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 XXII

Searches for a high-mass Higgs boson produced in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV

Direttore della Scuola: Prof. Attilio Stella

Supervisore: Dott.ssa Donatella Lucchesi

Dottorando: Dott. Simone Pagan Griso

Searches for a high-mass Higgs boson produced in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV

Abstract

In this Thesis the electro-weak spontaneous symmetry breaking of the Standard Model (SM) is studied by searching for the scalar Higgs particle in the proton antiproton collisions of the Tevatron collider, at the center of mass energy $\sqrt{s} = 1.96$ TeV. We report on the inclusive search for the Higgs boson decaying to two W bosons in the final state characterized by two charged leptons (e or μ) and two neutrinos. The analysis of 4.8 fb^{-1} of data recorded by the Collider Detector at Fermilab (CDF) shows no evidence of its production. We set 95% Confidence Level upper limits on the production cross section as a function of the mass m_H in the range 110 to 200 GeV/c^2. For $m_H = 165$ GeV/c^2 the observed limit is 1.3 times the SM predicted cross section (σ_H); the expected limit, in absence of signal, is $1.2^{+0.6}_{-0.4}\sigma_H$. These limits are the most stringent ones set by a single experiment up to date, for $m_H > 130$ GeV/c^2.

Ricerca del bosone di Higgs ad alta massa prodotto in collisioni $p\bar{p}$ a $\sqrt{s} = 1.96$ TeV

Prefazione

In questa Tesi è studiato il meccanismo di rottura spontanea di simmetria del Modello Standard attraverso la ricerca della particella scalare di Higgs in collisioni protone antiprotone di Tevatron, ad un'energia del centro di massa $\sqrt{s} = 1.96$ TeV. Presentiamo la ricerca inclusiva del bosone di Higgs nel decadimento $H \to W^+W^-$ e con stato finale costituito da due leptoni carichi (e o μ) e due neutrini. L'analisi effettuata su 4.8 fb^{-1} di dati raccolti dall'esperimento Collider Detector at Fermilab (CDF) non mostra alcuna evidenza della sua produzione. Sono quindi calcolati i limiti al 95% di Livello di Confidenza sulla sezione d'urto di produzione di tale particella, in funzione della sua massa m_H, nell'intervallo tra 110 e 200 GeV/c^2. Per $m_H = 165$ GeV/c^2 il limite osservato è 1.3 volte la sezione d'urto predetta dal Modello Standard (σ_H), ed il limite atteso in assenza di segnale è $1.2^{+0.6}_{-0.4}\sigma_H$. I risultati mostrati in questa Tesi sono, ad oggi, i più stringenti limiti ottenuti da un singolo esperimento sulla produzione del bosone di Higgs per $m_H > 130$ GeV/c^2.

Contents

Introduction

The comprehension of the electro-weak symmetry breaking is one of the most important missing pieces of the particle physics as described by the Standard Model (SM). The SM predictions have been extensively tested in the past decades. The analysis of the data collected by the experiments of the e^+e^- collider LEP and of the $p\bar{p}$ collider Tevatron confirmed the SM predictions with an accuracy sometimes well below 1%. However the SM, in its simplest form, does not account for the mass of the observed particles. In particular the W and Z bosons, mediating the weak interaction, are experimentally found to have a mass of 80.399 ± 0.023 GeV/c^2(1) and 91.1875 ± 0.0021 GeV/c^2(2) respectively.

The Higgs mechanism provides a solution to the problem introducing a new scalar field into the theory. All the fermions and the weak quanta of the theory acquire mass through this mechanism. Moreover the Higgs mechanisms also predicts a new scalar particle, the Higgs boson (H), whose existence has yet to be experimentally confirmed. The experimental evidence of its existence would be a strong indication of the validity of the Higgs mechanism as the way that the Nature chose to give mass to the observed particles. The mass of this particle (m_H) is a free parameter of the theory, but it is deeply connected with the energy scale at which the underlying electro-weak symmetry is broken to give mass to the W and Z bosons. Direct searches performed by the LEP experiments found no evidence of the Higgs production and constrain $m_H > 114.4$ GeV at 95% C.L. Moreover indirect constraints from quantum corrections induced by the Higgs boson to other SM observables indicate an upper bound of $m_H \lesssim 186$ GeV/c^2.

In this Thesis we present the search for the Higgs boson produced in proton antiproton collisions at the center of mass energy $\sqrt{s} = 1.96$ TeV using the data collected by the Collider Detector at Fermilab (CDF). The small cross section predicted for the Higgs production at the Tevatron (on the order of $0.1 - 1$ pb depending on its mass) makes the search very challenging. Backgrounds with cross sections bigger than the Higgs signal by several orders of magnitude require the development of advanced analysis techniques to statistically disentangle the signal contribution in the observed data.

A brief review of the SM theory is presented in this Thesis in **Chapter 1**, with emphasis on the Higgs mechanism. We present also the current status of the Higgs searches and the existing constraints on its mass m_H.

We focus on the Higgs searches at hadron colliders in **Chapter 2**, describing the production and the decay properties as function of its mass at the Tevatron and, for comparison, at the upcoming Large Hadron Collider (LHC). The possible Higgs signatures are analyzed in the hypothesis of a mass greater than ~ 135 GeV/c^2 (High Mass Higgs), in which case the Higgs preferentially decays into two massive gauge bosons. The $H \to WW$ decay channel is shown to be the most promising high mass Higgs search at the Tevatron.

Chapter 3 describes the experimental apparatus used in the Thesis for this search: the Tevatron collider and the CDF detector.

We analyze 4.8 fb^{-1} of data collected by the CDF experiment in **Chapter 4** and study the event selections for the $H \to WW$ signature: two oppositely charged leptons and two neutrinos. The expected sample composition for the selected data is discussed. The search for the Higgs boson produced by gluon fusion, which is the dominant production mechanism at the Tevatron, is performed in **Chapter 5** using multivariate techniques to improve the discrimination of signal from backgrounds. We search the Higgs boson produced in association with a W or Z boson and by the W or Z fusion in **Chapter 6** with analogous techniques.

The results of this Thesis are combined and discussed in **Chapter 7**. We describe also the combination of the CDF and DØ results and discuss the future prospects for the high mass Higgs searches at the Tevatron in the next one or two years.

In **Chapter 8** we conclude the Thesis summarizing the results obtained and the significant impact they have to the current knowledge of the Higgs boson properties.

Chapter 1

The Standard Model and the Higgs mechanism

The *Standard Model* (SM) unifies the electro-magnetic, weak and strong interactions providing predictions that have succesfully been tested in the last decades.

In this Chapter we present a brief review of the electro-weak sector of this theory.

The Higgs mechanism is described in detail. It provides the most elegant and economic way known up to date to give mass to the weak quanta and to the observed fermions. The Higgs boson is predicted as a consequence of this mechanism, but its existence is still not experimentally confirmed.

In this Chapter we will also review the existing direct and indirect constraints on its mass, which is a free parameter of the theory.

1.1 The Standard Model of particle physics

The extraordinary predictive capability and the formal elegance that determined the success of *Quantum Electro-Dynamics* (QED), on which our present understanding of electromagnetic interactions is based, induced theoretical physicists to try to extend its formulation to an analogous gauge theory able to describe also the weak and strong nuclear interactions. These models are based on the simple idea that, by requiring the physical laws to be invariant with respect to opportune local gauge transformations, fermionic fields describing ordinary matter can be associated with gauge fields responsible of their interactions.

The Standard Model (3; 4; 5; 6) is a quantum field theory based on the gauge symmetry group $SU(3)_C \times SU(2)_L \times U(1)_Y$. The first gauge group $SU(3)_C$ is related to the description of the strong interactions which affect quarks only and are mediated by gluons. $SU(3)_C$ defines the *Quantum Chromo-Dynamic*, QCD, theory. On the other hand, $SU(2)_L \times U(1)_Y$ is the underlying symmetry which provides a theoretical description of electromagnetic and weak interactions.

Within the Standard Model framework matter is composed of elementary particles that are subject to interactions mediated by what we call *gauge particles*. The fundamental constituents of matter are fermions classified into leptons and quarks, which are further

	Generation		
Quarks	1^{st}	2^{nd}	3^{rd}
$Q = +\frac{2}{3}$	up (u) $M_u \approx 1.5 \div 4.5\ MeV$	charm (c) $M_c \approx 1.15 \div 1.35\ GeV$	top (t) $M_t = 173.1 \pm 1.3\ GeV$
$Q = -\frac{1}{3}$	down (d) $M_d \approx 4.0 \div 8.0\ MeV$	strange (s) $M_s \approx 80 \div 130\ MeV$	bottom (b) $M_b \approx 4.1 \div 4.9\ GeV$
Leptons	1^{st}	2^{nd}	3^{rd}
$Q = -1$	electron (e) $M_e = 0.511\ MeV$	muon (μ) $M_\mu = 106\ MeV$	tau (τ) $M_\tau = 1.78\ GeV$
$Q = 0$	electron neutrino (ν_e) $0 < M_{\nu_e} < 3\ eV$	muon neutrino (ν_μ) $0 < M_{\nu_\mu} < 0.19\ MeV$	tau neutrino (ν_τ) $0 < M_{\nu_\tau} < 18.2\ MeV$

Table 1.1: Fermions in the Standard Model. Mass values are from (7). In this Table and through all the Thesis we set $\hbar = c = 1$.

organized into three families, called *generations*. To each generation are associated particles and their corresponding anti-particles, the latter having the same properties as the partner particles, but opposite charges (the *charge* of the particle is the quantum number that defines the coupling of the particle to the force carriers). Table 1.1 summarize the elementary particles of the SM with their electric charge and mass.

1.1.1 Electro-Weak interactions

The electro-weak sector of the Standard Model is based on the gauge symmetry group $SU(2)_L \times U(1)_Y$. $SU(2)_L$ is the weak isospin group which acts on left-handed fermions; $U(1)_Y$ is the weak hypercharge group. $SU(2)_L \times U(1)_Y$ has four generators, three of which from $SU(2)_L$: $T_i = \frac{\sigma_i}{2}$ with $i = 1, 2, 3$; and the fourth one from $U(1)_Y$, $\frac{Y}{2}$. The commutation relations of the group read:

$$[T_i, T_j] = i\epsilon_{ijk}T_k; \quad [T_i, Y] = 0; \quad i, j, k = 1, 2, 3. \tag{1.1}$$

The left-handed fermions transform as doublets under $SU(2)_L$,

$$f_L \to e^{i\vec{T}\vec{\theta}} f_L; \quad f_L = \begin{pmatrix} \nu_L \\ e_L \end{pmatrix}, \quad \begin{pmatrix} u_L \\ d_L \end{pmatrix}, ... \tag{1.2}$$

whereas the right-handed fermions are singlets for $SU(2)_L$ transformations:

$$f_R \to f_R; \quad f_R = e_R, u_R, d_R, ... \tag{1.3}$$

First-generation fermion quantum numbers are provided in Table 1.2 and are related to each other by the following equation:

$$Q = T_3 + \frac{Y}{2}. \tag{1.4}$$

The number of associated gauge bosons, being equal to the number of the symmetry group generators, is four: W^i_μ $(i = 1, 2, 3)$ and B_μ, associated to $SU(2)_L$ and $U(1)_Y$ respectively.

Fermions	T	T^3	Q	Y
ν_L	1/2	1/2	0	−1
e_L	1/2	−1/2	−1	−1
e_R	0	0	−1	−2
u_L	1/2	1/2	2/3	1/3
d_L	1/2	−1/2	−1/3	1/3
u_R	0	0	2/3	4/3
d_R	0	0	−1/3	−2/3

Table 1.2: First-generation fermion quantum numbers within the Standard Model.

The Standard Model Lagrangian construction for the electro-weak sector follows the same rules as for any gauge theory. In particular, $SU(2)_L \times U(1)_Y$ symmetry is promoted from global to local by replacing the field derivatives by their corresponding covariant derivatives. For a generic fermion field f, the covariant derivative reads:

$$D_\mu f = \left(\partial_\mu - ig\vec{T}\cdot\vec{W}_\mu - g'\frac{Y}{2}B_\mu\right) f, \tag{1.5}$$

where g and g' are the coupling constants associated to $SU(2)_L$ and $U(1)_Y$, respectively.

Similarly to QED and QCD, the electro-weak Lagrangian includes kinetic terms for the gauge fields:

$$\mathcal{L}_G = -\frac{1}{4}W^i_{\mu\nu}W_i^{\mu\nu} - \frac{1}{4}B_{\mu\nu}B^{\mu\nu}, \tag{1.6}$$

where the field strength tensors are defined as follows:

$$W^i_{\mu\nu} = \partial_\mu W^i_\nu - \partial_\nu W^i_\mu + g\epsilon^{ijk}W_{\mu j}W_{\nu k} \tag{1.7}$$

$$B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu. \tag{1.8}$$

The physical gauge bosons $W^\pm_\mu$, Z_μ and A_μ are obtained from the electro-weak interaction eigenstates by the following expressions, in which θ is the weak mixing angle:

$$W^\pm_\mu = \frac{W_{1\mu} \mp iW_{2\mu}}{\sqrt{2}} \tag{1.9}$$

$$Z_\mu = W_{3\mu}\cos\theta - B_\mu\sin\theta \tag{1.10}$$

$$A_\mu = W_{3\mu}\sin\theta + B_\mu\cos\theta \tag{1.11}$$

The gauge invariant interactions and the fermion kinematics are generated by $\bar{f}iD_\mu\gamma^\mu f$ terms in the Lagrangian. The Standard Model Lagrangian of the electro-weak sector will thus be:

$$\mathcal{L}_{SM} = \mathcal{L}_f + \mathcal{L}_G + \mathcal{L}_{SSB} + \mathcal{L}_{YW}, \tag{1.12}$$

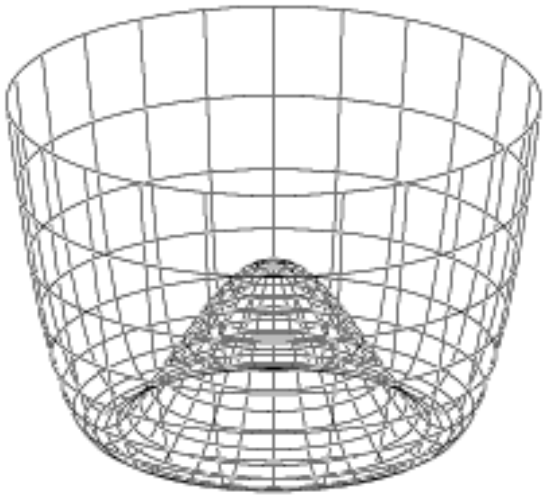


Figure 1.1: The Higgs potential

where the last two terms are the spontaneous symmetry breaking and Yukawa Lagrangians respectively, and are needed in order to properly introduce gauge bosons and fermion mass terms in a gauge invariant way. For $SU(2)_L \times U(1)_Y$ gauge invariance terms like $M_W^2 W_\mu W^\mu$, $M_Z^2 Z_\mu Z^\mu$ and $m_f^2 \bar{f} f$ are indeed forbidden. It will be the spontaneous symmetry breaking and the Higgs mechanism to provide the right mass generation as will be described in the following.

1.2 The Higgs mechanism

To resolve the mass generation problem, the spontaneous symmetry breaking (SSB) phenomenon is introduced. The SSB happens when the Lagrangian describing the dynamics of a physical system has a symmetry that is not preserved by the system ground states.

With the SSB, for a given gauge theory based on a local invariance with respect to a symmetry group G, given $H \subset G$ the symmetry group of the vacuum state, and being $dim(G) = N$ and $dim(H) = M$, $N - M$ massless Goldstone bosons will be absorbed by $N - M$ massive vector bosons. Therefore, in the $SU(2)_L \times U(1)_Y$, where $dim(G) = 4$ and $H = U(1)_{em}$, three vector bosons will realize the desired mass spectrum. This mechanism requires, as the simplest choice, the introduction of a doublet of complex fields, the Higgs field, of which three of the four degrees of freedom will be spent for the longitudinal polarization state of the massive bosons. The remaining degree of freedom is associated to the presence of the undetected Higgs particle, H_0.

The results of this theoretical environment is that the SSB mechanism is responsible for the reduction of the symmetry group of the theory from $SU(2)_L \times U(1)_Y$ to $U(1)_{em}$, the latter being related to the electric charge conservation.

The simplest SSB Lagrangian for the $SU(2)_L \times U(1)_Y$ can be written as:

$$\mathcal{L}_{SSB} = (D_\mu\Phi)^\dagger(D^\mu\Phi) - V(\Phi); \tag{1.13}$$

$$V(\Phi) = -\mu^2\Phi^\dagger\Phi + \lambda(\Phi^\dagger\Phi)^2; \quad \lambda > 0; \tag{1.14}$$

where $\Phi = \binom{\phi^+}{\phi^0}$ is a complex doublet with hypercharge $Y(\Phi) = 1$, and $V(\Phi)$ is the simplest renormalizable potential. For $(-\mu^2) < 0$, the minimum of the potential is reached on a circle of radius $v = \sqrt{\mu^2/\lambda}$ (see Fig. 1.1), and

$$| < 0|\Phi|0 > | = \begin{pmatrix} 0 \\ v/\sqrt{2} \end{pmatrix}. \tag{1.15}$$

Consequently, in the lowest energy state the system spontaneously chooses one vacuum expectation value which no longer reflects the symmetry of the potential $V(\Phi)$. The physical spectrum is then realized by performing "small oscillations" around the vacuum state. By parameterizing $\Phi(x)$ as

$$\Phi(x) = \exp\left(i\frac{\vec{\xi}(x)\vec{\sigma}}{v}\right)\begin{pmatrix} 0 \\ (v + H(x))/\sqrt{2} \end{pmatrix}, \tag{1.16}$$

and eliminating the unphysical fields $\vec{\xi}(x)$ by means of gauge transformations, the mass spectrum can be obtained from the following term of $\mathcal{L}_{SM}$:

$$(D_\mu\Phi')^\dagger(D^\mu\Phi') = \frac{g^2v^2}{4}W_\mu^+W^{-\mu} + \frac{1}{2}\frac{(g^2+g'^2)v^2}{4}Z_\mu Z^\mu + \ldots \tag{1.17}$$

$$V(\Phi') = \frac{1}{2}2\mu^2H^2 + \ldots \tag{1.18}$$

$$\mathcal{L}_{YW} = \lambda_e\frac{v}{\sqrt{2}}\bar{e'_L}e'_R + \lambda_u\frac{v}{\sqrt{2}}\bar{u'_L}u'_R + \lambda_d\frac{v}{\sqrt{2}}\bar{d'_L}d'_R + \ldots \tag{1.19}$$

Explicitly, the tree level mass predictions for gauge and Higgs bosons read:

$$\begin{aligned} M_{W_\mu^\pm} &= \frac{gv}{2} \\ M_{Z_\mu} &= \frac{\sqrt{g^2+g'^2}v}{2} \\ M_{A_\mu} &= 0 \\ M_{Higgs} &= \sqrt{2\lambda}v, \end{aligned} \tag{1.20}$$

where

$$v = \sqrt{\frac{\mu^2}{\lambda}} \tag{1.21}$$

is determined from the muon decay: $v = (\sqrt{2}G_F)^{-1/2} \sim 246\ GeV$ and it fixes the scale of the spontaneous symmetry breaking.

The mechanism here described, called the Higgs mechanism, gives mass terms for $W^{\pm}$, Z, as well as for quarks and leptons preserving the gauge invariance of the theory. On the other hand, it introduces a new scalar particle, not yet experimentally observed, the Higgs boson, whose mass and self-interaction are not theoretically determined.

1.3 Higgs searches: state of the art

Up to now the only unknown parameter of the Standard Model is the Higgs boson mass m_H. After its determination, production modes and corresponding cross sections of the Higgs boson, as well as its lifetime and decay branching ratios, can be predicted.

However limits on the Higgs boson mass exist from theoretical considerations and experimental observations.

1.3.1 Theoretical constraints

At high energy, the elastic scattering amplitude of massive weakly interacting bosons W, $WW \rightarrow WW$, increases indefinitely with the energy for longitudinally polarized particles, due to the linear dependence of the longitudinal wave-function W_L on particle energy. In particular, the amplitude of the partial wave for angular momentum $J = 0$, $A_0 = G_F s/8\pi\sqrt{2}$, presents a quadratic divergence that violates the unitarity limit for elastic scattering amplitude if

$$s > 4\pi\sqrt{2}/G_F \sim (1.2\ TeV)^2. \tag{1.22}$$

Unitarity can be restored by considering the exchange of the Higgs scalar particle. Imposing the s-wave initarity condition, we can get the Higgs mass m_H upper bound(8):

$$m_H < \sqrt{\frac{8\pi\sqrt{2}}{3\mathrm{G_F}}} \simeq 1\ \mathrm{TeV}. \tag{1.23}$$

Another upper limit on m_H is provided by the analysis of the triviality of the Higgs potential(9). The running coupling constant λ has a functional dependence on the energy scale Λ from the renormalization group equations. If m_H is large, the value of λ becomes infinite at some enrgy scale (Landau pole): the perturbation theory has ceased to be maningful long before. Similar bounds manifest as large cutoff effects in lattice calculations(10), which confirm or even tighten the m_H limits. The limits depend on the scale of energy Λ where these effects become large. For $\Lambda = 10^{19}$ GeV Ref. (9) finds $m_H < 180 \pm 6$ GeV. Less strict bounds are obtained if we assume that new physics modify SM theory at lower energy scales.

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. This is also refered to as stability bound. At large energy scales Λ it is well approximated by requiring the Higgs quartic self-coupling to remain positive: $\lambda(\Lambda) > 0$. Such an analysis has been carried out at two-loop level (11). If we expect the Standard Model is valid up to a scale $\Lambda = 10^{19}$ GeV the stability condition requires $m_H \gtrsim 135$ GeV.

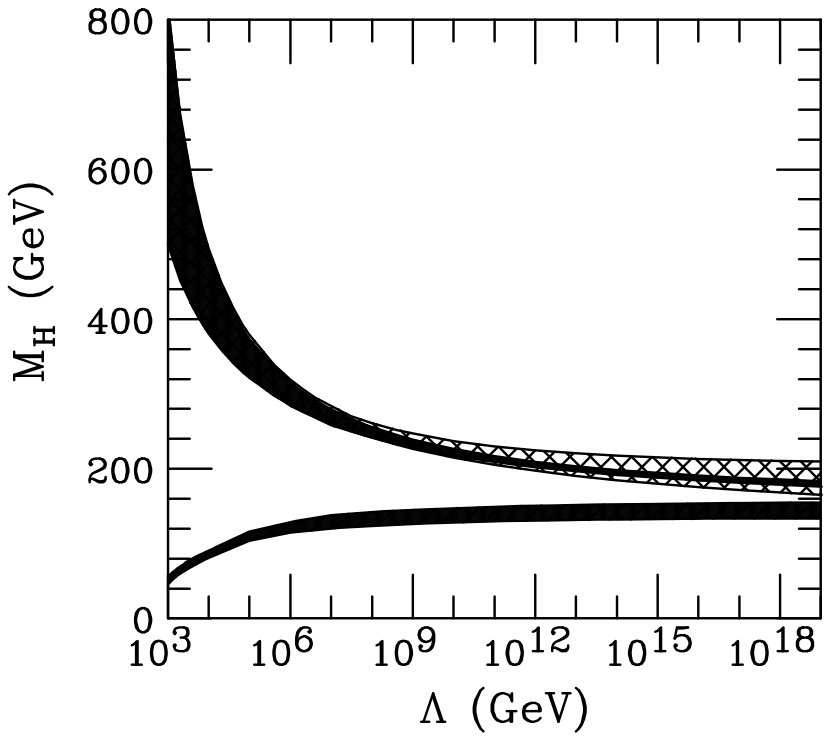


Figure 1.2: Summary of theoretical bounds on m_H. The upper solid band indicates the triviality upper bounds with the theoretical uncertainties. The cross-hatched shows the effect of varying the top mass quark used in the calculation in the range $150 - 200$ GeV; the nominal value used is $m_{top} = 175$ GeV. The lower solid area represents the theoretical uncertainties in the m_H lower bounds derived from stability requirements.

The triviality and stability bounds are showed as function of the energy scale Λ in Figure 1.2. The Figure shows that if we expect new physics contibutions to be sizeable before $\Lambda = 10^{19}$ GeV, the bounds on m_H becomes looser thant what perviously stated.

1.3.2 Experimental constraints

Direct searches at LEP-II

Direct searches for the production of the Higgs boson have been performed at the four experimets (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 and the vector boson fusion, which Leading Order Feynman diagrams are shown in Figure 1.3, the former being largely dominant. The LEP experiments performed detailed analyses on the $\sim 2461\ pb^{-1}$ of total integrated luminosity collected by the four detectors (at $189 \leq \sqrt{s} \leq 209\ GeV$) 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(12) enstablishes a lower bound of 114.4 GeV at 95% C.L. on its mass.

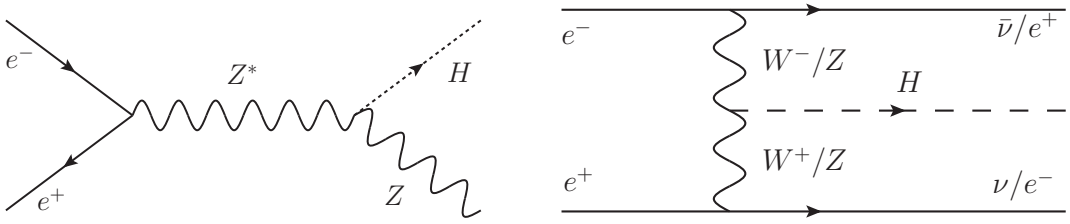


Figure 1.3: The Standard Model dominant Leading Order Feynman diagrams for the production of the Higgs boson at the LEP e^+e^- collider. Left: Higgs-strahlung from a virtual Z; Right: vector boson fusion.

Direct searches at the Tevatron Run-I

Direct searches for the Higgs boson production have been also performed at the Tevatron collider with $p\bar{p}$ collisions at the center of mass energy of $\sqrt{s} = 1.8$ TeV. Results from the CDF and DØ experiments, analyzing ~ 100 pb^{-1} of data collected during the Run I (1992 to 1995), were far from probing the Standard Model Higgs production above LEP limits, setting limits on the Higgs production cross section at 95% C.L. ~ 20 times larger than the expected cross section for the Higgs mass hypothesis $m_H \sim 115$ GeV(13).

Indirect constraints

In addition to direct searches, indirect constraints on the Higgs mass come from electro-weak measurements in the SM. The radiative corrections to the relative strength ρ of the charged and the neutral current depends in particular on the W mass m_W, the top mass m_t and the Higgs mass m_H(14):

$$\rho = \frac{m_W^2}{m_Z^2 \cos^2\theta_{\mathrm{W}}} = \frac{1}{1-\delta\rho}. \tag{1.24}$$

At three level $\delta\rho = 0$. One-loop corrections, for example due to the $t-b$ loop and the $H-W$ loop shown in Figure 1.4, contribute to $\delta\rho$:

$$\delta_\rho^{t-b} = \frac{3\mathrm{G_F}}{8\sqrt{2}\pi^2}(m_t^2 + m_b^2 - 2\frac{m_t^2 + m_b^2}{m_t^2 - m_b^2}\ln\frac{\mathrm{m_t^2}}{\mathrm{m_b^2}}), \tag{1.25}$$

$$\delta_\rho^{H-W} = \frac{3\mathrm{G_F}}{8\sqrt{2}\pi^2}(\frac{m_W^2 \sin^2\theta_{\mathrm{W}}}{\cos^2\theta_{\mathrm{W}}}\ln\frac{\mathrm{m_H^2}}{\mathrm{m_W^2}}). \tag{1.26}$$

The dependence of the correction on m_t and m_W is quadratic, while the Higgs contribution is only proportional to $\ln(\mathrm{m_H^2})$. A measurement of the parameter ρ can therefore translate into a constraint on the Higgs mass m_H, provided that the other parameters entering the Equations are known. It is clear that precise measurements of the top and W mass are crucial in contraining the Higgs mass.

Other SM electro-weak observables depend on the Higgs mass through radiative corrections, althought their dependence is in general milder than for the ρ parameter.

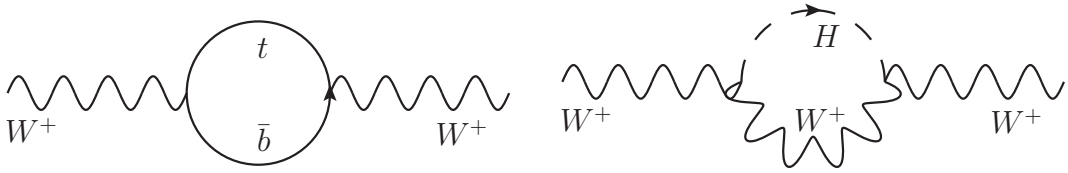


Figure 1.4: Feynman diagrams of one loope correction of the W mass due to the $t - b$ loop (Left) and the $H - W$ loop (Right).

Precision measurement results on electro-weak observables from the four LEP experiments, CDF, DØ, NuTeV, and SLD are combined by the LEP electro-weak working group (LEP-EWWG)(15). The ZFITTER(16) package is used to evaluate the dependence of the SM observables on the Higgs mass. A global SM fit as function of the Higgs mass is then performed using as input all the available observables (17). The minimum χ^2 of the fit divided by the number of degrees of freedom is χ^2_{min}/n.d.f. = 17.3/13 = 1.33, corresponding to a p-value of 18%. The corresponding Higgs mass is $m_H = 87^{+35}_{-26}$ GeV, and the 95% one-sided confidence level (C.L.) upper limit is $m_H < 157$ GeV. Figure 1.5 shows the $\Delta\chi^2 = \chi^2 - \chi^2_{min}$ of the fit as function of the Higgs mass.

Including the direct searches at LEP, the 95% one-sided C.L. upper limit from the electro-weak precision measurements becomes $m_H < 186$ GeV(17), thus cornering the Higgs mass in the range $114.4 < m_H < 186$ GeV.

A similar analysis is performed using the GFITTER(18) package. This package implements a Bayesian approach for performing the fit, compared to the frequentist method used by the LEP-EWWG(17); althought there are small differences in the treatment of the errors from the LEP-EWWG work, the two results are in agreement.

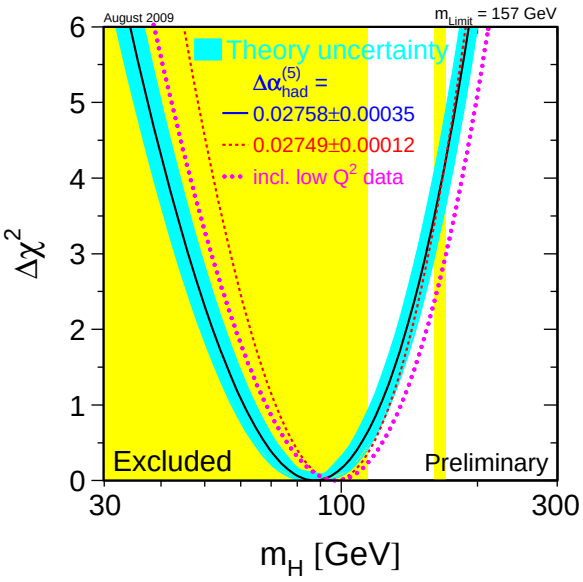


Figure 1.5: $\Delta\chi^2 = \chi^2 - \chi^2_{min}$ of the fit to electro-weak precision measurements performed by the LEP-EWWG. The line is the fit result; the band represents an estimate of the theoretical error due to missing higher order corrections. The vertical band shows the 95% exclusion limit on m_H from direct searches at LEP-II (up to 114 GeV) and Tevatron (which will be the topic of this Thesis). The dashed curve is the result obtained using a slightly different estimation of the hadronic contribution to the running of the QED fine structure constant. The dotted curve also includes low-Q^2 data in the fit(17).

Chapter 2

Hunting for the Higgs boson at hadron colliders

Nowadays hadron colliders provide the highest center of mass energy ($\sqrt{s}$) particle collisions. Therefore they are the ideal place to look for the production of the Standard Model Higgs boson if its mass lies in the range indicated by direct and indirect constraints. In this Chapter we summarize Higgs production and decay properties at the $p\bar{p}$ collider Tevatron, operating at $\sqrt{s} = 1.96$ TeV, and at the pp Large Hadron Collider (LHC), with the nominal $\sqrt{s} = 14$ TeV. Then we discuss the most promising signatures at the Tevatron, focusing on $m_H \gtrsim 130$ GeV, and we outline the possible search strategies. The future prospects at the upcoming high-energy collider LHC are described to provide a comparison.

2.1 Higgs boson production

In the Standard Model the Higgs boson couples preferentially to heavy particles, in particular to W and Z bosons and to top quark. This determines the dominant production processes. Figure 2.1 shows the Leading Order Feynman diagrams of the four main production processes at hadron colliders.

Direct production ($gg \to H$) via gluon fusion and the top quark loop is the dominant production mechanism. The production cross section is greatly enhanced by QCD radiative corrections(19). The latest available calculation(20; 21; 22), implements a full QCD Next-to-Next-to-Leading Order (NNLO) approximation. It also takes into account soft gluon resummation, NLO electro-weak corrections and NLO effects of the $b - quark$ in the fermionic loop. The Parton Distribution Functions (PDF) of the (anti-)proton used to evaluate the cross section for hadronic collisions are the NNLO MSTW2008(22). The uncertainty on the cross section calculation has mainly two origins: the one arising from our limited knowledge of the PDFs, and the one coming from the partonic cross section. The PDF are varied within their errors and the effect is propagated through the calculation to estimate the corresponding cross section uncertainty. The most important source of uncertainty on the partonic cross section comes from uncalculated higher order QCD diagrams. Their effect

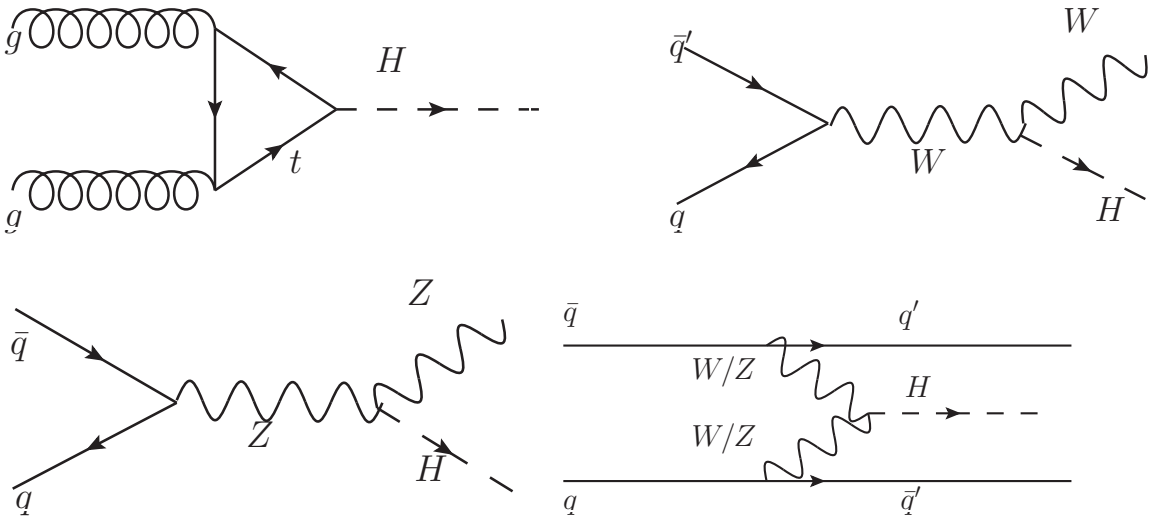


Figure 2.1: Leading order Feynman diagrams of the dominant Higgs production mechanisms at hadron colliders: gluon fusion (top left), associated production (top right and bottom left) and vector boson fusion (bottom right).

on the cross section calculation is estimated varying the normalization (μ_R) and factorization (μ_F) scale around the hard scale $\mu_R = \mu_F = m_H$. The total uncertainty is on the order of 10%(20).

Associated production : $q\bar{q} \to Z^0 H$ and $q\bar{q}' \to W^{\pm} H$. In this case, the Higgs boson is radiated by a W or Z boson. These are electro-weak processes which receive small ($\sim 1\%$) contributions by radiative corrections. Calculations at Next-to-Next-to-Leading Order (NNLO)(23) are available with an accuracy of less than 5%.

Vector Boson Fusion takes place when each of the incoming quarks interacts through charged or neutral current processes and the two Ws or Zs bosons couple to the Higgs: $q\bar{q} \to Hq\bar{q}$. This is a pure electro-weak process that is known at NLO with an estimated accuracy of $\sim 10\%$(23).

Figure 2.2 (Left) shows the expected production cross sections at the Tevatron. The four processes are shown separately. The gluon fusion has the cross section the order of $0.1 - 1$ pb in the range $100 < m_H < 200$ GeV, contributing for 78% to the inclusive Higgs production cross section. The associated production cross sections decrease from ~ 0.3 pb to ~ 0.01 pb for heavy Higgs. They contribute together for 15% of the total cross section. Vector boson fusion has behaviour similar to the associated production, and contributes $\sim 7\%$ to the inclusive cross section.

At the Large Hadron Collider (LHC) the Higgs production cross sections are about two orders of magnitude larger. Figure 2.2 (Right) shows the cross sections expected at LHC for the four production mechanisms considered. Direct production is $\sim 88\%$ of the inclusive cross section, ranging from 75 pb to 11 pb for $100 < m_H < 300$ GeV. Associated productions have smaller relative contribution to the total cross section, respect to Tevatron, being $\sim 3\%$ and rapidly decreasing for heavy Higgs. The production through vector

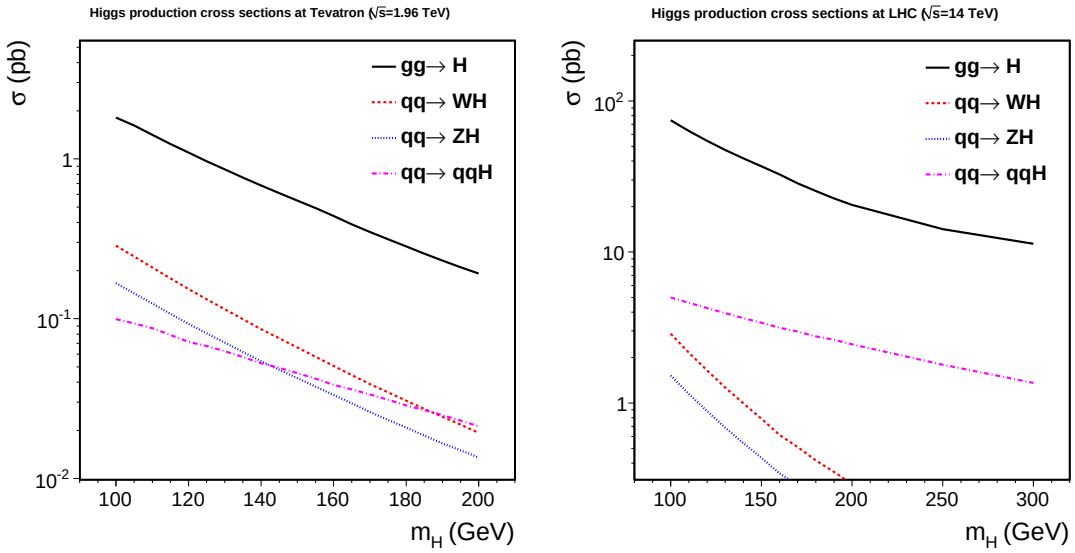


Figure 2.2: Production cross sections at the $p\bar{p}$ collider Tavatron (Left, $\sqrt{s} = 1.96$ TeV) and at the pp collider LHC (Right, $\sqrt{s} = 14$ TeV), as function of the Higgs mass m_H.

boson fusion has a significantly larger cross section respect to the associated production in all the mass range considered, and contributes 9% to the Higgs production.

2.2 Higgs boson decay

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 partial widths of Higgs decay channels are calculated at Next-to-Leading order using HDECAY(24), and they are shown in Figure 2.3, as functions of the Higgs mass, together with the expected branching fractions.

For $m_H < 2m_W$ 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. The total decay width is less than 10 MeV in this range. For $2m_W < m_H < 2m_Z$ there is a rapid increase of the decay widths into vector bosons, even if the ZZ decay channel is still partially suppressed. The resulting branching fractions are dominated by the WW decay. In the mass range above $m_H > 2m_Z$ both the decay channels to WW and ZZ produce on-shell vector bosons and represent the major part of the decay width of the Higgs, which rapidly increases to ~ 1 GeV. For very high Higgs masses, $m_H > 2m_{top}$, also the decay into a top pair becomes important and the total width reaches the 10 GeV. Higgs decays to leptons less often than to heavy quarks, given the lower mass. However the most favored leptonic decay channel $H \to \tau\tau$ reaches branching fractions of $\sim 7 - 8\%$ for $m_H = 100 - 120$ GeV. The Higgs does not couple directly to massless particles, but it can also decay through loop-mediated processes. This is the case, for example, of the $H \to \gamma\gamma$ decay channel, which has a significant branching

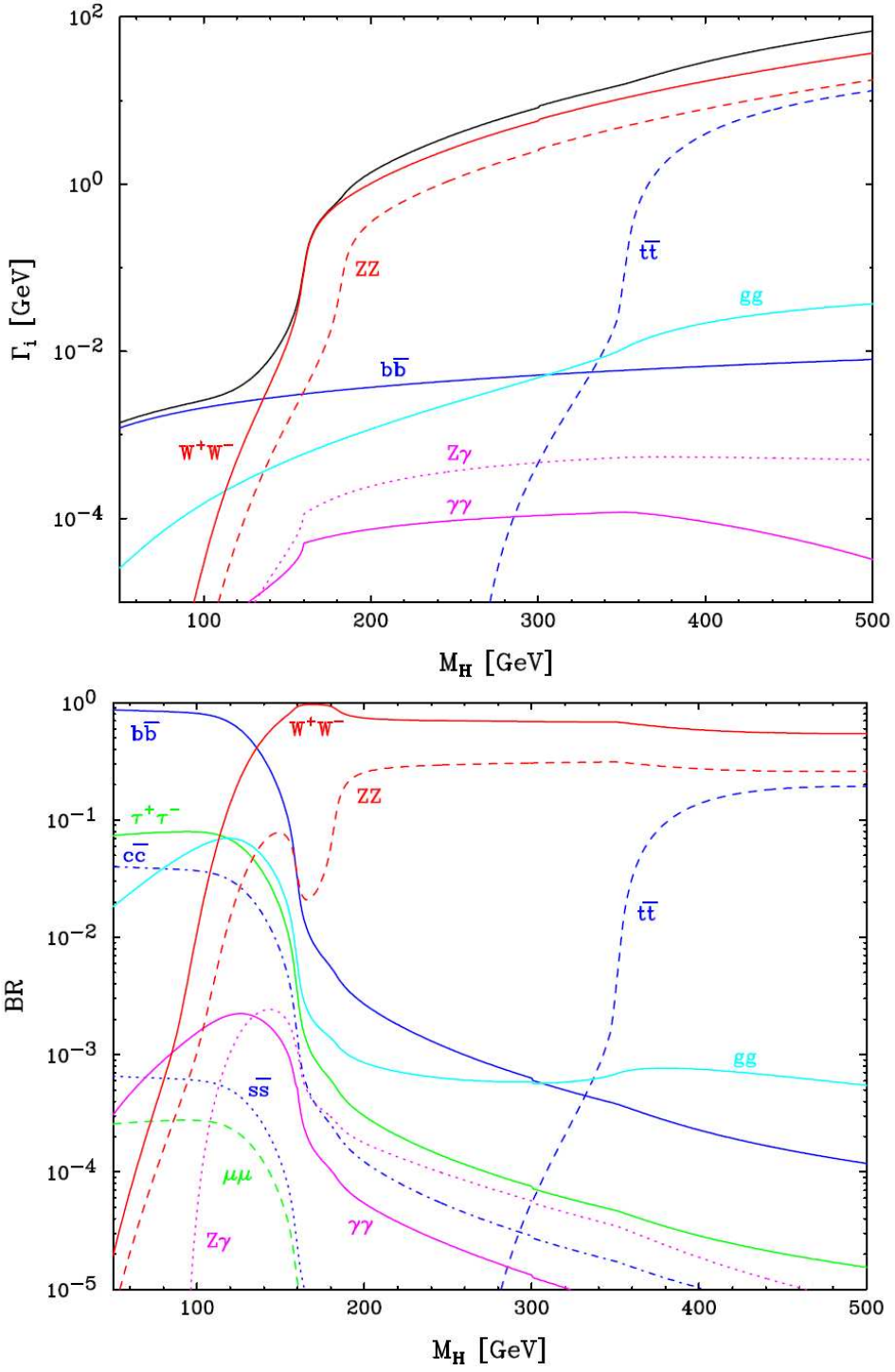


Figure 2.3: Top: partial widths of the Higgs boson decay channels in the Standard Model; the curve without label represent the total width. Bottom: Branching fractions.

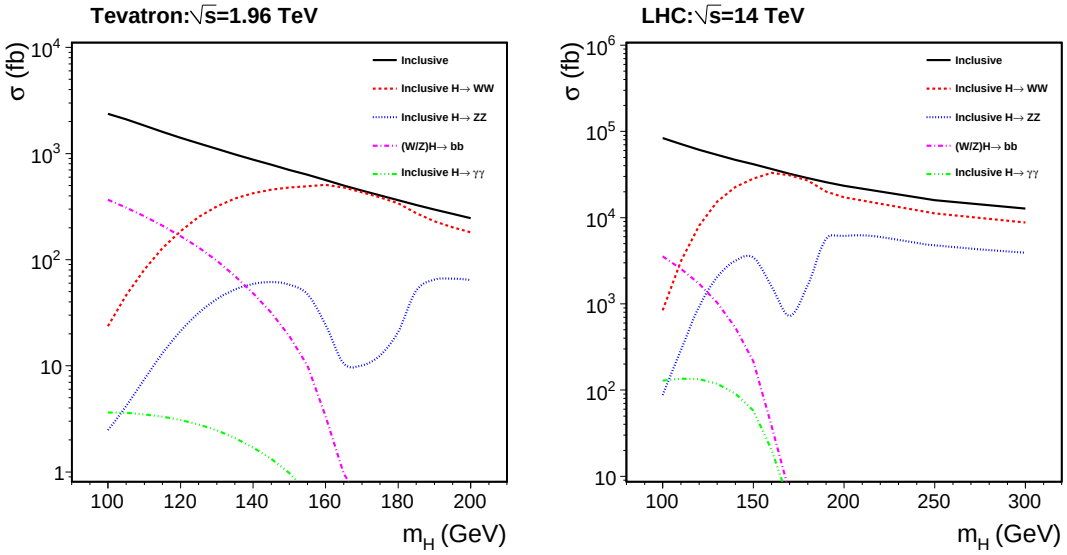


Figure 2.4: The cross section times branching fraction ($\equiv \sigma$) in fb for the main decay channels (i.e. the expected number of Higgs events in 1 fb^{-1} of data). The solid black line represents the inclusive cross section. Results refer: the left plot to the $p\bar{p}$ Tevatron collider with $\sqrt{s} = 1.96$ TeV, while the right plot to the pp LHC collider with $\sqrt{s} = 14$ TeV.

fractions for $m_H \lesssim 160$ GeV.

2.3 *High* and *low* mass Higgs region

Higgs boson search through the process $gg \to H \to b\bar{b}$ suffers of large background from non-resonant $b\bar{b}$ production that has the cross section several orders of magnitude larger.

Tevatron $H \to b\bar{b}$ searches exploit the associated production of the Higgs boson, with the W or Z bosons leptonic decays, to increase the expected signal to background ratio(25; 26; 27). However, for the rapid decrease of the $H \to b\bar{b}$ branching ratio and for the small cross section of the associated production, these searches are not powerful for $m_H \gtrsim 120$ GeV. Figure 2.4 shows that for $m_H = 135$ GeV the number of events obtained by associated production with $H \to b\bar{b}$ are four times less than $m_H \sim 100$ GeV.

$H \to \gamma\gamma$ at Tevatron suffers the low number of expected events and detectors not designed for having good photon reconstruction. The results in this decay channel are not competitive with $H \to b\bar{b}$ searches(28).

At the LHC searches in the $H \to b\bar{b}$ decay channel are even less sensitive than at Tevatron because of the more unfavorable signal to background ratio.

The $H \to \gamma\gamma$ decay channel provides instead one of the best ways to look for Higgs boson in the $110 - 140$ GeV mass range at the LHC. Figure 2.4 (right) shows the number of events expected as function of the Higgs mass. ATLAS Collaboration reports(29), for example, a 3σ evidence potential for Higgs boson discovery in the mass range $110 - 140$ GeV with ~ 10 fb^{-1} of data using the $H \to \gamma\gamma$ decay channel. Higgs searches in this

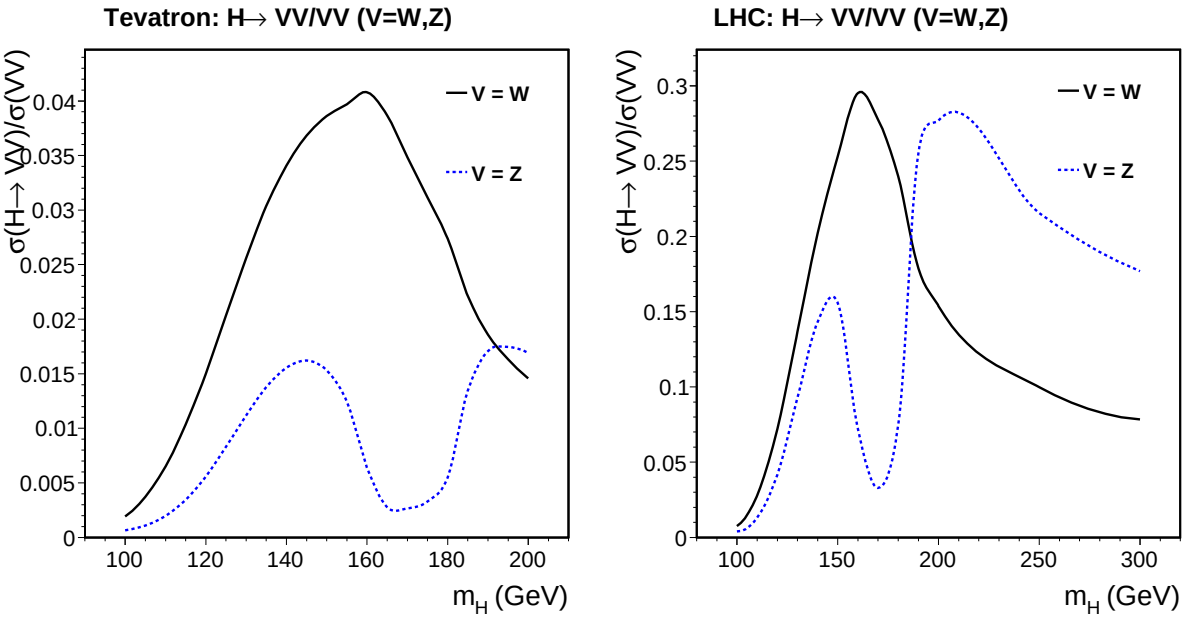


Figure 2.5: Ratio of resonant $H \to WW, ZZ$ and non-resonant WW, ZZ production cross sections as function of the Higgs mass. The results are shown for Tevatron (left) and LHC (right).

mass range, exploiting either the $b\bar{b}$ or $\gamma\gamma$ decay channels, are commonly referred to *Low mass Higgs* searches.

If the Higgs mass $m_H \gtrsim 120$ GeV the number of expected $H \to WW$ events is larger than that expected for $H \to b\bar{b}$ by associated production (see Figure 2.4). Higgs searches exploiting di-boson Higgs decay channels are commonly referred as *High Mass Higgs* searches. The main background contributor to these processes is the non-resonant WW and ZZ production. Figure 2.5 shows the ratio of the cross sections for the resonant $\sigma(H \to WW/ZZ)$ over the non-resonant $\sigma(WW/ZZ)$ production. While at the Tevatron at maximum $\sim 4\%$ of the produced WW events are expected to come from Higgs decay, at the LHC the ratio is much more favorable. The hypothesis $m_H = 160$ GeV provides the best signal over background ratio for the $H \to WW$ process. Figure 2.5 shows that $H \to ZZ$ is expected to compete with the WW decay channel for $m_H \gtrsim 190$ GeV.

2.4 High mass Higgs boson search strategies

Higgs searches exploiting the WW and ZZ decay channels are the most promising if the Higgs boson has $m_H \gtrsim 120$ GeV. The analysis strategy strongly depends on the decay mode of the gauge bosons, since the expected backgrounds are very different. In this Section we review the potential for Higgs boson searches of the main decay modes of WW and ZZ di-bosons, roughly estimating the sensitivity of such searches to the Higgs signal.

Here and in the rest of the Thesis the sensitivity of a search at some Confidence Level (C.L.) is defined as the ability to measure a signal, over the expected background, with a

1-C.L. $p-$value, i.e. the probability that the observed signal originated from a background fluctuation is 1-C.L. The C.L. can also be expressed in units of σ of a normal gaussian.

2.4.1 Higgs boson decaying to W^+W^-

Hadronic WW decays

The W boson decays to hadrons with a branching fraction of $\sim 68\%$. However final states involving only hadronic jets suffer a large background from multi-jet production, which has a cross section larger than signal by several orders of magnitude.

When one of the two W bosons decays to leptons and the other one to hadrons, we have a final state with one high-p_T lepton (l), two jets (j) and energy imbalance in the detector for the undetected neutrino: $ljj+ \not\!\!\vec{E}_T$. The latter quantity is referred to as missing transverse energy ($\not\!\!\vec{E}_T$): since we expect the transverse energy to be almost zero before the collision any imbalance in the measured energy of the decay products in the transverse plane can indicate undetected particles[1].

W+jets is the major contributor to this sample, with the cross section much larger than non-resonant di-boson production. Z+jets significantly contributes too, with $\not\!\!\vec{E}_T$ arising from one undetected lepton of the Z decay.

There are not previous Higgs searches exploiting this decay channel at the Tevatron. However we can roughly estimate the potential sensitivity to the Higgs signal using the recent measurement performed at the CDF experiment of the sum of the WW and WZ cross sections in the $ljj+ \not\!\!\vec{E}_T$ decay channel(30). The poor mass resolution of the jet pair does not allow to separately measure the WW and WZ cross sections. This analysis evaluates an expected sensitivity of 5.1σ for the Standard Model $\sigma(WW+WZ) = 15.9 \pm 0.85$ pb(31) using 2.7 fb^{-1} of data. For a fair comparison with later results, we're interested in evaluating the cross section to which this measurement would be sensitive to at 95% C.L. (1.96σ). We assume the same analysis technique and the same analysis efficiencies for the Higgs signal as for the di-bosons. Therefore we expect to be sensitive at 95% C.L. to processes with a cross section of ~ 15.9 pb $\cdot\, 1.96/5.1 = 6$ pb. This scaling assumes the same amount of background, that statistical gaussian errors dominate and that the sensitivity is proportional to the number of signal events expected. Although this is a rough and optimistic scaling, it shows that we expect to be sensitive to cross sections about ten times larger than the inclusive $H \to WW$ one, for the most favorable case $m_H = 160$ GeV.

However a big improvement for the Higgs search in this decay channel can be obtained exploiting associated production of the Higgs. In fact, a distinctive signature can arise when the Higgs boson is produced in association with a W or Z boson, which then decays leptonically. As depicted in Figure 2.6, we can have two leptons with the same charge in the final state, missing transverse energy and jets. Although CDF simulations show that the expected signal is just 1.9 ± 0.2 events in 4.8 fb^{-1}, after analysis selections, backgrounds are heavily suppressed. The most significant physics background arise from the WZ production. However we expect also contribution from $W + jets$, when one of

[1] See Section 3.5.4 for a detailed discussion on $\not\!\!\vec{E}_T$ and its corrections in the case of the CDF experiment.

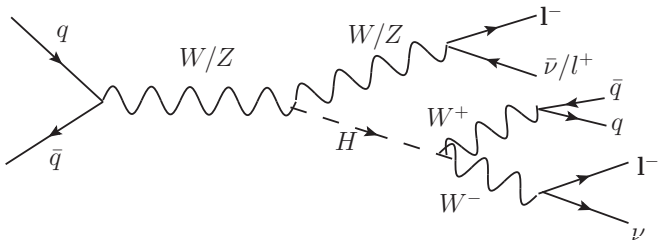


Figure 2.6: Associate production of Higgs boson with two leptons of the same charge in the final state.

the jets is erroneously identified as lepton. A detail study and results for this search at CDF are presented in Chapter 6 of this Thesis.

At the LHC, only high statistic samples allow significant results in this decay channel. The ATLAS collaboration performed a sensitivity study for the search of the Higgs boson produced by VBF and decaying to $H \rightarrow WW \rightarrow l\nu jj$(32), showing that at least 30 fb^{-1} of data are needed in order to reach a $> 5\sigma$ significant sensitivity for the most favorable high Higgs masses $m_H > 600$ GeV.

WW $\rightarrow$ lνlν

The fully leptonic decay channel offers the best signal/background ratio, with still a significant number of events expected. In fact the branching fraction of the WW to a pair of electrons or muons is about 6%, including leptonic tau decays. For $m_H = 160$ GeV it means that we expect about 30 events to be produced at Tevatron by all significant production mechanisms, and ~ 2000 at the LHC, for each fb^{-1} of data.

We decided to organize the analysis based on the production mechanism, since the expected backgrounds significantly differ.

Direct production of the Higgs boson has not jets in the final state at Leading Order. The signature of such process is then two opposite charged high$-p_T$ leptons, $\vec{\not{E}}_T$ coming from the two neutrinos and low jet activity in the detector. The main background process with the same signature is the non-resonant WW production. The Chapter 5 of this Thesis will describe the study and the results of the Higgs search with this signature.

However jets can be produced by initial state radiation, as depicted in Figure 2.7(Left). This can happen also for background processes, as the non-resonant WW(see Figure 2.7(Right)). While $gg \rightarrow H$ process is initiated by gluons, WW is produced mainly by quarks as initial state particles; it is thus expected that the Higgs signal will produce an additional initial state gluon more often than the WW process(33). In Chapter 6 we search for the Higgs boson in events with jet activity in the detector. We will see that half of the expected signal comes from the Higgs produced by gluon fusion with at least one jet from initial state radiation.

Measuring the WW cross section is a mandatory step for the Higgs search in the fully leptonic decay channel, since this process is the major background contributor. The result of the recent CDF measurement of the WW cross section(34) is compared to previous

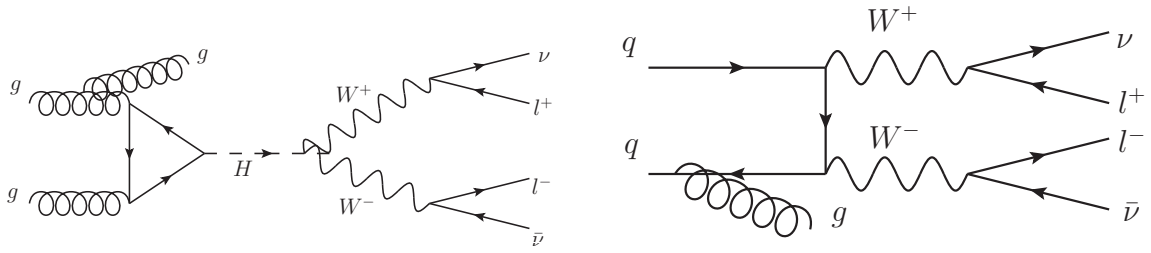


Figure 2.7: Left(Signal): Next-to-Leading Order contribution to direct Higgs production. Right(Background): Next-to-Leading Order contribution for WW production. Both processes lead to final state with two opposite sign leptons, neutrinos and one (or more) jets by initial state radiation.

Tevatron measurements in Figure 2.8. The latest analysis has been performed in 3.6 fb^{-1} of data using the same techniques developed for the Higgs searches and described in this Thesis. Therefore the good agreement of the measurement with the expected Standard Model $p\bar{p} \rightarrow WW$ cross section is a strong validation of the Higgs analysis too, and assure that we correctly understand the properties of the most important background for the Higgs search. Moreover the cross section measurement has a total relative error of 15%, compared to the 22% of the $\sigma(WW + WZ) \rightarrow ljj + \vec{\not{E}}_T$ measurement(30). We therefore already expect that the fully leptonic decay channel will have a better sensitivity for the Higgs search than the hadronic WW decay channels.

Associated production of the Higgs boson with a W or Z boson, is likely to have jets in the final state that come from the hadronic decay of the gauge boson, as depicted in Figure 2.9(Left). In this case the signature will be two opposite charged leptons, missing transverse energy coming from the leptonic decays of the W bosons, and the presence of additional jets in the detector. The **Vector Boson Fusion** (VBF) Higgs production process, which Leading Order Feynman diagram is shown in Figure 2.1, naturally has two jets in the final state. In addition to the non-resonant WW production, we can also expect significant background contribution from $t\bar{t}$ pairs in the sample. In fact it has the same final state as the signal when each top decays $t \rightarrow Wb \rightarrow l\nu b$, with two $b-$jets in the final state, as depicted schematically in Figure 2.9(Right). The search for the associated production of the Higgs boson and the vector boson fusion production is performed in Chapter 6 of this Thesis.

At the LHC, the pure leptonic decay of the WW pair provides the best signature for the $H \rightarrow WW$ signal, further favored by the higher ratio of its cross section to the non-resonant WW production (as shown in Figure 2.5). The CMS Collaboration reports a sensitivity study(35) which looks for direct production of the Higgs boson decaying to a W pair in the two opposite charged leptons and missing energy final state. Jets are vetoed in this study, which claims a 95% C.L. sensitivity to the Higgs signal for 140 GeV $< m_H <$ 190 GeV with just 1 fb^{-1} of collected data.

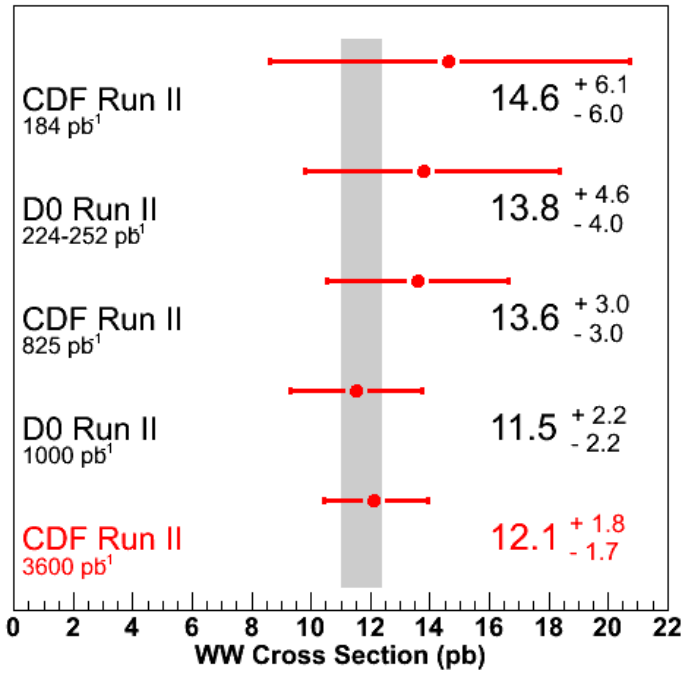


Figure 2.8: Comparison of $p\bar{p} \rightarrow WW$ cross section measurements at the Tevatron collider.

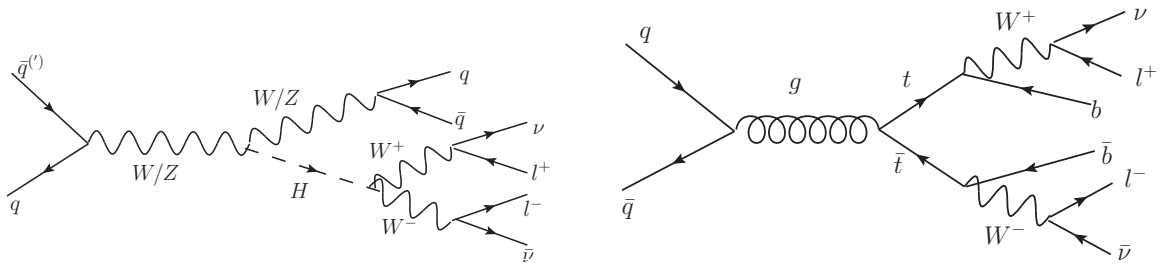


Figure 2.9: Left(Signal): Leading Order W/ZH production. Right(Background): Leading Order $t\bar{t}$ production. Both processes produce jets already at the Leading Order.

2.4.2 Higgs boson decaying to Z^0Z^0

Although the branching ratio $H \to ZZ$ is smaller than $H \to WW$, this decay channel can provide distinctive signatures and helps or even becomes the best way to look for the high mass Higgs at the LHC. There are not published or preliminary results at the Tevatron in this decay channel; in the following we will give an estimation of the sensitivity that is expected.

$\mathbf{ZZ \to lll'l'}$

The decay of both Z bosons into massive leptons happens with a branching fraction of only $\sim 0.6\%$. However it provides the cleanest signature, with a final state fully detectable by the detector. Moreover the good resolution achievable for the measurement of electron's and muon's momenta, compared to jets, allows to place tight constraints on the invariant mass of lepton pairs to separate signal from backgrounds.

The recent CDF observation of the ZZ production in the same decay channel(36) allows to estimate the sensitivity for the Higgs search. This analysis expects 4.7 ZZ events in 4.8 fb^{-1} after loose event selections, with residual background less than 0.1 events. Using the same analysis technique and selections, simulations show that we expect 0.19 ± 0.2 Higgs events in the same 4.8 fb^{-1} data sample. A simple counting experiment would have sensitivity at 95% C.L. for signal cross sections on the order of 1.3 pb, assuming the non-resonant ZZ as background and the sensitivity to be driven by $S/\sqrt{B}$, with S and B being the expected number of signal and background events respectively. This is about 25 times the expected inclusive $H \to ZZ$ cross section for the most favorable case of $m_H = 190$ GeV. We will see that results presented in this Thesis on the $H \to WW$ search have a sensitivity at 95% C.L. to Higgs signal cross sections that are about three times the one expected by the Standard Model for the same mass hypothesis(37). Even with tight cuts on the invariant mass of the four lepton system the statistics for the $H \to ZZ \to 4l$ search is simply too low to significantly contribute to the Higgs search at the Tevatron.

On the other hand the $H \to ZZ \to 4l$ decay channel will provide the best way to search for the high mass Higgs boson at the LHC collider. The dominant background contributor is still the non-resonant ZZ production. The inclusive Higgs signal cross section is larger than at the Tevatron and the ratio $\sigma(H \to ZZ)/\sigma(ZZ)$ is more favorable. CMS expects to have 95% C.L. sensitivity to the Standard Model inclusive Higgs production cross section in the range $m_H \sim 190 - 250$ GeV already with 1 fb^{-1}(38). Moreover in the LHC start-up phase, the 4-lepton resonance will provide an easier and unambiguous signature respect to $H \to WW$ searches, where a careful understanding and modeling of the missing transverse energy resolution is needed.

$\mathbf{ZZ \to ll\nu\nu}$

The Z boson decays to neutrinos with a branching fractions of 20%. Thus the final state involving two high$-p_T$ leptons and neutrinos offers a cross section times branching fractions more than five times bigger than the four leptons decay channel. Missing transverse energy can be used to separate signal from the overwhelming single Z production, but it still remains an important background.

CDF performed a measurement of the ZZ production cross section in the $ll\nu\nu$ decay channel using 1.9 fb^{-1}(39). Since no specific Higgs searches in this decay mode have been performed yet, we can use the cited analysis to roughly derive the expected significance to the Higgs signal. The main background contributors are WW and WZ production, and the latter is the most difficult to discriminate against Higgs signal. Residual backgrounds from single Z production and fake leptons are well separated from signal on a statistical basis. To make a comparison with the four massive lepton channel, we evaluated the expected signal and main backgrounds in 4.8 fb^{-1} using the same analysis selections of Reference (39). We expect $\sim$ 0.7 Higgs events in the most favorable mass hypothesis $m_H = 190$ GeV. This number has to be compared to $\sim$ 34 events expected from ZZ non-resonant production and $\sim$ 31 events from the WZ process. Using the same assumptions as for $ZZ \to 4l$ and considering the ZZ and WZ processes as the only backgrounds, we can derive what is the the expected Higgs cross section for which we have 95% C.L. sensitivity. We find a cross sections on the order of 23 times the Standard Model prediction.

However, while in the current $ZZ \to 4l$ decay channel the efficiency on Higgs signal events is already high and limited by detector coverage, the $ZZ \to ll\nu\nu$ analysis applies tight $\not\!\!\vec{E}_T$ requirements which are only $\sim$ 30% efficient for $H \to ZZ \to ll\nu\nu$ events. Reducing the $\not\!\!\vec{E}_T$ requirement to gain in signal efficiency will let more Z background to pass the selections. It is then important to have a good modeling of such background contribution and good techniques to discriminate it against signal. Studies on this decay channel at CDF are already ongoing and we expect they will significantly contribute to Higgs searches for $m_H \gtrsim 190$ GeV.

ZZ $\to$ llq$\bar{\text{q}}$, ZZ $\to \nu\nu$q$\bar{\text{q}}$ and Fully hadronic ZZ decays

If at least one of the Z bosons decays hadronically (Br($Z \to$hadrons$\sim$ 70%)) the branching ratio is larger than for the $ZZ \to ll\nu\nu$ decay. However we expect the $Z + jets$ process to become an important player. The first observation of di-boson production when one of the gauge bosons decays hadronically has been reported by the CDF collaboration in Ref. (40), selecting final states with $\not\!\!\vec{E}_T$ and two jets. Although this is an important result, the significance of the measurement is just above 5σ for the inclusive cross sections for all the di-boson processes (WW+WZ+ZZ), not having the mass resolution to distinguish each of them. A possible development of this measurement would be to require jets to be identified as $b - jets$, taking advantage of the more favorable ratio of $\sigma(Z \to b\bar{b})/\sigma(b\bar{b})$ respect to inclusive jets. CDF collaboration already reported the first step in this direction with the observation of the $Z \to b\bar{b}$ signal in 584 pb^{-1} of data(41): Figure 2.10 shows the expected and observed distribution of the invariant mass of the jets for the selected events with the Z signal contribution over the mostly $b\bar{b}$ background. However the relative small branching fraction $Br(Z \to b\bar{b}) \sim 15\%$ and the huge $b\bar{b}$ background contribution makes the measurement very challenging for the search of the Higgs boson.

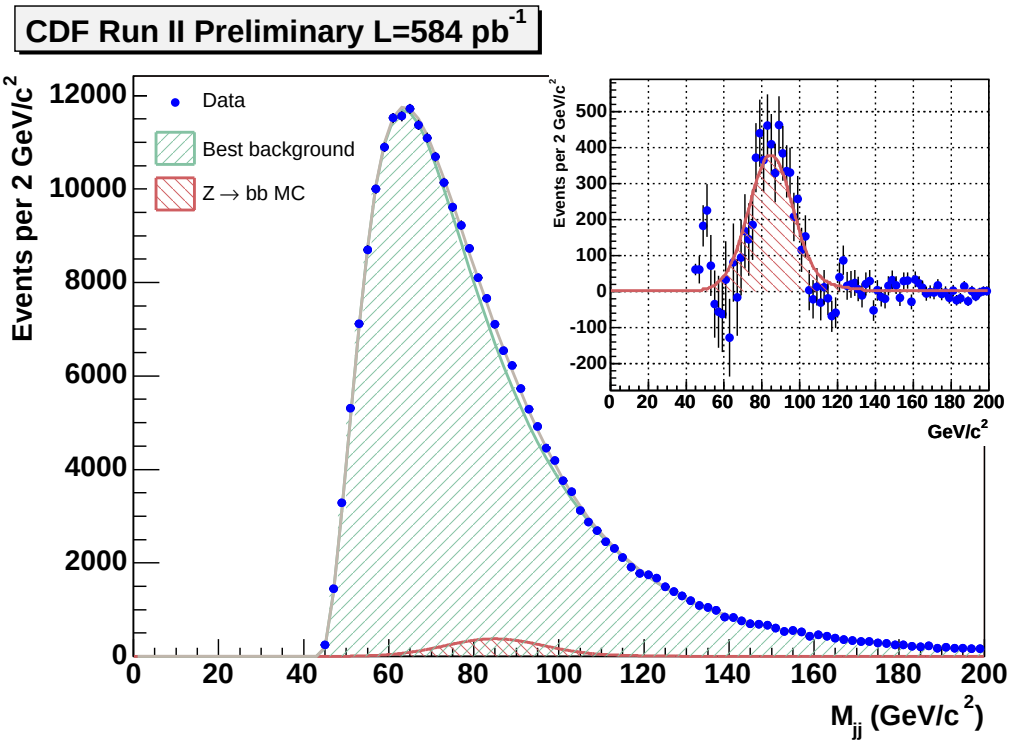


Figure 2.10: Invariant mass distribution of jets and $Z \rightarrow b\bar{b}$ signal contribution. The measured cross section is $\sigma(Z \rightarrow b\bar{b}) = 1578^{+636}_{-410}$ pb(41).

Chapter 3

Experimental environment

3.1 The Tevatron accelerator complex

The Tevatron(42) is a proton-antiproton synchrotron accelerator hosted at the Fermi National Accelerator Laboratory and producing $p\bar{p}$ collisions at the center-of-mass energy of $\sqrt{s} = 1.96\ TeV$. The final proton and antiproton beams are the results of a complex acceleration apparatus which involves different stages, spanning from proton and antiproton production, their acceleration and transfer towards different sub-systems, to their actual collision in designed interaction points where the CDF and DØ particle detectors are installed. A schematic view of the Tevatron acceleration chain is provided in Figure 3.1. In the following a description of the acceleration apparatus will be given.

3.1.1 The proton source

The process leading to $p\bar{p}$ collisions begins in a Cockroft-Walton chamber in which H^- gas is produced by hydrogen ionization. H^- ions are immediately accelerated by a positive voltage to an energy of 750 KeV and transported through a transfer line to the linear accelerator, the Linac.

The Linac(43) picks up the H^- ions at energy of 750 KeV, and accelerates them up to the energy of 400 MeV.

The Booster(44) takes the 400 MeV negative hydrogen ions from Linac and strips the electrons off, which leaves only protons. The Booster is the first circular accelerator in the Tevatron chain, and consists of a series of magnets arranged around a 75-meter radius circle with 18 radio frequency cavities interspersed. When the bare protons are collected in the Booster, they are accelerated to the energy of 8 GeV by the conventional method of varying the phase of RF fields in the accelerator cavities, and subsequently injected into the Main Injector. The final "batch" will contain a maximum of 5×10^{12} protons divided among 84 bunches spaced by 18.9 ns, each consisting of 6×10^{10} protons.

3.1.2 The Main Injector

The Main Injector (MI)(45) is a circular synchrotron seven times the circumference of the Booster that plays a central role in linking the Fermilab acceleration facilities: the

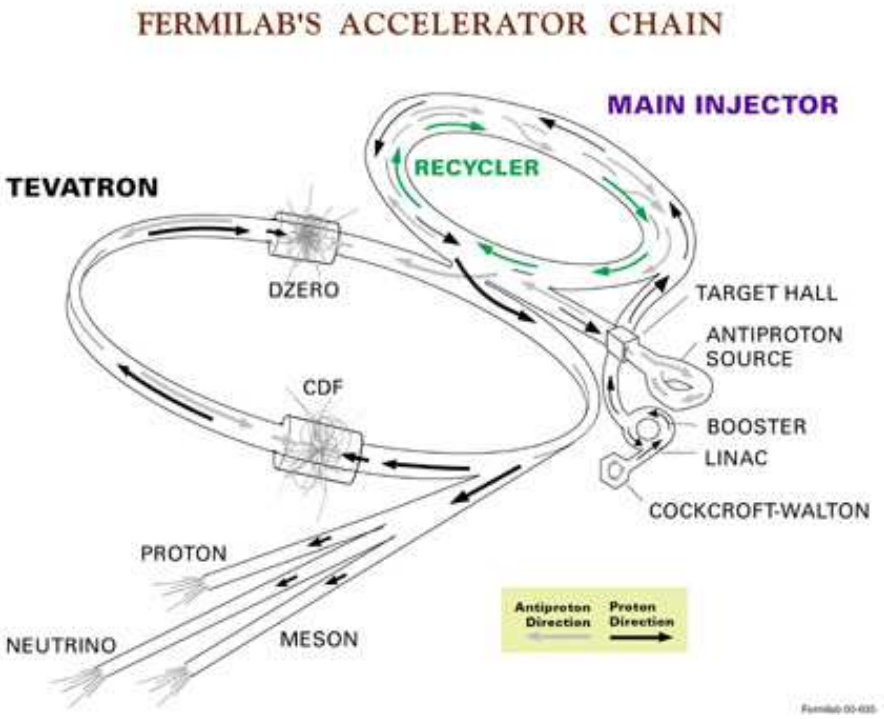


Figure 3.1: The FERMILAB's accelerator chain.

Main Injector can accelerate or decelerate particles, by means of radio frequency (RF) systems, between the energies of 8 GeV and 150 GeV. The sources of these particles and their final destination are variable, depending on the Main Injector operation mode: it can accept 8 GeV proton from the Booster or antiprotons from the Recycler and it can accelerate protons up to 120 GeV for antiproton production. The beam energy, for both proton and antiproton, can reach 150 GeV during the collider mode when particles are injected to the Tevatron for the last stage of the acceleration.

Antiproton production

Providing beam to the antiproton production target is one of the simplest tasks of the Main Injector. In this mode, a single batch of protons is accepted from the Booster, accelerated up to 120 GeV and extracted towards the target, which yields 8 GeV antiprotons as will be described below. Recently upgrades focused in doubling the number of protons on the target station. The new procedure, called "slip stacking" allows to merge two batches from the Booster before sending them to the target station.

Collider operations

Collider Mode is the most complex scenario that the Main Injector has to cope with: in addition to supplying 120 GeV protons for antiproton production, the Main Injector must also feed the Tevatron protons and antiprotons at 150 GeV. The protons and antiprotons need to be filled into super-bunches more intense than any individual bunch that can be

accelerated by the Booster. A process called coalescing has been developed for this task; coalescing takes place at Main Injector *flat top* (*i.e.* the maximum energy at which the machine can keep the particles for an extended time).

The sequence of steps needed during a shot (the scientific term for loading protons and antiprotons) can be described as follows:

- One batch (84 bunches) of protons is accelerated to 8 GeV by the Booster;
- Only 7 of the batch bunches are extracted to the Main Injector to be then accelerated to 150 GeV.
- At flat top the bunches are coalesced, *i.e.* pushed together to form a narrow, high intensity bunch.
- The coalesced bunch is injected to the Tevatron.
- Previous steps are repeated until 36 coalesced bunches are sent to the Tevatron.
- Meanwhile, the 8 GeV antiprotons from the production target have been stored in the Accumulator, waiting to be injected in opposite direction with respect to the protons to the Main Injector.
- When this happens, antiprotons are accelerated to 150 GeV, coalesced and injected in the Tevatron in the opposite direction taken by protons.
- Main Injector drops back to 8 GeV for another group of antiproton bunches. The processes repeat until 36 antiproton bunches have been delivered to Tevatron.

After these steps the Tevatron ring hosts 36 bunches of protons and 36 of antiprotons at the energy of 150 GeV.

3.1.3 The antiproton source

The number of antiprotons available has always been an important limiting factor in producing the high luminosity desired for Tevatron physics. They are difficult, or at least time-consuming, to produce. The performance of the proton source, moreover, greatly affects the quality and duration of the physics run of the Tevatron. Anyway colliding protons and antiprotons has greate advantages. For example the antiprotons can be accelerated in the same ring used for protons, because of the opposite charge, thus reducing the cost of the magnets for a second ring. Moreover, the production rate for a number of interesting processes is higher in $p\bar{p}$ collisions at $\sqrt{s}$ up to 3 TeV compared with pp collisions at the same energy.

The Fermilab antiproton source(46) is comprised of a target station, two rings called the Debuncher and Accumulator, and the transfer lines between these rings and the Main Injector. In general, the following steps are taken in order to produce an antiproton beam suitable for collisions in the Tevatron.

- A single batch of protons with an intensity up to $4 - 5 \times 10^{12}$ is accelerated by the Main Injector at 120 GeV. Proton beam intensities up to 7×10^{12} have been achieved.

- After the extraction the proton beam proceeds to the target area where its spot size is reduced by means of quadrupole magnets and is made to collide with a nickel target producing showers of secondary particles. Immediately downstream of the target station is located the collection Lithium lens module, in which a solenoidal magnetic field focuses the negative secondaries. Lithium was chosen because it is the least-dense solid conductor which in turn minimizes particles scattering and absorption.

- A pulsed dipole magnet follows the lens. Its purpose is to select 8 GeV negative charged particles and to force them towards the Debuncher. Most of particles with wrong charge-to-mass ratio are filtered out of the beam and collected by a graphite-core beam-dump.

- Surviving particles, typically 1 or 2 antiprotons for every 10^5 protons striking the target, are then injected in to the Debuncher where the momentum spread is reduced using stochastic and momentum cooling(46). The reduction of the momentum spread of incoming particles is needed in order to improve the Debuncher to Accumulator transfer efficiency, because of the limited momentum aperture of the Accumulator at injection. The Debuncher is a rounded triangular-shaped synchrotron with a mean radius of 90 meters.

- Just before the next pulse arrives from the target, the antiprotons are extracted from the Debuncher and injected to the Accumulator. As its name implies, the purpose of the Accumulator is to accumulate antiprotons. It is also a triangular-shaped synchrotron of radius 75 meters and is housed in the same tunnel as the Debuncher. It is the storage ring for the antiprotons; all of the collected antiprotons are stored here at 8 GeV and cooled until needed. Both RF and stochastic cooling systems are used in the momentum stacking process. The RF decelerates the recently injected pulses of antiprotons from the injection energy to the edge of the stack tail. The stack tail momentum cooling system sweeps the beam deposited by the RF away from the edge of the tail and decelerates it towards the dense portion of the stack, known as the core. Additional cooling systems keep the antiprotons in the core at the desired momentum and minimize the transverse beam size.

- When enough antiprotons have been accumulated in the Acculumator, their transfer starts. Antiproton beam destination can be either the Main Injector or the Recycler ring.

Figure 3.2 shows the general layout of the antiproton source at the Tevatron. Figure 3.3 shows the antiproton production complex with the addition of the functionalities provided by the Recycler ring.

Overall it can take from 10 to 20 hours to build up a stack of $\sim 3.5 \cdot 10^{12}$ antiprotons, which is then used in the Tevatron collisions.

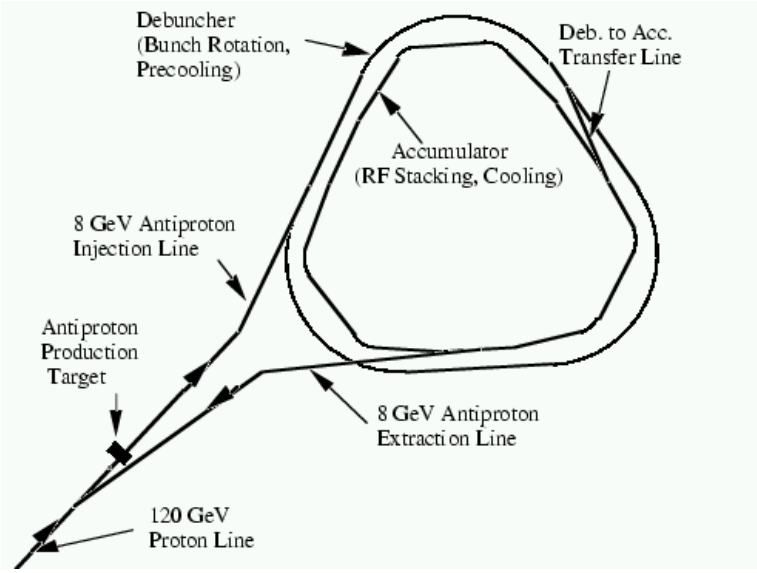


Figure 3.2: Layout of the antiproton source.

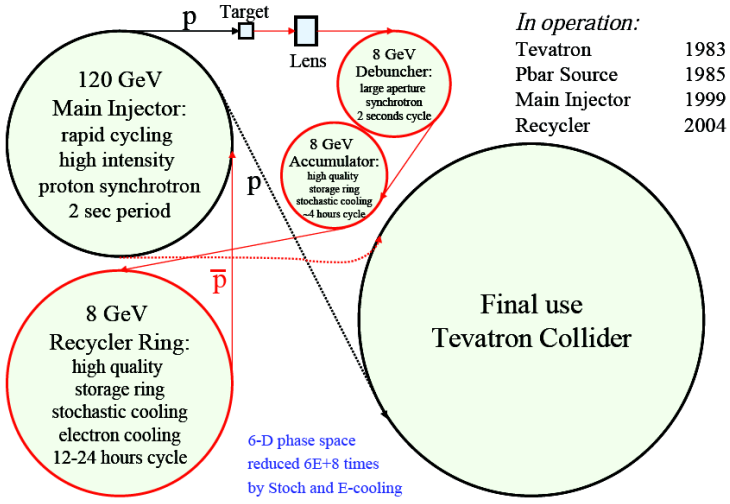


Figure 3.3: Antiproton production complex.

The Recycler ring

The Recycler(47) is a 3.3 Km-long storage ring of fixed 8 GeV kinetic energy, and is located directly above the Main Injector. It is composed solely by permanent gradient magnets and quadrupoles. Three main missions were designed for the Recycler operations: first, it allows antiprotons left over at the end of Tevatron Collider stores to be re-cooled and re-used; secondly, since the antiproton production rate decreases as the beam current in the Accumulator ring rises, the Recycler is designed to act as a post-Accumulator cooler ring, allowing the Accumulator to operate optimally. Finally, permanent magnets were chosen in the construction of the Recycler construction in order to dramatically reduce the probability of unexpected losses of antiprotons. In fact, the ring has been designed so that Fermilab-wide power could be lost for an hour with the antiproton beam surviving.

Recycling the antiprotons left over after the end of Tevatron collisions is a bit involved, since the antiprotons are at 1 TeV. The procedure for accomplish this task takes time and cause huge losses in the remaining number of $\bar{p}$. For these reasons the Recycler is not used anymore for this purpose. Instead the Recycler takes up the role of the Accumulator as the final storage for 8 GeV antiprotons, allowing the existing Antiproton Source to perform more efficiently and to produce antiprotons with higher rate.

3.1.4 The Tevatron ring

The Tevatron is the last stage of the Fermilab accelerator chain. The Tevatron is a 1 km radius synchrotron able to accelerate the incoming 150 GeV beams from Main Injector to 980 GeV, providing a center of mass energy of 1.96 TeV. The accelerator employs superconducting magnets throughout, requiring cryogenic cooling and consequently a large scale production and distribution of liquid helium. The Tevatron operates at the 36×36 mode, which refers to the number of bunches in each beam.

The antiprotons are injected after the protons have already been loaded. Just before the antiproton injection a set of electrostatic separators are used to create a pair of non-intersecting helical closed orbits. When the Tevatron loading is complete, the beams are accelerated to the maximum energy and collisions begin. There are 72 regions along the ring where the bunch crossing occurs. While 70 of these are parasitic, in the vicinity of CDF and DØ detectors additional focusing and beam steering is performed, to maximize the chance the proton strikes an antiproton. The focusing, driven by quadrupole magnets, reduces the beam spot size and thus increases the *luminosity*. The instantaneous luminosity, a quantity proportional to number of collisions per unit time, is given approximatively by:

$$\mathcal{L} = \frac{N_B N_{\bar{p}} N_p f}{2\pi(\sigma_p^2 + \sigma_{\bar{p}}^2)}, \tag{3.1}$$

where N_B is the number of bunches, $N_{(\bar{p}),p}$ is the number of the (anti)protons per bunch, f is the revolution frequency, and $\sigma^2_{(\bar{p})p}$ is the effective width of the (anti)proton beam. Clearly, the smaller $\sigma^2_{(\bar{p})p}$ the larger the rate of collisions.

For the data taking period considered in this Thesis the instantaneous luminosity was in the range $0.1 \times 10^{32}\ cm^{-2}s^{-1}$ and $3 \times 10^{32}\ cm^{-2}s^{-1}$.

During collisions the instantaneous luminosity decreases in time as particles are lost and the beams begin to heat up. Meanwhile, new antiprotons are stored in the Accumulator. When the luminosity becomes too low (approximately after 15-20 hours) it becomes beneficial dumping the current store and start a new cycle. Table 3.1 summarizes the accelerator parameters.

Figure 3.4 shows the Tevatron peak luminosity as a function of the time. The blue squares show the peak luminosity at the beginning of each store. The red triangle displays a point representing the last 20 peak values averaged together. Continous improvements in the accelerator complex led to the rapid increase of the initial instantaneous luminosity and, more important, to the increase of the rate of integrated luminosity delivered to the experiments. This can be seen in Figure 3.5, that reports the weekly and total integrated luminosity to date as function of the time.

Parameter	Value
Particles collided	$p\bar{p}$
Maximum beam energy	0.980 TeV
Time between collisions	0.396 μs
Crossing angle	0 μrad
Energy spread	0.14×10^{-3}
Bunch length	57 cm
Beam radius	39μm for p, 31μm for $\bar{p}$
Filling time	30 min
Injection energy	0.15 TeV
Particles per bunch	30^{10} for p; 9.7×10^{10} for $\bar{p}$
Bunches per ring per species	36
Average beam current	82μA for p, 27μA for $\bar{p}$
Circumference	6.12 Km
$\bar{p}$ source accumulation rate	$25 \times 10^9/hr$
Max number of $\bar{p}$ in accumulation ring	2.4×10^{12}

Table 3.1: Accelerator parameters for Run II configuration.

Figure 3.6 shows the total Tevatron luminosity delivered compared to the total luminosity recorded by the experiments (in the Figure CDF as an example) as a function of the time. An average efficiency of $\sim$ 85% is reached by the experiments in collecting the delivered luminosity, as explicitly shown in Figure 3.7. About 5% of inefficiency arises from trigger dead-time. In fact in order to maximize the physics of the experiment the trigger is run at its limit, where some dead-time is unavoidable. Another 5% comes from beam conditions where losses are too high and do not allow to operate the detector properly. The last 5% is either from small detector problems or operational decision to dedicate part of a store to detector studies.

3.2 The CDF detector

The Collider Detector at Fermilab (CDF) (48; 49) is a general-purpose detector designed to detect particles produced from the Tevatron $p\bar{p}$ collisions. It is located at one of the two intercation points along the Tevatron collider ring as shown in Figure 3.1. As illustrated in Figure 3.8, the detector has a cylindrical layout centered on the accelerator beamline. Tracking detectors are installed in the region directly around the interaction point to reconstruct charged-particle trajectories inside a 1.4 T uniform magnetic field (along the proton beam direction). The field is produced by a 5 m long superconducting solenoid located at the outer radius of the tracking region (1.5 m). Calorimeter modules are arranged in a projective tower geometry around the outside of the solenoid to provide energy measurements for both charged and neutral particles. The outermost part of the detector consists of a series of drift chambers used to detect muons, which are minimum-ionizing particles that typically pass through the calorimeter with almost no interactions.

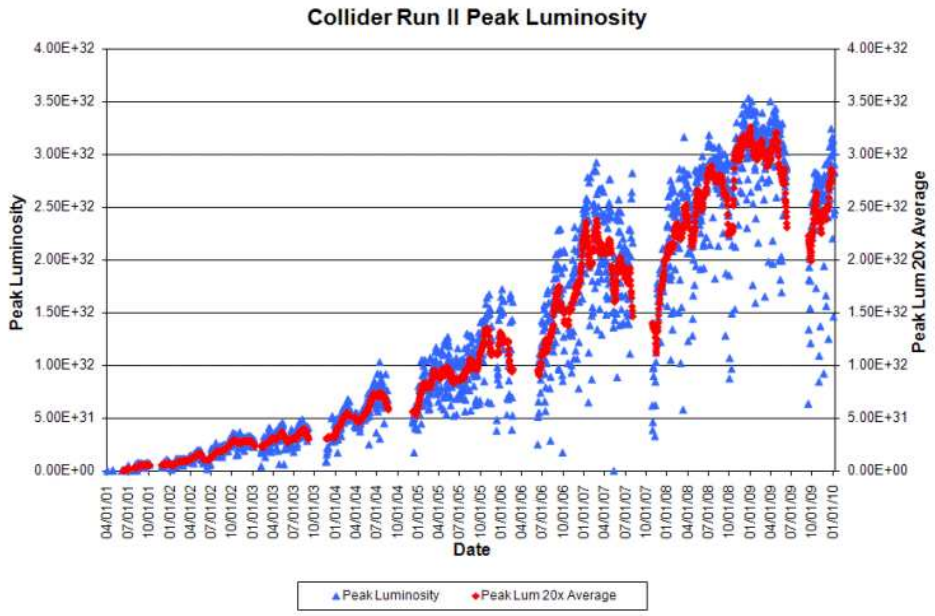


Figure 3.4: CDF initial instantaneous luminosity as function of the time.

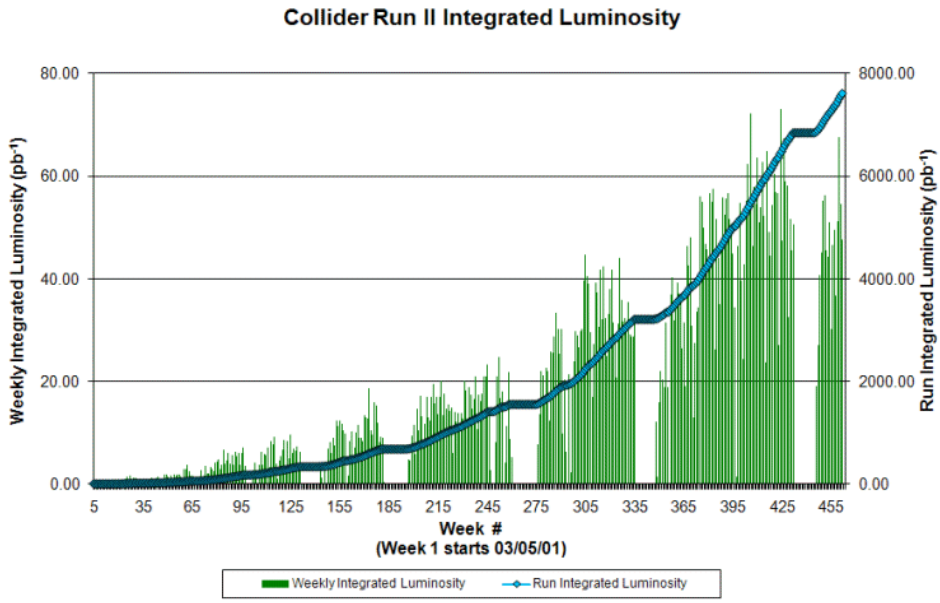


Figure 3.5: Weekly and total integrated luminosity delivered by the Tevatron collider to the CDF and DØ experiments.

3.2.1 CDF Coordinate system

CDF uses a Cartesian coordinate system centered in the nominal point of interaction, with the z axis coincident with the beamline and oriented parallel to the motion of the

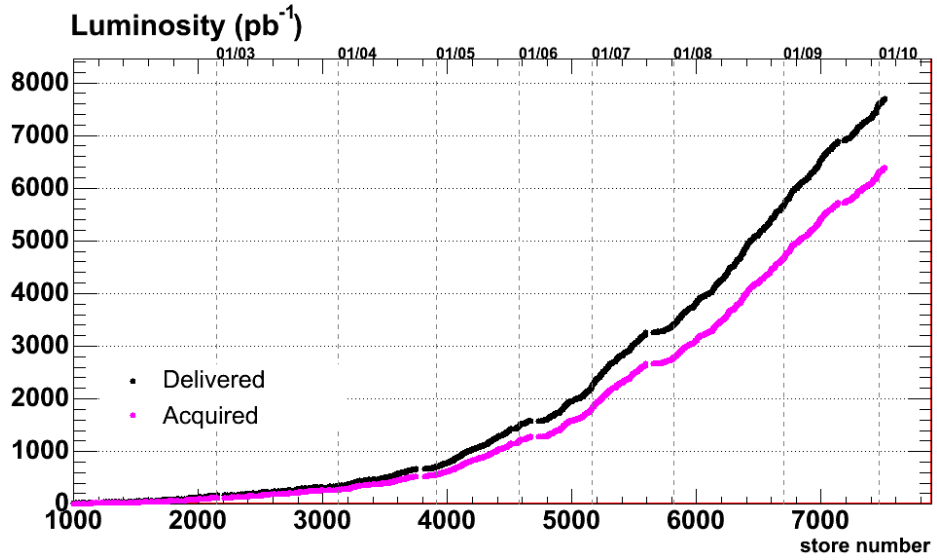


Figure 3.6: Total integrated luminosity delivered by the Tevatron collider as a function of the time. The acquired integrated luminosity by the CDF detector is also shown.

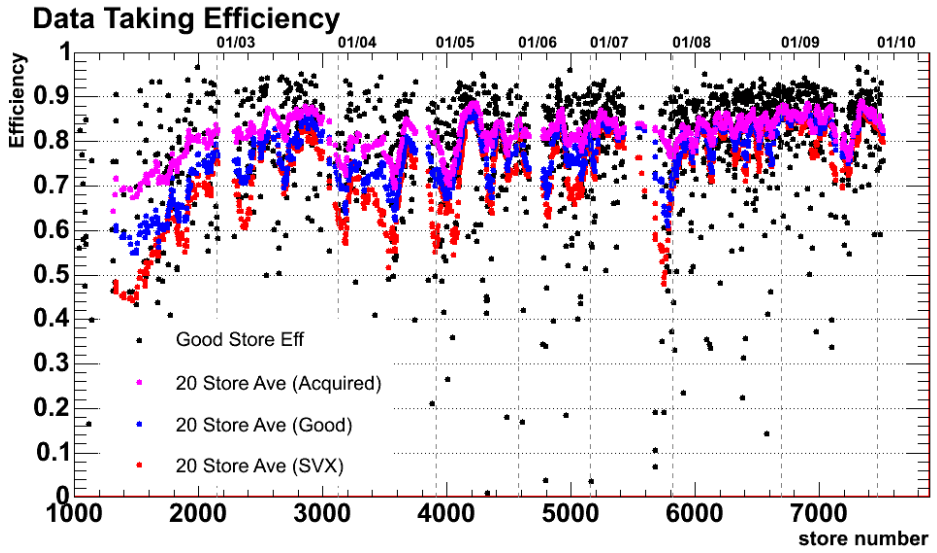


Figure 3.7: Data taking efficiency as function of the time (store number).

proton beam. The x axis is in the horizontal plane of the accelerator ring, pointing radially outward, while the y axis points vertically up (see Figure 3.9).

For the simmetry of the detector, it is often convenient to work with cylindrical (z, r, ϕ) or polar (r, θ, ϕ) coordinates. The azimuthal angle ϕ is measured in the $x-y$ plane starting from the x axis, and it is defined positive in the anti-clockwise direction; on the other side, the polar angle θ is measured from the positive direction of the z axis. The coordinate r defines the transverse distance from the z axis. Another important coordinate that can be used instead of the polar angle θ, is called *pseudorapidity* and it

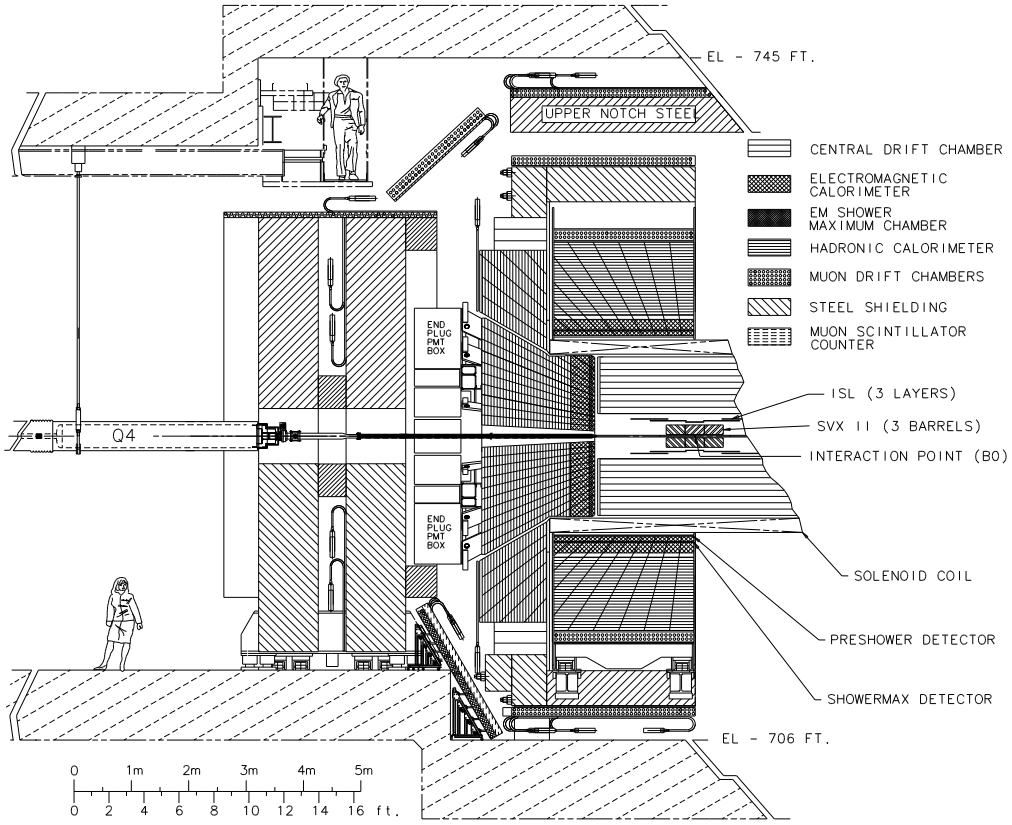


Figure 3.8: Elevation view of half of the CDF Run II detector.

is defined as:

$$\eta = -\log \tan \frac{\theta}{2} \tag{3.2}$$

In some cases we also use a detector pseudorapidity variable, η_{det}, to refer to fixed locations within the detector. This variable is based on the standard definition of pseudorapidity given above where the angle θ is redefined in the context of a fixed location as $\theta = \arctan(\frac{r}{z_{det}})$, where z_{det} is the distance from the detector center in the direction along the beam axis.

The pseudorapidity is usually preferred to θ at hadron colliders, where events are boosted along the beamline, since it transforms linearly under Lorentz boosts, i.e. η intervals are invariant with respect to boosts. For these reasons, the detector components are chosen to be as uniformly segmented as possible along η and ϕ coordinates.

3.2.2 Tracking system

The inner part of the CDF II is devoted to tracking systems, whose volume is permeated by an uniform magnetic field of magnitude $B = 1.4\,T$, oriented along the z-axis. This feature constrains charged particles to an helicoidal trajectory by means of the Lorentz force, whose radius, measured in the transverse plane $(x - y)$ is directly related to the particles transverse momentum, p_T.

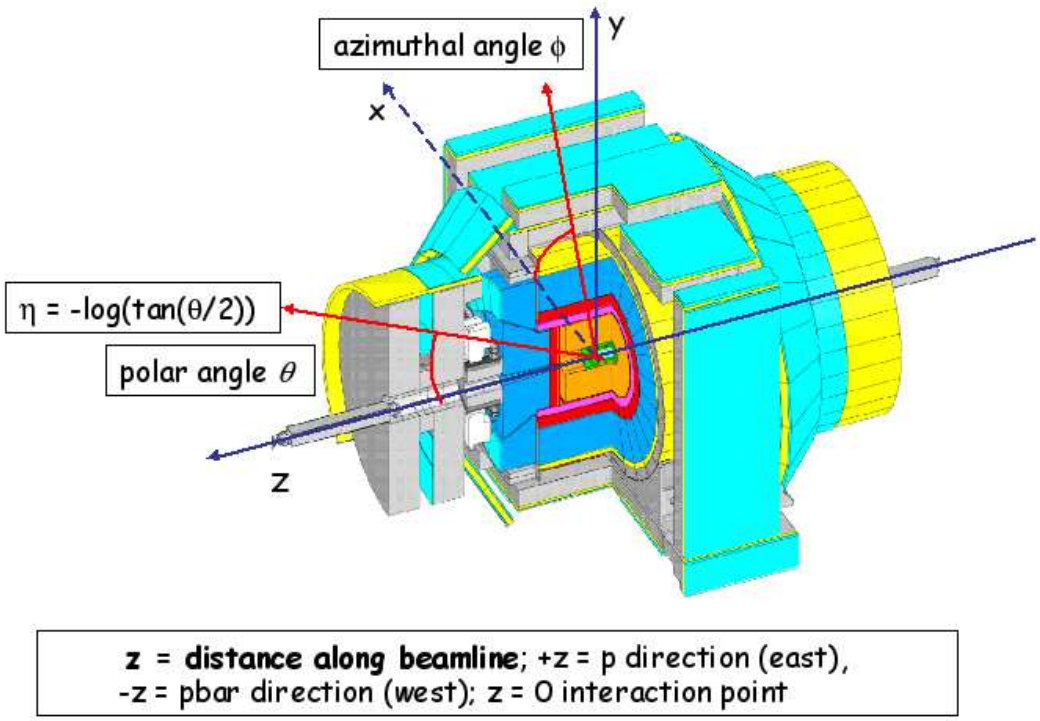


Figure 3.9: Isometric view of the CDF II Detector and its coordinate system.

Particle trajectories can be completely described by five parameters:

- z_0 : the z coordinate of the closest point to the z axis;
- d_0 : the impact parameter defined as the distance between the point of closest approach to z axis and the z axis;
- ϕ_0 : the ϕ direction of the transverse momentum of the particle (tangential to the helix) at the point of the closest approach to the z axis;
- $\cot\theta$: the helix pitch, defined as the ratio of the helix step to its parameter;
- C : the helix curvature.

Actually, the impact parameter and the curvature are signed quantities defined by :

$$C = \frac{q}{2R}, \tag{3.3}$$

$$d_0 = q(\sqrt{x_c^2 + y_c^2} - R), \tag{3.4}$$

where q is the charge of the particle, $(x_c^2 + y_c^2)$ is the center of the helix as projected onto the x-y plane and R is its radius. A graphical view of these variables together with the ϕ_0 is shown in Figure 3.10.

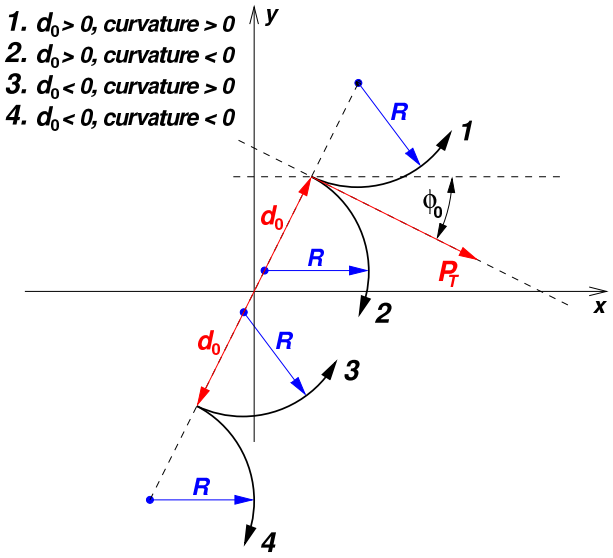


Figure 3.10: Illustration of helix track parametrization.

From helix parameters one can easily derive particle transverse and longitudinal momenta as:

$$P_T = \frac{cB}{2|C|}, \tag{3.5}$$

$$P_z = P_T \cot\theta. \tag{3.6}$$

CDF inner tracking system consists of three silicon detectors responsible for high precision measurements and a drift chamber devoted to add further information for track reconstruction; going in more detail and starting from the interaction point we find:

- **Layer 00 (L00)**
  A single sided silicon micro-strip detector located immediately outside the beam pipe, at a radius of approximately 1.6 cm and covering $|\eta| \leq 4.0$ (Figure 3.11) (50). It provides more precise track measurements and better b quark tagging efficiency compared to the other tracking devices.

- **Silicon Vertex Detector (SVXII)**
  A double sided silicon micro-strip detector, located outside *L00*, extending from r = 2.1 cm to r = 17.3 cm and covering $|\eta| \leq 2.0$ (Figure 3.12). It consists of 5 layers and its strips are aligned axially to the beam on one side, while in the other side are aligned with a small (1.2°) angle stereo (layers 2 and 4) and with a 90-degree stereo (layers 0, 1 and 3). It provides high precision tracking and secondary vertex detection.

- **Intermediate Silicon Layer (ISL)**
  A double sided silicon micro-strip detector, with axial strips on one side and small angle stereo strips on the other side. It consists of three layers, positioned at different radii (Figure 3.13): central layer is at r = 22 cm while forward and backward layers

are respectively at $r = 20$ cm and $r = 28$ cm. They also have different η coverage: $|\eta| \leq 1.0$ for central layer and $1.0 \leq |\eta| \leq 2.0$ for the others.

- **Central Outer Tracker (COT)**
  An open cell drift chamber with argon-ethane gas in a 50/50 mixture. It's located outside *SVX* from $r = 40$ cm to $r = 137$ cm, covering $|\eta| \leq 1.0$, thus providing tracking in the central regions of the detector (Figure 3.14). Its 2520 cells are divided into 8 super-layers, each containing twelve layers of sense wires. The odd super-layers have wires parallel to the beam (*axial* super-layers) while the even have wires at a small (2.0°) stereo angle (*stereo* super-layers).

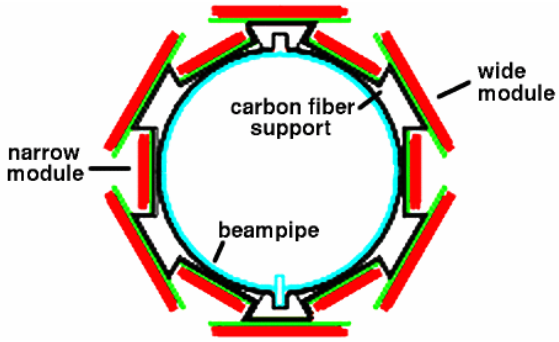


Figure 3.11: Schematic view of *L00*

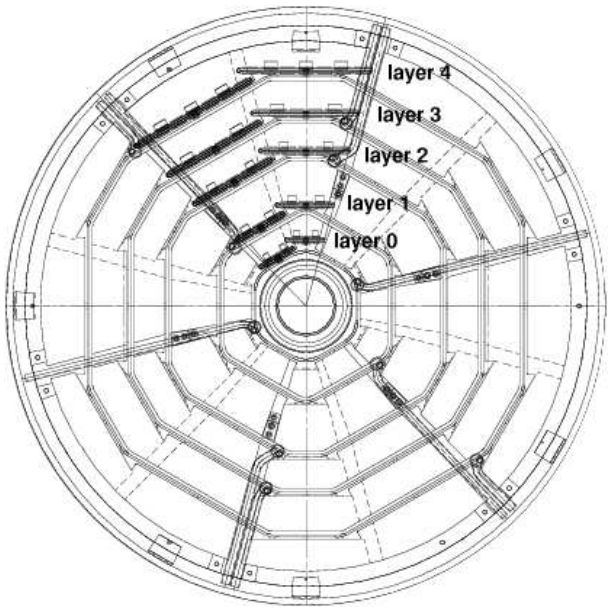


Figure 3.12: Schematic view of *SVXII*

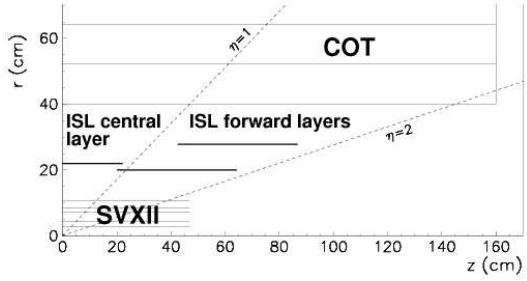


Figure 3.13: *ISL* position.

3.2.3 Time-of-flight detector

Located outside the COT (at $r = 140$ cm), this sub-detector provides time-of-flight information to improve particle identification capabilities in the central detector (especially for K-π discrimination) (51). It consists of 216 scintillator bars, each running the length of COT and arranged cilindrically around it, with a PMT at each end of each bar.

3.2.4 Calorimeteric systems

The CDF calorimetry system has been designed to measure energy and direction of neutral and charged particles leaving the tracking region.

Particles hitting the calorimeter can be divided into two classes according to their interaction with matter: electromagnetically interacting particles, such as electrons and photons, and hadronically interacting particles, such as mesons or baryons produced in hadronization processes. To detect these two classes of particles, two different calorimetric parts have been developed: an inner electromagnetic and an outer hadronic section, providing coverage up to $|\eta| < 3.6$. The calorimeter is also segmented in $\eta - \phi$ sections, called

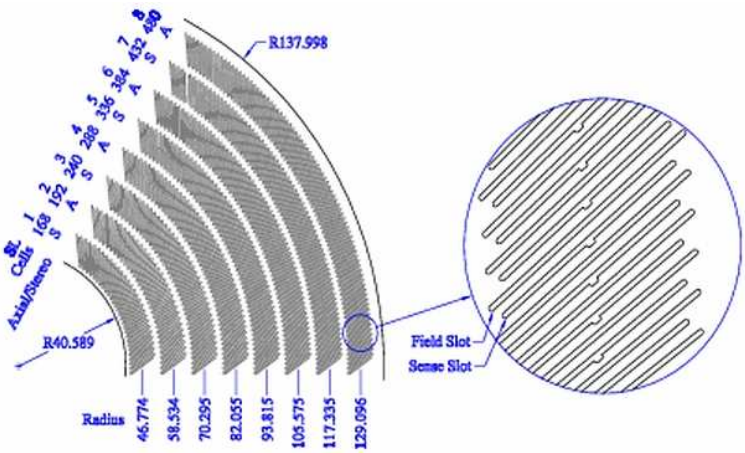


Figure 3.14: COT section: the eight superlayers (left) and the alternation of field plates and wire planes (right)

	CEM	CHA	WHA	PEM	PHA
η coverage	< 1.1	< 0.9	$0.7 < \|\eta\| < 1.3$	$1.3 < \|\eta\| < 3.6$	$1.3 < \|\eta\| < 3.6$
n. of modules	48	48	48	24	24
η towers/mod	10	8	6	12	10
n. of channels	956	768	676	960	864
Absorber (mm)	Pb (3.0)	Fe (25.4)	Fe (50.8)	Pb(4.6)	Fe (50.8)
Thickness	$19X_0, 1\Lambda_0$	$4.5\Lambda_0$	$4.5\Lambda_0$	$21X_0, 1\Lambda_0$	$7\Lambda_0$
Position res.	0.2×0.2	10×5	10×5		
Energy res.	$\frac{13.5\%}{\sqrt{E_T}} \oplus 1.7\%$	$\frac{75\%}{\sqrt{E_T}} \oplus 3\%$	$\frac{80\%}{\sqrt{E_T}}$	$\frac{16\%}{\sqrt{E_T}} \oplus 1\%$	$\frac{80\%}{\sqrt{E_T}} \oplus 5\%$

Figure 3.15: Geometry, parameters and performance summary of CDF Calorimetric System. The position resolution is given in $r \cdot \phi \times z$ cm^2 and is measured for a 50 GeV incident particle.

towers, projected towards the geometrical center of the detector, in order to supply information on particle positions. Each tower consists of alternating layers of passive material and scintillator tiles. The signal is read out via wavelength shifters (WLS) embedded in the scintillator and light from WLS is then carried by light guides to photomultiplier tubes.

We refer to the energy E of a given tower as the sum of the energy deposited in that tower of the electromagnetic calorimeter and the corresponding one in the hadronic calorimeter: $E = E_{EM} + E_{HAD}$. The transverse energy E_T is obtained as $E \cdot \sin\theta$, where θ is the polar angle for the given tower.

The calorimetric system is subdivided into three regions, central, wall and plug, in order of increasing pseudorapidity ranges, with the following naming convention: Central Electromagnetic (CEM), Central Hadronic (CHA), Wall Hadronic (WHA), Plug Electromagnetic (PEM) and Plug Hadronic (PHA); an inner commented view of the detector is shown in Figure 3.16. Table in Figure 3.15 summarizes the most important characteristics of each part of the calorimeter.

The Central Calorimeter

The Central Electro-Magnetic calorimeter (CEM) is segmented in $\Delta\eta \times \Delta\phi$=0.11×15° projective towers consisting of alternate layers of lead and scintillator, while the Central and End Wall Hadronic calorimeters (CHA and WHA respectively), whose geometric tower segmentation matches the CEM one, use iron layers as radiators. A perspective view of a central electromagnetic calorimeter module, a *wedge*, is shown in Figure 3.17.

Two position detectors are embedded in each wedge of the CEM:

- The Central Electromagnetic Strip chamber (CES) (see Figure 3.18) is a two-dimensional strip/wire chamber located at the radial distance 184 cm. It measures the charge deposition of the electromagnetic showers, providing information on their pulse-height and position with a finer azimuthal segmentation than the calorimeter towers. This results in an increased purity of electromagnetic object reconstruction.

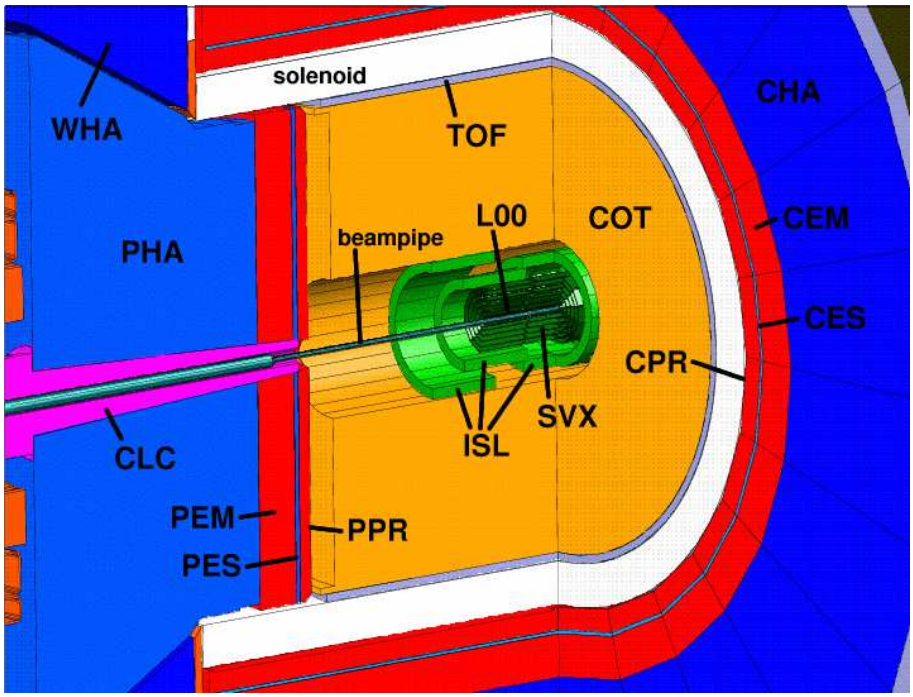


Figure 3.16: Schematic view of the inner parts of CDF detector.

- The Central Pre-Radiator (CPR) consists of two wire chamber modules placed immediately in front of the calorimeter. It acts as pre-shower detector and with its 3072 channels collects charge deposit by showers originated by interaction of particles with tracking system and solenoid material. It can help in discriminating pions from electron and photons, because the latter deposit a greater amount of energy in the chamber.

The plug calorimeter

The plug calorimeter, shown in Figure 3.19, covers the η region from 1.1 to 3.6. Both electromagnetic and hadronic sectors are divided in 12 concentric η regions, with $\Delta\eta$ ranging from 0.10 to 0.64 according to increasing pseudorapidity, each segmented in 48 or 24 (for $|\eta| < 2.1$ or $|\eta| > 2.1$ respectively) projective towers.

As in the central calorimeter, there is a front electromagnetic compartment and a rear hadronic compartment (PEM and PHA). Projective towers consist of alternating layers of absorbing material (lead and iron for electromagnetic and hadronic sectors respectively) and scintillator tiles. The first layer of the electromagnetic calorimeter acts as a pre-shower detector; to this scope, the first scintillator tile is thicker (10 mm instead of 6 mm) and made of a brighter material.

As in the central calorimeter, a shower maximum detector is also included in the plug electromagnetic calorimeter (PES). The PES consists of two layers of 200 scintillating bars each, oriented at crossed relative angles of 45^o ($\pm 22.5^o$ with respect to the radial

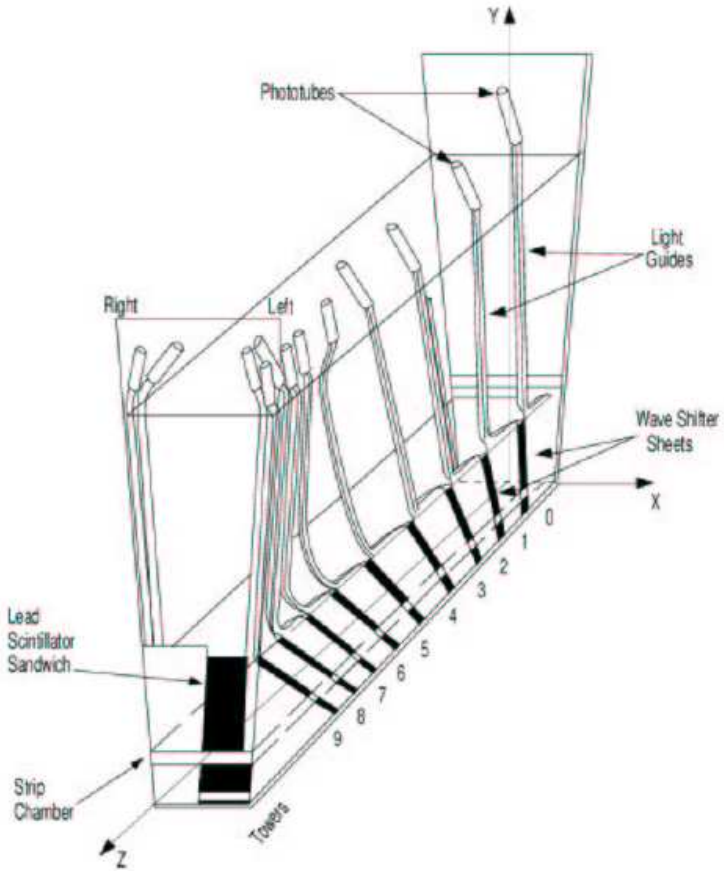


Figure 3.17: Perspective view of a CEM module.

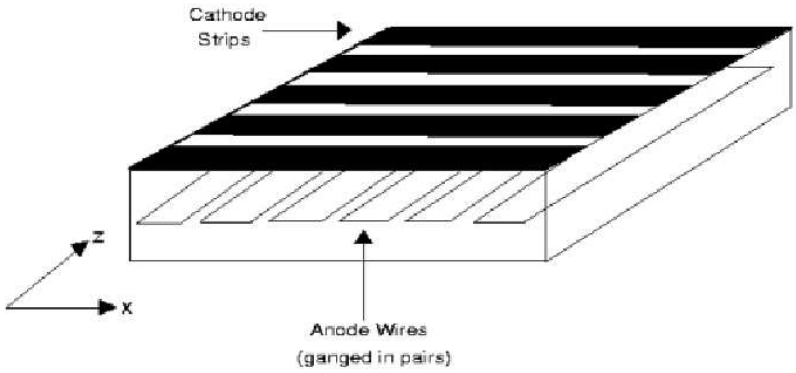


Figure 3.18: The CES detector in CEM. The cathode strips run in the x direction and the anode wires run in the z direction providing x and $(r \cdot \phi)$ measurements.

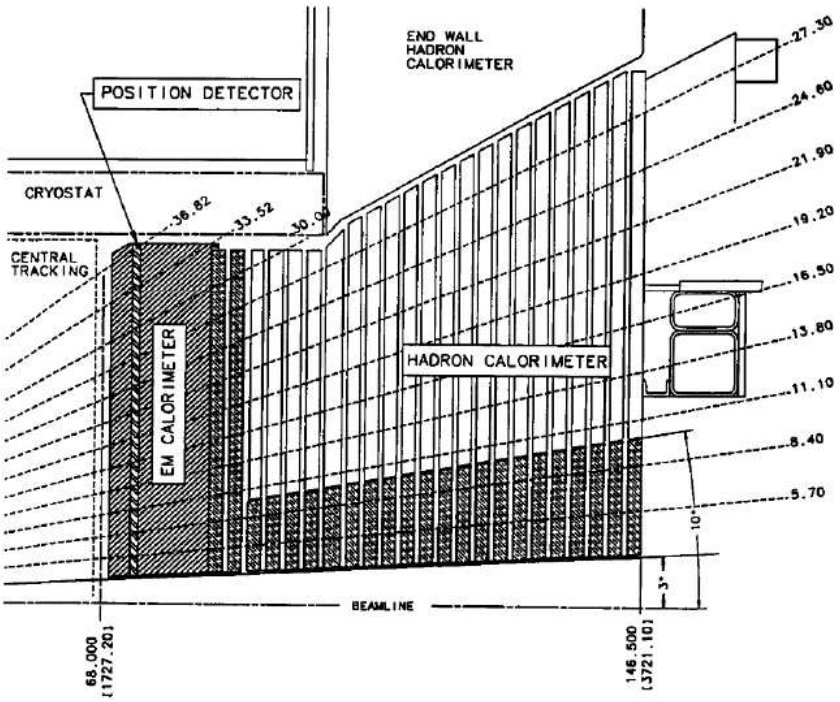


Figure 3.19: Plug Calorimeter (PEM and PHA) inserted in the Hadron End Wall calorimeter WHA and into the solenoid.

direction). The position of a shower on the transverse plane is measured with a resolution of $\sim 1\ mm$.

3.2.5 Muon detectors

The CDF muon detector system(52) consists of drift chanmbers and scintillator counters, altogether covering a rapidity range $|\eta| < 2$. The muon detectors are behind the most of the CDF detector material and many of them are further behind thick steel. The muon detectors geometry coverages are available in Figures 3.8 and 3.20.

The central muon detector (**CMU**) is located right outside the CHA behind 5.5 nuclear interaction lengths (λ_0) of detector material. It has a rapidity range of $|\eta| < 0.68$. It is a barrel with inner and outer radii $r_i = 347$ cm and $r_o = 396$ cm respectively, consiting of 4 drift tube layers sectioned by wedge matching the CHA towers: 3 sections of 4 tubes per layer per 15° wedge. Each tube operates in proportional mode, with a maximum drift time of $0.8\mu s$. The transtube multiple scattering resolution is $12/(p[\text{GeV}])$ cm and the longitudinal resolution is $\delta z \simeq 10$ cm.

The central muon upgrade detector (**CMP**) is located outside the CMU behind $7.8\lambda_0$ of detector material that includes additional 60 cm thick steel slabs. The CMP contains four layers of rectangularly arrayed drift tubes. The rapidity extension of the CMP detector is $|\eta| < 0.68$. The CMP gas operation mode is proportional. The maximum drift time is $1.4\mu s$ and it has a transtube multiple scattering resolution of $15/(p[\text{GeV}])$ cm.

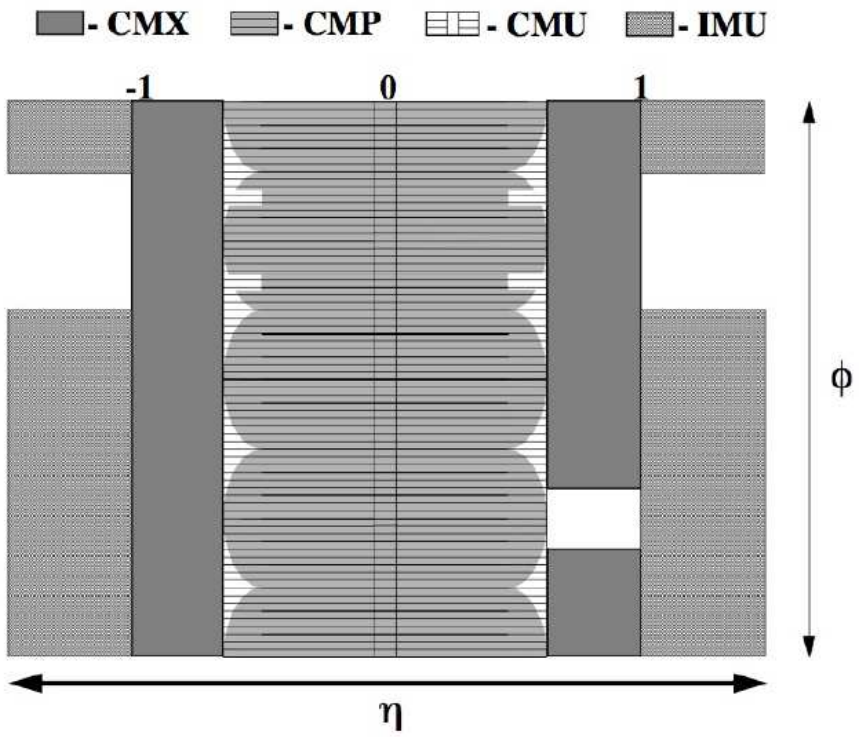


Figure 3.20: $\eta - \phi$ coverage of the muon detector system. The shape is irregular because of the obstruction by systems such as cryo pipes or structural elements.

A layer of scintillators (**CSP**) is mounted onto the outside surface of the CMP to provide timing information with a resolution of $1 - 2$ ns.

The central muon extension detector (**CMX**) consists of conical sections facing toward the interaction point behind $6.2\lambda_0$ of detector material, extending the central muon detector rapidity coverage $0.65 < |\eta| < 1.0$, excluding the east top 30° in azimuth. The CMX contains two folds of 4 layers of rectangular drift tubes. The transtube multiple scattering resolution is $13/(p[\text{GeV}])$ cm and the longitudinal position resolution is $\delta z \simeq 14$ cm.

Two layers of scintillators **CSX** are mounted with one layer to the upper surface and the other to the lower surface of the CMX to provide timing information for the CMX.

The intermediate muon detector (**IMU**) is built behind $6.2 - 20\lambda_0$ of material, depending on the rapidity. It consists of two barrels extending the CDF geometric muon acceptance in the range $1.0 < |\eta| < 1.5$. It contains four layers of proportional drift tubes (BMU), with a maximum drift time of $0.8\mu s$. The transtube multiple scattering resolution is $13 - 25/(p[\text{GeV}])$ cm and the longitudinal position resolution is $\delta z \simeq 16.5$ cm.

Three layers of scintillators (**BSU-F**, **BSU-R** and **TSU**) are mounted outside the BMU and provide timing information.

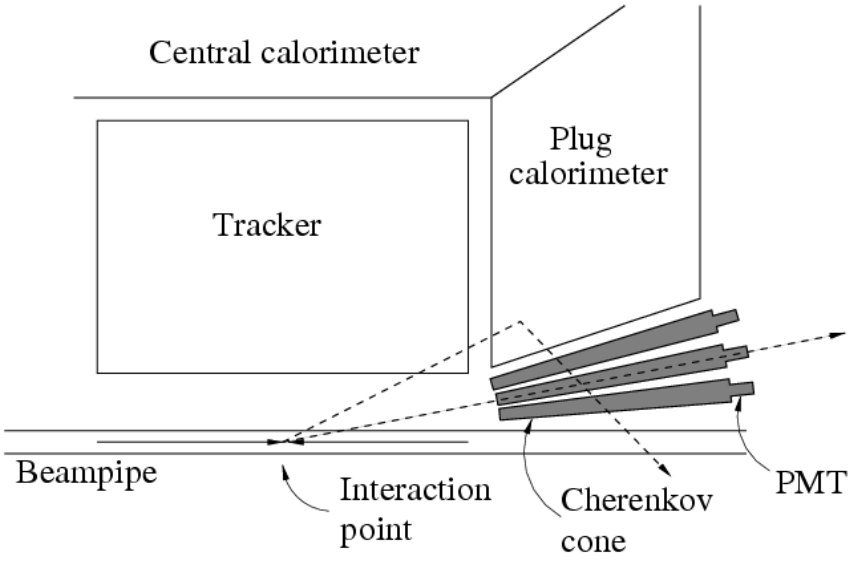


Figure 3.21: Schematic view of the luminosity monitor inside a quadrant of CDF.

3.2.6 Cherenkov luminosity counters

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 (53). They are located inside the endplug calorimeters, in the forward and backward regions ($3.7 < |\eta| < 4.7$). Each module consists 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.

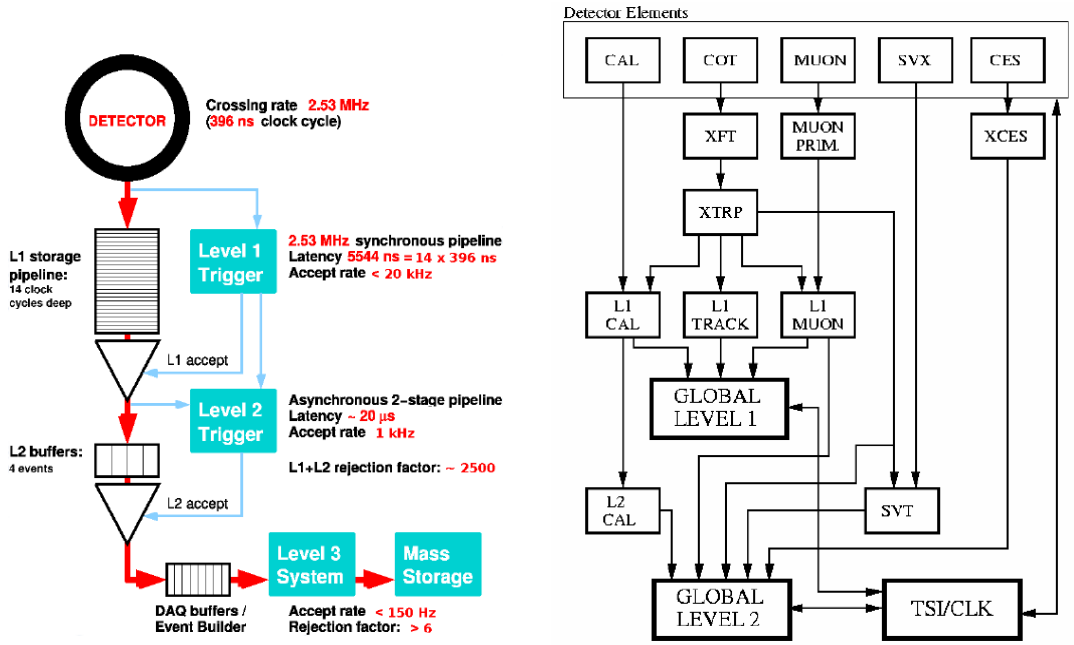
The counters measure the average number of interactions per bunch crossing μ, which is used to provide a measurement of the instantaneous luminosity $\mathcal{L}$:

$$\mu \cdot f_{bc} = \sigma_{p\bar{p}} \cdot \mathcal{L}, \tag{3.7}$$

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 method measures the luminosity with about the 6% systematic uncertainty.

3.3 CDF Trigger and data acquisition system

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 order $10^{32}\ cm^{-2}s^{-1}$, results in a collision rate of about 6 MHz, while the tape writing speed is only of ~ 100 events per second. The role of the trigger is to efficiently select the most interesting



(a) CDF readout functional diagram. (b) Block diagram of the CDF trigger system.

Figure 3.22: CDF trigger system.

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 minimal deadtime (see Figure 3.22). The front-end electronics of all detectors is interfaced to a syncronous pipeline where up to 42 subsequent events can be stored for 5.5 μs while the hardware is taking a decision. Level 1 (L1) always occurs at a fixed time < 4 μs 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 μs: 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 ~ 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 signal one wants to isolate, specific sets of requirements are established by exploiting the physics objects (*primitives*) available for each trigger level. Successively, links across different levels are established by defining *trigger paths*: a trigger path identifies a unique combination of a L1, a L2, and a L3 trigger; data sets (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 data sample such trigger paths are *prescaled* by a factor N, meaning that just one event out of N is accepted among the ones that pass trigger selections. When the prescale factor is fixed the prescale is called *static*. Instead the prescale is called *dynamic* when it is varied during the data taking to exploit at the most the available bandwith.

3.3.1 Level 1 primitives

Tracks

The most significant tool for L1 trigger is the possibility of track finding by means of a hardwired algorithm named e**X**tremely **F**ast **T**racker (XFT). The XFT has been designed to work with COT signals at high collision rates, returning track p_T and ϕ_0 by means of a fast r-ϕ 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.

Calorimetric primitives

At L1 calorimetric towers are merged in pairs along η 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 event ΣE_T and $\vec{\not{E}}_T$.

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.

3.3.2 Level 2 primitives

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(48). L2 clusters can be used to build object triggers by applying a cut on their transverse energy and position (provided from η-ϕ address of the seed towers), and global triggers by selecting on the number and $\sum E_T$ of clusters.

SVT tracks

The Silicon Vertex Tracker (SVT) (54) exploits the potential of a high precision silicon vertex detector to trigger on tracks with large impact parameter.

The architecture of SVT is shown in Figure 3.23. 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 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 essentially unchanged with respect to L1, the only difference consists in an improved ϕ-matching (within 1.25^o) between XFT tracks and track segments (*stubs*) formed with hits in the muon chambers. In the case of electrons, a finer ϕ-matching can be instead performed at L2 thanks to the information from central and plug shower maximum detectors.

3.3.3 Level 3 primitives

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

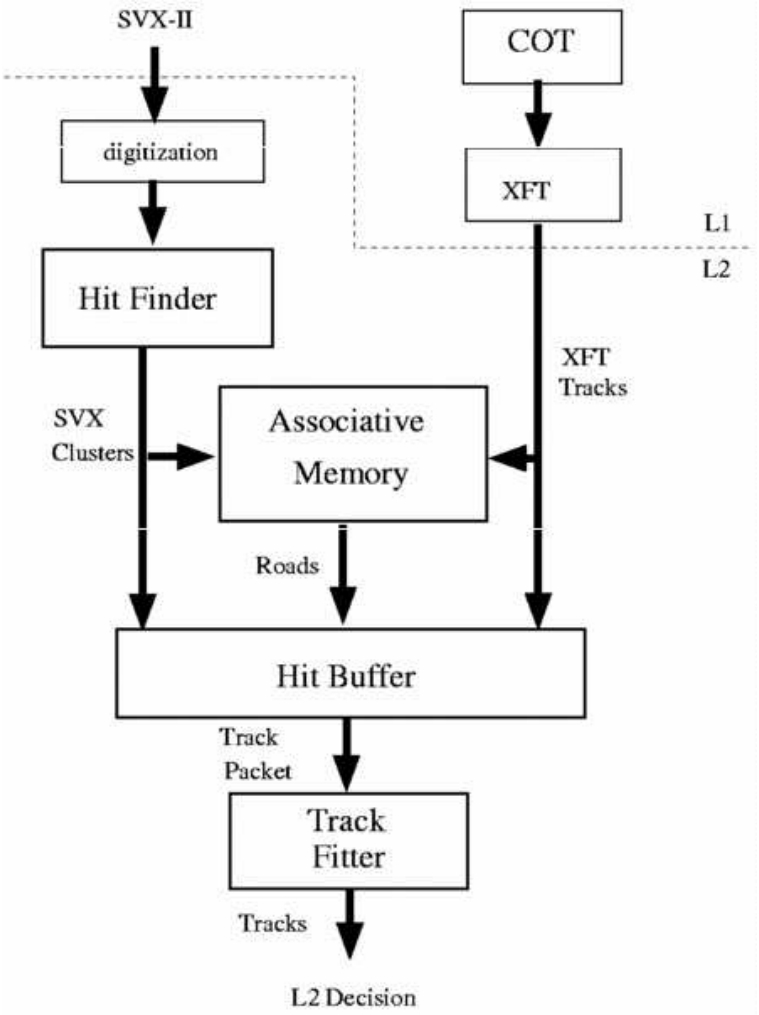


Figure 3.23: The SVT architecture.

are permanently stored on tapes for further offline analisys. Each stored event is about 250 kB large on tape. Further offline processing is then performed on the selected events.

3.3.4 Trigger Upgrades

CDF has recently undergone two major trigger upgrades 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 (55; 56).

XFT upgrade 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 (σ_{z_0} = 11 cm).

The upgraded L2CAL system uses a *fixed cone* cluster finding algorithm which prevents fake cluster formation and exploits full 10-bit trigger tower energy information for $\vec{\not{E}}_T$ and

ΣE_T calculation (the old system, due to hardware limitations, used only 8-bit tower information). A jet is formed 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.

3.4 Offline data processing

The raw data flow from L3 triggers, segmented into streams according to trigger sets tuned to a specific physics process, is then stored on fast-access disks in real time (on-line), as the data are collected. All other manipulations with data are referred to as off-line data handling. The most important of these operations is the so-called "production" which stands for the complete reconstruction of the collected data. At this stage raw data are unpacked and physics objects suitable for analysis, such as tracks, vertices, leptons and jets are generated. The procedure is similar to what is done at L3, except that it is done in a much more elaborate fashion, applying the most up-to-date detector calibrations, using the best measured beamlines, etc. The output of the production is further categorized into data sets which are used as input to physics analyses.

Homogeneous data are already grouped during online acquisition in *run numbers*. Offline we then group several run numbers in *run periods*, each one with an integrated luminosity of the order of $\sim 100\text{pb}^{-1}$. Table 3.2 reports the conventional classification of the data acquired up to date. Data used in this Thesis has been collected through March 21st 2009 (Run Period 23 included).

3.5 High-p_T objects identification

The main physics processes of interest for our searches were described in section 2.4. They involve leptons and quarks in the final state of the hard scattering, in addition to gluons and photons that can be produced by both initial and final state radiations.

Lepton reconstruction depends on the type of lepton and on its direction inside the detector. The subdetectors have different segmentation resulting in a different resolution that combined with a non-homogeneous background distribution force to use specific reconstruction algorithms. We optimized the identification selections for leptons with $p_T > 10$ GeV. This requirement, or a tighter one, is always applied. Identification criteria for electrons, muons and neutrinos are described in Sections 3.5.1, 3.5.2 and 3.5.4; a brief summary of lepton categories used in this analysis is also reported for convenience in Table 3.3.

Tau leptons are classified depending of their decay mode. Leptonic decays are identified using electrons or muons and used in this analysis via their standard reconstruction. Hadronic decays of τs are handled separately(57), but they are not considered in this Thesis.

Period	Run numbers	Online dates	Luminosity [pb^{-1}]	Cumulative luminosity[pb^{-1}]
26	282976-284843	15 Sep 09 - 25-Oct-09	189	5971
25	275873-277511	05 May 09 - 13-Jun-09	236	5782
24	274123-275848	22 Mar 09 - 04-May-09	283	5546
Data sample used in this Thesis:				
23	272470-274055	15 Feb 09 - 21-Mar-09	232	5263
22	271072-272214	2 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	5 Dec 07 - 27 Jan 08	159	2939
14	252836-254683	28 Oct 07 - 3 Dec 07	44.5	2780
13	241665-246231	13 May 07 - 4 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 3.2: Data acquired by the CDF detector. The table shows the conventionally attributed run period, the run number ranges grouped into a given run period, the starting and ending dates when data were collected and the integrated luminosity collected in each run period.

LBE	Central electron with $\|\eta_{det}\| \leq 1.1$ identified with a Likelihood-based algorithm
PHX	Forward electron relying on Silicon tracking with $1.2 \leq \|\eta_{det}\| \leq 2.0$
CMUP	Central muon with hits in both CMU and CMP chambers: $\|\eta_{det}\| \leq 0.6$
CMX	Central muon with hits in the CMX Arches muon chambers: $0.65 \leq \|\eta_{det}\| \leq 1.0$
CMU	Central muon with hits in the CMU chamber pointing to regions not covered by CMP chambers.**Not currently used for technical reasons.**
CMP	Central muon with hits in the CMP chamber pointing to regions not covered by CMU chambers.
CMXMsKs	Central muon with hits in the CMX Miniskirts or Keystone detectors.
BMU	Forward muon with hits in the IMU detector: $1.1 \leq \|\eta_{det}\| \leq 1.5$.
CMIOCES	Central muons which does not satisfy muon chambers hits requirement but is identified using calorimetric energy selections.
CMIOPES	Forward muons which does not satisfy muon chambers hits requirement but is identified using calorimetric energy selections.
CrkTrk	High-p_T isolated track pointing to not fully instrumented regions of the detector.It is assumed to be either an electron or a muon.

Table 3.3: Summary of lepton types (categories) used in this Thesis

.

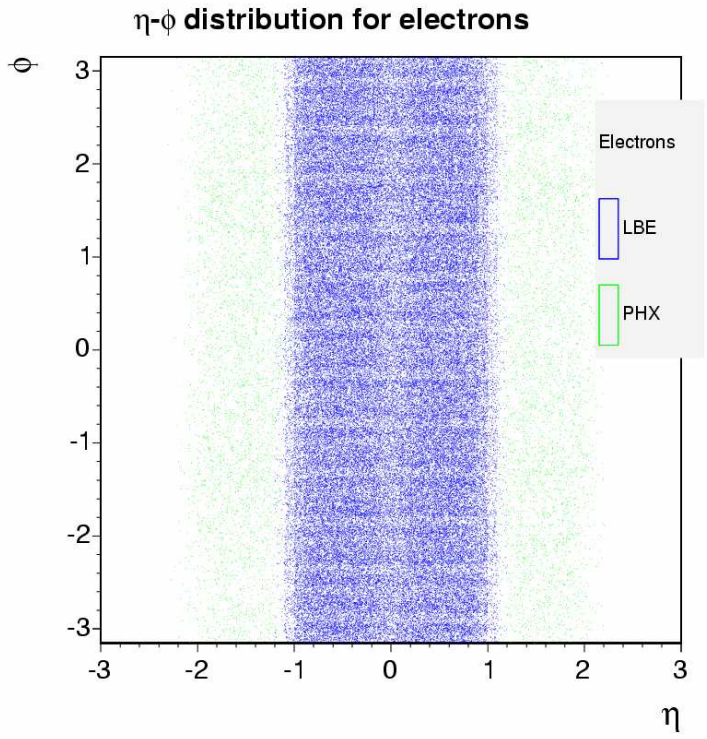


Figure 3.24: Distribution in the $\eta - \phi$ plane of a sub-sample of the reconstructed central (LBE) and forward (PHX) electrons in data.

Quarks and gluons are detected via jets; the jets reconstruction algorithm is described in Section 3.5.3.

For the studies performed in this thesis, an explicit identification of photons is not required, and will not be discussed.

3.5.1 Electron identification

Electrons are identified by requiring a track matched to an energy cluster in the calorimeter with an appropriate shower profile. In order to reconstruct the particle 4-momentum $(E, \vec{p})$, electrons are assumed massless particles. Track information are used to set the three dimensional direction $\frac{\vec{p}}{|\vec{p}|}$, while the calorimetric energy measurement determines the magnitude $E \equiv |\vec{p}|$.

We distinguish two electron categories:

- **LBE**: Central ($|\eta| \lesssim 1$) Electrons,
- **PHX**: Forward ($1 \lesssim |\eta| \lesssim 2$) Electrons.

Figure 3.24 shows a scatter plot of a small subsample of reconstructed electrons in our data sample as function of track η and ϕ.

	LBE
Region	central
E_{HAD}/E_{EM}	≤ 0.125
Track p_T	≥ 10 GeV (5 GeV if $E_T < 20$ GeV)
Track $\lvert z_0 \rvert$	≤ 60 cm
$CalIso$	≤ 0.3
Conversion	false
Track	Beam constrained
Likelihood ($\mathcal{L}$)	$\mathcal{L} > 0.99$

Table 3.4: Identification criteria for Central Electrons

Central Electrons: Likelihood-Based Electrons (LBE)

The electron cluster is made from an electromagnetic (EM) seed tower and at most one additional tower that is adjacent to the seed tower in η_{det} and within the same ϕ wedge. The seed tower must have $E_T > 2$ GeV and a reconstructed track which extrapolates to that tower. The hadronic energy E_{HAD} in the corresponding towers of hadronic calorimeter is required to be less than 0.125 times the electromagnetic energy E_{EM} of the cluster. Both the track and the cluster are required to lie within the well-instrumented regions of the central electromagnetic calorimeter (CEM).

The calorimetric isolation is defined using the energy of the electron cluster (E_T^e):

$$CalIso \equiv \frac{E_T^{\Delta R=0.4} - E_T^e}{E_T^e}, \tag{3.8}$$

where $E_T^{\Delta R=0.4}$ is the energy in a cone of radius $\Delta R = 0.4$ around the electron cluster. At this level $CalIso < 0.3$ is required.

The electron track is reconstructed constraining its origin in the $x - y$ plane to the beam position. It is required to have a z position at the closest approach to the beamline $|z_0| < 60$ cm. The transverse momentum p_T has to be greater than 10 GeV(5 GeV) for candidate electrons with $E_T > 20$ GeV ($10 < E_T < 20$ GeV); the different thresholds are due to the trigger requirements. The conversion veto is applied to reject electrons coming from processes like $e^{\pm}\gamma \to e^{\pm}e^{\pm}e^{\mp}$ or $\gamma \to e^{\pm}e^{\mp}$.

These requirements are summarized in Table 3.4.

In order to enhance the fake electron rejection maintaining an high efficiency, we develop a likelihood to combine additional information in a single discriminant constructed starting from the following quantities (*identification variables*):

- E_{HAD}/E_{EM}: ratio of the energy deposited in the hadronic calorimeter to the electromagnetic calorimeter by the electron. Real electrons are expected to deposit most of their energy in the EM calorimeter.

- E/p: ratio of the EM cluster transverse energy to the track transverse momentum as measured by the tracking system. This is expected to be near ~ 1 for real electrons.

- L_{shr}: the lateral shower profile compares the distribution of the EM towers energies adjacent to the seed tower as a function of its energy to the derived distribution from electron test-beam data.

- $CalIso$: calorimetric isolation as defined in Equation 3.8.
- $TrkIso$: track-based isolation. This variable evaluates the isolation using tracking detectors information:

$$TrkIso \equiv \frac{\sum_i p_T^{i,\Delta R=0.4} - p_T}{p_T} < 0.1. \tag{3.9}$$

  where the sum index i runs over all tracks with $p_T > 500$ MeV and $|\eta| < 1.5$ in a cone of $\Delta R = 0.4$ around the electron track.

- $Q \times \Delta x_{CES}$: distance in the $r - \phi$ plane between the extrapolated track and the nearest cluster reconstructed in the CES detector, multiplied by the charge of the track to account for asymmetric tails originated from bremsstrahlung radiation.
- Δz_{CES}: distance in the r-z plane between the extrapolated track and the best matching CES cluster.
- $NCotHitsAx$: number of Axial COT super-layers with at least 5 hits belonging to the track associated to the candidate electron.
- $NCotHitsSt$: number of Stereo COT super-layers with at least 5 hits belonging to the track associated to the candidate electron.
- χ^2_{COT}: χ^2 of the track fit computed using only the COT hits belonging to the track.
- $NSvxHits$: number of SVX hits belonging to the track associated to the candidate electron.

These variables are then used to form the likelihood:

$$\mathcal{L}(\vec{x}) = \frac{L_{sig}}{L_{sig} + L_{bckg}} = \frac{\prod_{i=1}^{N} P_i^{sig}(x_i)}{\prod_{i=1}^{N} P_i^{sig}(x_i) + \prod_{i=1}^{N} P_i^{bckg}(x_i)} \tag{3.10}$$

where

- x_i: is the i-th identification variable used in the likelihood. These are the 11 variables described above.
- N: the number of ID variables, in our case 11.
- $P_i^{sig}(x_i), P_i^{bckg}(x_i)$: are the functions that give the probability to obtain the value x_i for the i-th ID variable given a real (signal) or a fake (background) electron, in the following 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 identified electron and a looser one (*probe*) with an invariant mass within $76 < m(ll) < 106$ GeV and opposite

	TCE
Region	central
Track p_T	≥ 10 GeV (5 GeV if $E_T < 20$ GeV)
Track $\lvert z_0 \rvert$	≤ 60 cm
# Ax SL (5hits)	≥ 3
# St SL (5hits)	≥ 2
Conversion	false
E_{HAD}/E_{EM}	≤ 0.055 GeV $+ 0.00045 * E$
$CalIso$	≤ 0.1
Lshr	≤ 0.2
E/P	< 2.5 GeV $+ 0.015 * E_T$
$\Delta X \cdot Q$	$-3 \leq q * \Delta X \leq 1.5$ cm
CES $\lvert \Delta Z \rvert$	< 3 cm
Track	Beam constrained

Table 3.5: Definition of cut-based Tight Central Electron (TCE).

charge. The fully identified electron is selected using standard CDF cut-based selections listed in Table 3.5 (TCE) and based on the same variables previously described. The second electron is required to have an identified electron-like cluster associated with a track satisfying looser selections reported in Table 3.6. This sample is dominate by Drell-Yan events with real leptons. The ID variables of probe objects are used to make the signal templates. When both legs are identified as TCE both are used in the template. The resulting likelihood distribution is shown in Figure 3.25(Left).

Fake electrons are obtained from a jets data sample. This sample is selected by requiring at least one jet with energy greater than 20 GeV (*JET20* data set) [1]. One reconstructed jet with $E_T > 20$ GeV and a loose electron (*fakeable*, see Table 3.7) per event are required, with an invariant mass out the Z peak region to remove Drell-Yan contribution: $m(ll) < 76$ GeV or $m(ll) > 106$ GeV. 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 is shown in Figure 3.25(Right).

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 (first bins of the Left plot and last bins of Right one) in Figure 3.25. 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. We find and subtract a contribution the order of $\sim 1\%$.

[1] See Appendix B for a detailed description of *JET20* requirements.

	Central electron probe
Region	Central
Track p_T	≥ 5 GeV
Track Z_0	≤ 60 cm
Conversion	false

Table 3.6: Definition of central electron probe.

	Fakeable Central Electron
Region	Central
E_{HAD}/E_{EM}	≤ 0.125 GeV $+ 0.00045 * E$
$CalIso$	< 0.3
Conversion	false

Table 3.7: Definition of central electron fakeable object.

Forward Electrons

Electron candidate clusters in the plug calorimeter are made starting from a seed tower and adding neighboring towers within two towers in η_{det} and ϕ from the seed. The hadronic energy of the cluster is required to be less than 0.05 times the electromagnetic energy. Plug electrons have to be reconstructed in a well-instrumented region of the detector, defined as $1.2 \leq |\eta| \leq 2.0$ as measured by the PES sub-system (η_{det}^{PES}). To improve cluster-matching track quality, we require a track with hits in at least 3 layers of the silicon detector, given the COT limited coverage in the plug region. The track is further required to have $|z_0| < 60$ cm. Additional requirements are summarized in Table 3.8. For historical reasons these electrons are commonly referred as PHOENIX electrons (**PHX**).

3.5.2 Muon identification

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. A muon is reconstructed starting from a track and adding track segments (stubs) formed with hits in the muon

	PHX
Region	Plug
$\lvert\eta_{det}^{PES}\rvert$	$1.2 \leq \lvert\eta_{det}^{PES}\rvert \leq 2.0$
E_{HAD}/E_{EM}	≤ 0.05
PEM 3×3 χ^2	≤ 10
PES5 $\times$ 9U	≥ 0.65
PES5 $\times$ 9V	≥ 0.65
$CalIso$	≤ 0.1
$\Delta R(PES, PEM)$	≤ 3.0
NSiHits	≥ 3
Track $\lvert z_0\rvert$	≤ 60 cm

Table 3.8: Definition of forward electrons (PHX).

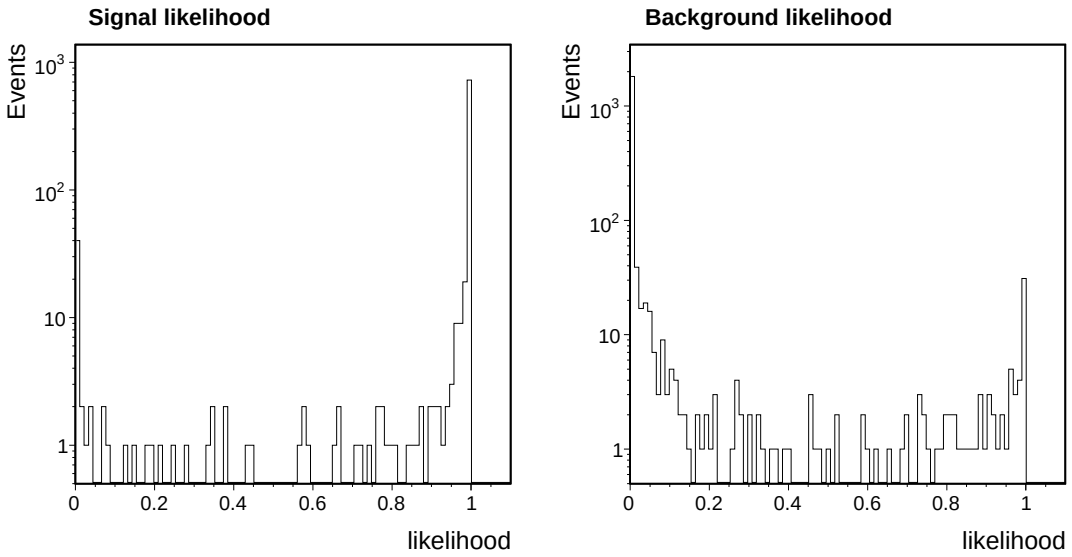


Figure 3.25: Likelihood distribution for real electrons (left) and fake electrons (right).

drift chambers. The track origin in the $x - y$ plane is constrained to the beam position. The muon 4-momentum $(E, \vec{p})$ is determined by measuring the track $\vec{p}$ and assuming a massless particle: $E \equiv |\vec{p}|$.

Muons are reconstructed in eight non-overlapping categories. Six of them require the muons to have hits in one muon detector subsystem, and we call them *stubbed muons*. The other two are designed to recover muons that do not have a stub and are referred as *stubbless muons*. To further increase acceptance we also define an extra lepton category in which both electrons and muons can fall. This category accepts leptons which fall in regions of the detector not fully instrumented (*cracks*) and fails electron and muon selections. Figure 3.26 shows the distribution in the $\eta - \phi$ plane of the reconstructed muons for a small sub-sample of analyzed data. The different categories are summarized in Table 3.3 and described in more detail in the following.

Stubbed Muons

Stubbed muon candidates are required to have a reconstructed track with a fit $\chi^2/n.d.f. < 3$. This requirement is relaxed to $\chi^2/n.d.f. < 4$ for data taken before run 186598. The track is required to have at least three Axial and two Stereo COT super-layers with at least 5 hits. The track $|z_0|$ has to be less than 60 cm. For forward muons (BMU) we have a limited coverage from the COT drift chamber, so we removed the hits requirement and we instead require a number of hits in the chamber that is at least 60% of the expected one, based on the track direction; we also require at least three hits in the Silicon detectors and a curvature significance $C/\sigma(C) > 12$. For all muons, the distance at the closest approach of the extrapolated track to the primary vertex, which is called impact parameter (d_0), has to be $|d_0| < 0.2$ cm, and it is tightened to be $|d_0| < 0.02$ cm if the track has also hits in the Silicon detectors, giving a much precise measurement of the impact parameter. To reject background we also require the track to be isolated: $TrkIso < 0.1$, as defined in

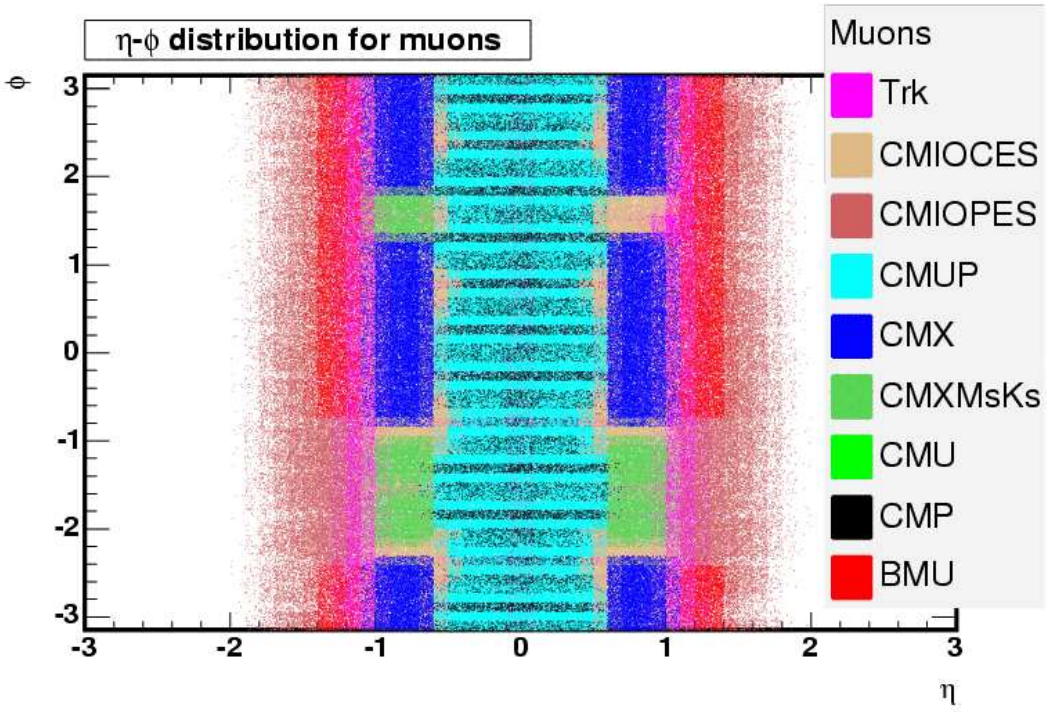


Figure 3.26: Distribution in the $\eta - \phi$ plane of a sub-sample of the reconstructed muons in data.

Equation 3.9.

The reconstructed track is then required to be compatible with a minimum ionizing particle (*m.i.p.*) by cutting on the energy deposited in the EM and HAD towers hit by the extrapolated track. We also require calorimetric isolation, which is defined in analogy of equation 3.8:

$$CalIso \equiv \frac{E_T^{\Delta R=0.4}}{E_T^{\mu}} \tag{3.11}$$

where the $E_T^{\mu} = p_T^{\mu}$.

We then differentiate stubbed muons in six categories, 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. The cuts applied to each category are summarized in Tables 3.9-3.12. and described in detail in Appendix A.

Stubless Muons

Stubless muons are recovered by selecting an isolated high-p_T track which deposited in the calorimeter an amount of energy consistent with a minimum ionizing track. Two

	CMUP
Region	Central
Track $\chi^2/n.d.f.$	≤ 3(≤ 4 if run ≤ 186598)
NAxL(5 hits)	≥ 3
NStL(5 hits)	≥ 2
Track z_0	≤ 60 cm
Track d_0	≤ 0.2 cm(≤ 0.02 cm $if\ NSiHits > 0$)
TrkIso	≤ 0.1
CalIso	≤ 0.1
E_{EM}	$\leq 2 + \max(0, (p - 100) \cdot 0.0115)$ GeV
E_{HAD}	$\leq 6 + \max(0, (p - 100) \cdot 0.028)$ GeV
ΔX_{CMU}	≤ 7 cm
ΔX_{CMP}	$\leq \max(6.0, 150/p_T[GeV])$ cm
Fiduciality	x-fid$_{CMU} < 0$ cm z-fid$_{CMU} < 0$ cm
	x-fid$_{CMP} < 0$ cm z-fid$_{CMP} < -3$ cm

Table 3.9: Definition of CMUP central muons.

categories are defined:

- **CMIOCES**: Central stubless muons which are fiducial to the central calorimeter.
- **CMIOPES**: Central stubless muons which are fiducial to the plug calorimeter.

The requirements of these two categories are reported in Table 3.13 and described in more detail in Appendix A.

Identification of tracks of unknown lepton type: Crack Tracks (CrkTrk)

To further increase the acceptance the leptons that fall in the crack of the detectors are recovered. To specifically select these leptons and to preserve the uniqueness of the categories the track has not to be fiducial to the central or plug calorimeter and we explicitly veto overlaps with stubbed muon categories.

The selection criteria are similar to CMIOCES muons, but we do not require any deposit in the calorimeter, since these tracks point to or near the cracks in the calorimeter. However we still require a low calorimeter and tracking activity near the track ($CalIso$ and $TrkIso$) which is fundamental in order to reduce fake lepton identification. We further require the track not to be consistent with a conversion electron.

Table 3.14 lists the detailed selection criteria applied. Since both electrons and muons can enter this category, we treat this category separately in the analysis.

3.5.3 Jet identification

Several algorithms have been developed for reconstructing jets. We use a fixed cone algorithm, JetClu(58), with a radius of $\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2} = 0.4$. This algorithm is based on calorimetric information and it starts considering all the calorimeter towers with $E_{EM} \cdot sin\theta_{EM} + E_{HAD} \cdot sin\theta_{HAD} > 1$ GeV, where E_{EM} and E_{HAD} are the energy of the electromagnetic and hadronic calorimeter tower respectively, and $\theta_{EM}, \theta_{HAD}$ are

	CMU
Region	Central
Track $\chi^2/n.d.f.$	≤ 3(≤ 4 if run ≤ 186598)
NAxL(5 hits)	≥ 3
NStL(5 hits)	≥ 2
Track z_0	≤ 60 cm
Track d_0	≤ 0.2 cm(≤ 0.02 cm $if\ NSiHits > 0$)
TrkIso	≤ 0.1
CalIso	≤ 0.1
E_{EM}	$\leq 2 + \max(0, (p - 100) \cdot 0.0115)$ GeV
E_{HAD}	$\leq 6 + \max(0, (p - 100) \cdot 0.028)$ GeV
Fiduciality	x-fid$_{CMU} < 0$ cm z-fid$_{CMU} < 0$ cm
	Not CMP Fiducial
	Not CMX Fiducial
ΔX_{CMU}	≤ 7 cm
Good Trigger	$Run > 270062$
	CMP
Region	Central
Track $\chi^2/n.d.f.$	≤ 3(≤ 4 if run ≤ 186598)
NAxL(5 hits)	≥ 3
NStL(5 hits)	≥ 2
Track z_0	≤ 60 cm
Track d_0	≤ 0.2 cm(≤ 0.02 cm $if\ NSiHits > 0$)
TrkIso	≤ 0.1
CalIso	≤ 0.1
E_{EM}	$\leq 2 + \max(0, (p - 100) \cdot 0.0115)$ GeV
E_{HAD}	$\leq 6 + \max(0, (p - 100) \cdot 0.028)$ GeV
Fiduciality	x-fid$_{CMP} < 0$ cm z-fid$_{CMP} < -3$ cm
	Not CMU Fiducial
ΔX_{CMP}	$\leq \max(6.0, 150/p_T[GeV])$ cm
$\phi - gaps$	$\phi \bmod 15^\circ \leq 2$ OR ≥ 13
Good Trigger	$Run > 229764$

Table 3.10: Definition of CMU and CMP central muons. CMU muons are not used up to now, since a trigger bug was discovered which affected the majority of the current data sample. This bug is fixed from run 270062 and will possibly be included in future versions of the analysis.

	CMX
Region	Central
Track $\chi^2/n.d.f.$	≤ 3(≤ 4 if run ≤ 186598)
NAxL(5 hits)	≥ 3
NStL(5 hits)	≥ 2
Track z_0	≤ 60 cm
Track d_0	≤ 0.2 cm(≤ 0.02 cm $if\ NSiHits > 0$)
TrkIso	≤ 0.1
CalIso	≤ 0.1
E_{EM}	$\leq 2 + \max(0, (p - 100) \cdot 0.0115)$ GeV
E_{HAD}	$\leq 6 + \max(0, (p - 100) \cdot 0.028)$ GeV
Fiduciality	x-fid$_{CMX}$ $<= 0$ cm z-dif$_{CMX}$ $<= -3$ cm
	$0° < \phi < 75°$ OR
	$105° < \phi < 225°$ OR
	$315^{circ} < \phi < 360°$
	$\rho_{COT} > 140$ cm
ΔX_{CMX}	$< \max(6.0, 125.0/p_T)$
Good Trigger	Run $>=$ 227704
	CMXMsKs
Region	Central
Track $\chi^2/n.d.f.$	≤ 3(≤ 4 if run ≤ 186598)
NAxL(5 hits)	≥ 3
NStL(5 hits)	≥ 2
Track z_0	≤ 60 cm
Track d_0	≤ 0.2 cm(≤ 0.02 cm $if\ NSiHits > 0$)
TrkIso	≤ 0.1
CalIso	≤ 0.1
E_{EM}	$\leq 2 + \max(0, (p - 100) \cdot 0.0115)$ GeV
E_{HAD}	$\leq 6 + \max(0, (p - 100) \cdot 0.028)$ GeV
Fiduciality	x-fid$_{CMX}$ $<= 0$ cm z-dif$_{CMX}$ $<= -3$ cm
	$(75° < \phi < 105°$ AND $\lvert\eta\rvert < 0)$
	OR
	$(225° < \phi < 315°)$
	$\rho_{COT} > 140$ cm
ΔX_{CMX}	$< \max(6.0, 125.0/p_T)$
Good Trigger	Run $>=$ 227704

Table 3.11: Definition of CMX and CMXMsKs central muons.

	BMU
Region	Forward
COT hit fraction	≥ 0.6
NSiHits	≥ 3
Track z_0	≤ 60 cm
Track d_0	≤ 0.02 cm
$C/\sigma(C)$	> 12
TrkIso	≤ 0.1
CalIso	≤ 0.1
E_{EM}	$\leq 2 + \max(0, (p - 100) \cdot 0.0115)$ GeV
E_{HAD}	$\leq 6 + \max(0, (p - 100) \cdot 0.028)$ GeV
$E_{EM} + E_{HAD}$	> 0.1 GeV
Fiduciality	x-fid$_{BMU} < 0$ cm z-fid$_{BMU} < -3$ cm PES fiducial (471.6 <= z_{BMU} <= 766.6 OR -433.0 <= z_{BMU} <= -764.7)
$N_{\text{stub hits}}$	> 2

Table 3.12: Definition of BMU forward muons.

	CMIOCES
Region	Central
NAxL(5 hits)	≥ 3
NStL(5 hits)	≥ 3
Track $\chi^2/n.d.f.$	≤ 3
Track z_0	≤ 60 cm
Track d_0	≤ 0.02 cm
TrkIso	≤ 0.1
CalIso	≤ 0.1
E_{EM}	$\leq 2 + \max(0, (p - 100) \cdot 0.0115)$ GeV
E_{HAD}	$\leq 6 + \max(0, (p - 100) \cdot 0.028)$ GeV
$E_{EM} + E_{HAD}$	> 0.1 GeV
Fiduciality	CES Fiducial
Uniqueness	Not a CMUP, CMU, CMP, CMX, CMXMsKs

	CMIOPES
Region	Plug
COT hit fraction	≥ 0.6
NSiHits	≥ 3
Track z_0	≤ 60 cm
Track d_0	≤ 0.02 cm
$C/sigma(C)$	> 12
TrkIso	≤ 0.1
CalIso	≤ 0.1
E_{EM}	$\leq 2 + \max(0, (p - 100) \cdot 0.0115)$ GeV
E_{HAD}	$\leq 6 + \max(0, (p - 100) \cdot 0.028)$ GeV
$E_{EM} + E_{HAD}$	> 0.1 GeV
Fiduciality	PES fiducial
Uniqueness	Not a BMU

Table 3.13: Definition of stubless central (CMIOCES) and forward (CMIOPES) muons.

	CrkTrk
NAxL(5 hits)	≥ 3
NStL(5 hits)	≥ 3
Track $\chi^2/n.d.f.$	≤ 3
Track z_0	≤ 60 cm
Track d_0	≤ 0.02 cm
TrkIso	≤ 0.1
CalIso	≤ 0.1
	OR
	≤ 0.1 using nearest EM cluster with $\Delta R = 0.05$
$E_{EM} + E_{HAD}$	> 0.1 GeV
Fiduciality	Not PES fiducial
	Not CES fiducial
Conversion	false
Uniqueness	Not a CMUP, CMU, CMP, CMX, CMXMsKs

Table 3.14: Definition of Crack Tracks.

the azimuthal angles of the corresponding tower. Starting from the highest E_T tower, adjacent ones within ΔR are clumped together to form a *pre-cluster*. Each tower can only be assigned to one pre-cluster. After the first iteration all towers with energy greater than 0.1 GeV are considered and added to each pre-cluster if within $\Delta R = 0.4$ from the E_T-weighted centroid of the pre-cluster. A new centroid for the cluster is recomputed each time that new towers are added. At this stage an overlap among pre-clusters is evaluated, defined as the ratio between the sum of transverse energies of the common towers and the smaller cluster E_T between the overlapping ones. If the overlap is above 0.75 the two clusters are merged, otherwise common towers are reclumped into the cluster with the nearest centroid. Final clusters are commonly referred as *raw jets*.

Jet Energy corrections

The energy scale of the electromagnetic calorimeter is set by reconstructing $Z \rightarrow e^+e^-$ decays, and imposing the reconstructed mass peak to be consistent with the world average mass of the Z boson. The ratio of the calorimeter energy and track momentum measurements, E/p, for well-identified electrons is used to apply additional relative calibration for each tower in order to improve the energy measurement resolution. The calibrations for the hadronic calorimeter are derived from a test beam with 50 GeV charged pions. The definition of the energy scale of the calorimeter is often referred as the Level 0 correction to the jets energy.

Even after the scale calibrations the response of the calorimeter is not uniform in pseudorapidity. The response dependencies on η arise from the separation of calorimeter components at $\eta = 0$, where the two halves of the central calorimeter join, and at $\eta \sim 1.1$, where the plug and central calorimeter are merged. The η-dependent corrections are obtained by requiring that the two leading jets in dijet events are balanced in p_T in absence of hard QCD radiation.

At high instantaneous luminosity more than one $p\bar{p}$ interaction may occur in the same bunch crossing. Given the Tevatron characteristics, the average number of interactions

is one for $\mathcal{L} = 0.4 \times 10^{32}\ cm^{-2}s^{-1}$, and increases to 3 and 8 for $\mathcal{L} = 1 \times 10^{32}\ cm^{-2}s^{-1}$, and $\mathcal{L} = 3 \times 10^{32}\ cm^{-2}s^{-1}$, respectively. These extra $p\bar{p}$ interactions increase the energy of the jets from the hardest scatter if their final state hadrons accidentally overlap with the jet originating from the primary interaction. The number of primary vertexes in the event, N_{vtx} is a good indicator of additional interactions occurring in the same bunch crossing. The transverse energy in a random cone is measured in minimum-bias data and parametrized as a function of N_{vtx} in the event. This procedure allows to extract the average energy each extra vertex in the event is adding, and then to correct jet energies accordingly.

An additional correction is applied to account for the jet modeling in the simulations. It depends on the correct simulation of the multiplicity and the p_T spectrum of the particles inside the jet. The correction is derived minimizing the average difference between the parton-level and the reconstructed calorimeter-level jets. After this correction the energy scale of the jets is independent of the CDF detector.

All these corrections are described in detail in Ref. (58). We consider only jets that have $E_T > 15$ GeV, after these corrections, and $|\eta| \leq 2.5$. Jets are also required to be away (with $\Delta R \geq 0.4$) from identified leptons.

$b-$jet identification

The high position resolution provided by the silicon vertex detectors can be exploited to identify secondary vertexes originated inside a jet by decays of long lifetime particles produced in heavy quark hadronization. This approach has been followed by the SECondary VerTeX (SECVTX) tagging algorithm(59) to provide discrimination of high-p_T b-jets from jets originated by light quarks or gluons. It relies on the displacement of secondary vertexes relative to the primary event vertex to identify b hadron decays.

The SECVTX algorithm operates on a per-jet basis, where only tracks within the jet cone are considered for each jet in the event. Displaced tracks in the jet are selected based on the significance of their impact parameter with respect to the primary vertex and are used to reconstruct secondary vertexes in the jet. Once a secondary vertex is found, the two-dimensional decay length L_{xy} is calculated as the projection onto the jet axis, in the $x - y$ plane, of the vector pointing from the primary vertex to the secondary vertex. Secondary vertexes corresponding to the decay of b hadrons are expected to have large positive L_{xy} while the secondary vertexes from random mis-measured tracks are expected to be less displaced from the primary vertex. To reduce the background from false secondary vertexes, a good secondary vertex is required to have $|L_{xy}/\sigma_{L_{xy}}| > 7.5$, where $\sigma_{L_{xy}}$ includes the error on the primary vertex and is of $\mathcal{O}(100\ \mu m)$. A b-tagged jet is defined to be a jet containing a good secondary vertex.

3.5.4 Missing transverse energy

Neutrinos weakly interact with the surrounding material. They escape the detection carrying away some amount of energy which cannot be directly measured, thus creating an energy imbalance in the detector. The $z-$component of the initially interacting partons is unknown on an event-basis, so the amount of missing energy cannot be determined.

However the momentum in the transverse plane is in good approximation zero, which allows to define the transverse component of the missing energy as:

$$\vec{E}_T^{raw} = -\sum_i E_T^i \hat{n}_i, \tag{3.12}$$

E_T^i represent the transverse energy measured in the $i-$th tower of the calorimeter and $\hat{n}_i$ is the projection of the versor pointing from the event vertex to the $i-$th calorimeter tower onto the plane perpendicular to the beam axis.

Corrections are applied to the raw missing transverse energy.The same corrections used for the raw jet energies discussed in Section 3.5.3 are applied. The most important correction is due to the reconstructed muons: we subtract the expected amount of energy left in the calorimeter and add back the measured p_T.

3.6 Trigger paths

One of the key ingredients for a successful physics program at hadron colliders is being able to trigger an an handful of events among the several millions of collisions occurring every second.

We exploit the three-level triggering system described in Section 3.3 in order to select events where at least one high-p_T electron or muon was produced in the collision and reconstructed by the CDF detector. At each level electrons and muons can be reconstructed with increased resolution.

In more detail we exploit the following six trigger paths:

- *ELECTRON_CENTRAL_18*, selects events with at least one central electron with $E_T \geq 18$ GeV;
- *MET_PEM*, which requires one forward electron and missing energy to reduce the trigger cross section to manageable levels;
- *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 derive 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 detail requirements at trigger level for all the trigger paths used in this Thesis are described in Appendix B.

3.6.1 Trigger efficiency

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 Section 3.5. Electrons and stubbed muons are triggerable leptons if $E_T > 20$ GeV, to avoid significant turn-on effects, while stubless muons and CrkTrk are not considered triggerable leptons. Each triggerable lepton can be triggered by only one trigger path. Table 3.15 summarize the triggerable leptons types and the associated trigger paths. This object is the denominator of the efficiency measurements. The numerator is the number of triggerable leptons which fired the corresponding trigger path.

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 a 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 detector and trigger conditions are stable and we perform just one measurement.

Central and Forward 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 $m(ll) \in [76, 106]$ GeV. 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 if 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 lower than 1%. However we assign a moderately inflated systematic uncertainty of 2%.

Appendix B describes the detailed calculations and results 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 as will be described in detail in Section 4.3. Data events are always required to have at least one reconstructed triggerable lepton compatible with the online trigger path that selected the event.

Lepton category	Trigger path
LBE	ELECTRON_CENTRAL_18
PHX	MET_PEM
CMUP	MUON_CMUP18
CMP	MUON_CMP18_PHIGAP
CMU	MUON_CMU18_ETAGAP
CMX	MUON_CMX18
CMXMsKs	MUON_CMX18

Table 3.15: Triggerable lepton types with the associated trigger paths.

Chapter 4

Di-lepton event selection and modeling

Chapter 2 outlined the most promising strategies to discover the Higgs boson in the region $m_H \gtrsim 135$ GeV.

In this Chapter we study the basic event selections and the expected sample composition for the most promising signature: events with exactly two opposite charged leptons and a significant amount of missing transverse energy. The simulations and data-driven methods used to model the selected data samples are also described.

4.1 Event selection

We start selecting events collected by one of the high$-p_T$ lepton trigger paths described in Section 3.6, corresponding to an integrated luminosity of $\sim 4.8 fb^{-1}$. In order to avoid double-counting every event triggered by more than one trigger path is accepted only once. While CDF detector is now steadily taking data, the running conditions unavoidably change with time; in particular some parts of the detector may be not properly working for part of the collected data. The total integrated luminosity analyzed depends therefore on the requirements (*good run lists*) applied on subdetectors online status, which vary with the reconstructed leptons in the event. Table 4.1 shows the amount of integrated luminosity analyzed for different lepton types and the corresponding detector requirements. We require each event to be in the most strict good run list which is needed by the reconstructed leptons in the event.

Good run list	Requires proper operation of		$\mathcal{L}[\text{pb}^{-1}]$
	Silicon	Muon chambers	
Central Electron	no	no	4829
Forward Electron	yes	no	4549
Central Muon	no	yes	4662
Forward Muon	yes	yes	4395

Table 4.1: Integrated luminosities ($\mathcal{L}$) collected by the different good run lists.

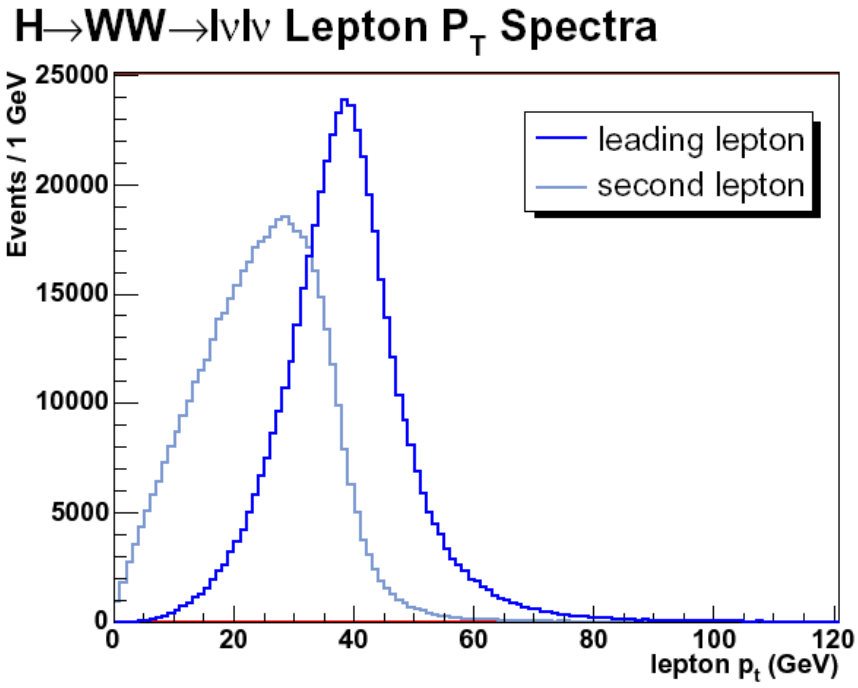


Figure 4.1: Leading and sub-leading lepton p_T distribution for signal $H \to WW \to ll\nu\nu$ events, for an Higgs mass $m_H = 160$ GeV.

Leptons are reconstructed according to the algorithms described in Section 3.5 and summarized in Table 3.3. The triggerable lepton is required to have $E_T > 20$ GeV, while the other lepton in the event is only required to have $p_T > 10$ GeV. Figure 4.1 shows the expected p_T distribution for the leading and sub-leading lepton for $H \to W^+W^- \to ll\nu\nu(l = e, \mu)$ signal events, with $m_H = 160$ GeV. Accepting a second lepton with lower p_T increases the acceptance to signal events by $28 - 510\%$ depending on the Higgs mass (ranging from $m_H = 110$ GeV to $m_H = 200$ GeV).

Opposite charge between the two leptons is required to suppress $W + \gamma/jets$ backgrounds, where the photon or jet is mis-identified as a lepton.

In the following we will refer to electrons as a an object identified by the LBE or PHX selections; a muon is any object identified by one of the following identification algorithms: CMUP, CMP, CMX, CMXMsKs, CMIOCES, CMIOPES; finally we treat separately objects identified as CrkTrk which we simply call isolated tracks (trk), since both real electrons and real muons can easily enter this category. We classify the selected events by the flavor of the lepton pair according to:

e − e: two identified electrons;

$\mu - \mu$: two identified muons;

e − μ: one identified electron and one muon;

e − trk: one identified electron and one CrkTrk;

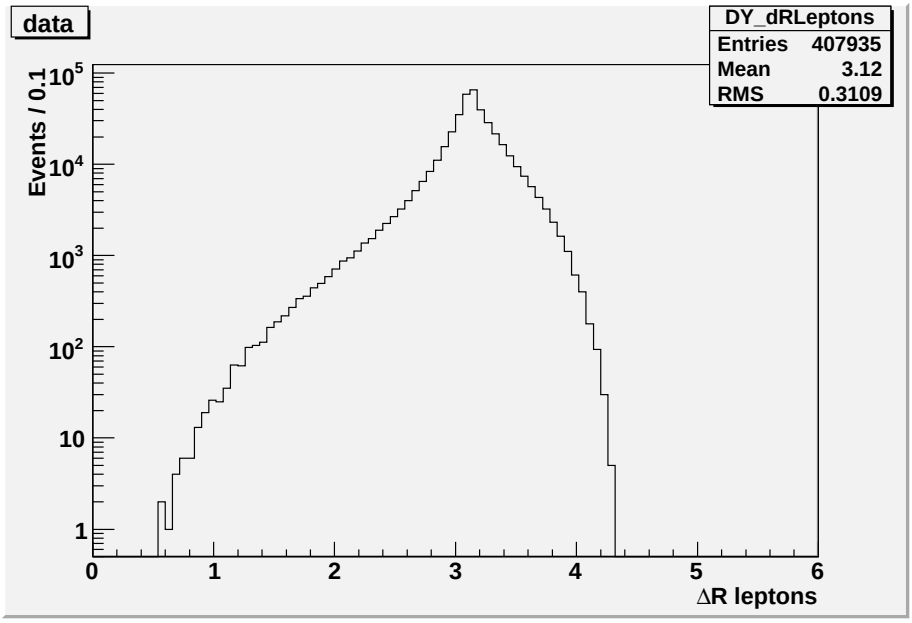


Figure 4.2: ΔR distribution for selected $Z \to ll$ events.

$\mu -$ **trk:** one identified muon and one CrkTrk.

High$-p_T$ muons from cosmic rays are removed from the data sample(60). Nearly 100% of cosmic ray events are rejected with a negligible loss of efficiency.

The two leptons are required to be consistent with originating from the same $p\bar{p}$ interaction cutting on the distance between leptons in the extrapolated z_0: $|\Delta z_0| < 4$ cm.

A minimum distance in the $\eta - \phi$ plane of $\Delta R > 0.05$ is requested between the two leptons. This requirement is actually already much tighter. In fact the calorimetric and tracking isolation required in the lepton identification selections ensure $\Delta R \gtrsim 0.4$. This can be seen for example in Figure 4.2 where we show the ΔR distribution for a sample of $Z \to ll$ data candidate events.

We also require the lepton pair to have an invariant mass $m(ll) > 16$ GeV. This selection is needed to get rid of heavy-flavour hadrons decays, especially Υ resonances, $B-$cascade decays and $b\bar{b}$ QCD production. This cut is also very effective in reducing $W\gamma$ contribution as can be seen in Figure 4.3. This Figure shows the expected and observed invariant mass distribution of the two leptons. The different components are stacked on top of each other, and the contribution of a 160 GeV Higgs boson with yields scaled up by a factor of 10 is overlaid. The fraction of signal rejected by the invariant mass requirement is between 2% and 13% depending on its mass.

After these selections data sample is dominated by Drell-Yan events. Drell-Yan (DY) events have in principle no missing transverse energy, since the decay products can be reconstructed in the detector. However, detector resolution effects, especially for events where initial/final state radiation produces jet activity, can give significant missing transverse energy. On the other hand, signal processes have neutrinos in the final state which are expected to produce a significant energy imbalance in the transverse plane. We define

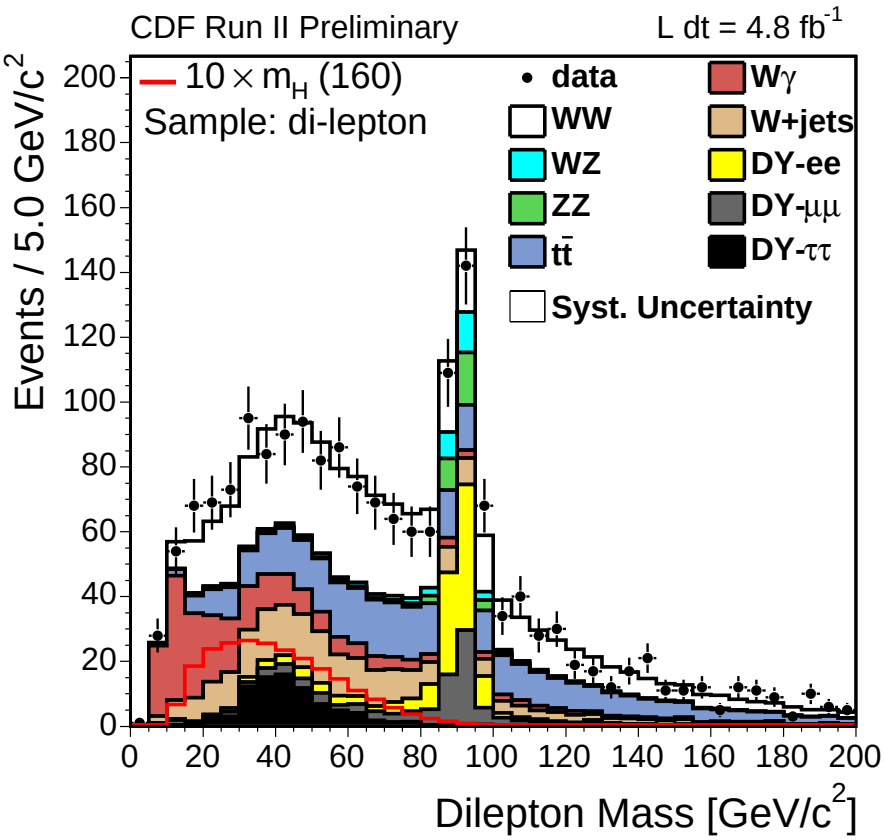


Figure 4.3: Expected and observed invariant mass distribution of the two selected leptons.

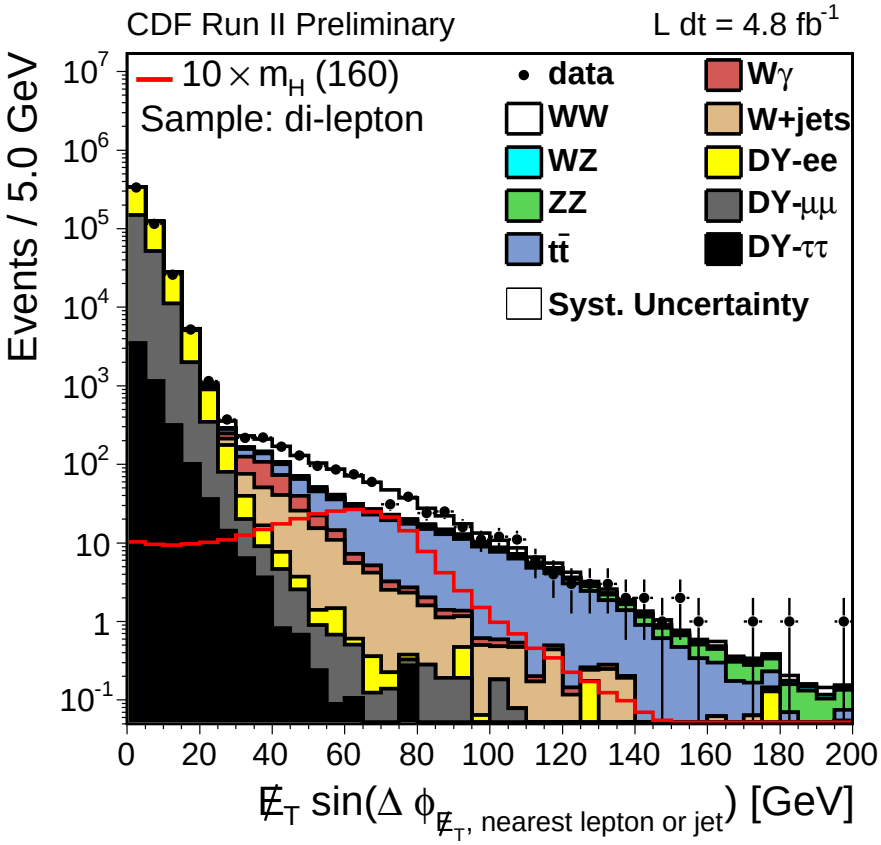


Figure 4.4: The Expected and observed $\vec{\not\!\!E}_T^{\rm Spec}$ distribution for the selected sample before the $\vec{\not\!\!E}_T^{\rm Spec}$ requirement.

the $\vec{\not\!\!E}_T^{\rm Spec}$ to get the best separation from the signal and the Drell-Yan background:

$$\vec{\not\!\!E}_T^{Spec} = \begin{cases} \vec{\not\!\!E}_T \cdot \sin\Delta\phi(\vec{\not\!\!E}_T, l \text{ or } jet) & \text{if } \Delta\phi(\vec{\not\!\!E}_T, l \text{ or } jet) < \pi/2 \\ \vec{\not\!\!E}_T & \text{if } \Delta\phi(\vec{\not\!\!E}_T, l \text{ or } jet) >= \pi/2 \end{cases} \tag{4.1}$$

If the $\vec{\not\!\!E}_T$ is collinear with a lepton or a jet is likely that it comes from a mis-measurement of the lepton or the jet, so we consider only the transverse component respect to the lepton or jet direction. We require $\vec{\not\!\!E}_T^{\rm Spec} > 25$ GeV; this requirement is loosened to $\vec{\not\!\!E}_T^{\rm Spec} > 15$ GeV for $e-\mu$ events, since Drell-Yan contribution is significantly smaller. Figure 4.4 shows the distribution of the expected and the observed number of events as function of $\vec{\not\!\!E}_T^{\rm Spec}$ before the described cut. The 99.994% of Drell-Yan events are rejected after the cut with an 85% efficiency on the signal.

In summary these are the requirements applied so far:

1. unique events collected by one of the high$-p_T$ lepton trigger paths;
2. exactly two opposite charge leptons;
3. one trigger lepton with $p_T > 20$ GeV, the other with $p_T > 10$ GeV;

4. data events in the appropriate good run list

5. cosmic ray veto;

6. $|\Delta z_0| < 4$ cm between leptons;

7. $\Delta R(ll) > 0.05$;

8. invariant mass of the lepton pair $m(ll) > 16$ GeV;

9. $\not\!\!E_T^{\rm Spec} > 15$ GeV for $e-\mu$ events, $\not\!\!E_T^{\rm Spec} > 25$ GeV otherwise

In 4.8 fb^{-1} of data 1727 events survived these cuts.

4.2 Sample composition

The physics processes that we expect to enter the selected sample are summarized in Table 4.2 and described in the following:

WW : the non-resonant production $p\bar{p} \to W^+W^-$ represents 33% of the sample, being the main contributor. The production cross section for this process has been calculated at Next-to-Leading Order (NLO) with MCFM(31). The result is $\sigma(WW) = 12.4 \pm 0.8$ pb, which is about a factor that goes from 25 to 150 times bigger than the predicted $H \to W^+W^-$ resonant production cross section, depending on the mass of the Higgs boson.

Other di-bosons : $p\bar{p} \to WZ$ and $p\bar{p} \to ZZ$ production also enter the data sample, with an expected contribution of about 4%. The process $ZZ \to ll\nu\nu$ passed the events selections as well as the $WZ \to l\nu ll$ decay with one of the leptons lost. The production cross section for these processes has been also evaluated with MCFM at NLO to be $\sigma(WZ) = 3.5 \pm 0.2$ pb and $\sigma(ZZ) = 1.51 \pm 0.09$ pb.

Drell-Yan : $p\bar{p} \to Z/\gamma^*$ contribute at the level of $\sim 22\%$ after the cuts, given the high cross section ($\sigma(p\bar{p} \to Z/\gamma^* \to l^+l^- \sim 490$ pb for $m(Z/\gamma^*) > 20$ GeV at NNLO(61)).

Wγ : While the production of single W bosons does not contribute directly to the data sample, the production of a photon in association with the W vector boson contributes $\sim 7\%$ after the cuts if the photon is mis-identified as a lepton.

t$\bar{\rm t}$: the production cross section for $t\bar{t}$ is 7.88 ± 0.79 pb(62) [1]. Top quark decays $\sim 100\%$ of the times into a W boson and a $b-$quark. These events pass the selections if both Ws decay leptonically. The presence of two $b-$jets in the final state can be helpful in discriminating these events from the signal. In the di-lepton sample $t\bar{t}$ contributes for 20%.

[1] This value assume $m_{top} = 172.4 \pm 1.2$ GeV

CDF Run II Preliminary	$\int \mathcal{L} = 4.8$ fb^{-1}		
$t\bar{t}$	364	$\pm$	60
DY	396	$\pm$	75
WW	595	$\pm$	63
WZ	45.2	$\pm$	6.2
ZZ	40.4	$\pm$	5.5
W+jets	247	$\pm$	61
$W\gamma$	132	$\pm$	21
Total Background	1820	$\pm$	130
Data	1727		

OS All Jets

Table 4.2: Observed and Expected number of events in ~ 4.8 fb^{-1} after di-lepton selections of Section 4.1. Errors include systematic uncertainties that will be discussed in Section 5.3.

W/Z + jets and QCD: Jets coming from initial state radiation in the W or Z production can erroneously be mis-identified as leptons. Although the probability is low, the large cross section respect to signal processes makes them to enter significantly in our sample together with the multi-jet production. We commonly indicate these contributions with $W + jets$, since it is the main component. All together they constitute 14% of the data sample.

Figure 4.5 graphically shows the relative contributions of these processes to the selected sample.

Table 4.2 shows that we expect a total of 1820 ± 130 events from the outlined physics processes, which is in good agreement with the selected data sample.

For the most favorable case of an Higgs boson with a mass $m_H = 165$ GeV, the number of expected signal events in this sample is 28.3 ± 5.2: 21.6 ± 4.7 from gluon fusion production process, 3.18 ± 0.41 and 1.78 ± 0.23 from production in association with a W or a Z, and 1.73 ± 0.28 from Vector Boson Fusion.

4.3 Monte Carlo simulations

The full simulation of the physics processes of interest can be divided into three main components:

- The physics process: a program that simulates the proton-proton interactions and the produced particles;
- Detector: simulates the interactions of particle with the CDF detector and the effects of the online conditions of data taking;
- Trigger and analysis simulation: in charge of reproducing trigger requirements and the analysis selections.

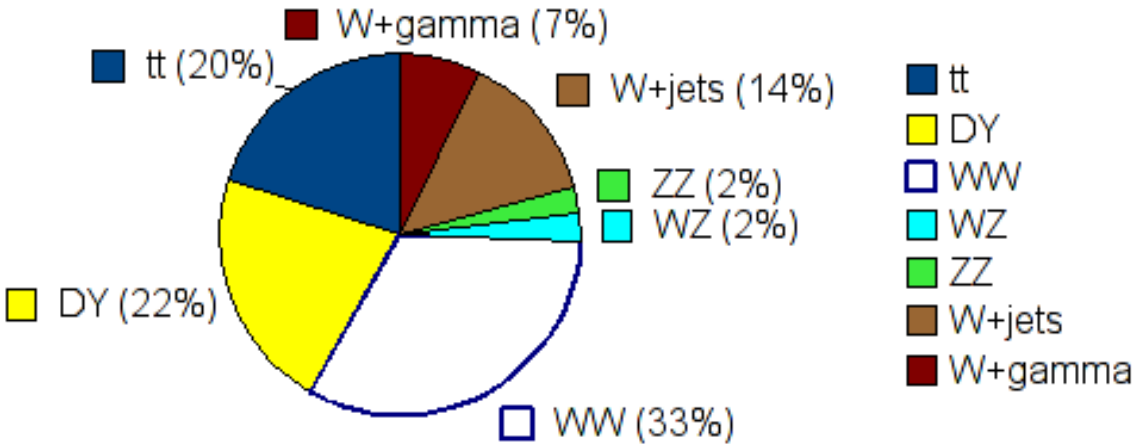


Figure 4.5: Relative contributions of physics processes in the di-lepton sample.

In this Thesis the whole process is often for brevity just mentioned as Monte-Carlo (MC) simulation for a given physics process, since these steps are heavily based on the Monte-Carlo methods to evaluate quantities of interest.

Physics process

The underlying physics process is simulated by using PYTHIA(63) which provides a Leading Order (LO) description of the hard-scattering between partons inside the (anti-)protons. PYTHIA is also used to describe the Leading-Logarithm resummation (showering) for each process in a generic way, which practically generates initial and final state radiation, and the hadronization of color-charged particles. Effects coming from the underlying event are also included in this simulation.

Since these simulations require a big amount of time especially to simulate detector effects, to save computing time we often require at generator level at least two leptons with a minimum p_T of the order of few GeV. Efficiencies for these requirements are taken into account in the normalization of the samples and studies show that they do not significantly bias the sample after analysis selections are applied.

$p\bar{p} \to WZ, ZZ, t\bar{t}$ and Drell-Yan processes are fully simulated using PYTHIA. We use settings which are commonly referred to as Tune A(64). Drell-Yan has been generated with a cut on the di-lepton invariant mass $m(ll) = m(Z/\gamma^*) > 20$ GeV; additional samples are created to fill the gap $10 < m(ll) < 20$ GeV which have a small contribution in the signal regions and negligible effect on the final results.

For the simulation of $W\gamma$ we use the generator described in Ref. (65) for the hard scattering, which provides a better description of the QED radiation. We generate only events with $p_T(\gamma) > 8$ GeV and $\Delta R(lepton, \gamma) > 0.35$.

Since $p\bar{p} \to WW$ is our main background and it heavily contributes to events with at

Process	Generator	$\sigma \times Br$ pb	Accuracy	ϵ_{filter}	Run Period
Main simulation samples					
WW	MC@NLO	12.4	NLO	1.0	0-7
WZ	PYTHIA	3.46	NLO	0.076	0-23
ZZ	PYTHIA	1.51	NLO	0.233	0-23
$Z/\gamma^* \to ll$	PYTHIA	490	NNLO	1.0	0-23
$W\gamma \to l\nu\gamma$	BAUR	18.6	NLO	1.0	0-11
$t\bar{t} \to W(\to l\nu)bW(\to l\nu)b$	PYTHIA	0.809[a]	NLO	1.0	0-23
$gg \to H \to WW$	PYTHIA	0.396[c]	NNLL	1.0	0-23
$qq' \to WH \to WWW$	PYTHIA	0.046[c]	NLO	0.71	0-23
$q\bar{q} \to ZH \to WWZ$	PYTHIA	0.030[c]	NLO	0.72	0-23
$q\bar{q}(') \to q\bar{q}(')H \to q\bar{q}(')WW$	PYTHIA	0.042[c]	NLO	1.0	0-23
Additional simulation samples					
$W \to l\nu$	PYTHIA	2687	NLO	1.0	0-23
WW	PYTHIA	12.4	NLO	1.0	0-23
WW	PYTHIA	12.4	NLO	1.0	0-7
$Z/\gamma^* \to ll$	PYTHIA	920[b]	NNLO	0.016	0-11

[a] This sample includes all the decays into e, μ and τ. It assumes $m_{top} = 172.5$ GeV.
[b] $10 < m(ll) < 20$ GeV at generator level
[c] $m_H = 160$ GeV. Other Higgs samples only differ by the cross section and Branching ratio used in the normalization which are reported for the full sample list in Table 4.4.

Table 4.3: List of physics processes simulated. We also report the event generator, the cross-section times branching fraction used in the normalization, and the generator-level filter efficiency (ϵ_{filter}), if any. The run-period on which the simulation was tuned is also reported, as further discussed in the text. Processes with a final state indicating a generic lepton l are generated separately for e, μ and τ.

least one hard reconstructed jet, we use a full Next-to-Leading Order (NLO) simulation to model the hard-scattering between partons. In particular MC@NLO(66) interfaced with HERWIG(67) for the shower and hadronization parts is used.

The simulations are scaled from the generated cross sections, and branching fractions, to the most recent available calculations, ranging from NLO to Next-to-Next-to-Leading Order (NNLO) accuracy and beyond. Table 4.3 gives a short summary of the simulated samples, with the event generator, the cross section used to normalize them and the efficiency of generator-level filters, if any. We use fourteen Higgs signal samples, which only differ by the Higgs mass hypothesis. The Higgs mass values have been chosen to probe the $110 < m_H < 200$ GeV range with a step of 10 GeV, that is smaller than the expected resolution on the Higgs mass; a step of 5 GeV is used in the region around $m_H = 160$ GeV, where we expect the larger number of signal events. Table 4.4 summarizes the cross sections and branching fractions used to normalize the Higgs signal samples as function of the Higgs mass.

M_H (GeV/c^2)	$\sigma_{gg\to H}$ (pb)	σ_{WH} (pb)	σ_{ZH} (pb)	σ_{VBF} (pb)	$\mathrm{Br}_{\mathrm{H}\to\mathrm{WW}}$
110	1.413	0.208	0.124	0.084	0.044
120	1.093	0.153	0.093	0.072	0.132
130	0.858	0.114	0.071	0.061	0.287
140	0.682	0.086	0.054	0.052	0.483
145	0.611	0.075	0.048	0.048	0.573
150	0.548	0.065	0.042	0.045	0.682
155	0.492	0.057	0.037	0.041	0.801
160	0.439	0.051	0.033	0.038	0.901
165	0.389	0.044	0.029	0.035	0.957
170	0.349	0.039	0.026	0.033	0.965
175	0.314	0.034	0.023	0.031	0.951
180	0.283	0.031	0.021	0.028	0.935
190	0.231	0.024	0.017	0.024	0.776
200	0.192	0.019	0.014	0.021	0.735

Table 4.4: Cross sections and branching fractions used to normalize the Higgs signal samples, as function of the generated Higgs mass of the specific sample.

Detector simulation

The event generator gives as output a set of particles with their 4-momenta and the position coordinates. We then use GEANT3(68) to fully simulate the interaction of those particles with the CDF detector.

Data taking conditions are taken into account by the detector simulation. In particular online beam position, and detector status (especially for the silicon detectors) are reproduced by the simulation. Instantaneous luminosity profile is part of the simulation in order to model at the best the underlying event. This implies that simulated events are tuned to represent data collected in a given time range, even if all of them will then be scaled to represent the full data set analyzed (up to Run Period 23). The last column of Table 4.3 shows the run periods used to tune the simulation. An additional correction to the normalization is derived for the WW(MC@NLO) sample, that was not tuned to the full data set, as the ratio of the selections efficiency in the WW(PYTHIA,0-23) and WW(PYTHIA,0-7) samples. This correction is calculated as function of the reconstructed jets and shown in Table 4.5. The normalization of the $W\gamma$ sample will be measured in data, so that no corrections are attempted at this stage.

N. Jets	A
0	0.945 ± 0.002
1	0.991 ± 0.004
≥ 2	1.031 ± 0.008

Table 4.5: Acceptance scale factor for the time-dependent effects not completely simulated in the WW MCNLO sample. This scale factor is evaluated and applied as function of the number of jets reconstructed in the event.

Trigger and analysis

The output of the detector simulation has the same format as the real data. This means that every selection and reconstruction algorithm applied to the data sample is ran on simulated data.

In order to avoid double counting due to jets identified as leptons and to consider only contributions of real leptons we match each lepton to a true particle (electron, muon or photon) by requiring ΔR between the reconstructed and the generated particle to be less than 0.1.

The generated z of the primary vertex is required to be within ± 60 cm from the center of the detector. We apply a correction factor ϵ_{vtx} measured in minimum bias events as 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 effective trigger efficiency is evaluated for every event as the probability for that event to be triggered as function of the reconstructed triggerable leptons in the event. If only one triggerable lepton is found, then the trigger efficiency for that triggerable lepton is also the trigger efficiency for the event. For events where two triggerable leptons are present we calculate the probability that at least one of these leptons has fired the trigger as $\epsilon^{trg} = 1 - ((1 - \epsilon_1) + (1 - \epsilon_2))$, where $\epsilon_{1(2)}$ are the trigger efficiencies of the two leptons. The prescale that is eventually applied to the corresponding trigger path and the simultaneous presence of 3D and 2D triggers for part of the data taking periods is taken into account[2].

In order to account for possible differences between simulated and collected data on the lepton identification (ID), we apply a scale factor (s^{lep}) which is measured in $Z \to ll$ events (Section 4.3.1).

We take into account these corrections scaling the simulated samples on a per-event basis. At the same time we also rescale the events to represent the full data set, regardless the number of events generated for each sample.

The master formula used to evaluate the weight w given to each simulated event is:

$$w = \frac{\sigma \times Br \times \epsilon_{\text{filter}} \times \epsilon^{\text{trg}} \times s^{\text{lep}} \times \epsilon_{\text{vtx}} \times \mathcal{L}_i}{N_i^{\text{gen}}(|z_0^{P.V.}| < 60 \text{ cm})} \tag{4.2}$$

where

$\sigma \cdot Br$	is the cross-section times branching fraction of the physics process simulated as listed in Tables 4.3 and 4.4
ϵ_{filter}	is the filter efficiency applied to the generation process as listed in Table 4.3
ϵ^{trg}	is the effective trigger efficiency
s^{lep}	is the effective lepton identification scale factor
ϵ_{vtx}	is the efficiency of the z-vertex position requirement ($\lvert z_0^{P.V.}\rvert < 60$ cm)
$\mathcal{L}_i$	is the luminosity of $i-$th the good run list corresponding to the reconstructed leptons
$N_i^{\text{gen}}(\lvert z_0^{P.V.}\rvert < 60 \text{ cm})$	is the number of generated events with $\lvert z_0^{P.V.}\rvert < 60$ cm for the $i-$th good run list corresponding to the reconstructed leptons.

[2]See Appendix B.2 for details.

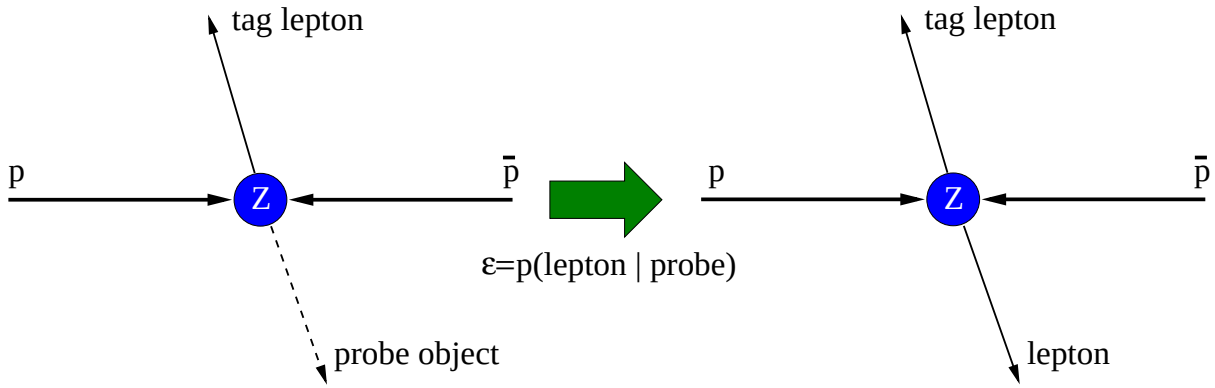


Figure 4.6: Schematic representation of the *tag and probe* method.

4.3.1 Lepton identification scale factors

In this Section we derive the scale factor s^{lep} of Equation 4.2, used to correct simulated events for the discrepancies in the lepton identification efficiencies between real (ϵ^{Data}) and simulated (ϵ^{MC}) data. The correction factor is given by the ratio $s^{\text{lep}} = \epsilon^{Data}/\epsilon^{MC}$ and it is measured for each lepton category and for several groups of run periods.

Real electrons and muons are obtained by selecting $Z \rightarrow ll$ events. We select events with one fully identified lepton (*tag*) and a second object, the *probe*, that passes looser and well simulated requirements. The tag and probe leptons are required to have opposite charge and an invariant mass $\pm 5\sigma$ around the Z peak: $76 < m(ll) < 106$ GeV. We also require $\not\!\!\vec{E}_T < 15$ GeV for $e\mu$ events and $\not\!\!\vec{E}_T < 25$ GeV for the other combinations [3], since we do not expect significant missing transverse energy in Drell-Yan events.

The identification efficiency is measured using the *tag and probe* method. It is measured for each lepton category (l) starting from the probe requirements and evaluating the probability to correctly identify the lepton, as depicted in Figure 4.6. The identification efficiency is:

$$\epsilon^{l} = \frac{N_{\text{l}}}{N_{\text{Probe(l)}}}, \tag{4.3}$$

where N_{l} is the number of leptons l in the selected Z events and $N_{\text{Probe(l)}}$ is the number of probe objects. The residual background due to fake leptons is estimated from the invariant mass sidebands $m(ll) \in [61, 76] \cup [106, 121]$ GeV, and then subtracted from N_{l} and $N_{\text{Probe(l)}}$. Figure 4.7 compares, as example, the invariant mass distribution for real and simulated data with two central electrons and shows the definition of the Z signal and sidebands regions.

Table 4.6 reports the tag and probe leptons used to measure the identification efficiency for each lepton category l. Each lepton probe is defined starting from generic central or forward electron/muon probe objects. The selections used to define the generic probes are listed in Tables 4.7-4.9.

For forward electrons we separately measure the efficiency for calorimetric (PHXPEM) and tracking (PHXTrk) requirements.

[3] this requirement assures that the sample is orthogonal to the ones used to perform the Higgs search.

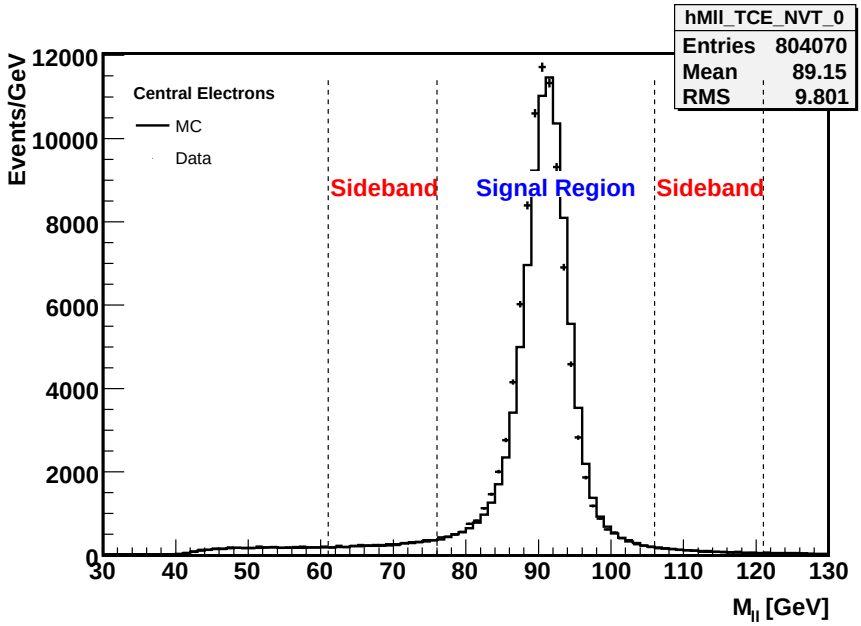


Figure 4.7: Invariant mass distribution of LBE-LBE pairs after the Z selections, comparing data and MC simulation.

The efficiency for CrkTrk leptons is measured separately for electrons and muons. When we apply the correction scale factor to simulated events we use either one or the other depending if the lepton matches an electron or muon at generator-level.

Table 4.10 shows an example of the efficiency measurements and scale factors for run periods 14 to 23. The scale factors for previous run periods differ by less than 10%.

The results show that central electrons have a good modeling in the simulation and their scale factors deviate less than $\sim 2\%$ from unity. The efficiency of the calorimetric requirements for forward electrons is 9% overestimated in our simulations. The central stubbed muons efficiencies tend to be overestimated in our simulation while for central stubless category we have the opposite effect. The isolated track scale factors are pretty similar for electrons and muons and indicate a slight overestimation of our simulation respect to data.

4.3.2 Drell-Yan control region

In this Section the Drell-Yan cross-section measurement is described. The comparison between data and Pythia MC simulations is the strongest check of the overall modeling of the leptons in the MC simulations.

We select Z events using the same selection of Section 4.3.1.

In order to check the agreement of the main kinematic properties of the events, data and MC have been compared. Figures 4.8 and 4.9 show the expected and observed distribution of the number of reconstructed jets (N_{jets}), the di-lepton invariant mass ($m(ll)$), leptons p_T and η, $\Delta R(ll)$ and $\Delta\phi(ll)$, the missing transverse energy ($\vec{\not\!\!E}_T$ and $\vec{\not\!\!E}_T^{\rm Spec}$), the angle $\Delta\phi$ between $\vec{\not\!\!E}_T$ and the nearest lepton or jet and the angle $\Delta\phi$ between

Lepton Category l	Tag	Probe object
Electrons		
LBE	LBE	Central electron probe
PHX	LBE	Forward electron probes
Muons		
CMUP	CMUP	Central muon probe + fiducial to CMU and CMP chambers as in Table 3.9
CMP	CMUP	Central muon probe + fiducial to CMP and not to CMU as in Table 3.10
CMX	CMX	Central muon probe + fiducial to CMX arches and $\rho_{COT} > 140\ cm$ as in Table 3.11
CMXMsKs	CMUP	Central muon probe + fiducial to CMX Miniskirts or Keystone and $\rho_{COT} > 140\ cm$ as in Table 3.11
BMU	CMUP	Forward muon probe + fiducial to BMU, PES and z_{BMU} as in Table 3.12
CMIOCES	CMUP	Central muon probe + fiducial to CES as in Table 3.13
CMIOPES	CMUP	Forward muon probe + fiducial to PES as in Table 3.13
Isolated tracks		
CrkTrk$-\mu$	CMUP	Central muon probe + not fiducial to PES or CES as in Table 3.14
CrkTrk$-e$	LBE	Central muon probe + not fiducial to PES or CES as in Table 3.14

Table 4.6: Tag leptons and probe objects used to measure the efficiency of each of the lepton types.

	Central electron probe
Region	Central (CES fiducial)
Track p_T	≥ 5 GeV
Track Z_0	≤ 60 cm
Conversion	false

Table 4.7: Selections for the central electron probe object.

Forward electron tracking probe	
$\lvert\eta_{det}^{PES}\rvert$	$1.2 \leq \lvert\eta_{det}^{PES}\rvert \leq 2.0$
E_{HAD}/E_{EM}	≤ 0.05
$CalIso$	≤ 0.1
PEM 3×3 χ^2	≤ 10
PES5 $\times$ 9U	≥ 0.65
PES5 $\times$ 9V	≥ 0.65
$\Delta R(PES, PEM)$	≤ 3.0
Forward electron calorimetric probe	
$\lvert\eta_{det}^{PES}\rvert$	$1.2 \leq \lvert\eta_{det}^{PES}\rvert \leq 2.0$
E_{HAD}/E_{EM}	≤ 0.05
Track Z_0	≤ 60 cm
Match	Matched to calorimeter cluster

Table 4.8: Selections for the forward electron tracking and calorimeter probe object.

Central muon probe		Forward muon probe	
Fiduciality	Not PES fiducial	Fiduciality	PES fiducial
NAxL(5 hits)	≤ 3	COT hit fraction	≥ 0.6
NStL(5 hits)	≤ 2	Track z_0	≤ 60 cm
Track z_0	≤ 60 cm		

Table 4.9: Selections for the central and forward muon probe object. Fiduciality requirements are also requested on top of these selections for each muon category (see Table 4.6).

Run Period: 14-23			
	data	MC	Scale Fac
LBE	0.830 ± 0.002	0.842 ± 0.001	0.985 ± 0.002
PHXTrk	0.886 ± 0.002	0.866 ± 0.001	1.023 ± 0.002
PHXPEM	0.772 ± 0.002	0.847 ± 0.001	0.911 ± 0.002
CMUP	0.763 ± 0.004	0.875 ± 0.001	0.872 ± 0.005
CMP	0.784 ± 0.009	0.897 ± 0.002	0.874 ± 0.010
CMX	0.860 ± 0.007	0.887 ± 0.002	0.970 ± 0.008
CMXMsKs	0.795 ± 0.011	0.891 ± 0.002	0.892 ± 0.012
BMU	0.766 ± 0.010	0.696 ± 0.003	1.101 ± 0.015
CMIOCES	0.288 ± 0.003	0.248 ± 0.001	1.162 ± 0.013
CMIOPES	0.560 ± 0.006	0.574 ± 0.002	0.975 ± 0.011
CrkTrk e	0.762 ± 0.006	0.814 ± 0.002	0.936 ± 0.008
CrkTrk μ	0.545 ± 0.004	0.577 ± 0.001	0.945 ± 0.008

Table 4.10: Lepton identification efficiencies for data and simulation for periods 14 to 23. Errors are statistical only.

Category	DY MC events	Data events	Data/MC	DY x-sec (pb) ($\pm$ stat.)	Stat. Residual [σ]
LBE	318064.9	310180	0.98	245.08 $\pm$ 0.45	-1.8
PHX	89023.6	87108	0.98	245.90 $\pm$ 0.85	-0.0
CMUP	141091.4	139411	0.99	248.31 $\pm$ 0.68	3.5
CMP	17345.5	18062	1.04	261.69 $\pm$ 2.01	7.9
CMX	71943.5	69132	0.96	241.48 $\pm$ 0.95	-4.7
CMXMsKs	11586.0	11726	1.01	254.34 $\pm$ 2.42	3.5
BMU	10778.3	10296	0.96	240.06 $\pm$ 2.43	-2.4
CMIOCES	49080.6	50001	1.02	256.02 $\pm$ 1.18	8.5
CMIOPES	21707.0	21272	0.98	246.27 $\pm$ 1.73	0.2
CrkTrk	84097.9	80054	0.95	239.22 $\pm$ 0.87	-7.7
Total	407359.4	398621	0.98	245.91 $\pm$ 0.40	0.0
Theory			1.0	251.31	13.5

Table 4.11: Drell-Yan cross section comparison between data and MC for the di-lepton categories with at least one given lepton type. The *Total* row shows the inclusive cross section measurement. Errors are statistical only.

the $\vec{\not\!\!E}_T$ and the nearest lepton. Systematic uncertainties are also shown, which will be described in detail in Section 5.3. For all of them we find agood agreement between data and simulation, within the systematic uncertainties.

We measure the Drell-Yan cross section (σ^{DY}) counting the number of events that passed the Z selections:

$$\sigma^{DY} = \frac{N^{\text{obs}}}{\mathcal{L} \cdot \epsilon} = \frac{N^{\text{obs}}}{N^{\text{exp}}} \cdot \sigma^{\text{theory}}, \tag{4.4}$$

where $\mathcal{L}$ is the luminosity of the sample, ϵ the selections efficiency, which is evaluated from the MC simulation, and N^{obs} is the number of observed events. The number of expected events from the MC simulation is given by $N^{\text{exp}} = \mathcal{L} \cdot \epsilon \cdot \sigma^{\text{theory}}$, where σ^{theory} is the theoretical cross section used to normalize the simulation; therefore the formula is equivalent to the ratio of N^{obs} and N^{exp} multiplied by σ^{theory}.

We divide the selected sample depending on the pairs of lepton categories reconstructed and measure in each subsample the DY cross section. An example is shown in Figure 4.10 for all the pairs of leptons containing at least one CMUP muon. Results show a reasonable agreement for all categories within few percent. To summarize the cross section results and to check the behaviour of each lepton category, Table 4.11 shows the cross section measurement performed on events with at least one particular type of lepton. The ones that statistically deviate from the inclusive measurement (the *Total* row in the Table) are away few percent. The difference is well covered by our systematic errors, $\sim 5\%$ on average, mainly coming from uncertainties on the trigger efficiencies and the lepton scale factors. Table 4.11 shows also the comparison between the inclusive cross section measurement, obtained summing up all the di-lepton types, and the theory prediction. They well agree within the 6% systematic uncertainty due to the luminosity measurement.

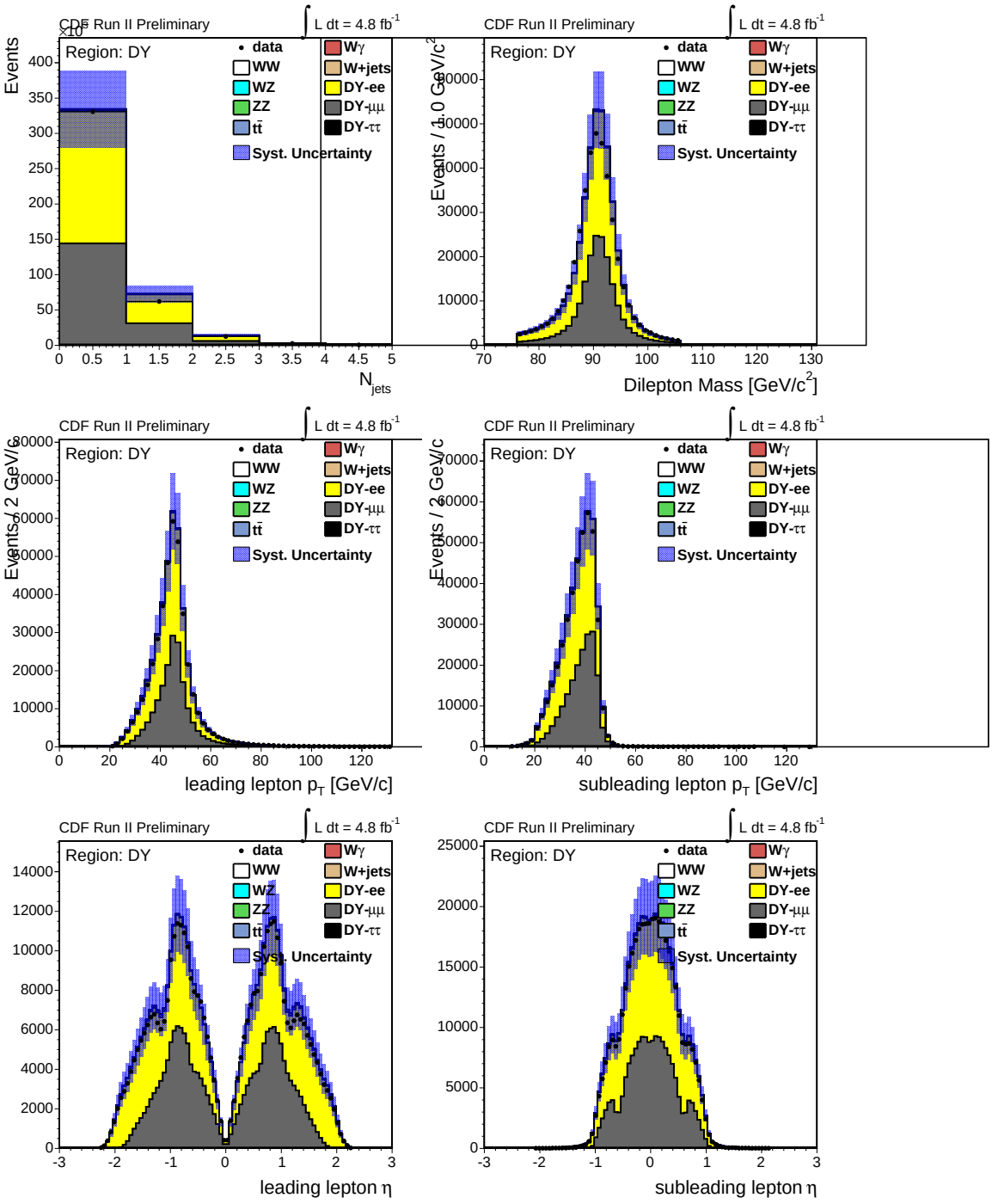


Figure 4.8: Kinematic distributions for events in the Drell-Yan ($Z \rightarrow l^+l^-$) control region.

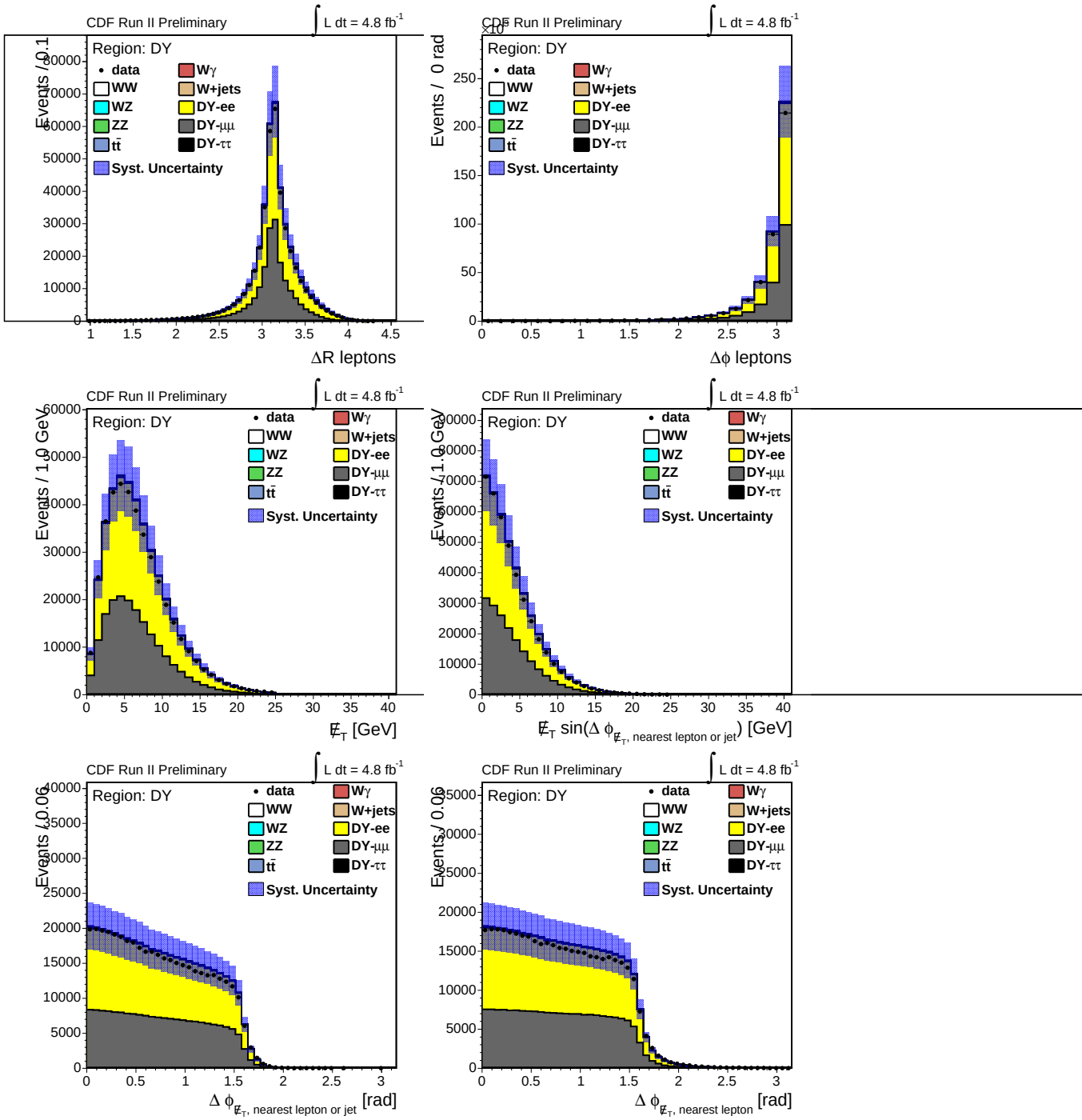


Figure 4.9: Kinematic distributions for events in the Drell-Yan ($Z \rightarrow l^+l^-$) control region.

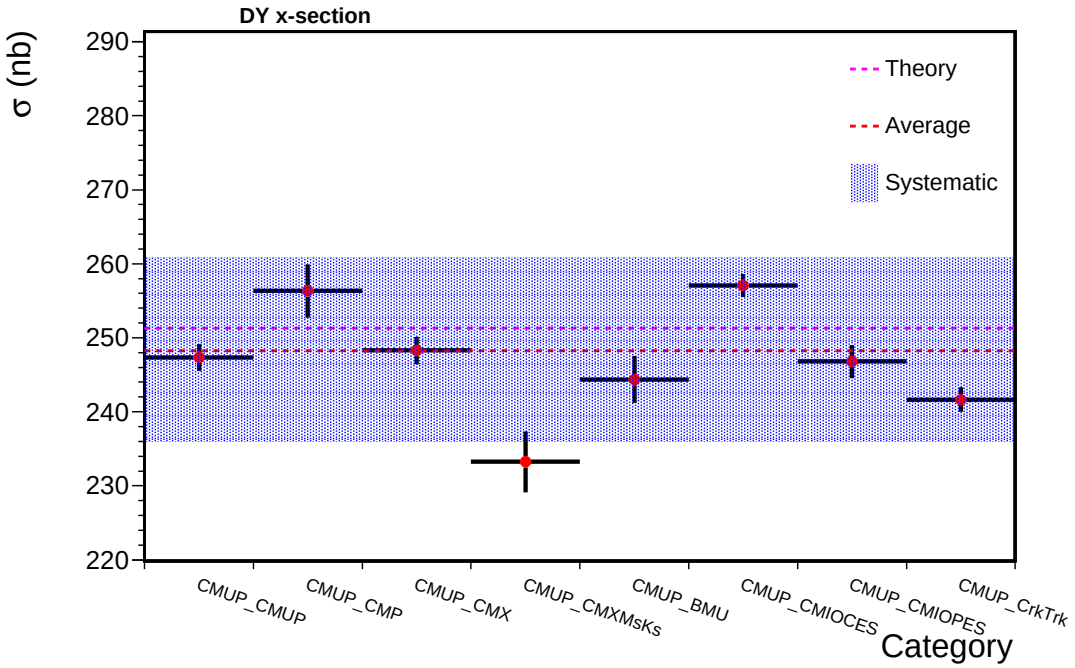


Figure 4.10: The Drell-Yan cross section measurement using events containing at least one CMUP muon.

4.4 Data driven modeling

The fake leptons background is estimated using a data driven method.

The *denominator object* is defined for each lepton category loosening the corresponding identification selections and vetoing the identified lepton. These selections define a sideband in the lepton identification variables. In the case of central electrons (LBE) it is represented by a central track with an energy deposit in the hadronic and electromagnetic calorimeter consistent with being an electron with loose isolation requirements. The selections are reported in Table 4.12. Table 4.13 reports the selections applied to have a denominator object for forward electrons (PHX): a track matched to a forward cluster with loose calorimetric requirements. In the case of muons and isolated tracks denominator objects are defined via common selections listed in Table 4.14. Analogously to the efficiency calculation, central muon denominators will be used for CMUP,CMP,CMX,CMXMsKs,CMIOCES and CrkTrk, while forward muon denominators are defined for CMIOPES and BMU leptons. All of them require a track with quality requirements and loose calorimetric selections. The effective denominator object for each muon category is obtained after the following requirements:

- fiduciality of the corresponding muon category;
- veto of being the corresponding lepton.

The vetoes assure that we select a sample that is orthogonal to the one that contains the corresponding identified leptons.

	Central Electron Denominator
Region	Central (CES)
E_{HAD}/E_{EM}	≤ 0.125 GeV
$CalIso$	< 0.3
Conversion	false
Veto	Not LBE

Table 4.12: Selections of a LBE denominator object.

	Forward Electron Denominator
Region	Plug (PES)
$\lvert\eta_{det}^{PES}\rvert$	$1.2 \leq \lvert\eta_{det}^{PES}\rvert \leq 2.0$
E_{HAD}/E_{EM}	≤ 0.125
$CalIso$	≤ 0.3
NSiHits	≥ 3
Track $\lvert z_0\rvert$	≤ 60 cm
Veto	Not PHX

Table 4.13: Selections of a PHX denominator object.

Central muon denominator	
Fiduciality	Not PES fiducial
NAxL(5 hits)	≤ 2
NStL(5 hits)	≤ 2
Track z_0	≤ 60 cm
Track d_0	≤ 0.2 cm(≤ 0.02 cm $if\ NSiHits > 0$)
Track $\chi^2/n.d.f.$	≤ 3(≤ 4 if run ≤ 186598)
CalIso	≤ 0.3
Forward muon denominator	
Fiduciality	PES fiducial
COT hit fraction	≥ 0.6
Track z_0	≤ 60 cm
Track d_0	≤ 0.2 cm(≤ 0.02 cm $if\ NSiHits > 0$)
Track $\chi^2/n.d.f.$	≤ 3(≤ 4 if run ≤ 186598)
CalIso	≤ 0.3

Table 4.14: Selections of a basic central and forward muon denominator object. Fiduciality requirements and specific vetoes are also requested on top of these selections for each muon category as described in the text.

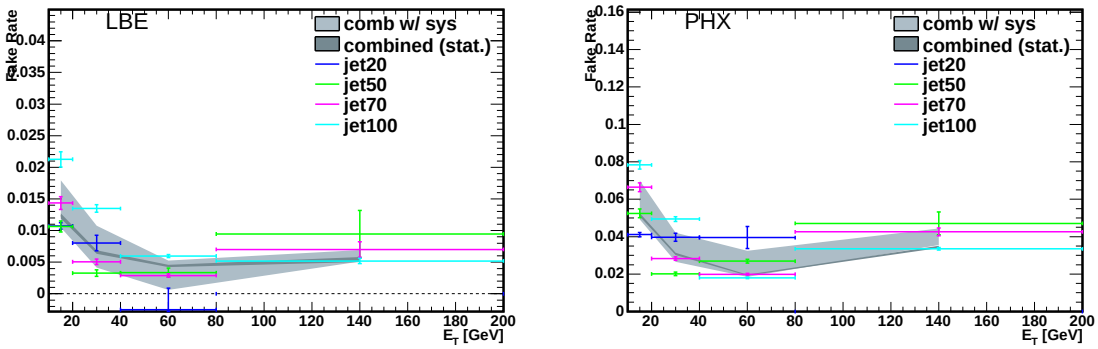


Figure 4.11: Electrons fake rates.

In a jet-enriched sample we then measure the ratio of the number of reconstructed denominator objects to the number of fully identified leptons, for each lepton category and as function of the E_T of the denominator. This number is the *fake rate*. This measurement is performed separately on data collected by JET20,JET50,JET70 and JET100 trigger paths; the weighted average is used while the spread among the various measurements is taken as systematic uncertainty. We correct for the small ($\sim 1\%$) fraction of real leptons in jet data, subtracting the expected number of W and Z evaluated with the simulation. Figures 4.11 and 4.12 show the measured fake rates for each lepton type. Results show that the JET samples cover different ranges in E_T, with significant overlap. The fake rates are of the order of few percents. As expected forward electrons and muons have the highest fake rate values, followed by CMP muons, which are reconstructed in regions not fully instrumented by the central calorimeter.

The fake rates are applied to a data sample consisting of one identified trigger lepton and one or more jets that have passed the denominator requirements. These events are then propagated through the analysis with a weight equal to the appropriate fake probability. If a denominator object can fake different lepton categories they are counted in each one. For each event only the highest E_T denominator object is considered, although multiple events are created in the cases where the denominator object can fake more than one lepton type.

Events with a single real lepton corresponding to any of the non-triggerable lepton types do not appear in the single lepton plus denominator samples used to generate fake events, since we require the lepton to have triggered the event. We obtain a correction for this effect from the inclusive simulated W sample. For each di-lepton category, we determine the ratio of the number of events in which the triggerable lepton originates from the promoted denominator and the number of events in which the real lepton from the W decay is in fact the triggerable lepton ($N^{MC}_{\text{nontrig}}/N^{MC}_{\text{triggerable}}$). The weight of each generated fake event is corrected scaling it up by $1 + N^{MC}_{\text{nontrig}}/\ N^{MC}_{\text{triggerable}}$, so that

$$N^{data}_{\text{total}} = N^{data}_{\text{triggerable}} + N^{data}_{\text{triggerable}} \frac{N^{MC}_{\text{nontrig}}}{N^{MC}_{\text{triggerable}}}. \tag{4.5}$$

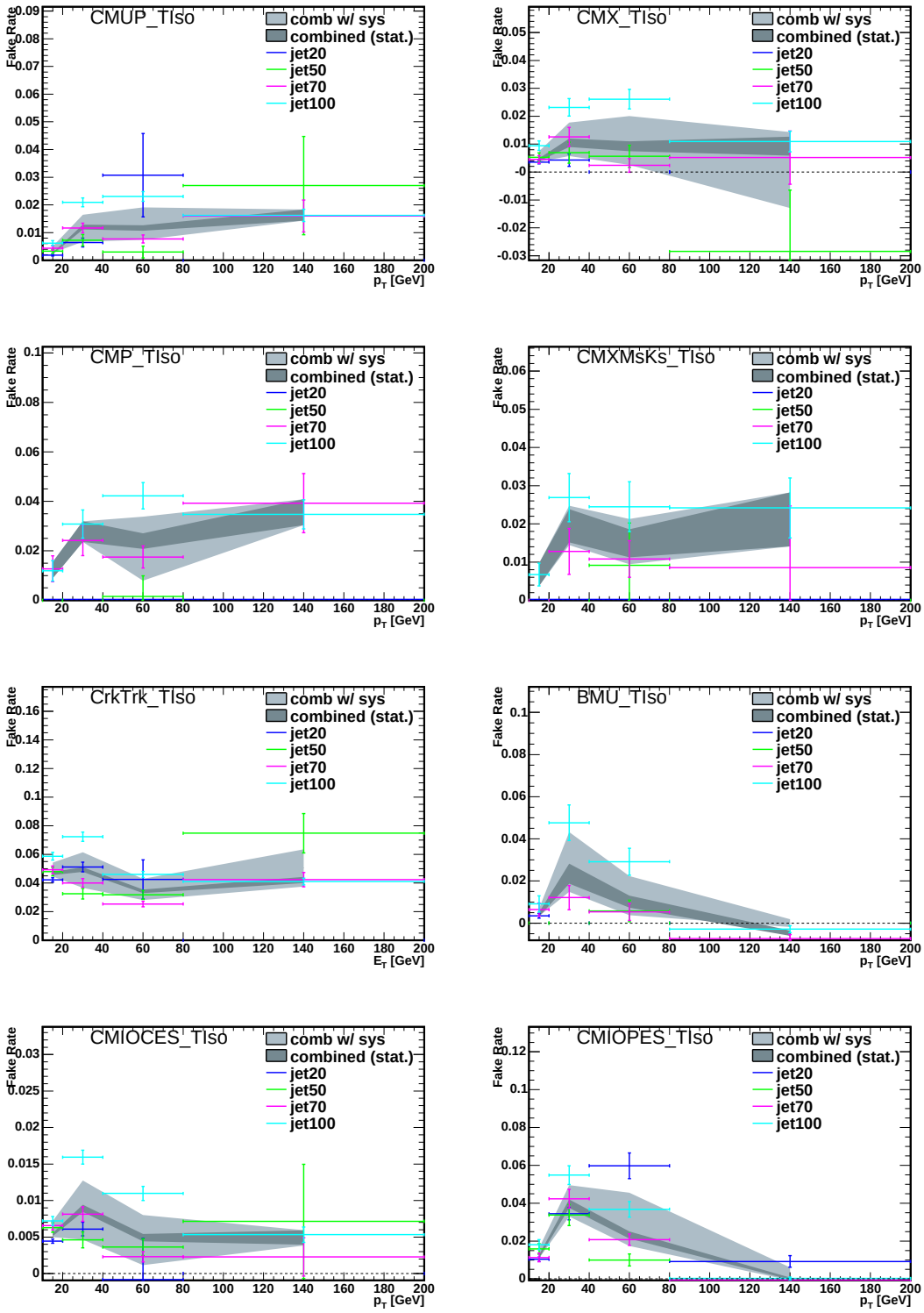


Figure 4.12: Muons fake rates.

Category	WW	WZ	ZZ	$t\bar{t}$	DY	$W\gamma$	W+jets	Total	Data
$e\ e$	4.8	1.2	0.8	0.0	5.2	62.3	50.5	124.7	109
$e\ \mu$	5.8	3.8	0.2	0.0	12.5	53.7	51.3	127.4	110
$\mu\ \mu$	0.0	1.9	0.2	0.0	0.1	0.0	4.2	6.4	6
e trk	1.7	0.9	0.2	0.0	1.9	13.0	9.5	27.3	32
μ trk	0.0	0.9	0.1	0.0	0.4	2.7	4.9	8.9	8
Total:	12.3	8.7	1.5	0.0	20.0	131.7	120.5	294.7	265

Table 4.15: Predicted and observed events in the W+jet control region.

The ratios $\frac{N^{MC}_{\text{nontrig}}}{N^{MC}_{\text{triggerable}}}$ are obtained for di-lepton category and are on the order of few percent, with larger values for the ones containing a CMIOPES or BMU muon.

This method accounts for all the sources of fake leptons coming from jets. The main contributor are $W + jets$ events, then $Z + jets$ gives a significant contribution when one of the leptons is lost. More rare backgrounds as multi-jet production and $WW + jet(s)$ are also included in this estimation.

Two main hypothesis underlay this method:

- no significant contribution of real leptons is in the denominators;
- the kinematic of the underlying physics processes do not depend on identification variables.

The efficiency of the full lepton selections formed by denominators plus specific lepton requirements are pretty high which assures that the number of real leptons which pass the denominator requirements will be small. We can further check both hypothesis directly looking at the predicted and observed yields and kinematic distributions of a control region dominated by fake leptons.

4.4.1 W+jet control region

The model of the W+jets fake backgrounds is checked by looking at events that satisfy all the Ti-lepton selection criteria with the exception of the opposite sign requirement which is reversed. To avoid overlap with the signal region used in Section 6.4 we require no jets in the event[4].

The W+jet control region is dominated by fake leptons originating mainly from jets. This region has significant contribution also from $W\gamma$; the ability of the MC to reproduce it will be checked in Section 4.5. The comparison of expected and observed number of events in the W+jet control region is shown in Table 4.15. We observe good agreement in the observed and predicted numbers of events as well as in the predicted and observed kinematic distributions as shown in Figures 4.13-4.14.

[4]Remember that jets are requested not to be close to identified leptons(Section 3.5). In case of a W+jet(s) event where one of the jets is mis-identified as a lepton, it means vetoing the presence of a second jet.

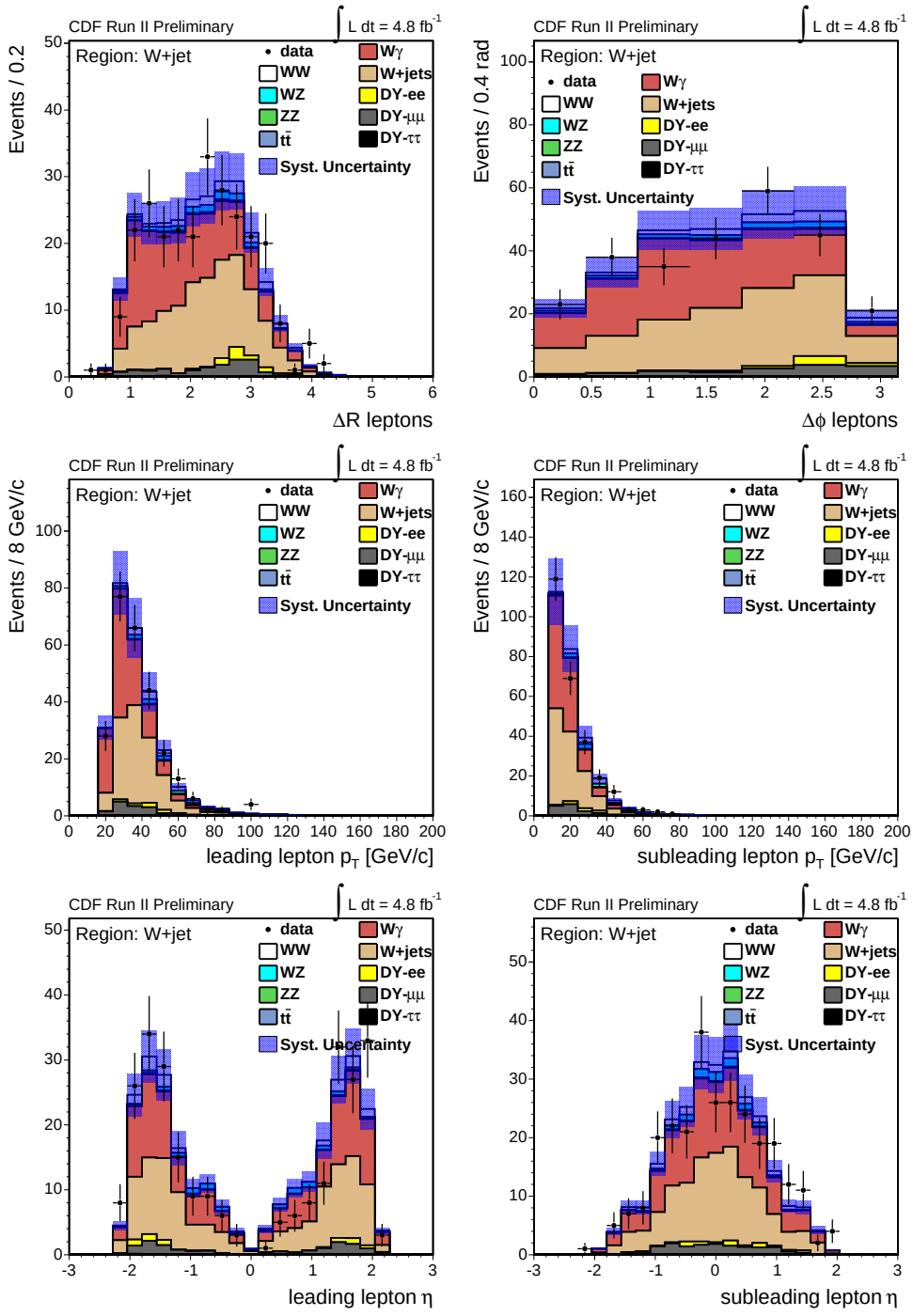


Figure 4.13: Kinematic distributions of the W+jet control region as defined in the Text.

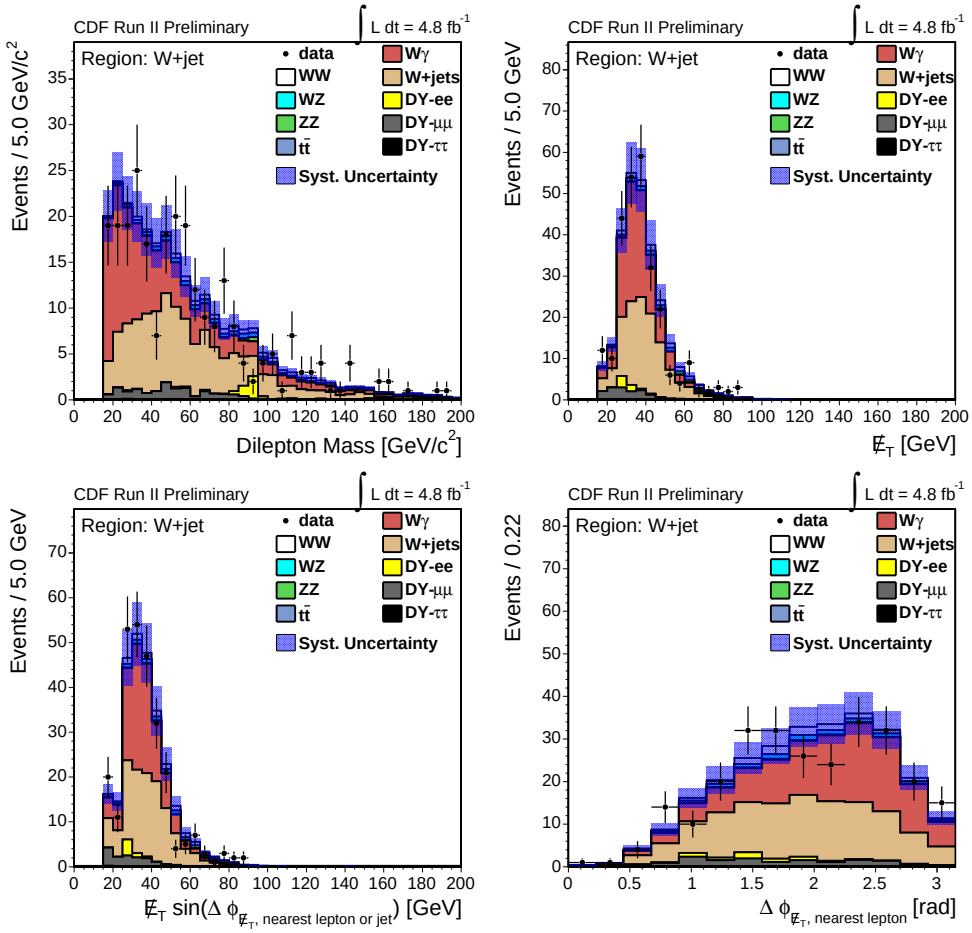


Figure 4.14: Kinematic distributions of the W+jet control region as defined in the Text.

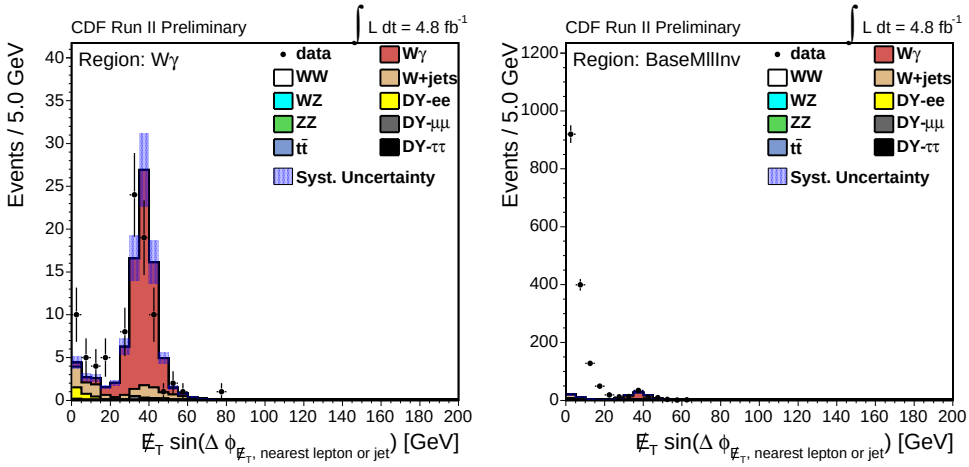


Figure 4.15: $\vec{\not\!\!E}_T^{\rm Spec}$ distribution for events in the $W\gamma$ control region (same sign leptons, Left) and for the same region but with opposite sign leptons (Right), before any requirement on $\vec{\not\!\!E}_T^{\rm Spec}$.

4.5 $W\gamma$ control region

How well the simulations can model fake leptons generated by photons can be verified selecting events that pass all the Dy-lepton selection criteria with two reversed requirements:

- the two leptons must have the same charge: a photon can be mis-identified as lepton without any correlation with the real lepton charge;
- the invariant mass of the leptons $m(ll) < 16$ GeV.

In fact, as can be noticed both from Figure 4.3 and 4.14, $W\gamma$ background tends to be large for low values of the invariant mass of the dig-lepton pair. Only events with zero or one additional reconstructed jets are considered in this region. The $\vec{\not\!\!E}_T^{\rm Spec}$ requirement effectively removes backgrounds from most of the $b-$related sources to negligible levels, and the same-sign requirement removes all the Υ-related ones. This can be seen in Figure 4.15 where the distribution of $\vec{\not\!\!E}_T^{\rm Spec}$ is displayed for opposite and same charged leptons before the $\vec{\not\!\!E}_T^{\rm Spec}$ requirement. The excess of data respect to background expectation is due to the $b-$related backgrounds, which are not simulated. We can see that $\vec{\not\!\!E}_T^{\rm Spec} > 25$ GeV removes this contribution. The result is an highly pure $W\gamma$ region.

The expected and observed number of events in this $W\gamma$ region is shown in Table 4.16. Kinematic distributions are shown in Figures 4.16 and 4.17.

The shapes of the observed kinematic distributions well agree with the expectation. Given the highly pure sample of $W\gamma$ events, we use this region to normalize our expectation for the $W\gamma$ process to reduce the big uncertainty that we have on the simulation of the rate a photon fakes leptons. There are 67 observed same-sign data event and subtracting the total predicted number of non-$W\gamma$ events we end up with 58.0 events in this region. We scale our $W\gamma$ simulation by the ratio r of the observed number of events

Category	WW	WZ	ZZ	$t\bar{t}$	DY	$W\gamma$	W+jets	Total	Data
$e\ e$	0.1	0.0	0.0	0.0	0.1	35.5	3.0	38.7	37.0
$e\ \mu$	0.0	0.1	0.0	0.0	1.1	21.6	2.2	25.1	20.0
$\mu\ \mu$	0.0	0.1	0.0	0.0	0.0	0.0	0.2	0.3	0.0
e trk	0.0	0.0	0.0	0.0	0.2	7.7	1.4	9.2	10.0
μ trk	0.0	0.0	0.0	0.0	0.1	2.0	0.3	2.5	0.0
Total:	0.1	0.3	0.0	0.0	1.5	66.8	7.2	75.8	67.0

Table 4.16: Expected and observed events in the $W\gamma$ control region.

divided by the simulation prediction: $r = 0.86$. The error on r is the quadrature sum of the statistical error and the systematic errors on the subtracted backgrounds prediction: $r = 0.86 \pm 0.10$.

4.6 Low Missing E_T control region

We select events that satisfy all di-lepton selection criteria with the exception of the $\vec{\not\!\!E}_T^{\rm Spec}$ requirement to check the background modeling in the tails of the resolution distribution of the $\vec{\not\!\!E}_T$. We look at the intermediate missing transverse energy region requiring $15 < \not\!\!E_T^{\rm Spec} < 25$ GeV (and $10 < \not\!\!E_T^{\rm Spec} < 15$ GeV for $e-\mu$ events) and zero or one jet reconstructed in the event. Figures 4.18 and 4.19 show the expected and the observed kinematic distributions in this region. These distributions tell us that the dominant process is the Drell-Yan, which pass our selections because the presence of missing energy due to detector resolution effects. This region is used to set systematic uncertainties on $\vec{\not\!\!E}_T$ modeling. The distribution of the angles between $\vec{\not\!\!E}_T$ and the nearest lepton and between leptons show a poor agreement between data and simulations. The disagreement is covered by the systematic uncertainties.

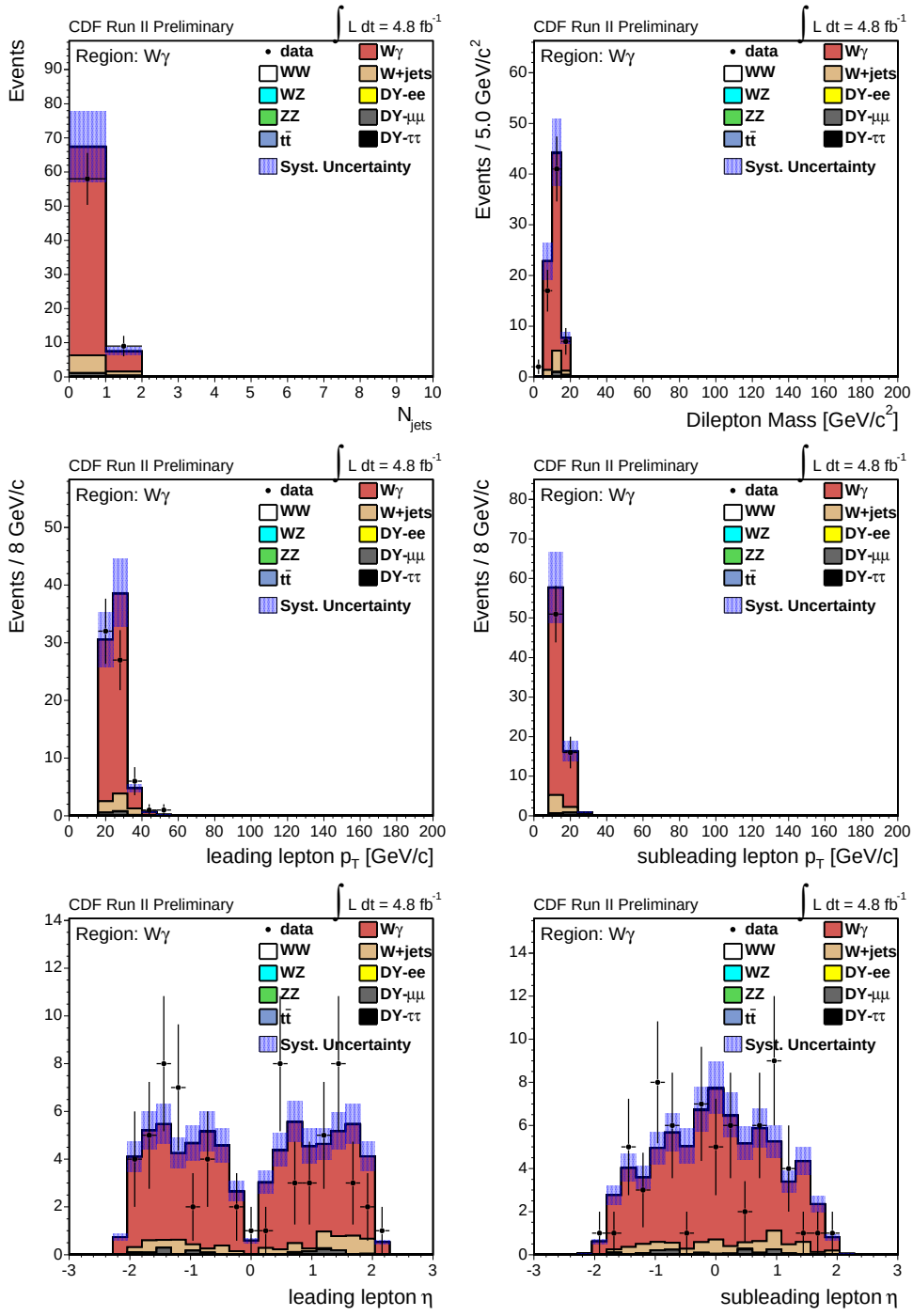


Figure 4.16: Kinematic distributions of events in the $W\gamma$ control region.

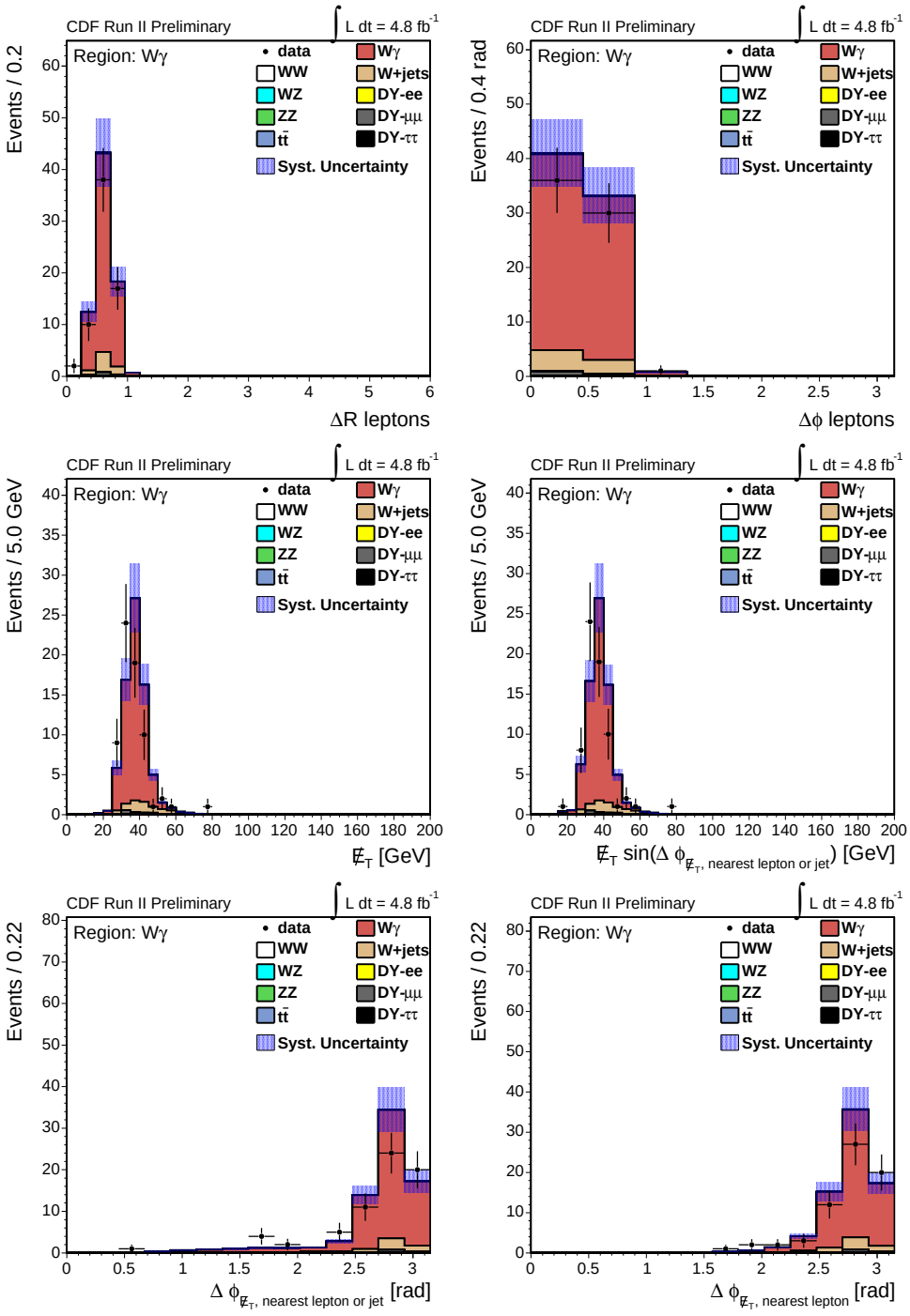


Figure 4.17: Kinematic distributions of events in the $W\gamma$ control region.

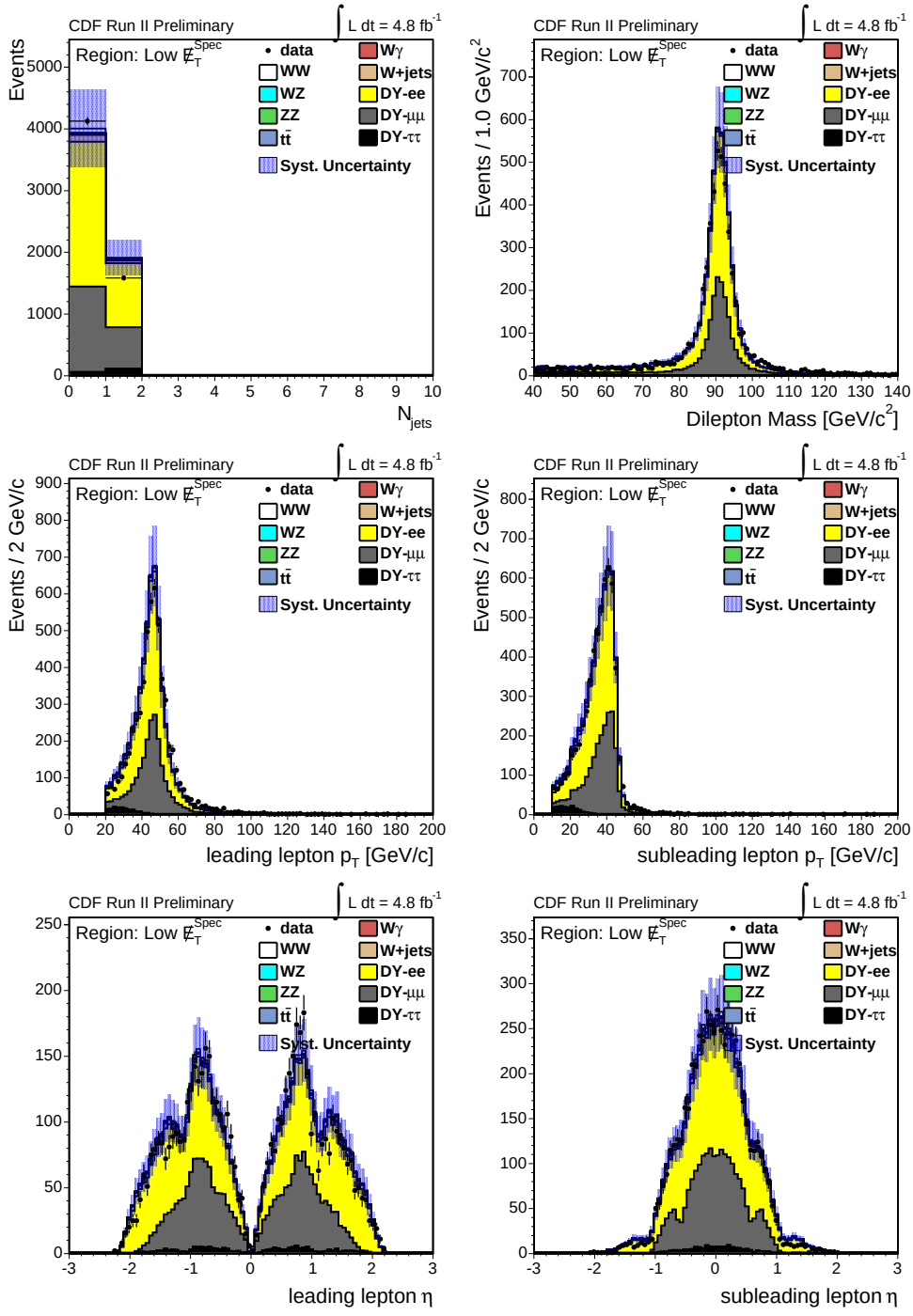


Figure 4.18: Kinematic distributions for events in the Low-$\vec{\not\!\!E}_T^{\mathrm{Spec}}$ control region.

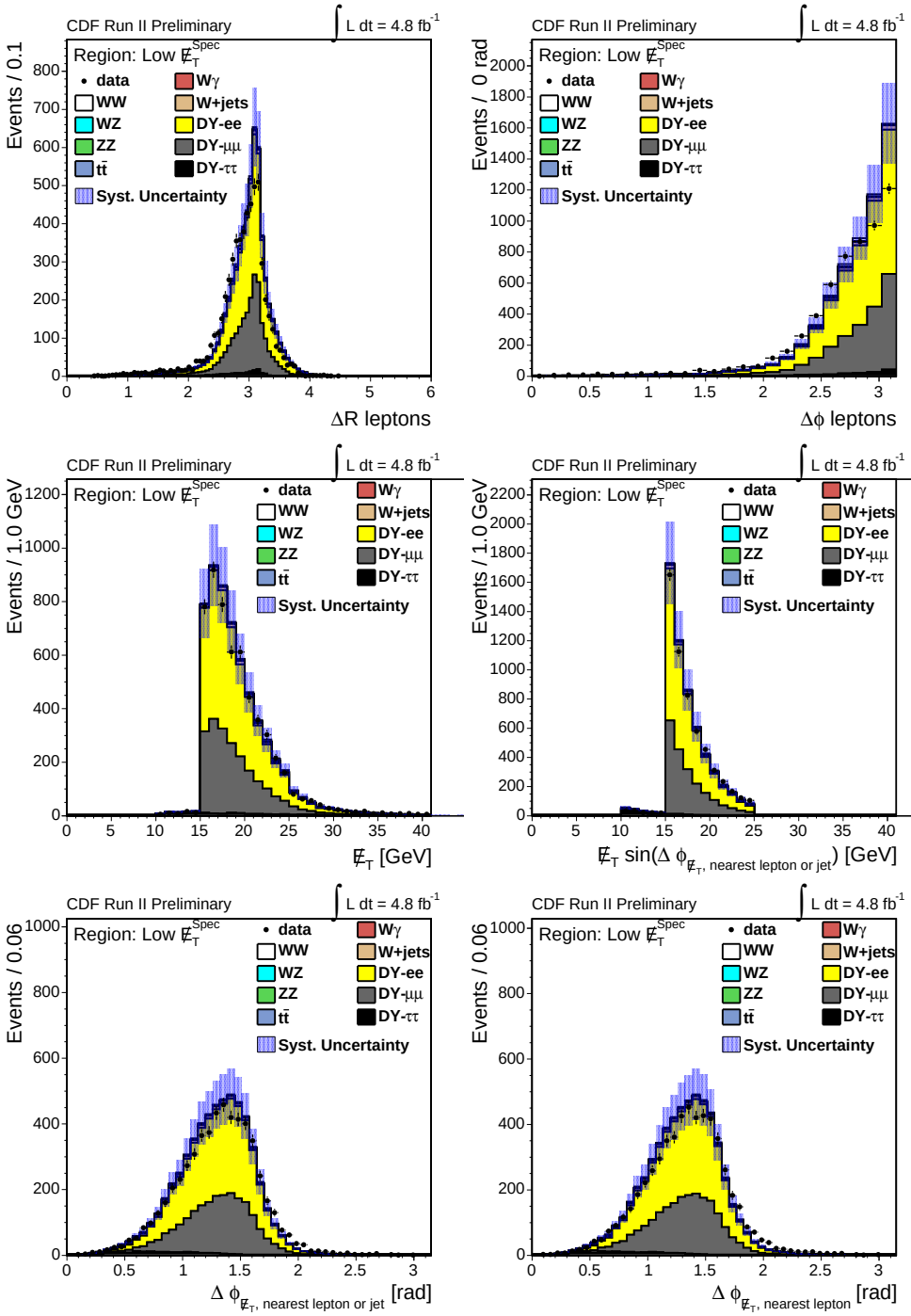


Figure 4.19: Kinematic distributions for events in the Low-$\vec{\not\!\!E}_T^{\rm Spec}$ control region.

Chapter 5

$gg \to H$: Analysis of events with low jet activity

We discussed in Section 2.4 that different Higgs production mechanisms produce different number of jets in the final state. As a consequence the selection criteria are optimized as a function of the number of jets, causing a different background composition.

In this Chapter we focus on the search for the Higgs boson produced through gluon fusion. This process has no jets at Leading Order. Therefore events with opposite sign leptons and zero reconstructed jets are selected.

5.1 Sample composition and kinematic

The expected background and the observed yield in the opposite-sign di-lepton signal region, defined in Section 4.1, for events containing no reconstructed jets with $E_T > 15$ GeV are shown in Table 5.1. Table also reports the expected signal contribution for direct Higgs production with $m_H = 165$ GeV.

The expected and the observed kinematic distributions for these events are shown in Figures 5.1 and 5.2. We stack the backgrounds on top of each other, and then for comparison we overlay the contribution of a 165 GeV Higgs boson with yields multiplied by a factor of 10.

The non-resonant production of WW contributes 50% to the selected sample. Fake leptons (W+jets and $W\gamma$) and Drell-Yan are important contributors too. Observed data are in very well agreement with the Standard Model background expectations.

The 95% C.L. sensitivity for a counting experiment to the $H \to WW$ process cross section is obtained using the procedure described in Section 5.4. From the expected signal and background events of Table 5.1, the sensitivity is $16.6^{+6.2}_{-5.0}$ times the Standard Model Higgs cross section, including the systematic uncertainties quoted in the Table and described in Section 5.3. In order to improve the sensitivity to the Higgs signal, we exploit the differences in the kinematic of the $H \to WW \to l\nu l\nu$ process compared to backgrounds.

We first divide the sample into two sub-samples: High S/B and Low S/B, to isolate events with an higher contribution from fake leptons. We assign to the Low S/B region

CDF Run II Preliminary $\int \mathcal{L} = 4.8$ fb^{-1}			
$M_H = 165$ GeV/c^2			
$t\bar{t}$	1.99	$\pm$	0.31
DY	128	$\pm$	30
WW	447	$\pm$	48
WZ	19.7	$\pm$	2.7
ZZ	29.9	$\pm$	4.1
W+jets	154	$\pm$	37
$W\gamma$	112	$\pm$	19
Total Background	893	$\pm$	79
$gg \to H$	12.6	$\pm$	1.7
Total Signal	12.6	$\pm$	1.7
Data	950		

OS 0 Jets

Table 5.1: Expected signal and background events in the opposite sign di-lepton sample with zero jets. Systematic errors discussed in Section 5.3 are included.

all the events with at least one forward electron (PHX) or one forward stubless muon (CMIOPES), and categories with the highest fake rates (CMP). Events with both an LBE and CMIOPES are considered in the High S/B region since LBE category has a small fake rate. Table 5.3 shows the expected and observed number of events in these two regions separately. We find good agreement between the observed and expected number of events in both regions. The Low S/B region slightly underestimate the background, but within the statistical and systematic uncertainties that will be discussed in Section 5.3. These Tables also show how we effectively separate the background coming from fake leptons, which is $\sim 17\%$ of the total expectation in the High S/B region and $\sim 50\%$ in the Low S/B region. The number of expected signal events as function of the Higgs mass is reported in Table 5.2.

In order to discriminate $H \to WW$ from the non-resonant WW production, the Higgs boson spin-0 property plays an important role(69). The Ws originate from the decay of a scalar and have opposite spin orientation. Due to the $V - A$ structure of the weak leptonic decay of the W, the left handed e^- (right handed e^+) is emitted along the W^- (W^+) spin. As result, one of the two charged leptons is emitted along the momentum direction of the W while the other one goes in the opposite direction (in the rest frame of the corresponding W). For the considered Higgs mass range, in the laboratory reference frame a small opening angle between the two charged leptons is expected for signal events while the backgrounds tend to have back-to-back charged leptons in the final states. This is visible in Figure 5.1, that shows the $\Delta\phi$ distribution between leptons. As direct consequence neutrinos have small $\Delta\phi$ giving an high transverse energy imbalance in the event. This can be seen in Figure 5.2 in the $\vec{\not\!\!E}_T$ and $\vec{\not\!\!E}_T^{\text{Spec}}$ distributions. Figure 5.2 also shows how signal leptons p_T spectra tend to be harder than Drell-Yan and fake leptons. In the same Figure we can notice that the invariant mass of the two leptons is peaking at

Mass (GeV/c^2)	gg→H		
	All	High S/B	Low S/B
110	0.70 ± 0.09	0.57 ± 0.08	0.13 ± 0.02
120	2.36 ± 0.32	1.91 ± 0.26	0.45 ± 0.06
130	5.20 ± 0.70	4.19 ± 0.57	1.01 ± 0.14
140	8.3 ± 1.1	6.73 ± 0.91	1.57 ± 0.21
145	9.4 ± 1.3	7.7 ± 1.0	1.76 ± 0.24
150	10.7 ± 1.4	8.7 ± 1.2	1.98 ± 0.27
155	12.0 ± 1.6	9.83 ± 1.3	2.19 ± 0.30
160	12.9 ± 1.7	10.6 ± 1.4	2.31 ± 0.31
165	12.6 ± 1.7	10.4 ± 1.4	2.24 ± 0.30
170	11.5 ± 1.6	9.49 ± 1.3	2.03 ± 0.27
175	10.2 ± 1.4	8.35 ± 1.1	1.89 ± 0.26
180	9.16 ± 1.2	7.50 ± 1.0	1.66 ± 0.22
190	6.22 ± 0.84	5.04 ± 0.68	1.18 ± 0.16
200	4.87 ± 0.66	3.96 ± 0.53	0.92 ± 0.12

Table 5.2: Expected number of signal events for several Higgs mass hypotheses.

values lower than those obtained from back-to-back leptons, for the small angle between leptons momenta. On the other hand, given the mass of the Higgs boson, the scalar sum of leptons energy and the missing transverse energy (H_T) is expected higher for signal than those obtained from other background sources (except for $t\bar{t}$, which is anyway not significant in events with no reconstructed jets).

Figure 5.3 shows how these kinematic properties vary for signal events with different Higgs mass hypothesis. We show normalized distributions to compare their shapes. The most interesting property is that the opening angle between leptons has the lowest average value for the Higgs mass $m_H \sim 160$ GeV, with both W being on-shell. In this case it offers the best discrimination power for signal events from backgrounds. For a lighter of heavier Higgs boson the distributions peak at higher values and become broader than for $m_H \sim 160$ GeV. The H_T and di-lepton invariant mass distributions peak at more and more high values as the Higgs mass increase. For heavy mass hypothesis the boost of the W bosons is larger than for lighter ones, and determines a large $\not\!\!\vec{E}_T$, increasing the separation of signal from backgrounds.

5.2 Multivariate analysis: Matrix Elements and Neural Networks

In order to exploit the differences in the kinematic properties of the signal respect to the backgrounds and to improve their separation, we use and then combine two different approaches:

- Matrix Element (ME) technique;

CDF Run II Preliminary $\int \mathcal{L} = 4.8$ fb^{-1}			
$M_H = 165$ GeV/c^2			
$t\bar{t}$	1.67	$\pm$	0.26
DY	82	$\pm$	19
WW	336	$\pm$	36
WZ	14.2	$\pm$	1.9
ZZ	22.4	$\pm$	3.1
W+jets	61	$\pm$	17
$W\gamma$	34.3	$\pm$	5.7
Total Background	551	$\pm$	52
$gg \to H$	10.4	$\pm$	1.4
Total Signal	10.4	$\pm$	1.4
Data	561		

OS 0 Jets, HighSB

CDF Run II Preliminary $\int \mathcal{L} = 4.8$ fb^{-1}			
$M_H = 165$ GeV/c^2			
$t\bar{t}$	0.319	$\pm$	0.050
DY	47	$\pm$	11
WW	111	$\pm$	12
WZ	5.54	$\pm$	0.75
ZZ	7.5	$\pm$	1.0
W+jets	93	$\pm$	20
$W\gamma$	78	$\pm$	13
Total Background	341	$\pm$	31
$gg \to H$	2.24	$\pm$	0.31
Total Signal	2.24	$\pm$	0.31
Data	389		

OS 0 Jets, LowSB

Table 5.3: Expected signal and background events in the High S/B (Top) and Low S/B (Bottom) regions for the opposite sign di-lepton sample with zero jets. Systematic errors discussed in Section 5.3 are included.

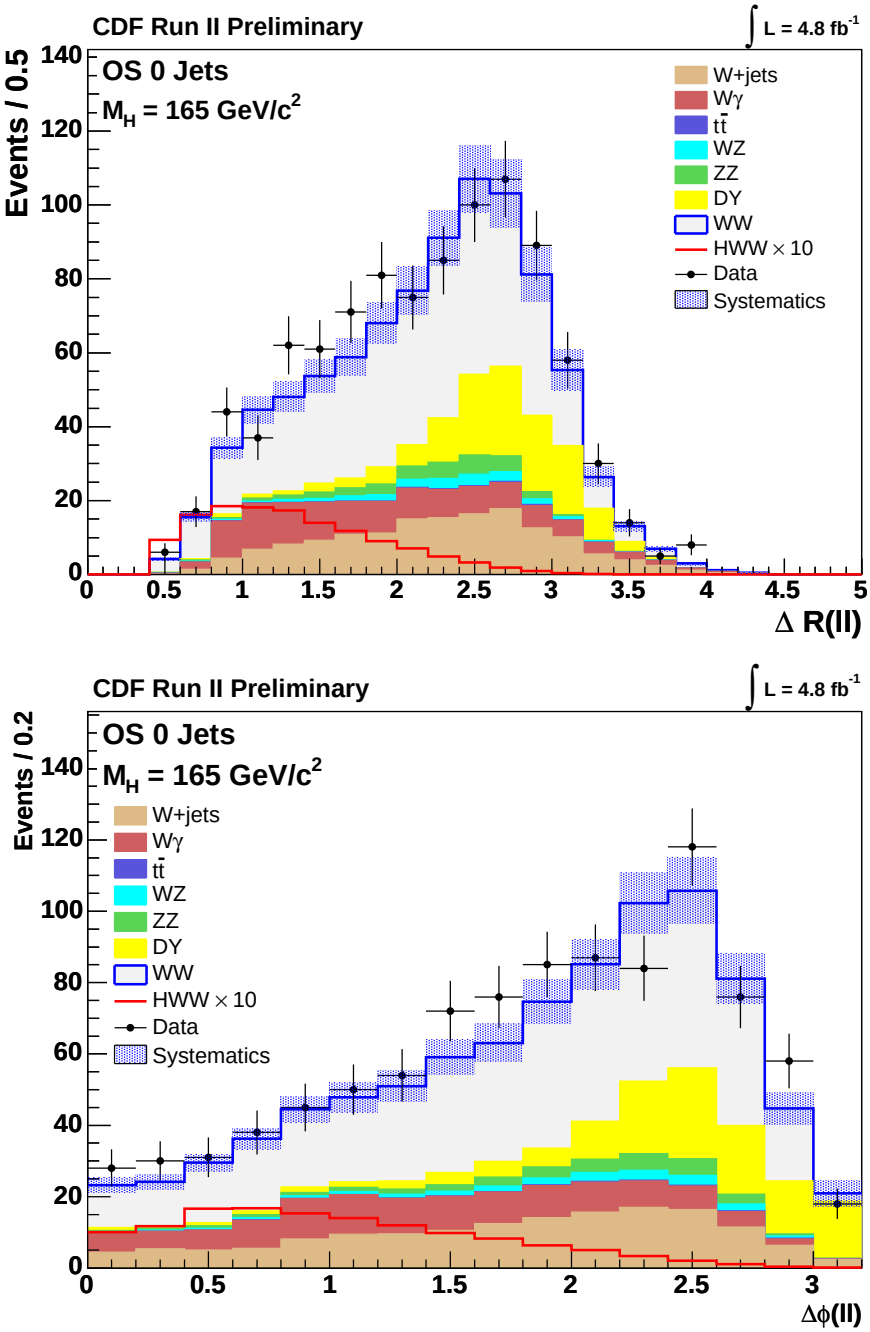


Figure 5.1: Expected and observed distributions of ΔR and $\Delta\phi$ between leptons in the opposite-sign region containing no reconstructed jets with $E_T > 15$ GeV.

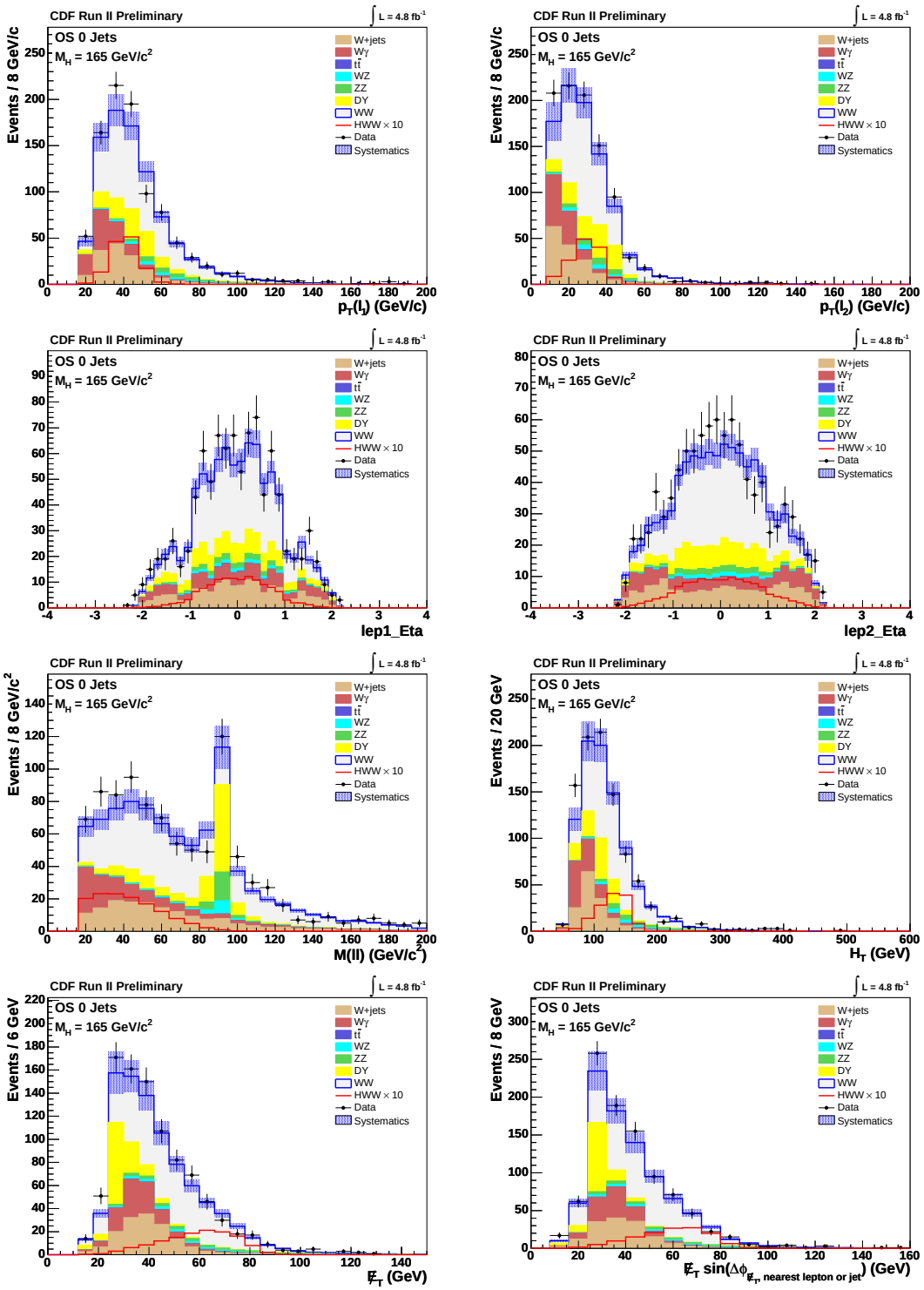


Figure 5.2: Signal and background distributions in the opposite-sign signal region for events containing no reconstructed jets with $E_T > 15$ GeV.

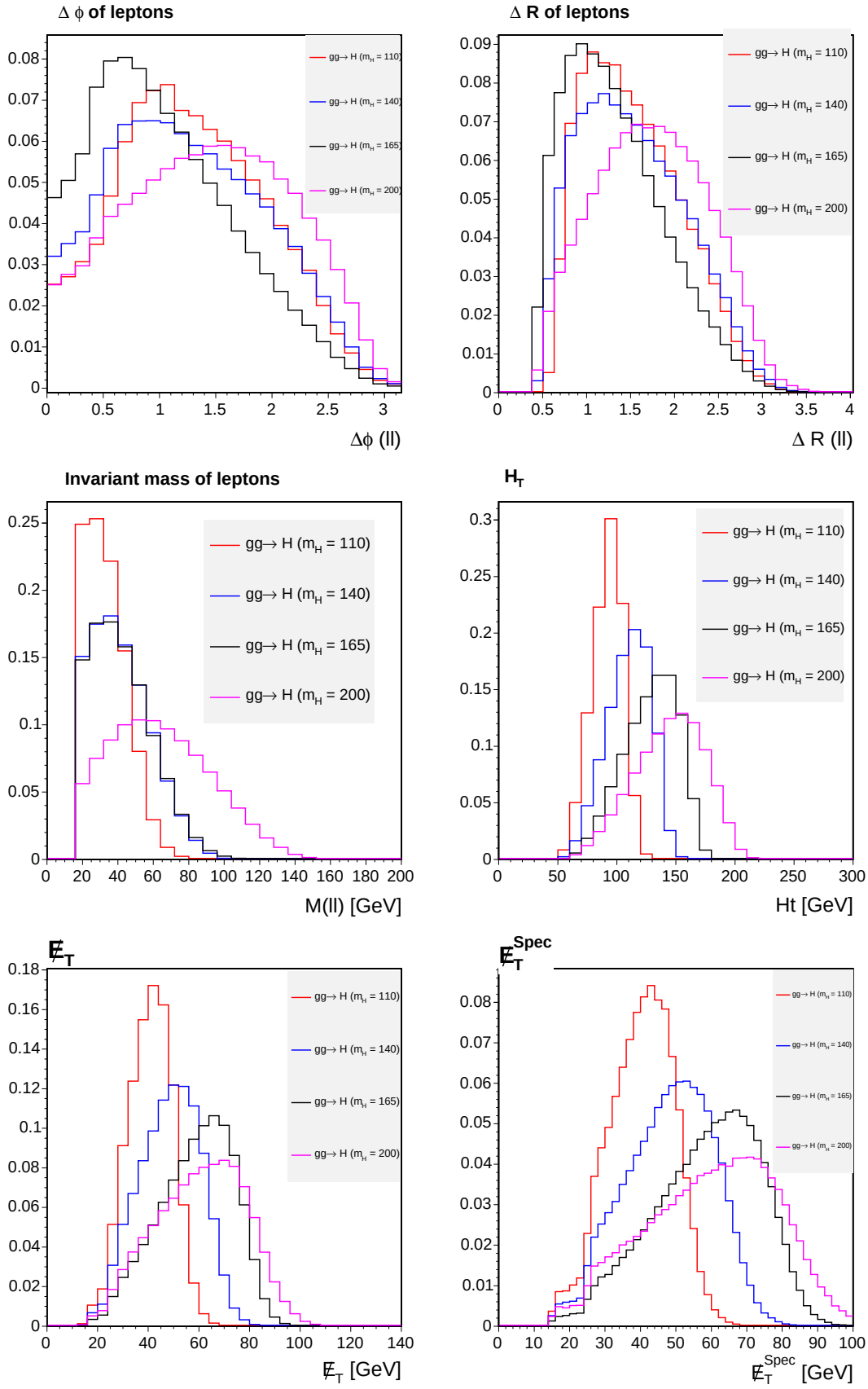


Figure 5.3: Unit-normalized kinematic distributions for signal events with different Higgs mass hypotheses: $m_H = 110$ GeV, $m_H = 140$ GeV, $m_H = 165$ GeV, $m_H = 200$ GeV.

- Artificial Neural Networks.

Both techniques end up assigning to each given event the probability to be compatible with signal or background processes. This probability is then evaluated for the expected and the observed events. The binned probability distributions are used for searching the Higgs boson.

5.2.1 Matrix Element technique

The principle behind this technique(70) is to compare the measured 4-momenta of the observed final state particles with the theoretical expectations obtained with a Leading-Order (LO) calculation (Matrix Element), taking into account the acceptances, the efficiencies and the resolution effects of the detector. The ME technique assigns to each event the probability that its kinematic is consistent with a given process based on the Matrix Element calculation of that process. The event probability for a process X is given by:

$$P_X(\vec{x}_{\text{obs.}}) = \frac{1}{\sigma^{\text{LO}}} \int \frac{\mathrm{d}\sigma^{\text{LO}}(\vec{y})}{\mathrm{d}\vec{y}}\ \epsilon(\vec{y})\ \mathrm{G}(\vec{x}_{\text{obs.}}, \vec{y})\ \mathrm{d}\vec{y}, \tag{5.1}$$

where $\vec{x}_{\text{obs.}}$ is the observed 4-momenta of the final state particles; σ^{LO} is the total production cross-section of the process X at Leading-Order, obtained from MCFM(31), and it is used as a normalization factor:

$$\int P_X(\vec{x}_{\text{obs.}})\mathrm{d}\vec{x}_{\text{obs.}} = 1. \tag{5.2}$$

The integral of Eq. 5.1 is the convolution of $\frac{\mathrm{d}\sigma^{\text{LO}}(\vec{y})}{\mathrm{d}\vec{y}}$, the Leading-Order differential cross section in function of the "true" 4-momenta of the final state particles, and the term $\epsilon(\vec{y}) \cdot \mathrm{G}(\vec{x}_{\text{obs.}}, \vec{y})$, the efficiency and detector resolution distribution functions which are tuned using the GEANT simulation of the CDF detector (see Section 4.3). The Matrix Elements for the calculations of $\frac{\mathrm{d}\sigma^{\text{LO}}(\vec{y})}{\mathrm{d}\vec{y}}$ are obtained from Ref. (71); the phase space factor is evaluated explicitly for each process and the parton distribution functions used for the (anti-)proton are the CTEQ5L(72).

The functional form of the probability (Eq. 5.1) of the processes: $H \to WW \to l\nu l\nu$, $WW \to l\nu l\nu$, $ZZ \to ll\nu\nu$, $W\gamma \to l\nu\gamma$ and $W + q \to l\nu q$ (usually referred as $W + jet$) is evaluated to discriminate among them.

For the signal process, we evaluate P_{HWW} as function of the Higgs mass (and the corresponding width) in the cases of on-shell and off-shell W final state.The charged leptons 4-momenta and the $\vec{\not\!\!E}_T$ are used as observables ($\vec{x}_{\text{obs.}}$) while we integrate over the remaining unobserved degrees of freedom. The same choice for $\vec{x}_{\text{obs.}}$ is made for WW and ZZ processes. The photon or the jet in $W\gamma$ and $W + jet$ events fakes the lepton, therefore we assign the measured leptons momenta to the lepton and to the photon/jet of the process. In this case there is only one neutrino in the final state, and we assume $\vec{\not\!\!E}_T$ to be a measurement of its transverse momentum.

These probabilities are then combined to form Likelihood-ratios in order to separate a given process (X) respect to the others:

$$LR(X) = \frac{P_X}{P_X + \sum_{i \neq X} k_i\ P_i}, \tag{5.3}$$

where the sum runs over all the listed processes and k_i is the relative yield of the process i to the other backgrounds such that $\sum_i k_i = 1$.

Figure 5.4 shows the expected and observed distributions of the Likelihood Ratios used to distinguish WW, ZZ, $W\gamma$ and $W+jet$. We show separately the High S/B and Low S/B regions of the opposite-sign di-lepton sample with no reconstructed jets. Figure 5.5 shows the expected and observed distributions of $LR(HWW)$ for $m_H = 165$ GeV. The Figure 5.6 shows the distributions as evaluated for few of the fourteen Higgs masses considered. The data distributions agree within uncertainties with background expectations.

In order to evaluate the separation between signal and background achieved with this technique and to compare to the simple counting experiment on the whole sample performed in Section 5.1, we select events in the highest bin of the $LR(HWW)$ distribution for the most favorable High S/B region. We expect a contribution of 3.17 ± 0.43 signal events for $m_H = 165$ GeV and 7.84 ± 0.79 events from SM backgrounds. Errors include the systematic uncertainties that are discussed in Section 5.3. The 95% C.L. sensitivity for a counting experiment on the selected events is $2.7^{+1.0}_{-0.8}$ times the SM Higgs cross section, corresponding to an improvement of a factor six compared to the sensitivity evaluated in Section 5.1.

It is important to point out that any approximation used in the P_X calculation reveals in a worse separation power of the LR variables, and it cannot per-se originate bias in the result.

This technique powerful separate the various components because the signal and the main backgrounds have no jets in the final state at Leading Order, and the calculation implemented in the Matrix Element technique properly describe the kinematic of these processes.

5.2.2 Artificial Neural Networks

Artificial Neural Networks (NN) are a well known method(73) to combine information into a single discriminant. The principle is to find the function f of the kinematic variables of each event that statistically maximize the separation between signal and backgrounds. The value assigned by f to each event is called the *score*. Ideally we want the function f of the input quantities to tell us if a given event is compatible with signal or backgrounds, with a score of $+1$ for signal and -1 for backgrounds. Neural Networks basically provide a robust approach to approximate f.

We use a *feed-forward* Neural Network. The structure is schematically shown in Figure 5.7. The building blocks, *nodes*, are functions $f_i : \mathcal{R} \to \mathcal{R}$ chosen to be a symmetric sigmoid function:

$$f_i(x) = \frac{2}{1 + \mathrm{e}^{-x}} - 1, \tag{5.4}$$

where x is the input of i-th node, while $f_i(x)$ is the output. These functions basically acts as a smooth step function and map $[-\infty, +\infty] \to [-1, 1]$. In the Figure 5.7 each line represents how the output of each node *feed* the input of the next one. When several lines enter the same node the input to that node is a linear combination of all the inputs I_i: $x = \sum_i w_i I_i$. w_i are weights which are the tunable parameters used to set the behavior of the NN. The network used can be divided in *layers*, where all the nodes in the same layer

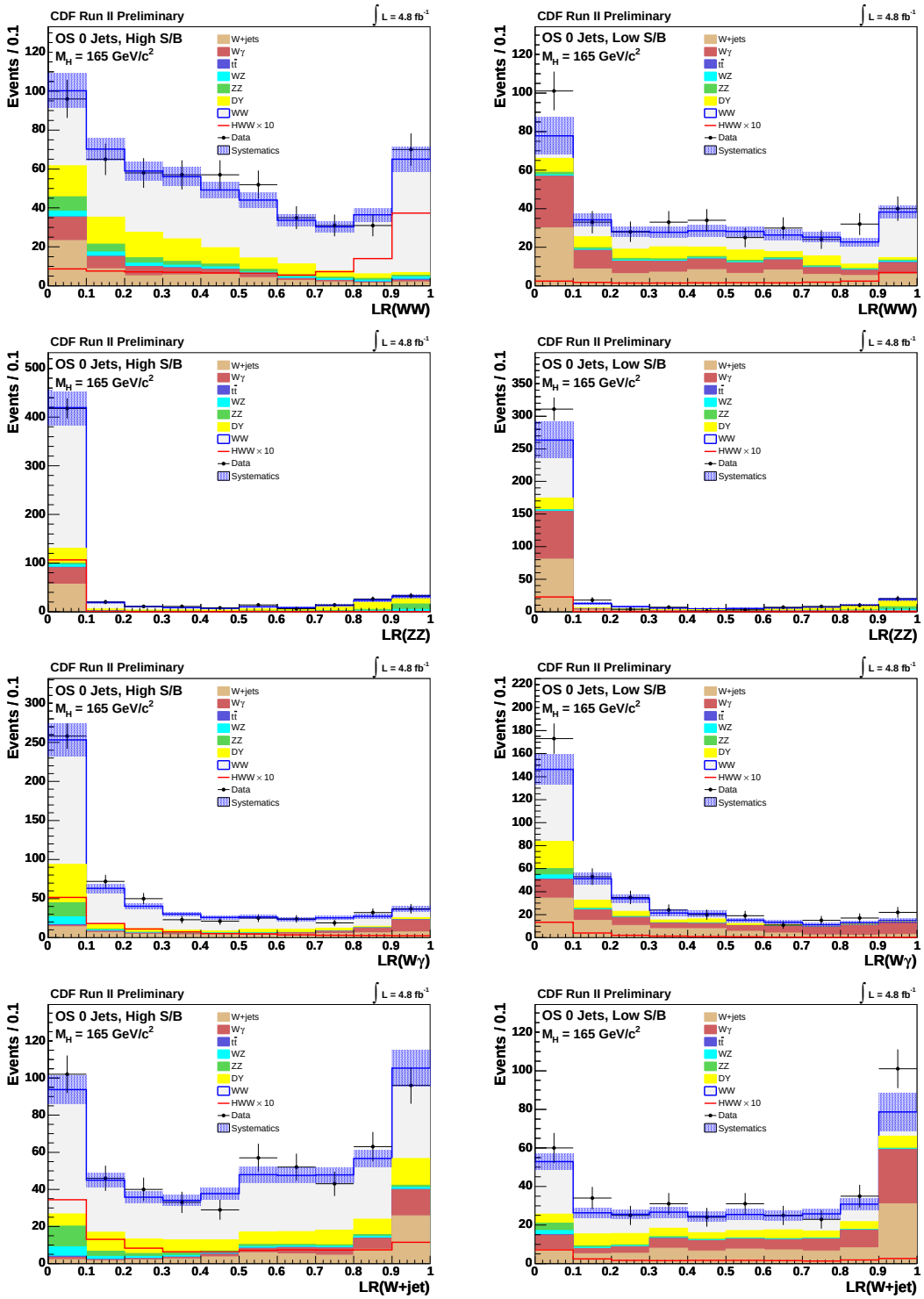


Figure 5.4: The expected and observed distributions of the $LR(WW)$, $LR(ZZ)$, $LR(W\gamma)$ and $LR(W + jet)$ variables for High (Left) and Low (Right) S/B regions.

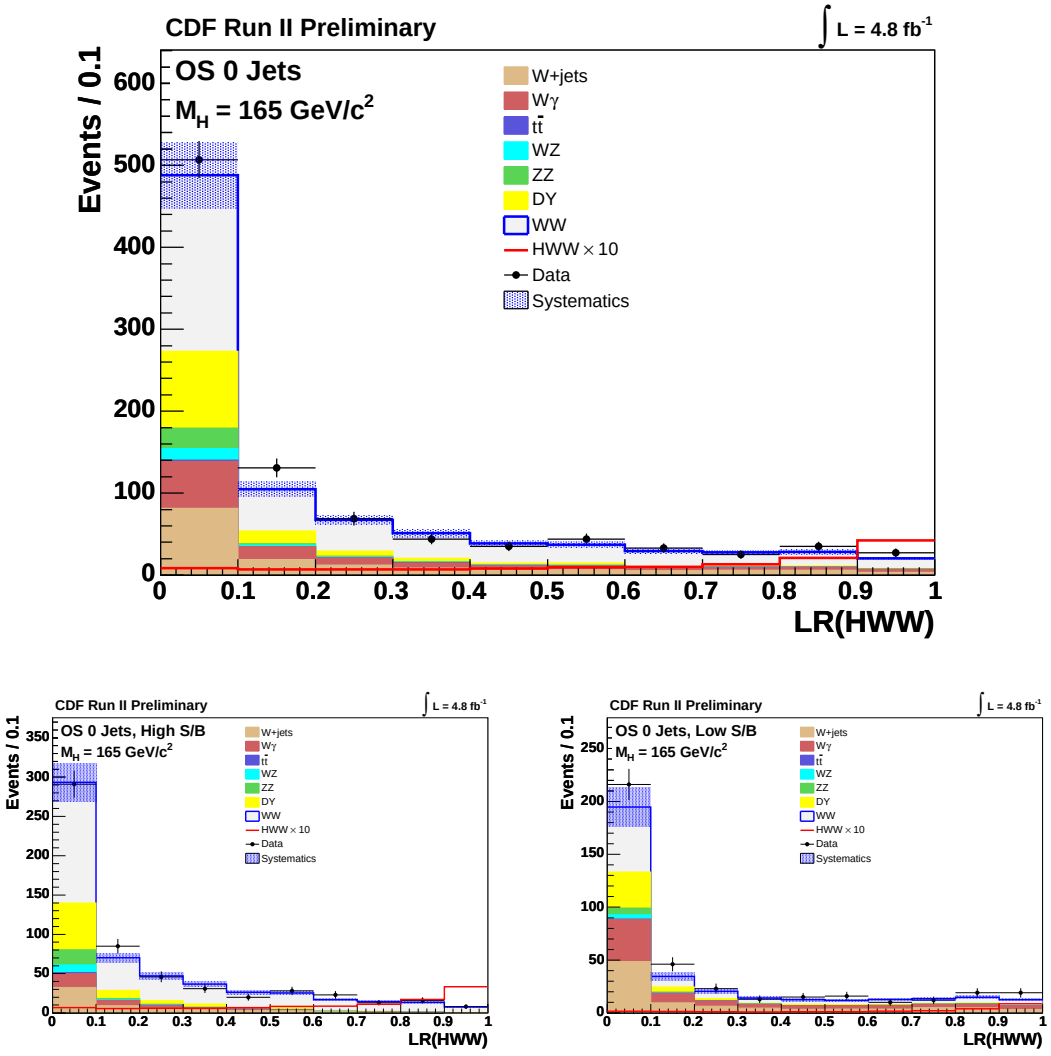


Figure 5.5: Expected and observed distribution of $LR(HWW)$ for High and Low S/B regions (bottom plots). The sum of the two regions is shown in the upper plot. Higgs mass is assumed to be $m_H = 165$ GeV.

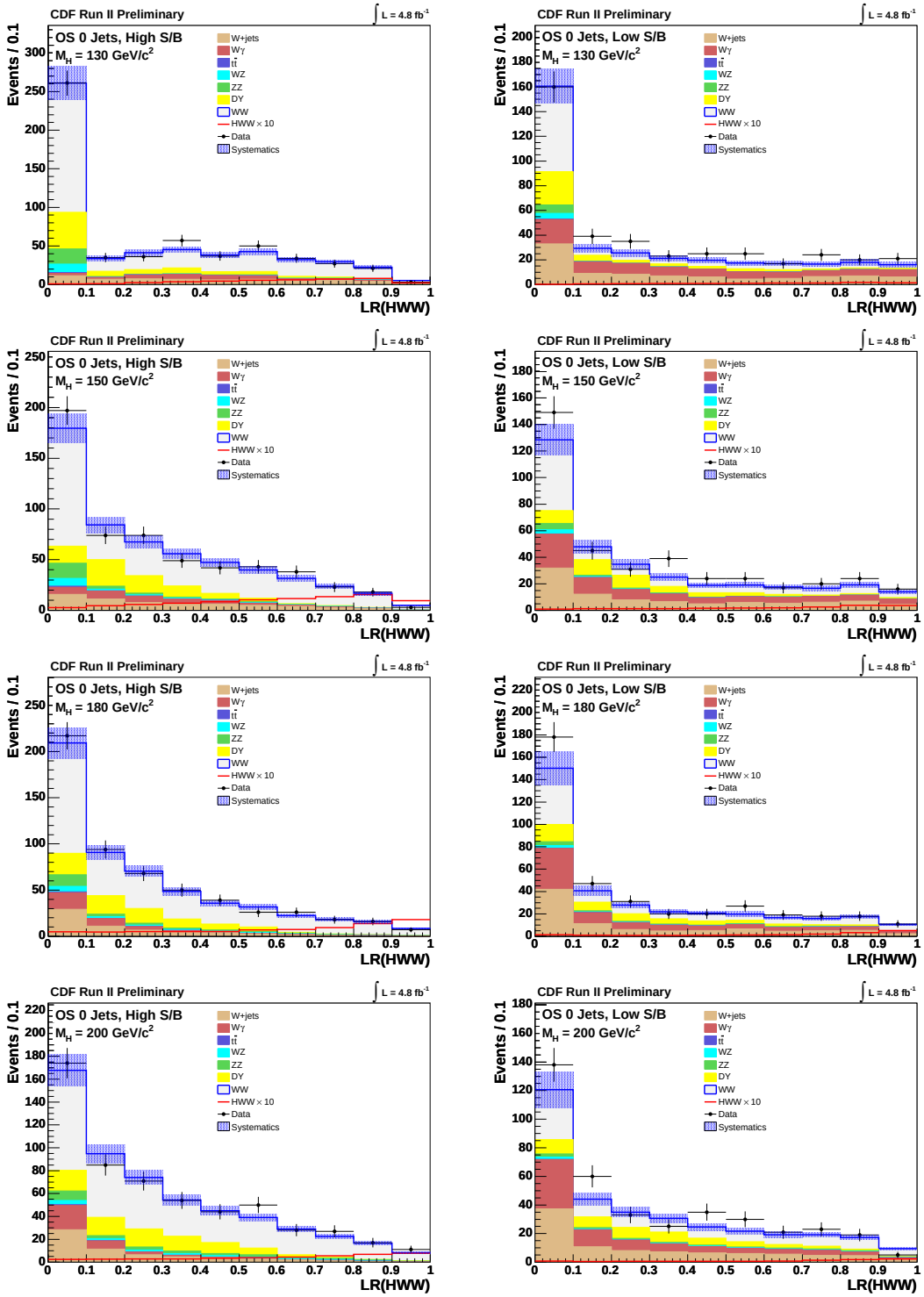


Figure 5.6: Expected and observed distribution of $LR(HWW)$ for High (Left) and Low (Right) S/B regions. Higgs masses of $m_H = 130$ GeV,$m_H = 150$ GeV,$m_H = 180$ GeV and $m_H = 200$ GeV are shown from top to bottom.

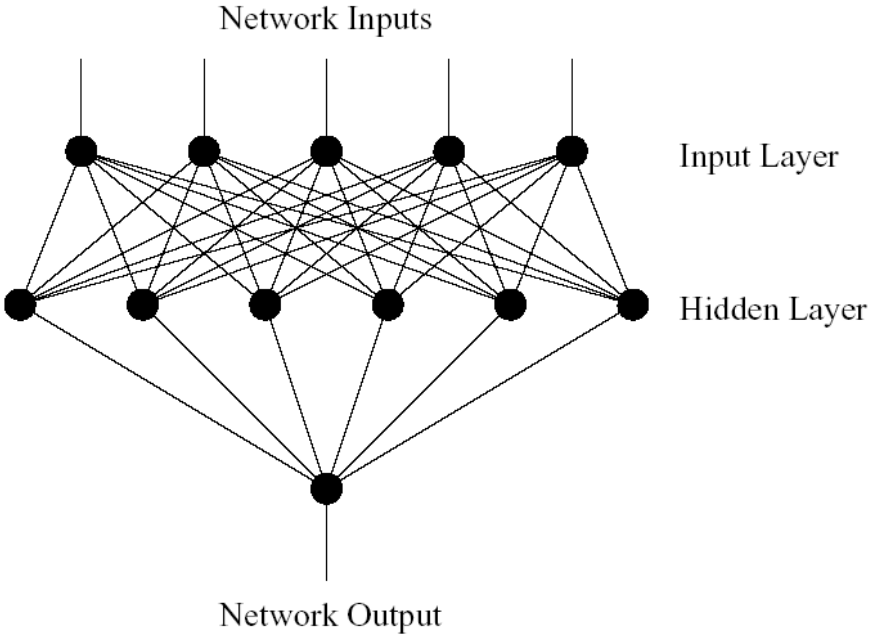


Figure 5.7: The Neural Network structure, consisting of 5 input nodes, 6 hidden nodes and one output node.

do not exchange information among them (none of the outputs is used as input to another node of the same layer): this structure is classified as a feed-forward NN. The nodes of the first layer have only one input each, corresponding to the measured kinematic variables. The second layer is known as hidden layer, since it only communicates with other nodes of the network. In our case the last layer is constituted by just one node. The output of that node is called the NN output or score.

The power of the feed-forward neural networks relies in the existence of a simple algorithm to choose the best weights w_i from empirical examples to approximate the function f. This well suits our conditions, were we have a description of the expected signal and of the background processes and we want to know from which each data event is more likely to come from. This algorithm is known as *back-propagation*(73) and it uses a gradient-descending minimization algorithm to minimize the squared distance between the NN output values and the target values for these outputs. The process of finding the best weights for a given classification problem is also known as the *training* of the NN.

The NeuroBayes®(74) package is used in this work. This package also offers instruments to make the training more robust against unphysical values of the inputs (if not defined for a given event) and to avoid the training process to be influenced by statistical fluctuation of the samples.

Five input variables are used as input to the Neural Network:

- LR(HWW): the Likelihood-ratio tuned to separate signal from background;
- LR(WW): the likelihood-ratio tuned to separate the WW process from other back-

grounds;

- $\Delta R(ll)$ between leptons;
- $\Delta\phi(ll)$ between leptons;
- H_T: the sum of leptons energy and $\vec{\not{E}}_T$.

The choice of input variables has been optimized to obtain the best expected sensitivity to the Higgs signal.

We train fourteen NNs, one for each signal sample depending on the generated Higgs boson mass to distinguish between it and a mixture of background samples weighted by their relative yield expectation. The samples are randomly divided in two sub-samples: the first one is used for training, and the second one is used to check the distribution of the scores. This assures that the NN is not sensitive to the statistical fluctuation inside the training sample.

Figure 5.8 shows the expected signal and background score distributions when the NN is trained on a MC simulation of the Higgs boson of mass $m_H = 165$ GeV (NN^{165}). The background distributions are plotted on top of each other. The Figure also compares the score distribution of data events with the one of the expected sample composition. As expected signal events tend to have a score near the target value of $+1$, while background events tend to be classified with lower scores, near -1. Figures 5.9, 5.10, 5.11 and 5.12 show the distribution of all the other thirteen NNs.

For all the fourteen NN output distributions, the data are well described by the Standard Model mixture of the expected background processes. It is also remarkable that for $m_H = 165$ GeV the expected number of signal events are comparable to the background expectation in the high-score bins of the NN output distribution, where $S/B \sim 0.5$ is expected. Then, the NN output distributions are used to set limits on the Higgs boson production cross section in Section 5.5.

5.3 Systematic uncertainties

This Section describes the sources of systematic uncertainties considered for the Higgs search. Systematic uncertainties affect the normalization (or *rate*) of the background and the signal estimations. However they may also change the *shape* of the kinematic distributions, and consequently the shape of the Neural Network output.

The rate uncertainties are divided into three categories.

cross section systematics: since the theoretical cross section of each process is used to normalized signal and background simulations, the error on that is used to calculate the systematic uncertainty on the number of expected events;

acceptance systematics: we group in this category all the sources of systematic uncertainties that vary the number of selected events either for the detector geometry or for the event selections applied;

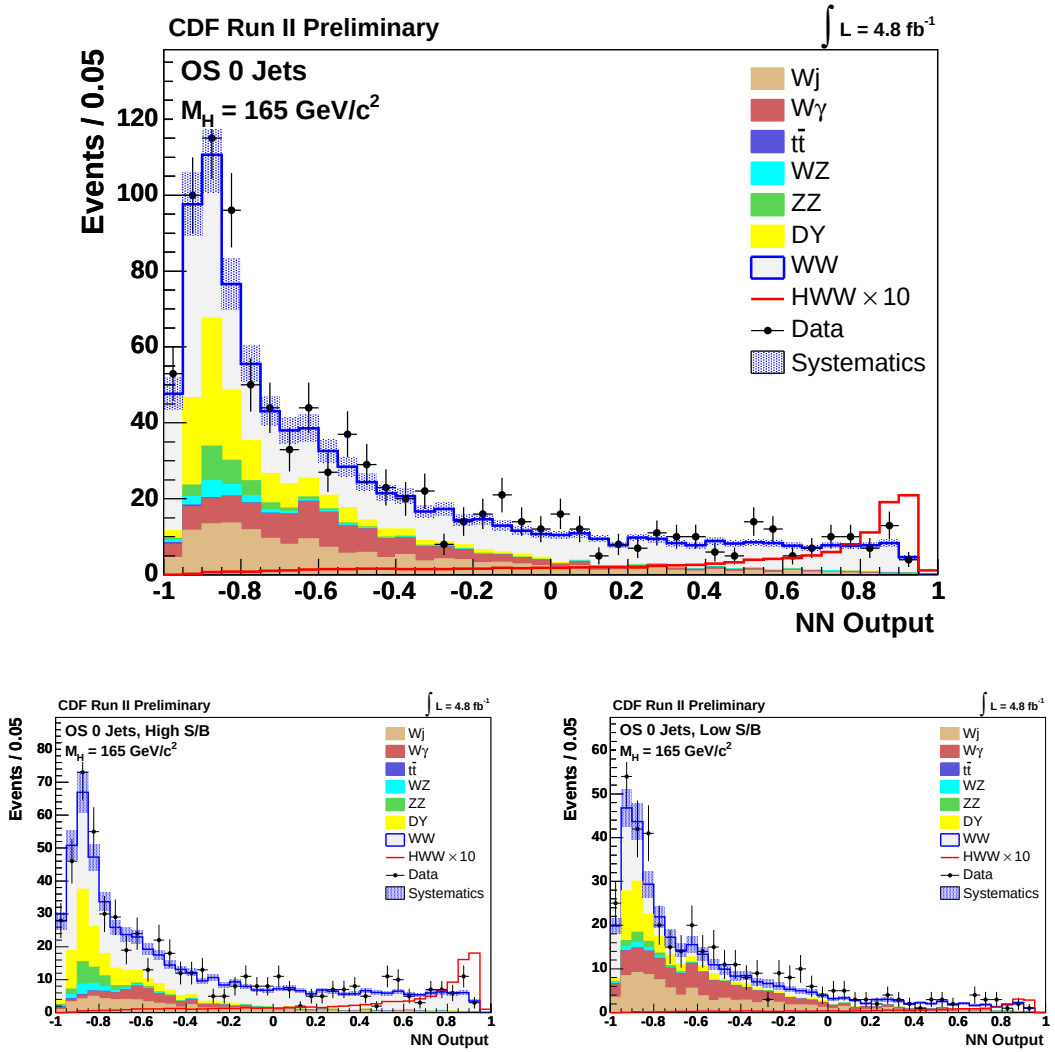


Figure 5.8: Top: Expected and observed NN^{165} output distribution in the opposite-sign region containing no reconstructed jets. Bottom: High S/B and Low S/B regions.

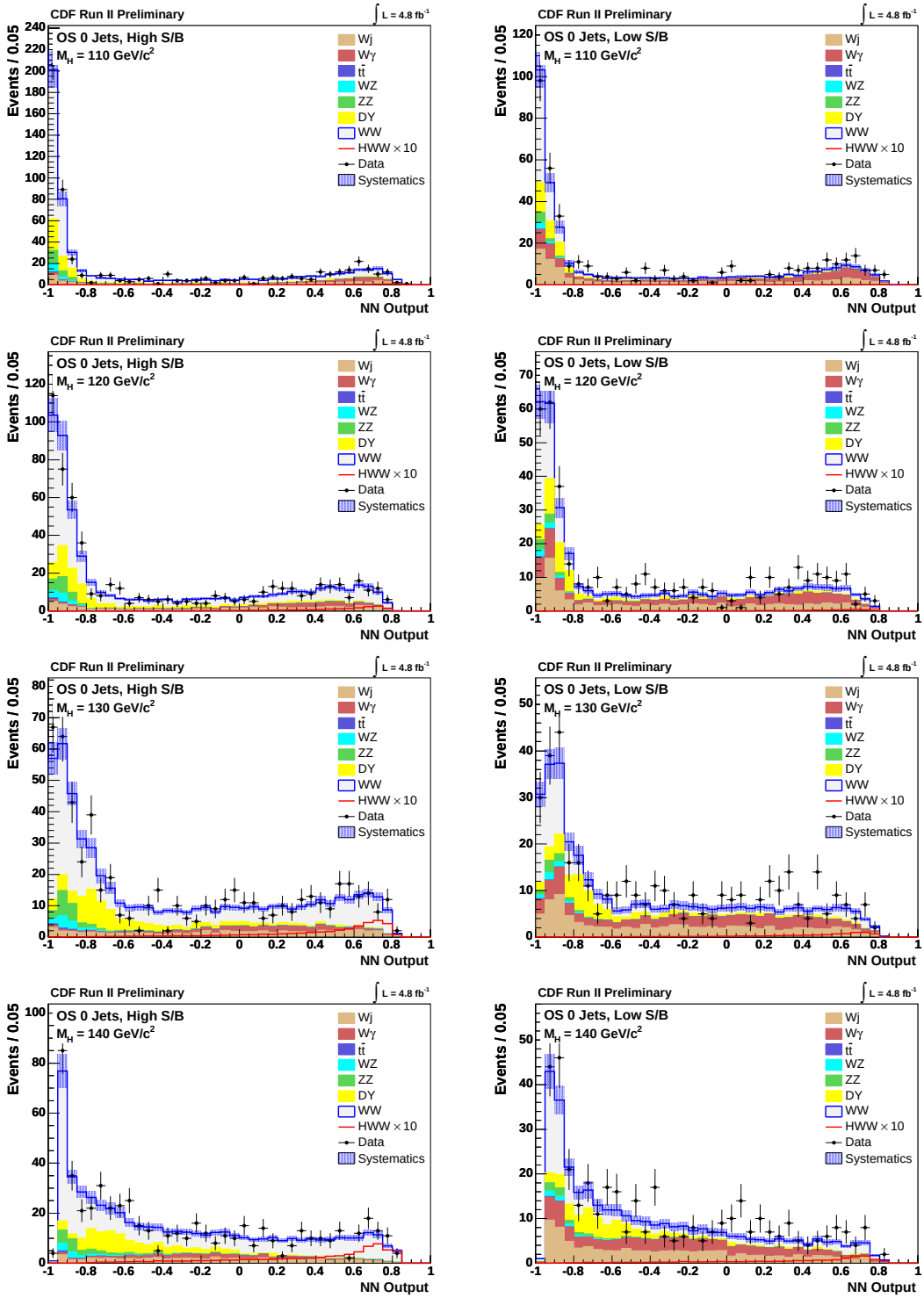


Figure 5.9: Expected and observed NN^{110}, NN^{120}, NN^{130} and NN^{140} output distributions in the opposite-sign region containing no reconstructed jets. Distributions separated in High S/B and Low S/B regions are shown.

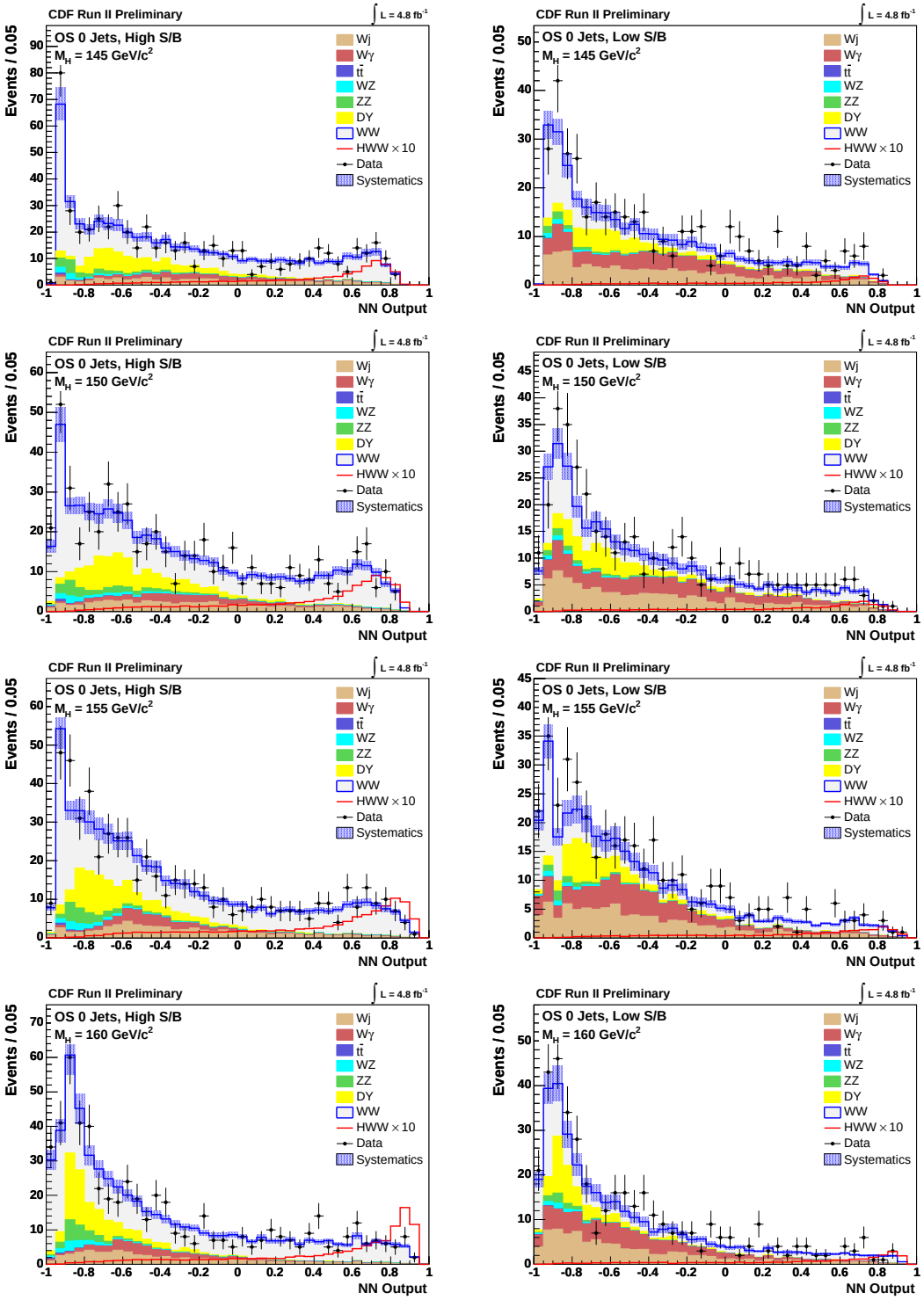


Figure 5.10: Expected and observed NN^{145}, NN^{150}, NN^{155} and NN^{160} output distributions in the opposite-sign region containing no reconstructed jets. Distributions separated in High S/B and Low S/B regions are shown.

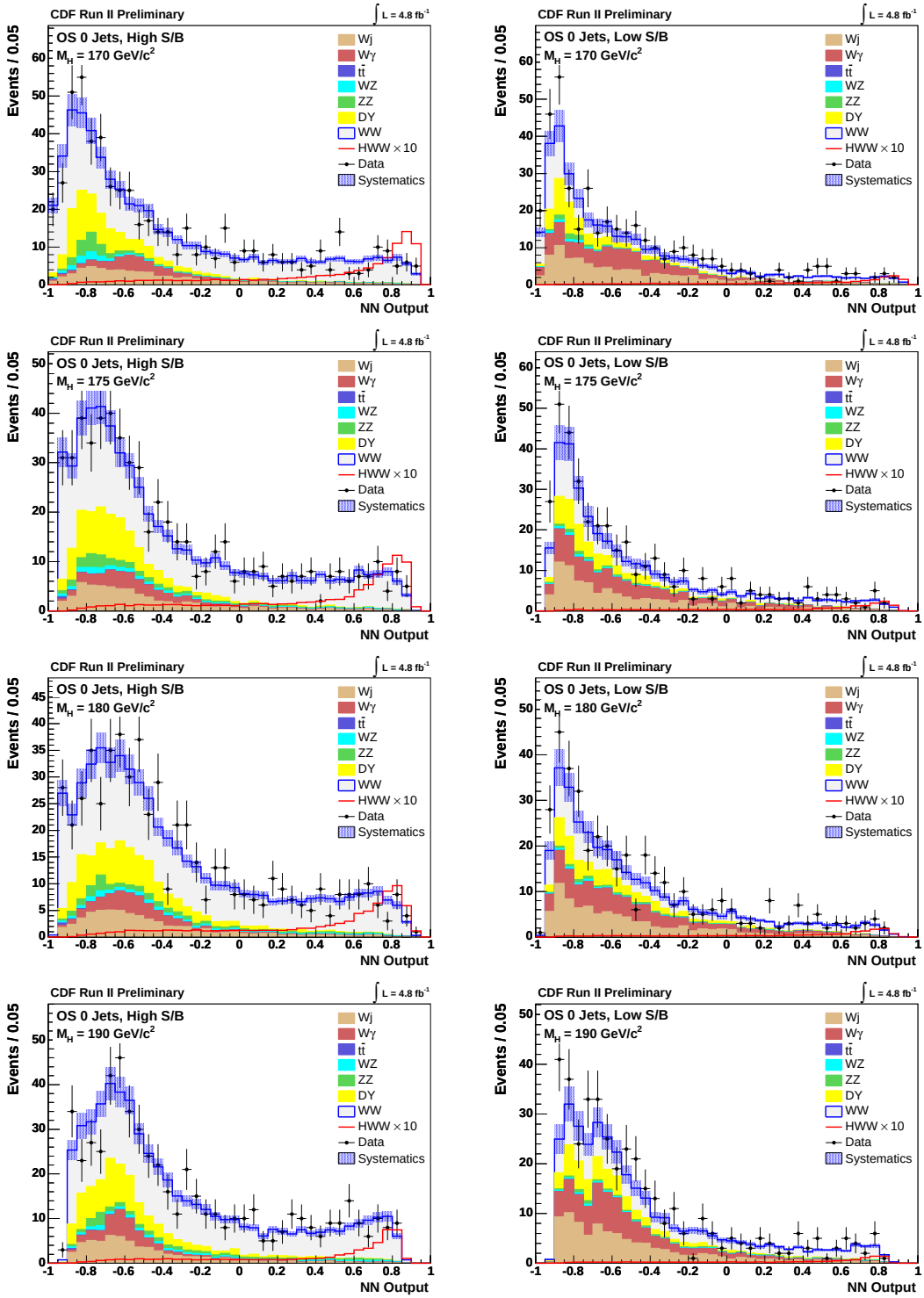


Figure 5.11: Expected and observed NN^{170}, NN^{175}, NN^{180} and NN^{190} output distributions in the opposite-sign region containing no reconstructed jets. Distributions separated in High S/B and Low S/B regions are shown.

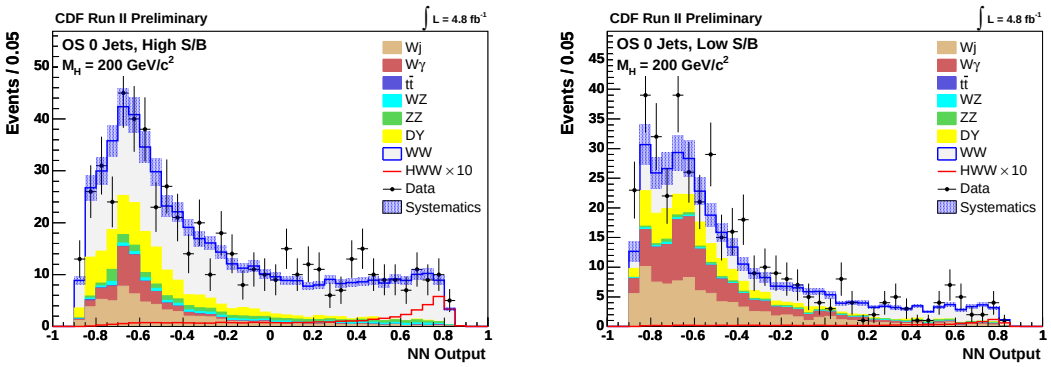


Figure 5.12: Expected and observed NN^{200} output distribution in the opposite-sign region containing no reconstructed jets. Distributions separated in High S/B and Low S/B regions are shown.

luminosity uncertainty: accounts for the error on the luminosity measurement which affect the number of events predicted by the simulations.

Table 5.4 lists the systematic uncertainties on the rate: each row represent a source of systematic uncertainty, while columns show the relative, expressed in %, effect of that systematic uncertainty on the number of expected events for the given processes. The calculation of these uncertainties is discussed below. Systematic uncertainties that are in *italic* in the same row of the Table are considered 100% correlated. The negative values are important when combining different results together, and they indicate anti-correlation among the analyses (see Chapter 7). All the other systematic uncertainties are assumed uncorrelated.

Section 5.3.4 studies the shape systematics on the NN score distributions and their effect in the limit calculation.

5.3.1 Cross section systematics

The $t\bar{t}$ cross section is calculated at NNLO in Ref. (62) and the estimated uncertainty is 10%. For the di-boson processes we use MCFM(31) to perform a NLO calculation with an uncertainty of 6%. Drell-Yan is known at NNLO and the associated systematic is 5%(61). These uncertainties include contributions from scale and PDF variation, and are reported in the row *Total* of the systematic table.

For the production cross section calculation of the Higgs boson through gluon fusion, we consider uncertainties due to PDF and scale variation separately. In fact uncertainties on the differential cross section as function of the jet multiplicity are expected to differ respect to an inclusive calculation(75). Following the approach outlined in Ref. (75), we used HNNLO(76) 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. The peculiarity of HNNLO is that it allows the user to apply kinematic selections on the final state leptons and quarks of the hard scattering, before the hadronization. We vary the renormalization and factorization scale of the calculation, differentiating final

Uncertainty Source	WW	WZ	ZZ	$t\bar{t}$	DY	$W\gamma$	W+jet(s)	$gg \to H$
Cross Section								
Scale	⇊	⇊	⇊	⇊	⇊			7.0
PDF Model	⇊	⇊	⇊	⇊	⇊			7.7
Total	*6.0*	*6.0*	*6.0*	10.0	5.0			10.4
Acceptance								
Scale (leptons)								2.5
Scale (jets)								4.6
PDF Model (leptons)	1.9	2.7	2.7	2.1	4.1			1.5
PDF Model (jets)								0.9
Higher-order Diagrams	*5.0*	*10.0*	*10.0*	10.0		11.0		
Missing Et Modeling					21.0			
$W\gamma$ Scaling						12.0		
Jet Fake Rates (Low/High S/B)							21.5/27.7	
Jet Modeling	-1.0					*-4.0*		
MC Run Dependence	2.8							
Lepton ID Efficiencies	2.0	1.7	2.0	2.0	1.9			1.9
Trigger Efficiencies	2.1	2.1	2.1	2.0	3.4			3.3
Luminosity	*5.9*	*5.9*	*5.9*	*5.9*	*5.9*			*5.9*

Table 5.4: Systematic uncertainties for the opposite-sign di-lepton analysis with zero jets. All the values are symmetric and expressed in percent. Each row represent a source of systematic uncertainty, while columns represent the relative effect of that systematic uncertainty on the number of expected events for the given process.

states by jet multiplicity. For events with no reconstructed jets we obtain ±7.0% variations, consistent with Ref. (75). We take the PDF uncertainties from Ref. (75) to be 7.7%, which are calculated varying the PDFs within their errors as recommended in Ref. (22).

5.3.2 Acceptance systematics

Scale and PDF model

The uncertainties on the PDF reflect on the acceptance of each process. We estimate their effect on the simulations varying their independent parameters according to the reported errors(72). The resulting PDF are then propagated through the analysis to obtain their effect on the background normalization.

We investigated the effect of scale and PDF systematic uncertainties on the signal acceptance using HNNLO. We varied PDF parameters, renormalization and factorization scales to derive different p_T and η distribution of the Higgs boson. We then re-weight the Pythia samples to match these calculations. The differences in acceptance are taken as systematics uncertainty.

The effect of the Higgs p_{T} re-weight depends on the number of jets in the final state. By re-weighting the Higgs p_{T} spectrum, we change the vector sum of the p_{T} of the jets that are recoiling against the Higgs. The Pythia Higgs p_{T} spectrum is harder than that obtained with HNNLO. Thus when we re-weight Pythia to match HNNLO, the events that previously had one or more jets can be moved into a lower jet bin. Because the re-weighting is moving events among the 0, 1, and two or more jet bins, the acceptance systematics *Scale (jets)* and *PDF Model (jets)* are anti-correlated among the jet bins.

The Higgs η distribution in Pythia is more central than the one generated by HNNLO, thus when we re-weight Pythia to match HNNLO, the leptons can be moved outside the forward acceptance. These acceptance systematics are labeled *Scale (leptons)* and *PDF Model (leptons)* in the systematics table.

Higher order diagrams

An other systematic is due to the differences between LO and higher order calculations on the acceptance. We determine the size of higher order approximation by comparing a LO Pythia WW sample to the WW MC@NLO sample. We take half of this difference (5%) as a systematic error on the WW process; the full difference (10%) is assigned to the processes simulated only by LO Pythia. This systematic is called *Higher Order Diagrams* in the systematic table.

$\not\!\!\vec{E}_T$ and Jet modeling

The uncertainties on the modeling of the $\not\!\!\vec{E}_T$ is very small for processes with real $\not\!\!\vec{E}_T$. However, the DY process which has no real $\not\!\!\vec{E}_T$ may have important systematic. To determine that we use the Low $\not\!\!\vec{E}_T^{\mathrm{Spec}}$ control region described in Section 4.6. Based on the rate discrepancy seen in this control region, we assign a DY $\not\!\!\vec{E}_T$ and jet modeling systematic of 21% for events with no jets.

We use this estimation to infer a systematic uncertainty of 20% for the model of the simulations in events with extra jets from initial or final state radiation. Then we evaluate the number of events moved along different number of reconstructed jets. Since WW is already simulated at NLO, this systematic is assigned to be 20% to events with at least two reconstructed jets, which correspond to 1% for events with no reconstructed jets in the final state. This systematic is indicated as *Jet Modeling* in the systematic table.

MC Run dependence

The WW MC@NLO sample has a restricted run range on which the simulation was tuned on. Then we assign an additional uncertainty which is determined as half of the correction applied in Section 4.3.

Fake background normalization

The only uncertainty on the W+jets process is due to the fake rate. We estimate this uncertainty by varying the fake rates within their assigned uncertainties (shown by the solid light gray bands in Figures 4.11-4.12). These variations are calculated separately in the High S/B and Low S/B regions given the different fake contribution.

The $W\gamma$ normalization is measured using data, as explained in Section 4.5. It has 12% systematics, which accounts for the uncertainties in the normalization measurement and for the uncertainty in the extrapolation of the normalization to the signal region. This is referred as $W\gamma$ scaling in the Systematic Table.

Lepton ID and Trigger efficiencies

The lepton ID scale factors and the trigger efficiencies are varied coherently $\pm 1\sigma$ to determine their effect on the acceptance. Both uncertainties have a small effect, around 2%, on the signal and background processes.

5.3.3 Luminosity measurement

The integrated luminosity of the sample is measured as described in Section 3.2.6. The uncertainty on this measurement is estimated 5.9%, with mainly two sources. The inelastic $p\bar{p}$ cross section contributes for 4%, while another 4.2% comes mainly from systematic uncertainties on the geometry description of the detector in the simulation. The luminosity uncertainty is taken 100% correlated for all the processes which do not normalize to data.

5.3.4 Systematic uncertainties on NN shape

We consider two possible effects that can modify the shape of the NN output distributions: the jet energy scale (JES) and the initial state radiation (ISR) or scale variations.

The uncertainty on the jet energy scale described in Ref. (58) is evaluated by changing ± 1 σ the energy of the jets. We then evaluate the NN output distribution for both background and signal expectations. The effects on the output shape are generally less

than $\sim 5\%$, measured as the maximum difference from the nominal shape in each bin of the the NN output distribution. The effect on the sensitivity of the measurement has been found to be negligible.

Pythia Tune A is used as our default for initial and final state radiation (for the process $gg \to H$, there is only ISR). The limited knowledge of these processes is expressed by a systematic error determined using Pythia $gg \to H$ samples with the Higgs mass of 160 GeV, where the ISR is either increased and decreased by $\pm 1\sigma$ respect to the default value. Both effects have a small impact on the Neural Network output and no consequences on the sensitivity.

5.4 Limit calculation on the Higgs production cross section

Neural Network output distributions derived in Section 5.2.2 and showed in Figures 5.8, 5.9 and 5.10 are compatible with Standard Model predictions with no $H \to WW$ component. It is thus desirable to quote an allowed upper limit on the possible contribution from the signal to the data at some confidence level. The confidence level chosen is 95%, which is a standard choice. In this Section is described the method used to evaluate the cross section upper limits. Results will be obtained in Section 5.5.

We adopt a pure Bayesian(77) method to rigorously estimate the upper limit at 95% C.L. on the Higgs production cross section, as the ratio to the Standard Model prediction, for the fourteen Higgs mass hypothesis considered in the range $110 - 200$ GeV.

The method combines counting experiments performed in statistically independent samples and it has many advantages including an easy extension for the combination of several searches and the natural inclusion of systematic uncertainties in the calculation.

The input is the binned Neural Network output histogram. For each Higgs boson mass hypothesis we build a likelihood $\mathcal{L}$ that is the product of Poisson probabilities of observing in the $i-$th bin of the input histogram n_i events, given an expectation of μ_i:

$$\mathcal{L}(R|\vec{\theta},\vec{n}) = \prod_i \frac{\mu_i^{n_i} \mathrm{e}^{-\mu_i}}{n_i!} \cdot \prod_k \mathrm{e}^{-\frac{\theta_k^2}{2}}, \tag{5.5}$$

with

$$\mu_i = R \cdot s_i(\vec{\theta}) + b_i(\vec{\theta}). \tag{5.6}$$

The expected number of events in the $i-$th bin (μ_i) can also be written as the sum of signal $s_i(\vec{\theta})$ and background $b_i(\vec{\theta})$ expectations. The artificial parameter R multiplies the number of expected signal events, that can also be written as $s_i = L \cdot \epsilon \cdot \sigma_H^{SM}$, where L is the integrated luminosity of the collected sample, ϵ the detector acceptance and analysis efficiency and σ_H^{SM} is the Standard Model Higgs cross section. Therefore R is a multiplicative factor on σ_H^{SM}. We assume no experimental information on the Higgs production cross section, and assign a flat prior to R. Each of the signal and background predictions depend on the parameters θ_k. They are used to account for

systematic uncertainties in the predictions:

$$s_i(\vec{\theta}) = s_i \cdot \prod_k (1 + u_i^k \cdot \theta_k) \tag{5.7}$$

$$b_i(\vec{\theta}) = b_i \cdot \prod_k (1 + u_i^k \cdot \theta_k), \tag{5.8}$$

where u_i is the relative systematic error that is associated to the prediction s_i/b_i. The product runs over the index k, which accounts for different sources of systematic uncertainties. θ_k is called *nuisance* parameter and is constrained with a gaussian prior in Equation 5.5. Since the systematic uncertainties u_i depend on i, it is easy to implement uncertainties either on the normalization of all the histograms bins together (*rate systematics*) or independently on each bin (*shape systematics*).

This technique easily adapts for combining results of several analyses. In fact when more of one histogram is used as input the index i of Equation 5.5 runs over each bin of each input histogram. In this way it's also straightforward to correlate systematic uncertainties among them.

The Likelihood is integrated over each nuisance parameter θ_k and evaluated for the observed number of events $\vec{n}$; the only left dependence is on the parameter R. The result, normalized to unit area, is the posterior $\mathcal{L}(R)$. An example of the posterior for the parameter R is shown in Figure 5.13(left). The 95% upper limit R^{95} is then calculated solving the equation:

$$\frac{\int_0^{R^{95}} \mathcal{L}(R)\mathrm{d}R}{\int_0^{+\infty} \mathcal{L}(\mathcal{R})\mathrm{d}\mathcal{R}} = 0.95. \tag{5.9}$$

It corresponds to the limit on the production cross section of the Higgs boson, normalized to the Standard Model prediction. In the example of Figure 5.13(left) the 95% of the posterior area lies before $R^{95} = 1.18$: it means that a signal with a cross section of $1.18\sigma_H^{SM}$ or larger is excluded at least at 95% C.L. Moreover the posterior peaks at $R = 0$, indicating that the preferred hypothesis is the one without any signal component.

To estimate the sensitivity of the analysis before looking at data events, we generate pseudo-experiments in a background-only hypothesis: *pseudo*-data n_i are randomly generated for each sample i from the Poisson distribution with the mean given by the background expectation b_i. Systematic uncertainties are taken into account randomly fluctuating b_i by its error. Each of them is then used to evaluate the 95% C.L. limit in the same way as for real data. An example of the resulting distribution of the calculated limits is shown in Figure 5.13(right). We define the expected limit as the median 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.

5.5 Results

We use the approach described in Section 5.4 to set 95% C.L. limits on the production cross section of the Higgs boson. We use as input the High S/B and Low S/B NN output distributions. They are statistically independent samples, and their systematic uncertainties are assumed correlated.

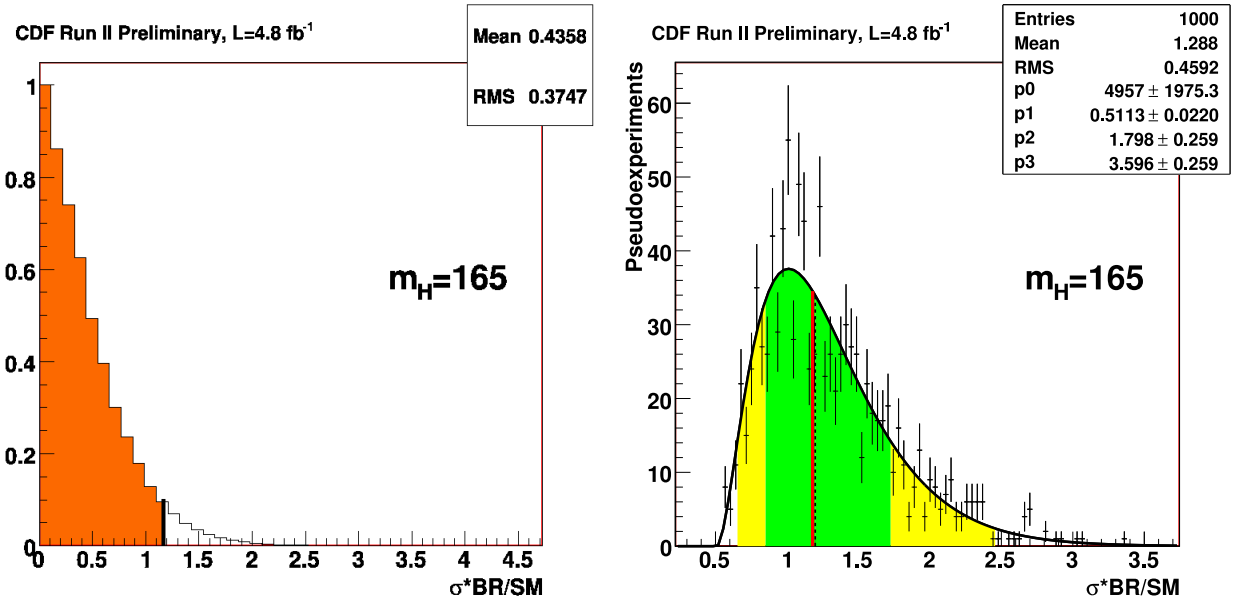


Figure 5.13: Left: Example of a normalized posterior distribution $\mathcal{L}(\mathcal{R})$. In the plots R is also indicated as $\sigma * BR/SM$, i.e. as the ratio of the signal cross section times branching ratio $\sigma * BR$ divided by the Standard Model SM expectation. The filled region corresponds to 95% of the total area. Right: Example of the expected distribution of 95% C.L. limits in the background-only hypothesis. The red line shows the median of the distribution, while the green and yellow bands represent the $\pm 1, 2\sigma$ intervals.

Table 5.5 shows the expected (median) and the observed cross section upper limits, divided by the expected Standard Model cross section. The 1σ and 2σ spread of the expected limit distribution is also reported. The limits are evaluated independently for each of the fourteen Higgs mass hypothesis ranging from 110 GeV to 200 GeV, and include the effects of systematic uncertainties. The same results are graphically shown in Figure 5.14.

The maximum sensitivity is reached for $m_H = 165$ GeV, which is consistent with the maximum of the cross section times branching fraction for the $H \to WW$ signal. For this mass hypothesis the 95% C.L. sensitivity is 1.96 times the expected SM cross section σ_H^{SM}. This result greatly improves the sensitivity obtained with the simple counting experiments described in Sections 5.1 and 5.2.1, highlighting the impact of the multivariate techniques to the analysis. The NN distribution of data events sets an observed upper limit of 2.76 σ_H^{SM}, which is consistent with that expected if the sample composition is the admixture of SM backgrounds. Although the sensitivity is near to the expected σ_H^{SM}, this analysis is not yet sufficient to probe the Higgs signal at the Standard Model level.

For heavier Higgs mass hypotheses the branching fraction $H \to WW$ changes less than 10% while the cross section decreases more rapidly determining the loss of sensitivity. On the other hand although the expected Higgs cross section increases for a lighter Higgs mass, the branching fraction $H \to WW$ rapidly decreases, since one of the W becomes off-shell. The net effect is a rapid loss of sensitivity.

Observed limits are within the 1σ spread of the expected distribution calculated in

OS 0 Jets	110	120	130	140	145	150	155
$-2\sigma/\sigma_{SM}$	25.02	7.91	4.03	2.54	2.16	1.87	1.48
$-1\sigma/\sigma_{SM}$	33.88	10.85	5.48	3.47	2.98	2.56	2.01
Median/σ_{SM}	**48.21**	**15.44**	**7.73**	**4.90**	**4.20**	**3.59**	**2.84**
$+1\sigma/\sigma_{SM}$	69.19	21.97	11.09	6.95	5.98	5.10	4.08
$+2\sigma/\sigma_{SM}$	96.93	30.42	15.45	9.58	8.29	7.19	5.67
Observed/σ_{SM}	**50.99**	**17.39**	**9.18**	**6.54**	**5.39**	**3.97**	**3.18**
	160	165	170	175	180	190	200
$-2\sigma/\sigma_{SM}$	1.08	1.03	1.30	1.50	1.88	3.00	4.29
$-1\sigma/\sigma_{SM}$	1.45	1.40	1.74	2.06	2.56	4.10	5.79
Median/σ_{SM}	**2.06**	**1.97**	**2.44**	**2.91**	**3.62**	**5.86**	**8.26**
$+1\sigma/\sigma_{SM}$	2.92	2.79	3.53	4.16	5.20	8.41	11.96
$+2\sigma/\sigma_{SM}$	4.14	3.95	4.94	5.83	7.16	11.81	16.55
Observed/σ_{SM}	**2.63**	**2.76**	**2.85**	**3.55**	**4.02**	**6.64**	**11.03**

Table 5.5: Expected and Observed limits on Higgs production cross section in the opposite sign di-lepton sample with zero jets as function of the Higgs mass. Limits are in units of SM cross section.

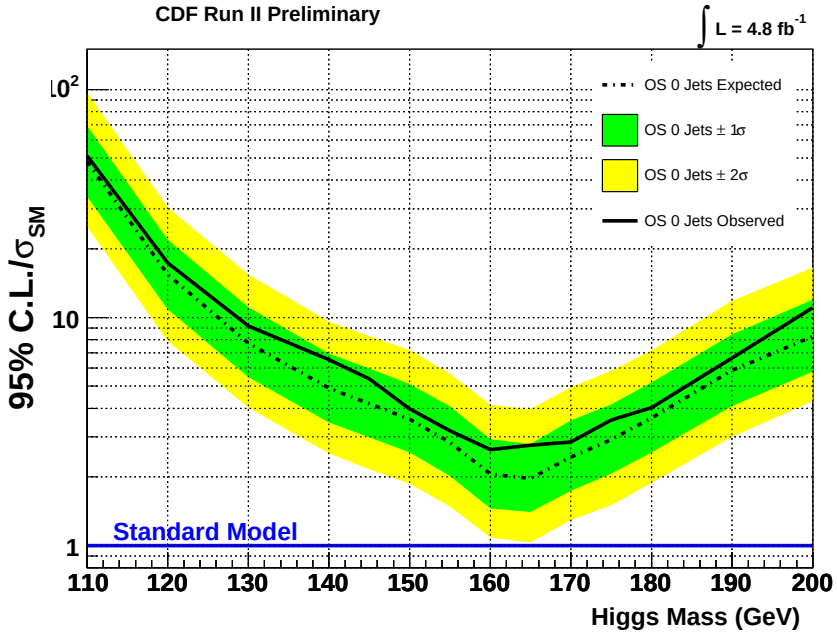


Figure 5.14: Expected and Observed limits on Higgs production cross section in the opposite sign di-lepton sample with zero jets in several bins of Higgs mass. Limits are in units of SM cross section.

the background-only hypothesis for all the mass hypotheses. Since the data sample used to set limits is the same for each Higgs mass hypothesis, the observed limits are highly correlated.

Chapter 6

Analysis of associate Higgs production

In the Chapter 5 we described the search for the Higgs boson produced by gluon fusion analyzing di-lepton events with zero reconstructed jets in the final state. In this Chapter we discuss the search for the associate production and Vector boson fusion, with the subsequent $H \rightarrow WW$ decay, where likely there are extra jets in the final state.

Both signatures, the opposite and same sign leptons, are presented and the results are discussed.

6.1 Signal and background composition

Direct Higgs production mainly has no jets in the event, although initial state radiation from gluons gives an important contribution to higher jet multiplicity bins. Associated production and Vector Boson Fusion are significant contributors to events with one jet in the final state and are they are the major signal contributors for events with at least two jets. Figure 6.1 (Left) and Table 6.1 report the number of expected Higgs signal events in the opposite-sign di-lepton sample as function of the number of reconstructed jets.

In addition Figure 6.1 (Right) shows the number of expected background events in the selected sample as function of the number of jets reconstructed in the event. Observed data events are also shown. Non-resonant WW production is half of the di-lepton sample with no reconstructed jets in the final state, and it also significantly contributes to events with one reconstructed jet in the final state. For these events the Drell-Yan is an important background. Events with at least two jets in the final state are dominated by $t\bar{t}$ production.

6.2 Events with one reconstructed jet

Events in the opposite sign di-lepton sample with one reconstructed jet are divided into two sub-samples: High S/B and Low S/B. The classification is the same applied in Chapter 5 for events with no jets.The expected number of events in both regions agrees with the observed data as listed in Table 6.2. Signal contribution for $m_H = 165$ GeV is also shown.

N_{Jets}	$gg \to H$	WH	ZH	VBF	Total
0	12.63	0.31	0.32	0.11	13.37
1	6.41	0.87	0.34	0.57	8.18
≥ 2	2.53	2.00	1.12	1.06	6.70
Total:	21.56	3.18	1.78	1.73	28.25

Table 6.1: Expected Higgs events as function of the number of jets in the final state in the di-lepton sample. The most important production mechanisms are shown separately. An Higgs mass $m_H = 165$ GeV is assumed.

6.2.1 Kinematics

The strong angular correlation between leptons of signal events holds for all the production mechanisms, since it only depends on the Higgs decay chain. The $\Delta\phi$ and ΔR between leptons are the best quantities to discriminate between signal and background. Good discrimination between signals and Drell-Yan is given by the di-lepton invariant mass, even if a few ZH signal events peak at the Z mass, when both leptons comes from the Z decay, and they do not contribute significantly to the sensitivity in the Higgs search. To improve separation between signals and backgrounds from di-boson and Drell-Yan production, we compute the transverse invariant mass of the two leptons and the missing transverse energy ($m_T(ll\ \vec{\not{E}}_T)$), which is expected to be larger for signal events than for backgrounds. The H_T variable is re-defined as the scalar sum of the leptons energy, the missing transverse energy and the jet energy. It is a good discriminant variable between signal and $t\bar{t}$ events. Figure 6.2 shows the expected and observed kinematic distributions for opposite sign di-lepton events with one jet. Signal contribution is overlaid, with the direct production and the sum of the associated production and the vector boson fusion showed separately. The signal events are multiplied by a factor 10 and a $m_H = 165$ GeV is assumed.

The Matrix Element technique is expected to be not powerful for events with an hard jet radiation since they involve LO kinematic calculations. In fact the boost of the Higgs given by the initial state radiation is not accounted in the probability distributions calculated with this technique, resulting in a worse separation between signal and background. For this reason we decided to use only the Neural Network technique. In order to separate signal from background we trained fourteen Neural Networks, one for each Higgs mass hypothesis. The signal comes from the sum of the Higgs production mechanisms considered, with their Standard Model relative contribution. Eight variables are used as input to the Neural Network:

- ΔR between leptons;
- $M_T(ll\ \vec{\not{E}}_T)$: transverse invariant mass of the leptons and the missing energy;
- $m(ll)$ invariant mass of the two leptons;
- $\vec{\not{E}}_T^{\text{Spec}}$ as defined in Equation 4.1;
- $p_T(l_1)$, $p_T(l_2)$, $E(l_1)$: p_T and energy of the leptons;

CDF Run II Preliminary	$\int \mathcal{L} = 4.8$ fb^{-1}		
$M_H = 165$ GeV/c^2			
$t\bar{t}$	40.3	$\pm$	6.3
DY	95	$\pm$	30
WW	93.6	$\pm$	9.8
WZ	14.0	$\pm$	1.9
ZZ	5.99	$\pm$	0.82
W+jets	22.0	$\pm$	6.9
$W\gamma$	4.48	$\pm$	0.99
Total Background	275	$\pm$	37
$gg \to H$	5.3	$\pm$	1.4
WH	0.721	$\pm$	0.094
ZH	0.283	$\pm$	0.037
VBF	0.477	$\pm$	0.076
Total Signal	6.8	$\pm$	1.5
Data	270		

OS 1 Jet, HighSB

CDF Run II Preliminary	$\int \mathcal{L} = 4.8$ fb^{-1}		
$M_H = 165$ GeV/c^2			
$t\bar{t}$	8.1	$\pm$	1.3
DY	38	$\pm$	12
WW	27.7	$\pm$	2.9
WZ	6.06	$\pm$	0.83
ZZ	1.98	$\pm$	0.27
W+jets	37.3	$\pm$	8.3
$W\gamma$	11.7	$\pm$	2.6
Total Background	131	$\pm$	16
$gg \to H$	1.06	$\pm$	0.28
WH	0.150	$\pm$	0.020
ZH	0.0562	$\pm$	0.0073
VBF	0.088	$\pm$	0.014
Total Signal	1.36	$\pm$	0.29
Data	123		

OS 1 Jet, LowSB

Table 6.2: Expected signal and background in the High S/B (Top) and Low S/B (Bottom) regions for the opposite sign di-lepton sample with one jet. Systematic errors are included as discussed in Section 6.2.2.

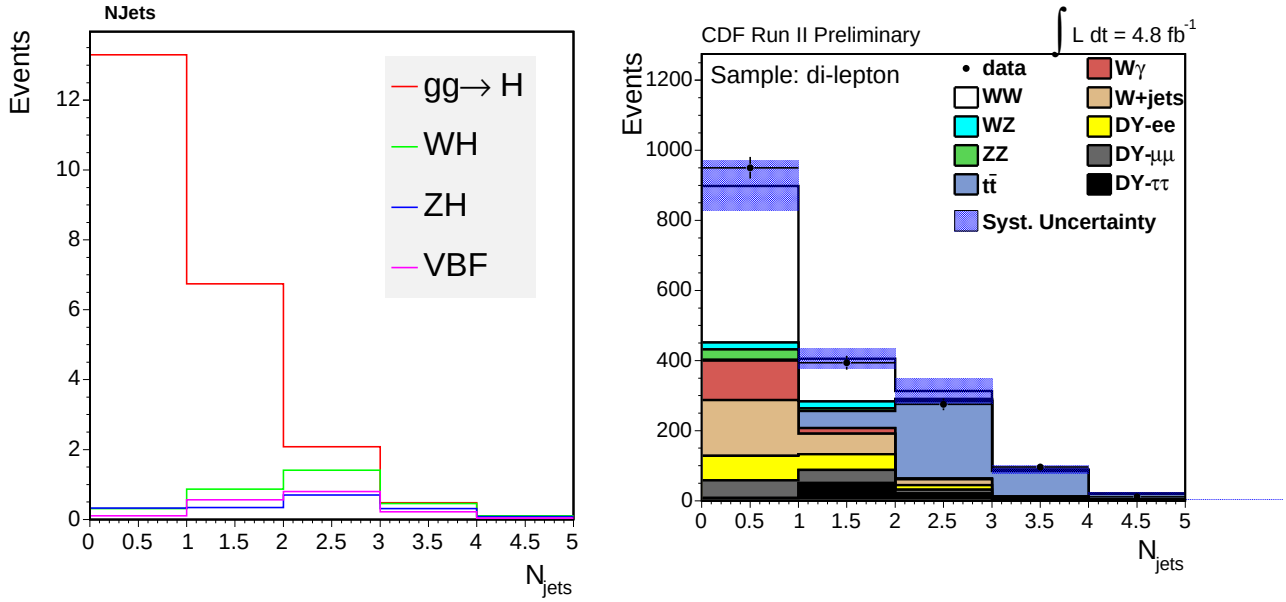


Figure 6.1: Left: the expected signal events from direct, associate and Vector Boson Fusion production in the di-lepton sample for an Higgs mass of $m_H = 165$ GeV as function of the number of jets. Right: the expected (background only) and observed events in the di-lepton sample as function of the number of jets in the event.

- H_T: scalar sum of leptons transverse momenta, $\vec{\not\!\!E}_T$ and jet E_T.

The hidden layer is made of nine nodes and the output by one. Figure 6.3 shows the Neural network output distribution varying the Higgs mass hypothesis for expected and observed events in the High S/B and Low S/B regions. Signal contribution sums up all the production mechanisms and it is shown multiplied by a factor 10.

Neural Network output distributions show that non resonant WW events are still the major background, sitting at high values of the NN score distribution.

6.2.2 Systematics

The search for the Higgs in the di-lepton plus one jet data set suffers systematic uncertainties mainly coming from the same sources discussed in Section 5.3.

We adopt the same categorization, and evaluate the effect of each systematic source on events with one reconstructed jet. The effect of the systematic uncertainties is also evaluated on the signal coming from associated production and vector boson fusion. In this Section we only describe the systematic uncertainties which have significantly different effects respect to the case of zero reconstructed jets.

Associate production cross sections are electro-weak processes known at NNLO; the associated uncertainties are 5% (23), much less than for direct production. VBF is only known at NLO and the systematic uncertainty is set to 10% (23). Using the same method described in Section 5.3 we find ±23.5% change on the Higgs direct production cross section for the scale variation. Instead the PDF systematic is assumed to be the same as

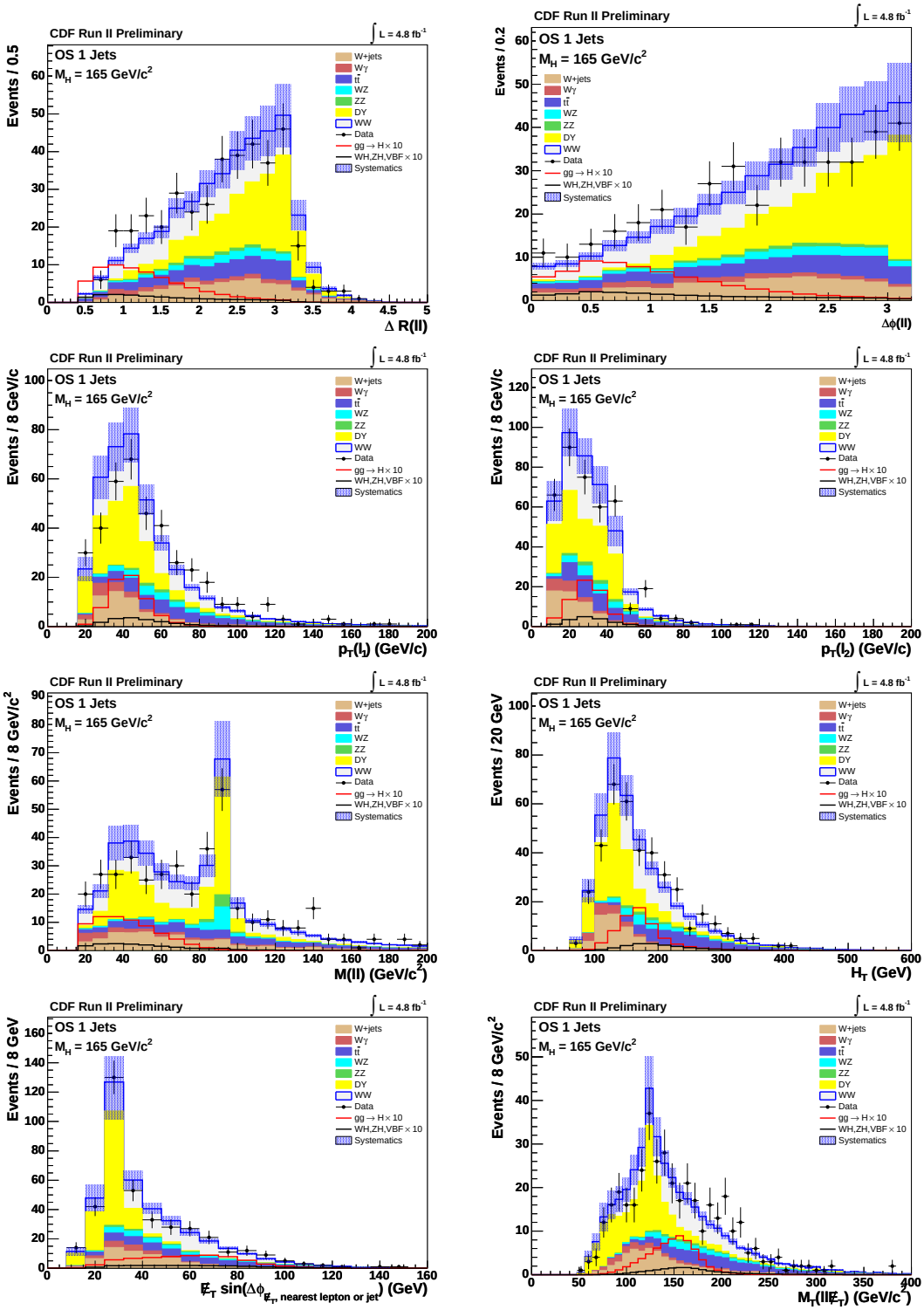


Figure 6.2: Signal and background kinematic distributions in the opposite-sign signal region of events containing one reconstructed jet with $E_T > 15$ GeV.

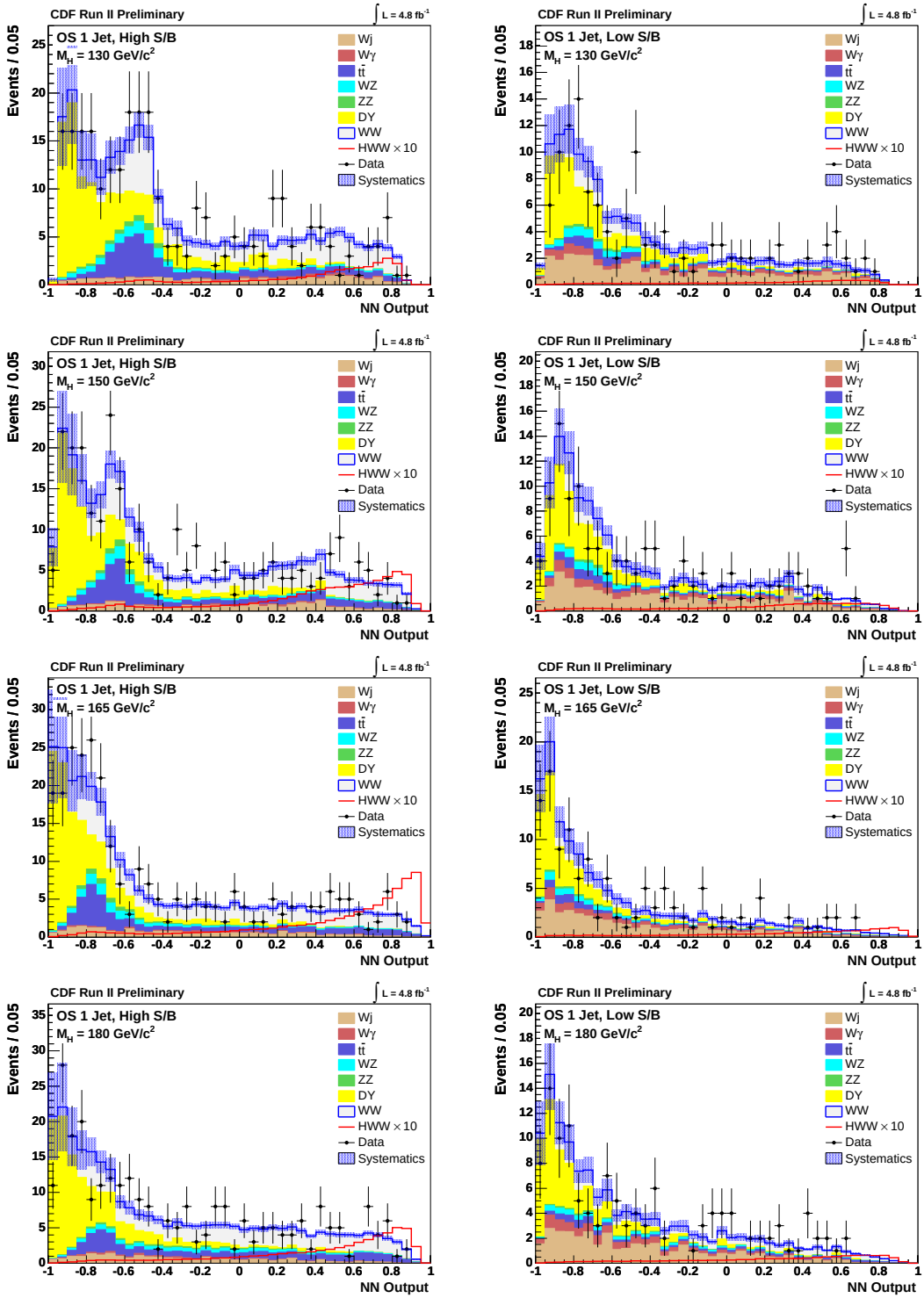


Figure 6.3: Expected and observed NN^{130}, NN^{150}, NN^{165} and NN^{180} output distributions in the opposite-sign region containing one reconstructed jet. Distributions on the left are High S/B and on the right for Low S/B.

for events with no reconstructed jets.

The effect of the higher order diagrams systematic is evaluated as in Section 5.3. We also assign the same 10% uncertainty to the Higgs associated production and vector boson fusion.

The $\vec{\not\!\!E}_T$ Modeling and Jet Modeling uncertainties are derived from the low $\vec{\not\!\!E}_T^{\rm Spec}$ control region as in Section 5.3. However their effect is found to be much larger for events with one reconstructed jet, with a variation of $\pm 30\%$ in the rate of Drell-Yan events expected.

Table 6.3 summarizes the systematic uncertainties for this search and their relative effect, in percent, on the number of expected signals and background events. In this table italic and bold values in the same row indicate separately correlated systematic uncertainties among different processes. All the other systematics are assumed uncorrelated.

6.2.3 Results

We use the Neural Network output distribution to set 95% upper limits on the Higgs production cross section. The technique used is described in Section 5.4 and already used for events with no reconstructed jets. The expected (median) and the observed limits are shown in Table 6.4 and in Figure 6.4 as function of the Higgs mass. The Standard Model production cross section is the sum of the four production mechanisms considered.

Even if the number of expected signal events is less than the case of zero jets, the expected limits are comparable. In fact the WW process is less likely to radiate a jet, since it is initiated by quarks, respect to gluons in the case of the Higgs direct production. Thus the ratio $\sigma(gg \to H \to WW)/\sigma(WW)$ is more favorable.

The observed limits agree within 1σ with the expected ones, with a small downward fluctuation at $m_H = 160$ GeV. The best limit is 1.82 times the expected Standard Model cross section for the most sensitive mass hypothesis $m_H = 165$ GeV.

6.3 High jet multiplicity events

The Higgs boson produced in association with W/Z and by their fusion, is likely to have jets in the final state. In this Section we describe the selection of events with opposite-sign leptons and with at least two jets.

6.3.1 Sample composition and signal kinematic

The di-lepton sample with $N_{jets} > 1$ is mainly constituted by $t\bar{t}$ events (see Figure 6.1), where both top quarks decay via $t \to Wb \to l\nu b$. The requirement of two jets is satisfied by the presence of the two $b-$quarks jets in the final state. On the other hand signal events are not expected to have big contribution from $b-$jets. Thus, to reduce this background we reject events that contain one or more b-tagged jets. The SECVTX(59) algorithm is used to identify b-jets as described in Section 3.5.3. This requirement removes 53% of the $t\bar{t}$ events, with only 4.7% loss of signal. Table 6.5 shows the expected and observed number of events in the signal region after the $b-$jet veto. Still $t\bar{t}$ remains the main background,

Uncertainty Source	WW	WZ	ZZ	$t\bar{t}$	DY	$W\gamma$	W+jet(s)	$gg \to H$	WH	ZH	VBF
Cross Section											
Scale	⇓⇓	⇓⇓	⇓⇓	⇓⇓	⇓⇓			23.5	⇓⇓	⇓⇓	⇓⇓
PDF Model	⇓⇓	⇓⇓	⇓⇓	⇓⇓	⇓⇓			7.7	⇓⇓	⇓⇓	⇓⇓
Total	*6.0*	*6.0*	*6.0*	10.0	5.0			24.7	**5.0**	**5.0**	10.0
Acceptance											
Scale (leptons)								2.8			
Scale (jets)								-5.1			
PDF Model (leptons)	1.9	2.7	2.7	2.1	4.1			1.7	1.2	0.9	2.2
PDF Model (jets)								-1.9			
Higher-order Diagrams	*5.0*	*10.0*	*10.0*	10.0		11.0			**10.0**	**10.0**	**10.0**
Missing Et Modeling					30.0						
$W\gamma$ Scaling						12.0					
Jet Fake Rates (Low/High S/B)							22.2/31.5				
Jet Modeling	-1.0					*15.0*					
MC Run Dependence	1.0										
Lepton ID Efficiencies	2.0	2.0	2.2	1.8	2.0			1.9	1.9	1.9	1.9
Trigger Efficiencies	2.1	2.1	2.1	2.0	3.4			3.3	2.1	2.1	3.3
Luminosity	*5.9*	*5.9*	*5.9*	*5.9*	*5.9*			*5.9*	*5.9*	*5.9*	*5.9*

Table 6.3: Systematic uncertainties for the opposite-sign di-lepton analysis with one reconstructed jet. All the values are symmetric and expressed in percent. Each row represent a source of systematic uncertainty, while columns represent the relative effect of that systematic uncertainty on the number of expected events for the given process.

OS 1 Jet	110	120	130	140	145	150	155
$-2\sigma/\sigma_{SM}$	28.43	9.33	4.42	2.90	2.53	2.06	1.73
$-1\sigma/\sigma_{SM}$	38.99	12.88	6.16	4.02	3.46	2.86	2.35
Median/σ_{SM}	**55.60**	**18.37**	**8.78**	**5.78**	**4.97**	**4.05**	**3.34**
$+1\sigma/\sigma_{SM}$	81.00	26.73	12.68	8.35	7.19	5.94	4.88
$+2\sigma/\sigma_{SM}$	117.74	38.55	18.06	12.02	10.15	8.48	6.91
Observed/σ_{SM}	**61.42**	**17.50**	**8.06**	**5.68**	**4.63**	**4.04**	**2.89**
	160	165	170	175	180	190	200
$-2\sigma/\sigma_{SM}$	1.37	1.30	1.52	1.81	2.13	3.37	4.55
$-1\sigma/\sigma_{SM}$	1.87	1.76	2.05	2.46	2.91	4.65	6.29
Median/σ_{SM}	**2.65**	**2.49**	**2.90**	**3.49**	**4.11**	**6.62**	**9.04**
$+1\sigma/\sigma_{SM}$	3.83	3.62	4.23	5.08	5.97	9.66	13.25
$+2\sigma/\sigma_{SM}$	5.43	5.19	5.96	7.18	8.64	13.91	18.79
Observed/σ_{SM}	**1.82**	**1.94**	**2.42**	**2.99**	**3.82**	**8.66**	**13.46**

Table 6.4: Expected and Observed limits on Higgs production cross section in the opposite sign di-lepton sample with one jet for different Higgs masses. Limits are in units of expected SM cross section.

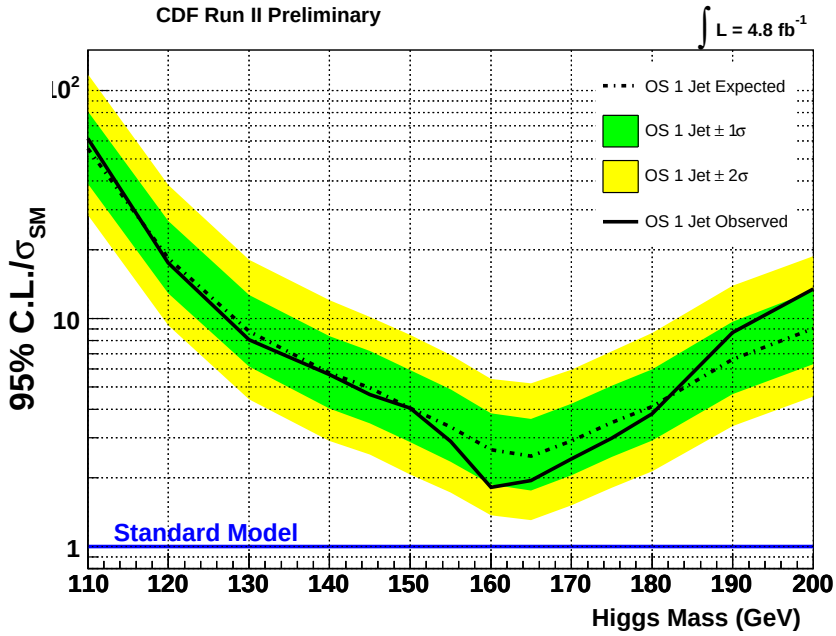


Figure 6.4: Expected and Observed limits on Higgs production cross section in the opposite sign di-lepton sample with one jet as function of the Higgs mass. Limits are in units of expected SM cross section.

CDF Run II Preliminary $\int \mathcal{L} = 4.8$ fb^{-1}			
$M_H = 165$ GeV/c^2			
$t\bar{t}$	145	$\pm$	24
DY	51	$\pm$	17
WW	25.6	$\pm$	5.8
WZ	5.30	$\pm$	0.73
ZZ	2.36	$\pm$	0.32
W+jets	21.9	$\pm$	5.9
$W\gamma$	2.72	$\pm$	0.67
Total Background	254	$\pm$	33
$gg \to H$	2.5	$\pm$	1.7
WH	1.90	$\pm$	0.25
ZH	0.99	$\pm$	0.13
VBF	1.04	$\pm$	0.17
Total Signal	6.4	$\pm$	1.8
Data	224		

OS 2+ Jets

Table 6.5: Expected and observed events in the opposite sign di-lepton sample with two or more reconstructed jets. Events with identified $b-$jets are removed.

being more than half of the selected sample. Since the fake background contribution is not important in this data set as it was for the other searches, we do not divide the sample into High S/B and Low S/B regions.

We use the Neural Network technique to increase the separation between signal and backgrounds. The training settings are analogous to what described for events with one reconstructed jet. The neural network input variables are:

- ΔR and $\Delta\phi$ between leptons;
- leptons p_T;
- $m(ll)$: invariant mass of the leptons;
- H_T: scalar sum of the leptons p_T, $\vec{\not\!\!E}_T$ and the jets E_T;
- $\Delta\phi$ between the vector sum of lepton momenta and $\vec{\not\!\!E}_T$;
- $p_T(j_1 j_2)$: the transverse component of the vector sum of the moments of the first and the second jets.

PYTHIA does not provide an accurate description of the angles between jets, when they both come from initial or final state radiation. Since some backgrounds and the Higgs signal produced by gluon fusion may pass the selections with two jets originating by initial state radiation, we explicitly avoided the usage of angular variables among jets as inputs of the NN.

The expected and observed kinematic distributions for the NN input variables in the selected sample are shown in Figure 6.5. We overlay the signal contribution separating direct production and the sum of associate production and the vector boson fusion. All the signals are multiplied by a factor 10.

$\Delta\phi$ and ΔR distributions between leptons provide the best discrimination between signal and backgrounds. The leptons p_T and their invariant mass show the same features outlined for events with one reconstructed jet in the final state. The angle between the vector sum of the two leptons p_T and the $\vec{\not{E}}_T$ improves the discrimination of signal from Drell-Yan events.

The H_T variable and the p_T of the di-jet system are very powerful to discriminate the remaining $t\bar{t}$ background, because they are expected larger in a top pair decay than in signal events. Direct Higgs production tends to have jets with lower energy respect to associate production and vector boson fusion, since they come from initial state radiation, resulting in smaller values of H_T.

The resulting expected and observed Neural Network output distributions are shown in Figure 6.6 for different mass hypotheses. They show that $t\bar{t}$ remains the dominant background contributor for the highest score bins of the NN output distributions.

6.3.2 $t\bar{t}$ control region

Events rejected by the $b-$quark veto provide a sample that can be used to verify the capability to model the $t\bar{t}$ background. Table 6.6 shows the number of expected and observed events with at least one identified $b-$jet. The sample is heavily dominated by top events, with a very small contribution from other backgrounds. The number of observed events agrees with the expectations. Figure 6.7 shows the behaviour of the kinematic distributions used as input of the neural network described in Section 6.3.1. The description of these variables provided by the simulation agrees with data.

6.3.3 Systematics

We considered the same sources of systematic errors described for the analysis of events with one reconstructed jet (see Section 6.2.2) and we evaluated their effect in the case of the two or more jets selections. It is remarkable that the Higgs direct production has an huge error on the theoretical cross section calculation, since in some sense it is now only a Leading Order calculation of the $gg \to H + 2jets$ process. An additional small systematic uncertainty comes from the modeling of $b-$jets rejection requirement, which affect $t\bar{t}$ normalization by $\pm 5.4\%$. The effect of the same sources of shape systematics considered in Section 5.3.4 are evaluated and found to give negligible effects on the limit calculation. Table 6.7 summarize the source of rate systematic uncertainties considered for this data set and their impact on the normalization of signal and backgrounds.

6.3.4 Results

As for the other selected samples we calculate the limits at 95% C.L. using the neural network output distributions as input of the calculation described in Section 5.4.

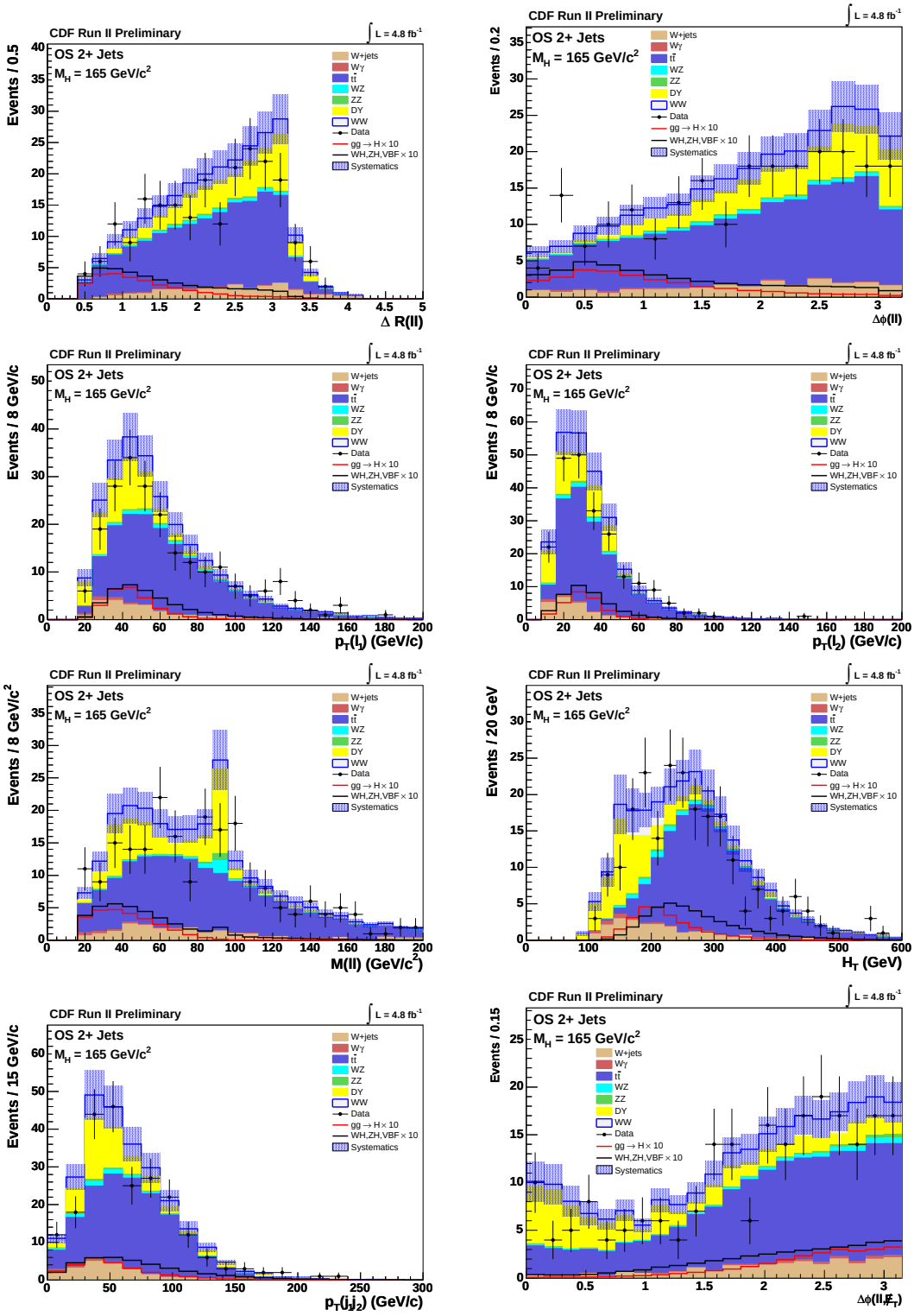


Figure 6.5: Signal and background kinematic variables in the opposite-sign signal region for events containing two or more reconstructed jet with $E_T > 15$ GeV and no identified $b-$jets.

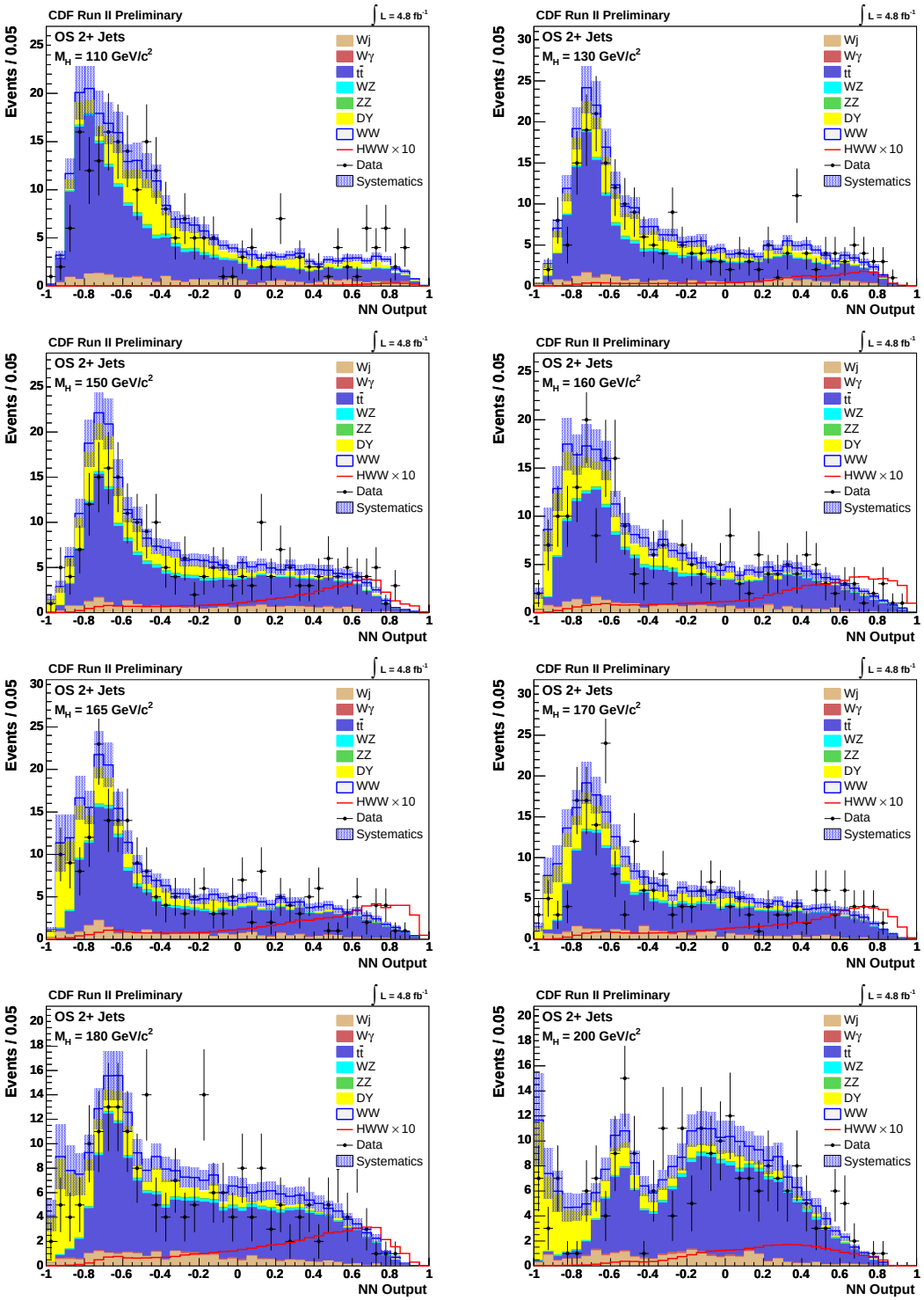


Figure 6.6: Expected and observed NN^{110}, NN^{130}, NN^{150}, NN^{160}, NN^{165}, NN^{170}, NN^{180} and NN^{200} output distributions in the opposite-sign region containing two or more reconstructed jets.

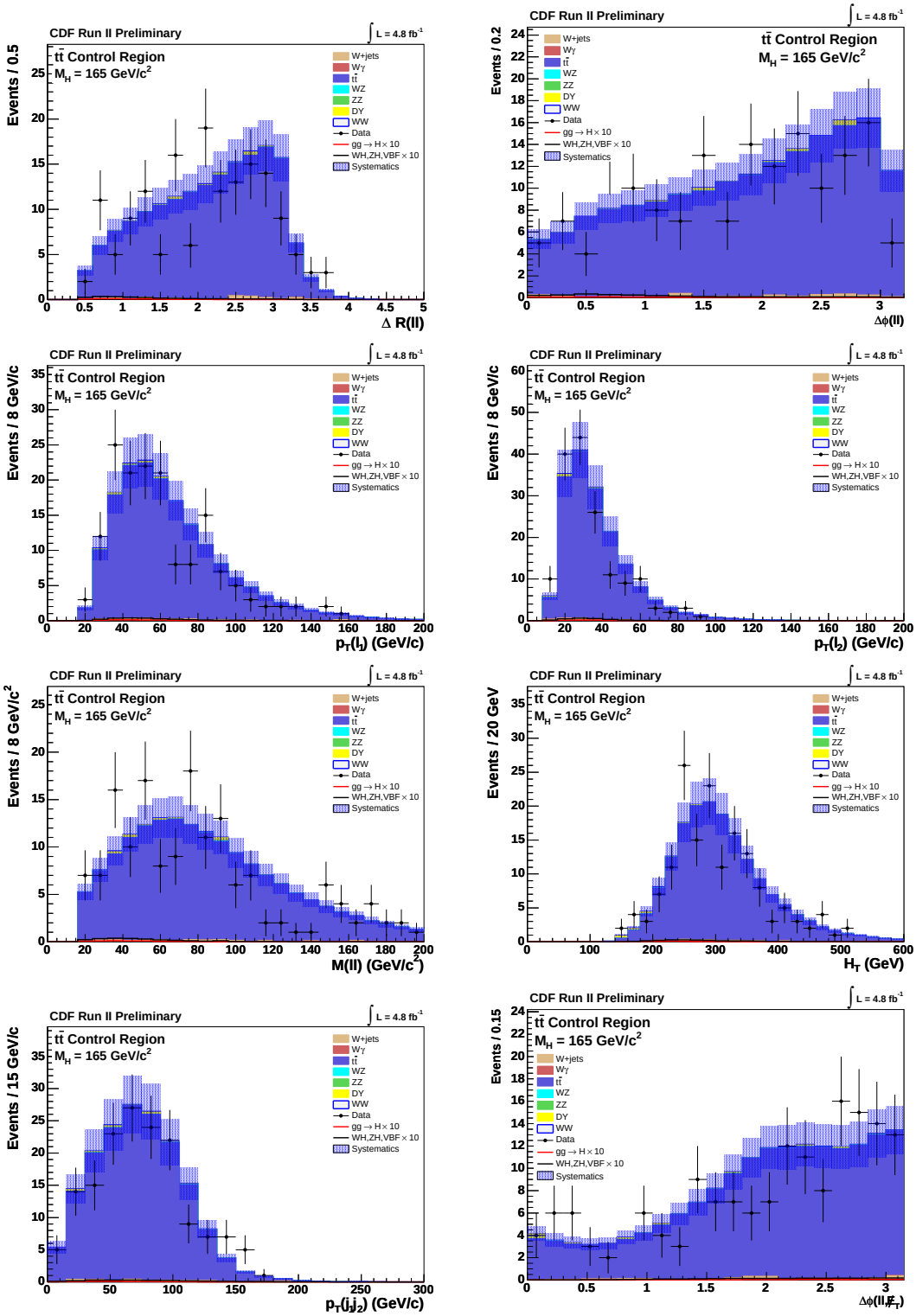


Figure 6.7: Kinematic distribution of expected and observed events in the $t\bar{t}$ control region.

CDF Run II Preliminary	$\int \mathcal{L} = 4.8$ fb^{-1}		
$M_H = 160$ GeV/c^2			
$t\bar{t}$	166	$\pm$	28
DY	1.24	$\pm$	0.41
WW	0.62	$\pm$	0.14
WZ	0.131	$\pm$	0.018
ZZ	0.135	$\pm$	0.018
W+jets	2.81	$\pm$	0.76
$W\gamma$	0.070	$\pm$	0.017
Total Background	171	$\pm$	28
$gg \rightarrow H$	0.053	$\pm$	0.037
WH	0.096	$\pm$	0.012
ZH	0.131	$\pm$	0.017
VBF	0.0135	$\pm$	0.0022
Total Signal	0.293	$\pm$	0.049
Data	159		

$t\bar{t}$ Control Region

Table 6.6: Expected and observed number of events in the $t\bar{t}$ control region.

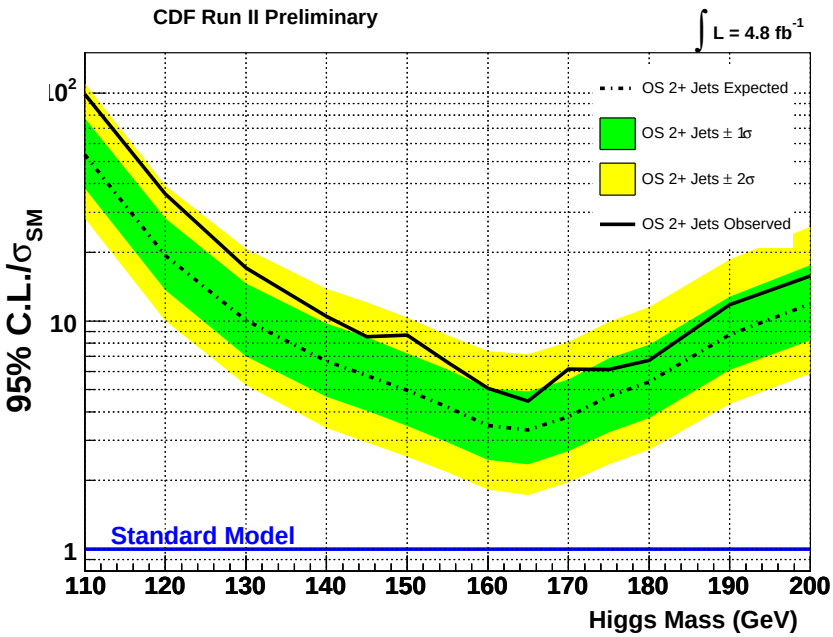


Figure 6.8: Expected and Observed limits on Higgs production cross section in the opposite sign di-lepton sample with two or more jets as function of the Higgs mass. Limits are in units of expected SM cross section.

Uncertainty Source	WW	WZ	ZZ	$t\bar{t}$	DY	$W\gamma$	W+jet(s)	$gg \to H$	WH	ZH	VBF
Cross Section											
Scale	⇊	⇊	⇊	⇊	⇊			67.5	⇊	⇊	⇊
PDF Model	⇊	⇊	⇊	⇊	⇊			7.7	⇊	⇊	⇊
Total	*6.0*	*6.0*	*6.0*	10.0	5.0			67.9	**5.0**	**5.0**	10.0
Acceptance											
Scale (leptons)								3.1			
Scale (jets)								-8.7			
PDF Model (leptons)	1.9	2.7	2.7	2.1	4.1			2.0	1.2	0.9	2.2
PDF Model (jets)								-2.8			
Higher-order Diagrams	*5.0*	*10.0*	*10.0*	10.0		11.0			**10.0**	**10.0**	**10.0**
Missing Et Modeling					32.0						
$W\gamma$ Scaling						12.0					
Jet Fake Rates							27.1				
Jet Modeling	20.0					*18.5*					
b-tag veto				5.4							
MC Run Dependence	1.5										
Lepton ID Efficiencies	1.9	2.9	1.9	1.9	1.9			1.9	1.9	1.9	1.9
Trigger Efficiencies	2.1	2.1	2.1	2.0	3.4			3.3	2.1	2.1	3.3
Luminosity	*5.9*	*5.9*	*5.9*	*5.9*	*5.9*			*5.9*	*5.9*	*5.9*	*5.9*

Table 6.7: Systematic uncertainties for the opposite-sign di-lepton analysis with two or more reconstructed jets. All the values are symmetric and expressed in percent. Each row represent a source of systematic uncertainty, while columns represent the relative effect of that systematic uncertainty on the number of expected events for the given process.

OS 2+ Jets	110	120	130	140	145	150	155
$-2\sigma/\sigma_{SM}$	28.23	10.03	5.22	3.39	2.93	2.54	2.17
$-1\sigma/\sigma_{SM}$	38.12	13.68	7.01	4.66	4.04	3.47	2.93
Median/$\sigma_{\mathbf{SM}}$	**53.54**	**19.36**	**10.05**	**6.68**	**5.77**	**4.97**	**4.19**
$+1\sigma/\sigma_{SM}$	77.35	28.17	14.59	9.73	8.48	7.20	6.12
$+2\sigma/\sigma_{SM}$	110.06	39.27	20.71	13.83	12.07	10.34	8.68
Observed/$\sigma_{\mathbf{SM}}$	**98.70**	**36.06**	**17.12**	**10.48**	**8.54**	**8.66**	**6.60**
	160	165	170	175	180	190	200
$-2\sigma/\sigma_{SM}$	1.83	1.72	1.96	2.35	2.74	4.33	5.88
$-1\sigma/\sigma_{SM}$	2.46	2.35	2.68	3.24	3.75	6.09	8.23
Median/$\sigma_{\mathbf{SM}}$	**3.49**	**3.33**	**3.80**	**4.67**	**5.41**	**8.70**	**11.87**
$+1\sigma/\sigma_{SM}$	5.07	4.90	5.54	6.85	7.88	12.77	17.58
$+2\sigma/\sigma_{SM}$	7.38	7.13	8.09	9.92	11.49	18.62	25.88
Observed/$\sigma_{\mathbf{SM}}$	**5.07**	**4.45**	**6.14**	**6.12**	**6.71**	**11.80**	**15.74**

Table 6.8: Expected and Observed limits on Higgs production cross section in the opposite sign di-lepton sample with two or more jets for different Higgs masses. Limits are in units of expected SM cross section.

Results are shown in Table 6.8 and Figure 6.8. The observed limits are $1-2\sigma$ higher than the expected ones, depending on the Higgs mass. This can be expected by looking at the neural network output distributions of Figure 6.6, where a small excess of events in the last bins of the distributions is evident. Still this effect is well within expected statistical fluctuations; Figure 6.9 shows for example the distribution of the expected limits for the $m_H = 165$ GeV hypothesis, in units of SM cross section, and the line indicates the observed one. The probability of having this or a higher observed limit if data contains no signal is 20%. The observed limits for other Higgs mass hypothesis are highly correlated, since they share the same data sample.

6.4 Same sign events

When the Higgs boson is produced in association with a W or Z it naturally can have two like-sign or same-sign (SS) leptons. These occur in the $ZH \rightarrow ZWW$ production, when the Z and one or both W bosons decay leptonically, and in the $WH \rightarrow WWW$ production, when the associated W and one or both W bosons from the Higgs decay leptonically.

The event selection for same sign di-lepton events focuses on reducing backgrounds from fake leptons and leptons with mis-measured charge. Because the PHX electrons have an high rate of charge mis-measurement, they are not used.

The Drell-Yan contributes to the same-sign region only when the lepton charge is mis-measured, and it is not as large contributor as in the opposite-sign regions. Therefore we remove the $\not\!\!\vec{E}_T^{\rm Spec}$ requirement. To reduce the number of fake leptons in the sample, the minimum $p_{\rm T}$ of the second lepton is raised from 10 GeV/c to 20 Gev/c. To reduce

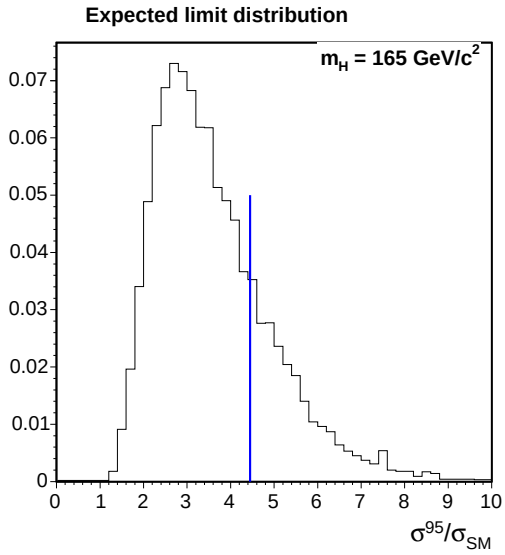


Figure 6.9: Distribution of 95% C.L. expected limits (σ^{95}) for $m_H = 165$ GeV hypothesis in units of the SM cross section (σ_{SM}). The blue line show the observed outcome.

the W+jets and the $W\gamma$ contributions, we only use events with one or more jets with $E_{\mathrm{T}} > 15$ GeV and $|\eta| < 2.5$. Same sign events with zero jets are used as a control sample, as described in Section 4.4.1.

Table 6.9 shows the number of expected and observed events after selections.

6.4.1 Background modeling

After the cleaning requirements the W+jets component accounts for about 50% of the background in the same-sign data sample. The Drell-Yan contributes for 25% of the total background. Other background contributions come from WW, WZ, ZZ, $W\gamma$, and $t\bar{t}$ processes.

The WW contribution is expected to be very small. In 4.8 fb^{-1} only ~ 0.04 WW events are expected. However, since it is expected to look very signal-like we have to include it in the background. To model WW kinematic properties we use the WW MC@NLO sample, but very few events pass the selection, not enough to build a smooth NN distribution. To make the NN distribution, we instead use events which satisfy the same-sign requirements, except the same signs. We then scale the template obtained with these events by the charge fake rate, i.e. the probability that the charge of the lepton is mis-measured. The charge fake rate (CFR) is measured using the MC $Z \to ee$, $Z \to \mu\mu$ and WW Pythia samples. Using the MC truth information, we count the number of lepton pairs with one lepton which charge is wrongly assigned by the detector simulation, and divide it by the total number of events. This ratio is the charge fake rate. Table 6.10 shows the charge fake rate for each di-lepton type. Although we do not use PHX electrons, their CFR are included for completeness: their CFRs show that the PHX inclusion would lead to significant increase of the background coming from mis-measured leptons.

CDF Run II Preliminary	$\int \mathcal{L} = 4.8$ fb^{-1}		
$M_H = 165$ GeV/c^2			
$t\bar{t}$	0.242	$\pm$	0.068
DY	26.7	$\pm$	8.1
WW	0.039	$\pm$	0.010
WZ	9.5	$\pm$	1.3
ZZ	1.98	$\pm$	0.27
W+jets	34	$\pm$	10
$W\gamma$	4.34	$\pm$	0.99
Total Background	76	$\pm$	13
WH	1.61	$\pm$	0.21
ZH	0.261	$\pm$	0.034
Total Signal	1.87	$\pm$	0.24
Data	81		

SS 1+ Jets

Table 6.9: Number of expected and observed events in Same-Sign region with one or more reconstructed jets.

di-lepton type	Opposite-sign events	Same-sign events	Charge fake rate
LBE-LBE	1804738	914	5.06e-4
LBE-Central μ	793007	270	3.40e-4
LBE-Forward μ	119401	38	3.18e-4
LBE-CrkTrk	1223739	490	4.00e-4
Central μ-Central μ	2350202	1	4.25e-7
Central μ-Forward μ	822095	54	6.57e-5
Central μ-CrkTrk	1109896	35	3.15e-5
PHX-PHX	442445	120832	0.273
PHX-Central μ	266784	34292	0.129
PHX-Forward μ	43481	5922	0.136
PHX-CrkTrk	574412	77449	0.135
LBE-PHX	1781103	228155	0.128

Table 6.10: Charge fake rates for each di-lepton type. "Central μ" includes CMUP, CMP, CMX, CMXMsKs, CMIOCES, and CMP muons, while "Forward μ" includes CMIOPES and BMU muons.

6.4.2 Kinematic properties

In the same-sign signal events the leptons are not any more correlated as they were in the opposite sign events. In fact they are not expected to come both from the decay of the Ws which come from the Higgs. Thus, the $\Delta\phi(ll)$ distribution now shows a back-to-back signature for signal events too. This is visible in Figures 6.10 and 6.11, that show the expected and observed distributions of important kinematic variables in the selected sample. The leptons p_T and energy, the di-lepton invariant mass and the p_T of the di-lepton system help to discriminate fake background from signal. The jets energy spectrum is expected to be harder in signal events respect to backgrounds; this is visible in the distributions of the energy of the highest E_T (first) jet and of the sum of jets E_T and H_T. Finally, $\vec{\not\!\!E}_T$ related variables offer also a good discrimination against fake and DY background.

We use these variables as input to a Neural Network trained to distinguish WH and ZH signal events from backgrounds. The output distributions for some Neural Networks out of the fourteen trained, one for each Higgs mass hypothesis, are shown in Figure 6.12. They show that WZ contribution is an important background too, when one of the Z leptons is lost, sitting at high values of the NN score distribution.

6.4.3 Systematics

Systematic uncertainties for the same-sign analysis are summarized in Table 6.11. Most of the them are the same of the opposite sign channels. Because we do not make a $\vec{\not\!\!E}_T$ cut, we do not have a systematic due to $\vec{\not\!\!E}_T$ modeling.

WW, Drell-Yan, and $t\bar{t}$ contribute to the same-sign analysis only when the charge of a lepton is mis-measured. These backgrounds are estimated using Monte Carlo samples; this assumes that the MC accurately models how often lepton charges are mis-measured in the data. To test this assumption, we measure the charge fake rate from the data. We select events in the Drell-Yan control region, described in Sec. 4.3.2, which should have minimal contamination from other backgrounds such as W+jets, where same-sign events come from QCD rather than charge mis-measurement. We then count the number of opposite-sign events and same-sign events to obtain the charge fake rates. Due to poor statistics these measurements do not give a reliable estimate for most of the di-lepton categories. We take the maximum difference between the charge fake rate measured in data and MC as systematic uncertainty.

6.4.4 Results

We see no significant excess of data respect to background expectation in the neural network outputs of Figure 6.12. Thus we set 95% C.L. limits on the production cross section of the Higgs boson for each of the fourteen mass hypotheses, using the same technique as for opposite sign analyses. Table 6.12 and Figure 6.13 show the expected and observed limits.

The maximum sensitivity is obtained for $m_H = 160$ GeV, with an expected limit 6.2 σ_H^{SM}. It is interesting to compare the sensitivity of this search to the one obtained

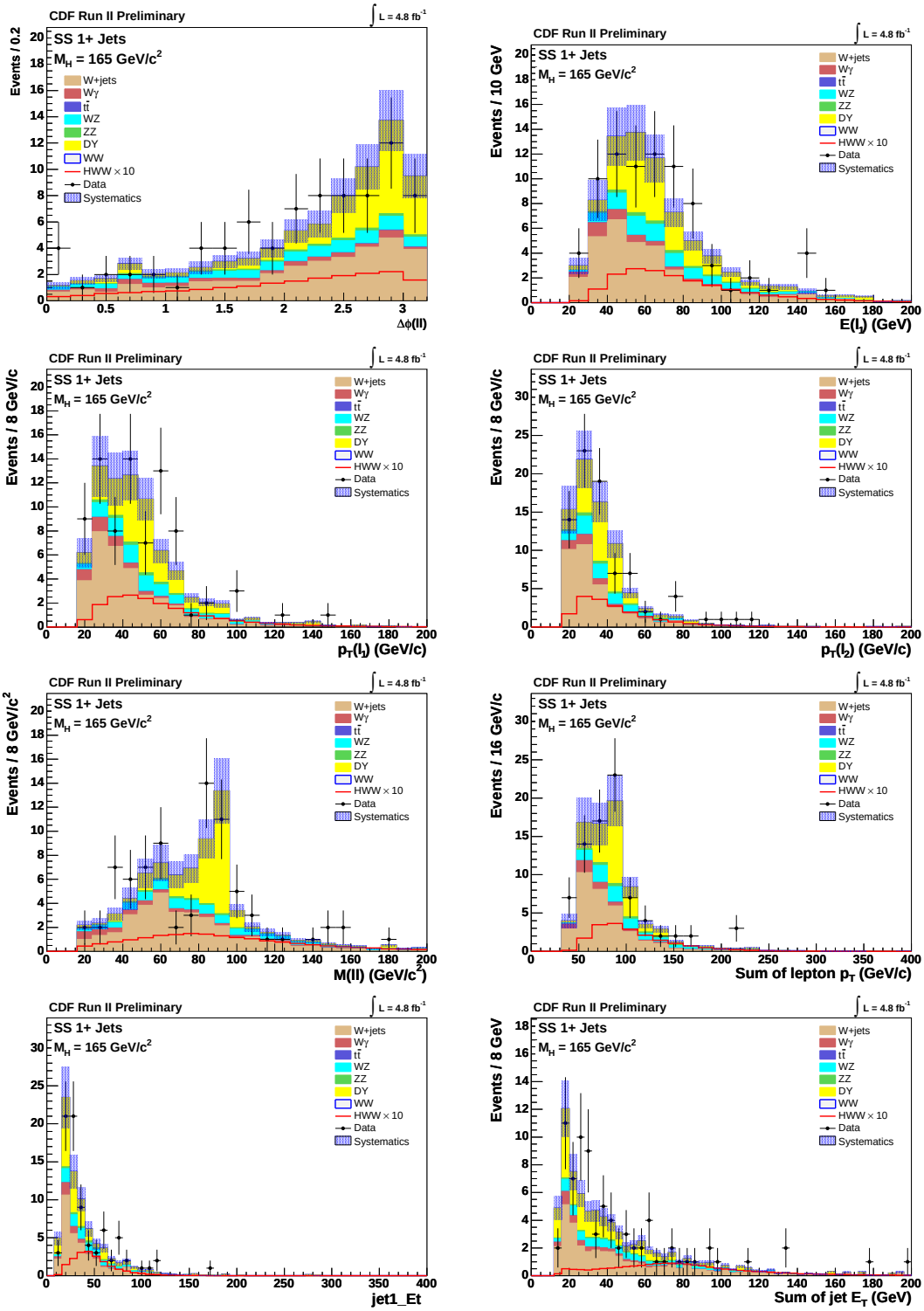


Figure 6.10: Kinematic distributions of expected and observed events in the Same-Sign di-lepton sample with one or more jets.

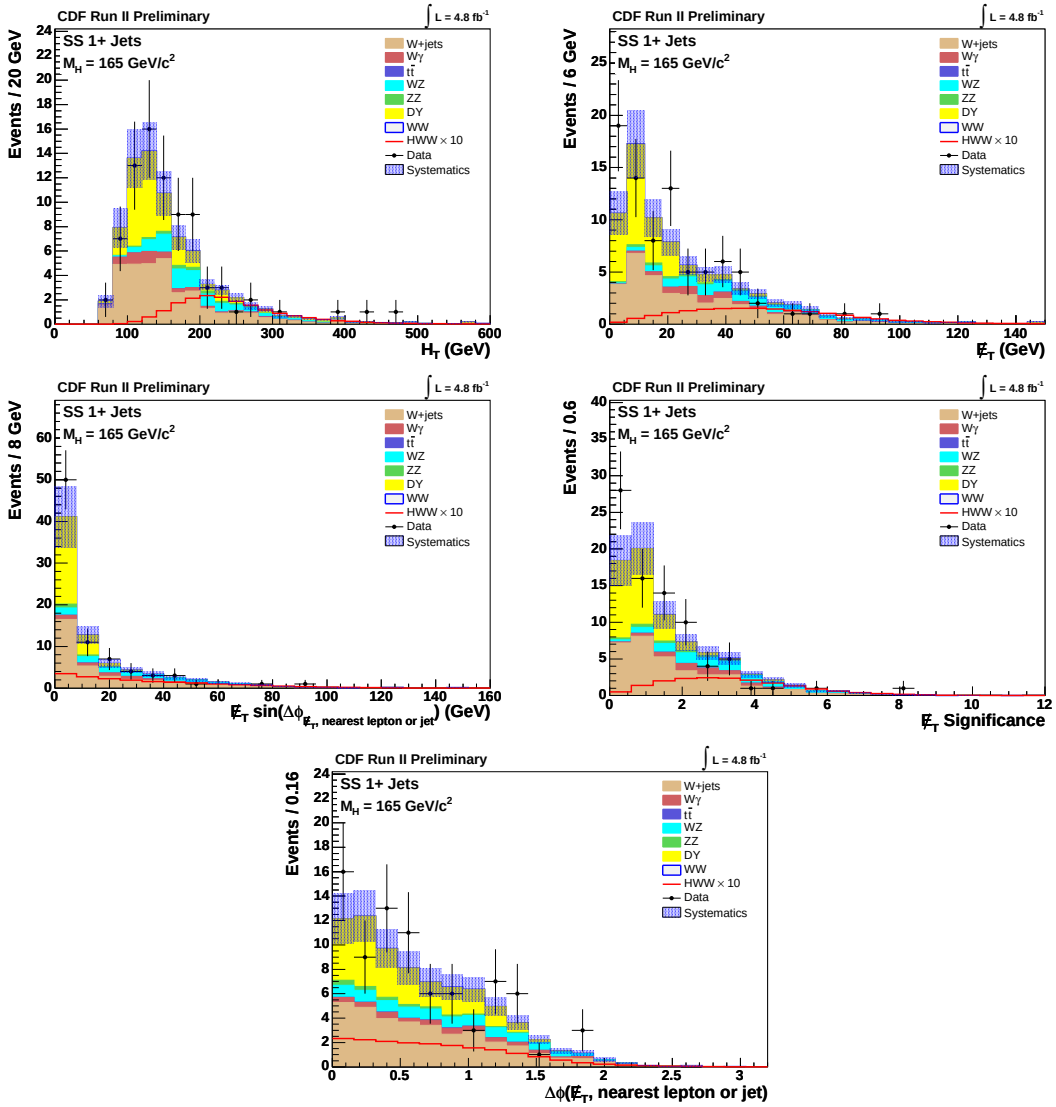


Figure 6.11: Kinematic distributions of expected and observed events in the Same-Sign di-lepton sample with one or more jets.

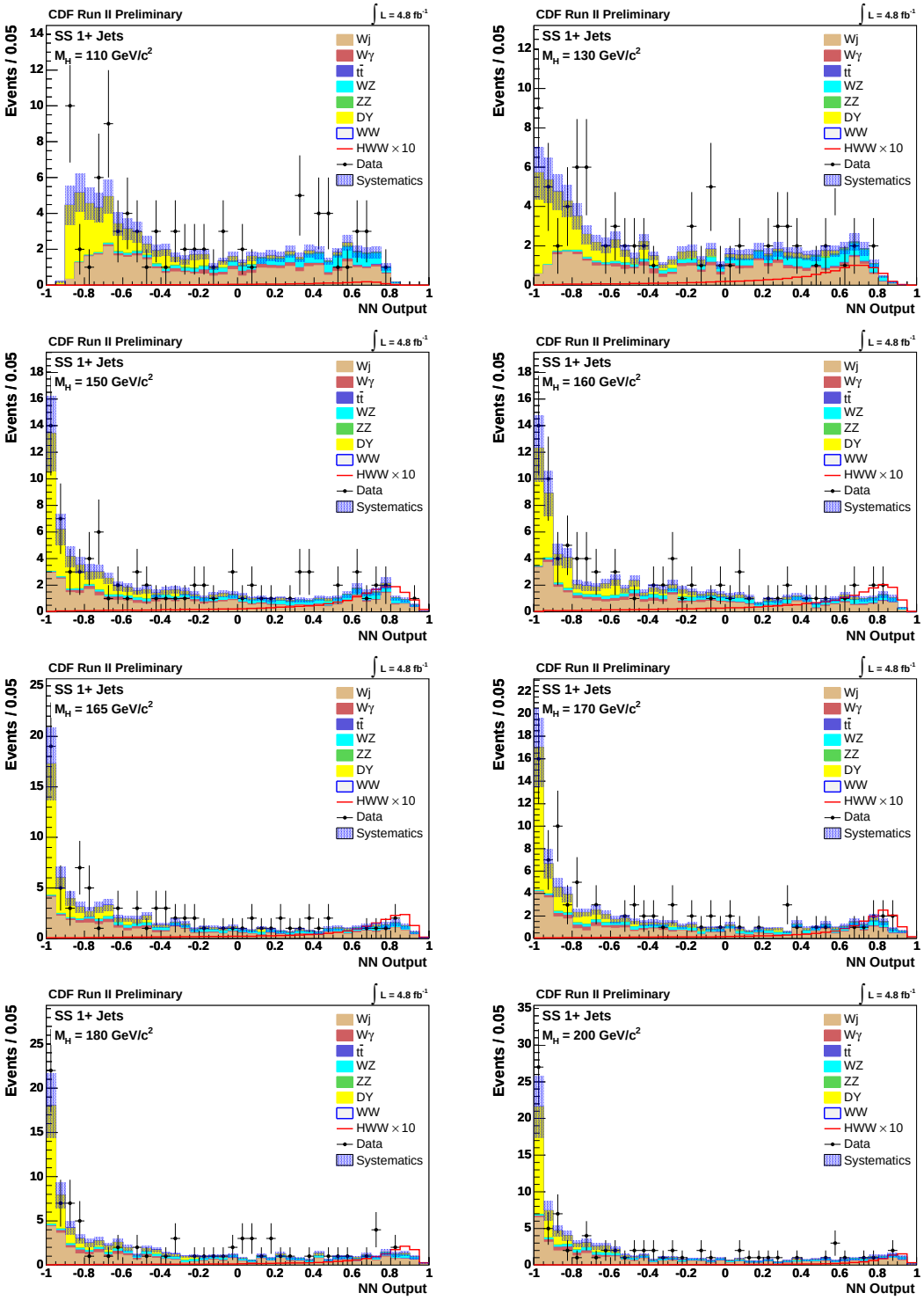


Figure 6.12: Expected and observed NN^{110}, NN^{130}, NN^{150}, NN^{160}, NN^{165}, NN^{170}, NN^{180} and NN^{200} output distributions in the Same-Sign region containing one or more reconstructed jets.

Uncertainty Source	WW	WZ	ZZ	$t\bar{t}$	DY	$W\gamma$	W+jet(s)	WH	ZH
Cross Section									
Scale	⇓⇓	⇓⇓	⇓⇓	⇓⇓	⇓⇓			⇓⇓	⇓⇓
PDF Model	⇓⇓	⇓⇓	⇓⇓	⇓⇓	⇓⇓			⇓⇓	⇓⇓
Total	*6.0*	*6.0*	*6.0*	10.0	5.0			**5.0**	**5.0**
Acceptance									
Scale (leptons)									
Scale (jets)									
PDF Model (leptons)	1.9	2.7	2.7	2.1	4.1			1.2	0.9
PDF Model (jets)									
Higher-order Diagrams	*5.0*	*10.0*	*10.0*	10.0		11.0		**10.0**	**10.0**
$W\gamma$ Scaling						12.0			
Jet Fake Rates							30.0		
Jet Modeling	3.0				17.0	16.0			
Charge Misassignment	16.5			16.5	16.5				
MC Run Dependence	1.0								
Lepton ID Efficiencies	2.0	2.0	2.0	2.0	2.0			2.0	2.0
Trigger Efficiencies	2.1	2.1	2.1	2.0	3.4			2.1	2.1
Luminosity	*5.9*	*5.9*	*5.9*	*5.9*	*5.9*			*5.9*	*5.9*

Table 6.11: Systematic uncertainties for the same-sign di-lepton analysis. All the values are symmetric and expressed in percent. Each row represent a source of systematic uncertainty, while columns represent the relative effect of that systematic uncertainty on the number of expected events for the given process.

SS 1+ Jets	110	120	130	140	145	150	155
$-2\sigma/\sigma_{SM}$	40.99	14.55	6.99	4.68	4.19	3.67	3.39
$-1\sigma/\sigma_{SM}$	57.25	20.13	9.52	6.46	5.75	5.03	4.64
Median/$\sigma_{\mathbf{SM}}$	**81.63**	**29.21**	**13.56**	**9.27**	**8.20**	**7.14**	**6.68**
$+1\sigma/\sigma_{SM}$	119.94	42.49	19.59	13.47	11.91	10.35	9.84
$+2\sigma/\sigma_{SM}$	168.11	58.93	27.14	19.01	16.79	14.57	13.77
Observed/$\sigma_{\mathbf{SM}}$	**94.78**	**27.79**	**13.33**	**9.97**	**8.13**	**6.67**	**7.11**
	160	165	170	175	180	190	200
$-2\sigma/\sigma_{SM}$	3.19	3.26	3.45	3.63	4.01	5.25	6.44
$-1\sigma/\sigma_{SM}$	4.37	4.46	4.68	4.94	5.40	6.90	8.64
Median/$\sigma_{\mathbf{SM}}$	**6.20**	**6.41**	**6.65**	**6.99**	**7.70**	**9.74**	**12.14**
$+1\sigma/\sigma_{SM}$	8.90	9.23	9.60	10.03	11.07	14.00	17.70
$+2\sigma/\sigma_{SM}$	12.56	13.20	13.59	14.31	15.50	19.42	24.92
Observed/$\sigma_{\mathbf{SM}}$	**5.72**	**4.87**	**6.67**	**6.46**	**8.12**	**7.81**	**10.74**

Table 6.12: Expected and Observed limits on Higgs production cross section in the same sign di-lepton sample with one or more jets for fourteen Higgs mass hypotheses. Limits are in units of expected SM cross section.

in the opposite sign analysis in events with zero jets. For the most sensitive hypothesis $m_H \sim 160$ GeV the same-sign analysis has an expected limit a factor ~ 3 larger than the opposite sign analysis. However, for light ($m_H \sim 110 - 130$ GeV) and heavy ($m_H \sim 180 - 200$ GeV) Higgs, the expected limits are larger only by a factor of two. There are three main effects concurring to this result. Firstly the backgrounds are completely different, and the discrimination between signal and fake background is more and more easy the heavier is the Higgs boson. Secondly the cut on the di-lepton invariant mass for the opposite sign analysis is less efficient for light Higgs, being for example 90% efficienct if $m_H = 120$ GeV and 95% if $m_H = 160$ GeV ; in the same-sign analysis the efficiency of this cut is flatter as function of the Higgs mass, since leptons do not come from the Higgs decaying to Ws. Finally the opening angle between leptons of the opposite sign analysis is expected to be a strong discriminant between signal and backgrounds if both W are near to be on-shell ($m_H \sim 160$ GeV); for light Higgs the discrimination power is less, causing the expected limit to become larger (see Figure 5.3). This effect is not present in the same-sign analysis and it does not help in discriminating signal from backgrounds, causing the expected limit to have a flatter behaviour as function of the Higgs mass than for the opposite-sign ones.

We therefore expect the same-sign analysis to significantly contribute to the Higgs search in the $m_H \sim 110 - 130, 180 - 200$ GeV mass region if combined with the other ones.

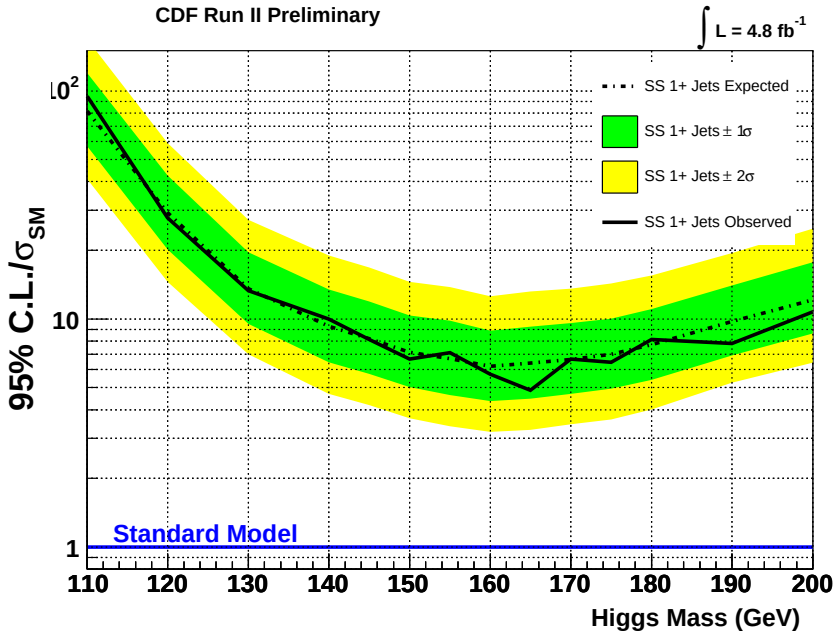


Figure 6.13: Expected and Observed limits on Higgs production cross section in the same sign di-lepton sample with one or more jets as function of the Higgs mass. Limits are in units of expected SM cross section.

Chapter 7

Discussion of the results

The searches for the Higgs boson performed in Chapters 5 and 6 are combined together in this Chapter. The combination of all CDF Higgs searches and of CDF and DØ results are also described. We will discuss the impact in the combination of the results presented in this Thesis, the improvements expected in the next one or two years and the final CDF reaches for the high mass Higgs.

7.1 The $H \to W^+W^-$ combined production cross-section limits

The combination of the results obtained in this Thesis is performed with the same method used to extract the individual results, and described in detail in Section 5.4. We combine the search for the Higgs boson:

- produced by gluon fusion and reconstructed in events with two opposite sign leptons and no jets in the final state (Chapter 5);
- produced by gluon fusion, associated productions and vector boson fusion, and reconstructed in events with opposite sign leptons and one jet in the final state (Section 6.2);
- produced by gluon fusion, associated productions and vector boson fusion, and reconstructed in events with opposite sign leptons and two or more jets in the final state (Section 6.3);
- produced in association with a W or Z, and reconstructed in events with two same sign leptons with one ore more jets in the final state (Section 6.4).

These searches, in the following also referred to as *channels*, have been performed on statistically independent samples. Therefore we extract the 95% C.L. upper limits using as input of equation 5.5 the Neural Network output distributions of each channel for the corresponding mass hypothesis. The total number of channels combined is six, separately including the High S/B and Low S/B regions for the searches in the samples with opposite-sign leptons an zero or one jet. The limits are evaluated for the fourteen mass hypotheses ranging from $m_H = 110$ GeV to $m_H = 200$ GeV.

As described in Section 5.4, this method easily accounts for systematic uncertainties correlations. Thus it is important to define the systematics of Tables 5.4, 6.3, 6.7 and 6.11 that are correlated among the independent searches performed. The systematic uncertainties that have the same name are assumed correlated. The only exception is the $\vec{\not\!\!E}_T$ Modeling systematic uncertainty, which is assumed uncorrelated among the different jet multiplicities. The systematics with opposite sign values in the Tables are considered anti-correlated. One example is the "Scale (jets)" systematic, which is evaluated re-weighting the Higgs p_T distribution and it causes events migration from one jet multiplicity to another: the total number of events does not change, so that if some event moves from having, for example, zero to one reconstructed jet above $E_T > 15$ GeV the errors on the zero and one jet analysis are anti-correlated. The same happens for the "PDF Model (jets)" systematic and for the Jet Modeling systematics, all of which just lead to events migration from one jet multiplicity to another.

The combined upper limits on the ratio of the Higgs production cross section to the Standard Model expectation are presented for each mass hypothesis in Table 7.1 and in Figure 7.1. The combination sets a 95% C.L. limit that is 1.3 times higher than the expected Standard Model cross section for the most sensitive hypothesis $m_H = 165$ GeV, with a median expected limit of 1.2 σ_H^{SM} in a background only hypothesis. The observed limits well agree with the expectations for all the Higgs mass hypotheses. It is remarkable to see that the median sensitivity of the combination is below two times the expected Standard Model cross section for the range $150 < m_H < 180$ GeV.

Figure 7.1 also shows, for comparison, the median expected upper limits for the individual searches entering the combination. The expected limits for opposite sign analyses have analogous behaviour as function of the Higgs mass. However the comparison shows that the $2+Jets$ analysis becomes less sensitive to an heavy Higgs boson faster than 0-jets and 1-jet analyses. In fact, the main background is $t\bar{t}$, which is discriminated mainly for having larger H_T (see Figure 6.5); the heavier is the Higgs boson the larger is the value at which the H_T distribution peaks for signal events (see Figure 5.3 as an example), decreasing the separation between signal and $t\bar{t}$ background. As expected from the discussion in Section 6.4.4, the same-sign analysis particularly contributes to increase the combination sensitivity for heavy Higgs mass hypotheses.

Overall the combined result greatly improves the sensitivity of the individual searches, and set the new world best limits by a single experiment on the direct search for the High mass Higgs boson.

7.2 The CDF Higgs analyses combination

The Higgs searches performed at CDF are combined together to probe the Higgs production in the mass range accessible at the Tevatron(77). The decay branching fraction of the Standard Model Higgs boson to the different decay channels is function of its mass m_H; thus the different search channels contribute in a complementary way to the sensitivity at different m_H.

The sensitivity for the low mass Higgs boson search strongly take advantage of the combination, since none of the existing analysis dominates. The most sensitive ones search

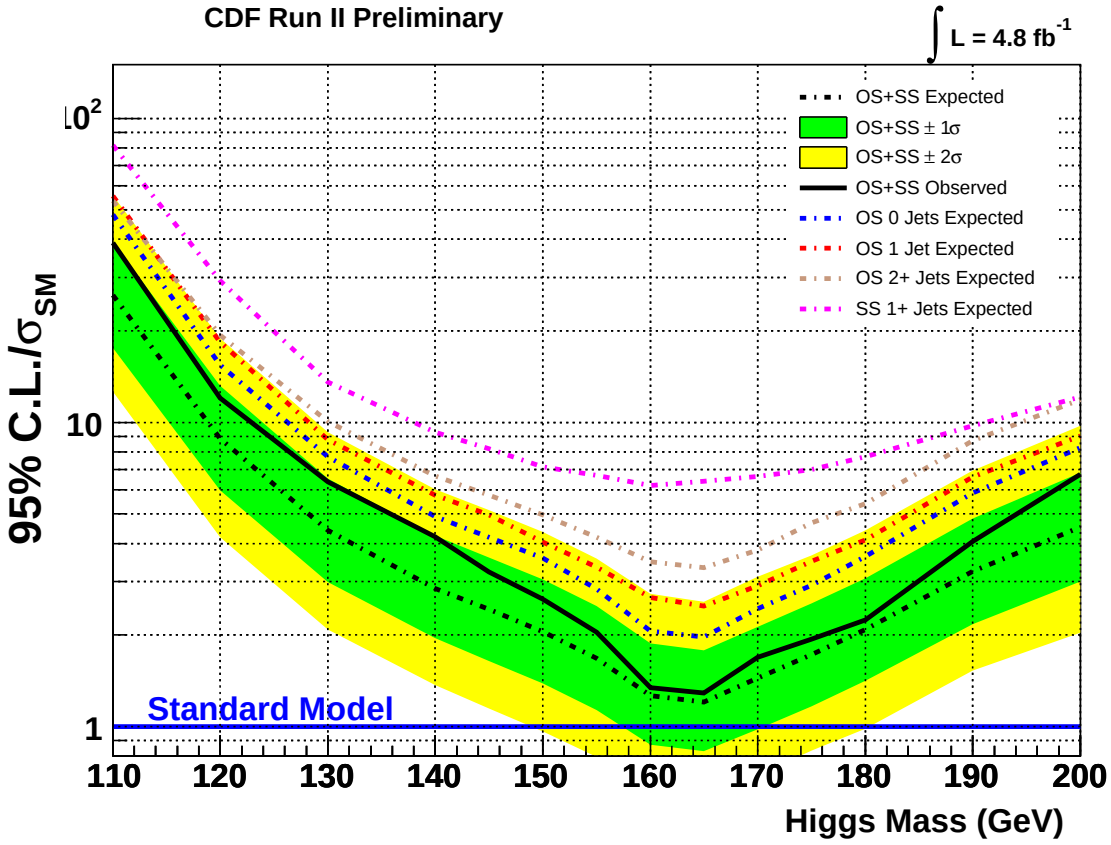


Figure 7.1: The expected and observed limits on the Higgs production cross section determined with the combination of the opposite sign (OS) and same sign (SS) di-lepton analyses as function of the Higgs mass. The median expected limits of the individual analyses entering the combination is shown for comparison. The limits are in units of expected SM cross section.

OS+SS	110	120	130	140	145	150	155
$-2\sigma/\sigma_{SM}$	12.65	4.19	2.09	1.37	1.15	0.96	0.79
$-1\sigma/\sigma_{SM}$	17.57	5.97	2.99	1.95	1.64	1.39	1.13
Median/$\sigma_{\mathbf{SM}}$	**26.27**	**8.85**	**4.41**	**2.85**	**2.43**	**2.05**	**1.67**
$+1\sigma/\sigma_{SM}$	39.09	13.19	6.54	4.24	3.60	3.05	2.49
$+2\sigma/\sigma_{SM}$	55.73	18.92	9.28	6.03	5.15	4.36	3.57
Observed/$\sigma_{\mathbf{SM}}$	**38.89**	**12.04**	**6.38**	**4.21**	**3.23**	**2.62**	**2.04**
	160	165	170	175	180	190	200
$-2\sigma/\sigma_{SM}$	0.62	0.59	0.69	0.83	0.98	1.53	2.03
$-1\sigma/\sigma_{SM}$	0.87	0.83	0.98	1.16	1.41	2.17	3.00
Median/$\sigma_{\mathbf{SM}}$	**1.26**	**1.20**	**1.44**	**1.72**	**2.09**	**3.24**	**4.53**
$+1\sigma/\sigma_{SM}$	1.88	1.78	2.12	2.54	3.07	4.82	6.84
$+2\sigma/\sigma_{SM}$	2.73	2.57	3.11	3.65	4.42	6.95	9.74
Observed/$\sigma_{\mathbf{SM}}$	**1.34**	**1.29**	**1.69**	**1.94**	**2.24**	**4.06**	**6.74**

Table 7.1: The expected and observed limits on the Higgs production cross section determined by combining the opposite sign (OS) and same sign (SS) di-lepton analyses for fourteen Higgs mass hypotheses. Limits are in units of expected SM cross section.

for the $H \to b\bar{b}$ produced in association with a W or Z. Then they differentiate in the decay channel of the associated gauge boson (25)(26)(27). Not considering the range of Higgs masses already excluded by LEP, these analyses have the best sensitivity for the Higgs mass $m_H \sim 115$ GeV, and loose sensitivity for Higgs boson heavier than that, due to the rapid decrease of the $H \to b\bar{b}$ branching fraction and of the Higgs production cross section.

The technique described in Section 5.4 is used to combine the CDF results. Most of the systematic uncertainties of the analyses are taken as uncorrelated, except the luminosity measurement uncertainty and the cross section uncertainties, which are supposed 100% correlated. The combination is performed for 19 Higgs mass hypotheses ranging from $m_H = 100$ GeV to $m_H = 200$ GeV, usually with a 5 GeV step. Low mass Higgs analyses are considered only for $m_H < 150$ GeV. The High mass Higgs searches are considered in the entire mass range; however for low Higgs mass values some of the mass hypotheses are missing. To fill these gaps, the missing Higgs mass hypotheses histograms are extrapolated from the neighbor masses(77).

Figure 7.2 shows the combined 95% C.L. upper limit on the Higgs production cross section; it also shows the single median expected and observed limits for the analyses entering into the combination.

The CDF combination includes the $H \to WW$ "lowMll" analysis in the High mass region, in addition to the searches presented in this Thesis (shown in the Figure as opposite sign and same sign $H \to WW$). The "lowMll" analysis searches for the Higgs direct production with the decay into a WW pair and the subsequent leptonic W decays, in analogy of what presented in Chapter 5. This analysis complements the ones presented in this Thesis selecting events in the $m(ll) < 16$ GeV region, taking advantage of the

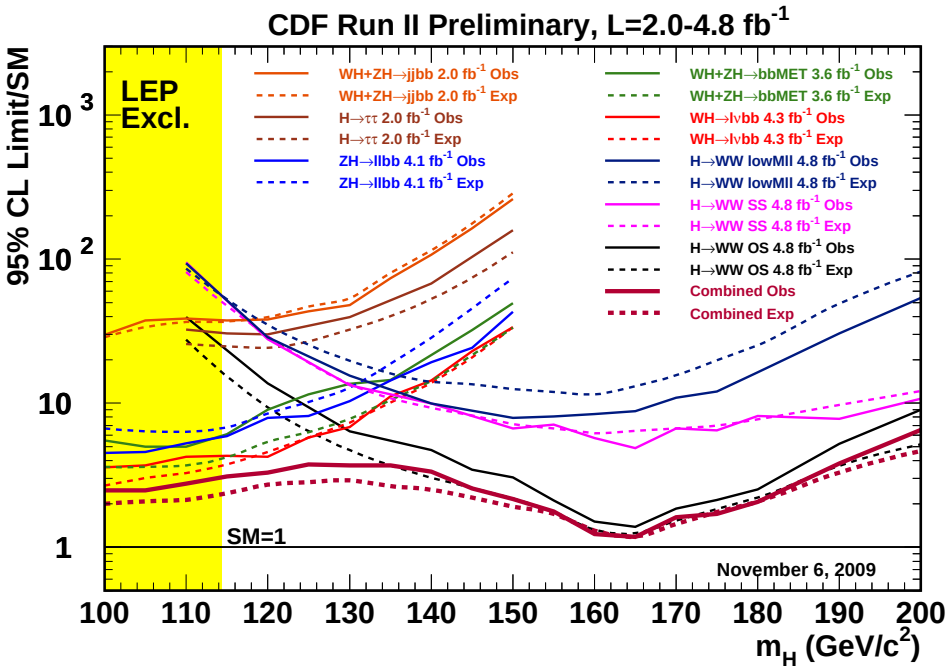


Figure 7.2: The expected and observed limits on the Higgs production cross section obtained by combining the CDF Higgs results as function of the Higgs mass. The limits are in units of expected SM cross section.

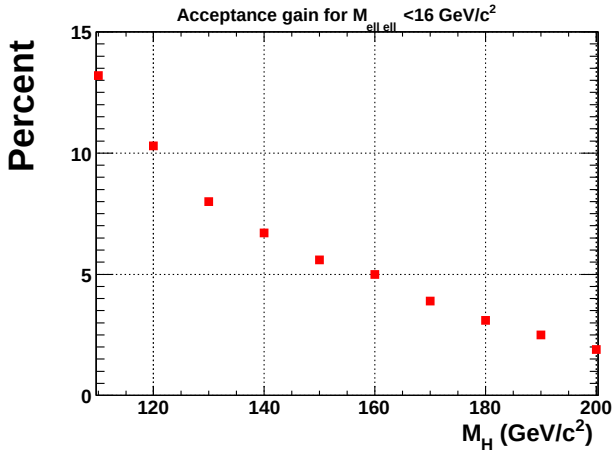


Figure 7.3: Expected fraction of reconstructed $gg \to H \to W^+W^- \to l^+l^-\nu\nu$ events with $m(ll) < 16$ GeV.

lower efficiency of this selection for a light Higgs, as shown in Figure 7.3. It improves the sensitivity of the $H \to WW$ analyses combination up to 6.3% for light Higgs, while it does not significantly contribute if $m_H \gtrsim 160$ GeV.

The branching fraction $H \to WW$ becomes greater than $H \to b\bar{b}$ only for $M_H \gtrsim 135$ GeV, as shown in Figure 2.3. However the comparison of the expected limits between the low and high mass analyses in Figure 7.2 shows that the $H \to WW$ search is the most sensitive analysis at $m_H \sim 125$ GeV. This can be expected from the Higgs production rates. In fact the number of $H \to b\bar{b}$ events produced by associated production, that are searched for in the low mass Higgs analyses, is greater than the one obtained by $H \to WW$ direct production for $m_H \gtrsim 120$ GeV (see Figure 2.4). It is therefore clear that although the $H \to WW$ analyses are commonly referred as high mass Higgs searches, they give an important contribution also for the search of a light Higgs, being the most sensitive analyses for most of the mass range bounded by the LEP direct searches and the indirect constraints.

7.3 Tevatron combination

The DØ collaboration performs searches of the Standard Model Higgs boson using analogous analysis techniques(78)(79). The analyzed data have a similar integrated luminosity and they are statistically independent from CDF. We can therefore combine the results obtained by both experiments into one single Tevatron result(80)(81). The combination technique uses the Bayesian approach described in Section 5.4. A modified frequentist approach(80; 82; 83) is used to independently calculate the 95% C.L. limits and to verify that the result does not depend on the details of the statistical formulation. They give similar results within 10%, with the modified frequentist approach giving more optimistic limits on average.

Though many sources of the systematic uncertainties differ between the experiments and the analyses, the appropriate correlations are taken into account in the combined

Tevatron	100	105	110	115	120	125	130	135	140	145	150
Expected/SM	1.52	1.58	1.73	1.78	2.1	2.2	2.2	2.1	1.91	1.75	1.49
Observed/SM	2.11	2.35	2.28	2.70	2.7	3.1	3.1	3.4	3.03	2.17	1.80
	155	160	165	170	175	180	185	190	195	200	
Expected/SM	1.30	0.96	0.89	1.07	1.32	1.63	2.1	2.6	3.2	3.7	
Observed/SM	1.96	1.09	0.94	1.29	1.24	1.44	2.3	2.5	4.3	5.0	

Table 7.2: The expected and observed limits on the Higgs production cross section obtained by combining CDF and DØ results for 21 Higgs mass hypotheses. The limits are in units of expected SM cross section.

limit calculation. The cross section uncertainties are taken as correlated among the two experiments. The luminosity uncertainty is taken as partially correlated between experiments.

The combined 95% C.L. upper limits on the Higgs production cross section are reported in Table 7.2 and Figure 7.4. The Tevatron combination sets the first direct constraint on the Standard Model Higgs mass above LEP limits, excluding at 95% C.L. the region 162 GeV $< m_H <$ 166 GeV. The combination has the sensitivity below the Standard Model Higgs predicted cross section in the range $159 < m_H < 169$ GeV. The exclusion ranges are determined by a linear interpolation of the expected and observed limits at each mass hypothesis. This extrapolation is expected to be valid since the resolution on the Higgs mass is much bigger than the 5 GeV step between the mass hypotheses. This is confirmed by the smooth behavior of the expected limits, mainly driven by the change in the Higgs production cross section and decay branching fraction.

The $H \to WW$ search presented in this Thesis gives a dominant contribution to the presented exclusion, being the most sensitive analysis entering the combination in the region where it has the sensitivity to probe the Standard Model predictions.

The impact of the results to the current knowledge on the Higgs boson is shown in the combination, performed by the GFITTER group(84), of the direct and indirect searches of the SM Higgs boson. A global fit to the Higgs mass is performed using several SM observables as input, in analogy of what already described in Section 1.3.2. The constraints coming from the direct searches are also included in the fit. The χ^2_{min} obtained for the fit is 17.76 over 14 degrees of freedom, corresponding to a p-value of 19.0 ± 0.004. It shows that, given the results of the current Higgs searches, the SM is still compatible with the presence of the Higgs boson. The $\Delta\chi^2 = \chi^2 - \chi^2_{min}$ of the fit as function of the Higgs boson in shown in Figure 7.5. The inclusion of the Teavatron results restricts the upper bound on the Higgs mass at 95% C.L. from 186 GeV to 156 GeV.

7.4 Future prospects

We have seen in Section 1.3.2 that the Standard Model Higgs boson, if it exists, is likely to have a mass in the range $114 < m_H <$ 186 GeV. The presented $H \to WW$ analysis already disfavor an Higgs boson with a mass $m_H \sim$ 160 GeV, and the Tevatron combination

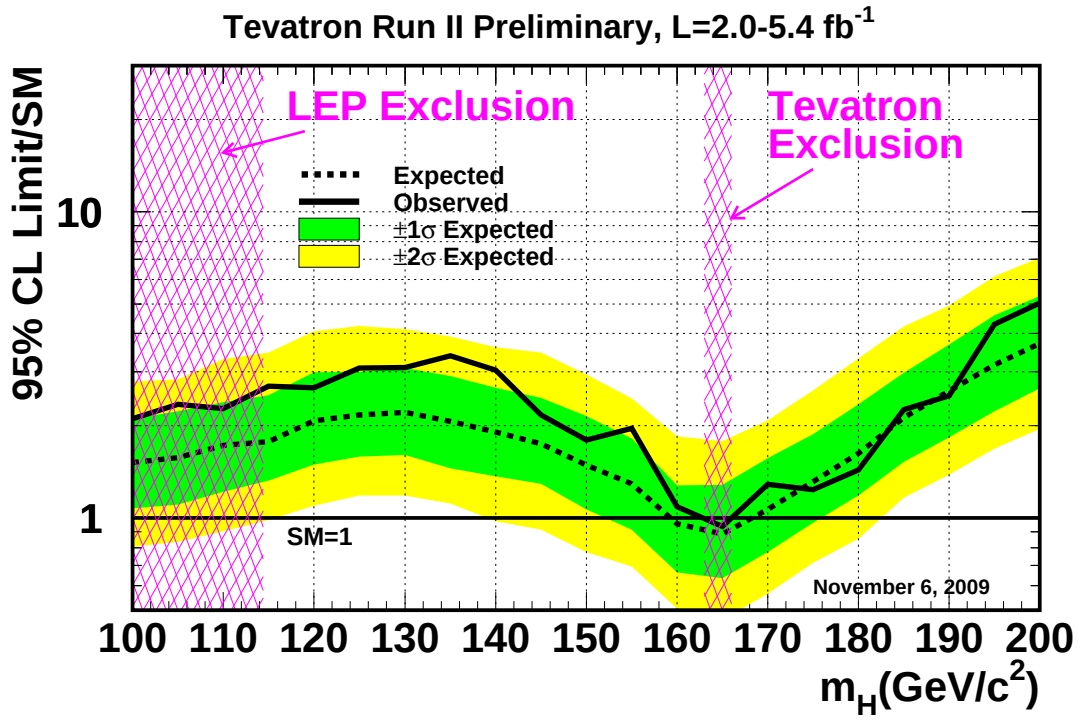


Figure 7.4: The expected and the observed limits on the Higgs production cross section obtained by combining CDF and DØ results as function of the Higgs mass. The limits are in units of expected SM cross section.

excludes it at 95% C.L. It is therefore mandatory to evaluate the future prospects for the Higgs searches in the whole mass range at the Tevatron before the LHC starts to be competitive.

The Tevatron collider has delivered about 7 fb^{-1} of integrated luminosity to both CDF and DØ experiments as of the end of year 2009. As described in Chapter 3, the data taking efficiency of the CDF detector is now stable, storing data with $\sim$ 85% efficiency, which corresponds to $\sim$ 6 fb^{-1} of integrated luminosity recorded on tape. Requirements on the proper operation of the CDF subdetectors reduce the available data set for the physics analyses by $\sim$ 10%. It means that $\sim$ 5.4 fb^{-1} of good data are available for physics analysis up to date. The Tevatron plans are to continue $p\bar{p}$ collisions at least through the year 2010. Figure 7.6 shows the Tevatron delivered luminosity as function of the operational year. While Tevatron performances became stable, we can conservatively assume that 2 fb^{-1} will be delivered in 2010. We can assume a data taking in 2011, which is still under discussion, with additional 2 fb^{-1} of integrated luminosity. Taking into account data taking efficiency it is expected $\sim$ 7 fb^{-1} of data by the end of 2010 and

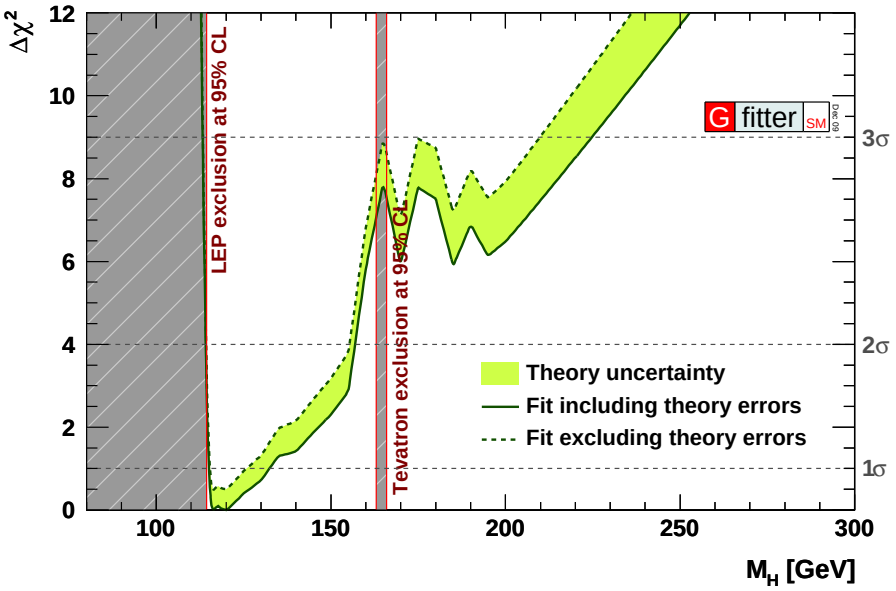


Figure 7.5: Fit to the Higgs mass performed by the GFITTER collaboration using direct and indirect searches of the Higgs boson. The difference of the χ^2 of the fit to the minimum χ^2_{min} is shown as function of the Higgs mass. The fit result is $m_H = 119.1^{+13.4}_{-4.0}$ GeV(84).

~ 9 fb^{-1} by the end of 2011. A more optimistic scenario estimates ~ 10 fb^{-1} available by the end of 2011. The results achieved in this Thesis are based on 4.8 fb^{-1} of data. In order to understand the reach of the current analyses, we estimate the scaling of the 95% C.L. sensitivity with the luminosity increase assuming gaussian statistical errors only. The sensitivity of the measurements is thus expected to to scale as:

$$\sigma^{95} = \sigma^{95}_0 \cdot \sqrt{\frac{L_0}{L}}, \tag{7.1}$$

with σ^{95}_0 being the current 95% C.L. sensitivity obtained with an integrated luminosity L_0 and σ^{95} is the expected sensitivity corresponding to an integrated luminosity L. The assumption of statistical errors only is a good approximation since they dominate, but results of the extrapolation tend to be a bit optimistic.

Improvements in the sensitivity of the searches can be achieved also with the optimization of the analyses techniques and with the search for final states not exploited by current analyses.

It is convenient to discuss the potential improvements, either due to the increase of the data sample analyzed or to the new analysis methods, in four different mass sub-ranges within the favored $114 < m_H < 186$ GeV.

$\mathbf{m_H \in [114, 120]}$ GeV

If the Higgs boson has a mass right above LEP limits, the high mass Higgs searches presented cannot significantly improve the current Tevatron sensitivity. The Tevatron

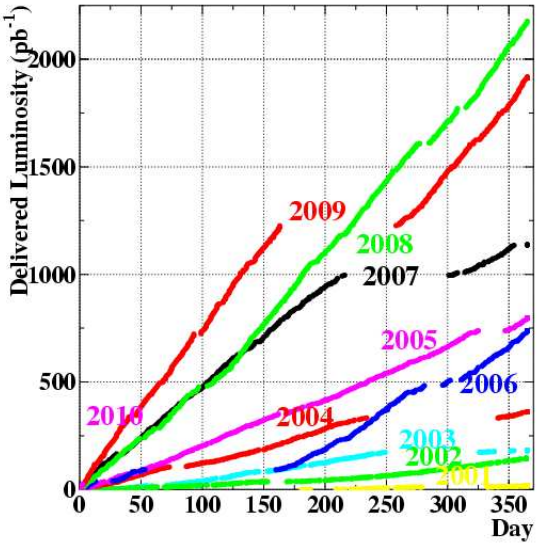


Figure 7.6: Tevatron delivered luminosity as function of the operational years.

combination of the analyses exploiting $H \to b\bar{b}$ decay channel will have to improve the sensitivity about a factor of two on average to probe the Standard Model Higgs production cross section.

The average luminosity currently used in the low mass Higgs CDF analyses, weighted by their expected limits, is approximately 3.6 fb^{-1}. Assuming the scaling of Equation 7.1 it means that the expected limit of the CDF combination for $m_H = 120$ GeV can improve down to $2.0\sigma_H^{SM}$ and $1.6\sigma_H^{SM}$ using respectively 7 and 10 fb^{-1} of data. The Tevatron combination is expected to scale similarly, with a 95% C.L. limit on the Higgs cross section of $1.3\sigma_H^{SM}$ using 10 fb^{-1} of data. Still at least another 30% average improvement is needed. This has to be obtained on the analyses techniques side. Two example are the improvement of $b-$tagging techniques to increase the selection efficiency on the Higgs signal, and the improvement of the di-jet invariant mass resolution, to increase the separation power of the $H \to b\bar{b}$ signal from the non-resonant backgrounds.

$\mathbf{m_H \in [120, 150]}$ GeV

The mass range 120 GeV $< m_H <$ 150 GeV receives contributions from the high and low mass Higgs searches. However both are loosing sensitivity in this region so that new improvements are needed in this region to achieve Standard Model sensitivity. For the purpose of this Thesis, we focus on how the high mass Higgs analyses can contribute.

One example is the $H \to WW$(low m_{ll}) analysis (shown with a blue line in Figure 7.2), which helps to increase the sensitivity especially in this mass region.

An additional improvement to $H \to WW$ analysis which helps in this mass range is connected with the $\vec{\not\!\!E}_T^{\text{Spec}}$ requirement on the di-lepton sample. Figure 7.8 shows that on average the $\vec{\not\!\!E}_T$ and $\vec{\not\!\!E}_T^{\text{Spec}}$ distributions have lower mean the lighter it is the Higgs boson.

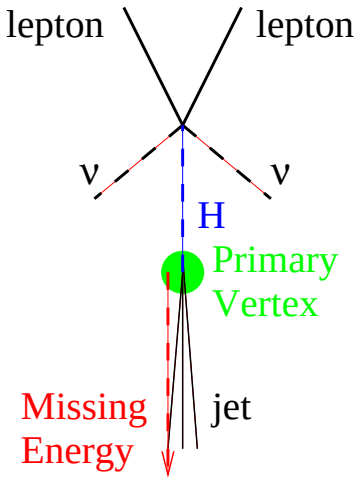


Figure 7.7: Schematic representation of the Higgs boson recoiling against one jet in the laboratory reference frame and decaying to two leptons and two neutrinos, via a WW pair. The $\vec{\not\!\!E}_T$ coming from neutrinos will tend to be aligned with the jet.

In events with at least one reconstructed jet this requirement is particularly inefficient. In fact in this case signal events tend to have the $\vec{\not\!\!E}_T$ aligned with the recoiling jet, as depicted in Figure 7.7. The resulting $\vec{\not\!\!E}_T^{\rm Spec}$, which is the transverse component of the $\vec{\not\!\!E}_T$ to the jet axis (see Equation 4.1), tends to be low, failing the event selections. Although relaxing the $\vec{\not\!\!E}_T^{\rm Spec}$ cut can provide gain in signal efficiency, the accepted Drell-Yan background grows fast too. However the trained neural networks can well separate Drell-Yan process from signal, as shown in this Thesis for the opposite sign analysis in events with one reconstructed jet. Studies are already ongoing to exploit the potential gain coming from loosening the $\vec{\not\!\!E}_T^{\rm Spec}$ requirement. The expected gain is on the order of 10% on the overall sensitivity.

Figure 7.9 shows the projections for the expected limits for CDF and Tevatron combinations with $m_H = 130$ GeV as function of the luminosity analyzed. Only the addition of data is not sufficient to provide enough sensitivity to probe Standard Model prediction, even for the Tevatron combination. About $\sim 10\%$ gain in the sensitivity of the analyses can be obtain in the near future improving the current analysis as just described, but it will not be sufficient. In this range the analyses need to improve their sensitivity by $\sim 50\%$ (the bottom of the orange band in the Figure) to probe the expected Standard Model cross section.

$\mathbf{m_H \in [150, 180]}$ GeV

The mass range $150 < m_H < 180$ GeV is the most sensitive region for $H \to WW$ searches. Already Tevatron combination has Standard Model sensitivity in the range $\sim 160-170$ GeV, while an improvement of less than 50% on the sensitivity will be needed to probe the Standard Model in the whole range. Assuming the scaling of Equation 7.1, Figure 7.10 shows the expected excluded mass range as function of the luminosity of the data sample. With ~ 10 fb^{-1} of data Standard Model Higgs can be excluded for

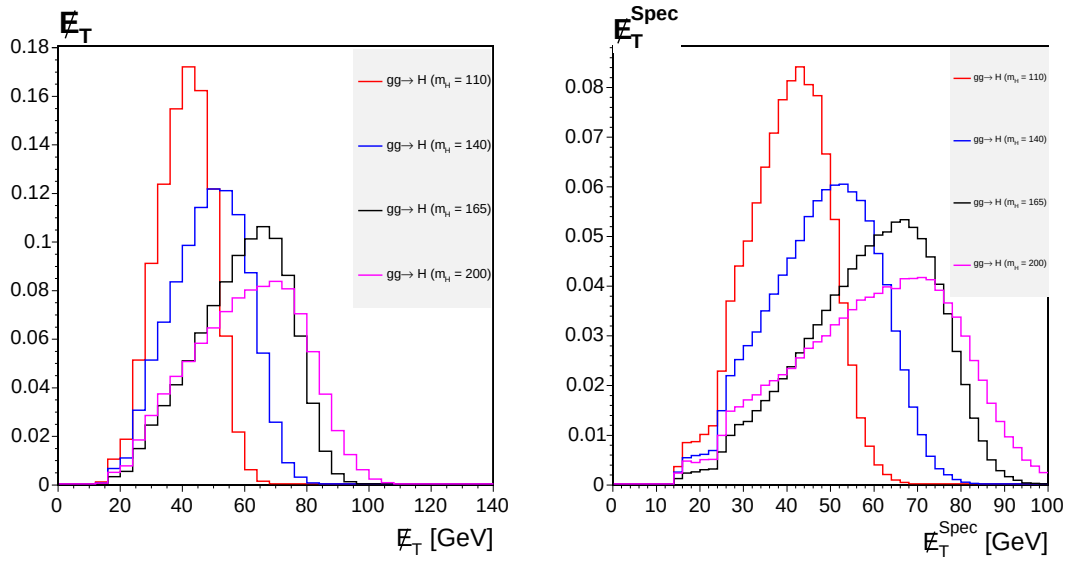


Figure 7.8: Unit-normalized $\vec{\not\!\!E}_T$ and $\vec{\not\!\!E}_T^{\text{Spec}}$ distribution for signal events with different Higgs mass hypotheses: $m_H = 110$ GeV, $m_H = 140$ GeV, $m_H = 165$ GeV, $m_H = 200$ GeV.

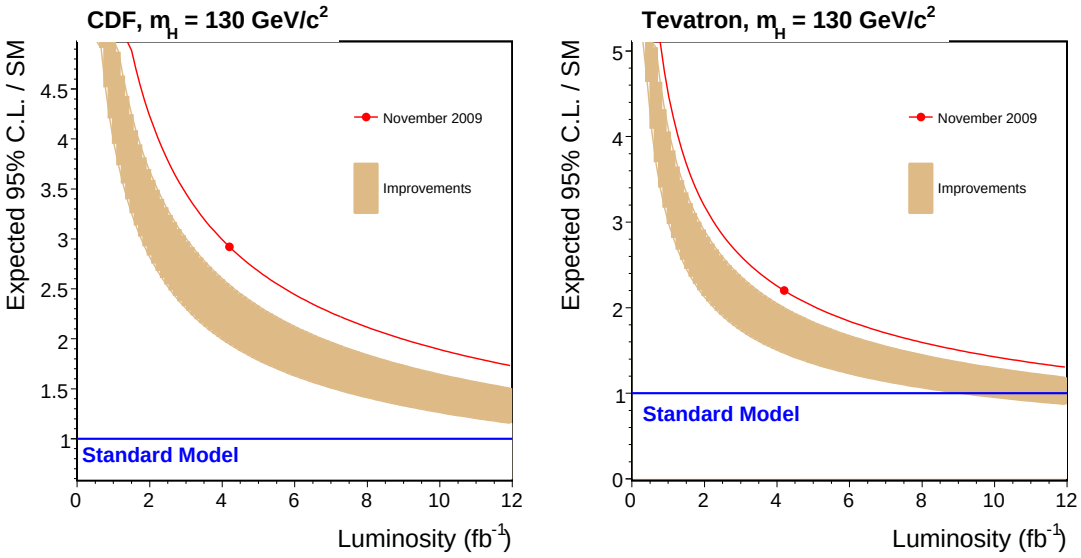


Figure 7.9: Plots show with a point the median 95% C.L. expected limit for the current CDF (left) or Tevatron (Right) combination as function of the average luminosity used in the combination. The curve represents the luminosity scaling of Equation 7.1. The orange band represents the effect of 10% to 50% improvements in the current sensitivity.

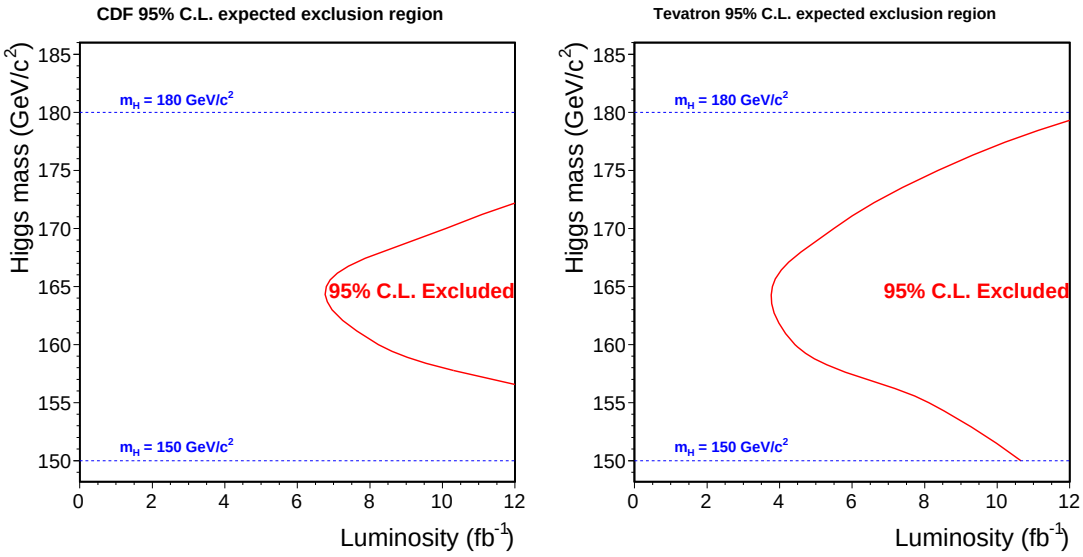


Figure 7.10: Expected Higgs mass exclusion region at 95% C.L. with the luminosity scaling of Equation 7.1. Left plot uses CDF $H \to WW$ analysis results. The right plot shows the projection for the Tevatron combination.

$m_H \in [152 - 178]$ GeV. Moreover improvements described in the other mass ranges will also contribute to this region, ensuring enough sensitivity to probe the Standard Model predictions.

$\mathbf{m_H \in [180, 200]}$ GeV

Above $m_H = 180$ GeV the sensitivity quickly decreases because of the low Higgs production cross section. It is important in this range to increase the acceptance and the efficiency to signal events in order to significantly improve the current limits.

The $H \to ZZ$ decay channel has a lower decay branching fraction than $H \to WW$; however the ratio of the resonant and the non-resonant cross sections $(\sigma(H \to ZZ)/\sigma(ZZ))$ becomes larger than $\sigma(H \to WW)/\sigma(WW)$ if $m_H \gtrsim 190$ GeV, as shown in Figure 7.11. This ratio is a significant indication of the sensitivity for the Higgs searches in these decay channels, since the non-resonant WW or ZZ production is expected to be the major background contributor. In particular, the decay chain $H \to ZZ \to l^+l^-\nu\bar{\nu}$ provides the best way to exploit $H \to ZZ$ at the Tevatron. However the $\vec{\not\!\!E}_T$ from neutrinos will be lower than in the $H \to WW$. For this reason studies to include it are deeply connected with the ones that aim to lower the $\vec{\not\!\!E}_T$ requirement mentioned before. Monte Carlo simulations show that we expect a comparable amount of $H \to ZZ$ events to those recovered in the $H \to WW$ lowering the $\vec{\not\!\!E}_T^{\text{Spec}}$ requirement. The ratio $H \to ZZ$ to the non-resonant ZZ background is also comparable with the $H \to WW$ case, as shown in Figure 7.11, but we expect the Drell-Yan background to be more difficult to separate, since we now have a real Z decaying to leptons. We expect a $\sim 10\%$ improvement in the Higgs sensitivity for $m_H = 190$ GeV from exploiting $H \to ZZ \to ll\nu\nu$ decay channel at the Tevatron.

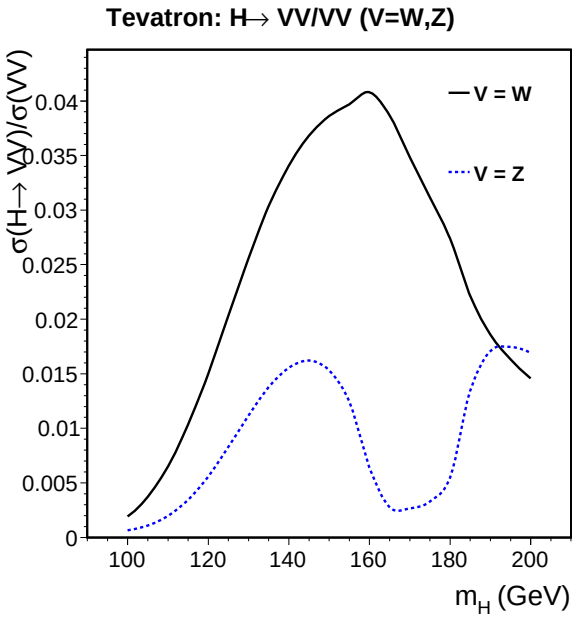


Figure 7.11: The ratio of resonant $H \to WW, ZZ$ and non-resonant WW, ZZ production cross sections as function of the Higgs mass. Same as Figure 2.5, reported here for convenience.

Figure 7.12 shows the projections for the expected limits for the CDF and Tevatron combination with $m_H = 190$ GeV. Only the addition of data is not sufficient to provide enough sensitivity to probe Standard Model prediction, even in the case of Tevatron results. The orange band shows the effect of improvements from 10% to 50% on the sensitivity. In this mass range only improvements of $\sim 70\%$ in the sensitivity and the analysis of 10 fb^{-1} of data can provide enough sensitivity to probe Standard Model predictions.

Observation probability

The presented extrapolations focus on the sensitivity for 95% C.L. upper limits. However it is important to evaluate also what is the probability of observing the Higgs boson as function of its mass. Figure 7.13 shows the probability of having at least a 3σ excess above the expected background as function of the Higgs mass. Results shown in the Figure assume a combination of the CDF and DØ results with equal-performing analyses (simply scaling the single result by $\sqrt{2}$). The analysis of 10 fb^{-1} per experiment and the improvement up to 50% in the analysis techinques give a probability of observing the Higgs signal in the range $114 - 186$ GeV greater than $\sim 30\%$. In the regions right above LEP limits $114 - 116$ GeV and $156 - 174$ GeV the probability is above 50%.

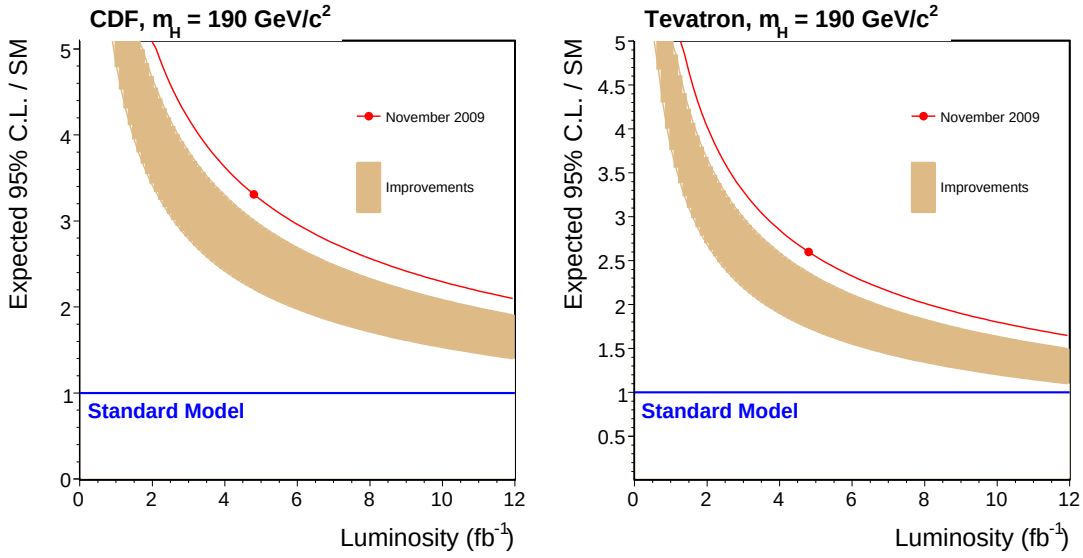


Figure 7.12: Plots show with a point the median 95% C.L. expected limit for the current CDF (left) or Tevatron (Right) combination as function of the average luminosity used in the analyses in the combination. The curve represents the luminosity scaling of Equation 7.1. The orange band represent the effect of 10% to 50% improvement in the current sensitivity.

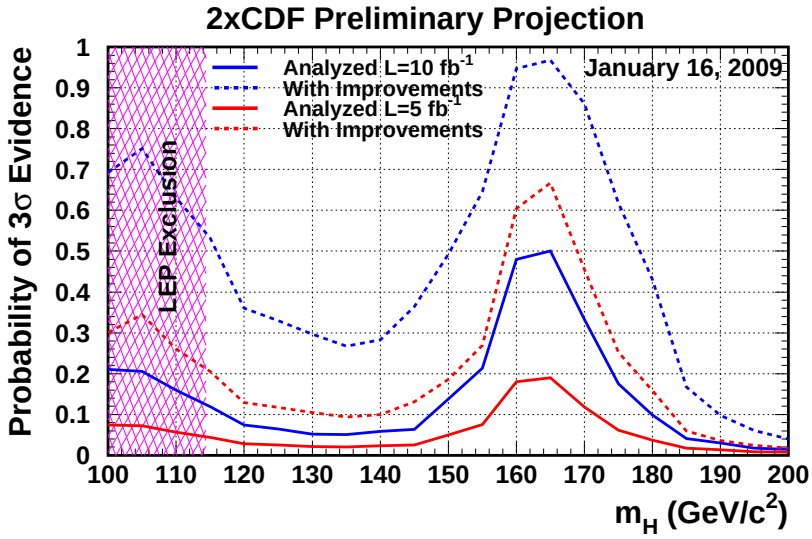


Figure 7.13: Probability of observing the Higgs boson signal over the expected background with a sensitivity of at least 3σ, as function of the Higgs mass. CDF results are scaled by $\sqrt{2}$ in this Figure, assuming CDF and D0 perform the same. Two scenarios are shown: the one in which channels have the same performance as for the current combination (solid lines), and for the case with another factor of 1.5 improvement in the sensitivity (dashed lines). The current observations are not taken into account, only expected sensitivities are shown. Sensitivities are expected to scale with sqrt(L) as in Equation 7.1.

Chapter 8

Conclusions

This Thesis reported on the study of the electro-weak symmetry breaking of the Standard Model (SM) by searching for a high mass Higgs boson produced by gluon fusion, associated production and vector boson fusion in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV.

Previous direct and indirect measurements constrain the Higgs mass in the range $114 - 186$ GeV/c^2. We showed that the $H \to WW$ decay channel with two leptons and two neutrinos in the final state is the most promising signature of its production if the Higgs mass $m_H \gtrsim 125$ GeV/c^2.

We analyzed 4.8 fb^{-1} of data collected by the CDF detector selecting events with two reconstructed leptons (e or μ) and significant missing transverse energy. We then divided the sample by the number of reconstructed jets in the event and by the relative charge of leptons for searching the Higgs boson in the different production mechanisms. We looked for direct production of the Higgs through gluon fusion in events with two opposite charged leptons and zero reconstructed jets. Associated and vector boson fusion production of the Higgs are the dominant signal contribution in events with higher jet multiplicities and in events with same charged leptons. The expected background composition is different in each sample and the analysis techniques are optimized separately for each production mechanism. The Matrix Element and Neural Network techniques are used and optimized also as a function of the Higgs mass (m_H) to discriminate signal from backgrounds. The expected and observed Neural Network output distributions are then used to search for the Higgs boson.

Observed data are consistent with the sample being the expected admixture of SM backgrounds without any signal contribution.

We set 95% C.L. upper limits on the production cross section of the Higgs boson as a function of the Higgs mass using a Bayesian approach. For $m_H = 165$ GeV/c^2 the combination of the searches performed in this Thesis sets an upper limit of 1.3 times the expected SM production cross section σ_H. The expected limit in a background-only hypothesis is $1.2^{+0.6}_{-0.4}\sigma_H(85)(37)$.

The combination of the CDF and the DØ Higgs results has an expected limit below the predicted SM cross section for $159 < m_H < 169$ GeV/c^2, and observed data excludes at 95% C.L. the Higgs boson in the mass range from 162 GeV/c^2 to 166 GeV/c^2(81). The analysis performed in this Thesis is the most sensitive one entering the combination for $m_H > 130$ GeV/c^2. The results provide the first direct constraint on the mass of the SM

Higgs boson beyond that obtained at LEP and, when combined with indirect searches, significantly restrict the upper bound on the Higgs mass to 156 GeV/c^2(84).

In the next one or two years, before LHC starts to be competitive, Tevatron Higgs searches are expected to further improve the presented results either by increasing the integrated luminosity analyzed and by improving analysis techniques. We estimate that within two years Tevatron results can reach sensitivity to the predicted SM Higgs cross section at 95% C.L. for most of the mass range allowed by the current direct and indirect constraints, while a 3σ observation has a probability to occur greater than 30% in the same range, if the SM Higgs boson exists.

Appendix A

Muon identification

In this Appendix we describe in detail the muon identification selections summarized in Tables 3.9-3.13.

A.0.1 CMUP

CMUP muons are required to be fiducial to both CMU and CMP detector volumes. For the CMP chamber We require the extrapolated track to be within the chamber volume in local x and at least 3 cm away from the edge of the chamber volume in local z, to select regions which are maximally efficient for hit-finding. The distance in the drift direction between the extrapolated track and the stub in the muon chamber (ΔX) to be respectively $\Delta X_{CMU} < 7$ cm and $\Delta X_{CMP} < \max(6.0, 150/p_T[GeV])$ for CMU and CMP chambers. Table 3.9 details all the selections.

A.0.2 CMP

CMP muons are required to have a candidate track fiducial to the CMP but not fiducial to the CMU detector, with $\Delta X_{CMP} < \max(6.0, 150/p_T[GeV])$. In fact gaps exist between each calorimeter wedge, and since CMU chambers are mounted in the edge of these wedges these gaps also affect CMU coverage. There is a 2.25° gap every 15° in ϕ in the CMU coverage. This can be clearly seen from the $\eta - \phi$ distributions of reconstructed CMP muons in Figure 3.26. We consider CMP muons mainly because we can trigger on single CMP-like muon events. The current implementation of this trigger as discussed in Appendix B.1.4 forces to select muons in these ϕ gaps, excluding a central gap at $\eta \sim 0$ that is for all ϕ range, but that represent less than $\sim 30\%$ of the reconstructed CMP muons[1]. Since we're interested in CMP muons which are as much similar as possible to the ones that triggered the event, we tightened the requirement with an explicit cut on the reconstructed ϕ of the track to lie inside these gaps, within 2°. We moreover define these leptons only for data that are collected from run number 229764,from which the corresponding trigger was properly working. Table 3.10 lists the detailed selections applied.

[1]muons reconstructed this gap are not included at trigger level because of the large fake rate which results in a too high trigger cross section.

A.0.3 CMU

CMU muons are required to have a candidate track fiducial to the CMU but not fiducial to the CMP detector, with $\Delta X_{CMU} < 7$ cm. We also explicitly require them to be not fiducial to CMX chambers to avoid a small overlap of the detectors. Table 3.10 details all the selections. When evaluating trigger efficiencies we discovered a technical problem in the implementation of the specific trigger for this type of muons (see Appendix B.1.5). For this reason we do not use CMU muons in the present thesis. The bug was recently fixed and they will soon be included in the future. There's no further loss in acceptance for this choice, since CMU muons are recovered as stubless central muons (CMIOCES).

A.0.4 CMX

Analogously to the CMP chambers, for **CMX** chamber we require the extrapolated track to be within the chamber volume in local x and at least 3 cm away from the edge of the chamber volume in local z, to select regions which are maximally efficient for hit-finding. We further require the ϕ of the candidate muon track to be $0° < \phi < 75°$ or $105° < \phi < 225°$ or $315° < \phi < 360°$ to select the Arches CMX detectors. Track-stub matching is required to have $\Delta X_{CMX} < \max(6, 125/p_T[GeV])$. To ensure the best tracking quality for these muons, we require that each muon track passes through all eight COT super-layers by making a minimum requirement on the exit radius of the track at the end-plates of the COT tracking chamber. The exit radius is defined as:

$$\rho_{COT} = (z_{COT} - z_0) \cdot \tan\theta, \tag{A.1}$$

where z_{COT} is the distance of the COT end-plates from the center of the detector (+155 cm for track with $\eta > 0$ and −155 cm for those with $\eta < 0$). Finally we consider CMX muons only for runs above 150144 to have fully working chambers. Table 3.11 details all the selection criteria.

A.0.5 CMXMsKs

We separately reconstruct muons in the Miniskirts and Keystone detectors (**CMXMsKs**) with similar selections respect to CMX muons. We only require the ϕ of the track to point to the corresponding sub-detectors and a run number greater than 227704 to ensure that the corresponding trigger path was working. Table 3.11 details all the selection criteria.

A.0.6 BMU

BMU muons are not only required to be fiducial to any of the IMU detectors but also to the to plug calorimeter, and in particular to the PES detector. This requirement is used to make these muons a subset of the forward stubless muons, which they share the same tracking requirement with. This choice is made for convenience, since it simplifies the understanding of these muons without a big loss of acceptance (which is recovered by the CrkTrk category). We also enforce fiduciality requirement to IMU detectors requiring that the extrapolated z_{BMU} of the track, corrected by the primary vertex position $z^{P.V.}$,

to the IMU detector radius ($\bar{r}_{BMU} = 391$ cm) is within the chambers. We evaluate z_{BMU} as:

$$z_{BMU} = Z^{P.V.} + \bar{r}_{BMU} \frac{1 - \mathrm{e}^{-2\eta}}{2\mathrm{e}^{-\eta}}. \tag{A.2}$$

Table 3.12 details all the selection criteria.

A.0.7 CMIOCES and CMIOPES

Stubless muons are required to have a track $|z_0| < 60$ cm and an impact parameter $|d_0| < 0.2$ cm ($|d_0| < 0.02$ cm if the track has hits also in the Silicon detectors). We then require to have a minimum amount of energy deposition in the calorimeter which is consistent with a minimum ionizing particle, calorimetric isolation ($CalIso$) and isolation in the tracking detector ($TrkIso$).

CMIOCES tracks are required to have at least three axial and stereo super-layers of the COT drift chambers with five or more hits. The $\chi^2/n.d.f.$ of the track fit is also required to be less than three.

For the limited η coverage of the COT, **CMIOPES** tracks require a number of hits in the chamber that is at least 60% of the expected one, based on the track direction; we also require at least three hits in the Silicon detectors and a curvature significance $C/\sigma(C) > 12$.

In order to preserve the uniqueness of each category each candidate muon object is give a chance to become a stubless muon only if it fails all the stubbed muon selections.

In particular a fully identified CMIOPES is also required to do not pass the BMU selections. A CMIOCES muon is required not to pass all the central stubbed muon identification criteria: CMUP, CMU, CMP, CMX, CMXMsKs.

Appendix B

Triggers

In this Appendix we describe in detail the trigger paths used in the analysis and outlined in Section 3.6. The trigger efficiency measurements are described in more detail for each of the trigger paths.

B.1 Trigger Paths description

B.1.1 ELECTRON_CENTRAL_18

The high$-E_T$ central electron trigger requires a track reconstructed in the COT drift chamber matched to a cluster in the central calorimeter with an appropriate shower profile and energy deposition requirements. The detail requirements at each trigger level are detailed in the following.

L1: CEM8_PT8 requires a trigger tower in the central calorimetric region with an electromagnetic transverse energy deposit of $E_T^{EM} > 8$ GeV and a ratio of hadronic to electromagnetic energy deposit of $E_{HAD}/E_{EM} > 0.125$; it also requires a track reconstructed by the eXtremely Fast Tracker (XFT) for the COT with 10 hits in 3 layers or 11 hits in 4 layers and with a transverse momentum $p_T > 8.34$ GeV.

L2: CEM16_PT8 requires a calorimetric cluster with $|\eta| < 1.317$, an electromagnetic transverse energy deposit in the seed tower $E_T^{EM,\,\mathrm{seed}} > 8$ GeV, in shoulder towers $E_T^{EM,\,\mathrm{shoulders}} > 7.5$ GeV and a total $E_T^{EM} > 16$ GeV, $E_{HAD}/E_{EM} < 0.125$ and an XFT track with 10 hits in 3 layers or 11 hits in 4 layers and with $p_T > 8$ GeV. Recently the E_T^{EM} requirement has been set to 18 GeV, taking advantage of a trigger upgrade(55) which provides a resolution in the energy measurement of the jets that is near to what is obtained in offline analysis.

L3: ELECTRON_CENTRAL_18 requires an offline central electromagnetic cluster with $E_T > 18$ GeV and $E_{HAD}/E_{EM} < 0.125$ and a COT track with $p_T > 9$ GeV. Since may 2005 it additionally required lateral shower profile $L_{shr} < 0.4$, matching between the CES cluster and the COT track $|\Delta z| < 8cm$, using the vertex z position of the interaction point for the transverse component and three instead of two towers for E_{HAD}/E_{EM} calculations.

B.1.2 MET_PEM

The plug electron plus missing transverse energy trigger requires a cluster in the plug calorimeter ($1.1 \leq |\eta| \leq 3.6$) compatible with an electron shower. Given the limited coverage of the COT chamber in this region, backgrounds are reduce requiring missing transverse energy. The detail requirements at each trigger level are:

L1: EM8_MET15 requires a calorimetric trigger tower either in the central or plug calorimeter with $E_T^{EM} > 8$ GeV and $E_{HAD}/E_{EM} < 0.125$ for central calorimeter or $E_{HAD}/E_{EM} < 0.0625$ for the plug calorimeter; it also requires raw missing transverse energy $\not\!\!\vec{E}_T > 15$ GeV.

L2: PEM20_MET15 requires a calorimetric cluster in the $1.1 < |\eta| < 3.6$ region with $E_T^{EM,\,\text{seed}} > 8$ GeV, $E_T^{EM,\,\text{shoulder}} > 7.5$ GeV, $E_T^{EM} > 20$ GeV and $E_{HAD}/E_{EM} < 0.125$.

L3: PEM20_MET15 requires an offline plug electromagnetic cluster with $E_T > 20$ GeV, $E_{HAD}/E_{EM} < 0.125$ and $\not\!\!\vec{E}_T > 15$ GeV.

B.1.3 MUON_CMUP18

The high$-p_T$ CMUP central muon trigger is aimed to collect central muons which have a well reconstructed track in the COT chamber matched to stubs in both the CMU and CMP detectors. This trigger changed significantly after the tracking XFT trigger upgrade(56). After this upgrade a full 3D matching (i.e. using also the z coordinate) between the track and the stub has been possible, drastically reducing the rate of fake muons being triggered. Main differences come in at Level 2. We label the old trigger path as *2D-trigger* and the new one as *3D-trigger*. Separate Level 2 requirements are listed below for these two versions; it's important to note that after 3D triggers came in the 2D version still existed, even if dynamically prescaled. The detail requirements at each trigger level are listed in the following.

L1: CMUP6_PT4 requires a stub in the CMU with $p_T > 6$ GeV as measured by the CMU and an XFT track pointing to the CMU with $p_T > 4.09$ GeV. It also requires hits in the CMP chambers.

2D-L2: TRK8_L1_CMUP6_PT4 additionally requires a 4-layers XFT track with a minimum p_T requirement that increased with time from $p_T > 8.34$ GeV up to $p_T > 14.77$ GeV. This trigger level is also dynamically prescaled.

3D-L2: L2_CMUP6_PT15_3DMATCH additionally to confirmation of the L1 requirements, it requires an XFT track with $p_T > 14.77$ GeV 3D-matched to the CMU stub.

L3: MUON_CMUP_18 requires a CMUP muon with $p_T > 18$ GeV and distances between the stub and the COT track extrapolated to the muon detector plane $\Delta x_{CMU} < 10$ cm and $\Delta x_{CMP} < 20$ cm.

B.1.4 MUON_CMP18_PHIGAP

This trigger path aims to collect muons which falls in the $\phi-$gaps of the CMU coverage. CMU detectors are installed on top of the calorimeter wedges and have small gaps in the ϕ coverage every $\sim 15°$. Also the region around $\eta \sim 0$ has a gap of CMU coverage, but no attempt is made to recover this region. In fact this region would add only an extra 30% recovered muons (respect to all other $\phi-$gaps) but the rate is too high and only a requirement on the ϕ of the track, which is available already at L1, is able to reduce the trigger cross section at manageable levels. The detailed requirements at each trigger level are listed in the following.

L1: L1_CMP3_PT15_3D_PHIGAP_DPS requires an XFT track with $p_T > 14.77$ GeV which points to the gaps of the CMU detectors. This trigger level is dynamically prescaled.

L2: L2_CMP3_PT15_3DMATCH_PHIGAP_CSP requires a 4-layers XFT track with $p_T > 14.77$ GeV pointing to CMU coverage cracks and with a 3D match to a CMP stub. Hits in the corresponding CSP are also required.

L3: L3_MUON_CMP18 requires a basic muon with $p_T > 18$ GeV and distance between the CMP stub and the COT track extrapolated to the muon detector plane $\Delta x_{CMX} < 20$ cm.

B.1.5 MUON_CMU18_ETAGAP

This trigger path is aimed to collect muons pointing to crack of the CMP coverage. In particular edges of the box-geometry of the CMP chambers are the main source of these gaps. The trigger requires a track to be 3D matched to a CMU stub in these cracks. The detailed requirements at each trigger level are listed in the following.

L1: L1_CMU6_PT4_&_TRK10_DPS requires a stub in the CMU with $p_T > 6$ GeV as measured by the CMU and an XFT track pointing to the CMU with $p_T > 4.09$ GeV. At least one track in the event is required to have $p_T > 10.11$ GeV. This trigger level is dynamically prescaled.

L2: L2_CMU6_PT15_3DMATCH_ETAGAP requires a 4-layer XFT track with $p_T > 14.77$ GeV with a 3D matched stub in the CMU detector.

L3: MUON_CMU18 requires a muon with $p_T > 18$ GeV and distance between the CMU stub and the COT track extrapolated to the muon detector plane $\Delta x_{CMX} < 10$ cm.

B.1.6 MUON_CMX18

The high$-p_T$ CMX muon trigger is aimed to collect central muons which have a well reconstructed track in the COT chamber matched to stubs in the CMX Arches detector. After run number 227704 it also allows stubs-track matching in the Miniskirts and Keystone detectors. As for the CMUP trigger, two versions of this trigger currently exist:

a 2D and a 3D trigger, which differ by the requirements made at Level-2. The detailed requirements at each trigger level are listed in the following.

L1: CMX6_PT8_CSX requires a stub in CMX with $p_T > 6$ GeV and an XFT track pointing to the CMX with $p_T > 8.34$ GeV. It also requires a hit in the CSX.

2D-L2: CMX6_PT15 additionally requires a 4-layers XFT track with a p_T requirement that increased with time from $p_T > 10.1$ GeV up to $p_T > 14.77$ GeV. A dynamic prescale is applied at this level.

3D-L2: L2_CMX6_PT15_3DMATCH additionally to confirmation of L1 requirements, it requires an XFT track with $p_T > 14.77$ GeV 3D-matched to the CMX stub.

L3: MUON_CMX_18 requires a CMX primitive muon with $p_T > 18$ GeV and distance between the stub and the COT track extrapolated to the muon detector plane $\Delta x_{CMX} < 10$ cm.

B.1.7 JET20,50,70,100

Jets-based triggers require a reconstructed jet with increasing energy threshold. Given the high rate of these events only JET100 is not prescaled. The detailed requirements at each trigger level are listed in the following, with reference to Table B.1 for the different threshold values.

L1: Level 1 requires at least one calorimetric tower with energy above a given threshold. Expect for JET100, a static prescale factor is applied.

L2: Level 2 requires at least a cluster within $|\eta| < 3.6$ and with a lower cut on its E_T. Some static prescale can also be applied to lower threshold triggers.

L3: At least one jet reconstructed with JetClu algorithm with a $\Delta R = 0.7$ cone above a given threshold.

Trigger Name	Level 1		Level 2		Level 3
	$E_T >$ [GeV]	Prescale	$E_T >$ [GeV]	Prescale	$E_T >$ [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 trigger requirements of JET20, JET50, JET70, JET100

B.2 Trigger efficiency measurement

B.2.1 Central Electrons

The ELECTRON_CENTRAL_18 trigger efficiency is evaluated for the cut-based TCE electrons, already defined in Section 3.5, with a further requirement of $E_T > 20$ GeV. We then assumed the same trigger efficiency for LBE electrons. This assumption has been checked on a data sample dominated by $WW \to e\mu$ events. The event selections for this sample are the same that will be detailed in Section 4.1, which select two opposite sign leptons, with significant missing transverse energy. We also require no reconstructed jets in the event to enhance WW contribution. Furthermore we select only $e\mu$ events where the electron is either a TCE or LBE, we require the muon to be a triggerable muon and the event to have been selected online by the corresponding muon trigger path. We then check how often the reconstructed TCE or LBE passes the trigger requirements (i.e. have been selected by the ELECTRON_CENTRAL_18 trigger path too). We obtain that 84 out of 88 $TCE - \mu$ events have also been selected by the electron trigger path, which corresponds to a trigger efficiency of $\sim 0.94 \pm 0.10$; For selected $LBE - \mu$ events 78 out of 83 events pass also the electron trigger, which corresponds to a trigger efficiency of $\sim 0.95 \pm 0.10$, perfectly compatible[1] with the one evaluated for TCE electrons and also with the standard calculation which is described in the following.

We measure separately the tracking and the calorimetric trigger efficiency of TCE electrons for the ELECTRON_18 trigger path.

The tracking efficiency is evaluated using data collected by a backup trigger path: W_NOTRACK. The trigger requirements for this trigger path are summarized in Table B.2. The peculiarity of this trigger path is that it has very similar calorimetric trigger requirements respect to ELECTRON_CENTRAL_18 (apart for a $\vec{\not\!\!E}_T$ requirement which will not bias our measurement) but no tracking requirements. We select $W \to e\nu$ candidate events collected by this trigger path requiring one high$-p_T$ electron (TCE) with $E_T \geq 25$ GeV, $\vec{\not\!\!E}_T \geq 25$ GeV and requiring the event not to be compatible with a cosmic ray. These selection provides an highly pure electron sample, with almost negligible background. The tracking trigger efficiency for the Level-1 of the ELECTRON_CENTRAL_18 is then given by:

$$\epsilon_{L1}^{\text{Tracking}} = \frac{\text{num}(W) \ \& \ L1^{EC18}}{\text{num}(W)}, \tag{B.1}$$

that is the number of events triggered by W_NOTRACK trigger which passes our W selections, num(W), and also the Level 1 of the ELECTRON_CENTRAL_18 trigger path, num(W)&$L1^{EC18}$, divided by num(W). $\epsilon_{L1}^{\text{Tracking}}$ is found to have a significant dependence of η which is taken into account with a linear parametrization plus a gaussian to account for a small drop ($\sim 6\%$) of efficiency in the $\eta \sim 0$ region. Analogously we can measure

[1] Errors are strongly correlated, still the two measurements are compatible after taking into account this correlation.

W_NOTRACK	
Trigger Level	Requirements
L1	≥ 1 tower $E_T \geq 8$ GeV $E_{HAD}/E_{EM} \leq 0.125$ $\not\!\!\vec{E}_T \geq 15$ GeV
L2	≥ 1 cluster $E_T \geq 20$ GeV $E_{HAD}/E_{EM} \leq 0.125$ $\|\eta\| \leq 1.1$ $\not\!\!\vec{E}_T > 15$ GeV
L3	≥ 1 Calorimetric object $E_T \geq 25$ GeV $E_{HAD}/E_{EM} \leq 0.125$ $\not\!\!\vec{E}_T \geq 25$ GeV

Table B.2: Trigger requirements of the W_NOTRACK trigger path.

the L2 and L3 trigger efficiencies as

$$\epsilon_{L2}^{\text{Tracking}} = \frac{\text{num}(W)\ \&\ L1^{EC18}\ \&\ L2^{EC18}}{\text{num}(W)\ \&\ L1^{EC18}} \tag{B.2}$$

$$\epsilon_{L3}^{\text{Tracking}} = \frac{\text{num}(W)\ \&\ L1^{EC18}\ \&\ L2^{EC18}\ \&\ L3^{EC18}}{\text{num}(W)\ \&\ L1^{EC18}\ \&\ L2^{EC18}} \tag{B.3}$$

$$\tag{B.4}$$

The total tracking efficiency is about $\epsilon^{\text{Tracking}} \sim 98\%$ with variations from one run period to another of the order of $\Delta\epsilon^{\text{Tracking}} \sim 2\%$.

The Level-1 calorimeter efficiency is evaluated in an unbiased sample triggered with muons, and is found to be 100%. Level-3 is also found to be 100% efficient. The Level-2 efficiency is determined using data collected by a backup trigger which has the same requirements of ELECTRON_CENTRAL_18 but without any requirement at Level-2. This trigger is heavily prescaled at L2 to maintain the rate under control, but can be used to evaluate the level-2 calorimeter trigger efficiency as the number of selected W events which pass this trigger and also the Level-2 of the ELECTRON_CENTRAL_18, divided by the total number of W events selected by the backup trigger. The efficiency is evaluated in function of the E_T of the electron, since it shows a turn-on curve up to energies of $E_T \sim 25$ GeV.

The total trigger efficiency is then dependent on η and E_T of the electrons. The average value is about 97%.

B.2.2 Forward Electrons

The technique to measure the MET_PEM efficiency divides the trigger requests into two main independent components. The first one is relative to the $\not\!\!\vec{E}_T$ selection while the second one is relative to the EM clustering in the plug region.

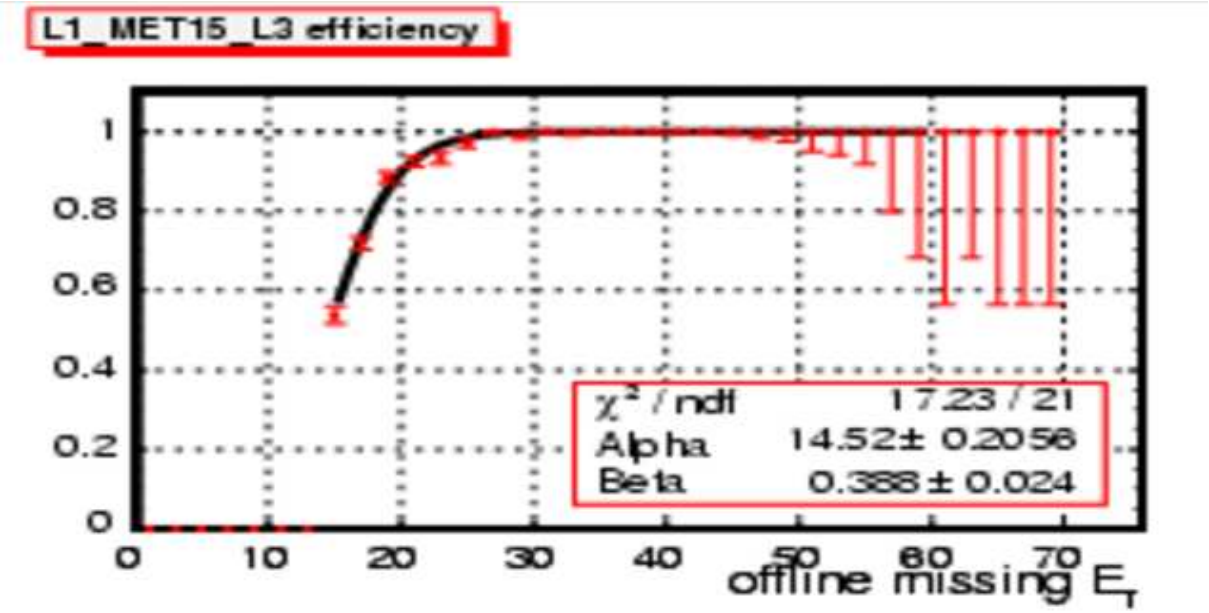


Figure B.1: Turn-on curve of the L1 $\vec{\not\!\!E}_T$ requirement of MET_PEM trigger in function of the offline $\vec{\not\!\!E}_T$ for data in run period 17.

The trigger efficiency of the $\vec{\not\!\!E}_T$ selections is evaluated identifying $W \to e\nu$ candidate events. We require a forward identified electron (PHX) and large missing transverse energy ($\vec{\not\!\!E}_T \geq 15$ GeV), which select an highly pure $W \to e\nu$ sample. We look at events collected by the PLUG_ELECTRON_20 trigger path, which has the same calorimetric requirements of MET_PEM but without the $\vec{\not\!\!E}_T$ requirements, and is prescaled by a static factor 25 to reduce its rate to manageable levels. We then evaluate the efficiency of the $\vec{\not\!\!E}_T$ requirements in function of the $\vec{\not\!\!E}_T$ as:

$$\epsilon_{\vec{\not\!\!E}_T}(\vec{\not\!\!E}_T) = \frac{num(W)\ \&\ PLUG\_ELECTRON\_20\ \&\ MET\_PETM}{num(W)\ \&\ PLUG\_ELECTRON\_20} \tag{B.5}$$

that is the number of W candidates, num(W), which fires the PLUG_ELECTRON_20 trigger path and also the MET_PEM trigger normalized to the number of Ws firing the PLUG_ELECTRON_20 trigger. The result shows a turn-on in function of the $\vec{\not\!\!E}_T$ which is shown in Figure B.1, as an example, for run period 17 and for the Level-1 $\vec{\not\!\!E}_T$ requirement, the dominant one, as function of the offline $\vec{\not\!\!E}_T$. The plateau is reached for offline $\vec{\not\!\!E}_T \gtrsim 25 - 30$ GeV, while the parametrization can still give a good description of the efficiency down to $\vec{\not\!\!E}_T$ $10 - 15$ GeV.

The efficiency of EM clustering trigger selections in the plug region is obtained using $Z \to ee$ candidate events collected by the high$-p_T$ central electron trigger. We further require one tight central electron (TCE) and one forward PHX electron with an invariant mass in the range 81 GeV $\leq m(ll) \leq$ 101 GeV. This selections provide a clean Z sample, which background contamination is estimated using sidebands of the invariant mass of the leptons and found to be negligible at the level of $< 1\%$. The efficiency is then evaluated counting how many Z events collected by the ELECTRON_CENTRAL_18 trigger also fire

the online calorimeter selections of the MET_PEM trigger[2]. The efficiency is parametrized as function of the lepton E_T and η.

B.2.3 Muons

Muons trigger efficiencies are measured selecting $Z \rightarrow \mu\mu$ events. We require an invariant mass of the di-lepton pair 76 GeV $\leq m(ll) \leq$ 106 GeV. CMX trigger efficiency is evaluated requiring one of the muons to be a triggerable CMUP, i.e. a CMUP with $p_T > 20$ GeV, and that the event is selected online by the corresponding trigger path: MUON_CMUP18; we then require the second muon to be a triggerable CMX and look how often the MUON_CMX18 trigger fired:

$$\epsilon_{CMX} = \frac{num(Z)\ \&\ MUON\_CMUP18\ \&\ MUON\_CMX18}{num(Z)\ \&\ MUON\_CMUP18}. \tag{B.6}$$

The same procedure can be applied to evaluate the trigger efficiencies of the MUON_CMP18_PHIGAP and MUON_CMU18_ETAGAP trigger paths using a triggerable CMP and CMU muon respectively. The trigger efficiency for CMXMsKs is assumed to be the same as for the CMX trigger, scaling down for a geometric factor to account for a part of the chambers which we cannot trigger on. The result is cross-checked and found in good agreement with a calculation similar to Eq. B.6 but with a triggerable CMXMsKs muon instead of a CMX one[3]. Analogously we evaluate the MUON_CMUP18 trigger efficiency requiring an event triggered by a CMX muon and looking how often it was also triggered by a triggerable CMUP muon. These procedure is used to evaluate efficiencies for 2D and 3D triggers separately. We cross-check the 3D calculation evaluating 3D trigger efficiency respect to the 2D version of the trigger, i.e. separately evaluating the loss of efficiency for the 3D requirement and multiplying it by the 2D efficiency. The two methods give compatible results and we use the former calculation for the analysis presented in this Thesis. Table B.3 report the measured trigger efficiencies in function of the run period. As already pointed out in Section 3.6, the 3D versions of the CMUP and CMX triggers, and the CMP, CMU trigger paths were introduced only after the XFT trigger upgrade. Numbers in brackets are used when the trigger was not working for the full period; for these cases the efficiency of the next run period is used. Trigger efficiencies after period 14 are evaluated without any further dependence on the run period, since data taking conditions are more stable from that point. CMU ETA-GAP trigger efficiency up to P21 is $< 0.1\%$, for a known implementation problem. Run period 21 has a trigger efficiency of 0.152 while a value of 0.989 and 0.969 is found for run periods 22 and 23 respectively. The efficiency reported in the Table is the total efficiency in Periods 14-23 which is a weighted average based on the integrated luminosity acquired in those periods. As a consequence we decided to do not use the CMU muon category, since the lepton acceptance coming from non-triggered CMU events is already recovered by stubless central muons.

[2]the online information about calorimetric requirements of the MET_PEM trigger is stored separately, even if only the complete MET_PEM requests are used to decide if the event has to be accepted or not.

[3]We note again that the trigger path is the same as for CMX muon. The only difference is the region in ϕ covered by the two muon categories.

Table B.4 also reports the livetimes of the 2D triggers which are dynamically prescaled. The livetime is defined as the inverse of the average prescale factor applied, and indicates which fraction of events which would have fired the trigger were actually stored. Given the continuous increase of average instantaneous luminosity delivered by Tevatron accelerator, the 2D triggers result more and more prescaled in the recent data, suffering a trigger rate that is too high for the current system. The CMX 3D trigger is also prescaled especially in run periods 9 to 13. From period 14 on a further optimization of the trigger selections (especially the 3D matching requirement) let a reduction of the prescale of this trigger path and also improved the efficiency for the CMUP 3D trigger path.

Run Period	CMUP 2-D	CMUP 3-D	CMX 2-D	CMX 3-D	CMP PHI-GAP	CMU ETA-GAP
P0	0.898		0.967			
P1-P4	0.919		0.949			
P5-P7	0.919		0.954			
P8	0.906		0.941			
P9	0.910		0.948	(0.927)		
P10	0.920		0.978	0.927	(0.948)	
P11	0.911	0.858	0.970	0.930	0.948	
P12	0.915	0.839	0.936	0.889	0.891	
P13	0.918	0.835	0.935	0.894	0.924	
P14-23	0.913	0.842	0.948	0.916	0.787	0.191

Table B.3: Trigger efficiencies for muon triggers as function of the run period. Empty cells correspond to run periods were the triggers were not implemented yet.

Livetimes	CMUP 2-D	CMUP 3-D	CMX 2-D	CMX 3-D	CMP PHI-GAP	CMU ETA-GAP
P0-P8	1.000		1.000			
P9	1.000		0.963	0.965		
P10	1.000		0.958	0.988	0.558	
P11	0.959	1.000	0.913	0.971	0.669	
P12	0.933	1.000	0.885	0.952	0.840	
P13	0.936	1.000	0.899	0.950	0.885	
P14-23	0.837	1.000	0.871	0.996	0.871	0.870

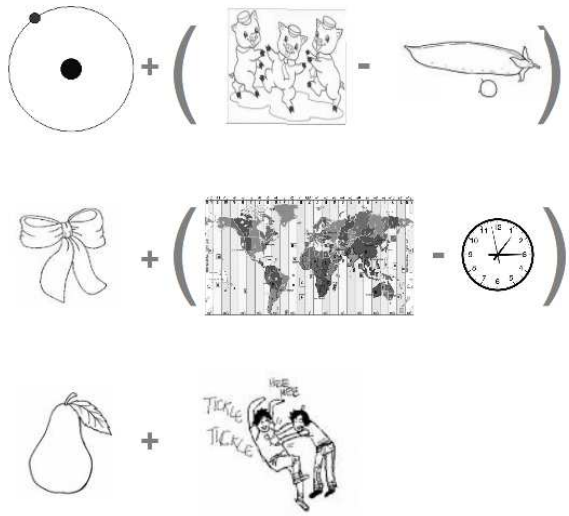
Table B.4: Trigger livetimes (inverse of average prescale factor) for muon triggers as function of the run period. Values of 1.000 correspond to run periods or triggers that were not prescaled. Empty cells correspond to run periods were the triggers were not implemented yet.

Muon trigger efficiencies show how the introduction of the 3D triggers increased trigger acceptance. In fact despite a small loss of efficiency they collect data without any prescale factor, and the net efficiency given by the product of the two numbers is greater for the 3D versions.

List of Figures

List of Tables

Acknowledgements



Bibliography

[1] CDF and DØ, "Updated Combination of CDF and D0 Results for the Mass of the W Boson," 2009, arXiv:0908.1374 [hep-ex].

[2] The ALEPH, DELPHI, L3, OPAL, SLD Collaborations, the LEP Electroweak Working Group, the SLD Electroweak and Heavy Flavour Groups, "Precision Electroweak Measurements on the Z Resonance," *Phys. Rept.*, vol. 427, p. 257, 2006, hep-ex/0509008.

[3] S. L. Glashow, "Partial Symmetries of Weak Interactions," *Nucl. Phys.*, vol. 22, pp. 579–588, 1961.

[4] S. Weinberg, "A Model of Leptons," *Phys. Rev. Lett.*, vol. 19, pp. 1264–1266, 1967.

[5] A. Salam, "Weak and Electromagnetic Interactions," 1968. Originally printed in *Svartholm: Elementary Particle Theory, Proceedings Of The Nobel Symposium Held 1968 At Lerum, Sweden*, Stockholm 1968, 367-377.

[6] G. 't Hooft, "Renormalizable lagrangians for massive Yang-Mills fields," *Nucl. Phys.*, vol. B35, pp. 167–188, 1971.

[7] W. M. Yao *et al.*, "Review of particle physics," *J. Phys. G*, vol. 33, pp. 1–1232, 2006.

[8] B. W. Lee, C. Quigg, and H. B. Thacker, "The Strength of Weak Interactions at Very High-Energies and the Higgs Boson Mass," *Phys. Rev. Lett.*, vol. 38, pp. 883–885, 1977.

[9] T. Hambye and K. Riesselmann, "Matching conditions and Higgs mass upper bounds revisited," *Phys. Rev.*, vol. D55, pp. 7255–7262, 1997, arXiv:hep-ph/9610272.

[10] W. Langguth and I. Montvay, "A Numerical estimate of the upper limit for the Higgs boson mass," *Z. Phys.*, vol. C36, p. 725, 1987.

[11] G. Altarelli and G. Isidori, "Lower limit on the Higgs mass in the standard model: An Update," *Phys. Lett.*, vol. B337, pp. 141–144, 1994.

[12] R. Barate *et al.*, "Search for the standard model Higgs boson at LEP," *Phys. Lett.*, vol. B565, pp. 61–75, 2003, arXiv:hep-ex/0306033.

[13] M. T. P. Roco and P. D. Grannis, "Higgs boson discovery prospects at the Tevatron," 1999. To be published in the proceedings of 13th Topical Conference on Hadron Collider Physics, Mumbai, India, 14-20 Jan 1999.

[14] D. C. Kennedy and B. W. Lynn, "Electroweak Radiative Corrections with an Effective Lagrangian: Four Fermion Processes," *Nucl. Phys.*, vol. B322, p. 1, 1989.

[15] LEP-EWWG, "The Lep Electro-Weak Working Group," 2009, URL `http://lepewwg.web.cern.ch/LEPEWWG/`.

[16] A. B. Arbuzov *et al.*, "ZFITTER: A semi-analytical program for fermion pair production in e^+ e^- annihilation, from version 6.21 to version 6.42," *Comput. Phys. Commun.*, vol. 174, pp. 728–758, 2006, arXiv:hep-ph/0507146.

[17] A. Collaboration *et al.*, "Precision Electroweak Measurements and Constraints on the Standard Model," 2009, arXiv:0911.2604 [hep-ex].

[18] GFITTER, "A Generic Fitter Project for HEP Model Testing," 2009, URL `http://gfitter.desy.de/`.

[19] S. Dawson and R. Kauffman, "QCD corrections to Higgs boson production: nonleading terms in the heavy quark limit," *Phys. Rev.*, vol. D49, pp. 2298–2309, 1994, arXiv:hep-ph/9310281.

[20] D. de Florian and M. Grazzini, "Higgs production through gluon fusion: updated cross sections at the Tevatron and the LHC," *Phys. Lett.*, vol. B674, pp. 291–294, 2009, arXiv:0901.2427 [hep-ph].

[21] C. Anastasiou, R. Boughezal, and F. Petriello, "Mixed QCD-electroweak corrections to Higgs boson production in gluon fusion," *JHEP*, vol. 04, p. 003, 2009, arXiv:0811.3458 [hep-ph].

[22] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt, "Parton distributions for the LHC," *Eur. Phys. J.*, vol. C63, pp. 189–285, 2009, arXiv:0901.0002 [hep-ph].

[23] U. Aglietti *et al.*, "Tevatron for LHC report: Higgs," 2006, arXiv:hep-ph/0612172.

[24] A. Djouadi, J. Kalinowski, and M. Spira, "HDECAY: a Program for Higgs Boson Decays in the Standard Model and its Supersymmetric Extension," 1997, URL `doi:10.1016/S0010-4655(97)00123-9`.

[25] T. Aaltonen *et al.*, "Search for a Higgs Boson in $WH \to \ell\nu b\bar{b}$ in $p\bar{p}$ Collisions at $\sqrt{s} = 1.96$ TeV," *Phys. Rev. Lett.*, vol. 103, p. 101802, 2009, arXiv:0906.5613 [hep-ex].

[26] T. Aaltonen *et al.*, "Search for the Higgs boson produced with $Z \to \ell^+\ell^-$ in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV," *Phys. Rev. Lett.*, vol. 101, p. 251803, 2008, arXiv:0807.4493 [hep-ex].

[27] T. Aaltonen *et al.*, "Search for the Higgs boson in events with missing transverse energy and b quark jets produced in proton- antiproton collisions at $\sqrt{s}$ =1.96 TeV," *Phys. Rev. Lett.*, vol. 100, p. 211801, 2008, arXiv:0802.0432 [hep-ex].

[28] V. M. Abazov *et al.*, "Search for Resonant Diphoton Production with the D0 Detector," *Phys. Rev. Lett.*, vol. 102, p. 231801, 2009, arXiv:0901.1887 [hep-ex].

[29] G. Aad *et al.*, "Expected Performance of the ATLAS Experiment - Detector, Trigger and Physics," 2009, arXiv:0901.0512 [hep-ex].

[30] CDF, "Measurement of WW/WZ cross section in lνjj decay using 3.9 fb-1," 2009, URL `http://www-cdf.fnal.gov/physics/ewk/2009/jjlnu/index.html`.

[31] J. M. Campbell and R. K. Ellis, "Radiative corrections to Z b anti-b production," *Phys. Rev.*, vol. D62, p. 114012, 2000, arXiv:hep-ph/0006304.

[32] ATLAS, "ATLAS Detector and physics performace Technical design report," 1999, URL `http://www-cdf.fnal.gov/physics/ewk/2009/jjlnu/index.html`. CERN/LHC 99-14.

[33] G. Altarelli and G. Parisi, "Asymptotic Freedom in Parton Language," *Nucl. Phys.*, vol. B126, p. 298, 1977.

[34] T. Aaltonen *et al.*, "Measurement of the WW Production Cross Section and Search for Anomalous WWg and WWZ Couplings." Submitted to PRL; *http://www-cdf.fnal.gov/physics/ewk/wwllll/index.html*.

[35] CMS, "Search strategy for a Standard Model Higgs boson decaying to two W bosons in the fully leptonic final state.," 2009. CMS PAS HIG-08-006.

[36] CDF, "Observation of $ZZ \to 4$ leptons at CDF.," 2009. CDF Public Note 9910.

[37] CDF, "CDF Search for Higgs to WW* Production using a Combined Matrix Element and Neural Network Technique," 2009. CDF Public Note 9887.

[38] CMS, "Search strategy for the Higgs boson in the ZZ^* decay channel with the CMS experiment," 2009. CMS PAS HIG-08-003.

[39] T. Aaltonen *et al.*, "First Measurement of ZZ Production in $p\bar{p}$ Collisions at $\sqrt{s}$ = 1.96-TeV," *Phys. Rev. Lett.*, vol. 100, p. 201801, 2008, arXiv:0801.4806 [hep-ex].

[40] T. Aaltonen *et al.*, "First Observation of Vector Boson Pairs in a Hadronic Final State at the Tevatron Collider," *Phys. Rev. Lett.*, vol. 103, p. 091803, 2009, arXiv:0905.4714 [hep-ex].

[41] J. Donini *et al.*, "Energy Calibration of b^- Quark Jets with $Z \to b\bar{b}$ Decays at the Tevatron Collider," *Nucl. Instrum. Meth.*, vol. A596, pp. 354–367, 2008, arXiv:0801.3906 [hep-ex].

[42] Fermilab, "Run II Handbook," 2003,
URL http://www-ad.fnal.gov/runII/index.html.

[43] F. B. Division, "The Linac rookie book," 2006,
URL http://www-bdnew.fnal.gov/operations/rookie_books/LINAC_v2.pdf.

[44] F. B. Division, "The Booster rookie book," 2009,
URL http://www-bdnew.fnal.gov/operations/rookie_books/Booster_V4.1.pdf.

[45] F. B. Division, "The Main Injector rookie book," 2006,
URL http://www-bdnew.fnal.gov/operations/rookie_books/Main_Injector_v1.pdf.

[46] F. B. Division, "The Antiproton Source rookie book," 2009,
URL http://www-bdnew.fnal.gov/operations/rookie_books/Pbar_v2.pdf.

[47] F. B. Division, "The Recycler Injector rookie book," 2010,
URL http://www-bdnew.fnal.gov/operations/rookie_books/Recycler_v1.41.pdf.

[48] R. Blair *et al.*, "The CDF-II detector: Technical design report," 1996. FERMILAB-PUB-96-390-E.

[49] F. Abe *et al.*, "The CDF detector: an overview," *Nucl. Instr. Meth.*, vol. A271, pp. 387–403, 1988.

[50] N. Timothy, "The CDF L00 detector," 2001. CDF/PUB/SEC_VTX/PUBLIC/5780.

[51] C. Grozis *et al.*, "A time-of-flight detector for CDF," *Int. J. Mod. Phys.*, vol. A16S1C, pp. 1119–1121, 2001.

[52] A. Artikov *et al.*, "Design and construction of new central and forward muon counters for CDF II," *Nucl. Instrum. Meth.*, vol. A538, pp. 358–371, 2005, arXiv:physics/0403079.

[53] D. Acosta *et al.*, "The CDF Cherenkov luminosity monitor," *Nucl. Instrum. Meth.*, vol. A461, pp. 540–544, 2001.

[54] A. Bardi *et al.*, "The CDF online silicon vertex tracker," *Nucl. Instrum. Meth.*, vol. A485, pp. 178–182, 2002.

[55] A. Bhatti *et al.*, "The CDF level 2 calorimetric trigger upgrade," *Nucl. Instrum. Meth.*, vol. A598, pp. 331–333, 2009.

[56] A. Abulencia *et al.*, "The CDF II 3D-Track Level 2 Trigger Upgrade," 2007. Presented at 15th IEEE Real Time Conference 2007 (RT 07), Batavia, Illinois, 29 Apr - 4 May 2007.

[57] T. Aaltonen *et al.*, "Search for Higgs bosons predicted in two-Higgs-doublet models via decays to tau lepton pairs in 1.96 TeV $p\bar{p}$ collisions," 2009, arXiv:0906.1014 [hep-ex].

[58] A. Bhatti *et al.*, "Determination of the jet energy scale at the collider detector at Fermilab," *Nucl. Instrum. Meth.*, vol. A566, pp. 375–412, 2006, arXiv:hep-ex/0510047.

[59] D. Acosta *et al.*, "Measurement of the $t\bar{t}$ production cross section in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV using lepton + jets events with secondary vertex b -tagging," *Phys. Rev. D*, vol. 71, p. 052003, Mar 2005.

[60] A. Taffard, "Run II Cosmic Ray tagger performances," *CDF Public Note 6255*, 2003.

[61] 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.*, vol. 103, p. 082001, 2009, arXiv:0903.2120 [hep-ph].

[62] S. Moch and P. Uwer, "Heavy-quark pair production at two loops in QCD," *Nucl. Phys. Proc. Suppl.*, vol. 183, pp. 75–80, 2008, arXiv:0807.2794 [hep-ph].

[63] T. Sjostrand, S. Mrenna, and P. Skands, "PYTHIA 6.4 Physics and Manual," *JHEP*, vol. 05, p. 026, 2006, arXiv:hep-ph/0603175.

[64] T. Aaltonen *et al.*, "Studying the underlying event in Drell-Yan and high transverse momentum jet production at Tevatron." To be submitted to PRD; *http://www-cdf.fnal.gov/physics/new/qcd/UEinDY_2008/analysis.html*.

[65] U. Baur and E. L. Berger, "Probing the weak boson sector in $Z\gamma$ production at hadron colliders," *Phys. Rev.*, vol. D47, pp. 4889–4904, 1993.

[66] S. Frixione and B. R. Webber, "Matching NLO QCD computations and parton shower simulations," *JHEP*, vol. 06, p. 029, 2002, arXiv:hep-ph/0204244.

[67] G. Corcella *et al.*, "HERWIG 6.4 release note," 2001, arXiv:hep-ph/0201201.

[68] R. Brun, R. Hagelberg, M. Hansroul, and J. C. Lassalle, "GEANT: Simulation program for particle physics experiments. User guide and reference manual," 1978. CERN-DD-78-2-REV.

[69] M. Dittmar and H. K. Dreiner, "How to find a Higgs boson with a mass between 155-GeV - 180-GeV at the LHC," *Phys. Rev.*, vol. D55, pp. 167–172, 1997, arXiv:hep-ph/9608317.

[70] V. M. Abazov *et al.*, "Measurement of the top quark mass in the lepton + jets final state with the matrix element method," *Phys. Rev.*, vol. D74, p. 092005, 2006, arXiv:hep-ex/0609053.

[71] J. M. Campbell and R. K. Ellis, "Update on vector boson pair production at hadron colliders," *Phys. Rev. D*, vol. 60, p. 113006, Nov 1999.

[72] H. L. Lai *et al.*, "Global QCD analysis of parton structure of the nucleon: CTEQ5 parton distributions," *Eur. Phys. J.*, vol. C12, pp. 375–392, 2000, arXiv:hep-ph/9903282.

[73] T. M. Mitchell, *Machine Learning*. McGraw-Hill Science/Engineering/Math, 1997.

[74] M. Feindt, "A Neural Bayesian Estimator for Conditional Probability Densities," 2004,
URL `http://www.citebase.org/abstract?id=oai:arXiv.org:physics/0402093`.

[75] C. Anastasiou, G. Dissertori, M. Grazzini, F. Stockli, and B. R. Webber, "Perturbative QCD effects and the search for a $H \to WW \to l\nu l\nu$ signal at the Tevatron," *JHEP*, vol. 08, p. 099, 2009, arXiv:0905.3529 [hep-ph].

[76] S. Catani and M. Grazzini, "An NNLO subtraction formalism in hadron collisions and its application to Higgs boson production at the LHC," *Physical Review Letters*, vol. 98, p. 222002, 2007.

[77] CDF, "Combined Upper Limit on Standard Model Higgs Boson Production," 2009. CDF Public Note 9999.

[78] DØ, "Combined Upper Limit on Standard Model Higgs Boson Production from the DØ experiment in 2.1-5.4 fb^{-1} of data," 2009. DØ Public Note 6008-CONF.

[79] T. D. Collaboration, "Search for Higgs boson production in dilepton and missing energy final states with 5.4 fb-1 of p-pbar collisions at sqrt(s) =1.96 TeV," 2010, arXiv:1001.4481 [hep-ex]. Submitted to Phys. Rev. Lett.

[80] CDF and DØ, "Combined CDF and D0 Upper Limits on Standard Model Higgs-Boson Production with 2.1 - 5.4 fb-1 of Data," 2009, arXiv:0911.3930 [hep-ex].

[81] C. Collaboration and D. Collaboration, "Combination of Tevatron searches for the standard model Higgs boson in the W+W- decay mode," 2010, arXiv:1001.4162 [hep-ex]. Submitted to Phys. Rev. Lett.

[82] T. Junk, "Confidence Level Computation for Combining Searches with Small Statistics," *Nucl. Instrum. Meth.*, vol. A434, pp. 435–443, 1999, arXiv:hep-ex/9902006.

[83] A. L. Read, "Presentation of search results: The CL(s) technique," *J. Phys.*, vol. G28, pp. 2693–2704, 2002.

[84] GFITTER, "GSM - A Gfitter Package for the Global Electroweak Fit," 2010,
URL `http://gfitter.desy.de/GSM/`.

[85] T. Aaltonen *et al.*, "Inclusive Search for Standard Model Higgs Boson Production in the WW Decay Channel using the CDF II Detector," 2010, arXiv:1001.4468 [hep-ex]. Submitted to Phys. Rev. Lett.