bar{t}$  contributions towards low output values while ZZ and  $WZ$  towards high ones.

Figure 5.25 reports the output of the different processes simulated contributions stacked with the proper normalization compared to the experimental data. The plot of the residuals, in particular, shows a good agreement between the data and the predictions.

### 5.2.5 Systematic Uncertainties

As we already discussed in Section 5.1.5, to measure the cross section we need to consider possible source of systematic uncertainties involved in the analysis method described so far. In the analysis strategy just described we model the expected contribution for the signal and background processes using Monte Carlo simulation: with respect to the analysis done in the  $\ell\ell\ell'\ell'$  decay mode, we have to take into account not only the possible sources of systematics that affect the predicted normalization (*rate*) of the MC samples, but also those that modify the *shape* of the relevant kinematic variable distributions. These latter have implications on the information used to train the Neural Network and can affect in particular the NN output distribution for the signal and background predictions.

Some of the sources of uncertainties are the same already discussed in Section 5.1.5; for those we will not report here the detailed description.

**Cross Section** This uncertainty is based on the theoretical uncertainty on the cross section of each process considered. This is evaluated for each process modeled using MC simulation. The  $t\bar{t}$  cross section is calculated at NNLO in Ref. [88] with an uncertainty of 10%. For the diboson processes, the NLO cross section calculation is done using MCFM [18] and has an uncertainty of 6%. The Drell-Yan cross section is known at NNLO and the associated error is 5% [82]. The quoted uncertainty for each process includes the effect of both normalization scale and PDF variation.

<sup>29</sup>The track momentum imbalance  $\not{p}_T$  is the result of the vectorial sum of the tracks reconstructed in the calorimeter, i.e.  $\not{p}_T \equiv -\sum_i \vec{p}_T^i$ .



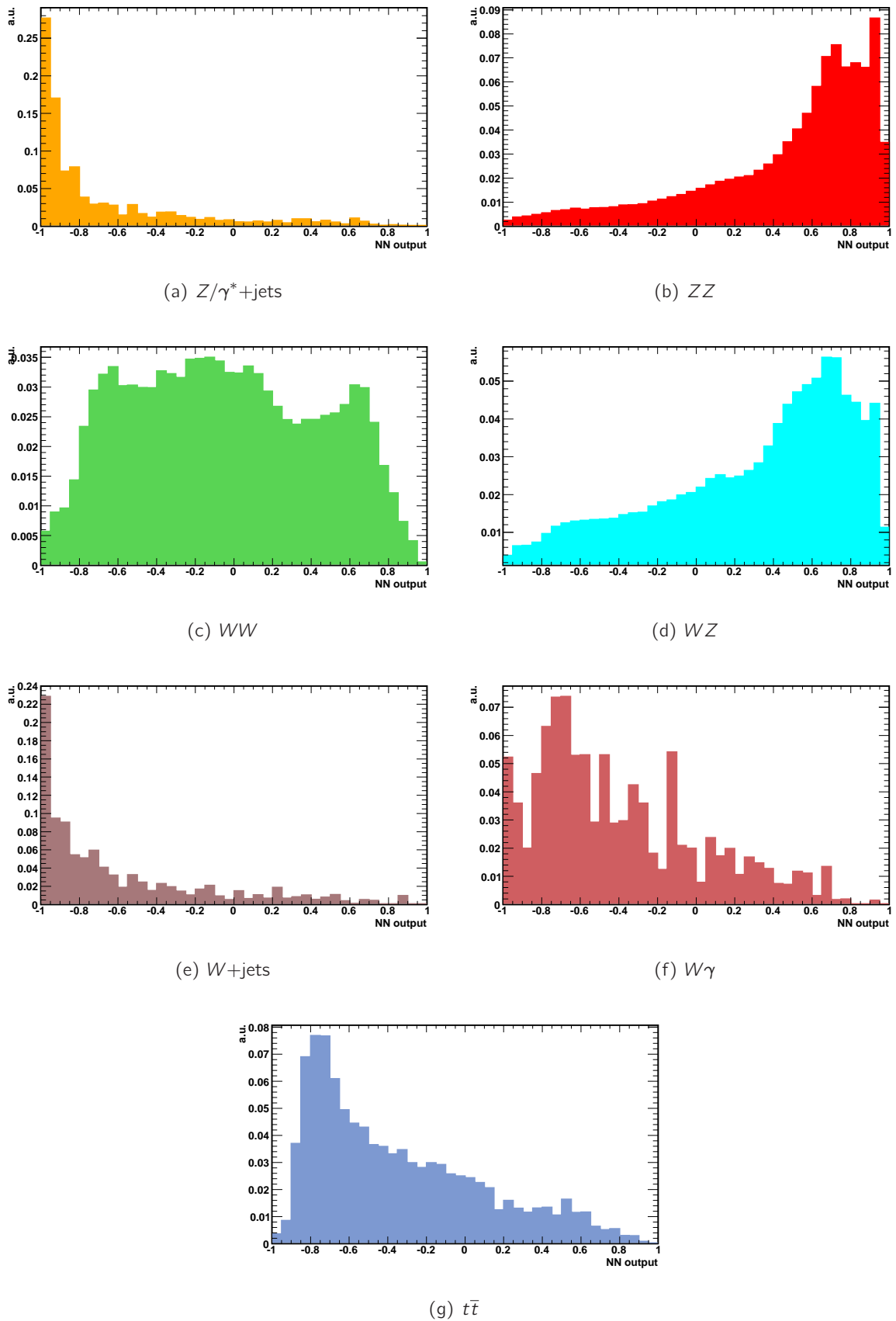
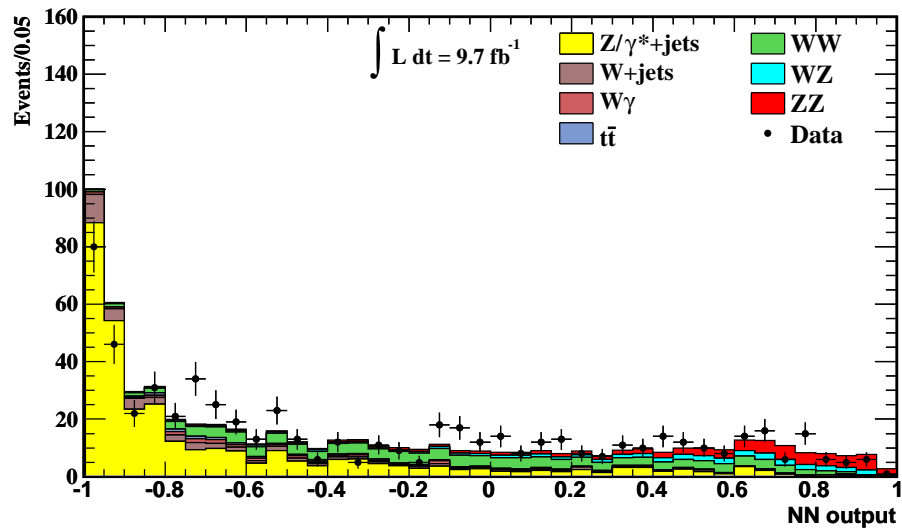
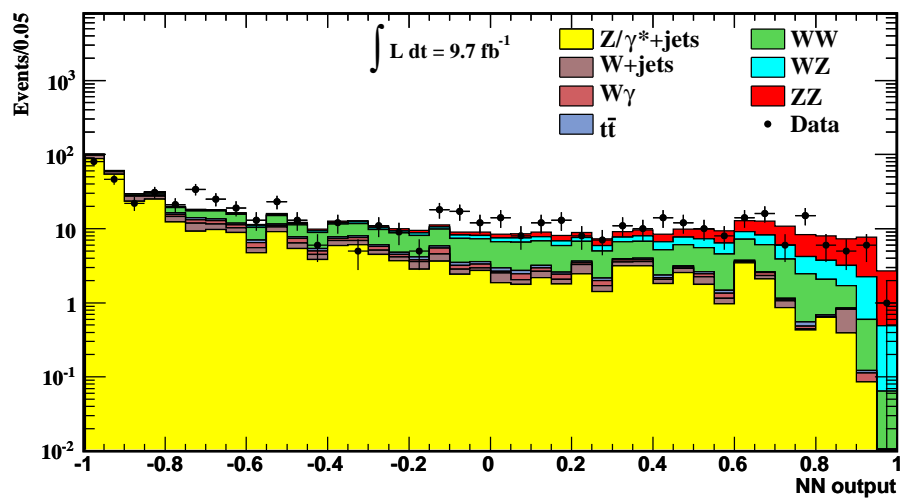


Figure 5.24: Neural network output distribution for the different processes, normalized to unit area.



(a)



(b)

Figure 5.25: Neural network output distribution for signal (in red), background expectations, with the data (dots) superimposed in (a) linear and (b) logarithmic scale.

**MC Run dependence** This systematic has been introduced to take into account the fact that the  $t\bar{t}$  Monte Carlo sample has been generated with no complete time dependency. Hence a small systematic effect can be due to the difference in the detector efficiency in different online data-taking conditions (mainly instantaneous luminosity). This difference has been evaluated using simulations for several processes, comparing the acceptance with detector conditions in the first and last periods of the data-taking<sup>30</sup>: this has been found to be of the order of few percent. A conservative 10% uncertainty is considered to include a possible systematic error from this effect.

**PDF Model** Calculated in the same way described in Section 5.1.5.

**Higher Order Diagram** Calculated in the same way described in Section 5.1.5.

**Luminosity** Calculated in the same way described in Section 5.1.5.

**Photon conversion** This systematic uncertainty takes into consideration the effect of photon conversions in the tracking detector material. This effect is considered only in the  $W\gamma$  sample since is expected to have a negligible effect in all the other processes. This photon conversion effect is included in the standard detector simulator but is known to present some discrepancies with respect to data. Using a di-photon sample we generate a *map* of the detector material reconstructing the vertex of the  $\gamma \rightarrow \ell^+\ell^-$  conversion. We evaluate the difference in the conversion vertex distribution between the experimental data and a simulation. This difference is included in the analysis as a systematic error of 10%.

**Jet Modeling** The uncertainties on the reconstruction of jets in the events can affect the acceptance of the analysis, in particular what concerns the reconstructed jet energy scale (JES). The uncertainty on the JES affects the number of reconstructed jets in the event (which can rise above the  $E_T$  threshold for the jet identification definition with JES increase and vice-versa), as well as the  $\cancel{E}_T$ , which contains informations on the hadronic activity in the event. The 2-way variation of the jet energy by its resolution uncertainty is propagated in the analysis to evaluate the effect on the expected number of events for the several processes. This has been included only for the processes but the  $W$ +jets which is extracted from experimental data.

**Jet Fake Rates** Calculated in the same way described in Section 5.1.5.

**Lepton ID efficiencies** Calculated in the same way described in Section 5.1.5.

**Trigger Efficiencies** Calculated in the same way described in Section 5.1.5.

**NN Output distribution** The mismodeling of any kinematic variable in the simulation of a given process (with respect to data) can reflect in the training of the Neural Network and lately in the shape of the NN output. Even if the NN separates quite well the background processes from the  $ZZ$  signal, a systematic variation of the shape of the NN output distribution can give a wrong estimate of the tail of background contribution in the high-NN output region, which would significantly affect the cross section measurement<sup>31</sup>.

<sup>30</sup> First and last periods of the data-taking represents the two extremal conditions both for the instantaneous luminosity and the detector aging effects.

<sup>31</sup> We focus only on the NN output distribution because this is the discriminant quantity from which we will measure the cross section, and includes the effect of the different input kinematic variables.

For this reason we check the NN output distribution for background simulation and data of events in the Control Regions defined in Section 5.2.3. We process the simulated events and the data in these control samples with the NN trained in the Signal Region and compare their NN output value distribution to data in the same samples.

The NN output distribution for events in the Drell-Yan control region is shown in Figure 5.26. From the data-to-simulation comparison we can see a significant difference in the shape of the NN output. From this discrepancy we extract a shape systematic uncertainty given by the bin-by-bin fractional difference between data and simulation. This quantity is assigned as a positive and negative bin-by-bin uncertainty on the Drell-Yan MC prediction in the Signal Region, as shown in Figure 5.27.

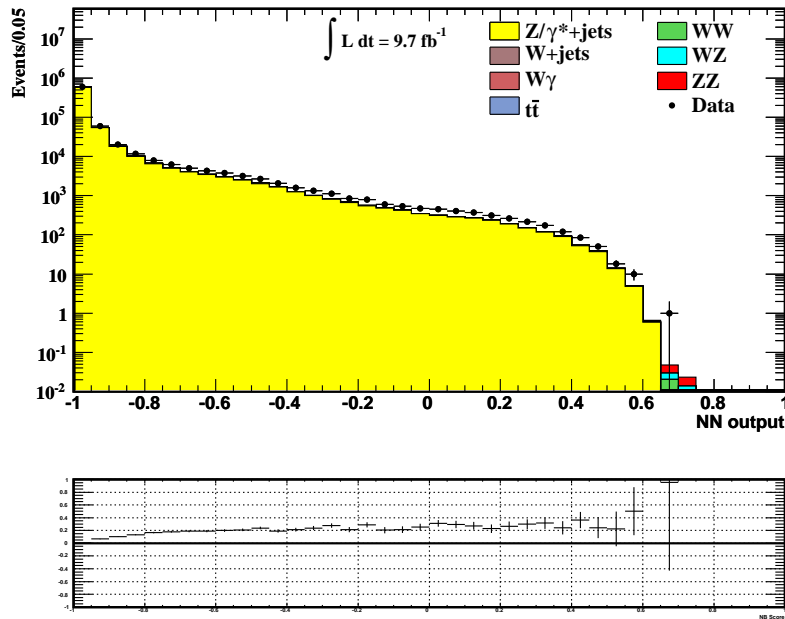


Figure 5.26: Neural Network output distribution for events in the Drell-Yan Control Region.

Figure 5.28 shows the NN output distribution for events in the  $e\mu$  control region. In this sample, dominated by  $WW$  production, we appreciate the good separation of this process toward low values of the NN output and the good agreement between the data and the MC simulation. The shape difference obtained from data-to-simulation comparison is not significant, hence we will not consider any shape systematic uncertainties on the  $WW$  background prediction.

The several sources of systematics that we consider for the  $ZZ$  production cross section are summarized in Table 5.10<sup>32</sup>.

<sup>32</sup>We recall that each row represent a source of systematic uncertainty with the percentual fraction that affects each process; systematic uncertainties in *italic* in the same row are considered 100% correlated, all the others are assumed to be uncorrelated.

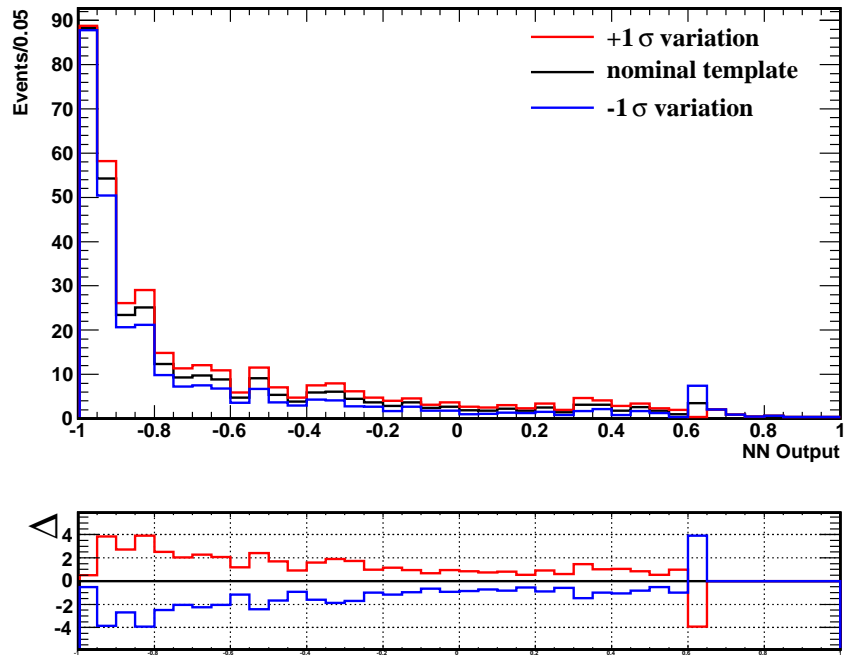


Figure 5.27: Drell-Yan background NN output distribution in the Signal Region, with the shape-varied distribution obtained from the data-to-simulation comparison in the control region.

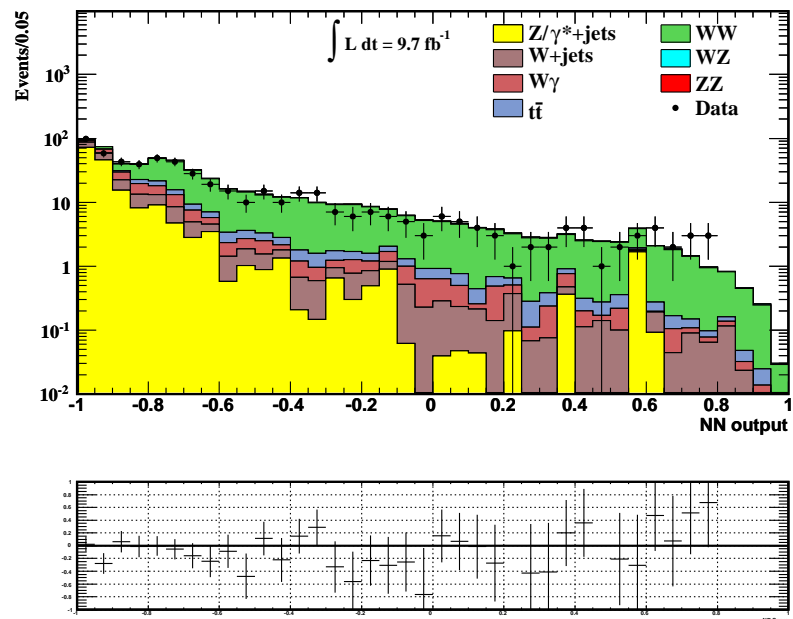


Figure 5.28: Neural Network output distribution for events in the  $e\mu$  Control Region.

Source	ZZ	WW	WZ	t $\bar{t}$	Drell-Yan	Wfl	W+jets
<b>Cross Section</b>		6 %	6 %	10 %	5 %	10 %	
<b>MC Run dependence</b>				10 %			
<b>PDF Model</b>	2.7 %	1.9 %	2.7 %	2.1 %	4.1 %	2.2 %	
<b>Higher-order Diagrams</b>	5 %		5 %	10 %		5 %	
<b>Luminosity</b>	5.9 %	5.9 %	5.9 %	5.9 %	5.9 %	5.9 %	
<b>Photon conversion</b>						10 %	
<b>Jet Modeling</b>	2.0%	1.6%	3.4 %	5.3%	6.2 %	2.0%	
<b>Jet Fake Rates</b>							16 %
<b>Lepton ID efficiencies</b>	3 %	3 %	3 %	3 %	3 %		
<b>Trigger efficiencies</b>	2 %	2 %	2 %	2 %			
<b>NN Output distribution</b>				<i>Shape uncertainty</i>			

Table 5.10: Systematics uncertainties for signal and background that contribute to the final data set for the  $ZZ \rightarrow \ell\ell\nu\nu$  cross section measurement.

### 5.2.6 Cross Section Measurement

We measure the  $ZZ$  production cross section with a binned likelihood fit to the NN output distribution of the data, considering all the possible relevant source of systematic errors described in Section 5.2.5. This is done with the same approach described in Section 5.1.6 for the cross section measurement in the four charged lepton decay channel, but exploiting the information from the NN output distribution.

We define a binned likelihood similar to the one in Equation 5.11, that describes the total posterior probability as the product of the bin-by-bin probabilities:

$$\mathcal{L} = \left( \prod_i \frac{\mu_i^{n_i} e^{-\mu_i}}{n_i!} \right) \cdot \prod_{j=syst.} e^{-\frac{\theta_j^2}{2}} \quad (5.19)$$

where  $\mu_i$  is the total number of events expected in the  $i$ -th bin and  $n_i$  is the number of the observed data events in the  $i$ -th bin.  $\mu_i$  takes into account also for the systematic errors and is given by the formula:

$$\mu_i = \sum_k \alpha_k \left[ \prod_j (1 + f_k^j \cdot \theta_j) \right] (N_k^{Exp})_i . \quad (5.20)$$

In this expression  $f_k^j$  is the fractional uncertainty due to the  $j^{th}$  systematic source for the physics process  $k$ .  $(N_k^{Exp})_i$  is the expected number of events of the physics process  $k$  in the  $i$ -th bin. This construction of the likelihood function takes into account properly the correlations of the systematic errors through the fitted distribution. The parameters  $\alpha_k$  represents the scaling factor of the predicted contribution for a given process: all of them are fixed to 1<sup>33</sup> but  $\alpha_{ZZ}$  which is free to float. Once integrated the likelihood function over all the other parameters (see Equation 5.10), the posterior p.d.f. of  $\mathcal{L}(\alpha_{ZZ})$  will give the cross section measurement in terms of the nominal  $ZZ$  cross section value.

The posterior p.d.f. of the parameter  $\alpha_{ZZ}$  is shown in Figure 5.29(a), from which we obtain as best value for the cross section:

$$\frac{\sigma(p\bar{p} \rightarrow ZZ)}{\sigma_{SM}} = 1.06_{-0.22}^{+0.24}(stat.) \pm 0.12(syst.) = 1.06_{-0.25}^{+0.27} . \quad (5.21)$$

This measurement is in good agreement with the SM NLO prediction and the accuracy of the measured cross section is limited by the statistical uncertainty.

The expected (observed) significance of the signal is evaluated by generating a large number of pseudo experiments, with the same technique described in 5.1.6. Figure 5.29(b) shows the distribution of the *test statistic* variable  $T_s$  (defined in Equation 5.15) for background only and signal+background pseudo experiments, as well as the value evaluated on the data and the median of the signal-like pseudo-experiments. From 100k pseudo-experiment we found that the investigated signal has an observed significance greater than  $4.5 \sigma$ .

<sup>33</sup> This means we assume that the predictions for these processes are the ones obtained from the simulation with the proper normalization.

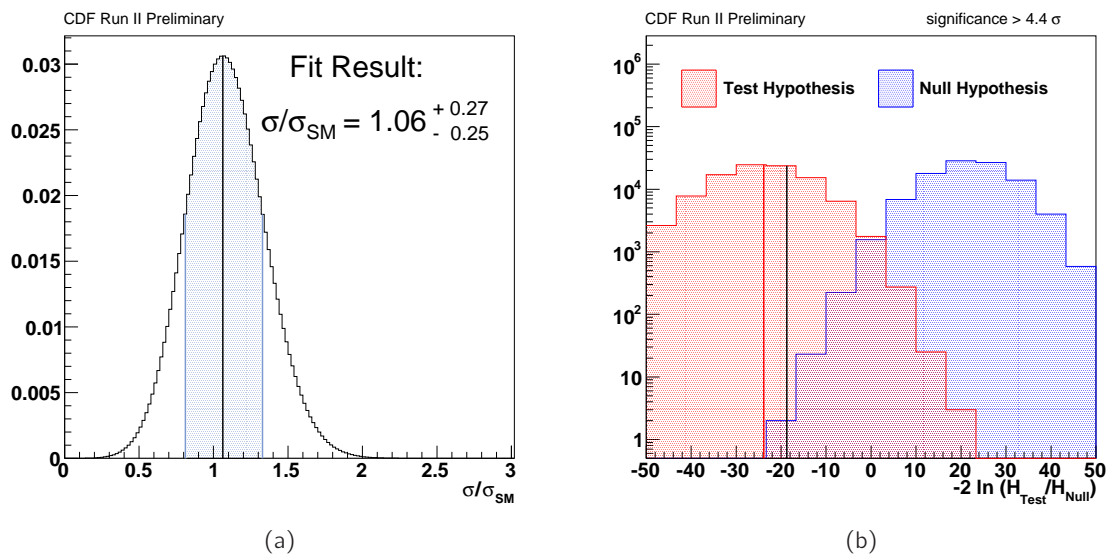


Figure 5.29: (a) Posterior p.d.f. of the parameter  $\alpha_{ZZ}$  after the integration over all the other parameters. The blue band represents the  $\pm 1\sigma$  confidence interval. (b) Test statistic distribution for signal-like (red) and background-only (blue) pseudo-experiments. The black line represents the observed test statistic value for the data.



# Standard Model Higgs Search in four lepton final state

## Contents

6.0.7	Event Reconstruction and Sample Composition . . . . .	110
6.0.8	Background Determination . . . . .	115
6.0.9	Event kinematics . . . . .	117
6.0.10	Systematic Uncertainties . . . . .	121
6.0.11	Limit Calculation on the Higgs production cross section . . . . .	122
6.0.12	Results . . . . .	124

The production of a pair of massive gauge bosons is a really interesting sector of the electroweak interactions because it is sensitive to deviations and anomalies with respect to the Standard Model of particle physics. As already mentioned in Section 2.3, given the relatively large mass of  $W$  and  $Z$  bosons, production of diboson is affected by modification of particle interactions, usually caused by new physics phenomenology. In particular, several simpler extensions of the Standard Model predict the existence of a massive particle  $X$  (of mass  $m_X$ ) which couples with SM particles proportionally to their mass. This may lead to strong interactions of these new particles with  $W$  and  $Z$  bosons, through production processes and decays often resulting in an experimental signature similar to the SM diboson production. One of the simplest examples is the Higgs boson<sup>1</sup>, *the most recently discovered particle*, of which the properties are still under investigation. The electroweak spontaneous symmetry breaking predicts the existence of this scalar particle, which, depending on the value of its mass ( $m_H$ ), decays with different relative strength to several particle-pair. Most of the interesting decay channels used to study the Higgs boson properties involves dibosons. Once understood properly the SM diboson production, it's interesting to search for hints of this kind of new physics in diboson events. Starting from the sample of events exploited for the  $ZZ \rightarrow \ell^+ \ell^- \ell'^+ \ell'^-$  production cross section measurement<sup>3</sup> we investigate CDF sensitivity to the  $H \rightarrow ZZ \rightarrow \ell \ell \ell' \ell'$  production. According to the SM prediction, a really small contribution from Higgs boson production is expected in the collected data sample; we exploit kinematic properties of this process to test CDF sensitivity and carefully evaluate statistical and systematic uncertainties to extract information on the Higgs production. If no

<sup>1</sup> See details about the Higgs boson in Chapter 1.

<sup>2</sup> For brevity we will often refer to  $ZZ \rightarrow \ell^+ \ell^- \ell'^+ \ell'^-$  as  $ZZ \rightarrow \ell \ell \ell' \ell'$ , omitting the charge specification.

<sup>3</sup> As specified in Chapter 5, we consider in this measurement  $\ell=e, \mu$ .

significant deviation is seen with respect to the predicted SM  $ZZ$  production we expect to be able to set an upper limit on the Higgs production cross section. In Section 6.0.7 is described the way events are selected and the expected contribution from several processes involving Higgs boson production. In Section 6.0.8 information about the modeling of the background processes contribution is reported while in Section 6.0.9 are shown the kinematic properties of the selected sample of events. In Section 6.0.10 we discuss the systematic uncertainties on the modeling of the signal and background processes that can affect this analysis, while Section 6.0.11 present the extracted upper limit on the Higgs production cross section.

Among the possible decays of the Standard Model Higgs boson we focus on its decay to a pair of  $Z$  boson with the consequent decay of each  $Z$  to a pair of massive leptons ( $\ell = e, \mu$ ). Even if the  $ZZ$  is not the dominant decay mode<sup>4</sup>, this process is fully reconstructed when considering the four-lepton final state: the good momentum resolution of reconstructed leptons improves the sensitivity in this decay process. A sample of four reconstructed lepton events is similar to the one described in Section 5.1.1, hence dominated by non resonant  $ZZ$  production. From the acquired information on that sample we can extract information on an eventual Higgs signal. At the Tevatron an Higgs boson is produced mostly through the fusion of a pair of gluons, present in the colliding protons and anti-protons. The process  $gg \rightarrow H \rightarrow ZZ \rightarrow \ell\ell'\ell''\ell'''$  should manifest as a resonant production of a  $Z$  boson pair and the four leptons coming from the  $Z$  decays should resonate at the mass of the Higgs,  $m_H$ . Even if a small number of  $H \rightarrow ZZ \rightarrow \ell\ell'\ell''\ell'''$  are expected to be produced (see Table 6.1), these should be detected as a narrow peak in the  $m_{4\ell}$  spectrum, over a continuous distribution of non resonant SM  $ZZ$  events. We search for a signal of a SM Higgs boson in a wide range of possible  $m_H$  values, from 115 to 300 GeV/ $c^2$ <sup>5</sup>. For  $m_H \gtrsim 2 \cdot m_Z$  the Higgs boson can effectively decay to a pair of on-shell  $Z$ , while for  $m_H \lesssim 2 \cdot m_Z$  it can only virtually decay to a pair of bosons, of which at least one has to be off-shell.

Figure 6.1 shows the predicted cross section for  $gg \rightarrow H$  at the Tevatron and LHC as well as the expected produced events of  $H \rightarrow ZZ \rightarrow \ell\ell'\ell''\ell'''$  produced in the CDF dataset, compared to the ones in 1 fb<sup>-1</sup> at the LHC.

### 6.0.7 Event Reconstruction and Sample Composition

At the Tevatron a Higgs boson is mainly produced through gluon fusion, but for a smaller fraction it is also produced in association with a vector boson ( $W, Z$ ) or through vector boson fusion (VBF). Considering the several possible decay of the Higgs boson, different processes can give a four-lepton signature in the final state, as shown in Figures 6.2.

The production of the Higgs through gluon fusion or VBF and its subsequent decay to a pair of  $Z$  bosons can give four resonant leptons in the final state when occurring both  $Z \rightarrow \ell\ell$  decays. The  $H \rightarrow ZZ \rightarrow \ell\ell'\ell''\ell'''$  decay can occur also when the Higgs is produced in association with a  $W$  or  $Z$  boson ( $VH \rightarrow VZZ$ ). Each of this processes can give four (or more) leptons in the final state, thanks to leptonic decays of the associated particle.  $ZH \rightarrow ZZZ$  can give four leptons in the final state either from the two  $Z$ s from the Higgs boson decay, or two from one of such  $Z$ s and other two from the  $Z$  produced in association to the Higgs boson<sup>6</sup>.  $WH \rightarrow WZZ \rightarrow \ell\nu\ell'\ell''\ell'''\ell''''$  contributes to the considered sample when one of the five leptons is not detected. The four identified leptons can resonate at the mass of the Higgs boson or not, depending on the fact that the missed lepton

<sup>4</sup> As seen in Section 1.2 for low values of the Higgs mass,  $m_H \lesssim 135$  GeV/ $c^2$ , the dominant decay channels are  $b\bar{b}$  and  $\gamma\gamma$ , while for higher mass values,  $m_H \gtrsim 135$  GeV/ $c^2$ , the decay to a  $W$  boson pair dominates.

<sup>5</sup> This is the Higgs mass range most favorable from experimental indirect constraints.

<sup>6</sup> Given the  $Z \rightarrow \ell\ell$  branching fraction a tiny fraction of  $ZH \rightarrow ZZZ$  events result in a six-lepton final state, hence we neglect this contribution in the discussion.

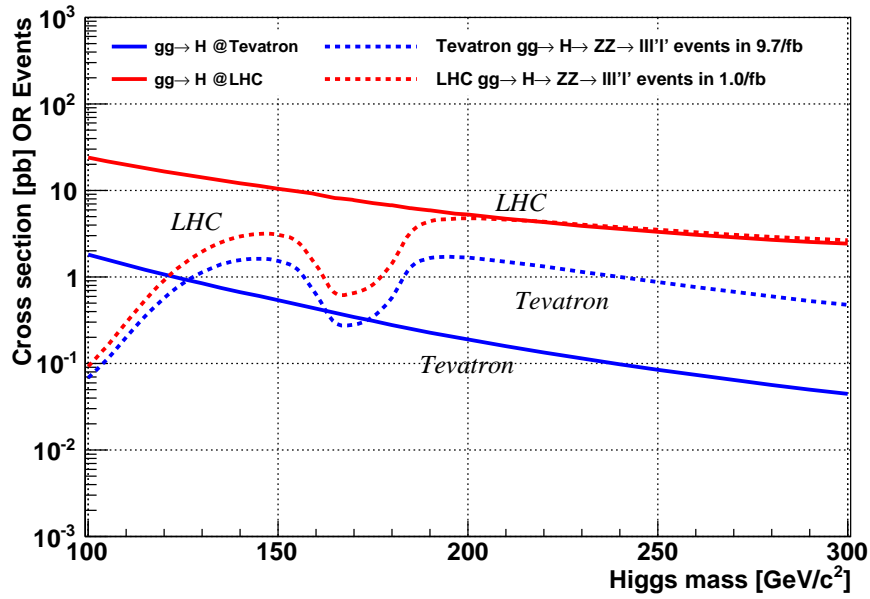


Figure 6.1: Predictions for Higgs production and decay to  $ZZ$  at the Tevatron and LHC.

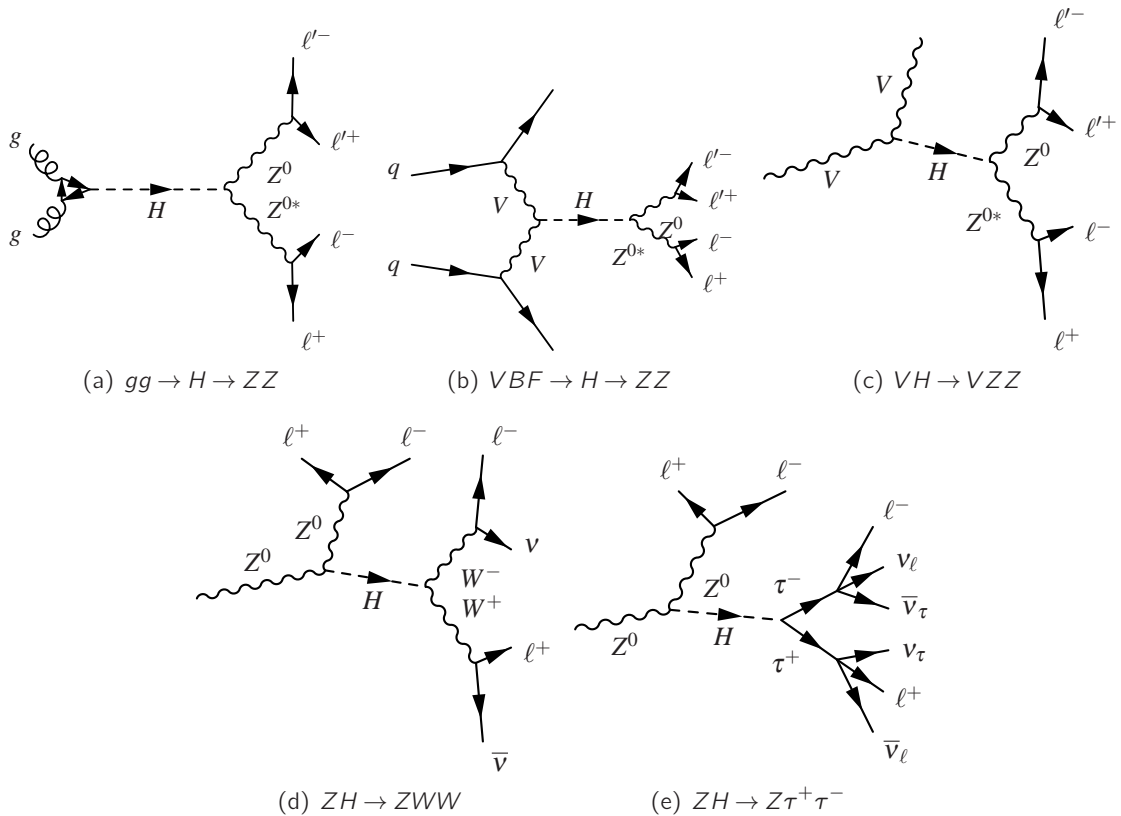


Figure 6.2: Feynman diagrams for the Higgs production processes that can give four charged leptons in the final state.

is the one from the  $W$  or from one of the  $Z$ s. At last,  $ZH \rightarrow ZWW \rightarrow ll' \nu \ell'' \nu$  decay gives four non-resonant charged leptons in the final state, as well as  $ZH \rightarrow Z\tau\tau \rightarrow ll' \nu_{\ell'} \nu_{\tau} \ell'' \nu_{\ell''} \nu_{\tau}$ .

All these processes involve the production of a Higgs boson, and will be the object of our search. Depending on the Higgs boson mass, their relative contributions vary, according to the SM prediction for the Higgs production cross section and branching ratios, reported in Table 6.1.

We model these processes using a MC simulation generated with PYTHIA[76], with a similar procedure as the one described in Section 5.1.2. We test different mass hypothesis by generating several sets of Higgs signal samples which differs only by the value of  $m_H$ . We probe the range  $120 \leq m_H \leq 300 \text{ GeV}/c^2$  in  $10 \text{ GeV}/c^2$  steps and we consider additional  $5 \text{ GeV}/c^2$  steps in the range  $120 \leq m_H \leq 200 \text{ GeV}/c^2$  where the CDF experiment expects to be more sensitive to the Higgs signal. We normalize the Higgs signal simulation using the best available NNLO calculation of the corresponding cross section, summarized in Table 6.1. Sample corresponding to gluon fusion simulation are normalized to the cross section calculated in [50, 51], while those corresponding to associated production uses the cross section calculated in [89]. The Higgs decay branching fractions are extracted from [54].

The sample of collected events exploited to search for the a Higgs signal is the same used for the  $ZZ$  cross section measurement in the  $\ell\ell'\ell'$  decay mode, described in Section 5.1.1. The events collected are then required to satisfy the following requirements:

- Exactly 4 leptons ( $e, \mu$ ) of the combination  $e^+e^-e^+e^-, \mu^+\mu^-\mu^+\mu^-, e^+e^-\mu^+\mu^-$ ;
- Each lepton with  $p_T \geq 10 \text{ GeV}$ ;
- At least one trigger lepton having  $p_T \geq 20 \text{ GeV}$ ;
- $\min(\Delta R_{\ell\ell}) > 0.1$ .

These requirements are optimized for the  $gg \rightarrow H \rightarrow ZZ \rightarrow \ell\ell'\ell'$  signal but have relevant acceptance also on Higgs boson production through the aforementioned processes contributing to the final state considered, characterized by the presence of four charged leptons in the final state. That is why we explicitly require that the four leptons can be paired having same flavor and opposite charge, assuming they come from  $Z \rightarrow \ell^+\ell^-$  decay. This requirement is not significant when considering associated production, for example  $ZH \rightarrow ZWW$ : for these processes we are sensitive only for the fraction of events passing the selections described. In the  $ZZ$  cross section measurement described in Section 5.1 we made an explicit requirement on the reconstructed dilepton invariant masses<sup>7</sup>. We release this requirement in the Higgs boson search because of the different event kinematic properties. For a Higgs boson with a mass below the  $ZZ$  production threshold ( $m_H \lesssim 180 \text{ GeV}/c^2$ ) at least one of the bosons has to be produced off-shell for kinematic reasons. Below this threshold most of these are actually three body decays  $H \rightarrow Z\ell\ell$  with subsequent  $Z \rightarrow \ell'\ell'$  decay; this will give a pair of off-shell Zs. This effect can be seen in Figure 6.3, where is shown the scatter plot of the two reconstructed masses ( $M_{\ell\ell}(1,2)$ ), comparing the SM  $ZZ$  production and the  $H \rightarrow ZZ$  signal, for  $m_H \lesssim$  and  $m_H \gtrsim 180 \text{ GeV}/c^2$ . While for SM  $ZZ$  production at least one of the two reconstructed  $Z$  is on-shell, a large fraction of  $H \rightarrow ZZ$  decays (for  $m_H \lesssim 180 \text{ GeV}/c^2$ ) has both reconstructed lepton pairs with an invariant mass smaller than the nominal  $m_Z$ . By applying any requirements on  $M_{\ell\ell}$  we would drastically reduce the acceptance on these decay, decreasing the sensitivity for Higgs boson signal in the mass range  $\lesssim 180 \text{ GeV}/c^2$ .

In the sample selected applying the aforementioned requirements we expect the contribution from the signal and background processes summarized in Table 6.2. Table 6.3 and Figure 6.4 show the expected contribution from the several signal processes as a function of the Higgs mass.

<sup>7</sup> For the cross section measurement analysis optimization we require  $76 \leq M_{\ell\ell-1} \leq 106 \text{ GeV}/c^2$  and  $40 \leq M_{\ell\ell-2} \leq 140 \text{ GeV}/c^2$ .

$m_H$ [GeV/ $c^2$ ]	$\sigma(gg \rightarrow H)$ [fb]	$\sigma(WH)$ [fb]	$\sigma(ZH)$ [fb]	$\sigma(VBF)$ [fb]	$\mathcal{BR}(H \rightarrow W^+W^-)$ [%]	$\mathcal{BR}(H \rightarrow ZZ)$ [%]	$\mathcal{BR}(H \rightarrow \tau^+\tau^-)$ [%]
120	1072.3	150.10	90.2	72.7	14.3	1.527	7.11
125	949.3	129.50	78.5	67.1	21.6	2.549	6.37
130	842.9	112.00	68.5	62.1	30.5	3.858	5.49
135	750.8	97.20	60.0	57.5	40.3	5.319	4.52
140	670.6	84.60	52.7	53.2	50.4	6.715	3.54
145	600.6	73.70	46.3	49.4	60.3	7.771	2.62
150	539.1	64.40	40.8	45.8	69.9	8.143	1.79
155	484.0	56.20	35.9	42.4	79.6	7.297	1.06
160	432.3	48.50	31.4	39.4	90.9	4.185	0.397
165	383.7	43.60	28.4	36.6	96.0	2.216	0.138
170	344.0	38.50	25.3	34.0	96.5	2.351	0.0920
175	309.7	34.00	22.5	31.6	95.8	3.204	0.0719
180	279.2	30.10	20.0	29.4	93.2	5.937	0.0587
185	252.1	26.90	17.9	27.3	84.4	14.86	0.0457
190	228.0	24.00	16.1	25.4	78.6	20.77	0.0376
195	207.2	21.40	14.4	23.7	75.7	23.66	0.0324
200	189.1	19.10	13.0	22.0	74.1	25.33	0.0287
210	158.9	15.20	10.5	19.1		27.16	
220	134.5	12.30	8.5	16.6		28.11	
230	114.7	9.90	7.0	14.5		28.70	
240	98.4	8.03	5.7	12.6		29.11	
250	85.0	6.53	4.7	11.0		29.43	
260	73.8	5.33	3.9	9.6		29.70	
270	64.5	4.37	3.2	8.4		29.92	
280	56.7	3.59	2.7	7.4		30.12	
290	50.1	2.96	2.2	6.4		30.29	
300	44.7	2.45	1.9	5.6		30.45	

Table 6.1: Production cross section for the several Higgs production mechanism at the Tevatron and  $\mathcal{BR}(H \rightarrow ZZ)$  in the considered Higgs mass range.

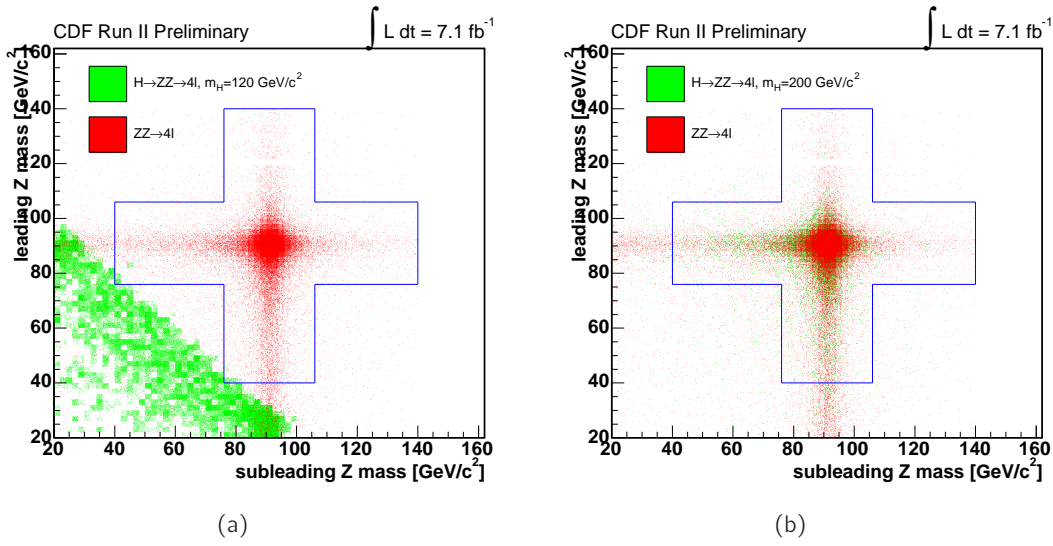


Figure 6.3: Scatter-plot of subleading  $p_T M_{ll}$  vs leading  $p_T M_{ll}$  for MC  $ZZ$  and  $H \rightarrow ZZ$  samples, where (a)  $m_H = 120 \text{ GeV}/c^2$  and (b)  $m_H = 200 \text{ GeV}/c^2$ . The blue lines represent the acceptance region of the  $M_{\ell\ell-1,2}$  requirements applied in the  $ZZ$  cross section measurement, which are released in this Higgs boson search.

$\int \mathcal{L} = 9.7 \text{ fb}^{-1}$

Contribution		Yields			
ZZ		$10.59 \pm 1.34$			
Fakes		$0.39 \pm 0.19$			
Total Background		$10.98 \pm 1.34$			
Higgs Process/Mass ( $\text{GeV}/c^2$ )	130	150	170	190	
$gg \rightarrow H \rightarrow ZZ$	0.09	0.23	0.06	0.39	
$WH \rightarrow WZZ$	0.01	0.01	0.00	0.01	
$ZH \rightarrow ZZZ$	0.01	0.02	0.01	0.04	
VBF $H \rightarrow ZZ$	0.01	0.02	0.01	0.03	
$ZH \rightarrow ZWW$	0.07	0.11	0.11	0.06	
$ZH \rightarrow Z\tau\tau$	0.03	0.01			
Total Signal	$0.22 \pm 0.02$	$0.41 \pm 0.03$	$0.18 \pm 0.01$	$0.55 \pm 0.05$	
Data	9				

Table 6.2: Expected and observed number of events in the considered sample.

The contribution from  $ZH \rightarrow ZWW$  is comparable with the one from direct  $H \rightarrow ZZ$  decay but while in the latter case are produced four resonant leptons in the final state, hence resulting in a peak in the  $m_{4\ell}$  spectrum, in the former case the uncorrelated production of the four lepton will have a broad  $m_{4\ell}$  spectrum. To exploit at best the sensitivity to this Higgs signal contribution we therefore need to exploit other kinematic properties of these decay modes.

Higgs (GeV/ $c^2$ )	Signal Contributions					Total
	$gg \rightarrow H \rightarrow ZZ$	$ZH \rightarrow ZWW$	$ZH \rightarrow ZZZ$	VBF $H \rightarrow ZZ$	$WH \rightarrow WZZ$	
120	0.029	0.036	0.002	0.003	0.002	0.07
125	0.054	0.052	0.004	0.006	0.003	0.12
130	0.095	0.068	0.008	0.010	0.005	0.19
135	0.143	0.082	0.012	0.015	0.008	0.26
140	0.185	0.095	0.016	0.020	0.010	0.33
145	0.226	0.103	0.020	0.024	0.012	0.38
150	0.232	0.109	0.021	0.024	0.012	0.40
155	0.202	0.118	0.019	0.021	0.010	0.37
160	0.113	0.125	0.011	0.011	0.005	0.27
165	0.056	0.120	0.006	0.006	0.003	0.19
170	0.056	0.109	0.006	0.005	0.002	0.18
175	0.072	0.101	0.008	0.007	0.003	0.19
180	0.128	0.087	0.014	0.012	0.005	0.25
185	0.302	0.075	0.034	0.028	0.012	0.45
190	0.389	0.064	0.045	0.035	0.015	0.55
195	0.400	0.056	0.047	0.035	0.014	0.55
200	0.389	0.049	0.047	0.034	0.013	0.53
210	0.340		0.043	0.028	0.011	0.42
220	0.303		0.040	0.024	0.008	0.38
230	0.270		0.038	0.021	0.007	0.33
240	0.241		0.035	0.018	0.005	0.30
250	0.216		0.033	0.015	0.004	0.27
260	0.191		0.030	0.013	0.003	0.24
270	0.174		0.028	0.011	0.002	0.21
280	0.161		0.027	0.010	0.001	0.20
290	0.141		0.025	0.008	0.001	0.17
300	0.132		0.024	0.007	0.000	0.16

Table 6.3: Expected number of signal events in  $9.7 \text{ fb}^{-1}$ . The uncertainty on the expected signal is about 10%. We are lacking  $ZH \rightarrow ZWW$  MC above  $200 \text{ GeV}/c^2$ .

## 6.0.8 Background Determination

The main background to this search is the SM non-resonant  $ZZ$  production, measured as described in Chapter 5, where both  $Z$ s decay to a pair of leptons. This process has been studied in Chapter 5, measuring its production cross section and checking the modeling of the most relevant kinematic variables. In the search for the Higgs boson, the  $ZZ$  production has been modeled using the PYTHIA MC simulation already described in Section 5.1.2.

As already mentioned in Section 5.1.3 a much smaller background contribution to this search comes from events with some misidentified jet mimicking one of the leptons in the final state. While in the  $ZZ$  cross section measurement that was the only background process, for the search of the Higgs boson this becomes almost negligible compared to the non-resonant  $ZZ$  background. The contribution to the data set considered coming from this detector effect is estimated with the data-driven method described in Section 5.1.3. The lepton plus denominator sample used to obtain this background estimation is not suitable for a precise modeling of the kinematic variables

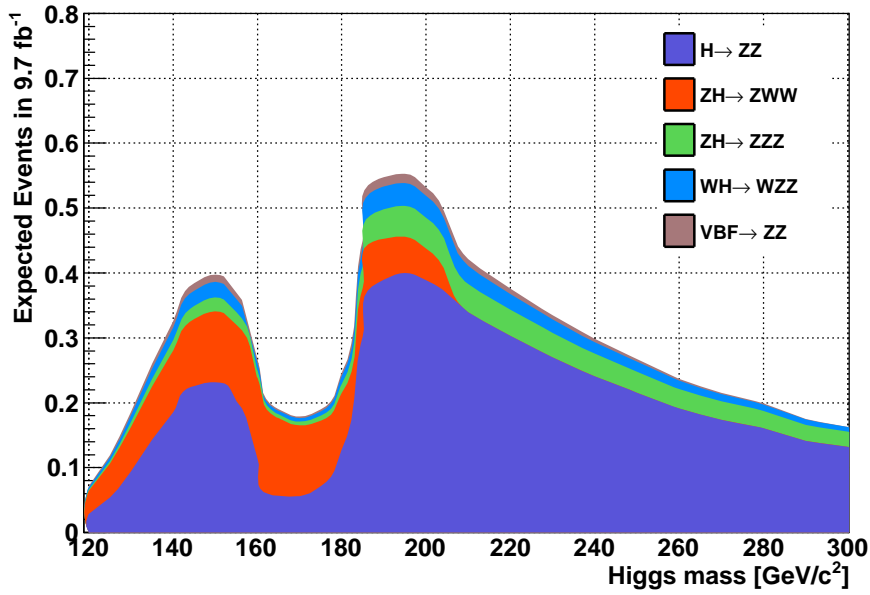


Figure 6.4: Expectations for the several Higgs signal contribution in the considered sample.

of interest in this search (e.g.  $m_{4\ell}$ ) because is composed of a small number of events<sup>8</sup>. For this reason we model this contribution with the help of a MC simulation of  $Z\gamma$  production, as explained in the following section.

### Fake background kinematic distribution

The method described in Section 5.1.3 gives a reliable estimate of the background contribution but it's based on a small statistic sample, that prevents from the possibility to have a description of the kinematic properties of this background which is not significantly affected by statistical fluctuation within the available sample. Hence, we extract the  $M_{4\ell}$  distribution for fake lepton events with the help of a Monte Carlo simulation, applying to a larger statistic sample the same procedure applied to data. The dominant contribution to this instrumental background is expected to come from  $Z\gamma$  production, where two leptons come from the  $Z$  decay and the other two can come from misidentification of the photon and of an additional jet, hence we consider a MC simulation of this process<sup>9</sup>. We weight simulated events containing 3(2) fully reconstructed leptons and 1(2) denominator objects with the appropriate mis-identification probabilities. To increase the sample obtained in this way we release some of the requirements in the definition of the denominator objects (see Appendix A.2 for standard definition). We first release the requirement on the Calorimeter Isolation (as defined in Section 4.3, Equation 4.7), which usually is  $Cal.Iso. < 0.3$  for denominator objects and  $Cal.Iso. < 0.1$  for real leptons. In addition to that, for a denominator electron candidate we ignore the requirement on  $E_{Had}/E_{Em}$  (see Appendix A.2 for the default requirement) and for a denominator muon candidate the requirement on  $E/P$  (which

<sup>8</sup> This sample is made of  $\sim 100$  events, which is enough statistic to evaluate the overall background contribution with a statistical uncertainty smaller than the systematic uncertainty associated to the fake rates. This is not suitable for a smooth modeling of the main event kinematic variables, which will be exploited to extract information on the Higgs boson production.

<sup>9</sup> The same procedure described in the following has been tested considering other minor contributing processes but found them to have a negligible effect to the modeling of the background.



is usually required to be  $< 1.0$ , see Appendix A.2). Weighting the simulated events containing one or more of such loosened denominator with the corresponding fake rates we obtain the  $M_{4\ell}$  distribution shown in Figure 6.5(a). This distribution is then fitted with a Landau function<sup>10</sup>, shown in Figure 6.5, to get a smooth shape of the  $M_{4\ell}$  for this background process<sup>11</sup>. The fitted parameters of the Landau function and the corresponding uncertainties will be considered in the measurement. The complete background estimation of the  $M_{4\ell}$  distribution is based on the MC fit for its shape and the data-driven method for the overall normalization.

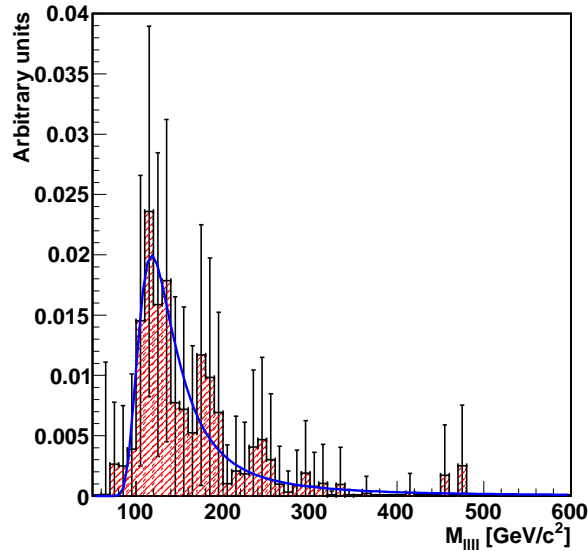


Figure 6.5: The four lepton invariant mass distribution for  $Z\gamma$  events containing one or two loosened denominator objects, with the fitted Landau function overlaid, used to extract the *fake* background modeling.

### 6.0.9 Event kinematics

We search for a signal of Higgs boson production and decay owing to four lepton final state in the 2D variable space of four lepton invariant mass ( $m_{4\ell}$ ) versus missing transverse energy ( $\cancel{E}_T$ ). Processes concurring to the  $4\ell$  production, like  $ZH \rightarrow ZWW$ , are characterized by the production of a significant amount of  $\cancel{E}_T$  due to the undetectable neutrinos. Moreover, unlike the four lepton production in the  $H \rightarrow ZZ^* \rightarrow 4\ell$  channel, where we expect a particular value of the four lepton invariant mass, the SM  $ZZ^*$  production channel is not expected to present  $m_{4\ell}$  production at any particular mass value. For this reason, in the search for the Higgs boson production in this final state it is important to study the correlation between the  $m_{4\ell}$  and the  $\cancel{E}_T$ . It is possible to see from the bi-dimensional distributions shown in Figure 6.6 and the corresponding one-dimensional projection shown in Figure 6.7 that the  $m_{4\ell}$ ,  $\cancel{E}_T$  and respective correlation, is different for the several signal and background processes considered. Considering the three most peculiar processes,

<sup>10</sup> The Landau function has found to model properly the turn-on and tail of the data distribution with sufficient precision. Other analytical function has been considered but found no significant difference with respect to the chosen one.

<sup>11</sup> This procedure relies on the fact that the loosened denominator objects preserve the same kinematic properties as the default ones, in particular for what concerns the  $M_{4\ell}$  distribution.

we can notice that  $ZZ$  and  $H \rightarrow ZZ$  production have both a small amount of  $\cancel{E}_T$  in the final state but the second one appears as a peak in the  $m_{4\ell}$  distribution. On the other hand, processes where the four leptons are not the result of the  $H \rightarrow ZZ \rightarrow \ell\ell'\ell'\ell'$  decay like  $ZH \rightarrow ZWW$ , present a broad  $m_{4\ell}$  spectrum but can be separated from the  $ZZ$  and  $H \rightarrow ZZ$  production thanks to the significant amount of  $\cancel{E}_T$  present in the final state. A bi-dimensional fit in the  $m_{4\ell}$ - $\cancel{E}_T$  plane will take into account all these information with the relative correlations to retrieve information on the Higgs boson production in the considered final state. Figures 6.8 display the projections of  $m_{4\ell}$  and  $\cancel{E}_T$  distributions, respectively, overlaid are expected signal and background, and data.

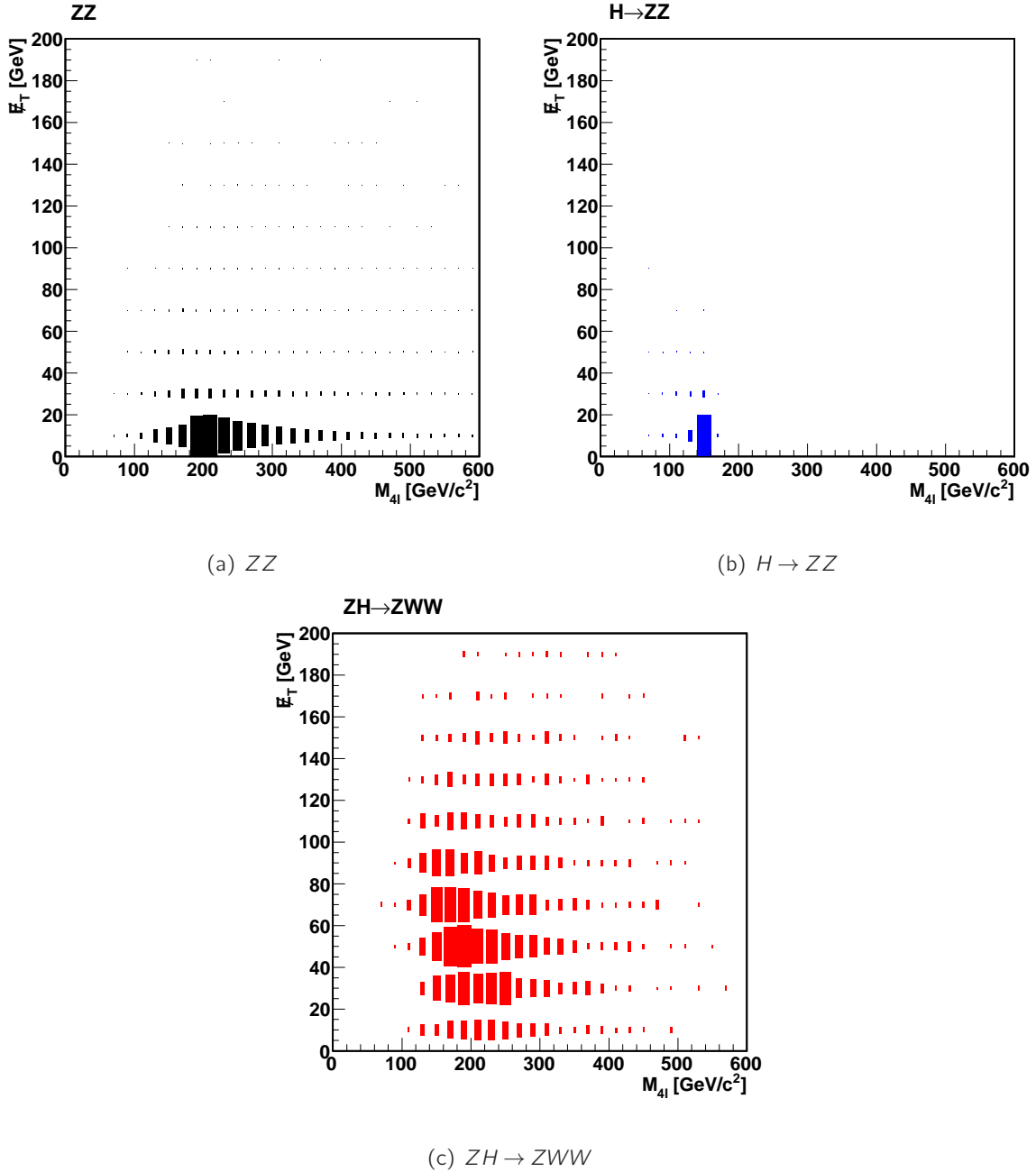


Figure 6.6: Correlated distribution of  $\cancel{E}_T$  vs.  $M_{4\ell}$  for the three main processes considered: (a) SM  $ZZ$  production, (b) gluon fusion  $H \rightarrow ZZ$ , and (c) associated production  $ZH \rightarrow ZWW$ . The simulation assumes  $m_H = 150$  GeV/c<sup>2</sup>.

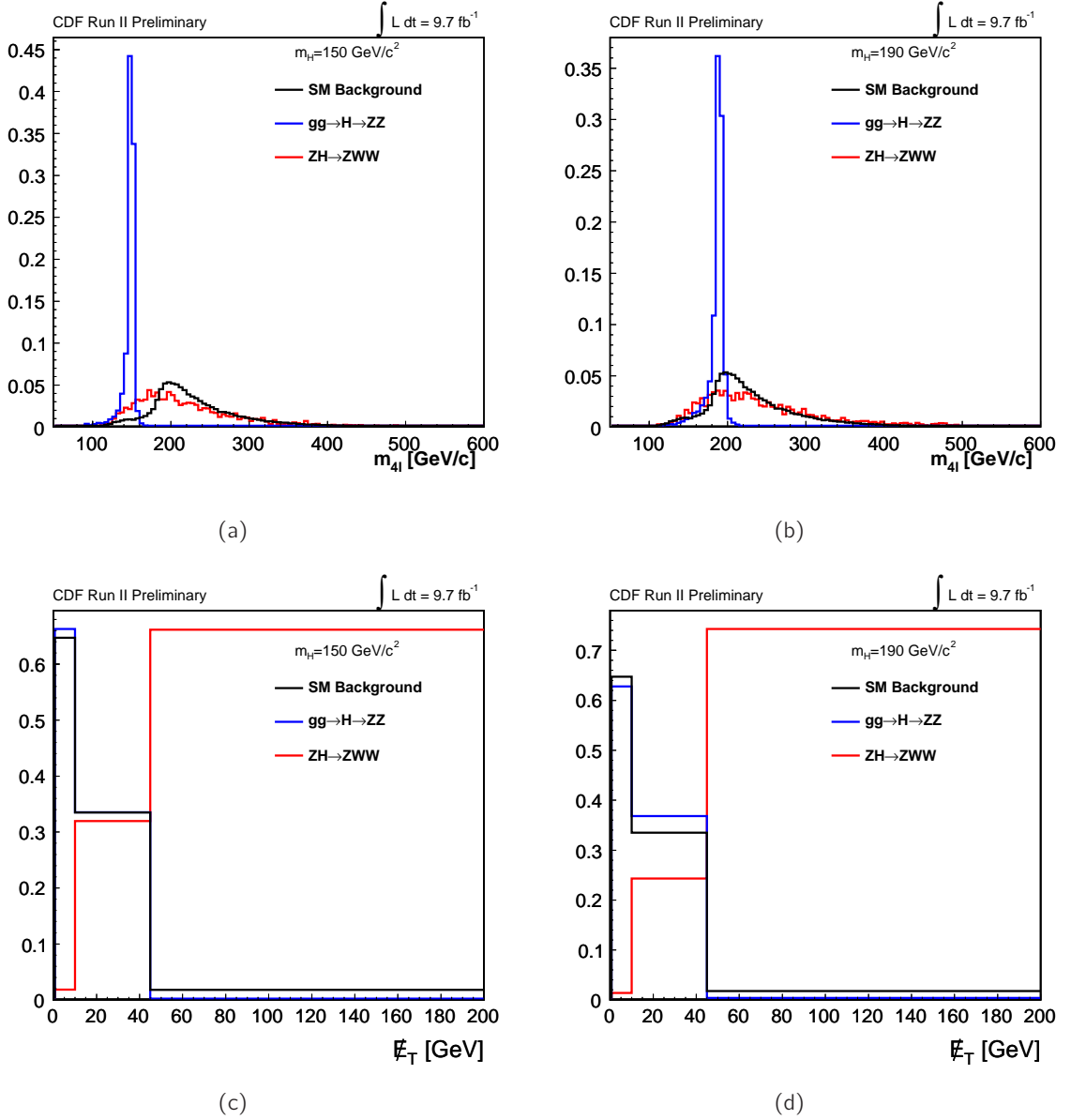


Figure 6.7: Kinematic distribution of the  $m_{4\ell}$ (a,b) and  $\cancel{E}_T$ (c,d) for assuming  $m_H = 150 \text{ GeV}/c^2$ (a,c) and  $m_H = 190 \text{ GeV}/c^2$ (b,d). In the plots are reported the predictions for the three most relevant contribution: in black is the SM non-resonant  $ZZ$  production, in blue the resonant  $gg \rightarrow H \rightarrow ZZ$  four lepton contribution, and in red the non-resonant  $ZH \rightarrow ZWW$  four lepton contribution.

From the data-to-simulation comparison in the  $m_{4\ell}$ - $\cancel{E}_T$  kinematic space is not seen any significant deviation from a background-only hypothesis, assuming no production of the Higgs boson. The expected Higgs signal in the considered sample, assuming the SM theoretical prediction on its production mechanisms, is small (as reported in Table 6.3) compared to the  $ZZ$  background contribution (see Table 6.2). For this reason we aim to extract from this 2D distribution comparison between data and simulated models an upper limit on the Higgs production cross section, for different Higgs mass hypotheses.

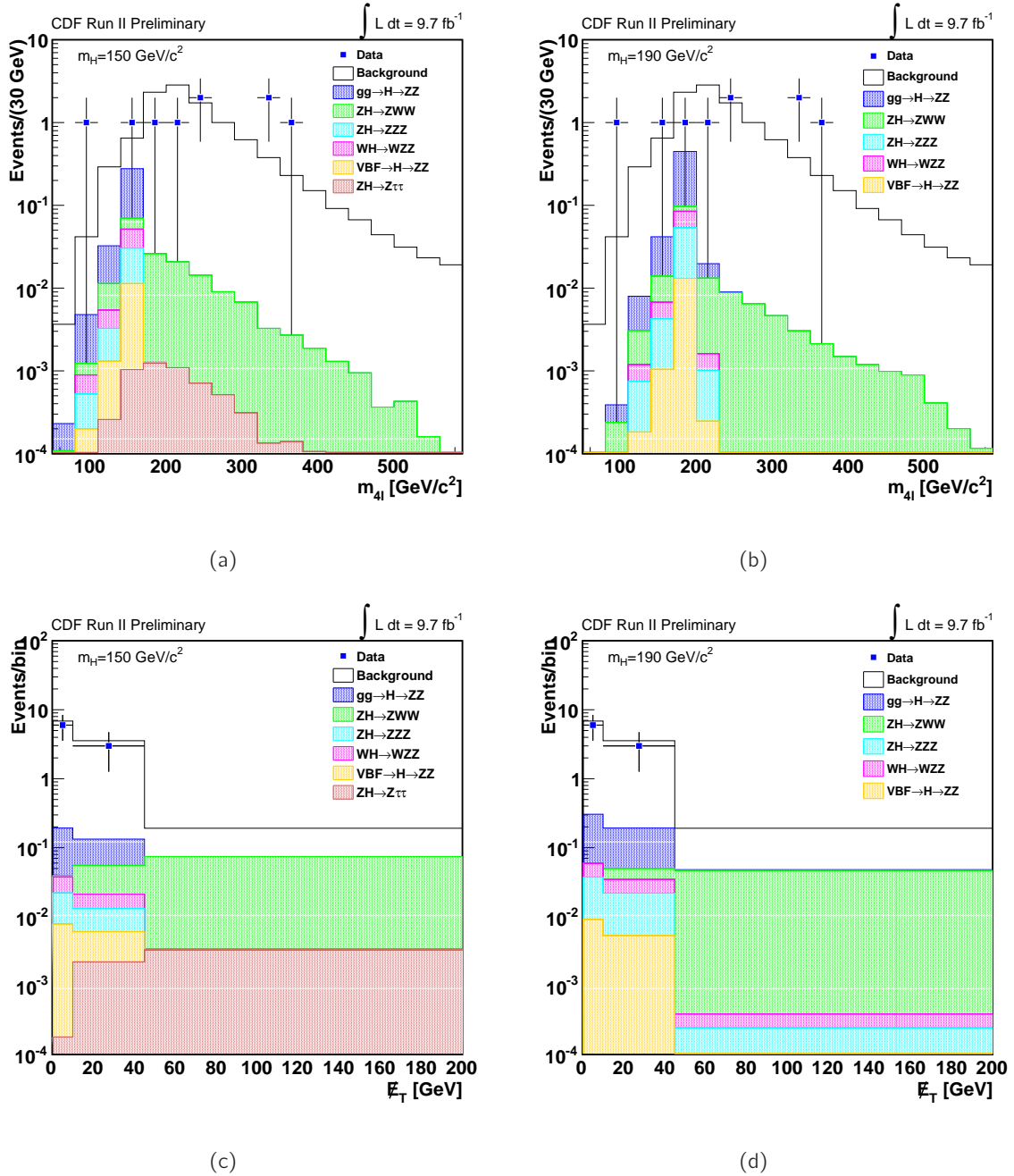


Figure 6.8: One dimensional projection of the  $m_{4\ell}$ (a,b) and  $\cancel{E}_T$ (c,d) distribution of the several processes predictions stacked over each other and the observed data overlaid.

### 6.0.10 Systematic Uncertainties

The calculation of the upper limit on the Higgs production cross section relies on some Monte Carlo simulations of the signal and background processes, which might be affected by systematic uncertainties. Since we extract information on the Higgs production from a data-to-simulation variable distribution comparison, we should take into account not only the sources of systematic uncertainty on the overall normalization of the simulated samples (*rate* systematics) but also the ones that might change the kinematic variable distributions (*shape* systematics). The search for the Higgs boson signal is performed exploiting the 2D kinematic distribution of  $m_{4\ell}$  vs.  $\cancel{E}_T$  hence we will focus on the systematic uncertainties that affects significantly these two kinematic variables. The  $m_{4\ell}$  distribution uncertainty is driven by the resolution in the single lepton momentum reconstruction, which is measured with  $\sigma(p_T) \sim 2\text{GeV}/c$  in the momentum range considered. This reflects in a resolution on the four lepton invariant mass of  $\sigma(m_{4\ell}) \sim 4\text{-}5\text{ GeV}/c^2$ , which is smaller or equal than the spacing in the tested Higgs mass hypotheses ( $5\div 10\text{ GeV}/c^2$ ). A mismodeling in the lepton momentum reconstruction will have a negligible effect in the data-to-simulation comparison in the  $m_{4\ell}$  distribution, hence we don't consider any source of shape variation for this kinematic variable. On the other hand the  $\cancel{E}_T$  distribution for signal and background processes might not be properly modeled in the MC simulation therefore we consider a systematic shape uncertainty for this kinematic variable.

A summary of the systematic uncertainties considered in the limit evaluation is reported in Table 6.4; most of the systematics considered in this case are the same already described in Section 5.1.5, then we discuss here only the relevant differences with respect to that measurement. The values in a same row of the Table 6.4 written in *italic* are expected to be correlated, and will be properly treated in the statistical approach to the limit calculation, that will be described in Section 6.0.11.

$\int \mathcal{L} = 9.7\text{ fb}^{-1}$						
Uncertainty Source	ZZ	$Z(\gamma^*)$	gg→H	WH	ZH	VBF
<b>Cross Section</b>						
Scale			7.0%			
PDF			7.7%			
Total	10%			5%	5%	10%
Branching Ratio			3%	3%	3%	3%
<b>Acceptance</b>						
Higher-order Diagrams	2.5%					
PDF	2.7%					
Luminosity	5.9%		5.9%	5.9%	5.9%	5.9%
Lepton ID Efficiencies	3.6%		3.6%	3.6%	3.6%	3.6%
Trigger Efficiencies	0.4%		0.5%	0.5%	0.5%	0.5%
Fake Rates		50%				
$\cancel{E}_T$ modeling			<i>Shape uncertainty</i>			

H → 4ℓ

Table 6.4: Summary of the systematic uncertainties considered in the limit calculation that affect signal and background predictions. Each row corresponds to a single source of systematic uncertainty while each column represent one of the processes that contributes to the selected data sample. Empty spaces represent source of systematic uncertainties that are not relevant for that particular process.

**Higgs Cross Section** The  $ZZ$  production cross section has been calculated using MCFM[18] at NLO with a resulting uncertainty of 10%. This uncertainty covers both the PDF and scale variation and is reported as *Total* systematic uncertainty on the cross section in Table 6.4. For the production of the Higgs boson through gluon fusion we consider separately the uncertainties due to PDF and scale variation. Following the approach outlined in [51, 50], we use HNNLO[90] to estimate these errors. HNNLO is a program to calculate the theoretical Higgs production cross section by gluon fusion that performs a NNLO QCD calculation. We estimate the systematic uncertainties by varying the renormalization and factorization scale ( $\mu_R, \mu_F$ ) used in the calculation, doubling and halving their values ( $\mu_{R,F} \equiv \mu_0 \rightarrow \mu_0/2, \mu_0 \cdot 2$ ), and found a variation of  $\pm 7.0\%$ , consistent with [51, 50]. We consider 7.7% uncertainty due to the PDF, as calculated in [51, 50], obtained varying the PDFs models used in the simulation within their quoted errors, as recommended in [52].

**Higgs Decay Branching Fraction** The uncertainty on the branching fraction of the Higgs decay mode considered are extracted from the one calculated in [54]. For the considered Higgs boson decay modes, the largest uncertainties are obtained for low Higgs mass hypotheses, for which the  $H \rightarrow b\bar{b}$  decay is relevant and drives the overall branching fraction uncertainties<sup>12</sup>. For simplicity we consider in this analysis an uncertainty independent from the Higgs boson mass, assigning to the Higgs production simulation the largest uncertainty value over the mass range considered.

**MC  $\cancel{E}_T$  modeling uncertainty** Some of the processes considered in the four lepton final state (in particular  $p\bar{p} \rightarrow ZZ, gg \rightarrow H \rightarrow ZZ$  and  $VBF \rightarrow ZZ$ ) does not produce any neutrino, hence they should not present any  $\cancel{E}_T$  as a measurement of high- $p_T$  undetected particles. For these kind of events the measured  $\cancel{E}_T$  is primarily caused by mis-measurement of other objects in the event. For the processes with no intrinsic  $\cancel{E}_T$  we don't expect the MC simulation to perfectly reproduce the  $\cancel{E}_T$  kinematic distribution. This resolution effect has been studied using a sample of  $Z \rightarrow \ell\ell$  events and found the  $\cancel{E}_T$  resolution being  $\sim 4 \div 5\text{GeV}$  in the kinematic range considered for the investigated processes. We take into account this effect by considering a simulated sample with the  $\cancel{E}_T$  shifted upward and downward by its resolution and propagating these through the analysis requirements. The variation in the  $\cancel{E}_T$  distribution for the shifted samples are shown in Figure 6.9 for the processes without intrinsic  $\cancel{E}_T$ . For the other processes we expect this uncertainty to be negligible. The relative bin-by-bin variation is then included as a shape systematic uncertainties in the final limit calculation.

### 6.0.11 Limit Calculation on the Higgs production cross section

In the  $m_{4\ell}$  and  $\cancel{E}_T$  distributions shown in Figure 6.8 the observed data look compatible with the predicted background contribution, with no significant deviations that might be due to the Higgs signal. Therefore it is desirable to quote an upper limit on the possible signal contribution to the data at some confidence level (C.L.). To establish the possible deviation of the data from a background-only scenario we use the same Bayesian approach applied in the cross section measurement described in Section 5.1.6. Following a standard choice, we estimate the upper limit at 95% C.L. on the Higgs production cross section, as the ratio to the Standard Model prediction, for the mass hypotheses considered in the range  $120 \leq m_H \leq 300 \text{ GeV}/c^2$ . This method has the

<sup>12</sup>The  $\mathcal{BR}(H \rightarrow b\bar{b})$  uncertainty is driven by the uncertainty on the  $b$ -quark mass and the corresponding coupling to the Higgs boson; more details can be found in [54].

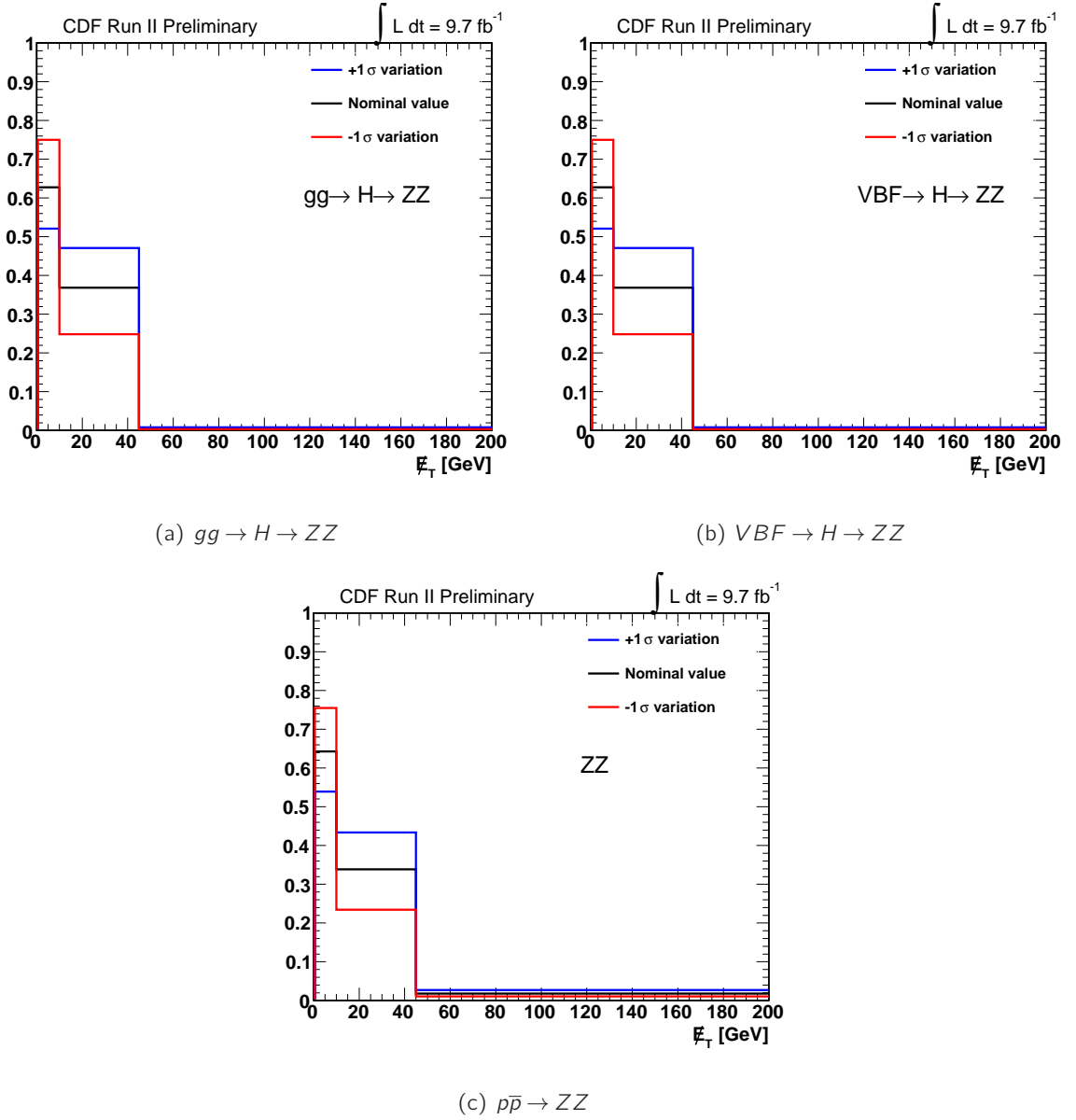


Figure 6.9: Representation of the  $\cancel{E}_T$  distribution variation considering the shifted samples. The  $\pm 1\sigma$  variation on the shape of the  $\cancel{E}_T$  distribution is considered as one of the systematic uncertainties in the limit calculation.

advantage of naturally including systematic uncertainties in the calculation and being flexible for an easy combination of several searches.

The input of the calculation is the 2 dimensional binned distribution of  $m_{4\ell}$  vs.  $\cancel{E}_T$ , and the procedure is replicated for each Higgs mass hypothesis. We apply the Likelihood definition introduced for the  $ZZ$  cross section measurement in Equation 5.19 to this case, building a likelihood function  $\mathcal{L}$  which is the product of the probabilities of observing in the  $i$ -th bin of the histogram  $n_i$  events, give an expectation of  $\mu_i$ :

$$\mathcal{L} = \prod_i \frac{\mu_i^{n_i} e^{-\mu_i}}{n_i!} \cdot \prod_{j=\text{syst.}} e^{-\frac{\theta_j^2}{2}} \quad (6.1)$$

where

$$\mu_i = \sum_k \alpha_k \left[ \prod_j (1 + f_k^j \cdot \theta_j) \right] (N_k^{Exp})_i \equiv R \cdot s_i(\vec{f}, \vec{S}) + b_i(\vec{f}, \vec{S}) \quad (6.2)$$

where the index  $k$  runs over the several background and signal processes that contribute to the considered sample.

The expectation in the  $i$ -th bin ( $\mu_i$ ) is expressed as the sum of all the background contributions ( $b_i(\vec{f}, \vec{S})$ ) and the signal contributions ( $s_i(\vec{f}, \vec{S})$ ). The parameter  $R$  multiplies the number of expected signal, that can be written as  $s_i = L \cdot \epsilon \cdot \sigma_H^{SM}$ , where  $L$  is the integrated luminosity of the collected sample,  $\epsilon$  the detector acceptance and analysis selection efficiency and  $\sigma_H^{SM}$  is the Standard Model predicted cross section for the Higgs production. In the confidence level calculation we assume no experimental information on the Higgs production cross section, and we assign a non-negative flat prior to  $R$ . The signal and background predictions dependent on the parameters  $\vec{S}$  and  $\vec{f}$ , that are used to account for the systematic uncertainties on the simulated samples, being  $\vec{S}$  the *nuisance* parameters and  $\vec{f}$  the fractional uncertainty associated to their prediction. In Equation 6.2 the product runs over the index  $c$ , which account for different sources of systematic uncertainties. The nuisance parameters  $S_c$  have a Gaussian constraint in the likelihood, as shown in Equation 6.1. This formulation allows the fractional uncertainties  $f$  to be different in each bin, being able to reproduce both *rate* and *shape* uncertainties on the predictions. In this binned likelihood the index  $i$  runs over the bins of the 2D distribution. The definition in Equation 6.2 assures to properly taking into consideration the correlations of the systematic uncertainties.

The likelihood is integrated over the nuisance parameters and evaluated for the observed number of events  $\vec{n}$ , with only residual dependence on the parameter  $R$ . The result, normalized to unit area, is the posterior p.d.f.  $\mathcal{L}(R)$ , of which an example is reported in Figure 6.10(a). The fact that the posterior p.d.f. peaks at zero means that the favorable hypothesis is the one with no Higgs signal in it.

The 95% upper limit  $R^{95}$  is evaluated from the posterior p.d.f. as the solution of

$$\frac{\int_0^{R^{95}} \mathcal{L}(R) dR}{\int_0^{+\infty} \mathcal{L}(R) dR} = 0.95. \quad (6.3)$$

This corresponds to the limit on the production cross section of the Higgs boson normalized to the SM prediction. If the 95% of the posterior p.d.f. area lies below  $R^{95}$  it means that a signal with a cross section of  $R^{95} \times \sigma_H^{SM}$  or larger is excluded at 95% C.L.

To estimate the sensitivity of the analysis before looking at the data events, we generate pseudo-experiments in a background-only hypothesis: *pseudo*-data  $\vec{n}$  are randomly generated for each sample from the Poisson distribution with the mean given by the background expectation ( $N^{Exp}$ ), fluctuated by its systematic error. Each of them is used to evaluate a 95% C.L. limit in the same way as the real data. An example of the resulting distribution is shown in Figure 6.10(b). We define the expected limit as the median value of this distribution and the  $\pm 1\sigma$  and  $\pm 2\sigma$  variation of the expected limit as the intervals containing respectively the 16-84% and 2.3-97.7% of the total area of the distribution.

## 6.0.12 Results

The result obtained from the limit calculation using the full CDF dataset is shown in Figure 6.11 and summarized in Table 6.5. With the optimization of this search and exploiting the two-dimensional kinematic variable space, we obtain similar sensitivity in the low mass region, i.e.



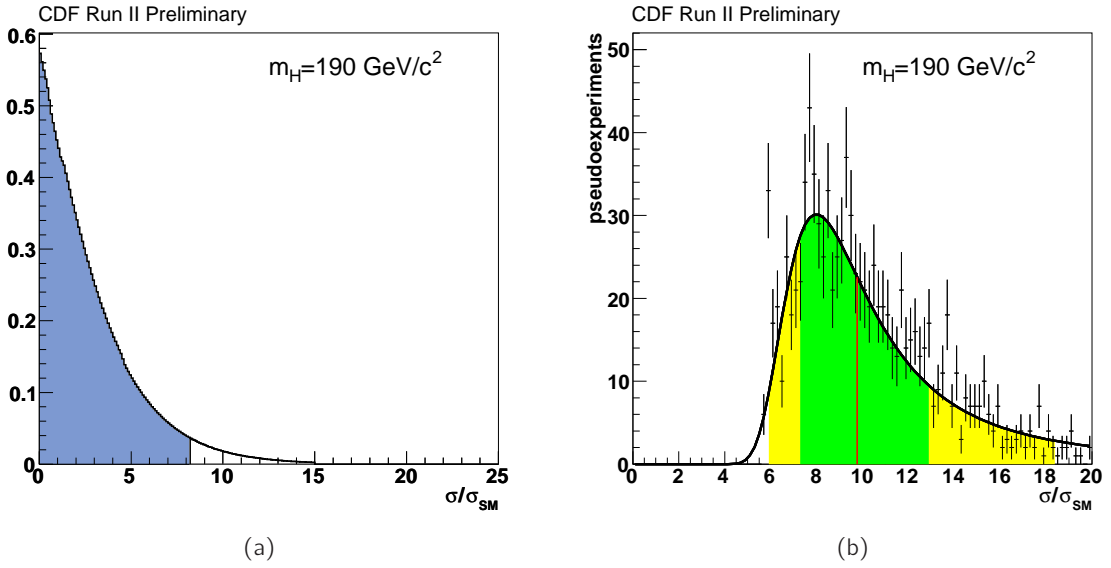


Figure 6.10: (a) Example of normalized posterior p.d.f. for  $\mathcal{L}(R)$ , indicates as  $\sigma/\sigma_{SM}$ , which is the ratio of the signal cross section to the standard model expectation. (b) Example of the expected distribution of 95% C.L. limits in the background-only hypothesis. The red line shows the median of the distribution, considered as the expected limit, while the green and yellow bands represent the  $\pm 1\sigma$  and  $\pm 2\sigma$  confidence interval respectively.

$m_H \sim 150 \text{ GeV}/c^2$  and in the  $ZZ$  on-shell region, i.e.  $m_H \sim 190 \text{ GeV}/c^2$ . The most stringent limit is expected for  $m_H = 150 \text{ GeV}/c^2$ ,  $9.44 \times \sigma_{SM}$ , where we set from the data a limit of  $9.49 \times \sigma_{SM}$ . At higher masses, e.g.  $m_H = 190 \text{ GeV}/c^2$ , the sensitivity is similar, with an expected(observed) limit of  $9.77(8.21) \times \sigma_{SM}$ . The most stringent limit is set for a Higgs with  $m_H = 200 \text{ GeV}/c^2$ , where we exclude a Higgs signal produced with a cross section of  $7.16 \times \sigma_{SM}$ . With this analysis we set limits on the Higgs production cross section in the four lepton final state, exploiting the best current sensitivity not only in the high Higgs boson mass region where both  $Z$ s are on-shell ( $> 190 \text{ GeV}/c^2$ ), but also in the lower mass region ( $\sim 150 \text{ GeV}/c^2$ ), where the sensitivity is significantly improved thanks to the inclusion of the  $H \rightarrow WW$  decay. This result [91] can be easily combined with searches of the Higgs boson in other final states carried out at CDF. The combination, that takes into account properly all the correlations, will take into account the result in the different final state and improve the sensitivity of each one into a single result.

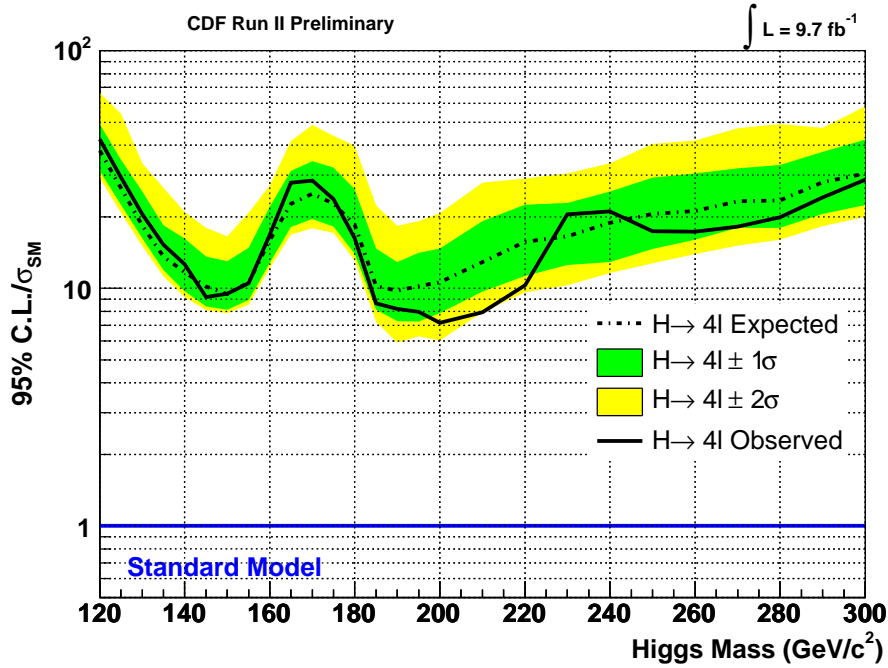


Figure 6.11: Expected and Observed 95% C.L. upper limit on the Higgs boson production cross section expressed as ratio to the SM prediction, in the Higgs mass range investigated.

		$\int \mathcal{L} = 9.7 \text{ fb}^{-1}$								
$H \rightarrow 4\ell$		120	130	140	150	160	170	180	190	200
$-2\sigma/\sigma_{SM}$		29.05	15.10	9.19	7.86	12.17	17.89	13.10	5.92	6.04
$-1\sigma/\sigma_{SM}$		31.20	16.21	9.73	8.11	12.80	19.60	13.98	7.29	7.88
<b>Median/<math>\alpha_{SM}</math></b>		<b>37.98</b>	<b>18.31</b>	<b>11.69</b>	<b>9.44</b>	<b>15.96</b>	<b>25.08</b>	<b>18.55</b>	<b>9.77</b>	<b>10.59</b>
$+1\sigma/\sigma_{SM}$		49.15	25.59	16.19	12.98	22.03	34.24	26.20	12.89	14.72
$+2\sigma/\sigma_{SM}$		66.59	33.62	20.96	16.58	27.12	48.78	39.90	18.38	20.72
<b>Observed/<math>\alpha_{SM}</math></b>		<b>42.38</b>	<b>20.51</b>	<b>12.64</b>	<b>9.49</b>	<b>16.83</b>	<b>28.45</b>	<b>16.27</b>	<b>8.21</b>	<b>7.16</b>

$H \rightarrow 4\ell$

		$\int \mathcal{L} = 9.7 \text{ fb}^{-1}$									
$H \rightarrow 4\ell$		210	220	230	240	250	260	270	280	290	300
$-2\sigma/\sigma_{SM}$		7.92	9.62	10.32	11.61	12.73	13.95	15.15	16.01	18.26	19.94
$-1\sigma/\sigma_{SM}$		9.69	11.26	12.52	12.91	14.67	16.02	17.99	17.96	20.60	22.45
<b>Median/<math>\alpha_{SM}</math></b>		<b>12.92</b>	<b>15.69</b>	<b>16.58</b>	<b>18.91</b>	<b>20.53</b>	<b>21.12</b>	<b>23.21</b>	<b>23.48</b>	<b>27.98</b>	<b>30.48</b>
$+1\sigma/\sigma_{SM}$		19.19	22.46	22.86	25.52	29.24	30.44	32.04	33.02	37.57	42.25
$+2\sigma/\sigma_{SM}$		27.82	28.96	30.46	33.57	40.62	41.95	47.21	49.16	47.36	58.28
<b>Observed/<math>\alpha_{SM}</math></b>		<b>7.92</b>	<b>10.31</b>	<b>20.46</b>	<b>21.11</b>	<b>17.36</b>	<b>17.30</b>	<b>18.17</b>	<b>19.93</b>	<b>24.09</b>	<b>28.64</b>

$H \rightarrow 4\ell$

Table 6.5: Expected and Observed 95% C.L. upper limit on Higgs production cross section as function of the Higgs boson mass. Limits are expressed in unit of the SM predicted cross section.

# Discussion of the Results

## Contents

7.1	Diboson measurements . . . . .	<b>127</b>
7.1.1	CDF $ZZ$ cross section measurement combination . . . . .	127
7.1.2	Dibosons at hadron colliders . . . . .	129
7.2	Higgs search summary . . . . .	<b>132</b>
7.2.1	Tevatron Higgs searches sensitivity . . . . .	132
7.2.2	LHC Higgs searches sensitivity . . . . .	132

## 7.1 Diboson measurements

### 7.1.1 CDF $ZZ$ cross section measurement combination

Considering the measurement of the  $ZZ$  production cross section obtained separately in the  $ll\ell'\ell'$  (see Section 5.1) and  $ll\nu\bar{\nu}$  (see Section 5.2) we can combine these two results taking into account the proper correlations between them. This can be done using the same Bayesian approach applied in the two aforementioned measurement. The two data samples exploited in the measurements are mutually exclusive, given the different explicit requirement on the number of reconstructed leptons in the final state. The correlations between the two measurements concern only the systematic uncertainties shared by the two analyses (e.g. theoretical prediction uncertainties, lepton reconstruction efficiencies, luminosity). In this case the likelihood function to maximize is obtained from the product of the likelihood functions in the two decay channels (i.e. Equation 5.11 and 5.19), with a common term representing the systematic uncertainties constrained to zero by a Gaussian distribution. The combined likelihood is then:

$$\mathcal{L} = \mathcal{L}_{ll\nu\bar{\nu}} \times \mathcal{L}_{4\ell} = \left[ \left( \prod_i \frac{\mu_i^{n_i} e^{-\mu_i}}{n_i!} \right)_{NN} \times \frac{\mu_{4\ell}^{n_{4\ell}} e^{-\mu_{4\ell}}}{n_{4\ell}!} \right] \cdot \prod_{j=syst.} e^{-\frac{\theta_j^2}{2}}. \quad (7.1)$$

The core of the likelihood is composed by the product of the probabilities in each bin of the Neural Network output and the one corresponding to the four lepton analysis. With such definition of the likelihood function all the systematic uncertainties parameters shared by the two decay modes are naturally correlated. Integrating the likelihood function over the nuisance parameters we can obtain from the posterior p.d.f the measured cross section (in units of the predicted  $ZZ$  SM cross

section):

$$\frac{\sigma(p\bar{p} \rightarrow ZZ)}{\sigma_{SM}} = 0.99 \pm 0.14(\text{stat.})_{-0.13}^{+0.14}(\text{syst.}) = 0.99_{-0.19}^{+0.20}, \quad (7.2)$$

which corresponds to a measured  $ZZ$  cross section:

$$\sigma(p\bar{p} \rightarrow ZZ) = 1.38 \pm 0.19(\text{stat.})_{-0.19}^{+0.20}(\text{syst.}) = 1.38_{-0.27}^{+0.28} \text{ pb} \quad (7.3)$$

which is in really good agreement with the Standard Model prediction  $\sigma_{ZZ}^{NLO} = 1.4 \pm 0.1 \text{ pb}$  at NLO. This result improves all the previous measurements performed at the Tevatron, summarized in Figure 7.1, and reduces the statistical and systematic uncertainties to the same level of  $\sim 15\%$ .

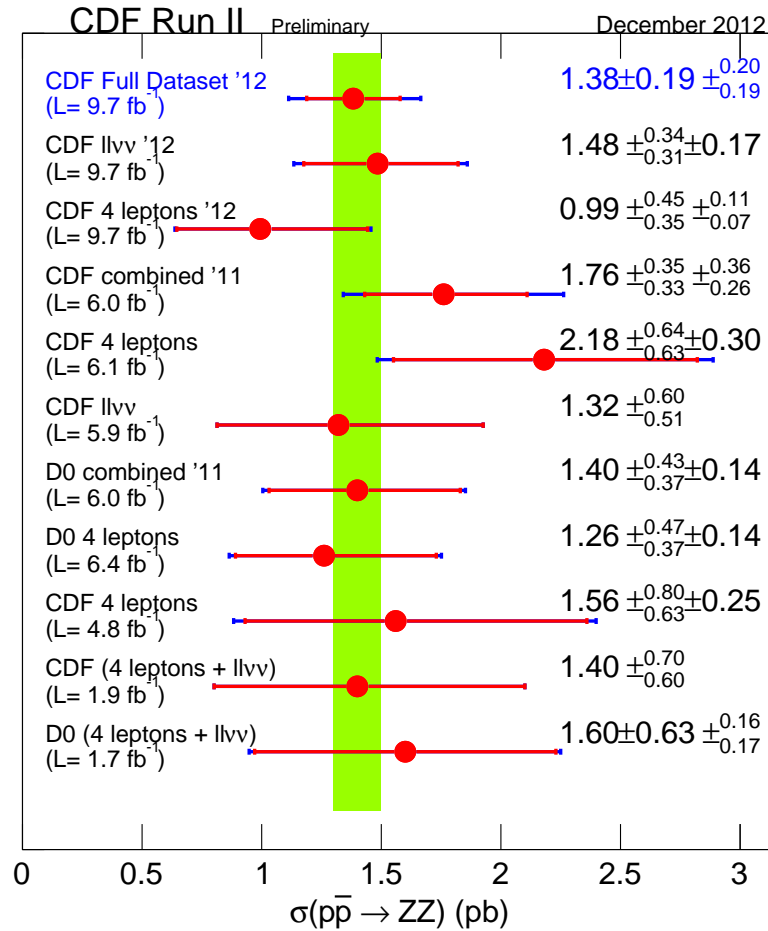


Figure 7.1: Summary of the  $ZZ$  cross section measurements done at the Tevatron (to date).

The statistical uncertainty of this result cannot be further reduced since the full CDF collected dataset is exploited; the systematic uncertainty is mainly driven by the uncertainty on the luminosity measurement (see Tables 5.4 and 5.10) which cannot be reduced, and on the theoretical cross sections used to normalize the MC simulated samples.

### Tevatron preliminary combination

To improve the experimental measurement of the  $ZZ$  cross section at  $\sqrt{s} = 1.96 \text{ TeV}$  it is possible to statistically combine the CDF measurement with analogous measurements carried out at the D0 experiment. As already mentioned in Section 2.2.2 (and shown in Figure 7.1) the D0

experiment performed a measurement in the  $ll'l'$  decay mode exploiting  $6.4 \text{ fb}^{-1}$  of data and in the  $ll\nu\nu$  decay mode exploiting  $8.6 \text{ fb}^{-1}$  of data. These measurements have larger uncertainties than the CDF one, and these could in principle can be slightly reduced extending the analyses to the whole data sample collected by the D0 experiment. The measurement obtained by CDF and D0 are not completely independent; some of the analysis inputs (e.g. process theoretical cross section and modeling) are shared by the four analyses. Although the combination of the four measurement would improve the precision of the experimental knowledge of  $ZZ$  production at the Tevatron energy.

*The combination procedure is in progress, thanks to the D0 collaboration, and will be officialized soon.*

### 7.1.2 Dibosons at hadron colliders

In the working period of the Tevatron collider the CDF and D0 experiments collected data corresponding to an integrated luminosity of  $\sim 10 \text{ fb}^{-1}$  each one. They both exploited these data samples to perform a set of measurement in the diboson sector. The most precise single process measurement of  $WW$ ,  $WZ$ , and  $ZZ$  have been obtained in the leptonic decay modes and are summarized in Figure 7.2.

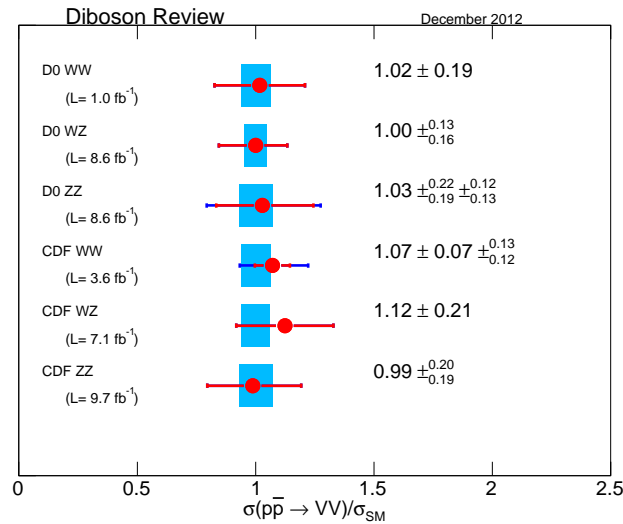


Figure 7.2: Summary of the main  $VV$  diboson measurements carried out at the Tevatron. The measured production cross section are normalized to the corresponding SM NLO prediction and the blue band represent the theoretical uncertainty for each one. The last measurement is the one described in Chapter 5 of this Thesis.

Both CDF and D0 exploited a fraction of the available dataset to measure the  $WW$  production cross section, which is the largest among the diboson ones, with an accuracy of 20%. The  $WZ$  cross section has been measured using  $7.1$  and  $8.6 \text{ fb}^{-1}$  from CDF and D0 respectively and both experiment reached a precision of 20% or better. CDF exploited the whole collected data sample to measure the  $ZZ$  production cross section, as reported in this Thesis, reaching the 20% level of precision, while D0 is currently updating its corresponding measurement. All the results achievable are limited by the finite size of the collected data sample, hence it will be hard to improve by a significant factor the precision of the measurements. Although a combination of the CDF and D0 best results of each process would provide the best achievable measurement at  $\sqrt{s} = 1.96 \text{ TeV}$ , contributing to the Tevatron Legacy.

At the LHC  $pp$  collider, providing collisions at  $\sqrt{s} = 7(8)$  TeV, the higher diboson production cross section (see Table 2.3) allowed the CMS and Atlas experiment to achieve (or improve) the Tevatron single-experiment precision with the current available dataset<sup>1</sup>. The main diboson results obtained by the Atlas and CMS experiment are reported in Figure 7.3.

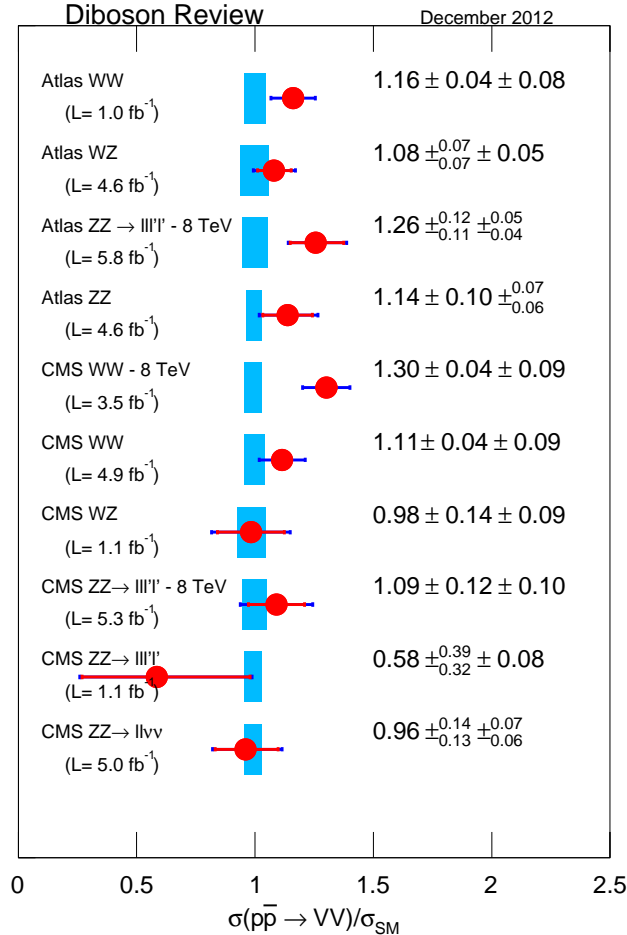
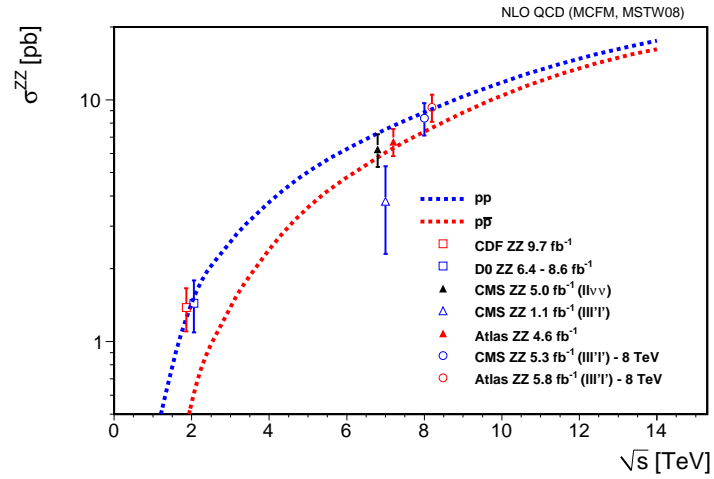


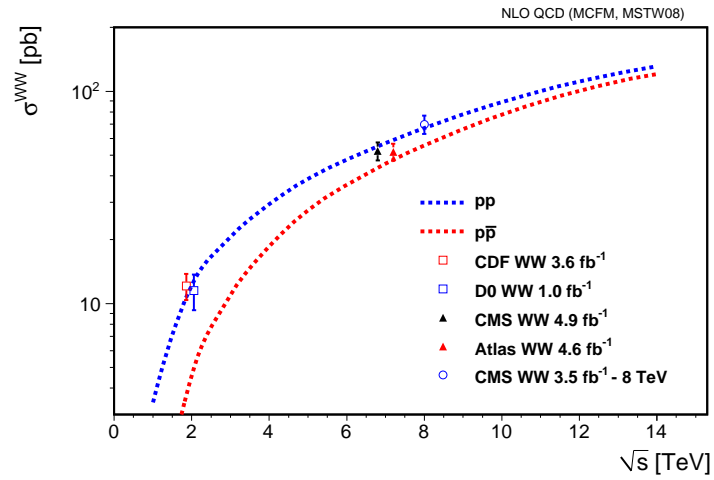
Figure 7.3: Summary of the main  $VV$  diboson measurements carried out at the LHC. The measured production cross sections are normalized to the corresponding SM NLO prediction and the blue band represent the theoretical uncertainty for each one.

In the past year both experiments measured the  $WW$ ,  $WZ$ , and  $ZZ$  production cross section at  $\sqrt{s} = 7$  TeV, while are still in progress some of the measurements at  $\sqrt{s} = 8$  TeV. If the LHC will upgrade to provide collision at higher center-of-mass energy, reaching the project working condition at  $\sqrt{s} = 14$  TeV, with a sample of data of the order of  $10 \text{ fb}^{-1}$ , it will be possible to complete a set of precise measurements in the diboson sector that will test the SM predictions in this electroweak sector. Figures 7.4 report the  $WW$ ,  $WZ$ , and  $ZZ$  production cross section as a function of the center-of-mass energy of the collisions, with the main measurements carried out at the Tevatron and LHC.

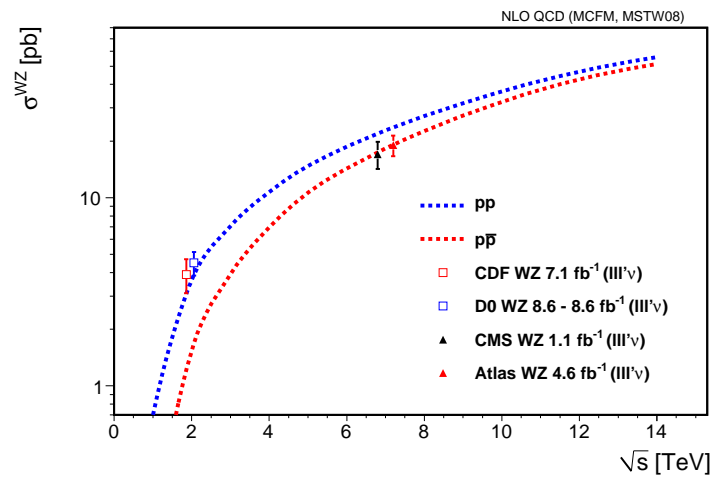
<sup>1</sup> The CMS and Atlas detector collected each one  $\sim 5 \text{ fb}^{-1}$  of data during 2010-'11 at  $\sqrt{s} = 7$  TeV and  $\sim 23 \text{ fb}^{-1}$  during 2012 at  $\sqrt{s} = 8$  TeV. For the diboson measurements only up to  $5 \text{ fb}^{-1}$  have been considered as of January 2013.



(a) ZZ production cross section



(b) WW production cross section



(c) WZ production cross section

Figure 7.4: Predicted and measured  $WW$ ,  $WZ$ , and  $ZZ$ , production cross section as a function of the center of mass energy  $\sqrt{s}$ , for  $pp$  and  $p\bar{p}$  collisions.

## 7.2 Higgs search summary

The  $ZZ$  cross section measurement presented in Chapter 5 set the basis for the search for the Standard Model Higgs boson in the  $ZZ$  decay mode, which has been presented in Chapter 6.

### 7.2.1 Tevatron Higgs searches sensitivity

In this Thesis we investigate the CDF sensitivity in the search for a SM Higgs boson in four-lepton final states. This search has been carried out analyzing the complete dataset collected by the CDF experiment and considering signal from  $H \rightarrow ZZ$  decay as well as contributions from several other processes involving the Higgs boson production producing four charged leptons in the final state. Since no significant hint of a SM Higgs boson signal has been seen in the four charged lepton data sample<sup>2</sup> it has been set an upper limit on its production cross section, as reported in Section 6.0.12. In the most sensitive  $m_H$  regions,  $m_H = 150, 190 \text{ GeV}/c^2$ , we expect to exclude the production of the Higgs boson with a cross section greater than  $\sim 10$  times the SM prediction, at 95% C.L.. This result is considered in a combination of similar SM Higgs searches performed at CDF considering different final states. Figure 7.5 shows the 95% C.L. upper limits (expected) set by CDF on the Higgs boson production cross section, for the different analyses performed. As we already mentioned, it is fundamental to study all the possible Higgs boson production and decay modes to retrieve information about the theoretical model it belongs to<sup>3</sup> through its coupling to the other SM known particles.

The various searches performed at CDF can be combined under the SM properties assumptions to achieve the best CDF sensitivity, shown in Figure 7.5, and with corresponding analyses carried out at D0. The SM Higgs boson Tevatron sensitivity is summarized in Figure 2.11, which shows that, exploiting the complete samples of collected data,  $\text{CDF} \oplus \text{D0}$ , reached the sensitivity to exclude the Higgs boson production at 95% C.L. in mass ranges between 147 and 180  $\text{GeV}/c^2$  [59]. With an expected sensitivity to Higgs production cross sections close to the SM ones, the Tevatron combination reported an excess of data in the range  $120 < m_H < 140 \text{ GeV}/c^2$ . The fit to the Higgs boson production cross section in the data, as a function of the mass  $m_H$ , is shown in Figure 7.6 and results in a Higgs boson production cross section compatible with the SM one in the aforementioned mass range.

### 7.2.2 LHC Higgs searches sensitivity

In the past two years of data collected at the LHC, several analyses has been performed to investigate the Higgs boson production considering similar production and decay modes to the Tevatron ones. The higher production cross section expected at  $\sqrt{s} = 7(8) \text{ TeV}$  (see Figure 2.8) made some Higgs boson decay modes more sensitive for the search of this particle. In particular, the  $H \rightarrow ZZ \rightarrow \ell\ell\ell'\ell'$  process is a *golden channel* to search for the Higgs boson for the CMS and Atlas experiments. The factor  $\sim 50$  in the  $gg \rightarrow H$  production cross section give a sizable amount of Higgs signal in the four lepton final state. The complete experimental determination of that process allows to reconstruct the Higgs boson mass from the four reconstructed leptons. In Figures 7.7 are shown the four-lepton invariant mass spectra of the  $H \rightarrow ZZ$  searches carried out at Atlas [60] and CB'S [61].

<sup>2</sup> The CDF experiment was not expected to see a significant Higgs boson signal in the final state considered.

<sup>3</sup> In particular it is fundamental to understand if a Higgs boson like particle observed has the properties predicted by the SM or has different properties, ascribable to an exotic model.



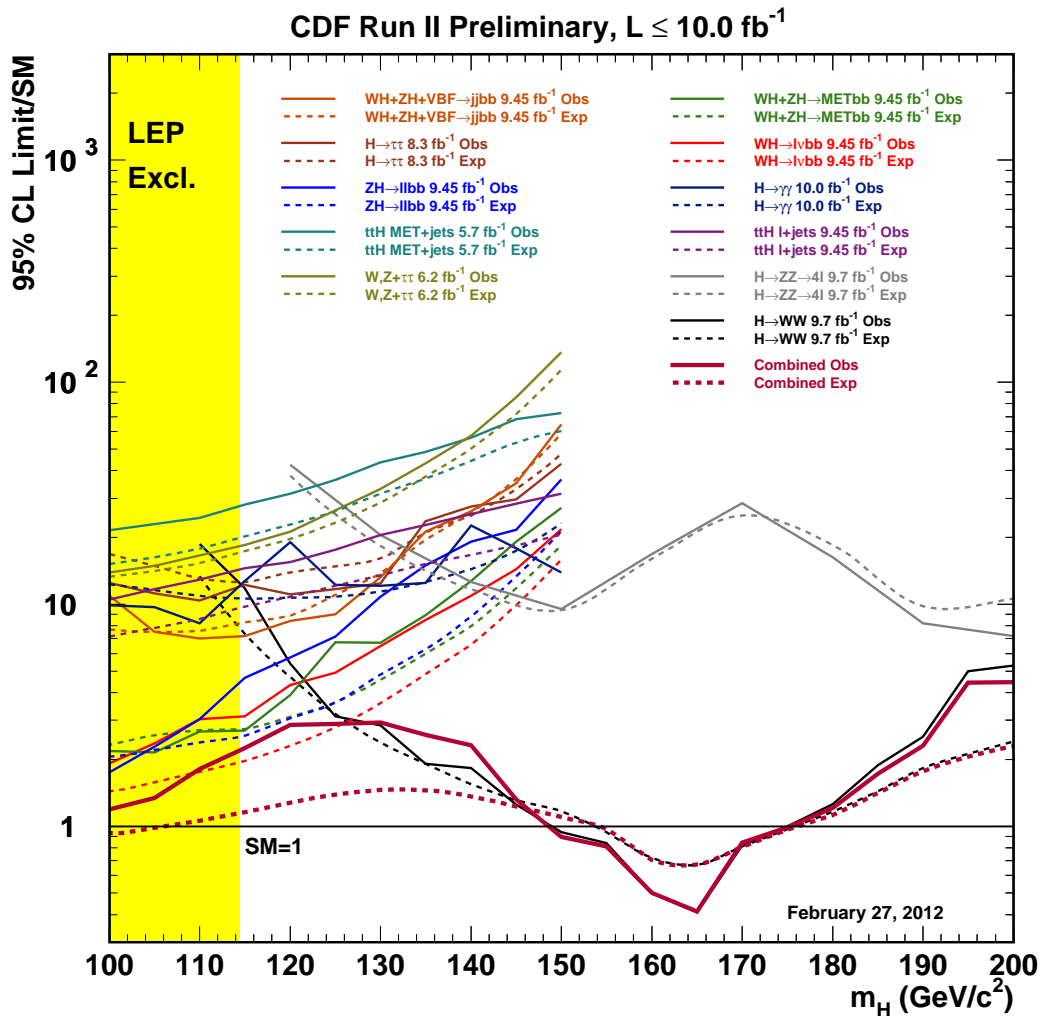


Figure 7.5: Summary of the CDF 95% C.L. upper limits on the Higgs production cross section in the different final state considered.

Both Atlas and CMS observed an excess of data in the  $m_{4\ell}$  spectrum around  $125 \text{ GeV}/c^2$ , that give a hint of the possible existence of the Higgs boson. An excess of data compatible with this has been found also in the analyses considering other final states. The Atlas and CMS overall sensitivity to the Higgs boson production is shown in Figure 7.8(a) and 7.8(b) respectively; the better sensitivity with respect to the Tevatron allowed to exclude at 95% C.L. almost the entire  $m_H$  range investigated (up to  $\sim 600 \text{ GeV}/c^2$ ), but the region  $m_H \sim 125 \text{ GeV}/c^2$ .

In this Higgs boson mass region both Atlas and CMS analyses reported an observed signal compatible with being due to a SM Higgs boson, as it is shown by the fit of the signal cross section in Figure 7.9. The investigation on the Higgs boson properties will continue at the LHC experiments with the aim of precisely measure its mass and the couplings to the SM particles. With a sample of the order of  $100 \text{ fb}^{-1}$  Atlas and CMS will have enough precision to determine these missing parameters of the SM.

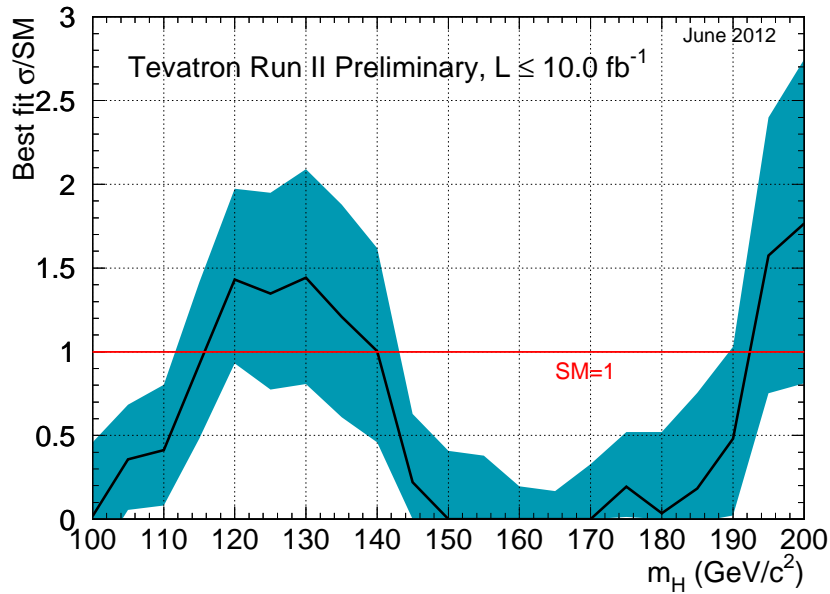
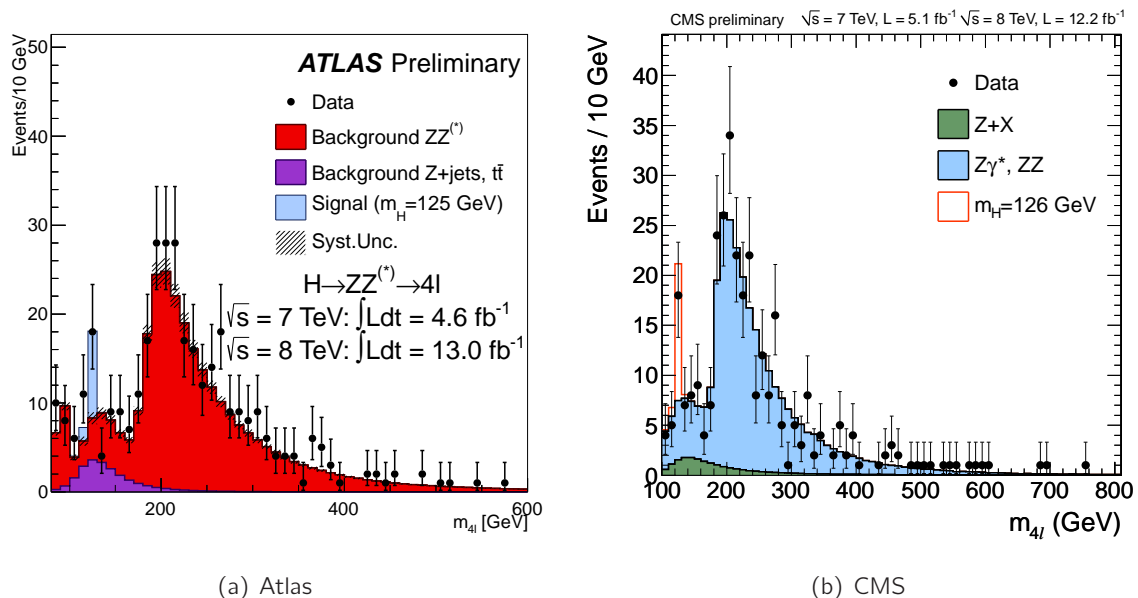


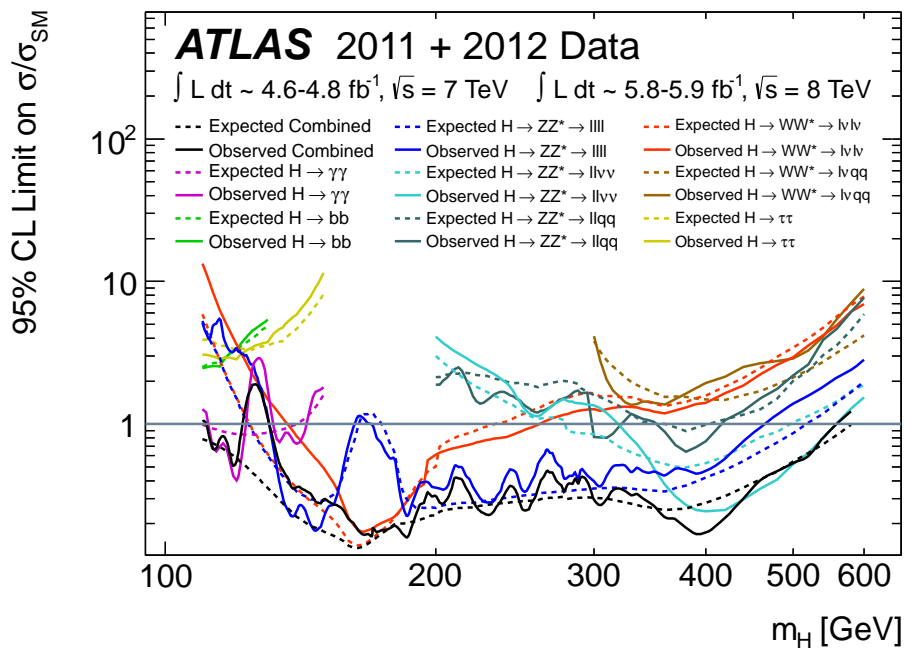
Figure 7.6: The best fit signal cross section of all the CDF and D0 search channels combined, shown as a ratio to the SM cross section as a function of the Higgs boson mass. The horizontal red line drawn at 1.0 correspond to the SM prediction.



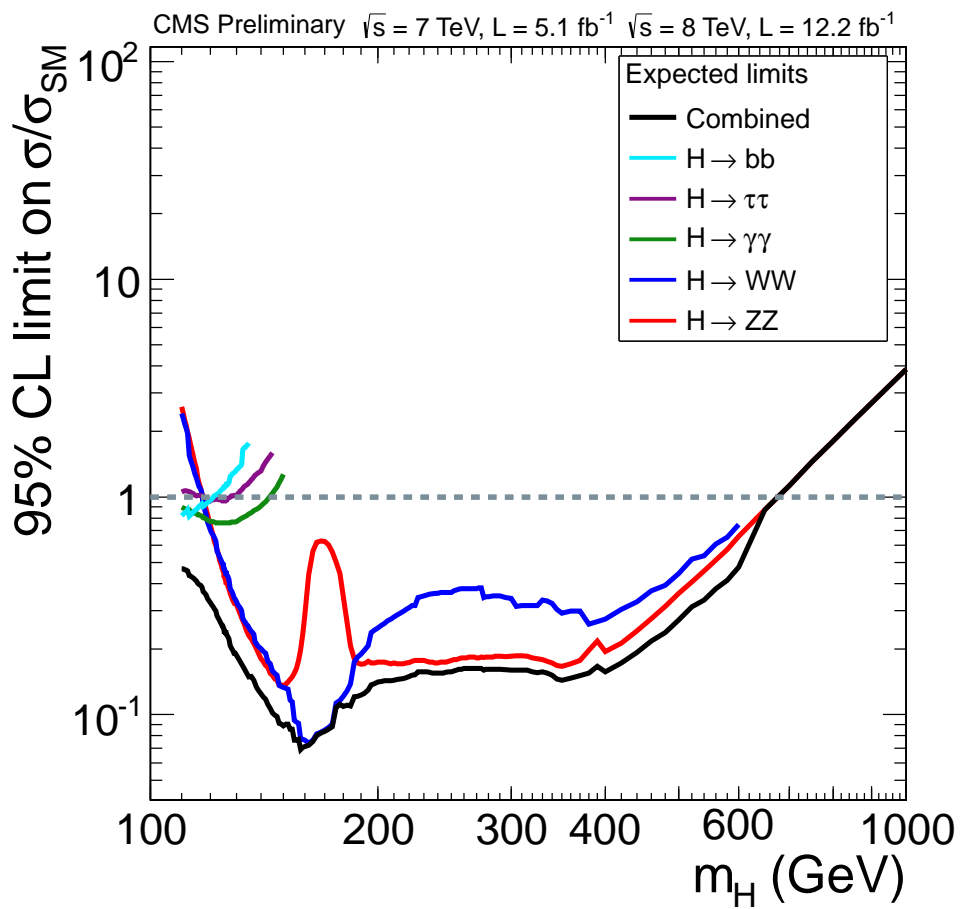
(a) Atlas

(b) CMS

Figure 7.7: Four lepton invariant mass distributions for the data collected at  $\sqrt{s} = 7$  and 8 TeV from the (a) Atlas and (b) CMS experiment, in the  $H \rightarrow ZZ$  search. In the plot is shown the expected signal for a Higgs boson having  $m_H \simeq 125 \text{ GeV}/c^2$ .



(a) Atlas



(b) CMS

Figure 7.8: Summary of the (a) Atlas and (b) CMS 95% C.L. upper limits on the Higgs production cross section in the different final state considered.

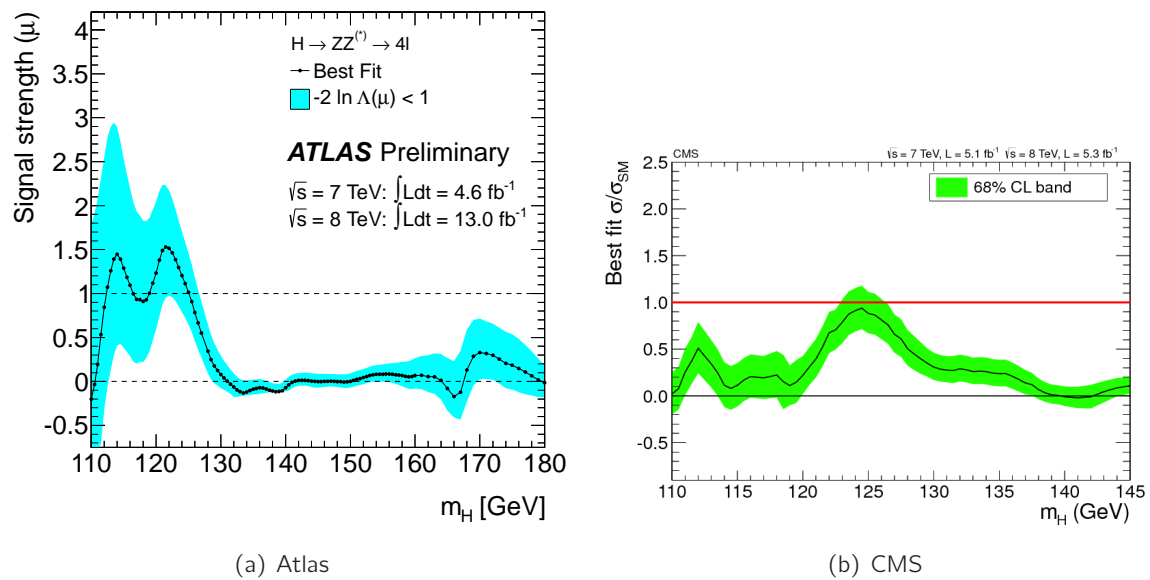


Figure 7.9: Atlas and CMS signal cross section fit, shown as a ration to the SM cross section as a function of the Higgs boson mass. The horizontal red line drawn at 1.0 correspond to the SM prediction.

## Conclusions

In this Thesis we reported the study of the production of a pair of massive  $Z$  bosons in proton antiproton collisions at  $\sqrt{s} = 1.96$  TeV. The full sample of data collected by CDF ( $9.7 \text{ fb}^{-1}$ ) has been exploited to measure precisely the production cross section of this process through its leptonic decays.

We measured the associated  $Z$  pair production cross section through two different leptonic decay modes, where one of the  $Z$  decays to a pair of charged leptons and the second one to either a pair of charged leptons or a pair of neutrinos. The statistical combination of these results give a measurement of the  $ZZ$  production cross section

$$\sigma(p\bar{p} \rightarrow ZZ) = 1.38 \pm 0.19(\text{stat.})_{-0.19}^{+0.20}(\text{syst.}) = 1.38_{-0.27}^{+0.28} \text{ pb}$$

which is in agreement with the Standard Model prediction and is the current best measurement of this process at the Tevatron. The combination of this result with the D0 results in the same final states would provide a measurement with an accuracy better than 20%, which will contribute to the Tevatron legacy.

The experiments collecting data from the LHC  $pp$  collisions at  $\sqrt{s} = 7$  and 8 TeV (Atlas and CMS) already collected and analyzed more than  $5 \text{ fb}^{-1}$  at each center-of-mass energy. Thanks to the higher expected cross sections for diboson production each experiment collected during the first years of data taking a remarkable sample of  $ZZ$  events that allowed a measurement of its production cross section with a similar precision to the Tevatron one. The production cross section measurement is although only the begin of a set of possible further studies of the  $ZZ$  diboson sample. In a clean  $ZZ \rightarrow \ell\ell'\ell'$  sample is possible to study the  $ZZ$  kinematic properties, given the complete reconstruction of the massive boson four-momenta in this final state. The study of the angular correlation between the two  $Z$  bosons as well as differential cross sections can provide information about the Trilinear Gauge Couplings (TGC) [1] and production mechanisms. To begin with, we exploited the information obtained in the sample of four charged lepton events to investigate the CDF sensitivity in the search for the Standard Model Higgs boson decaying to a pair of  $Z$  bosons,  $p\bar{p} \rightarrow H \rightarrow ZZ^{(*)} \rightarrow \ell\ell'\ell'$ .

Depending on the mass of the Higgs boson CDF expect to be sensitive to a Higgs boson signal produced with a cross section  $\sim 10$  times larger than the one predicted from the Standard Model. In the sample of data considered no evidence of the Higgs boson production has been seen, and we exclude at 95% C.L. its production with a cross section larger than 9.5 and 7.2 times the Standard Model prediction for  $m_H = 150$  and 200 GeV/ $c^2$  respectively [91]. This result has been combined with CDF and D0 searches for the Higgs boson that set altogether the Tevatron best

sensitivity to this process. A similar approach has been carried out at LHC, where the search for  $pp \rightarrow H \rightarrow ZZ^{(*)} \rightarrow 4\ell$  significantly contributed to the Higgs boson searches, that led up to the observation of a new particle compatible with the Standard Model Higgs boson with  $m_H \sim 125 \text{ GeV}/c^2$ .

The  $ZZ^{(*)} \rightarrow 4\ell$  decay is a really clean sample where to investigate all the properties of the newly observed particle. After the determination of the value of its mass it is fundamental to understand whether this particle behaves exactly as predicted from the Standard Model or if it can be identified as belonging to an extended theoretical model. In particular the kinematic properties of the four charged lepton decays can be used to investigate  $CP$  violation<sup>1</sup> in the massive boson sector (Higgs sector) [92]. The only neutral scalar of the Standard Model (the Higgs boson) is predicted to have  $J^{PC} = 0^{++}$ ; various extensions of the Standard Model predict several Higgs-like particles with different  $CP$  properties<sup>2</sup>. To determine precisely if the observed particle is *the* Standard Model Higgs boson is needed to establish the  $CP$  eigenvalues for the Higgs boson if  $CP$  is conserved and measure the mixing between  $CP$ -even and  $CP$ -odd states if it is not. A  $CP$  violation in this sector of the Standard Model may be an alternative source of  $CP$  violation that could explain the observed baryon-antibaryon asymmetry in the universe.

A broad variety of experimental measurements can be carried out starting from the results described in this Thesis, that will be objects of investigation in the next years at the high-energy colliders.

---

<sup>1</sup> The violation of the Charge and Parity symmetry has already been observed in the flavor sector of the Standard Model.

<sup>2</sup> For example Minimal SuperSymmetric Model (MSSM) has two  $CP$ -even and two  $CP$ -odd states.

---

# Lepton Reconstruction Categories

In the following sections we report the detailed information about the lepton identification in the detector, the different categories used in the analyses and mentioned in this Thesis, and *lepton-like* objects used for lepton ID, trigger, misidentification efficiency calculation.

## A.1 Reconstructed Lepton Categories

We exploit the maximum of the detector coverage for lepton identification by defining mutually exclusive categories of electrons and muons. Leptons belonging to each categories must satisfy different sets of requirements, optimized according to the reconstruction efficiency in particular regions of the detector. All the categories used are listed in Table A.1 and described in more details in the following sections.

### A.1.1 Tight Central Electrons

We reconstruct electrons in the central electromagnetic calorimeter ( $|\eta| < 1.1$ ) applying to track and energy cluster the requirements listed in Table A.2, where the different variables have the meaning described in the following.

**Region:** A flag indicating if the track is fiducial to the central or plug calorimeters. This flag is set by a specific routine during the offline processing software.

**Fiducial:** The geometric correspondence of a track extrapolation to a given subdetector.

**Track  $p_T$ :** The transverse component of the momentum which is measured explicitly using the track curvature.

**Track  $z_0$ :** The longitudinal ( $z$ ) position of the track where it intersects the beamline

**Axial and Stereo SL:** The number of axial and stereo superlayers in the COT which have at least 5 hits associated with this track.

**Conversion flag:** A routine is implemented to identify electrons which may have come from photon conversion. These electron candidates have their conversion flag set to one and rejected.

---

<b>TCE</b>	Central electron with $ \eta_{det}  \leq 1.1$ identified by a set of requirements
<b>LBE</b>	Central electron with $ \eta_{det}  \leq 1.1$ reconstructed with a Likelihood-based algorithm
<b>PHX</b>	Forward electron based on Silicon tracking and calorimeter information, with $1.2 \leq  \eta_{det}  \leq 2.0$
<b>PLBE</b>	Forward electron with $1.2 \leq  \eta_{det}  \leq 2.0$ , reconstructed with a Likelihood-based algorithm
<b>CMUP</b>	Central muon with hits in both the CMU and CMP chambers, and $ \eta_{det}  \leq 0.6$
<b>CMU</b>	Central muon with hits only in CMU chamber, pointing to detector regions not covered by CMP
<b>CMP</b>	Central muon with hits only in CMP chamber, pointing to detector regions not covered by CMU
<b>CMX</b>	Intermediate muon with hits in the CMX arches muon chambers, with $0.65 \leq  \eta_{det}  \leq 1.0$
<b>CMX-MsKs</b>	Intermediate muon with hits in the CMX Miniskirt and Keystone chambers, with $0.65 \leq  \eta_{det}  \leq 1.0$
<b>BMU</b>	Forward muon with hits in the IMU detector, with $1.1 \leq  \eta_{det}  \leq 1.5$
<b>CMIOCES</b>	Central muon with no reconstructed track segment in muon chambers, relying on energy deposit requirements in the central calorimeter
<b>CMIO PES</b>	Forward muon with no reconstructed track segment in muon chambers, relying on energy deposit requirements in the forward calorimeter
<b>CrkTrk</b>	High- $p_T$ isolated tracks pointing to uninstrumented regions of the detector. Considered in the analysis to be either electron or muon

Table A.1: Summary of the lepton types (*categories*) used in the analysis described in this thesis.

<b>Tight Central Electrons (TCE)</b>	
Region	Central ( $ \eta  < 1.1$ )
Fiducial	Track fiducial to CES
Track $P_T$	$\geq 10$ or $\geq 5$ if $E_T < 20$ (GeV)
Track $ z_0 $	$\leq 60$ cm
# Axial SL	$\geq 3$ with $\geq 5$ hits
# Stereo SL	$\geq 2$ with $\geq 5$ hits
Conversion Flag	$\neq 1$
Isolation/ $E_T$	$\leq 0.1$
$E_{HAD}/E_{EM}$	$< 0.055 + 0.00045 \cdot E$
$L_{shr}$	$\leq 0.2$
$E/P$	$< 2.5 + 0.015 \cdot E_t$
CES $\Delta X$	$-3 \leq q \cdot \Delta X \leq 1.5$
CES $\Delta Z$	$< 3$ cm

Table A.2: Tight Central Electron (TCE) identification requirements.

**Isolation** / $E_T$ : The energy deposited in the calorimeter in a cone of radius  $\Delta R \leq 0.4$  around the electron cluster excluding the energy of the electron cluster divided by the  $E_T$  ( $p_T$ ) of the



electron candidate.

$E_{HAD}/E_{EM}$ : The ratio of energy which is deposited in the hadronic (CHA or WHA) portion of the calorimeter to the energy deposited in the electromagnetic (CEM or PEM) portion of the calorimeter.

$L_{shr}$ : A variable that compares the lateral shower profile in towers next to the seed tower to an expected profile given by

$$L_{shr} = 0.14 \frac{\sum_i (M_i - P_i)}{\sqrt{(0.14\sqrt{E_{EM}})^2 + \sum_i (\Delta P_i)^2}} \quad (\text{A.1})$$

where  $i$  denotes the adjacent towers,  $M_i$  the measured energy, and  $P_i$  the predicted energy in the  $i$ -th tower.

**E/P**: The ratio of the energy measured in the calorimeter to the momentum calculated from the measurement of the track curvature.

**CES  $\Delta X$** : The difference in the  $r - \phi$  plane between the best CES match and the COT beam-constrained track extrapolation to the CES.

**CEM  $\Delta Z$** : The longitudinal difference between the best CES match and the COT track extrapolation to the CES.

**J<sub>PES</sub>**: The pseudo-rapidity as measured by the best matched PES cluster.

**PEM  $3 \times 3$  Fit**: A  $\chi^2$  fit to the electron test beam data of 9 PEM towers.

**$\sigma_{PES}^2$** : A  $\chi^2$  fit to the electron test beam data for shower-maximum profile.

**PES  $5 \times 9$  U/V**: The ratio of the central 5 tower energy to the total 9 tower energy.

**$\Delta R$  (PES, PEM)**: The difference in the  $r - \phi$  plane between the best PES match and the PEM measurement.

**Track Matched**: PHX electrons must have a track that is matched to the PEM cluster and event vertex.

**# of Silicon hits**: The number of the hits in the silicon detector associated with a specific track. The maximum number of hits is 8 (for L00, SVX and ISL combined).

The tight requirements listed in Table A.2 select a high-purity electron sample in the central part of the detector.

### A.1.2 Likelihood Based Electrons

To increase the acceptance on central electron reconstruction we loosen the tight cuts of TCE reconstruction to those summarized in Table A.3.

To keep the fake electron rate at a reasonably low level we combine all the information on electron candidates into a more powerful discriminating variable. We built a likelihood function from the *identification variables* described in Section A.1.1 of the form:

$$\mathcal{L}(\vec{x}) = \frac{L_{sig}}{L_{sig} + L_{bckg}} = \frac{\prod_{i=1}^N P_i^{sig}(x_i)}{\prod_{i=1}^N P_i^{sig}(x_i) + \prod_{i=1}^N P_i^{bckg}(x_i)} \quad (\text{A.2})$$

Central Electrons (LBE)	
Region	Central ( $ \eta  < 1.1$ )
$E_{HAD}/E_{EM}$	$\leq 0.125$
Track $p_T$	$\geq 10$ or $\geq 5$ GeV if $E_T < 20$ GeV
Track $ z_0 $	$\leq 60$ cm
Isolation/ $E_T$	$\leq 0.3$
Conversion	false
Likelihood ( $\mathcal{L}$ )	$\mathcal{L} > 0.9$

Table A.3: Likelihood-based central electron (LBE) identification requirements.

where

$\mathbf{x}_i$ : is the  $i$ -th identification variable used in the likelihood. These are the variable described in Section A.1.1

$\mathbf{N}$ : is the number of ID variables

$\mathbf{P}_i^{\text{sig}}(\mathbf{x}_i), \mathbf{P}_i^{\text{bckg}}(\mathbf{x}_i)$  are the functions that give the probability to obtain the value  $x_i$  for the  $i$ -th ID variable given a signal (*real*) or background (*fake*) electron, often referred as *templates*.

By definition, the value of  $\mathcal{L}$  is restricted within the range  $[0,1]$ . We use data to build the signal and background templates. This is done by selecting samples dominated by real high- $p_T$  electrons for signal templates and dominated by fake electrons for background templates. Real electrons are obtained selecting  $Z$  events requiring a fully reconstructed TCE and a looser one (*probe*) with an invariant mass in within  $76 \leq M_{\ell\ell} \leq 106$  GeV/ $c^2$  and opposite charge. The *probe* is required to have an identified electron-like cluster associated with a track satisfying looser selections reported in Table A.14. This sample is dominated by Drell-Yan events with real leptons. The ID variables of probe objects are used to make the signal template. When both leptons are identified as TCE, information of both are included in the template.

Fake electrons are obtained from a jets data samples. This sample is selected by requiring at least one jet with energy greater than 20 GeV (*JET20* dataset<sup>1</sup>). One reconstructed jet with  $E_T > 20$  GeV and a loose *fakeable* electron (see Section A.2 for definition) per event are required, with an invariant mass out of the  $Z$  peak region to remove Drell-Yan contribution:  $M(\ell\ell) < 76$  or  $M(\ell\ell) > 106$  GeV/ $c^2$ . The fakeable object is required not to be the leading  $E_T$  jet to remove any trigger bias. The ID variables of fakeable objects are used to build the background template.

The resulting likelihood distribution for real and fake electrons are shown in Figure A.1, left and right respectively.

At this stage both signal and background templates are not determined using pure real/fake electron samples as it can be seen from the small contamination in Figure A.1. We first measure and then subtract the contribution of fake electrons in the signal template normalizing the extracted likelihood background shape using the first bin of the signal distribution. This gives us an estimation of the shape and magnitude of the residual fake contribution in our signal template, which is found to be  $\sim 3\%$ . The residual contribution of real electrons in the background template is evaluated using inclusive  $Z$  and  $W$  Monte Carlo samples, which are expected to be the major source of real lepton in that sample. It is found and subtracted a contribution from real lepton of the order of  $\sim 1\%$ .

We select as LBE leptons those for which  $\mathcal{L} > 0.9$ .

<sup>1</sup> See Appendix B for details about *JET20* collected sample.

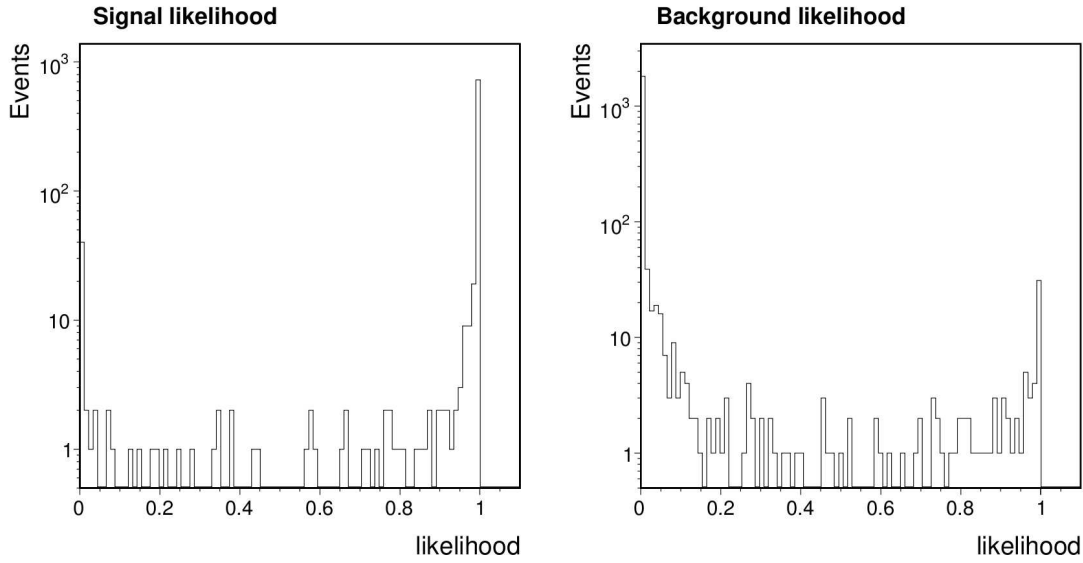


Figure A.1: Likelihood distribution for real (left) and fake electrons (right).

### A.1.3 Forward Electrons

Similar to what is done in the central part of the detector, we reconstruct two categories of electrons in the forward part of the detector, for  $1.2 \leq |\eta_{det}| \leq 2.0$ . Identification through the set of selections listed in Table A.4 allows the tight identification of forward electrons, commonly referred to as PHOENIX electrons (PHX). This reconstruction is based on an energy cluster with less than 5% of the total energy being in the hadronic calorimeter, with a matching track pointing to it. The considered  $\eta$  range is not fully covered by the COT tracking, hence the purity is kept high by requiring hits at least in three layers of the Silicon detector. Starting from the variables listed in Table A.4 we built a likelihood similar to the one described in Equation A.2 to identify loose forward electrons, referred to as Plug Likelihood-based Electrons (PLBE), selecting those that have  $\mathcal{L} > 0.9$ .

<b>Forward Electrons (PHX)</b>	
Region	Plug ( $1.1 <  \eta  < 2.0$ )
$E_{HAD}/E_{EM}$	$< 0.05$
PEM $3 \times 3$ Fit	true
$\chi^2_{PES}$	$\leq 10$
PES $5 \times 9$ U	$\geq 0.65$
PES $5 \times 9$ V	$\geq 0.65$
Isolation/ $E_T$	$\leq 0.1$
$\Delta R(\text{PES}, \text{PEM})$	$\leq 0.3$
Track matched	true
# of Silicon hits	$\geq 3$
Track $ z_0 $	$\leq 60$ cm

Table A.4: Forward (PHX) electron identification requirements.

### A.1.4 Central Muons

Muons traverse the entire CDF detector and leave hits in the outer muon chambers, in cases where they point to a region which is covered by them. In the central part of the detector,  $|\eta| \leq 1.0$ , the muon detector system is composed by several subdetectors. We reconstruct five categories of muons based on the inner track and outer segment in the detector. All these must satisfy the requirements summarized in Table A.5, where the several identification variables have the meaning described in the following.

Base Muon Selection	
$E_{EM}$	$< 2 + \max(0, (p - 100) \cdot 0.0115)$
$E_{HAD}$	$< 6 + \max(0, (p - 100) \cdot 0.028)$
Isolation/ $P_T$	$\leq 0.1$
# Axial SL	$\geq 3$ with $\geq 5$ hits
# Stereo SL	$\geq 2$ with $\geq 5$ hits
Track $ z_0 $	$< 60$ cm
Track $ d_0 $	$< 0.2$ cm ( $< 0.02$ cm for tracks with silicon)
$\chi^2/dof$	$< 3.0$ ( $< 4.0$ if Run $> 186598$ )

Table A.5: Base identification requirements for all muon categories.

**CM(U|P|X)**  $x_{fid}$ ,  $z_{fid}$ : The extrapolation of the track to the relevant muon detector is required to be fiducial to the detector and in the case of CMX must also not be within 3 cm in  $z_{fid}$  of the edge of the detector<sup>2</sup>.

**$\Delta X_{CM(U|P|X)}$** : The distance between the actual stub in a given muon detector and the track position extrapolated to that detector.

**$a_{COT}$** : The radius at which the track cross the front side of the COT. With this request we ensure that these muons exited the COT from its side surface and can pass CMX trigger track requirements.

**$z_{BMU}$** : The  $z$  position of the track at the BMU radius ( $r_{BMU} = 391$  cm) corrected for the track's  $z_0$  position according to the following

$$z_{BMU} = z_0 + r_{BMU} \cdot (1 - e^{-2 \cdot \eta}) / 2e^{-2 \cdot \eta}.$$

It is required to the track to point to the IMU muon chambers.

**# of Stub hits**: The number of hits in the fiducial muon chamber associated with a specific track.

**$d_0$** : The distance of closest approach of the fitted track to the beamline.

**$\emptyset^2$** : The chi-squared compares the fitted track to the hit information in the tracking detectors.

**Curvature significance**: The measured track curvature divided by the curvature error.

<sup>2</sup> These coordinates refer to the face of the specific muon detector and not the CDF II coordinate system.

In addition to this common requirements we define six different categories of muons, depending on the region of the detector that the extrapolated track is pointing to. We measure the location of an extrapolated muon track candidate with respect to the drift direction (local  $x$ ) and wire axis (local  $z$ ) of a given chamber. We do not take into account possible multiple scattering in the extrapolation. We refer to these requirements as *fiduciality* of the track to the given muon detector. The fiduciality requirements ensure that all the categories are non-overlapping: a given muon cannot be classified into two different categories. In the region  $|\eta| \leq 0.6$  tracks can point to the CMU or CMP chambers (see geometry in Section 3.2.3): muons with reconstructed track-segment in the CMU or CMP detector and not in the other one fall into two separate categories (respectively named CMU and CMP) while those with hits in both are called CMUP muons. These must satisfy the specific requirements summarized in Tables A.6, A.7, and A.8 respectively. A trigger path has always been active base on the coincidence of CMU and CMP hits (i.e. CMUP muons), while from a certain period on during data taking, trigger paths have been associated also to track segment in CMP and CMU separately (starting on Dec. '06 and Dec. '08 respectively).

<b>CMU Muon</b>	
CMU Fiduciality	$x_{fid} < 0, z_{fid} < 0$ cm <i>Not CMP Fiducial, not CMX Fiducial</i>
$\Delta X_{CMU}$	$< 7$ cm
Good trigger	run $\geq 270062$

Table A.6: CMU-*only* muons identification requirements.

<b>CMP Muon</b>	
CMP Fiduciality	$x_{fid} < 0, z_{fid} < -3$ cm <i>Not CMU Fiducial, not CMX Fiducial</i>
$\Delta X_{CMP}$	$< \max(6.0, 150.0/P_T)$ cm
$\phi$ -gaps	$\phi \pmod{15^\circ} < 2$ or $\phi \pmod{15^\circ} > 13$ No Bluebeam for run $\leq 154449$
Good trigger	run $\geq 229764$

Table A.7: CMP-*only* muons identification requirements.

<b>CMUP Muon</b>	
CMU Fiduciality	$x_{fid} < 0, z_{fid} < 0$ cm
CMP Fiduciality	$x_{fid} < 0, z_{fid} < -3$ cm
$\Delta X_{CMU}$	$< 7$ cm
$\Delta X_{CMP}$	$< \max(6.0, 150.0/P_T)$ cm

Table A.8: CMUP muons identification requirements.

### A.1.5 Intermediate Muons

In the intermediate  $\eta$  region of the detector ( $0.6 < |\eta| < 1.1$ ) the muon detection system is composed by four main arches (two for  $\eta > 0$ , two for  $\eta < 0$ ) of drift chambers (CMX), two

chambers in the lower part of the detector (one for  $\eta > 0$ , one for  $\eta < 0$ ) called MiniSkirts and a small chamber at the top (for  $\eta < 0$ ) called KeyStone. We select candidate muons fiducial these chambers, which have hits in them and pass the requirements reported in Table A.5 and additional ones in Table A.9; we categorize them as CMX or CMX-MsKs if their track points to the arches or to the Miniskirt/Keystone respectively. In particular, an explicit geometrical requirement is made to require that the particle track crosses all the layers of the COT, hence exiting from its lateral surface, to increase the track reconstruction precision. From November '06 a trigger path has been associated to CMX arches, miniskirt and keystone.

<b>CMX &amp; MsKs Muon</b>	
CMX Fiduciality	$x_{fid} < 0, z_{fid} < -3$ cm
$\Delta X_{CMX}$	$< \max(6, 125/P_T)$ cm
$\rho_{COT}$	$> 140$ cm
Good trigger	run $\geq 227704$
Arches	$0^\circ < \phi < 75^\circ$ OR $115^\circ < \phi < 225^\circ$ OR $315^\circ < \phi < 360^\circ$
Ms-Ks	$(75^\circ < \phi < 105^\circ$ and $\eta < 0)$ OR $225^\circ < \phi < 315^\circ$

Table A.9: CMX Arches and Miniskirt-Keystone muons identification requirements.

### A.1.6 Forward Muons

In the forward (and backward) part of the detector  $1.1 < |\eta| < 1.4$  we reconstruct as BMU muons the particles satisfying the requirements listed in Table A.10. These must have a good quality track with hits in the silicon tracker and in more than 60% of the COT wire layers. From the extrapolation of the track is required a small energy deposit in the plug calorimeters and a matching stub in the IMU drift chamber. No trigger path is associated with this lepton category, hence it is a non-triggerable one.

<b>BMU</b>	
$E_{HAD} + E_{EM}$	$> 0.1$ GeV
Isolation/ $P_T$	$\leq 0.1$
COT Hit Fraction	$> 0.6$
Curvature significance	$> 12$
# of Silicon hits	$\geq 3$
Fiduciality	PES Fiducial IMU Fiducial
# of Stub hits	$\geq 2$
$z_{BMU}$	$471.6 \lesssim z_{BMU} \lesssim 766.6$ OR $-433.0 \lesssim z_{BMU} \lesssim -764.7$

Table A.10: BMU muons identification requirements.

### A.1.7 Stubless Muons

The detector acceptance in muon reconstruction is increased by promoting to candidate muons objects with no reconstructed track segment in any of the muon chambers. The reconstruction of these *stubless* muons is based on the identification of a good quality track in the inner tracking system and m.i.p. calorimetric requirements. Two different categories of stubless muons are defined according to the fiduciality of the track to the central or plug calorimeter, named CMIOCES and CMIOCES respectively. These must pass the selections reported in Tables A.11 and A.12, and must belong to none of the stubbed muon categories defined before.

<b>CMIOCES Muon</b>	
Uniqueness	Not a CMUP/CMP CMX/MsKs
Fiducial	Track fiducial to CES
# Axial SL	$\geq 3$ with $\geq 5$ hits
# Stereo SL	$\geq 3$ with $\geq 5$ hits
$\chi^2/dof$	$< 3.0$
Track $z_0$	$\leq 60$ cm
Track $d_0$	$\leq 0.02$ cm
Iso/ $P_T$	$\leq 0.1$
$E_{EM} + E_{Had}$	$> 0.1$ GeV

Table A.11: CMIOCES muons identification requirements.

<b>CMIOCES Muon</b>	
Uniqueness	Not a BMU
Fiducial	Track fiducial to PES
COT hit fraction	$> 0.6$
# Stereo SL	$\geq 3$ with $\geq 5$ hits
Track $z_0$	$\leq 60$ cm
Track $d_0$	$\leq 0.02$ cm
Curvature significance	$> 12.0$
Iso/ $P_T$	$\leq 0.1$
$E_{EM}$	$2 + \max(0, (p - 100) \cdot 0.0115)$ GeV
$E_{Had}$	$6 + \max(0, (p - 100) \cdot 0.028)$ GeV
$E_{EM} + E_{Had}$	$> 0.1$ GeV

Table A.12: CMIOCES muons identification requirements.

### A.1.8 Isolated Tracks

To maximize the detector acceptance is defined a separate lepton category for tracks that fall in uninstrumented parts of the detector (*Cracks*). These candidate tracks has not to be fiducial to the central or plug calorimeters and must not belong to any stubbed muon category. The selection criteria, reported in Table A.13, are similar to those for CMIOCES muons. No explicit requirement on the calorimetric energy deposit is applied since these tracks point closely to calorimeter cracks; only a low calorimetric activity in the track surroundings is required. We

explicitly veto that the track is compatible with coming from a conversion electron. Both muons and electrons can belong to the category, hence this is kept separated from the other ones, and *CrkTrk* leptons will be considered in the analysis as having either flavors.

<b>CrkTrk</b>	
Isolation/ $P_T$	$\leq 0.1$ using CDF Muon or $\leq 0.1$ using nearest EM cluster, $\Delta R < 0.05$
# Axial SL	$\geq 3$ with $\geq 5$ hits
# Stereo SL	$\geq 3$ with $\geq 5$ hits
Track $ z_0 $	$< 60$ cm
Track $ d_0 $	$< 0.2$ cm ( $< 0.02$ cm with silicon)
$\chi^2/dof$	$< 3.0$
Uniqueness	Not a CMUP/CMP/CMX/MsKs/BMU
In Crack	Not CES or PES fiducial
Conversion	$\neq 1$

Table A.13: CrkTrk identification requirements.

## A.2 Lepton Probes and Denominator Objects

To evaluate the lepton Identification efficiencies outlined in Section 5.1.2 and the lepton misidentification rates (*fake rates*) described in Section 5.1.3 are exploited *loose lepton categories* referred to as *probe* and *denominator* objects respectively.

We defined central/forward electron probes, applying the requirements reported in Tables A.14 and A.15 respectively, and central/forward muon probes, applying the requirements reported in Tables A.16. As we mentioned in Section 5.1.2, the ID efficiency for forward electrons is evaluated separately for tracking and calorimeter identification, then combined; therefore we define a tracking and a calorimetric probe, satisfying the requirements reported in Table A.15.

<b>Central Electron Probe</b>	
Region	Central ( $ \eta  < 1.1$ )
Track $P_T$	$\geq 5$ GeV
Track $ z_0 $	$\leq 60$ cm
Conversion Flag	$\neq 1$

Table A.14: Central Electron Probe identification requirements.

In the calculation of each lepton category identification efficiency we match to each category the corresponding generic probe, with additional requirements on the fiduciality of this one to the considered region of the detector. The matching between each lepton category and the corresponding probe are summarized in Table A.17.

Similarly, for the denominator objects exploited in the lepton fake rate calculation (see Section 5.1.3), we define generic electron and muons *fakeable* objects, applying then specific fiduciality requirement depending on the lepton category considered (see Table A.17). The central and forward fakeable electrons and muons are required to satisfy the selections reported in Tables A.18, A.19, A.20, and A.21.



<b>Forward Electron Tracking Probe</b>		<b>Forward Electron Calorimetric Probe</b>	
Region	Plug ( $1.2 <  \eta  < 2.0$ )	Region	Plug ( $1.2 <  \eta  < 2.0$ )
$E_{HAD}/E_{EM}$	$< 0.05$	$E_{HAD}/E_{EM}$	$< 0.05$
Isolation/ $E_T$	$\leq 0.1$	Track $ z_0 $	$\leq 60$ cm
PEM $3 \times 3 \chi^2$	$\leq 10$	Track matched	true
PES $5 \times 9$ U	$\geq 0.65$		
PES $5 \times 9$ V	$\geq 0.65$		
$\Delta R(\text{PES, PEM})$	$\leq 0.3$		

Table A.15: Forward Electron Probe identification requirements.

<b>Central Muon Probe</b>		<b>Forward Muon Probe</b>	
Fiduciality	Not PES fiducial	Fiduciality	PES fiducial
# Axial SL	$\leq 3$ with $\geq 5$ hits	COT Hit Fraction	$\geq 0.6$
# Stereo SL	$\leq 2$ with $\geq 5$ hits	Track $ z_0 $	$\leq 60$ cm
Track $ z_0 $	$\leq 60$ cm		

Table A.16: Central and Forward Muon Probe identification requirements.

Lepton Category	Probe/Denominator Object
Electrons	
TCE	Central Electron Probe/Denominator
LBE	Central Electron Probe/Denominator
PHX	Forward Electron Probe/Denominator
PLBE	Forward Electron Probe/Denominator
Muons	
CMUP	Central Muon Probe/Denominator + fiducial to CMU and CMP chambers
CMU	Central Muon Probe/Denominator + fiducial to CMU and not to CMP
CMP	Central Muon Probe/Denominator + fiducial to CMP and not to CMU
CMX	Central Muon Probe/Denominator + fiducial to CMX arches and $\rho_{COT} > 140$ cm
CMX-MsKs	Central Muon Probe/Denominator + fiducial to CMX Miniskirt or Keystone and $\rho_{COT} > 140$ cm
BMU	Forward Muon Probe/Denominator + fiducial to BMU and PES
CMIOCES	Central Muon Probe/Denominator + fiducial to CES
CMIO PES	Forward Muon Probe/Denominator + fiducial to PES
Isolated Tracks	
CrkTrk( $e$ )	Central Electron Probe/Denominator + not fiducial to PES or CES
CrkTrk( $\mu$ )	Central Muon Probe/Denominator + not fiducial to PES or CES

Table A.17: Tight lepton and corresponding loose objects used to measure the lepton identification efficiencies and misidentification rates.

Fakeable Central Electron	
Region	Central ( $ \eta  < 1.1$ )
Conversion Flag	$\neq 1$
Isolation/ $E_T$	$\leq 0.3$
$E_{HAD}/E_{EM}$	$< 0.125 + 0.00045 \cdot E$

Table A.18: Fakeable Central Electron identification requirements.

<b>Fakeable Forward Electron</b>	
Region	Plug ( $1.2 <  \eta  < 2.0$ )
$E_{HAD}/E_{EM}$	$< 0.125$
Isolation/ $E_T$	$\leq 0.3$
# of Silicon hits	$\geq 3$
Track $ z_0 $	$\leq 60$ cm
Veto	Not a PHX

Table A.19: Fakeable Forward Electron identification requirements.

<b>Fakeable Central Muon</b>	
Fiduciality	Not PES fiducial
# Axial SL	$\leq 2$ with $\geq 5$ hits
# Stereo SL	$\leq 2$ with $\geq 5$ hits
Track $ z_0 $	$\leq 60$ cm
Track $ d_0 $	$\leq 0.2$ cm ( $< 0.02$ cm for tracks with Silicon)
Isolation/ $E_T$	$\leq 0.3$
$\chi^2/dof$	$\leq 3$ ( $\leq 4$ if Run $<186598$ )

Table A.20: Fakeable Central Muon identification requirements.

<b>Fakeable Forward Muon</b>	
Fiduciality	PES fiducial
COT Hit Fraction	$geq 0.6$
Track $ z_0 $	$\leq 60$ cm
Track $ d_0 $	$\leq 0.2$ cm ( $< 0.02$ cm for tracks with Silicon)
Isolation/ $E_T$	$\leq 0.3$
$\chi^2/dof$	$\leq 3$ ( $\leq 4$ if Run $<186598$ )

Table A.21: Fakeable Forward Muon identification requirements.

### A.3 Lepton reconstruction efficiencies and scale factors

Below are reported the lepton identification efficiencies evaluated on data and simulation for different grouped run periods, described in Section 5.1.2: Tables A.22, A.23, A.24, A.25, A.26, A.27, A.28, and A.29.

#### Run Period: 0

	data	MC	Scale Fac
LBE	$0.096 \pm 0.003$	$0.091 \pm 0.001$	$1.057 \pm 0.029$
TCE	$0.817 \pm 0.004$	$0.812 \pm 0.001$	$1.006 \pm 0.005$
PHXTrk	$0.862 \pm 0.004$	$0.865 \pm 0.001$	$0.997 \pm 0.005$
PHXPEM	$0.853 \pm 0.004$	$0.895 \pm 0.001$	$0.953 \pm 0.005$
PLBEPPEM	$0.087 \pm 0.003$	$0.075 \pm 0.001$	$1.158 \pm 0.044$
CrkTrk ( $e$ )	$0.668 \pm 0.006$	$0.760 \pm 0.001$	$0.879 \pm 0.009$
CMUP	$0.858 \pm 0.010$	$0.897 \pm 0.001$	$0.957 \pm 0.011$
CMU	$0.000 \pm 1.000$	$0.000 \pm 1.000$	$0.000 \pm 1.000$
CMP	$0.000 \pm 1.000$	$0.000 \pm 1.000$	$0.000 \pm 1.000$
CMX	$0.918 \pm 0.011$	$0.916 \pm 0.002$	$1.002 \pm 0.012$
CMXMsKs	$0.000 \pm 1.000$	$0.000 \pm 1.000$	$0.000 \pm 1.000$
BMU	$0.753 \pm 0.022$	$0.725 \pm 0.004$	$1.038 \pm 0.031$
CMIOCES	$0.364 \pm 0.007$	$0.347 \pm 0.001$	$1.049 \pm 0.020$
CMIOPEs	$0.651 \pm 0.012$	$0.683 \pm 0.002$	$0.954 \pm 0.018$
CrkTrk ( $\mu$ )	$0.709 \pm 0.010$	$0.757 \pm 0.002$	$0.937 \pm 0.014$

Table A.22: Lepton identification efficiencies for data and simulation for period 0. Errors are statistical only.

#### Run Period: 1 - 4

	data	MC	Scale Fac
LBE	$0.099 \pm 0.002$	$0.094 \pm 0.001$	$1.062 \pm 0.027$
TCE	$0.804 \pm 0.004$	$0.810 \pm 0.001$	$0.993 \pm 0.005$
PHXTrk	$0.878 \pm 0.004$	$0.874 \pm 0.001$	$1.004 \pm 0.004$
PHXPEM	$0.843 \pm 0.004$	$0.882 \pm 0.001$	$0.956 \pm 0.005$
PLBEPPEM	$0.096 \pm 0.003$	$0.078 \pm 0.001$	$1.241 \pm 0.041$
CrkTrk ( $e$ )	$0.688 \pm 0.006$	$0.779 \pm 0.001$	$0.883 \pm 0.008$
CMUP	$0.834 \pm 0.008$	$0.896 \pm 0.001$	$0.931 \pm 0.009$
CMU	$0.000 \pm 0.500$	$0.000 \pm 0.500$	$0.000 \pm 0.500$
CMP	$0.000 \pm 0.500$	$0.000 \pm 0.500$	$0.000 \pm 0.500$
CMX	$0.906 \pm 0.010$	$0.913 \pm 0.002$	$0.992 \pm 0.011$
CMXMsKs	$0.000 \pm 0.500$	$0.000 \pm 0.500$	$0.000 \pm 0.500$
BMU	$0.741 \pm 0.018$	$0.714 \pm 0.003$	$1.037 \pm 0.026$
CMIOCES	$0.355 \pm 0.006$	$0.339 \pm 0.001$	$1.048 \pm 0.018$
CMIOPEs	$0.579 \pm 0.012$	$0.601 \pm 0.002$	$0.964 \pm 0.020$
CrkTrk ( $\mu$ )	$0.759 \pm 0.009$	$0.786 \pm 0.002$	$0.966 \pm 0.012$

Table A.23: Lepton identification efficiencies for data and simulation for periods 1 to 4. Errors are statistical only.

**Run Period: 5 - 7**

	data	MC	Scale Fac
LBE	$0.105 \pm 0.003$	$0.097 \pm 0.001$	$1.091 \pm 0.034$
TCE	$0.791 \pm 0.005$	$0.803 \pm 0.001$	$0.984 \pm 0.006$
PHXTrk	$0.885 \pm 0.004$	$0.869 \pm 0.001$	$1.018 \pm 0.005$
PHXPEM	$0.832 \pm 0.005$	$0.876 \pm 0.001$	$0.949 \pm 0.006$
PLBEPem	$0.103 \pm 0.004$	$0.080 \pm 0.001$	$1.282 \pm 0.051$
CrkTrk ( $e$ )	$0.544 \pm 0.008$	$0.659 \pm 0.002$	$0.826 \pm 0.013$
CMUP	$0.820 \pm 0.011$	$0.893 \pm 0.002$	$0.919 \pm 0.013$
CMU	$0.000 \pm 0.577$	$0.000 \pm 0.577$	$0.000 \pm 0.577$
CMP	$0.000 \pm 0.577$	$0.000 \pm 0.577$	$0.000 \pm 0.577$
CMX	$0.907 \pm 0.013$	$0.908 \pm 0.002$	$0.999 \pm 0.015$
CMXMsKs	$0.000 \pm 0.577$	$0.000 \pm 0.577$	$0.000 \pm 0.577$
BMU	$0.738 \pm 0.023$	$0.715 \pm 0.004$	$1.031 \pm 0.033$
CMIOCES	$0.362 \pm 0.008$	$0.336 \pm 0.002$	$1.078 \pm 0.024$
CMIOPEs	$0.590 \pm 0.015$	$0.594 \pm 0.003$	$0.993 \pm 0.025$
CrkTrk ( $\mu$ )	$0.759 \pm 0.012$	$0.783 \pm 0.002$	$0.968 \pm 0.015$

Table A.24: Lepton identification efficiencies for data and simulation for periods 5 to 7. Errors are statistical only.

**Run Period: 8 - 10**

	data	MC	Scale Fac
LBE	$0.113 \pm 0.002$	$0.097 \pm 0.000$	$1.169 \pm 0.024$
TCE	$0.782 \pm 0.003$	$0.802 \pm 0.001$	$0.974 \pm 0.004$
PHXTrk	$0.878 \pm 0.003$	$0.878 \pm 0.001$	$1.000 \pm 0.004$
PHXPEM	$0.822 \pm 0.004$	$0.874 \pm 0.001$	$0.941 \pm 0.004$
PLBEPem	$0.112 \pm 0.003$	$0.082 \pm 0.001$	$1.359 \pm 0.035$
CrkTrk ( $e$ )	$0.522 \pm 0.006$	$0.647 \pm 0.001$	$0.806 \pm 0.009$
CMUP	$0.835 \pm 0.007$	$0.887 \pm 0.001$	$0.941 \pm 0.009$
CMU	$0.000 \pm 0.577$	$0.000 \pm 0.577$	$0.000 \pm 0.577$
CMPni	$0.851 \pm 0.027$	$0.911 \pm 0.003$	$0.934 \pm 0.030$
CMX	$0.900 \pm 0.009$	$0.908 \pm 0.001$	$0.991 \pm 0.010$
CMXMsKs	$0.763 \pm 0.029$	$0.896 \pm 0.003$	$0.853 \pm 0.033$
BMU	$0.720 \pm 0.015$	$0.719 \pm 0.003$	$1.000 \pm 0.021$
CMIOCES	$0.318 \pm 0.005$	$0.291 \pm 0.001$	$1.071 \pm 0.018$
CMIOPEs	$0.547 \pm 0.010$	$0.591 \pm 0.002$	$0.925 \pm 0.018$
CrkTrkMu	$0.671 \pm 0.008$	$0.719 \pm 0.001$	$0.936 \pm 0.012$
CrkTrk ( $\mu$ )	$0.670 \pm 0.008$	$0.720 \pm 0.001$	$0.935 \pm 0.012$

Table A.25: Lepton identification efficiencies for data and simulation for periods 8 to 10. Errors are statistical only.

**Run Period: 11 - 12**

	data	MC	Scale Fac
LBE	$0.122 \pm 0.003$	$0.101 \pm 0.001$	$1.208 \pm 0.030$
TCE	$0.761 \pm 0.004$	$0.790 \pm 0.001$	$0.964 \pm 0.006$
PHXTrk	$0.879 \pm 0.004$	$0.880 \pm 0.001$	$1.000 \pm 0.004$
PHXPEM	$0.816 \pm 0.004$	$0.859 \pm 0.001$	$0.950 \pm 0.005$
PLBEPem	$0.110 \pm 0.003$	$0.090 \pm 0.001$	$1.217 \pm 0.039$
CrkTrk ( $e$ )	$0.636 \pm 0.007$	$0.753 \pm 0.001$	$0.844 \pm 0.009$
CMUP	$0.795 \pm 0.009$	$0.883 \pm 0.001$	$0.900 \pm 0.011$
CMU	$0.000 \pm 0.707$	$0.000 \pm 0.707$	$0.000 \pm 0.707$
CMP	$0.793 \pm 0.019$	$0.907 \pm 0.002$	$0.874 \pm 0.021$
CMX	$0.856 \pm 0.012$	$0.900 \pm 0.002$	$0.951 \pm 0.014$
CMXMsKs	$0.822 \pm 0.027$	$0.896 \pm 0.003$	$0.917 \pm 0.030$
BMU	$0.712 \pm 0.020$	$0.719 \pm 0.004$	$0.990 \pm 0.028$
CMIOCES	$0.284 \pm 0.006$	$0.248 \pm 0.001$	$1.145 \pm 0.025$
CMIOPEs	$0.512 \pm 0.013$	$0.579 \pm 0.002$	$0.884 \pm 0.023$
CrkTrk ( $\mu$ )	$0.545 \pm 0.011$	$0.572 \pm 0.002$	$0.953 \pm 0.019$

Table A.26: Lepton identification efficiencies for data and simulation for periods 11 to 12. Errors are statistical only.

**Run Period: 13**

	data	MC	Scale Fac
LBE	$0.115 \pm 0.003$	$0.097 \pm 0.001$	$1.182 \pm 0.034$
TCE	$0.775 \pm 0.005$	$0.798 \pm 0.001$	$0.971 \pm 0.006$
PHXTrk	$0.859 \pm 0.005$	$0.857 \pm 0.001$	$1.002 \pm 0.006$
PHXPEM	$0.814 \pm 0.005$	$0.866 \pm 0.001$	$0.940 \pm 0.006$
PLBEPem	$0.113 \pm 0.004$	$0.082 \pm 0.001$	$1.376 \pm 0.052$
CrkTrk ( $e$ )	$0.649 \pm 0.008$	$0.761 \pm 0.002$	$0.852 \pm 0.010$
CMUP	$0.820 \pm 0.010$	$0.886 \pm 0.001$	$0.925 \pm 0.012$
CMU	$0.000 \pm 1.000$	$0.000 \pm 1.000$	$0.000 \pm 1.000$
CMP	$0.810 \pm 0.020$	$0.904 \pm 0.003$	$0.896 \pm 0.022$
CMX	$0.899 \pm 0.014$	$0.903 \pm 0.002$	$0.996 \pm 0.015$
CMXMsKs	$0.797 \pm 0.029$	$0.906 \pm 0.003$	$0.880 \pm 0.033$
BMU	$0.716 \pm 0.025$	$0.692 \pm 0.004$	$1.036 \pm 0.037$
CMIOCES	$0.287 \pm 0.007$	$0.250 \pm 0.001$	$1.149 \pm 0.028$
CMIOPEs	$0.476 \pm 0.018$	$0.516 \pm 0.003$	$0.922 \pm 0.035$
CrkTrk ( $\mu$ )	$0.542 \pm 0.012$	$0.578 \pm 0.002$	$0.938 \pm 0.021$

Table A.27: Lepton identification efficiencies for data and simulation for period 13. Errors are statistical only.

**Run Period: 14 - 21**

	data	MC	Scale Fac
LBE	$0.120 \pm 0.001$	$0.104 \pm 0.001$	$1.161 \pm 0.016$
TCE	$0.762 \pm 0.002$	$0.784 \pm 0.001$	$0.972 \pm 0.003$
PHXTrk	$0.892 \pm 0.002$	$0.869 \pm 0.001$	$1.027 \pm 0.002$
PHXPEM	$0.790 \pm 0.002$	$0.850 \pm 0.001$	$0.930 \pm 0.003$
PLBPEM	$0.124 \pm 0.002$	$0.095 \pm 0.001$	$1.310 \pm 0.022$
CrkTrk ( $e$ )	$0.633 \pm 0.003$	$0.738 \pm 0.001$	$0.858 \pm 0.005$
CMUP	$0.774 \pm 0.005$	$0.876 \pm 0.001$	$0.884 \pm 0.006$
CMU	$0.000 \pm 1.000$	$0.000 \pm 1.000$	$0.000 \pm 1.000$
CMP	$0.793 \pm 0.010$	$0.900 \pm 0.002$	$0.881 \pm 0.012$
CMX	$0.858 \pm 0.006$	$0.893 \pm 0.002$	$0.961 \pm 0.007$
CMXMsKs	$0.801 \pm 0.013$	$0.894 \pm 0.003$	$0.897 \pm 0.015$
BMU	$0.729 \pm 0.011$	$0.696 \pm 0.004$	$1.048 \pm 0.016$
CMIOCES	$0.287 \pm 0.003$	$0.248 \pm 0.001$	$1.156 \pm 0.014$
CMIOPEs	$0.542 \pm 0.007$	$0.571 \pm 0.002$	$0.949 \pm 0.013$
CrkTrk ( $\mu$ )	$0.553 \pm 0.006$	$0.567 \pm 0.002$	$0.975 \pm 0.010$

Table A.28: Lepton identification efficiencies for data and simulation for periods 14 to 21. Errors are statistical only.

**Run Period: 22 - 38**

	data	MC	Scale Fac
LBE	$0.122 \pm 0.001$	$0.102 \pm 0.001$	$1.199 \pm 0.012$
TCE	$0.752 \pm 0.001$	$0.787 \pm 0.001$	$0.956 \pm 0.002$
PHXTrk	$0.886 \pm 0.001$	$0.864 \pm 0.001$	$1.026 \pm 0.002$
PHXPEM	$0.769 \pm 0.002$	$0.851 \pm 0.001$	$0.904 \pm 0.002$
PLBPEM	$0.133 \pm 0.001$	$0.095 \pm 0.001$	$1.408 \pm 0.017$
CrkTrk ( $e$ )	$0.603 \pm 0.002$	$0.741 \pm 0.001$	$0.814 \pm 0.003$
CMUP	$0.746 \pm 0.003$	$0.877 \pm 0.001$	$0.850 \pm 0.004$
CMU	$0.767 \pm 0.006$	$0.878 \pm 0.003$	$0.874 \pm 0.008$
CMP	$0.752 \pm 0.007$	$0.902 \pm 0.002$	$0.834 \pm 0.008$
CMX	$0.853 \pm 0.004$	$0.898 \pm 0.002$	$0.950 \pm 0.005$
CMXMsKs	$0.772 \pm 0.009$	$0.890 \pm 0.003$	$0.867 \pm 0.010$
BMU	$0.727 \pm 0.007$	$0.695 \pm 0.004$	$1.047 \pm 0.012$
CMIOCES	$0.208 \pm 0.002$	$0.145 \pm 0.001$	$1.434 \pm 0.015$
CMIOPEs	$0.538 \pm 0.005$	$0.575 \pm 0.002$	$0.936 \pm 0.009$
CrkTrk ( $\mu$ )	$0.543 \pm 0.004$	$0.564 \pm 0.002$	$0.962 \pm 0.007$

Table A.29: Lepton identification efficiencies for data and simulation for periods 22 to 38. Errors are statistical only.





## Single Object Triggers

In this Appendix we describe in details the trigger paths exploited in the analyses presented in this Thesis and outlined in Section 4.7. To collect the data we exploit several high- $p_T(E_T)$  lepton triggers and to evaluate the jet-to-lepton misidentification we consider single jet triggers, as mentioned in Section 4.7.

The requirements for the lepton triggers considered, for each trigger level, are reported below.

- ELECTRON\_CENTRAL\_18

- L1 - A central electron cluster with  $E_T > 8$  GeV,  $E_{HAD}/E_{EM} < 0.125$ , and an associated  $p_T > 8.34$  GeV/c XFT track.
- L2 - A central electron cluster with  $E_T > 16$  GeV,  $E_{HAD}/E_{EM} < 0.125$ , and an associated  $p_T > 8$  GeV/c XFT track.
- L3 - A central electron cluster with  $E_T > 18$  GeV,  $E_{HAD}/E_{EM} < 0.125$ ,  $L_{SHR} < 0.4$ , and an associated  $p_T > 9$  GeV/c L3 track that extrapolates to the CES within 8 cm in  $z$  of the cluster position. The  $E_T$  calculation uses the track angle.

- MUON\_CMUP18

- L1 - An XFT track with  $p_T > 4$  GeV/c associated with both a CMU and a CMP stub.
- L2 - An XFT track with  $p_T > 14.77$  GeV/c associated with both a CMU and a CMP stub.
- L3 - A minimum ionising track with  $p_T > 18$  GeV/c associated with both a CMU and CMP stub, with  $|\Delta X_{CMU}| < 20$  cm,  $|\Delta X_{CMP}| < 10$  cm.

- MUON\_CMX18

- L1 - An XFT track with  $p_T > 8.34$  GeV/c associated with a CMX stub and CSX scintillator information.
- L2 - An XFT track with  $p_T > 14.77$  GeV/c associated with a CMX stub.
- L3 - A minimum ionising track with  $p_T > 18$  GeV/c associated with a CMX stub with  $|\Delta X_{CMX}| < 10$  cm.

- MUON\_CMP18\_PHIGAP

- L1 - An XFT track with  $p_T > 4$  GeV/c associated with a CMP stub but NOT a CMU stub.
-

- L2 - An XFT track with  $p_T > 14.77$  GeV/c associated with a CMP stub but NOT a CMU stub.
  - L3 - A minimum ionising track with  $p_T > 18$  GeV/c associated with a CMP stub but NOT a CMU stub, with  $|\Delta X_{CMP}| < 10$  cm.
- MUON\_CMU18\_ETAGAP
    - L1 - An XFT track with  $p_T > 4$  GeV/c associated with a CMU stub but NOT a CMP stub.
    - L2 - An XFT track with  $p_T > 14.77$  GeV/c associated with a CMU stub but NOT a CMP stub.
    - L3 - A minimum ionising track with  $p_T > 18$  GeV/c associated with a CMU stub but NOT a CMP stub, with  $|\Delta X_{CMU}| < 10$  cm.

From Period 9 (Sept. 2006) to the end of data taking (Sept. 2011), alternate paths with various combinations of prescales and luminosity enabling were introduced to control the muon trigger rates at high instantaneous luminosity. In this analysis we use the trigger paths that provide the largest integrated luminosity for a given run range considered.

In order to study data driven background prediction we exploit data collected by jet-based triggers: JET20, JET50, JET70, JET100. All these trigger paths are based on the reconstruction of at least one jet (at trigger level) with a given energy threshold: 20, 50, 70, and 100 GeV. For each trigger paths is required:

- L1 - One calorimetric tower with energy above a given threshold
- L2 - At least a clusters within  $|\eta| < 3.6$  and with a lower cut on its  $E_T$
- L3 - At least one jet reconstructed with the JETCLU algorithm with a  $\Delta R = 0.7$  cone above a given threshold.

Given the rate of jet production at the Tevatron, all these trigger paths are prescaled but the JET100. The different threshold applied for the jet trigger paths are summarized in Table B.1.

Trigger Name	Level 1		Level 2		Level 3
	$E_T > [\text{GeV}]$	Prescale	$E_T > [\text{GeV}]$	Prescale	$E_T > [\text{GeV}]$
JET20	5	50	20	100	20
JET50	5	50	40	2	50
JET70	10	8	60	1	70
JET100	20	1	90	1	100

Table B.1: Thresholds for JET20, JET50, JET70, JET100 trigger requirements.

## B.1 Trigger efficiencies

To evaluate trigger efficiencies we use an approach based on the reconstructed objects in the event. Since we are using single-lepton triggers, we separately evaluate the probability for each trigger to be fired by the corresponding offline-selected lepton. The lepton categories connected

with a trigger used to collect data are called *triggerable* categories and are CMUP, CMP, CMU, CMX, CMX-MsKs, TCE, and LBE. We evaluate the trigger efficiency for a given event as the probability that at least one of the reconstructed objects in the event has fired its associated trigger. For the efficiency measurement we require always leptons to have  $p_T (E_T) > 20$  GeV/c in order to be exploited in the efficiency calculation, since trigger requirements usually apply a  $p_T (E_T) > 18$  GeV/c online cut to the XFT tracks. In this section we describe in details how muon and electron trigger efficiencies are calculated.

To calculate trigger efficiencies for muons we start selecting  $Z \rightarrow \mu\mu$  candidate events in the  $76 < M_{\mu\mu} < 106$  GeV/c<sup>2</sup> dimuon invariant mass range. Among these events we count those with one of the two muons that fired a reference trigger path (*tag* muon) and the other is fiducial to the detector we're measuring the trigger efficiency (*probe* muon). We then check whether the other muon has fired the trigger or not. We separate calculation for triggers that require *track-stub* match in the  $r-\phi$  plane (*2D*) and those that require also a stereo matching, XFT *3D* matching requirement (available after trigger upgrade, described in [93]). Trigger efficiencies for 2D triggers are trivial and are calculated with the following formula (for CMUP 2D trigger paths, similar for others):

$$\epsilon_{CMUP-2D} = \frac{\#(CMUP-2D \ \& \ CMX-2D)}{\#CMX-2D} \quad (B.1)$$

where  $\#CMUP-2D$  ( $\#CMX-2D$ ) is the number of events where the corresponding trigger path has fired. The requests for a 3D trigger are the same as the corresponding 2D trigger with an additional requirement on the stereo track match; to calculate the efficiencies of 3D triggers we can use two different methods. The first is calculating the efficiency relative to the corresponding 2D trigger path, evaluating in this way just the efficiency of the 3D part of the requirements. The second method is analogous to the 2D method and evaluates directly the entire 3D efficiency.

$$\epsilon_{CMUP-3D}^*(I) = \frac{\#(CMX-2D \ \& \ CMUP-2D \ \& \ CMUP-3D)}{\#(CMX-2D \ \& \ CMUP-2D)} \quad (B.2)$$

$$\epsilon_{CMUP-3D}(I) = \epsilon_{CMUP-2D} \cdot \epsilon_{CMUP-3D}^*(I) \quad (B.3)$$

$$\epsilon_{CMUP-3D}(II) = \frac{\#(CMUP-3D \ \& \ CMX-3D)}{\#CMX-3D} \quad (B.4)$$

Similar formula are used to calculate CMX-2D (3D) efficiency. The two methods agree very well and we use the second one when available, since it gives a slightly more accurate estimation. To calculate the trigger efficiency we take into account the fraction of the active *lives* of the triggers, obtaining a corrective scale factor for each efficiency. Results are summarized as a function of run period in Table B.2; certain periods are grouped given the homogeneous data taking conditions and trigger behavior in that period. In this Table we notice a drop in CMUP-3D efficiency in period 18, which is due to a technical known problem.

Livetime corrections listed in Table B.2 are applied to account for different prescales of the trigger paths. Trigger efficiencies and livetimes are then applied to MC simulations on a per-event basis, in order to model online triggering effects on the simulated samples, as described in Section 5.1.2.

The efficiency for the electron trigger CENTRAL-ELECTRON\_18 is separately calculated for the tracking trigger and for the calorimeter trigger. The tracking trigger efficiency is calculated using a backup trigger that have the same (or tighter) calorimeter requirements that CENTRAL-ELECTRON\_18 but no tracking requirement: W\_NOTRACK. With this trigger we select  $W \rightarrow e\nu$  events and check that the electron is pointing to the central electron trigger. We can

Run Period	P8	P9	P10	P11	P12	P13	P14	P15	P16	P17
<b>Trigger <math>\epsilon</math></b>										
CMUP 2-D	0.906	0.910	0.920	0.911	0.915	0.918	0.908	0.921	0.906	0.921
CMUP 3-D (I)				0.858	0.839	0.835	0.881	0.882	0.860	0.875
CMUP 3-D (II)				0.863	0.838	0.838	0.888	0.885	0.852	0.878
CMX ARCH 2-D	0.941	0.948	0.978	0.970	0.936	0.935	0.915	0.950	0.947	0.940
CMX ARCH 3-D (I)		(0.927)	0.927	0.930	0.889	0.894	0.892	0.928	0.929	0.905
CMX ARCH 3-D (II)				0.931	0.894	0.880	0.895	0.928	0.926	0.912
CMX MS/KS 2-D		(0.849)	0.849	0.814	0.773	0.750	0.847	0.861	0.889	0.816
CMX MS/KS 3-D (I)			0.835	0.784	0.745	0.737	0.847	0.834	0.879	0.800
CMX MS/KS 3-D (II)				0.780	0.758	0.766	0.849	0.807	0.864	0.791
CMP PHI-GAP			(0.948)	0.948	0.891	0.924	0.889	0.764	0.774	0.845
<b>LiveTimes</b>										
CMUP 2-D	1.000	1.000	1.000	0.959	0.933	0.936	0.945	0.939	0.924	0.923
CMUP 3-D	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
CMX 2-D	1.000	0.963	0.958	0.913	0.885	0.899	0.941	0.900	0.845	0.845
CMX 3-D	1.000	0.965	0.988	0.971	0.952	0.950	0.951	0.961	0.989	1.000
CMP PG	0.000	0.000	0.558	0.669	0.840	0.885	0.968	0.977	0.954	0.946

Run Period	P14-17	P18	P19	P14-19	P20	P21	P22	P23	P24-38
<b>Trigger <math>\epsilon</math></b>									
CMUP 2-D	0.917	0.924	0.907	0.917	0.916	0.902	0.905	0.912	0.913
CMUP 3-D (I)	0.875	0.736	0.861	0.826	0.867	0.851	0.857	0.869	0.842
CMUP 3-D (II)	0.876	0.744	0.862	0.828	0.871	0.853	0.852	0.868	0.843
CMX ARCH 2-D	0.942	0.956	0.927	0.944	0.921	0.954	0.967	0.956	0.948
CMX ARCH 3-D (I)	0.916	0.922	0.901	0.916	0.898	0.919	0.937	0.899	0.916
CMX ARCH 3-D (II)	0.916	0.900	0.876	0.901	0.883	0.896	0.923	0.885	0.899
CMX MS/KS 2-D	0.849	0.813	0.785	0.828	0.802	0.826	0.774	0.863	0.822
CMX MS/KS 3-D (I)	0.831	0.792	0.768	0.810	0.783	0.813	0.752	0.811	0.801
CMX MS/KS 3-D (II)	0.814	0.777	0.774	0.794	0.753	0.787	0.775	0.808	0.787
CMP PHI-GAP	0.803	0.758	0.693	0.769	0.794	0.802	0.797	0.838	0.787
<b>LiveTimes</b>									
CMUP 2-D	0.930	0.851	0.803	0.876	0.830	0.776	0.800	0.813	0.837
CMUP 3-D	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
CMX 2-D	0.871	0.768	0.719	0.802	0.750	0.667	0.701	0.711	0.750
CMX 3-D	0.981	1.000	1.000	0.992	1.000	1.000	1.000	1.000	0.996
CMP PG	0.959	0.929	0.819	0.920	0.876	0.813	0.800	0.811	0.871

Table B.2: Trigger efficiencies and livetimes of each trigger for each run periods.

then evaluate the electron tracking trigger efficiency for the different trigger level (see Section 3.2.5) with the following formulas:

$$\epsilon(L1_{trk}) = \frac{N_W \& \text{passed } L1}{N_W} \quad (\text{B.5})$$

$$\epsilon(L2_{trk}) = \frac{N_W \& \text{passed } L1 \& \text{passed } L2}{N_W \& \text{passed } L1} \quad (\text{B.6})$$

$$\epsilon(L3_{trk}) = \frac{N_W \& \text{passed } L1 \& \text{passed } L2 \& \text{passed } L3}{N_W \& \text{passed } L1 \& \text{passed } L2} \quad (\text{B.7})$$

These efficiencies are calculated as a function of several kinematic variables:  $p_T$ ,  $\phi$ ,  $\eta$ ,  $z_0$ ,  $Callso(\Delta R = 0.4)$ ,  $TrkIso(\Delta R = 0.1)$ , number of jets. The only relevant dependence found are for  $\eta$  and  $z_0$ . Near  $\eta \sim 0$  there is a significant inefficiency as a result of the COT space bars and also due to the charge collection inefficiency coming from the shorter particle path length. The efficiency has been fitted with the function

$$\epsilon = A - \frac{C}{2\pi\sigma} e^{-\frac{\eta^2}{2\sigma^2}} \quad (\text{B.8})$$

to take into account for this effect in MC simulation. The  $z_0$  efficiency increases for  $z_0 \sim 0$ ; this dependence for the tracking trigger efficiency comes from quite the same kinematic reasons of  $\eta$  inefficiency.

The L1 calorimeter efficiency is calculated from events with one electron that passed the L3 tracking trigger electron and is found to be 100 %.

To obtain an unbiased measurement of L2 calorimeter trigger efficiency events that passed L1 selections with no other request are selected. Since the rate is pretty high, those events are selected with a prescaled trigger, `L2_PS100_L1_CEM8_PT8`, that randomly extract one event every hundred of events that passed L1. From those events the efficiency is then calculated counting the number of events with at least one identified electron (TCE) satisfying the L1 cuts except for the requirement of  $E_T > 18$  GeV (instead of 20 GeV) to study the *turn-on* profile of the trigger as a function of the  $E_T$  and the isolation ratio<sup>1</sup> less than 0.1 to reduce background. The corresponding formula is

$$\epsilon(L2_{cal}) = \frac{N_{el} \& \text{passed } L1_{cal} \& \text{passed } L2_{PS} \& \text{passed } L2_{cal}}{N_{el} \& \text{passed } L1_{cal} \& \text{passed } L2_{PS}} \quad (\text{B.9})$$

The L2 calorimeter trigger efficiency depends on the  $E_T$  of the lepton as a result of the tower clustering algorithm, tower energy calculation and corrections. The turn-on curve reaches 100% at about 30 GeV. This curve is fitted to obtain the efficiency as a function of  $E_T$  with the following relation:

$$\epsilon = A - B e^{-CE_t} \quad (\text{B.10})$$

To measure L3 calorimeter trigger efficiency is used a calibration trigger path, that is similar to the `ELECTRON_CENTRAL_18` trigger but with a lower L1 cut on the jet transverse energy,  $E_T > 8$  GeV and the same prescaled trigger used for L2 calorimeter trigger efficiency `ELECTRON_CENTRAL_8_NO_L2`. The efficiency is then calculated counting the number of identified

<sup>1</sup> Isolation ratio is defined as the ratio of calorimeter isolation (*Callso*) in *EM* and *Had* calorimeter.

electrons that pass this trigger path and the L2 calorimeter trigger with the following formula

$$\epsilon(L3_{cal}) = \frac{N_{el} \& \text{ passed } EL\_CENT\_8\_NO\_L2 \& \text{ passed } L2_{cal} \& \text{ passed } L3_{cal}}{N_{el} \& \text{ passed } EL\_CENT\_8\_NO\_L2 \& \text{ passed } L2_{cal}} \quad (\text{B.11})$$

Since the offline cut on transverse energy is  $E_T > 20$  GeV and L3 cut is  $E_T > 18$  GeV, this efficiency is expected to be close to 100 % with possible small inefficiency due to energy determination (since L3 clustering algorithm is almost identical to the offline algorithm). The L3 calorimeter trigger efficiency is in fact found to be 100 % in the full  $E_T$  range.

The total calorimeter efficiency is then evaluated convoluting the electron  $E_T$  distribution of the sample of interest with the  $E_T$  dependence function. When we use this trigger efficiency to weight MC events a trigger scale factor is calculated event-per-event according to the triggered electron  $E_T$ .

## Lepton momentum corrections

The Monte Carlo simulation we use for this analysis<sup>1</sup> poorly models the electron and muon transverse momentum, which affects at first the dilepton invariant mass distribution  $M_{\ell\ell}$ . This is clearly noticeable when comparing simulation to data for an inclusive sample of Drell-Yan events, as can be seen in Figures C.3. The detector simulation used in the MC production overestimate the track fit resolution, which got worse during the years with the aging of the tracking drift chamber (COT). This impacts particularly the resolution of the track curvature measurement, which in the simulation is more precise than it is on average in the collected data. To better simulate the real data, curvature resolution in the MC simulation has to be deteriorated, introducing an additional uncertainty on this fit parameter.

Secondly, the generated  $Z \rightarrow e^+e^-$  sample presents a small intrinsic shift in the  $Z$  mass peak mean value with respect to the nominal  $Z$  mass [3].

We compensate these effects introducing a correction on the track momentum parameters, which is parametrized by modifying it so as:

$$p_T \rightarrow p_T \cdot C, \quad \frac{1}{p_T} \rightarrow \frac{1}{p_T} + f \quad (\text{C.1})$$

where  $C$  is a *fixed-value* parameter that sets a correction to the momentum scale, and  $f$  is a random value parameter chosen from a Gaussian spectrum<sup>2</sup> with mean 0 and width  $\sigma$ , that sets an additional smearing to the track curvature resolution.

The values of the  $p_T$  scale correction  $C$  and curvature smearing  $\sigma$  are tuned separately for electrons and muons in a  $Z \rightarrow \ell\ell$  dominated sample of events containing exactly 2 reconstructed leptons and a small amount of missing transverse energy ( $\cancel{E}_T \leq 20$  GeV). The tuning values are chosen by comparing the  $M_{\ell\ell}$  distribution for data and simulation in a narrow invariant mass region around the  $Z$  peak, i.e. in the range from 86 and 96 GeV/ $c^2$ . This is technically done by producing several dilepton mass template of the simulation, corrected for different values of  $C$  and  $\sigma$ , and comparing them to the data distribution. The  $\chi^2$  between Data and the resulting MC distribution

<sup>1</sup> Which is the same used in most of the CDF published analyses.

<sup>2</sup> In fact this correction holds under the assumption that a Gaussian smearing can reproduce the curvature resolution degradation, which is an approximation suitable for the purpose of the analysis described.

is evaluated separately as a function of  $\sigma$  and  $C$  according to the following equation:

$$\chi^2 = \sum_i^{Mass\ bins} \left( \frac{N_{Data}^i - N_{MC}^i}{\sqrt{N_{Data}^i}} \right)^2. \quad (C.2)$$

In the calculation of the  $\chi^2$  the MC spectra are normalized to the area of the data spectrum. The *best parameters* and their errors are extracted from a second-degree polynomial fit to the  $\chi^2$  distribution, considering the minimum and half of the width of the curve at minimum +1 as the parameter best value and parameter error, respectively. The  $\eta$  dependence of the track reconstruction resolution correction is taken into account by considering separately central and forward leptons, according to their fiduciality to the central or plug calorimeters<sup>3</sup>. The parameter tuning is done first for events with two central leptons; then, these parameter are fixed for central leptons and tuned for plug leptons in events containing one central and one plug lepton. The results of the parameter tuning are shown in Figures C.1 and C.2 and summarized in Table C.1.

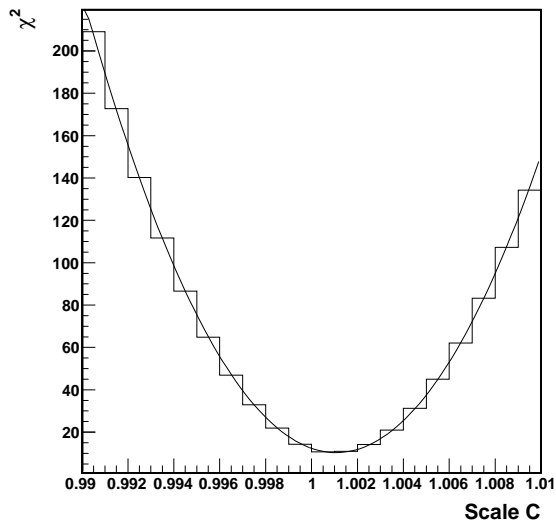
Parameters	Scale (C)	Resolution width ( $\sigma$ ) [GeV <sup>-1</sup> ]
<b>Muons</b>		
Central	0.999±0.001	(4.3±0.3)×10 <sup>-4</sup>
Forward	0.998±0.004	(20.1±2.0)×10 <sup>-4</sup>
<b>Electrons</b>		
Central	1.001±0.001	(0.0±0.0)×10 <sup>-4</sup>
Forward	0.995±0.001	(6.6±1.0)×10 <sup>-4</sup>

Table C.1: Scale and resolution correction parameters obtained for central/forward electrons and muons.

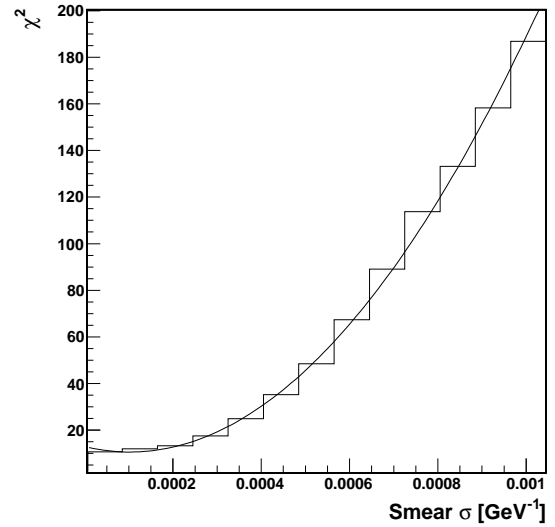
These parameters are used to correct the lepton transverse momentum according to Equation C.1 in all the MC simulated samples, observing the major effect in the Drell-Yan simulated sample. These corrections affect not only the dilepton invariant mass distribution but also other kinematic properties of the reconstructed  $Z \rightarrow \ell\ell$  decay (such as  $p_T(\ell\ell)$ ,  $\Delta\phi(\ell\ell)$ ) and the  $\cancel{E}_T$ , which is corrected for the contribution of each muon momentum (see Figure C.5). The corrections applied improve the agreement between data and simulation but are not expected to completely fix the mismodeling. In the analyses described in this Thesis a residual discrepancy between data and simulation is noticed and taken into account among the systematic uncertainties.

<sup>3</sup> The several lepton categories defined in Appendix A.1 are exclusively fiducial to the central or forward region of the detector, hence the central/forward separation is done assigning each lepton category to one of the two topology.

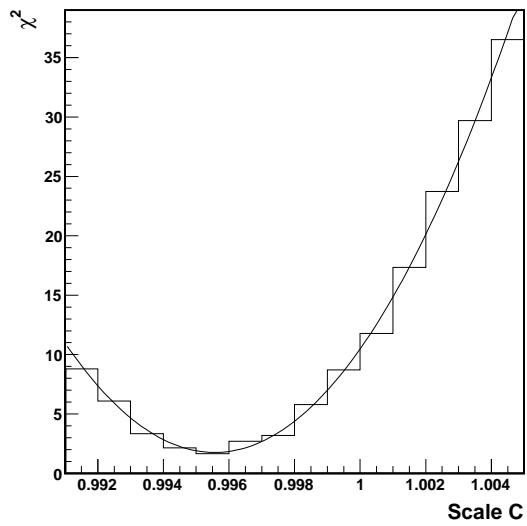




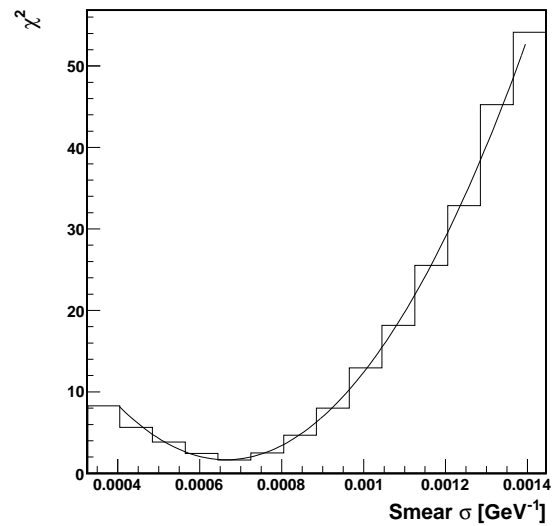
(a) Central e scale



(b) Central e smearing



(c) Forward e scale



(d) Forward e smearing

Figure C.1: Scale and smearing parameter obtained from the tuning procedure for electrons.

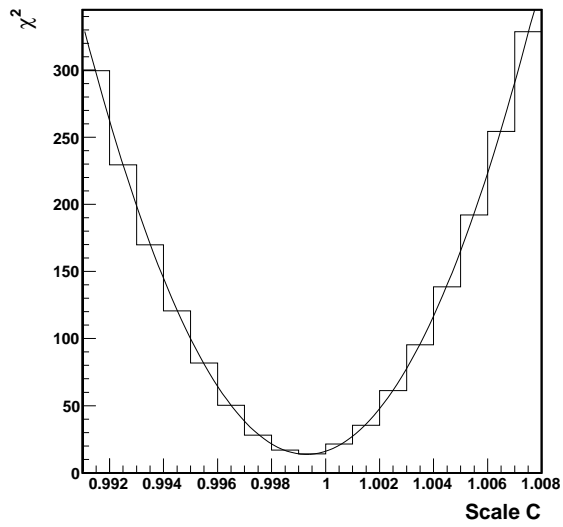
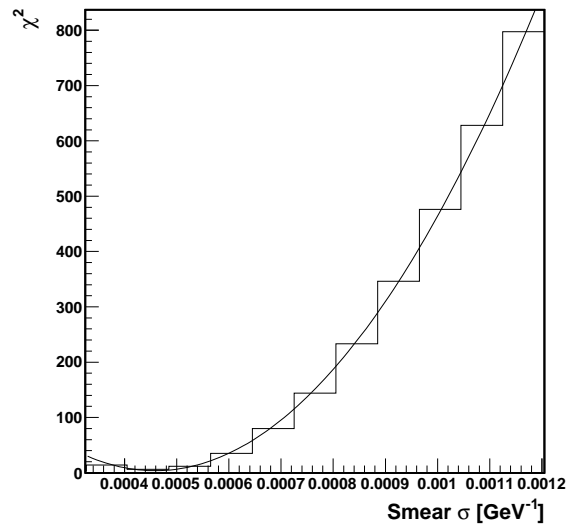
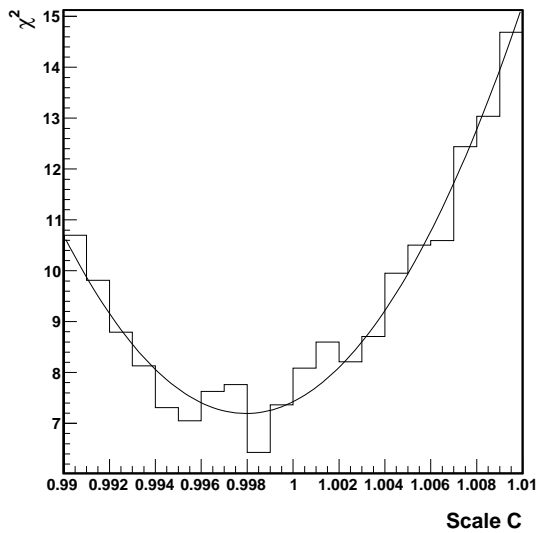
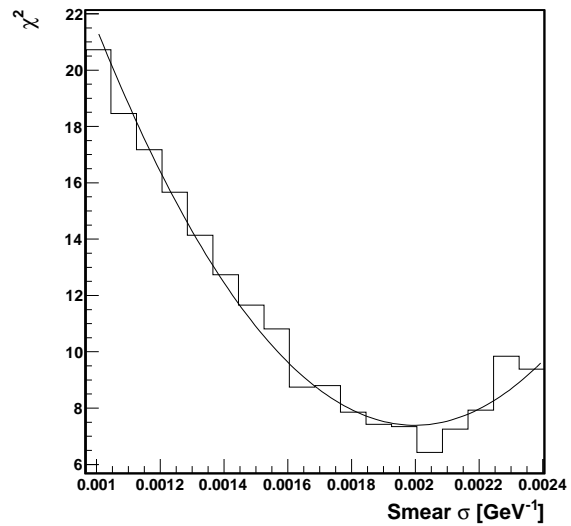
(a) Central  $\mu$  scale(b) Central  $\mu$  smearing(c) Forward  $\mu$  scale(d) Forward  $\mu$  smearing

Figure C.2: Scale and smearing parameter obtained from the tuning procedure for muons

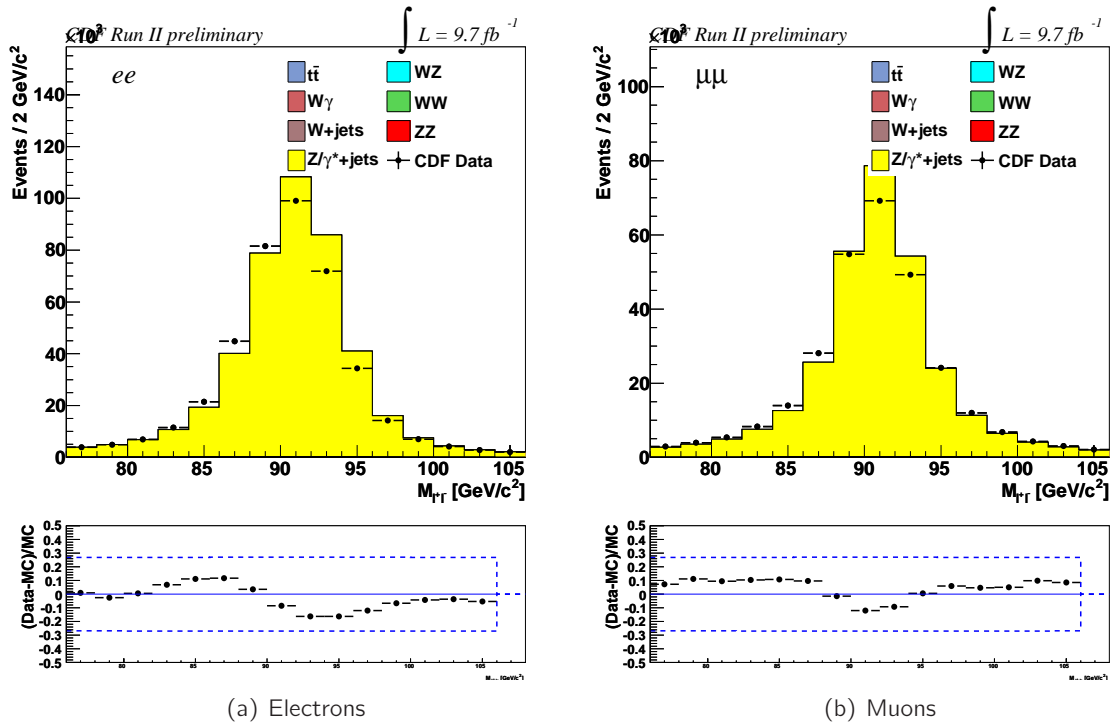


Figure C.3:  $M_{\ell\ell}$  distribution for (a)  $e^+e^-$  and (b)  $\mu^+\mu^-$  events in the Drell-Yan sample used for the tuning of the track resolution parameters.

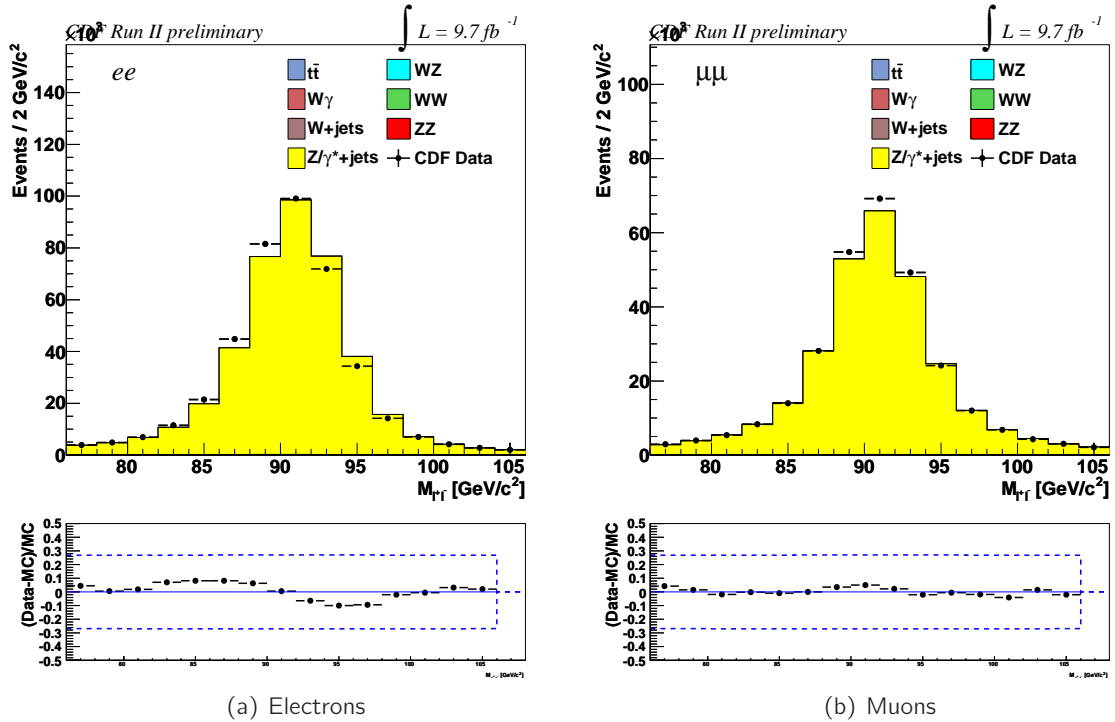


Figure C.4:  $M_{\ell\ell}$  distribution for (a)  $e^+e^-$  and (b)  $\mu^+\mu^-$  events in the Drell-Yan sample after correcting the leptons momentum.

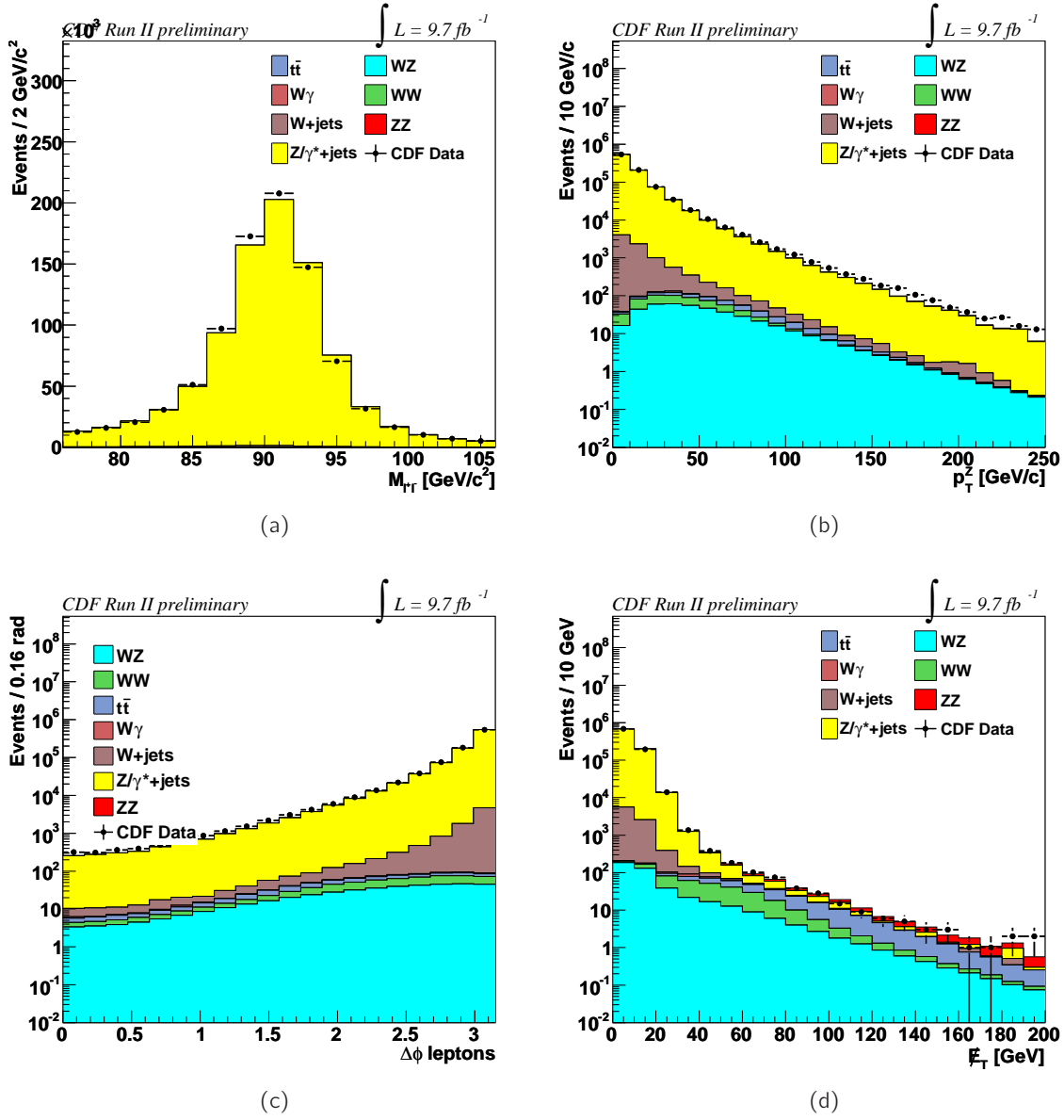


Figure C.5: (a)  $M_{\ell\ell}$ , (b)  $p_T(\ell\ell)$ , (c)  $\Delta\phi(\ell\ell)$ , and (d)  $E_T^{\ell}$  distribution after correcting the scale and resolution of the lepton momenta.

# Acknowledgments

This Thesis has been the result of lots of effort during the past three years and I would like to thank everyone that made this possible, either in a way or another.

I would like to thank Donatella, my supervisor, who guided me till this point: she definitely taught me the way people do research and why they usually dedicate so much of their life to that.

Writing this Thesis has been really painful, because I'm not too good at that, hence I would like to thank really much also who read these pages many times to spot all the mistakes, trying to improve the final shape.

Thanks a lot to all the people of the CDF collaboration (in its glorious times) at Fermilab, in particular to Eric, Sergo, Simone, Larry, Mark, Doug, Pasha, who helped me and gave me a lot of feedbacks on the analyses during the past years.

A huge thank goes also to Silvia, my office mate in the recent good and hard times: I owe you more than a couple of beers for the support you provided me when I was almost going crazy.

I would like to thank Tony, for the precious experience he shared with me; I really loved collaborating with you and hope to have another of such a chance in the future.

Many many thanks to all the guys with whom I spent great times in Chicago and suburbs: Matteo, Sabatino, Mirco, Patrizia, Angelo, Viviana, Francesco, Paola, Caterina, Stefano, Marino, Francesco, Elisabetta, Marco, Fabrizio, Federico, Lucia, Jorge, Camacho, Jesus, and the ones that I surely forgot.

I will always be grateful to the friends in Padova University, with whom I shared the beginning of my studies in physics and that I carry with me even if they are physically far from me most of the time: Alessandro, Ennio, Lucio, Luca, Claudio, Luca. Be patient if I don't hear from me so often, as usual I promise I will work on this.

Thanks to the people that has nothing to do with physics, but lot to do with my life: Ilaria, Mirko, Alessandra, Luca, Paola, Elisa, Tomaso, Riccardo, Francesco, Silvia; you're a kind of open window to the everyday life, even if you probably don't know.

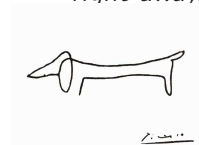
Thanks to Giulia, who had been caught into this *adventure* almost by chance but revealed herself to be more than what I deserved. She fed me during late hours in the office, has been a great mate of sleepless nights and restless days (together with all the party crew). The time I spent with you has been amazing.

Thanks to Roberta and Carlotta, who indirectly took care of my mum when I was not there, and when, even if I was there, I was not kind enough with her.

Really special thanks goes to my Mom and Dad, who did a lot for me, and, above all, always allowed me to do what I really like in my life. They have been present during the past 27 years almost everyday, each one in a different way, always proving their love for me. I know they've been probably the people I neglected the most during bad times but I promise I'll try to recover.

Thanks to everyone else contributed to this work and to my life. That's all folks!

*When old  
we will remember the good things,  
it is worth to forget the bad ones  
right away.*





# List of Figures

1.1	The primary vertexes of QCD. . . . .	4
1.2	Higgs potential $V(\phi)$ for $\mu < 0$ . . . . .	7
2.1	The $e^+e^-$ annihilation cross-section to hadrons, from initial low energies in early colliders to the maximum energy at LEP. . . . .	13
2.2	Precision observables of the electroweak part of the Standard Model, measured at LEP (when not specified) and SLC (by the SLD experiment). . . . .	14
2.3	Tree level Feynman diagrams for $ZZ$ production at the LEP collider. . . . .	15
2.4	Combined cross-sections measured from data collected in 1997-2000. The errors shown are sums in quadrature of the statistical and systematic errors. . . . .	16
2.5	(a) $W$ pair production production cross section as a function of the center-of-mass energy $s$ at LEP. (b) Ratio of the measured cross section to the prediction. . . . .	17
2.6	Tree-level $t$ -channel(a) and $s$ -channel(b) Feynman diagram for $ZZ$ diboson production at the Tevatron. . . . .	17
2.7	Leading order Feynman diagrams of the dominant Higgs production mechanisms at colliders: (a) gluon fusion, (b) associated production, and (c) vector boson fusion at hadron colliders, and (d) associated production and (e) vector boson fusion at $e^+e^-$ colliders. . . . .	20
2.8	Production cross sections at the (a) $p\bar{p}$ collider Tevatron ( $\sqrt{s}= 1.96$ TeV) and at the (b) $pp$ collider LHC ( $\sqrt{s}= 7$ TeV), as function of the Higgs boson mass. . . . .	21
2.9	Higgs boson decay branching ratios [54]. . . . .	21
2.10	$\Delta\chi^2 = \chi^2 - \chi_{min}^2$ of the fit to electroweak precision measurements performed by the LEP-EWWG. . . . .	23
2.11	Mass exclusion limits (95% C.L.) obtained from the combination of all the Tevatron searches for a SM Higgs boson [59]. Exclusion limits obtained from the CMS [61] and Atlas [60] Collaborations are overlaid. . . . .	24
3.1	The accelerator complex of the Fermi National Accelerator Laboratory. . . . .	26
3.2	Run 2 Integrated luminosity as a function of time. . . . .	29
3.3	Run 2 Peak luminosity in the stores and its average, as a function of time. . . . .	29
3.4	Elevation view of the CDF II detector. . . . .	31
3.5	Isometric view of the CDF II detector. . . . .	32
3.6	Longitudinal view of the Tracking System of the CDF II detector. . . . .	32
3.7	(a) Sketch of the silicon detector in a x/y projection. (b) Cutaway transverse to the beam of the three inner tracker sub-systems. . . . .	33
3.8	Isometric (a) and end-view (b) of the CDF Silicon Vertex Detector. . . . .	33
3.9	Slots housing the wire holding fixtures of one quadrant of a COT endplate (a). Drift and field wires in three cells (b). The horizontal arrow shows the radial direction. . . . .	34

---

3.10	Equipotential lines in a COT cell. . . . .	34
3.11	Light-shifter plates connected to light guides and to photomultipliers of the front electromagnetic sector of a central calorimeter wedge. . . . .	35
3.12	Longitudinal view of Plug Calorimeters. . . . .	37
3.13	Coverage of muon detectors in the $(\eta, \phi)$ space. . . . .	38
3.14	Schematic view of the luminosity monitor inside a quadrant of CDF. . . . .	39
3.15	CDF trigger system. . . . .	41
3.16	The SVT architecture. . . . .	43
4.1	CDF track parameters and coordinate system. . . . .	49
4.2	$E_{Had}/E_{EM}$ (left) and isolation (right) distribution of central (top) and plug (bottom) calorimeter electron selection from unbiased, second legs of $Z^0 \rightarrow e^+e^-$ candidate events in Data [71]. . . . .	51
4.3	Distribution of different electron types in $\eta - \phi$ plane. . . . .	53
4.4	Distribution of different muon types in $\eta - \phi$ plane. . . . .	54
4.5	A parton originated from a hard scattering hadronizes and gives origin to a collimated spray of particles, a jet. . . . .	55
4.6	Calorimeter deposit in the $\eta - \phi$ plane as represented in the CDF event display. EM deposits are green (light-colored cluster on the right) while HAD deposits are red (darker-colored cluster on the left). . . . .	56
5.1	Sketch of the relative $ZZ$ fraction for the different decay modes. . . . .	62
5.2	Schematic representation of the <i>tag and probe</i> method. . . . .	67
5.3	Invariant mass distribution of central muon pairs, comparing data and MC simulation. . . . .	67
5.4	Fake rate probabilities as a function of the $E_T(p_T)$ of the denominator object for the several electron(muon) categories. . . . .	70
5.5	Fake rate probabilities as a function of the $p_T$ of the denominator object for the several muon categories. . . . .	71
5.6	Distribution of the sub-leading $Z$ vs. leading $Z$ $M_{\ell\ell}$ for the predicted signal, fake lepton background contribution. . . . .	74
5.7	Comparison of the MC simulation prediction and observation for some kinematic variables of the events in the collected sample. . . . .	75
5.8	Comparison of the MC simulation prediction and observation for some kinematic variables of the events in the collected sample. . . . .	76
5.9	(a) Posterior p.d.f. of the parameter $\alpha_{ZZ}$ after the integration over all the other parameters. The blue band represents the $\pm 1\sigma$ confidence interval. (b) Test statistic distribution for signal-like(red) and background-only(blue) pseudo-experiments. The black line represent the observed test statistic value for the data, while the red one is the median value of the signal-like pseudo-experiments. . . . .	80
5.10	Simulation-to-data comparison for events in the preselected sample for (a)leading lepton $p_T$ , (b)sub-leading lepton $p_T$ , (c)dilepton invariant mass ( $M_{\ell\ell}$ ), (d)reconstructed $Z$ transverse momentum ( $p_T(Z) \equiv p_T(\ell\ell)$ ). . . . .	84
5.11	Simulation-to-data comparison for events in the preselected sample for (a) $\Delta\phi(\ell\ell)$ , (b) $\Delta R(\ell\ell)$ , (c)missing transverse energy ( $\cancel{E}_T$ ), (d) the number of jets. . . . .	85
5.12	Distribution of $\Delta\phi(j, Z)$ for event in the preselected sample. A comparison of the data distribution with the prediction is shown in (a)linear and (b)logarithmic y-scale. . . . .	86
5.13	The predicted and observed number of jets in the event in the preselected sample(a) and after applying the veto on the presence of a recoiling jet(b). . . . .	86



5.14	Schematic draw of the $\cancel{E}_T$ projection along the $Z$ axis. . . . .	87
5.15	(a) $\cancel{E}_T^{Ax}$ distribution for event in the preselected sample with no recoiling jets, comparing data to simulation. The red lines over-impressed represent the $ZZ$ signal shape, magnified $\times 5$ with respect to the proper normalization. (b) A shape comparison of the same kinematic variable for the several concurrent processes, normalized to unit area. . . . .	88
5.16	Data-to-simulation comparison of some relevant kinematic variable distributions for the events passing the signal region requirements. . . . .	90
5.17	Data-to-simulation comparison of some relevant kinematic variable distributions for the events passing the signal region requirements. . . . .	91
5.18	Kinematic variables distributions in the Drell-Yan Control Region: simulated predictions compared to data. . . . .	94
5.19	Kinematic variables distributions in the Drell-Yan Control Region: simulated predictions compared to data. . . . .	95
5.20	Kinematic variables distributions in the $e\mu$ Control Region: simulated prediction compared to data. . . . .	96
5.21	Kinematic variables distributions in the $e\mu$ Control Region: simulated prediction compared to data. . . . .	97
5.22	The Neural Network structure, consisting of $N$ input nodes, $N + 2$ hidden nodes and a single output node. . . . .	99
5.23	Example of the Neural Network output distribution for signal and background sample. . . . .	99
5.24	Neural network output distribution for the different processes, normalized to unit area. . . . .	101
5.25	Neural network output distribution for signal (in red), background expectations, with the data (dots) superimposed in (a) linear and (b) logarithmic scale. . . . .	102
5.26	Neural Network output distribution for events in the Drell-Yan Control Region. . . . .	104
5.27	Drell-Yan background NN output distribution in the Signal Region, with the shape-varied distribution obtained from the data-to-simulation comparison in the control region. . . . .	105
5.28	Neural Network output distribution for events in the $e\mu$ Control Region. . . . .	105
5.29	(a) Posterior p.d.f. of the parameter $\alpha_{ZZ}$ after the integration over all the other parameters. The blue band represents the $\pm 1\sigma$ confidence interval. (b) Test statistic distribution for signal-like (red) and background-only (blue) pseudo-experiments. The black line represent the observed test statistic value for the data. . . . .	108
6.1	Predictions for Higgs production and decay to $ZZ$ at the Tevatron and LHC. . . . .	111
6.2	Feynman diagrams for the Higgs production processes that can give four charged leptons in the final state. . . . .	111
6.3	Scatter-plot of subleading $p_T$ $M_{ll}$ vs leading $p_T$ $M_{ll}$ for MC $ZZ$ and $H \rightarrow ZZ$ samples, where (a) $m_H = 120 \text{ GeV}/c^2$ and (b) $m_H = 200 \text{ GeV}/c^2$ . The blue lines represent the acceptance region of the $M_{\ell\ell-1,2}$ requirements applied in the $ZZ$ cross section measurement, which are released in this Higgs boson search. . . . .	114
6.4	Expectations for the several Higgs signal contribution in the considered sample. . . . .	116
6.5	The four lepton invariant mass distribution for $Z\gamma$ events containing one or two loosened denominator objects, with the fitted Landau function overlaid, used to extract the <i>fake</i> background modeling. . . . .	117

---

6.6	Correlated distribution of $\cancel{E}_T$ vs. $M_{4\ell}$ for the three main processes considered: (a) SM $ZZ$ production, (b) gluon fusion $H \rightarrow ZZ$ , and (c) associated production $ZH \rightarrow ZWW$ . The simulation assumes $m_H = 150 \text{ GeV}/c^2$ . . . . .	118
6.7	Kinematic distribution of the $m_{4\ell}$ (a,b) and $\cancel{E}_T$ (c,d) for assuming $m_H = 150 \text{ GeV}/c^2$ (a,c) and $m_H = 190 \text{ GeV}/c^2$ (b,d). In the plots are reported the predictions for the three most relevant contribution: in black is the SM non-resonant $ZZ$ production, in blue the resonant $gg \rightarrow H \rightarrow ZZ$ four lepton contribution, and in red the non-resonant $ZH \rightarrow ZWW$ four lepton contribution. . . . .	119
6.8	One dimensional projection of the $m_{4\ell}$ (a,b) and $\cancel{E}_T$ (c,d) distribution of the several processes predictions stacked over each other and the observed data overlaid. . .	120
6.9	Representation of the $\cancel{E}_T$ distribution variation considering the shifted samples. The $\pm 1\sigma$ variation on the shape of the $\cancel{E}_T$ distribution is considered as one of the systematic uncertainties in the limit calculation. . . . .	123
6.10	(a) Example of normalized posterior p.d.f. for $\mathcal{L}(R)$ , indicates as $\sigma/\sigma_{SM}$ , which is the ratio of the signal cross section to the standard model expectation. (b) Example of the expected distribution of 95% C.L. limits in the background-only hypothesis. The red line shows the median of the distribution, considered as the expected limit, while the green and yellow bands represent the $\pm 1\sigma$ and $\pm 2\sigma$ confidence interval respectively. . . . .	125
6.11	Expected and Observed 95% C.L. upper limit on the Higgs boson production cross section expressed as ratio to the SM prediction, in the Higgs mass range investigated. . . . .	126
7.1	Summary of the $ZZ$ cross section measurements done at the Tevatron (to date). . . . .	128
7.2	Summary of the main $VV$ diboson measurements carried out at the Tevatron. The measured production cross section are normalized to the corresponding SM NLO prediction and the blue band represent the theoretical uncertainty for each one. The last measurement is the one described in Chapter 5 of this Thesis. . .	129
7.3	Summary of the main $VV$ diboson measurements carried out at the LHC. The measured production cross sections are normalized to the corresponding SM NLO prediction and the blue band represent the theoretical uncertainty for each one. . .	130
7.4	Predicted and measured $WW$ , $WZ$ , and $ZZ$ , production cross section as a function of the center of mass energy $\sqrt{s}$ , for $pp$ and $p\bar{p}$ collisions. . . . .	131
7.5	Summary of the CDF 95% C.L. upper limits on the Higgs production cross section in the different final state considered. . . . .	133
7.6	The best fit signal cross section of all the CDF and D0 search channels combined, shown as a ration to the SM cross section as a function of the Higgs boson mass. The horizontal red line drawn at 1.0 correspond to the SM prediction. . . . .	134
7.7	Four lepton invariant mass distributions for the data collected at $\sqrt{s} = 7$ and 8 TeX from the (a) Atlas and (b) CMS experiment, in the $H \rightarrow ZZ$ search. In the plot is shown the expected signal for a Hies boson having $m_H \simeq 125 \text{ GeV}/c^2$ . . .	134
7.8	Summary of the (a) Atlas and (b) CMS 95% C.L. upper limits on the Higgs production cross section in the different final state considered. . . . .	135
7.9	Atlas and CMS signal cross section fit, shown as a ration to the SM cross section as a function of the Higgs boson mass. The horizontal red line drawn at 1.0 correspond to the SM prediction. . . . .	136
A.1	Likelihood distribution for real (left) and fake electrons (right). . . . .	143

---

C.1	Scale and smearing parameter obtained from the tuning procedure for electrons.	165
C.2	Scale and smearing parameter obtained from the tuning procedure for muons . . .	166
C.3	$M_{\ell\ell}$ distribution for (a) $e^+e^-$ and (b) $\mu^+\mu^-$ events in the Drell-Yan sample used for the tuning of the track resolution parameters. . . . .	167
C.4	$M_{\ell\ell}$ distribution for (a) $e^+e^-$ and (b) $\mu^+\mu^-$ events in the Drell-Yan sample after correcting the leptons momentum. . . . .	167
C.5	(a) $M_{\ell\ell}$ , (b) $p_T(\ell\ell)$ , (c) $\Delta\phi(\ell\ell)$ , and (d) $\cancel{E}_T$ distribution after correcting the scale and resolution of the lepton momenta. . . . .	168

---

# List of Tables

1.1	Flavor, charge and mass [3] of the quarks. . . . .	2
1.2	Charge, lepton number and mass [3] for the leptons. . . . .	2
1.3	Summary of the force carriers and their masses [3] . . . . .	3
1.4	Weak isospin, hypercharge and electric charge for the quarks and leptons. . . . .	6
2.1	Main $Z$ decay modes. . . . .	12
2.2	Cross section obtained by DELPHI measurements. The first errors are statistical while the seconds are systematics. . . . .	16
2.3	Summary of the recent measurements of the diboson production cross section in leptonic and semi-leptonic final states. The reference to the individual measurements are reported in the table as well as the used integrated luminosity; the statistical uncertainty appears before the systematic uncertainties (when both are available); the theoretical predictions are calculated at NLO [18, 45]. The results marked with † are obtained at the LHC working at $\sqrt{s} = 8$ TeV; for all the other results the measurement from CDF and D0 are assumed to be obtained at $\sqrt{s} = 1.96$ TeV and from CMS and Atlas at $\sqrt{s} = 7$ TeV. . . . .	19
3.1	Data acquired by the CDF detector. The table shows conventional separation into run periods, and for each the online range information and luminosity acquired. . . . .	45
4.1	Triggerable lepton types with the associated trigger paths. . . . .	60
5.1	Good run lists considered for different detector conditions with the corresponding integrated luminosity. The first three good run lists doesn't have any requirement on the silicon detector conditions, while the latter three explicitly requires the proper working of the silicon tracking subdetector. . . . .	63
5.2	Lepton identification efficiencies for data and simulation for periods 22 to 38. Errors are statistical only. . . . .	68
5.3	Expected events in the four-lepton collected sample for the $ZZ$ signal and background from fake lepton events, compared to the observed data. . . . .	74
5.4	Systematics uncertainties for signal and background that contribute to the final data set. . . . .	77
5.5	Summary of the properties of the processes concurring to the $ll\nu\nu$ signature. . . . .	82
5.6	List of physics processes simulated. We report the event generator, the cross section times branching fraction used in the normalization and the generator-level filter efficiencies ( $\epsilon_{\text{filter}}$ ), if any. Processes with a final state indicating a generic lepton $\ell$ are generated separately for $e$ , $\mu$ , and $\tau$ . . . . .	83
5.7	Expected and observed number of events in $9.7 \text{ fb}^{-1}$ of integrated luminosity for the several process contribution. The uncertainty includes statistical and systematic errors added in quadrature. . . . .	89

---

5.8	Expected and observed number of events in $9.7 \text{ fb}^{-1}$ in the Drell-Yan Control Region. The errors include statistical and systematic uncertainties added in quadrature.	92
5.9	Expected and observed number of events in $9.7 \text{ fb}^{-1}$ in the $e\mu$ Control Region.	93
5.10	Systematics uncertainties for signal and background that contribute to the final data set for the $ZZ \rightarrow \ell\nu\nu$ cross section measurement.	106
6.1	Production cross section for the several Higgs production mechanism at the Tevatron and $\mathcal{BR}(H \rightarrow ZZ)$ in the considered Higgs mass range.	113
6.2	Expected and observed number of events in the considered sample.	114
6.3	Expected number of signal events in $9.7 \text{ fb}^{-1}$ . The uncertainty on the expected signal is about 10%. We are lacking $ZH \rightarrow ZWW$ MC above $200 \text{ GeV}/c^2$ .	115
6.4	Summary of the systematic uncertainties considered in the limit calculation that affect signal and background predictions. Each row corresponds to a single source of systematic uncertainty while each column represent one of the processes that contributes to the selected data sample. Empty spaces represent source of systematic uncertainties that are not relevant for that particular process.	121
6.5	Expected and Observed 95% C.L. upper limit on Higgs production cross section as function of the Higgs boson mass. Limits are expressed in unit of the SM predicted cross section.	126
A.1	Summary of the lepton types ( <i>categories</i> ) used in the analysis described in this thesis.	140
A.2	Tight Central Electron (TCE) identification requirements.	140
A.3	Likelihood-based central electron (LBE) identification requirements.	142
A.4	Forward (PHX) electron identification requirements.	143
A.5	Base identification requirements for all muon categories.	144
A.6	CMU- <i>only</i> muons identification requirements.	145
A.7	CMP- <i>only</i> muons identification requirements.	145
A.8	CMUP muons identification requirements.	145
A.9	CMX Arches and Miniskirt-Keystone muons identification requirements.	146
A.10	BMU muons identification requirements.	146
A.11	CMIOCES muons identification requirements.	147
A.12	CMIOPEs muons identification requirements.	147
A.13	CrkTrk identification requirements.	148
A.14	Central Electron Probe identification requirements.	148
A.15	Forward Electron Probe identification requirements.	149
A.16	Central and Forward Muon Probe identification requirements.	149
A.17	Tight lepton and corresponding loose objects used to measure the lepton identification efficiencies and misidentification rates.	150
A.18	Fakeable Central Electron identification requirements.	150
A.19	Fakeable Forward Electron identification requirements.	151
A.20	Fakeable Central Muon identification requirements.	151
A.21	Fakeable Forward Muon identification requirements.	151
A.22	Lepton identification efficiencies for data and simulation for period 0. Errors are statistical only.	152
A.23	Lepton identification efficiencies for data and simulation for periods 1 to 4. Errors are statistical only.	152

---

A.24	Lepton identification efficiencies for data and simulation for periods 5 to 7. Errors are statistical only. . . . .	153
A.25	Lepton identification efficiencies for data and simulation for periods 8 to 10. Errors are statistical only. . . . .	153
A.26	Lepton identification efficiencies for data and simulation for periods 11 to 12. Errors are statistical only. . . . .	154
A.27	Lepton identification efficiencies for data and simulation for period 13. Errors are statistical only. . . . .	154
A.28	Lepton identification efficiencies for data and simulation for periods 14 to 21. Errors are statistical only. . . . .	155
A.29	Lepton identification efficiencies for data and simulation for periods 22 to 38. Errors are statistical only. . . . .	155
B.1	Thresholds for JET20, JET50, JET70, JET100 trigger requirements. . . . .	158
B.2	Trigger efficiencies and livetimes of each trigger for each run periods. . . . .	160
C.1	Scale and resolution correction parameters obtained for central/forward electrons and muons. . . . .	164

---

# Bibliography

- [1] K. Hagiwara, R. D. Peccei, D. Zeppenfeld, and K. Hikasa. Probing the weak boson sector in  $e^+e^- \rightarrow W^+W^-$ . *Nuclear Physics B*, 282:253 – 307, 1987. pages iii, 137
  - [2] Martin Kober, Benjamin Koch, and Marcus Bleicher. First Order Calculation of the Inclusive Cross Section  $pp$  to  $ZZ$  by Graviton Exchange in Large Extra Dimensions. *Phys. Rev.*, D76:125001, 2007. pages iii
  - [3] J. Beringer, *et al.* *Phys. Rev. D*, 86:010001, 2012. pages 2, 3, 11, 12, 16, 62, 73, 163, 176
  - [4] Y. Fukuda, *et al.* Evidence for oscillation of atmospheric neutrinos. *Physical Review Letters*, 81(8):1562 – 1567, Aug 1998. pages 2
  - [5] I. J. R. Aitchison and A. J. G. Hey. *Gauge Theories in Particle Physics, 2 Volume Set*. Taylor & Francis, 3 edition, January 2004. pages 4
  - [6] R. Ellis, W.J. Stirling, and B.R. Webber. *QCD and Collider Physics*. Cambridge University Press, 2003. pages 4
  - [7] S. Narison. *QCD as a Theory of Hadrons*. Cambridge University Press, 2007. pages 4
  - [8] Peter W. Higgs. Broken symmetries and the masses of gauge bosons. *Physical Review Letters*, 13(16):508+, October 1964. pages 7
  - [9] Donald H. Perkins. *Introduction to High Energy Physics*. Cambridge University Press, April 2000. pages 8
  - [10] Experimental observation of isolated large transverse energy electrons with associated missing energy at  $\sqrt{s}=540$  GeV. *Physics Letters B*, 122(1):103 – 116, 1983. pages 11
  - [11] M. Banner and O. Kofoed-Hansen. Observation of single isolated electrons of high transverse momentum in events with missing transverse energy at the cern  $pp$  collider. *Physics Letters B*, 122(5-6):476–485, 1983. pages 11
  - [12] T. Aaltonen, *et al.* Precise Measurement of the  $W$ -Boson Mass with the CDF II Detector. *Phys. Rev. Lett.*, 108:151803, Apr 2012. pages 11
  - [13] Kenji Abe, *et al.* Direct measurement of  $A_b$  and  $A_c$  at the  $z^0$  pole using a lepton tag. *Phys. Rev. Lett.*, 83:3384–3389, Oct 1999. pages 13
  - [14] J. Abdallah *et al.*  $ZZ$  production in  $e^+e^-$  interactions at  $\sqrt{s} = 183$  GeV to 209 GeV. *Eur. Phys. J.*, C30:447–466, 2003. pages 15
-

- [15] Ansgar Denner, S. Dittmaier, M. Roth, and D. Wackerroth.  $O(\alpha)$  corrections to  $e^+e^- \rightarrow WW \rightarrow$  four fermions ( $+\gamma$ ): First numerical results from RACOON  $WW$ . *Phys.Lett.*, B475:127–134, 2000. pages 16
- [16] S. Jadach, W. Placzek, M. Skrzypek, B.F.L. Ward, and Z. Was. Final state radiative effects for the exact  $O(\alpha)$  YFS exponentiated (un)stable  $W^+W^-$  production at and beyond LEP-2 energies. *Phys.Rev.*, D61:113010, 2000. pages 16
- [17] Raimund Strohmer. Review of the properties of the  $W$  boson at LEP, and the precision determination of its mass. *Int.J.Mod.Phys.*, A18:5127–5220, 2003. pages 16
- [18] John Campbell and R. K. Ellis. Radiative corrections to  $Zb\bar{b}$  production. *Phys. Rev. D*, 62:114012, Oct 2000. pages 16, 19, 77, 100, 122, 176
- [19] V. M. Abazov, *et al.* Measurement of the  $WW$  Production Cross Section in  $p\bar{p}$  Collisions at  $\sqrt{s} = 1.96$  TeV. *Phys. Rev. Lett.*, 94:151801, Apr 2005. pages 16, 19
- [20] T. Aaltonen, *et al.* Measurement of the  $W^+W^-$  Production Cross Section and Search for Anomalous  $WW\gamma$  and  $WWZ$  Couplings in  $p\bar{p}$  Collisions at  $\sqrt{s} = 1.96$  TeV. *Phys. Rev. Lett.*, 104:201801, May 2010. pages 16, 19
- [21] A. Abulencia, *et al.* Observation of  $WZ$  Production. *Phys.Rev.Lett.*, 98:161801, 2007. pages 17
- [22] T. Aaltonen, *et al.* Measurement of the  $WZ$  Cross Section and Triple Gauge Couplings in  $p\bar{p}$  Collisions at  $\sqrt{s} = 1.96$  TeV. *Submitted to Phys. Rev. D - R.C.*, 2012. pages 17, 19
- [23] V. M. Abazov, *et al.* A measurement of the  $WZ$  and  $ZZ$  production cross sections using leptonic final states in  $8.6 \text{ fb}^{-1}$  of  $p\bar{p}$  collisions. *Phys.Rev.*, D85:112005, 2012. pages 17, 19
- [24] T. Aaltonen, *et al.* Measurement of the  $WW + WZ$  Production Cross Section Using the lepton + jets Final State at CDF II. *Phys. Rev. Lett.*, 104:101801, Mar 2010. pages 18, 19
- [25] T. Aaltonen, *et al.* Measurement of the  $WW + WZ$  production cross section using a matrix element technique in lepton + jets events. *Phys. Rev. D*, 82:112001, Dec 2010. pages 18, 19
- [26] V. M. Abazov, *et al.* Measurements of  $WW$  and  $WZ$  Production in  $W$  + jets Final States in  $p\bar{p}$  Collisions. *Phys. Rev. Lett.*, 108:181803, May 2012. pages 18, 19
- [27] T. Aaltonen *et al.* First Measurement of  $ZZ$  Production in  $p\bar{p}$  Collisions at  $\sqrt{s} = 1.96$  TeV. *Phys. Rev. Lett.*, 100:201801, 2008. pages 18
- [28] V. M. Abazov *et al.*  $ZZ \rightarrow \ell^+\ell^-\nu\bar{\nu}$  production in  $p\bar{p}$  collisions at  $\sqrt{s} = 1.96$  TeV. *Phys. Rev.*, D78:072002, 2008. pages 18
- [29] V. M. Abazov *et al.* Observation of  $ZZ$  production in  $p\bar{p}$  collisions at  $\sqrt{s} = 1.96$  TeV. *Phys. Rev. Lett.*, 101:171803, 2008. pages 18
- [30] T. Aaltonen, *et al.* Measurement of  $ZZ$  Production in Leptonic Final States at  $\sqrt{s}$  of 1.96 TeV at CDF. *Phys. Rev. Lett.*, 108:101801, Mar 2012. pages 18, 19
-



- [31] Georges Aad et al. Measurement of  $W^+W^-$  production in  $pp$  collisions at  $\sqrt{s} = 7$  TeV with the ATLAS detector and limits on anomalous  $WWZ$  and  $WW\gamma$  couplings. 2012. pages 18, 19
- [32] G. Aad, *et al.* Measurement of  $W^\pm Z$  production in proton-proton collisions at  $\sqrt{s} = 7$  TeV with the ATLAS detector. *The European Physical Journal C*, 72:1–24, 2012. pages 18, 19
- [33] Georges Aad et al. Measurement of  $ZZ$  production in  $pp$  collisions at  $\sqrt{s} = 7$  TeV and limits on anomalous  $ZZZ$  and  $ZZ\gamma$  couplings with the ATLAS detector. 2012. pages 18, 19
- [34] Measurement of the  $ZZ$  production cross section in the four-lepton channel using  $5.8 \text{ fb}^{-1}$  of ATLAS data at  $\sqrt{s} = 8$  TeV, 2012. Atlas-Conf-012-090. pages 18, 19
- [35] S. Chatrchyan, *et al.* Measurement of  $w^+w^-$  production and search for the higgs boson in  $pp$  collisions at  $\sqrt{s} = 7$  tev. *Physics Letters B*, 699(1 - 2):25 – 47, 2011. pages 18, 19
- [36] Measurement of the  $ww$ ,  $wz$  and  $zz$  cross sections at cms, 2011. CMS-PAS-EWK-11-010. pages 18, 19
- [37] Technical report. pages 18, 19
- [38] Measurement of  $ww$  production rate, 2012. CMS-PAS-SMP-12-013. pages 18, 19
- [39] Measurement of  $zz$  production cross section in  $zz$  to  $2\ell 2\ell'$  decay channel in  $pp$  collisions at  $\sqrt{s} = 8$  tev, 2012. CMS-PAS-SMP-12-014. pages 18, 19
- [40] S. Chatrchyan, *et al.* Measurement of the sum of  $WW$  and  $WZ$  production with  $W$ +dijet events in  $pp$  collisions at  $\sqrt{s} = 7$  TeV. Technical Report arXiv:1210.7544. CMS-SMP-12-015. CERN-PH-EP-2012-311, CERN, Geneva, Oct 2012. Submitted to Eur. Phys. J. C. pages 18, 19
- [41] CDF Collaboration. Evidence for  $WZ$  and  $ZZ$  production in final states with  $b$ -tagged jets at CDF, 2012. CDF Public Note 10805. pages 19
- [42] Evidence for  $WZ$  and  $ZZ$  production in final states with  $b$ -tagged jets. 2011. pages 19
- [43] TEVNPH Working Group for the CDF and D0 Collaborations. Combined CDF and D0 measurement of  $WZ$  and  $ZZ$  production in final states with  $b$ -tagged jets, 2012. pages 19
- [44] T. Aaltonen, *et al.* First observation of vector boson pairs in a hadronic final state at the tevatron collider. *Phys. Rev. Lett.*, 103:091803, Aug 2009. pages 19
- [45] John Campbell and R. K. Ellis. Next-to-leading order corrections to  $W + 2$  jet and  $Z + 2$  jet production at hadron colliders. *Phys. Rev. D*, 65:113007, Jun 2002. pages 19, 176
- [46] Benjamin W. Lee, C. Quigg, and H. B. Thacker. Strength of Weak Interactions at Very High Energies and the Higgs Boson Mass. *Phys. Rev. Lett.*, 38:883–885, Apr 1977. pages 19
- [47] Thomas Hambye and Kurt Riesselmann. Matching conditions and Higgs mass upper bounds revisited. *Phys.Rev.*, D55:7255–7262, 1997. pages 19
- [48] W. Langguth and I. Montvay. A numerical estimate of the upper limit for the Higgs boson mass. *Z.Phys.*, C36:725, 1987. pages 19
-

- 
- [49] G. Altarelli and G. Isidori. Lower limit on the Higgs mass in the standard model: An update. *Physics Letters B*, 337(1-2):141 – 144, 1994. pages 20
- [50] D. de Florian and M. Grazzini. Higgs production through gluon fusion: updated cross section at the Tevatron and the LHC. *Phys. Lett.*, B674:291–294, 2009. pages 20, 112, 122
- [51] C. Anastasiou, R. Boughezal, and F. Petriello. Mixed QCD-electroweak corrections to Higgs boson production in gluon fusion. *JHEP04*, 003, 2009. pages 20, 112, 122
- [52] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt. Parton distributions for the LHC. *Eur. Phys. J.*, C63:189–285, 2009. pages 20, 122
- [53] U. Aglietti, *et al.* Tevatron for LHC report: Higgs. hep-ph/0612172, 2006. pages 20
- [54] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt. Handbook of LHC Higgs Cross Sections: 1. Inclusive Observables, 2011. pages 21, 112, 122, 171
- [55] R. Barate *et al.* Search for the standard model Higgs boson at LEP. *Phys.Lett.*, B565:61–75, 2003. pages 22
- [56] D.C. Kennedy and B.W. Lynn. Electroweak Radiative Corrections with an Effective Lagrangian: Four Fermion Processes. *Nucl.Phys.*, B322:1, 1989. pages 22
- [57] LEP-EWWG. The lep electro-weak working group, 2009. <http://lepewwg.web.cern.ch/LEPEWWG/>. pages 22
- [58] J. Alcaraz. Precision Electroweak Measurements and Constraints on the Standard Model. 2009. pages 22
- [59] Tevatron New Physics and Higgs Working Group - CDF and D0 Collaborations. Updated Combination of CDF and D0 Searches for Standard Model Higgs Boson Production with up to  $10.0 \text{ fb}^{-1}$  of Data. 2012. pages 22, 24, 132, 171
- [60] G. Aad, *et al.* Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. *Physics Letters B*, 716(1):1 – 29, 2012. pages 23, 24, 132, 171
- [61] S. Chatrchyan, *et al.* Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. *Physics Letters B*, 716(1):30 – 61, 2012. pages 23, 24, 132, 171
- [62] R. Blair *et al.* The CDF-II detector: Technical design report. FERMILAB-PUB-96-390-E. pages 30, 42
- [63] G. Apollinari, Konstantin Goulianos, P. Melese, and M. Lindgren. Shower maximum detector for the CDF plug upgrade calorimeter. *Nucl. Instrum. Meth.*, A412:515–526, 1998. pages 37
- [64] D. Acosta, *et al.* The CDF Luminosity Monitor. *Nucl. Instrum. Meth.*, A461:540–544, 2001. pages 39
- [65] A. Bardi, *et al.* The CDF online silicon vertex tracker. *Nucl.Instrum.Meth.*, A485:178–182, 2002. pages 42
- [66] A. Bhatti, *et al.* The CDF level 2 calorimetric trigger upgrade. *Nucl.Instrum.Meth.*, A598:331–333, 2009. pages 43
-

- [67] David Cox. The CDF L2 Track Trigger Upgrade. 2008. Presented at 15<sup>th</sup> IEEE Real Time Conference 2007 (RT 07), Batavia, Illinois, 29 Apr - 4 May 2007. pages 43
- [68] J. Marriner. Secondary vertex fit with mass and pointing constrains, 1993. CDF/DOC/SEC\_VTX/PUBLIC/1996. pages 48
- [69] P. Gatti. *Performance of the new tracking system at CDF II*. PhD thesis, Università degli Studi di Padova, 2001. [www-cdf.fnal.gov/thesis/cdf5561\\_tracking\\_performance.ps.gz](http://www-cdf.fnal.gov/thesis/cdf5561_tracking_performance.ps.gz). pages 48, 49
- [70] C. Hays et al. Inside-Out Tracking at CDF. *Nucl. Instrum. Meth.*, A538:249, 2005. pages 50
- [71] CDF Collaboration. Measurement of the inclusive  $W$  and  $Z$  cross section in  $p\bar{p}$  collisions at  $\sqrt{s} = 1.96$  TeV. *Jour. of Phys.*, G34:2457, 2007. pages 51, 172
- [72] G.C. Blazey, et al. Run II Jet Physics, 2000. CDF/PUB/JET/PUBLIC/5293. pages 55
- [73] F. Abe, et al. Topology of three-jet events in  $p\bar{p}$  collision at  $\sqrt{s} = 1.8$  TeV. *Phys. Rev.*, D45:1448, 1992. pages 57
- [74] J. Adelman, et al. Determination of the Jet Energy Scale at the Collider Detector at Fermilab. *Nucl. Instr. Meth.*, A566:375, 2006. pages 58
- [75] A. Bhatti, et al. Determination of the jet energy scale at the collider detector at Fermilab. *Nucl.Instrum.Meth.*, A566:375–412, 2006. pages 58
- [76] Torbjorn Sjostrand, Stephen Mrenna, and Peter Skands. Pythia 6.4 physics and manual. *JHEP*, 05:026, 2006. pages 64, 82, 112
- [77] T. Aaltonen, et al. Studying the Underlying Event in Drell-Yan and High Transverse Momentum Jet Production at the Tevatron. *Phys.Rev.*, D82:034001, 2010. pages 65, 82
- [78] R. Brun, R. Hagelberg, M. Hansroul, and J.C. Lassalle. GEANT. pages 65
- [79] H.L. Lai et al. Global QCD analysis of parton structure of the nucleon: CTEQ5 parton distributions. *Eur.Phys.J.*, C12:375–392, 2000. pages 77
- [80] CDF Collaboration. Combined upper limit on standard model higgs boson production, 2009. CDF Public Note 9999. pages 78
- [81] W. S. Kendall, F. Liang, and J.-S. Wang. *Markov Chain Monte Carlo - Innovations and Applications*. Word Scientific, 2005. pages 78
- [82] S. Catani, L. Cieri, G. Ferrera, D. de Florian, and M. Grazzini. Vector boson production at hadron colliders: A Fully exclusive QCD calculation at NNLO. *Phys.Rev.Lett.*, 103:082001, 2009. pages 81, 100
- [83] Stefano Frixione and Bryan R. Webber. Matching nlo qcd computations and parton shower simulations. *JHEP*, 06, 2002. pages 82
- [84] G. Corcella, et al. HERWIG 6.4 release note. 2001. pages 82
- [85] U. Baur and E. L. Berger. Probing the weak-boson sector in  $Z\gamma$  production at hadron colliders. *Phys. Rev. D*, 47:4889–4904, Jun 1993. pages 82
-

- 
- [86] M. Feindt. A neural bayesian estimator for conditional probability densities. arXiv:physics/0402093v1, 2004. pages 98
- [87] T. M. Mitchell. *Machine Learning*. McGraw-Hill Science/Engineering/Math, 1997. pages 98
- [88] S. Moch and P. Uwer. Heavy-quark pair production at two loops in QCD. *Nucl.Phys.Proc.Suppl.*, 183:75–80, 2008. pages 100
- [89] J. Baglio and A. Djouadi. Predictions for Higgs production at the Tevatron and the associated uncertainties. *JHEP10*, 064, 2010. pages 112
- [90] S. Catani and M. Grazzini. An NNLO subtraction formalism in hadron collision and its application to Higgs boson production at the LHC. *Phys. Rev. Lett.*, 98:222002, 2007. pages 122
- [91] T. Aaltonen, *et al.* Novel inclusive search for the Higgs boson in the four-lepton final state at CDF. *Phys.Rev.*, D86:072012, 2012. pages 125, 137
- [92] P. Avery, *et al.* Precision Studies of the Higgs Golden Channel  $H \rightarrow ZZ^* \rightarrow 4\ell$ . Part I. Kinematic discriminants from leading order matrix elements. 2012. pages 138
- [93] E. J. Thomson, *et al.* Online track processor for the CDF upgrade. *IEEE Trans.Nucl.Sci.*, 49:1063–1070, 2002. pages 159
-