Search for the Standard Model Higgs boson in the decay mode $H \to W^+W^- \to \ell^+\nu\ell^-\nu$ at the DØ experiment

Dissertation zur Erlangung der Doktorwürde
Vorgelegt von
Björn Penning

Albert-Ludwigs-University of Freiburg
Faculty for Mathematics and Physics

Dekan:	Prof. Dr. Kay Königsmann
Leiter der Arbeit:	Prof. Dr. Karl Jakobs
Referent:	Prof. Dr. Karl Jakobs
Koreferent:	Prof. Dr. Gregor Herten
Tag der Verkündigung des Prüfungsergebnisses:	7. September 2009

for Gilles

Contents

1 Introduction

The question of the nature and principles of the universe and our place in it is the driving force of science since Mesopotamian astronomers glanced for the first time at the starry sky and Greek atomism has been formulated. During the last hundred years modern science was able to extend its knowledge tremendously, answering many questions, opening entirely new fields but as well raising many new questions. Particularly Astronomy, Astroparticle Physics and Particle Physics lead the race to answer these fundamental and ancient questions experimentally.

Today it is known that matter consists of fermions, the quarks and leptons. Four fundamental forces are acting between these particles, the electromagnetic, the strong, the weak and the gravitational force. These forces are mediated by particles called bosons. Our confirmed knowledge of particle physics is based on these particles and the theory describing their dynamics, the *Standard Model of Particles*. Many experimental measurements show an excellent agreement between observation and theory but the origin of the particle masses and therefore the electroweak symmetry breaking remains unexplained. The mechanism proposed to solve this issue involves the introduction of a complex doublet of scalar fields which generates the masses of elementary particles via their mutual interactions. This *Higgs mechanism* also gives rise to a single neutral scalar boson with an unpredicted mass, the *Higgs boson*.
During the last twenty years several experiments have searched for the Higgs boson but so far it escaped direct observation. Nevertheless these studies allow to further constrain its mass range. The last experimental limits on the Higgs mass have been set in 2001 at the LEP collider, an electron positron machine close to Geneva, Switzerland. The lower limit set on the Higgs boson mass is $m_H > 114.4$ GeV/c^2 [1] and remained for many years the last experimental constraint on the Standard Model Higgs Boson due to the shutdown of the LEP collider and the experimental challenges at hadron machines as the $\mathrm{TEVATRON}$.
This thesis was performed using data from the DØ detector located at the Fermi National Accelerator Laboratory in Batavia, IL. Final states containing two electrons or a muon and a tau in combination with missing transverse energy were studied to search for the Standard Model Higgs boson, utilizing up to $4.2\ \mathrm{fb}^{-1}$ of integrated luminosity.

In 2008 the CDF and DØ experiments in a combined effort were able to reach for the first time at a hadron collider the sensitivity to further constrain the possible Standard Model Higgs boson mass range. The research conducted for this thesis played a pivotal role in this effort. Improved methods for lepton identification, background separation, assessment of systematic uncertainties and new decay channels have been studied, developed and utilized. Along with similar efforts at the

CDF experiment these improvements led finally the the important result of excluding the presence of a Standard Model Higgs boson in a mass range of $m_H = 160 - 170$ GeV/c^2 at 95% Confidence Level.
Many of the challenges and methods found in the present analysis will probably in a similar way be ingredients of a Higgs boson evidence or discovery in the near future, either at the TEVATRON or more likely at the soon starting Large Hadron Collider (LHC). Continuing to pursue the Higgs boson we are looking forward to many exciting results at the TEVATRON and soon at the LHC.

In Chapter 2 an introduction to the Standard Model of particle physics and the Higgs mechanism is given, followed by a brief outline of existing theoretical and experimental constraints on the Higgs boson mass before summarizing the Higgs boson production modes. Chapter 3 gives an overview of the experimental setup. This is followed by a description of the reconstruction of the objects produced in proton-antiproton collisions in Chapter 4 and the necessary calorimeter calibrations in Chapter 5. Chapter 6 follows with an explanation of the phenomenology of the proton-antiproton collisions and the data samples used. In Chapter 7 the search for the Standard Model Higgs boson using a di-electron final state is discussed, followed by the analysis of the final states using muons and hadronic decaying taus in Chapter 8. Finally a short outlook for the prospects of Higgs boson searches is given in Chapter 9.

2
The Standard Model and the Higgs Boson

The following chapter briefly outlines the theoretical principles used in this thesis. The visible matter in the universe is made out of quarks and leptons. The properties and interactions of these elementary particles are described by the Standard Model (SM) of particle physics. Although the SM predicts many phenomena correctly the masses of the fermions are not predicted and are free input parameters. A short introduction to the Standard Model is given followed by a more detailed explanation of the Higgs mechanism, the simplest proposed mechanism to generate fermion and boson masses in the Standard Model. The last sections of this chapter gives a brief overview of production and decay of the Higgs boson concluding with a summary of the status of the Higgs boson searches performed so far.

2.1 The Standard Model

Since the time of Democritus and his mentor Leucippus [2], almost 2500 years ago, various concepts were suggested to explain the nature of matter as consisting out of smallest, not further dividable parts. This idea - initially more a philosophically reasoning rather than an empirical observation - persisted and evolved into the modern science of particle physics. The modern theory, the *Standard Model*, explains all phenomena in particle physics in terms of properties and interaction of a small number of particles. Fermions with spin $\frac{1}{2}$ are interacting via fundamental forces.

Force	**Gauge Boson**	**Range**	**Interacts with**
electromagnetic	Photon (γ)	infinite	charged particles
weak	$W^{\pm}, Z$	$\sim 10^{-18}$m	quarks, leptons, $W^{\pm}, Z$
strong	Gluons (g)	$\sim 10^{-15}$m	quarks, gluons

Table 2.1: Overview of the fundamental forces.

In the Standard Model the fermions are arranged in families of quarks and leptons. Three types of charged leptons exist, the electron (e), muon (μ), and the tau (τ). For each charged lepton there is an uncharged partner, the neutrino (ν). The three lepton families are: (ν_e, e^-), (ν_μ, μ^-), (ν_τ, τ^-).
The forces, the electromagnetic, the weak and the strong interaction are carried by gauge bosons, particles with an integer spin. An overview of these forces is given in Table 2.1. Although the electromagnetic and weak force appear to be very different at low energies these two forces can

be unified at energies above about 100 GeV into the electroweak interaction. The fourth known fundamental force, gravity, cannot yet be consistently included in the theory. Since the strength of the gravitational force is many orders of magnitude smaller than the remaining forces it can be neglected in most particle physics calculations and is not part of the Standard Model. In the following sections it will be shown how the interactions of the bosons with the fermions emerge naturally as a consequence of the principle of local gauge invariance.

The stable matter in the universe is made out of electrons, protons and neutrons. The protons and neutrons consist out of quarks bound by gluon fields. There are three quarks bearing an electric charge of $-\frac{1}{3}$, *down (d), strange (s), bottom (b)* and three with a positive charge of $+\frac{2}{3}$, *up (u), charm (c), top (t)*. As in the case of the leptons they can be arranged in three families: (u,d), (c,s), (t,b). In addition to the strong and the electromagnetic force quarks also interact via the weak force. Quarks carry another charge called color charge. There are three types of color charge, named blue, green and red. Each of them is complemented by an anticolor carried by the anti-quarks. Quarks are never observed individually but always in combinations such that the final compound state is colorless. This feature of the strong interaction is called *confinement*. Even if the quarks are bound within nuclei by the strong force, they can act as free particles when probed at a sufficiently high energy scale. This intriguing property of the strong interaction is called *asymptotic freedom*.

Every fermion has a corresponding anti-fermion with the same mass but the additive quantum numbers as for example spin, baryon- or lepton number are of opposite sign. For the neutral neutrinos it is not clear yet if neutrino and anti-neutrinos are different particles or if they are their own anti-particles. In the first case one would speak of Dirac neutrinos, in the latter of Majorana neutrinos [3] but this would require a not Standard Model theory to be described.

It appears that all neutrinos are left handed and all anti-neutrinos show right handed chirality. As such they are handled in the Standard Model. However, the recent observation of neutrino oscillations [4] points to the existence of a neutrino mass. Because chirality and helicity are the same for massless particles but not for particles with mass, massive neutrinos cannot be easily accommodated in the Standard Model. Together with astrophysical evidence supporting the existence of dark matter [5] and dark energy [6, 7] this is one of the indications for physics beyond the Standard Model.

2.2 Symmetries and Quantum Numbers

Quantum numbers and their conservation are a very useful concept in particle physics. Most conserved quantum numbers can be traced back to an invariance of the theory under some transformation, for example angular momentum conservation follows from the invariance under rotations in the four dimensional Minkoswki space. Other transformation acting on internal degrees of freedom may as well lead to conservation laws. One of the best examples is the electric charge whose conservation follows from the invariance under local $U(1)$ gauge transformations.
This is formulated as *Noether's Theorem*: To every differentiable symmetry generated by local actions, there corresponds a conserved current.
The properties of any set of symmetry operations must have the defining properties of a group. These groups can be finite or infinite which leads likewise to discrete or continuous symmetries and

Fermion	**Particle**			**Q[e]**	$\lvert\vec{\mathbf{T}}\rvert$	$\mathbf{T_3}$	$\mathbf{SU(3)_C}$	$\mathbf{SU(2)_L}$	$\mathbf{U(1)_Y}$
leptons	$\begin{pmatrix}\nu_e\\ e\end{pmatrix}_L$	$\begin{pmatrix}\nu_\mu\\ \mu\end{pmatrix}_L$	$\begin{pmatrix}\nu_\tau\\ \tau\end{pmatrix}_L$	0 -1	$\frac{1}{2}$ $\frac{1}{2}$	$\frac{1}{2}$ $-\frac{1}{2}$	1	2	-1
	e_R	μ_R	τ_R	-1	0	0	1	1	-2
quarks	$\begin{pmatrix}u\\ d'\end{pmatrix}_L$	$\begin{pmatrix}c\\ s'\end{pmatrix}_L$	$\begin{pmatrix}t\\ b'\end{pmatrix}_L$	$\frac{2}{3}$ $-\frac{1}{3}$	$\frac{1}{2}$ $\frac{1}{2}$	$\frac{1}{2}$ $-\frac{1}{2}$	3	2	$\frac{1}{3}$ $\frac{1}{3}$
	u_R	c_R	t_R	$\frac{2}{3}$	0	0	3	1	$\frac{4}{3}$
	d_R	s_R	b_R	$-\frac{1}{3}$	0	0	3	1	$-\frac{2}{3}$

Table 2.2: The fermions of the Standard Model. The mixing of left handed (L) and right handed (R) eigenstates of the electroweak interaction results in the mass eigenstates. The primed symbols d', s' and b' indicate the eigenstates of the electroweak interaction which are connected by the Cabibbo-Kobayachi-Maskawa-Matrix to the mass eigenstates d, s, and b. The discovery of neutrino oscillations requires neutrinos to be massive, therefore they also have right-handed contributions. T and T_3 are the total weak-isospin and its third component. Q is the electric charge.

are usually mathematically represented by special unitary or matrices of dimension n, sometimes with the additional requirement that their determinant is 1. Those groups are referred to as $SU(n)$. The Standard Model is based on the $SU(3)_C \times SU(2)_L \times U(1)_Y$ symmetry group, where C, L and Y refer to the various charges, namely color, weak isospin and hypercharge. These charges will be explained more in detail in the following. Table 2.2 gives all the fermions of the Standard Model along with their quantum numbers for charge, hypercharge and isospin as explained in Sec. 2.3. The bosons are listed n in the same fashion in Table 2.3.

Bosons	**Field**	**Q [e]**	$\mathbf{T_3}$	$\mathbf{SU(3)_C}$	$\mathbf{SU(2)_L}$	$\mathbf{U(1)_Y}$
$U(1)_Y$ gauge field	B	0	0	1	1	0
$SU(2)_L$ gauge field	W^+, W^-, W^3	$1, 0, -1$	$1, 0, -1$	1	3	0
$SU(3)_c$ gauge field	$G^1 ... G^3$	0	0	8	1	0
Higgs field Φ	$\Phi^+,\ \Phi^0$	$1, 0$	$\frac{1}{2}, -\frac{1}{2}$	1	2	1

Table 2.3: The bosons of the Standard Model. The gauge fields carry spin 1 whereas the Higgs field carries spin 0.

2.3 Electroweak Interaction

The electroweak interaction as the Standard Model itself is described by a gauge theory. The corresponding symmetry group is the combination of $SU(2)_L$ and $U(1)_Y$. The index L denotes that only left handed fermions interact weakly. Left handed fermions can be represented as doublet of the weak isospin $|\vec{T}| = \frac{1}{2}$ with the eigenvalues $T_3 = \pm\frac{1}{2}$. Right handed fermions are in contrast singlet states with $\vec{T} = T_3 = 0$.

Just as Q generates the group $U(1)_{em}$, the group $U(1)_Y$ is generated by the hypercharge Y which can be calculated using the Gell-Mann-Nishijima relation.

$$Q = T_3 + \frac{Y}{2} \tag{2.1}$$

The hypercharge is the same for all particles within the same multiplet.

The Lagrangian $\mathcal{L}_{EW}$ has to be invariant under transformations of the group $SU(2)_L \times U(1)_Y$, therefore one obtains an isotriplet of vector fields W^i_ν $(i = 1, 2, 3)$ coupled with the strength g to the weak isospin. Additionally one obtains a single vector field B_μ coupling with g' to the weak hypercharge. Therefore the Lagrangian is given by

$$\mathcal{L}_{EW} = \bar{\Psi}\gamma_\mu D^\mu \Psi - \frac{1}{4}\left[\mathbf{W}_{\mu\nu}\mathbf{W}^{\mu\nu} + B_{\mu\nu}B^{\mu\nu}\right] \tag{2.2}$$

where D_μ is the covariant derivative

$$D_\mu = i\partial_\mu - g\mathbf{T}\cdot\mathbf{W}_\nu + g'YB_\nu \tag{2.3}$$

and the field tensors are given by

$$\mathbf{W}_{\mu\nu} = \partial_\mu \mathbf{W}_\nu - \partial_\nu \mathbf{W}_\mu - g\mathbf{W}_\mu \mathbf{W}_\nu, \tag{2.4}$$

$$B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu \tag{2.5}$$

The two neutral fields mix in such a way that the physical state, the mass eigenstates, are

$$A_\mu = B\mu\cos\theta_W + W^3_\nu \sin\theta_W \qquad \text{massless} \tag{2.6}$$

$$Z_\mu = -B\mu\sin\theta_W + W^3_\nu \cos\theta_W \qquad \text{massive} \tag{2.7}$$

here θ_W is the Weinberg or weak mixing angle. It has been measured to be $\sin^2\theta_W = 0.231$ [8]. The two remaining fields W^1_μ and W^2_μ are mixing as well, forming the charged gauge boson:

$$W^\pm_\mu = \frac{1}{2}\left(W^1_\mu \mp W^2_\mu\right) \tag{2.8}$$

The fields A_μ and Z_μ can be identified with the photon and the Z boson. These fields couple to both, left and right handed fermions contrary to the charged gauge bosons $W^\pm$ which couple only

the left handed fermions, e.g particles with $|\vec{T}| \neq 0$. By writing out the neutral current interaction and due to the fact that the photon field couples with strength e to charged fermions it can be shown that:

$$g \sin\theta_W = g' \cos\theta_W = e. \tag{2.9}$$

The coupling of the Z^0 boson is given by

$$-i\frac{g}{\cos\theta_W}\gamma^\mu \frac{1}{2}\left(c_v^f - c_A^f \gamma^5\right) \tag{2.10}$$

here c_v^f and c_A^f represent the vector and axial couplings, respectively:

$$c_v^f = T_3^f - 2\sin^2\theta_W Q_f \tag{2.11}$$

$$c_A^f = T_3^f \tag{2.12}$$

The vertex of the charged bosons is described by:

$$-i\frac{g}{\sqrt{2}}\gamma^\mu \frac{1}{2}\left(1 - \gamma^5\right) \tag{2.13}$$

Because particles of interest, e.g leptons, quarks and likewise the gauge bosons $Z^0, W^\pm$ are not massless the symmetry has to be broken and the particles acquire mass.

2.4 Quantum Chromodynamics and Strong Interaction

The strong interaction or *Quantum Chromodynamic* (QCD) is based on the symmetry group $SU(3)_C$, the index C stands for 'color'. The Lagrangian of the QCD interaction is given by

$$\mathcal{L}_{QCD} = \bar{q}\gamma_\mu D^\mu q - \mathbf{G}^a_{\mu\nu}\mathbf{G}^{\mu\nu}_a \tag{2.14}$$

and the covariant derivative by

$$D_\mu = i\partial_\mu - g_s \mathbf{T}^a \cdot \mathbf{G}^a_\mu. \tag{2.15}$$

The field tensor $\mathbf{G}^a_{\mu\nu}$ represents:

$$\mathbf{G}^a_{\mu\nu} = \partial_\mu \mathbf{G}^a_\nu - \partial_\nu \mathbf{G}^a_\mu - g_s f_{abc}\mathbf{G}^b_\mu \mathbf{W}^c_\nu. \tag{2.16}$$

The coupling is represented by g_s and f_{abc} (with $a, b, c = 1...8$) are the structure constants of the $SU(3)$ group. The color charge is ordered in color triplets. Color charges are carried by quarks and the mediating particles of the strong force, the gluons. The gluons are forming a color-octet and show self-interaction because it is a non-Abelian gauge theory. This self-interaction leads to an increasing strength of the coupling with increasing distance of the color charges, therefore making it impossible to observe free quarks but only colorless states as *mesons* ($q\bar{q}$) and *baryons* (qqq). This is called *confinement*. In contrary the coupling decreases for very small distances, called *asymptotic freedom*, and can be calculated using perturbative theory.

2.5 The Higgs Mechanism

The theory of weak interaction is predicting massless gauge bosons in contrast to the theoretical observation where all gauge boson except of the photons have a substantial mass. Unfortunately the manual introduction of mass terms like $\frac{M^2}{2}\mathbf{W}_{\mu\nu}\mathbf{W}^{\mu\nu}$ to the Lagrangian $\mathcal{L}$ leads to unrenormizable divergences, rendering the theory meaningless. The masses of particles can be generated by introducing the Higgs mechanism as opposed to putting them in by hand. Here an additional potential is added to the Lagrangian which has two minima, leading to *spontaneous symmetry breaking* and generating the masses of the particles.

2.5.1 The Higgs Potential

The Lagrangian shown in Eq. 2.2 remains gauge invariant when adding a potential of the form

$$\mathcal{L}_{Higgs} = (D_\mu \Phi)^\dagger (D^\mu \Phi)^\dagger - V(\Phi). \tag{2.17}$$

Here Φ is a 2 dimensional scalar field given by

$$\Phi(x) = \begin{pmatrix} \Phi^+ \\ \Phi^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \Phi_1 + i\Phi_2 \\ \Phi_3 + i\Phi_4 \end{pmatrix} \tag{2.18}$$

and $V(\Phi)$ by

$$V(\Phi) = \mu^2 \Phi^\dagger \Phi + \lambda \left(\Phi^\dagger \Phi\right)^2 \tag{2.19}$$

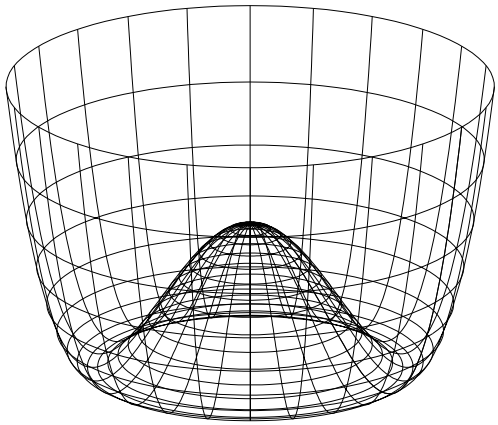


Figure 2.1: Example of the Higgs potential of a scalar field [9].

When choosing $\lambda > 0$ and $\mu^2 < 0$ one obtains two minima for the potential. These satisfy:

$$\Phi^\dagger \Phi = \frac{1}{2}\left(\Phi_1^2 + \Phi_2^2 + \Phi_3^2 + \Phi_4^2\right) = -\frac{\mu^2}{2}\lambda = v^2 \tag{2.20}$$

Setting $\Phi_1 = 0,\ \Phi_2 = 0,\ \Phi_3 = 0$ and $\Phi_4 = v$ the vacuum expectation value is given by

$$\langle \Phi \rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix}, \tag{2.21}$$

Perturbative calculations should involve expressions around the classical minimum, $\Phi = \pm v$ and one therefore writes

$$\Phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix} \tag{2.22}$$

The Higgs particle is interpreted as a space-time dependent radial fluctuation $h(x)$ of the field Φ near the vacuum configuration. Due to the local gauge invariance each point satisfies Eq. 2.22 because rotations of the isospin can performed at each of this points. The only field remaining is the Higgs field $h(x)$.

The Lagrangian, expressed in terms of the vacuum expectation value v and the physical state $h(x)$ describes a scalar particle with the mass $m_h = \sqrt{2\lambda}v$

$$\mathcal{L} = \frac{1}{2}\partial_\mu \partial^\nu h - \lambda v^2 h^2 - \lambda v h^3 - \frac{\lambda}{4} h^4. \tag{2.23}$$

The scalar particle described by Eq. 2.23 is referred to as the *Standard Model Higgs boson*. This neutral state after electroweak symmetry breaking is the only component left in Eq. 2.18. The other three components are the longitudinally polarized components of the weak vector boson. This boson carries no electrical charge and it couples proportional to fermions and heavy gauge bosons. The coupling is proportional to the mass of the particles.

2.5.2 Masses of the Gauge Bosons

In the Lagrangian the mass terms are given by representations like $\frac{1}{2} M \Psi^\dagger \Psi$ and therefore are quadratic in their fields. By inserting the vacuum expectation value of Eq. 2.21 into the Lagrangian Eq. 2.17 one is able to study its structure. The covariant derivative then is reduced to the form [10]:

$$D_\mu \langle \Phi \rangle = -\left(\frac{ig}{2} \begin{pmatrix} W_\mu^3 & W_\mu^1 - iW_\mu^2 \\ W_\mu^1 + iW_\mu^2 & W_\mu^3 \end{pmatrix} + \frac{ig'}{2} B_\mu \right) \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix} \tag{2.24}$$

By using that in the Lagrangian one obtains additional quadratic field terms from the expression $(D_\mu \langle \Phi \rangle)^\dagger (D_\mu \langle \Phi \rangle)$:

$$(D_\mu \langle \Phi \rangle)^\dagger (D_\mu \langle \Phi \rangle) = \frac{1}{8} v^2 g^2 \left((W_\mu^1)^2 + (W_\mu^2)^2 \right) + \frac{1}{8} v^2 (g' B_\mu - g W_\mu^3)^2 \tag{2.25}$$

Using Eq. 2.6 - 2.9 one obtains:

$$\frac{1}{8} v^2 g^2 \left((W_\mu^1)^2 + (W_\mu^2)^2 \right) + \frac{1}{8} v^2 \left(g' B_\mu - g W_\mu^3 \right) = \left(\frac{1}{2} v g \right)^2 W_\mu^+ W^{-\mu}_\mu + \frac{1}{4} v^2 (g^2 + g'^2) Z_\mu Z^\mu \tag{2.26}$$

The symmetry breaking leads to three additional terms with quadratic fields. These mass terms can be identified with the gauge fields W^+, W^- and Z^0. All gauge boson except the photon acquired mass. The photon remains massless because Φ is invariant under transformations of $U(1)_Y$, generated by Q.

Equation 2.26 contains the gauge boson masses:

$$M_W^2 = \frac{1}{4}v^2 g^2, \qquad M_Z^2 = \frac{1}{4}v^2(g^2 + g'^2), \qquad M_\gamma^2 = 0 \tag{2.27}$$

Using Eq. 2.6 and Eq. 2.7 one can show that the W and Z mass are connected via the following relation:

$$\cos\theta_W = \frac{M_W}{M_Z} \tag{2.28}$$

which is in very good agreement with experimental results [8]. The vacuum expectation value v of the Higgs potential can be derived from the Fermi constant: $G_F = 1.667 \cdot 10^{-5}\text{GeV}^{-2}$

$$v^2 = 4\frac{M_W^2}{g^2} = \frac{1}{\sqrt{2}G_F} \approx (250\ \text{GeV})^2 \tag{2.29}$$

2.6 Theoretical Constraints on the Higgs Boson Mass

Several issues arising in the scalar sector of the Standard Model link the mass of the Higgs boson to the energy scale where the validity of the Standard Model is expected to fail. Below that scale, the Standard Model is the extremely successful effective field theory that emerges from the electroweak precision tests. Above that scale the Standard Model has to be embedded into some more general theory that gives origin to a wealth of new physics phenomena. Therefore we are able to connect the Higgs boson mass and the scale of new physics Λ.

2.6.1 Triviality and Vacuum Stability

The requirement that the Higgs' quartic coupling has to remain finite at high energy scales is called *triviality*. In the scalar sector of the Standard Model the quartic coupling is

$$\lambda = \frac{m_H^2}{2\nu^2} \tag{2.30}$$

The coupling λ changes with the effective energy scale Q due to the self interactions of the scalar field [11, 12]:

$$\frac{d\lambda}{dt} = \frac{3\lambda^2}{4\pi^2} \tag{2.31}$$

where $t = \log(Q^2/Q_0^2)$ and Q_0 is some reference scale. By solving Eq. 2.31 one sees that $\lambda(Q)$ becomes infinite for $\lambda \to \infty$, the Landau pole. This is independent of the reference scale $\lambda(Q_0)$. Alternatively $\lambda(Q_0) \to 0$ for $Q \to 0$ with $\lambda(Q) > 0$. This behavior is shown in Fig. 2.2.

When requiring the quartic coupling to be finite at high scales of Λ [11]:

$$\frac{1}{\lambda(\Lambda)} > 0 \tag{2.32}$$

can be interpreted as a bound on the Higgs boson mass .

$$m_H^2 = \frac{8\pi^2 v^2}{3\log\left(\Lambda^2/v^2\right)} \tag{2.33}$$

By requiring the Standard Model to be valid up to the scale of grand unified theories, $\Lambda \sim 10^{16}$ GeV, the approximate upper bound on the Higgs boson mass is [12]:

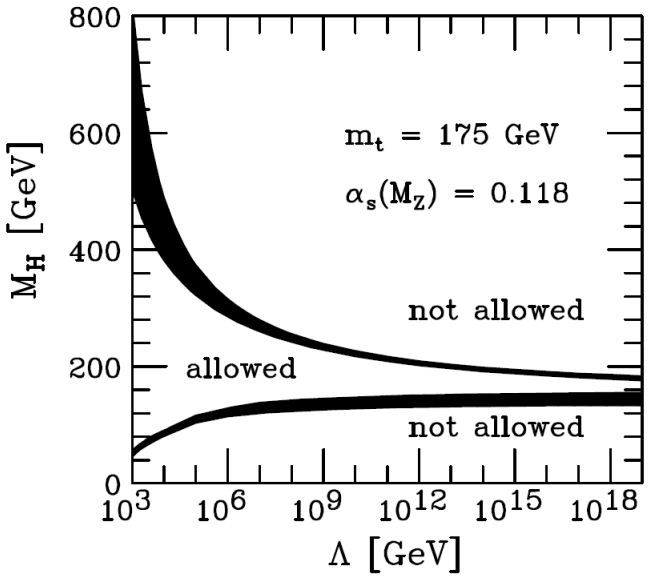


Figure 2.2: Theoretical limits on the Higgs boson mass as a function of the scale of new physics, Λ. The allowed region is between the curves [12].

$$m_H < 160 \text{ GeV} \tag{2.34}$$

As the scale of Λ becomes smaller the limit on the Higgs boson mass becomes progressively weaker. Λ is often interpreted at the scale of new physics because the Standard Model cannot be extended beyond that scale. Another bound can be derived because the top quark Yukawa coupling gives a negative contribution to the beta function of the scalar self-coupling. This causes the scalar self-coupling to decrease, eventually becoming negative if the Higgs boson mass is too low. Then the value of the effective potential becomes negative, dropping lower than the standard model vacuum [13]. This is equivalent to the statement that λ remains positive at all scales of Λ, $\lambda(\Lambda) > 0$ because if λ becomes negative the potential has no state of minimum energy and the

vacuum is not stable any more. It can be shown that if the Standard Model is valid up to the scales of $\sim 10^{16}$ GeV one obtains [11]:

$$m_H(\text{GeV}) = 130 + 2(m_t - 170) \tag{2.35}$$

Therefore the mass of the Higgs boson is restricted between $\sim 126-$ ~ 160 GeV. It should be pointed out that this is exactly the mass range preferred by the electroweak precision measurements as discussed in Sec. 2.7.2.

2.6.2 Naturalness

One of the most striking theoretical inadequacies of the Standard Model arises when computing the quantum corrections to the Higgs boson mass. The one-loop corrections depend quadratically on a high scale cut-off of Λ. Assuming the allowed mass region for the electroweak precision measurements one can deduce that the Λ cut off scale has to be at the order of TeV in order to avoid large corrections. This is the *hierarchy problem*: Why should the particular choice of Λ be at the TeV scale and not the Planck scale?
Various models have been suggested to avoid the hierarchy problem. The basic approach is to postulate new particles contributing to the Higgs boson mass renormalization and canceling the Standard Model contributions. Best known are *Supersymmetric Models* which postulate scalar partners to the know fermions with the right coupling to cancel the Standard Model contributions. Little Higgs models on the other side acquire the necessary cancellation by using particles with the same spin as the Standard Model particles.

2.6.3 Unitarity

Another limit is obtained by studying scattering processes at very high energies.

Considering the scattering of longitudinal gauge bosons $W_L^+W_L^- \rightarrow W_L^+W_L^-$ and applying unitarity requirements for $s >> m_H^2$ one obtains [11]:

$$m_H < 870 \text{ GeV}. \tag{2.36}$$

By considering coupled channels like $W_L^+W_L^- \rightarrow Z_LZ_L$ the bound can be lowered to

$$m_H < 710 \text{ GeV}. \tag{2.37}$$

This means for heavier Higgs boson masses perturbation theory is not valid. In this case the Higgs boson is either not heavier or another mechanism has to regulate the $W_L^+W_L^-$ scattering cross section.

By studying the case that $s << m_H^2$ one obtains the critical scale $\sqrt{s_c}$ at which perturbative unitarity has to be violated:

$$\sqrt{s_c} < 1.8 \text{ TeV}. \tag{2.38}$$

And using more constraining channels this bound can be reduced to

$$\sqrt{s_c} < 1.2 \text{ TeV}. \tag{2.39}$$

Here one sees again a strong indication for new physics at the TeV scale. The region has been started to be probed at the TEVATRON and will be further explored at the LHC.

2.7 Previous Higgs Boson Searches

2.7.1 Direct Higgs Boson searches

Even before the Large-Electron-Positron (LEP) era many Higgs boson searches were performed, the most important are the search in π decays up to 110 MeV and the result that $m_H \neq 0$ from electron-deuteron scattering [14]. However, the most stringent bound on the Higgs boson mass has been set during the very successful run of the LEP collider at CERN. LEP was the most powerful lepton collider built up to today and operated from 1989 to 2000. During the first data taking period LEP was able to push the limits on the Higgs boson mass of $m_H > 66$ GeV and after a successful upgrade the final LEP limit on the Higgs boson mass reached $m_H > 114.4$ GeV at 95% confidence level (C.L.). This remarkable results remained the last experimental constraint for another 8 years until the Summer 2008. Combining the results from the two Tevatron experiments CDF and DØ it was possible for the first time to obtain sensitivity to Standard Model Higgs boson in the region not yet excluded by LEP [15].

At LEP2 the Higgs boson was mainly produced by associated production of the Higgs boson with a Z boson as shown in Fig. 2.3. This requirement of associated production introduces a kinematic limit on the mass of a possible Higgs boson at $m_{tresh} = \sqrt{s} - m_Z$. The reach of the LEP Higgs boson search was critically dependent on the LEP energy and luminosity.

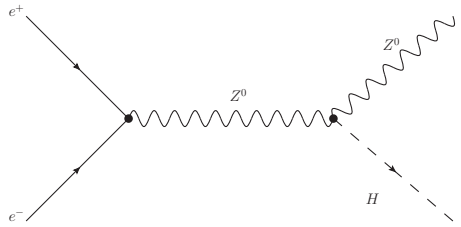


Figure 2.3: Higgs-Strahlung, dominant Higgs boson production mode at LEP

At the energies accessible at LEP (Fig. 2.12) the Higgs boson decays dominantly into $b\bar{b}$ pairs. The Standard Model Higgs boson search was performed for a set of channels. The main search channel at LEP was the 4 jet channel where also the Z decayed into jets, considering also the less frequent decays into cc, gluons and $\tau\tau$. The second best channel is the Z decay in neutrinos, resulting in a final state characterized by 2 jets and missing transverse momentum. The main background came from multijet, WW and ZZ production.

LEP was not able to find any evidence for a Higgs boson although a fluctuation had been observed. This fluctuation, initially up to 2.9 σ, was stemming mainly from candidate events in the ALEPH detector and one event in the L3 detector. The final observed LEP limit on the Higgs boson mass is $m_H > 114.4$ GeV. Figure 2.4. shows the 95% CL_s value as function of the Higgs boson mass m_H, leading to the mentioned limit.

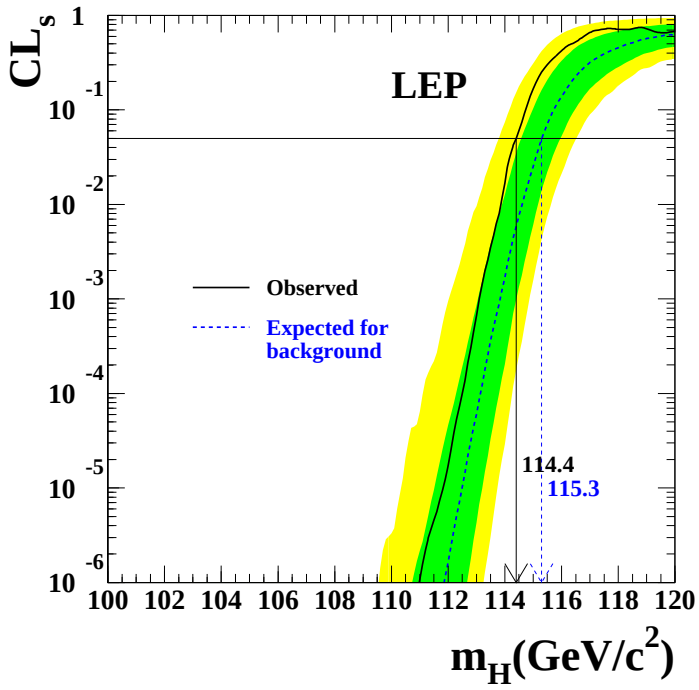


Figure 2.4: Final LEP limit on the Higgs boson mass [16]. The dark and light shaded bands around the median expected line correspond to the 68% and 95% probability bands. The intersection of the horizontal line for $CL_s = 0.05$ with the observed curve is used to define the 95% confidence level lower bound on the mass of the Standard Model Higgs boson.

2.7.2 Indirect Higgs Boson Mass Constraints

The existence of a Higgs boson would not only lead to a renormalizable and therefore meaningful theory, it as well has significant impact on electroweak parameters. For example the Higgs boson contributes to radiative corrections on the top quark and W boson masses. Therefore precision measurements of the electroweak parameters can be used to obtain constraints on the Higgs boson mass. The *1-loop effects* of the top quark and Higgs boson on the W boson mass are shown in Fig. 2.5.

The three main electroweak observables are the electroweak ρ-parameter, the effective leptonic weak mixing angle $\sin^2 \Theta_{eff}^{lept}$ and the W boson mass defined as

$$\rho = \frac{m_W^2}{m_Z^2} \left(1 - \sin^2 \Theta_W\right) = 1 + \Delta r \tag{2.40}$$

The term Δr is sensitive to the existence of heavy particles in the Standard Model, in particular

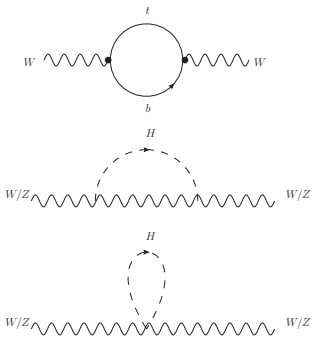


Figure 2.5: The masses of top quark, W and Higgs boson are related via radiative loop corrections.

the top quark and the Higgs boson [17]. These radiative correction Δr can be written as:

$$\Delta r = \frac{3G_F}{8\pi^2\sqrt{2}} m_t^2 + \frac{\sqrt{2}G_F}{16\pi^2} m_t^2 \left[\frac{11}{3} \ln\left(\frac{m_H^2}{m_W^2}\right) + \cdots \right] + \cdots \tag{2.41}$$

Hence one can deduce the Higgs boson mass from precision measurements of the top and W mass, assuming the Standard Model is the correct theory to describe nature. These quantities have been measured in several experiments at LEP, SLC and Tevatron as shown in Fig. 2.6

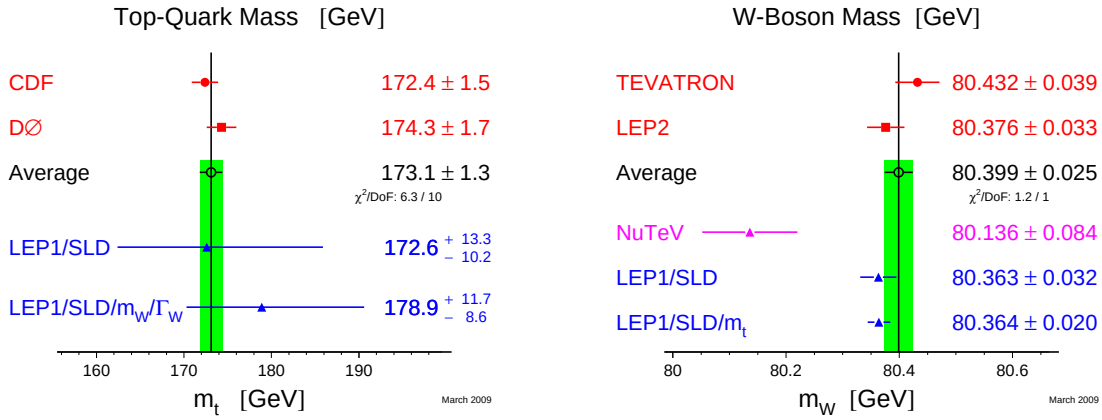


Figure 2.6: World average and individual measurements of the top (left) and W mass (right) [18] including both direct measurements of the top mass as well as indirect measurements via higher order corrections to the W boson mass.

The combined result between top and W mas is plotted for both indirect constraints and direct measurements in Fig. 2.7. Increased precision in the measured mass of the top quark and W masses have further reduced the uncertainty on these constraints over the last few years. Ultimately the direct comparison of the Higgs boson searches and the precision electroweak measurements is a powerful test of the self-consistency of the SM. Any deviation between the measured values for the top and W masses and the band predicted by the SM for different Higgs boson masses would indicate the need for new physics.

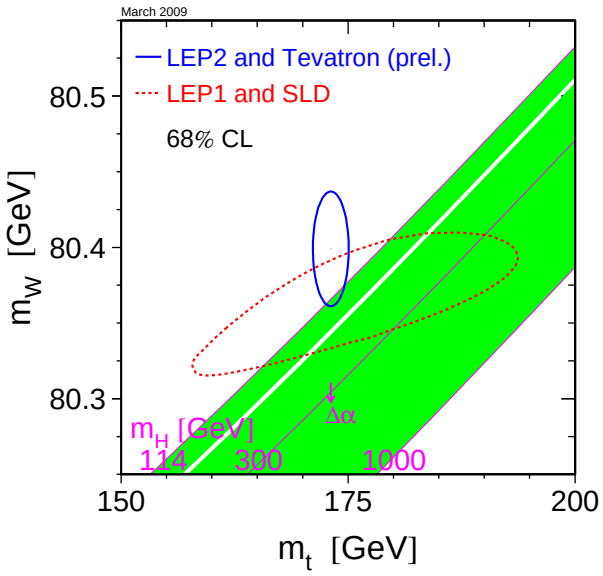


Figure 2.7: Combined results between m_t and m_W. The plots shows the 68% C.L. for the indirect measurements of LEP1 and SLD (dashed ellipses) and direct observations from LEP2 and Tevatron (solid ellipses). The shaded band shows the SM prediction based on the value for G_F for various values of the Higgs boson mass and fixed $\Delta\alpha^{(5)}_{had}(m_Z)$. Varying the hadronic vacuum polarization yields an additional uncertainty on the SM prediction indicated by the arrow labeled $\Delta\alpha$.

Figure 2.8 shows the constraints from the electroweak fits on a potential Standard Model Higgs boson mass as $\Delta\chi^2 = \chi^2 - \chi^2_{min}$ vs. m_H curve. The line is the results of the fit using high-Q^2 data while the band represents an estimate of the theoretical error due to missing higher order corrections. The vertical yellow bands show the direct exclusion limits by LEP2 and Tevatron. The dotted curve shows the corresponding fits using also low-Q^2 data and the dashed respectively solid curves various evaluations of the hadronic correction from light quark contributions to $\alpha(m_Z^2)$ [19]. The Higgs boson mass preferred by the fit is

$$m_H = 90^{+36}_{-27} \text{ GeV} \tag{2.42}$$

At 95% C.L. the Higgs boson mass is lower than 163 GeV [20]. This limit increases to 191 GeV when taking into account the LEP2 direct search limits. Summarizing the constraints of direct and indirect Higgs boson searches one obtains:

$$114.4 < m_H < 191 \text{ (GeV/c}^2\text{)} \quad \text{at 95\% C.L.} \tag{2.43}$$

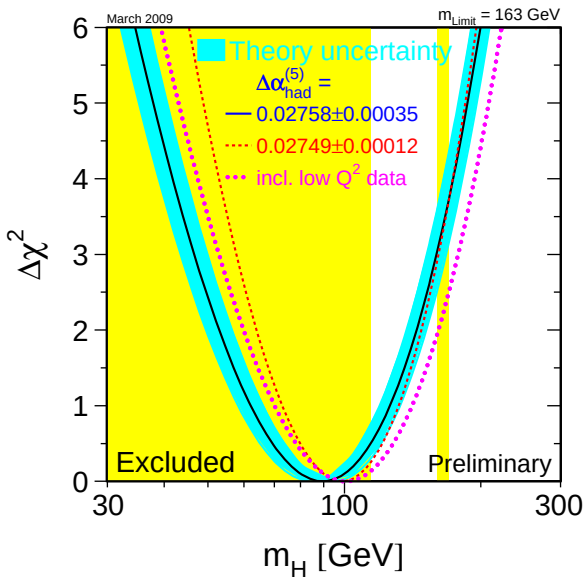


Figure 2.8: Constraint on the Higgs boson mass as $\Delta\chi^2 = \chi^2 - \chi^2_{min}$ vs m_H curve. The line is the results of the fit using high-Q^2 data whereas the band represents an estimate of the theoretical error due to missing higher order corrections. The vertical yellow bands show the direct exclusion limits by LEP2 and Tevatron. The dotted curves shows the corresponding fits using low-Q^2 data and the dashed respectively solid curves various evaluations of the hadronic correction from light quark contributions to $\alpha(m_Z^2)$ [19].

2.8 Higgs Boson Production & Decay at the Tevatron

2.8.1 Higgs Boson Production

The are four dominant production processes for Higgs bosons in $p\bar{p}$ collisions at the Tevatron, two Higgs-strahlung processes, namely associated WH and ZH production and gluon-gluon fusion and vector-boson fusion production. The Feynman diagrams of these processes are shown in Figs. 2.9 and 2.10. In contrary to the gluon fusion $gg \to H$, the W or Z boson produced in associated production can be used for easier identification.

The cross sections of these processes at the TEVATRON are shown in Fig. 2.11. The gluon fusion production has the largest cross section, ranging from 0.1 pb at $m_H = 200$ GeV to about 1.0 pb at $m_H = 115$ GeV. This if followed by the WH associated production with about ten times smaller cross sections but still about twice as large (depending on the Higgs boson mass) as the ZH and vector boson fusion (VBF) production processes. Details to the calculation of the production cross section is given in Sec. 6.3.2.

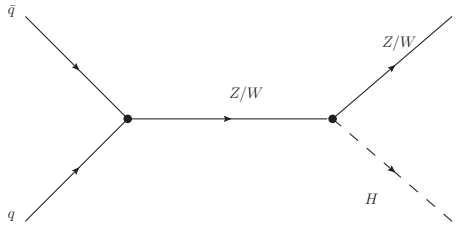


Figure 2.9: The two Higgs-strahlung channel at the Tevatron, the most dominant production channels for low mass Higgs boson searches

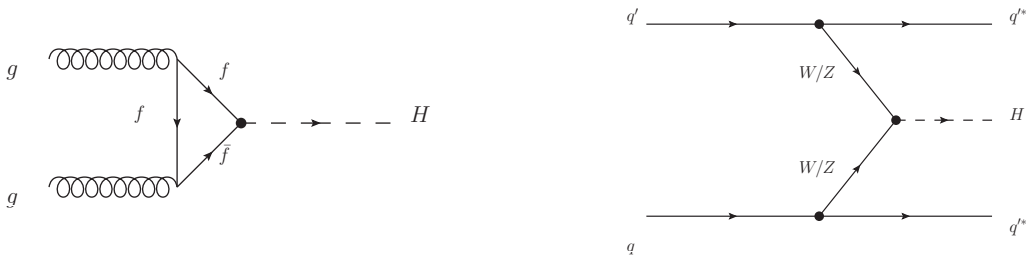


Figure 2.10: The gluon-fusion and vectorboson-fusion channel, the main production channels for this analysis. The gluon-fusion on the left is shown with the dominant and clean decays mode via to Ws into two leptons.

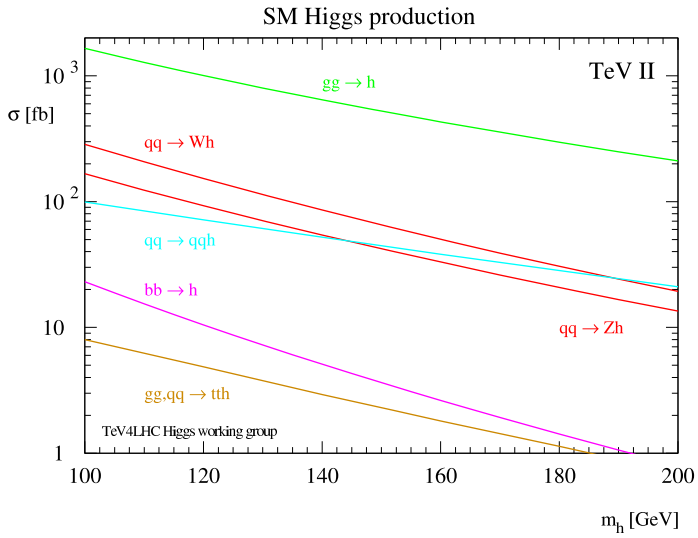


Figure 2.11: The main production for a Standard Model Higgs boson at the Tevatron [21]. All four leading production mechanisms, gluon fusion, associated production with a heavy gauge boson and vector boson fusion, plotted in green, red and cyan respectively, are considered in the present analysis.

2.8.2 Decays of the Higgs Boson

To conduct a promising Higgs boson search not only the production mode but as well the decay is of uttermost importance. The particular decay mode is dependent of the Higgs boson mass as shown in Fig. 2.12. For small masses below $m_H \simeq 140$ GeV the Higgs boson decays mainly into $b\bar{b}$ pairs, for larger massed the $H \to W^+W^-$ decay becomes dominant. One has to consider all possible decay channels to get optimal sensitivity. Given the production cross sections as shown in Fig. 2.11 ones expects ~ 4000 or more Higgs boson events to be recorded per detector at the TEVATRON with the current luminosity. Detector acceptance and selection efficiencies lead to the rejection of some of these Higgs boson events, another inefficiency is introduced by the fact that not all channels can be utilized for all possible Higgs boson masses. At low Higgs boson masses the dominant gluon-fusion channels cannot be distinguished from the overwhelming $b\bar{b}$ production with a cross section of 10 μb. Therefore one has to use the WH and ZH production channels because the production cross section of $Wb\bar{b}$ is $\simeq 30$ pb. For higher Higgs boson masses above $m_H \simeq 140$ GeV the background form multijet production can be reduced by using the decay $H \to W^+W^- \to \ell^+\nu\ell^-\nu$, making this the most promising channel. This channel has particularly good sensitivity for Higgs boson masses $m_H \sim 160$ GeV where the WW branching ratio becomes maximal.

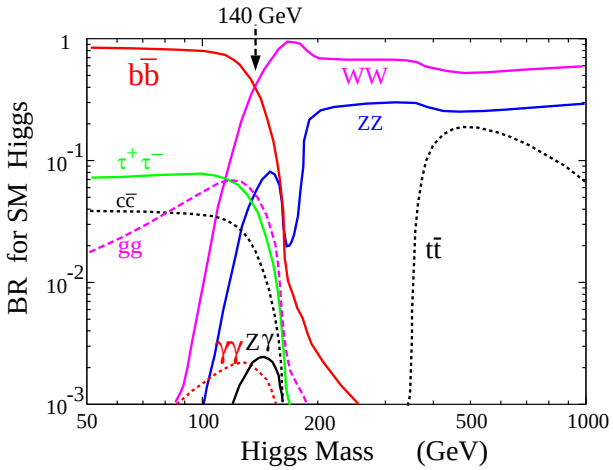


Figure 2.12: Branching ratios of the Higgs boson as function of the Higgs boson mass [22]. One sees that the Higgs boson decays dominantly to $b\bar{b}$ pairs for Higgs boson masses less than $m_H \simeq 140$ GeV and for higher masses the dominant decay takes place via two W bosons.

The Higgs boson decay branching ratio predictions are calculated with HDECAY [22]. The results are shown in Fig. 2.12. The most important search channels for $H \to WW$ involve decays in the $ee\nu\nu$, $e\mu\nu\nu$ and $\mu\mu\nu\nu$ final states which retain contributions of the tau branching modes such as $H \to WW \to e\tau\nu\nu$, $\mu\tau\nu\nu$ and $\tau\tau\nu\nu$. Initial Higgs boson sensitivity studies neglected hadronic tau decay modes [23]. In order to recover sensitivity these channels are now considered for Higgs boson searches by both experiments at the TEVATRON.

3
The DØ Experiment and the Tevatron

Collisions of particles at very high energies are necessary for the study of the production and properties of well established particles like quark, gluons and leptons and for the search for new particles. The collection of a sufficient numbers of these high energy collisions requires large experimental setups involving a system to accelerate and store the elementary particles (a collider) and one or more detectors which will identify and measure the particles produced in these collisions. The accelerator which currently reaches the highest center-of-mass energy ($\sqrt{s} = 1.92$ TeV) is the Tevatron , a machine located in Batavia, IL, about 40 miles west of Chicago.

Founded in November 1967 Fermilab has been home to many important discoveries in high energy physics. Among them was the discovery of the bottom quark 1977, observation of the top quark in 1995 and the tau neutrino in 2000.

3.1 The Tevatron

The Tevatron is the last stage of an accelerator chain for protons and anti-protons and is used to accelerate and collide these particles. The whole chain consists of a Cockcroft-Walton accelerator, a linear accelerator (LINAC), the Booster synchrotron ring, the Recycler, the Main Injector and finally the Tevatron. The antiprotons are produced in an separate complex. The sequence of the various elements in the accelerator chain is shown schematically in Fig. 3.1.

The acceleration process starts with the ionization of hydrogen atoms to H which are subsequently accelerated to energies of 750 keV using the Cockcroft-Walton accelerator. The H^- then reach the linear accelerator (LINAC) where they are further accelerated up to energies of 400 MeV. After being stripped of their electrons the remaining protons are accelerated to 8 GeV in the Booster and then injected in the Main Injector. Here the protons are accelerated to 120 GeV. A fraction of these nuclei is directed on a Ni-Cu target to create antiprotons. Using a lithium lens particles and anti-particles are separated, the anti-particles collected and then stored at an energy of 8 GeV in the accumulator When enough anti-protons have been accumulated ("stacking") both protons and anti-protons are accelerated again in the Main Injector and in the Recycler to 150 GeV before finally entering the Tevatron where both beams get separately accelerated to 0.98 TeV.

The particles are kept on their designated trajectory using 4.2 T magnetic fields during their acceleration in the 6.3 km storage ring. The bunches of particles are getting collided in two of the six possible interaction points at the Tevatron where the CDF and DØ detector are located. The main limiting factor for the luminosity, a measure of the frequency of beam's collisions, is the low efficiency for producing, collecting and storing the antiprotons. At the order of 10^5 protons

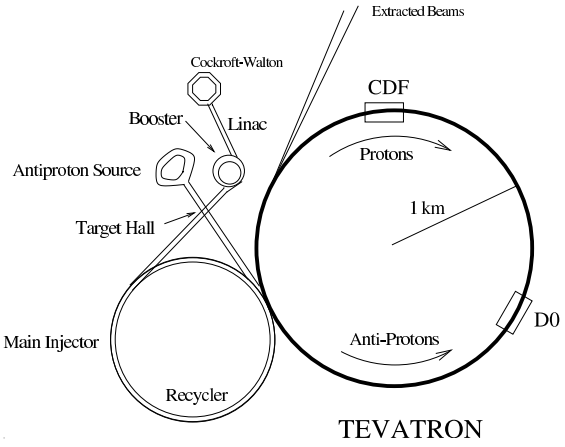


Figure 3.1: The accelerator chain with the Tevatron as the the last stage. In the Tevatron the proton and anti-proton beams are accelerated to 1.96 TeV.

are necessary to create one anti-proton. The performance of the Tevatron has been very good in the last few years, exceeding regularly the projections for the instantaneous and total luminosity integrated. The total delivered and recorded luminosity since beginning of Run II in April 2002 to now can be found in Fig. 3.2.

3.2 The DØ Detector

The DØ detector, named this way because it is located at the D0 section of the Tevatron, is a typical high p_T collider detector. Its main design criteria were precision measurement of electrons, jets and missing transverse energy. The detector consists out of several sub-detectors, a vertex detector, a tracking system, a liquid argon (LAr) calorimeter and a muon spectrometer. The subdetectors are read out by the data acquisition following a three level trigger system. The original detector design from Run I was improved in preparation for Run II. The improvement consists mainly out of a new tracking system embedded in a 2 T solenoid field newly installed for Run II. During a long shutdown in spring 2007 the tracking system was enhanced by the addition of a new silicon microstrip layer placed outside as innermost layer outside the beam pipe. This compensates for the loss of efficiency of the other Si detectors caused by radiation damage.

The DØ detector uses a right-handed coordinate system. The positive z axis points along the proton beam, the positive x-axis points horizontally inside the Tevatron ring, and the y-axis points upwards. The transverse plane is defined by the x and y-axis. A more useful coordinate system than the Cartesian (x, y, z) coordinates is the cylindrical coordinate system with (r, ϕ, η), with

$$r = \sqrt{x^2 + y^2}$$

$$\phi = \tan \frac{x}{y}$$

The pseudo rapidity η is a quantity frequently used at hadron colliders instead of the polar angle

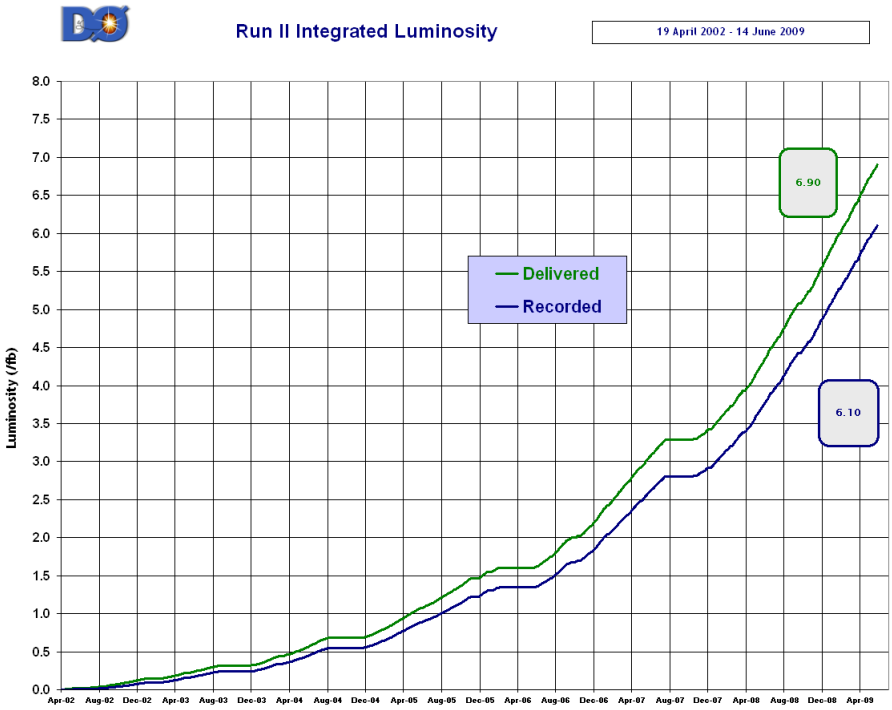


Figure 3.2: Total luminosity delivered to DØ by the Tevatron. The green line gives the total delivered luminosity whereas the blue line gives the recorded luminosity. The plateaus correspond to down times and scheduled shutdowns.

θ:

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

.

Figure 3.3 shows a cut-away of the DØ detector. A short overview of several sub-systems important for this thesis will be given in the following sections.

3.2.1 The Tracking System

The tracking system measures the momentum, sign and direction of the charged particles produced in the collisions. A particle with charge q and a momentum of p will follow a helix-shaped trajectory in the solenoidal field The radius r of this helix is given by

$$r[m] = \frac{p_T[GeV]}{0.3 \cdot B[T]} \quad \text{with} \quad p_T = \sqrt{p_x^2 + p_y^2}, \tag{3.1}$$

where B is the strength of the magnetic field. The tracking system surrounds directly the interaction point and is the first subdetector transversed by particles produced in the collisions. It consists of the silicon microvertex tracker (SMT), and of the central fiber tracker (CFT) around the SMT. The whole tracking system is surrounded by a solenoidal superconducting magnet providing

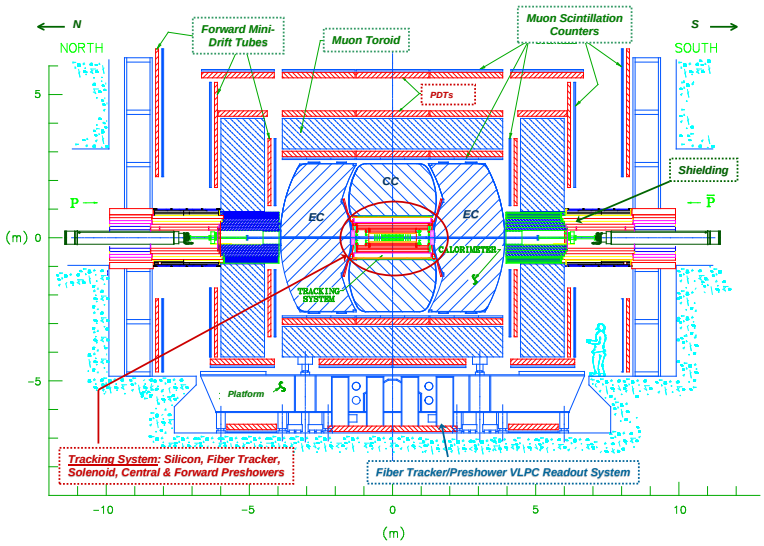


Figure 3.3: Schematics of the DØ detector and it's various sub-detectors [24]. The upgrades for Run II are circled in red.

field of 2 T. A schematic view of the tracking system is shown in Fig. 3.4. Figure 3.6 shows the $\vec{B}$ field produced by the solenoid and toroid magnets. The tracking system has a momentum resolution of

$$\frac{\delta p_T}{\text{GeV}} = 0.002 \frac{p_T}{\text{GeV}}^2 \tag{3.2}$$

Since this resolution is better than that of the muon system, the tracking system is also important for the measurement of muons.

The transverse momentum of the particle is measured by the curvature of the track in the $r-\phi$ plane. The direction of the track in the $r-z$ plane completes the 3d momentum measurement of the particle. Another important feature of the SMT is the identification of particles created at secondary vertices rather than in the primary vertex. This is used to reconstruct the decay vertex of long-lived particles such as hadronically decaying b quarks ("b-tagging"). This is an important tool for selecting final states with b-quarks and commonly used in top physics or low mass Higgs searches.

The Silicon Microvertex Tracker

Figure 3.4 shows the Silicon Microvertex Detector (SMT), the innermost detector closest to the beam and interaction point. Its main purpose is the detection of primary and secondary vertices. The readout modules of the SMT are arranged in 6 barrels and 16 discs. The main purpose of the barrel detector is to measure the $r-\phi$ coordinate, which is used to determine the transverse momentum of the track. The discs improve the resolution for particles emitted at small angles relative to the beam direction. Each barrel has four layers of silicon readouts. The silicon modules in the barrels are called ladders. The first and second layer consist of 12 ladders respectively, while the third and fourth layer have 24 ladders each. Each layer consists of two sub-layers. The central

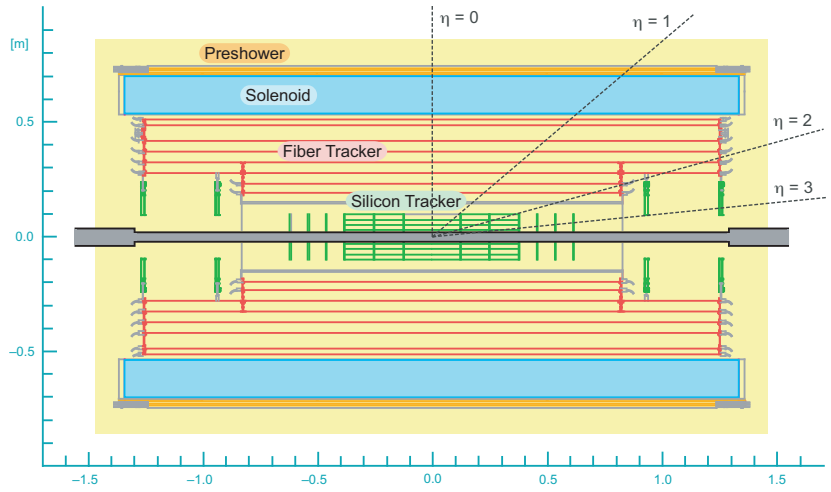


Figure 3.4: Schematic view of the central tracking system.

four barrels use double-sided double-metal (DSDM) detectors in layer 1 and 3. Layer 2 and 4 in all barrels use a single-sided (SS) technology. In the outermost barrels double-sided (DS) detectors are used. The DSDM sensors consist of axial and stereo strips with a stereo angle of 2°. Stereo layers provide the y direction, the SS detectors are axial only. The discs that intersperse the barrels are the F-discs that consist of twelve double-sided wedge-detectors. In the forward region two H-discs are placed providing tracking information for high $|\eta|$ up to $|\eta| = 3$. The H-discs consist of 24 wedges which are made of two back-to-back single-sided wedges. The discs are planar modules. The double-sided F-wedges have an effective stereo angle of 30°. The two single sided detectors of the H-wedges form together a double-sided sensor that provides an effective stereo angle of 15°
In 2006 a new layer, called layer 0, was installed in the SMT [25]. It is now the layer closest to the interaction point and fits inside layer 1. The vertex resolution is improved by use of Layer 0.

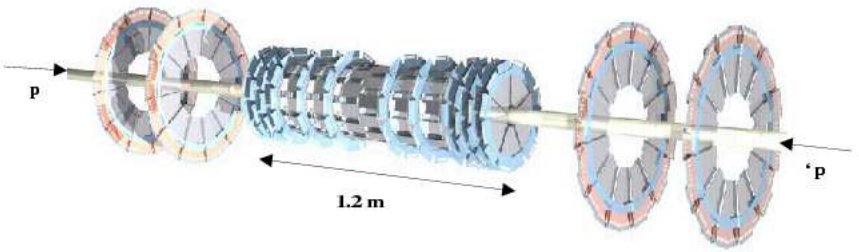


Figure 3.5: The silicon microvertex detector [24].

The central fiber tracker

The central fiber tracker (CFT) surround the SMT. The CFT covers the region $|\eta| < 1.62$ and consists of 76800 scintillating fibers mounted on eight concentric cylinders. The two innermost cylinders have a length of 1.66 m, the six outer cylinders are 2.52 m long. Every cylinder consists of a doublet layer of fibers in axial direction (axial layer) and a doublet layer of fibers with a stereo angle in ϕ of $\pm 3^\circ$ (stereo layer). The stereo layers on the first, third, fifth and seventh cylinders are oriented in $+3^\circ$ direction, the stereo layers on the other cylinders in the -3°. The scintillating fibers are coupled to clear fiber waveguides which conduct the light to visible light photon counters (VLPCs) where the light signal is converted into an electric signal and is read out. Each fiber has a diameter of 835 μm and is 1.66m or 2.52m long. They provide a cluster resolution of about 100 μm Only one end of the scintillating fibers is connected to a waveguide. The opposite end is mirrored with sputtered aluminum coating, which provides reflectivity of about 90%. A minimum ionizing particle produces about ten photoelectrons in a fiber. The VLCP have a quantum efficiency of 75% and a high gain. They convert each photon to 22000 to 65000 electrons. The photon counters are located in a liquid hydrogen cryostat and operated at 9K.

The solenoid magnet

The momentum of a charged particle is determined by measuring the curvature of it's track within the 2 Tesla magnetic field. This field is created by a superconducting solenoid magnet of 2.7 m length. The magnet consists out of two layers with an average radius of 60 cm. The energy contained by the magnetic field corresponds to 5 MJ. The operating temperature of the solenoid magnet is 10K. The magnetic field for each charged particle is uniform to 0.5 % in the tracking volume. The material of the solenoid and the cryostat wall corresponds to 1.1 X_0 [1]. The magnetic field lines are shown in Fig. 3.6

Performance of Track Reconstruction

Hits registered in both tracking subsystems are combined to form a reconstructed track. The momentum resolution of the tracker for a minimum ionizing particle can be parametrized as follows:

$$\sigma(p^{-1}) = \frac{\sqrt{(S \cdot \sqrt{\cosh \eta})^2) + (C \cdot p_T)^2}}{p}, \tag{3.3}$$

where p corresponds to the particle momentum, η to the pseudo-rapidity and S is a term describing multiple scattering. C gives the resolution of the system.

3.2.2 The Calorimeter

The purpose of the calorimeter is to identify and to measure the energy of electrons, photons, jets and determine whether there is an energy imbalance in the event indicating the production of

[1] The radiation length X_0 is a characteristic of a material, related to the energy loss of high energy, electromagnetic-interacting particles with it. It is both the mean distance over which a high-energy electron loses all but $1/e$ of its energy by bremsstrahlung, and $7/9$ of the mean free path for pair production by a high-energy photon.

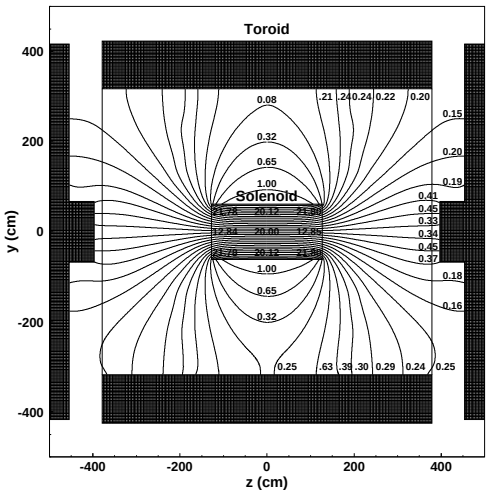


Figure 3.6: The field of the solenoid and toroid magnet [26].

particles which do not interact in the detector. The calorimeter consists of three parts: the central calorimeter (CC) and the two endcap calorimeters (EC) on the north (ECN) and south (ECS) sides of the detector, as shown in Fig. 3.7. The CC covers the range $|\eta| < 1.1$ while the ECN and ECS cover the range from $1.4 < |\eta| < 4$. The gap between $1.1 < |\eta| < 1.3$ is covered by the Inter Cryostat Detectors (ICD), described in Sec. 3.2.2. Each of the three calorimeter parts shows a radial segmentation in four electromagnetic layers (EM1-4), forming the innermost calorimeter section, surrounded by three fine and one coarse hadronic layers. The DØ calorimeter is a sampling calorimeter using liquid argon as the sensitive medium. The absorber plates are made of from depleted uranium with a thickness of 3 (4) mm in the CC (EC) regions of the EM calorimeter. The coarse hadronic modules contain thicker absorber plates, made out of 46.5 mm of copper in the CC and stainless steel in the EC. An electric field of 2.0 keV is applied to the 2.3 mm wide gap between the absorber plates. This field causes the collection of the charges created by charged particles traversing the liquid argon, with a collection time of 450 ns.

Figure 3.8 shows a schematic of a calorimeter cell. Figure 3.9 shows the coverage of the calorimeter as well as the segmentation of the absorber plates.

Calorimeter readout cells form projective towers as also shown in Fig. 3.9. Each tower is further divided into several longitudinal "depths". Those towers are called "pseudo-projective" since the cell boundaries are perpendicular to the boundaries of the absorber plates whereas the center of cells of increasing shower depths lie on rays projecting from the interaction region. The modules are arranged to form 32 wedge-shaped modules. The segmentation of the fine, the coarse hadronic and the first, second and fourth layer of the electromagnetic calorimeter in the $\eta - \phi$ space is $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$. This is comparable to the transverse size of showers. The segmentation of the third layer of the EM calorimeter, supposed to be located close to the maximum of the shower development, is twice as fine in both η and ϕ to ensure a more precise determination of the position of the EM shower. At large η cell sizes increase in η and ϕ to avoid very small cells. In the central region the EM and HAD sections compensate approximately about $20.0X_0$ and about $21.6X_0$ in

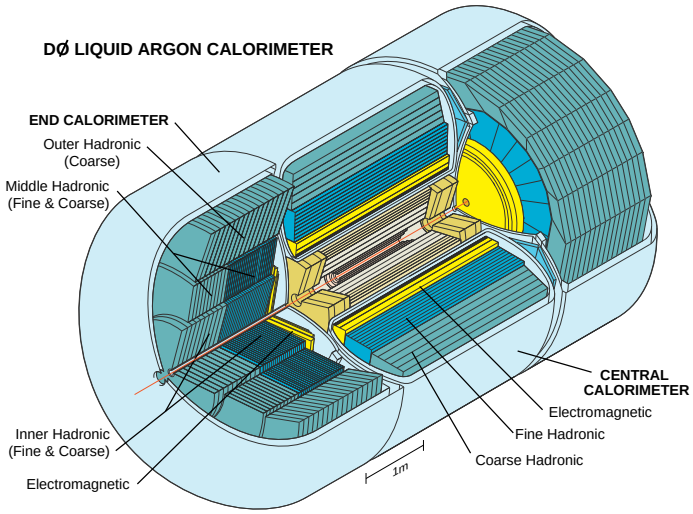


Figure 3.7: Cut-away of the DØ calorimeter [24].

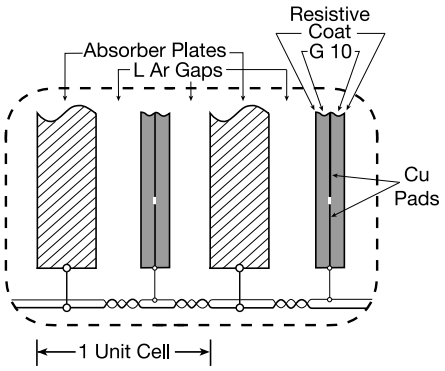


Figure 3.8: Schematic of a calorimeter cell

the endcaps. The various detector components between the interaction region and the the first active liquid argon corresponds to about $\sim 4.0X_0$ at $\eta = 0$ and $\sim 4.4X_0$ at $\eta = 2$. Electromagnetic particles are completely absorbed in the EM calorimeter. In contrast hadronic jets start to shower within the EM, but the showering mainly takes place in the fine and coarse hadronic calorimeter. This enables us to distinguish electromagnetic and hadronic objects. The width of the shower can be used to distinguish further electromagnetic and hadronic objects as well.

Muons loose only a small faction of their energy through ionization and therefore penetrate the entire detector leaving only a small signal in the calorimeter. Particles likes neutrinos which do not interact with via the electromagnetic or hadronic interaction have a negligible probability of loosing energy in the calorimeter. The absence of an energy deposition from these particles causes an imbalance in the energy flow in the transverse plane. This imbalance is called missing transverse energy ($\not\!\!E_{\mathrm{T}}$) and is the experimental signature of neutrinos and other new interacting particles.

The relative uncertainty on the energy of electromagnetic and hadronic objects in the calorimeter

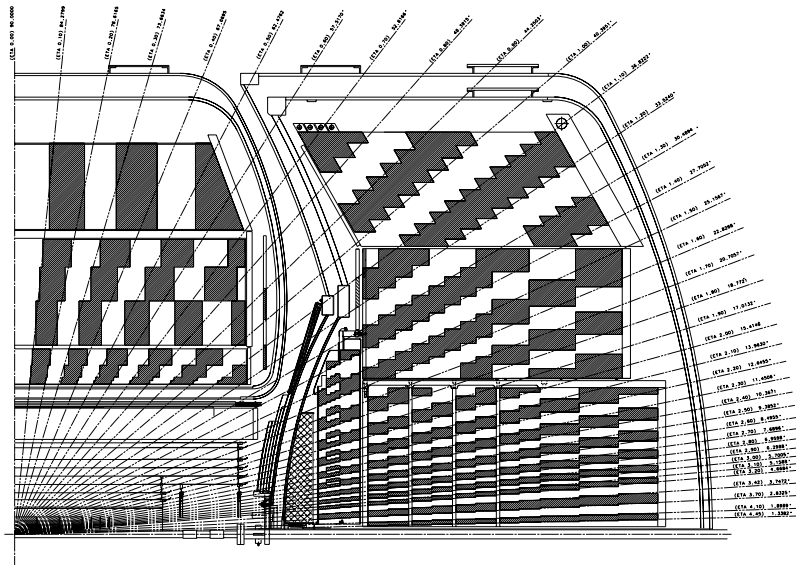


Figure 3.9: Cut-away of the DØ calorimeter [24] showing segmentation of the calorimeter, the pseudo-projective towers and the separation in a central and end-cap part.

can be parametrized as

$$\frac{\Delta E}{E} = \sqrt{\frac{S^2}{E/\text{GeV} + \frac{N^2}{E^2/\text{GeV}^2} + C^2}} \tag{3.4}$$

where N describes instrumental effects like uranium noise and pedestal subtraction, S describes fluctuations in the deposition of energy and C represents a constant term in the resolution which can be attributed to uncertainties in the calibration and non-uniformities in the detector response. The parameters N, C and S measured from data are listed in Table 3.1. In contrast to the tracking system the energy resolution improves with increasing energy depositions.

Object	C	S [$\sqrt{\text{GeV}}$]	N [$\sqrt{\text{GeV}}$]
Electrons, Photons	0.041	0.15	0.29
Jets	0.036	1.05	2.13

Table 3.1: Energy resolution parameters [27, 28].

The Calorimeter Read-Out Electronics

The readout chain of the calorimeter data is shown in Fig. 3.10. Each of the 47032 sensitive cells of the calorimeter is read out via electronic chain of three staged. The readout takes place in three

stages:

- Signals from the detector are first transported to charge preamplifiers located on the cryostats.
- Signal from the preamplifiers are then transported to the signal shaping and storage circuits on baseline subtractor boards (BLS).
- The signals from the BLSs are transmitted on an analog bus to the analog digital converter (ADC).

From the ADCs the signal enters the data acquisition system for the Level 3 trigger decision and storage on tape. To account for the decreased bunch spacing in Run II both preamplifiers and BLSs were completely rebuilt.

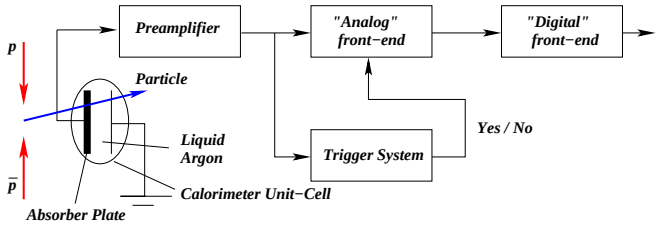


Figure 3.10: Data flow of the calorimeter read out.

The Inter-Cryostat-Detector and Massless Gaps

Due to the separation of the calorimeter in central part and endcaps the coverage between $0.8 < |\eta| < 1.4$ is incomplete. Additional layers of sampling material have been added in this region to recover the energy deposited in the cryostat walls which are not instrumented. These are the Massless Gaps (MG) and the Inter-Cryostat Detector (ICD). The central massless gaps are positioned inside the central cryostat after the stainless steel end plates of the modules and cover the region of $0.7 < \eta < 1.7$. The endcap massless gaps are positioned after the endcap cryostat walls and cover the region $0.7 < \eta < 1.4$. The Massless Gaps are divided in two separate layers. The Massless Gap in layer 8 covers a range in η of $0.7 < \eta < 1.4$ and is referred to as the Endcap Massless Gap (ECMG). Layer 10 covers $0.7 < \eta < 1.3$ and is referred to as Central Calorimeter Massless Gap (CCMG). The MG cells are constructed in the same way as the regular calorimeter cells. In the ICD, which covers the region $1.1 < |\eta| < 1.4$ the sensitive medium is instead made out of 0.5" thick scintillating tiles, each covering $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ for a total of 378 channels.

3.2.3 The Preshower Detector

The preshower detector is arranged around the solenoid. It consists of the central preshower detector (CPS) covering the region $|\eta| < 1.3$ and the forward preshower detector (FPS) covering $1.5 < |\eta| < 2.5$. The purpose of the preshower detectors is to improve the electron and photon identification as well as the background rejection. The CPS is located around a 5.6 mm lead

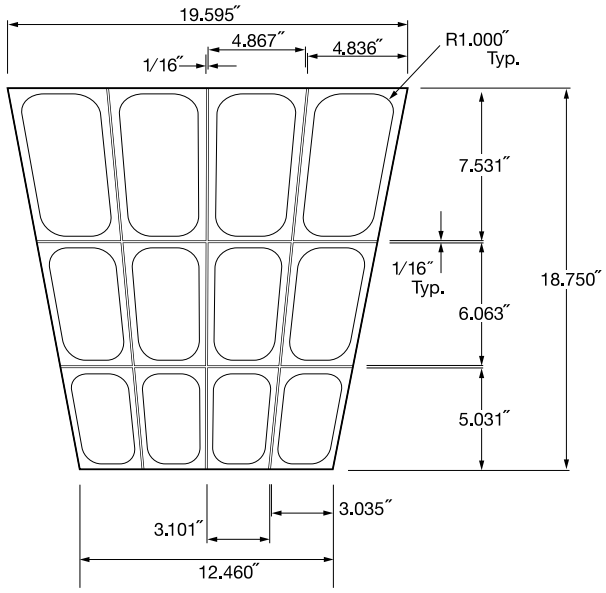


Figure 3.11: An ICD tile with it's 12 subtiles [26].

radiator, which corresponds to about one radiation length. Together with the $0.9X_0$ thick solenoid it forms at least two radiation length of material, increasing to four radiation lengths depending on the angle of the entering particle. The CPS consists of three layers of scintillator strips, each layer consisting of 1280 strips. Each strip is read out by two wavelength-shifting (WLS) fibers. The two FPSs on the north and south side of the detector are mounted on the calorimeter cryostats. Both detectors consist of two layers of double planes of scintillators, separated by a 11 mm ($2X_0$) thick lead-stainless-steel absorber. The innermost layer is called MIP [2] layer, the outermost layer is called shower layer. In the MIP layer charged particles passing through the detector deposit minimum ionizing signals. In the shower layer charged particles and photons produce a shower signal.

3.2.4 The Muon System

The outermost system of the detector is the muon system. Its purpose is to identify and measure muons. In order to have a stand-alone muon-system momentum measurement, a toroidal magnet with a magnetic field of 1.8 T is located within the muon system. The magnetic field lines are shown in Fig. 3.6. The strategy of muon identification is based in the fact that muons are not subject to hadronic interactions. They are losing only energy by ionization. The thickness of the calorimeter and of the toroid ensure that all hadrons are absorbed. Therefore each charged particle reaching the muon system will be identified as a muon.

The muon system consists of two parts: The wide angle muon system or central muon system (WAMUS) and the forward muon system (FAMUS). The central muon system covers the range

[2] minimum ionizing particle (MIP): particle whose mean energy loss rate through matter is close to the minimum. When a fast charged particle passes through matter, it ionizes or excites the atoms or molecules that it encounters, losing energy in small steps. Most relativistic particles (e.g., cosmic-ray muons) are minimum ionizing particles.

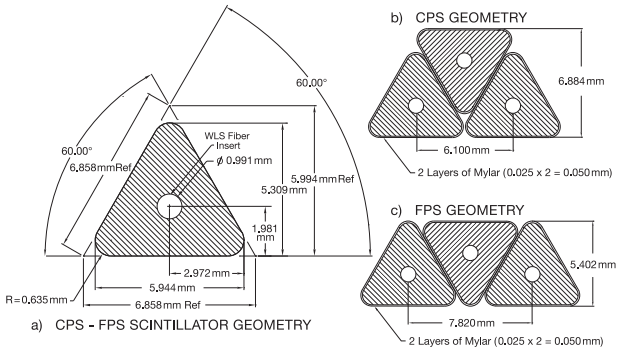


Figure 3.12: Schematics of the Front Preshower Detector (FPS) [26].

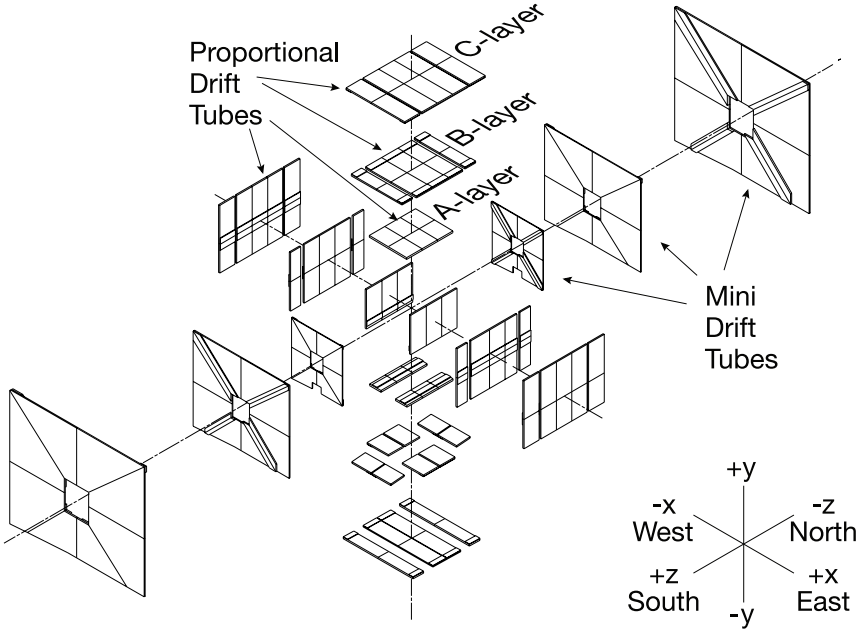


Figure 3.13: The Muon System [26].

up to $|\eta| \simeq 1$ and uses proportional drift tubes (PDTs) whereas the forward muon system covers $1 \leq |\eta| \leq 2$ with mini drift tubes (MDTs). Scintillating counters are used as for trigger purposes in both systems. Three layer of muon system detectors are present in each system. One before the toroidal magnet (A-Layer) and two placed outside the toroid (B- and C-Layers).

A view of the PDTs and MDTs in the central and forward muon system is shown in Fig. 3.13. Figure 3.14 shows the positions of the trigger scintillation counters within the muon system.

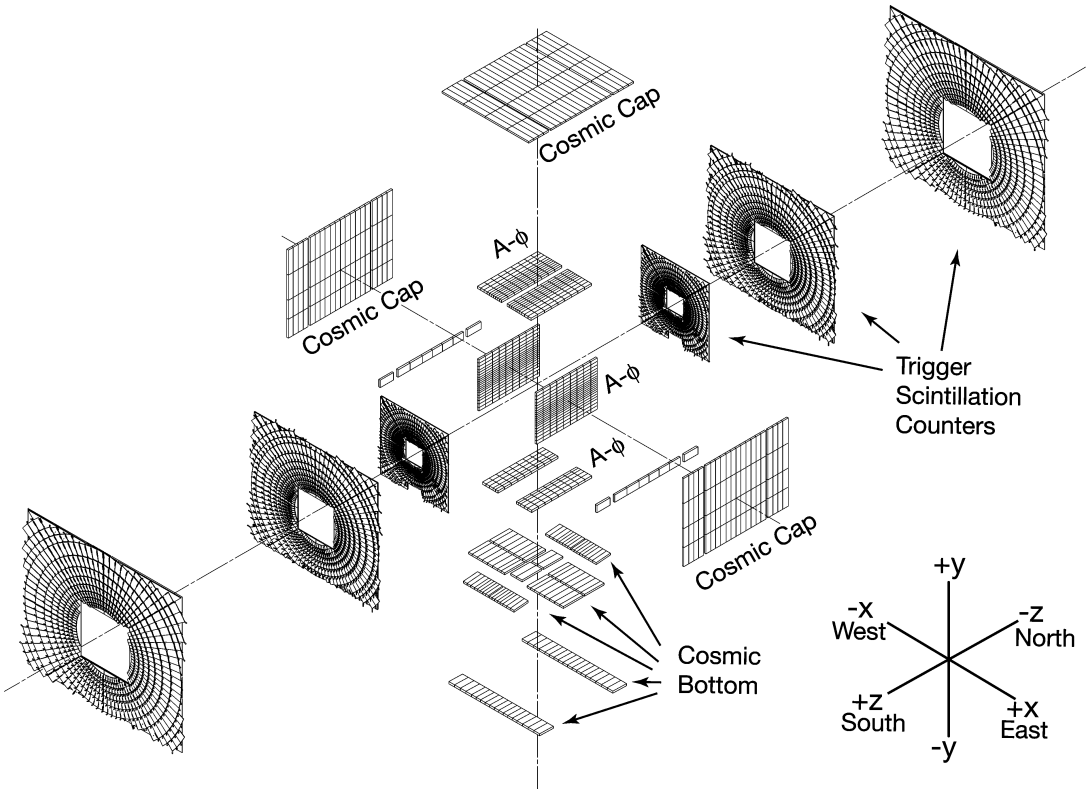


Figure 3.14: The Muon System [26]

3.2.5 The Toroid Magnet

The toroid magnet enables a momentum measurement of the muons. The magnet is located within the first and second layer of the muon system, in about 318 cm distance of the beamline. In order to access the central detectors the toroid was constructed in three sections. The center-bottom section is fixed where the the two out C-shaped sections can be moved perpendicular to the beamline. The magnet is wound using twenty coils of ten turns each. The magnets are operated at current of 1500A and the magnetic field is about 2 T.

3.2.6 The Luminosity System

In order to measure the Tevatron luminosity at the DØ interaction point, luminosity monitors (LM) are located at $|z| = 140$, as shown in Fig. 3.15. The LMs detect inelastic $p\bar{p}$ collisions. Each LM consists of an array of 24 plastic scintillation counters. Each counter is 15 cm long and covers the rapidity $2.7 < |\eta| < 4.4$. The timing resolution of the scintillators is about 0.3 ns. The high timing resolution enables the discrimination between particles coming from the interaction region and the beam halos.

The inelastic proton-antiproton counting rate is used to determine the instantaneous luminosity

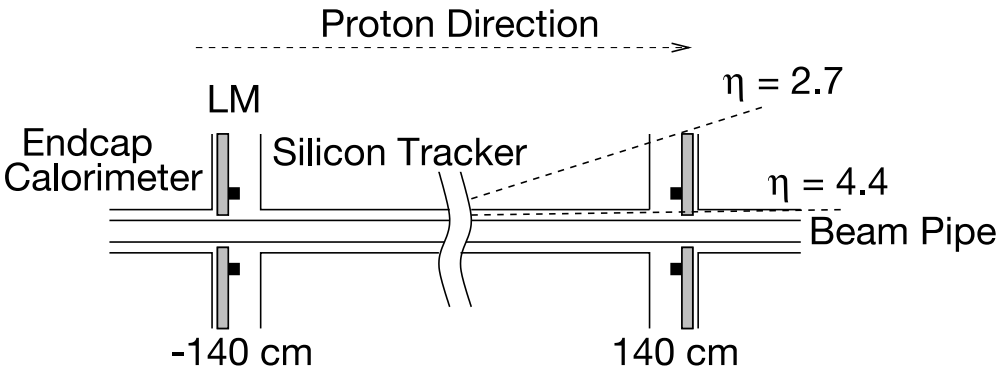


Figure 3.15: Location of the luminosity monitors on the z-axis [26].

[29], [30], [31]

$$\mathcal{L} = \frac{1}{\sigma_{eff}} \frac{dN}{dt} \tag{3.5}$$

where σ_{eff} is the effective inelastic cross section measured by the LM. The effective cross section is derived from the inelastic cross section [32] $\sigma_{inelastic}(1.96 \text{ TeV}) = 60.7 \pm 2.4$mb, taking into account acceptance effects and the efficiency of the LM detector. In order to properly distinguish $p\bar{p}$ interactions from beam halo interactions the z coordinate of the interaction vertex is calculated from the difference in time-of-flight between the north and south part of the LM. Beam halo particles have a larger time-of-flight difference than inelastic $p\bar{p}$ collisions.

The integrated luminosity is calculated in luminosity blocks. Each luminosity block, which builds the fundamental unit of time for the luminosity measurement, is indexed by a luminosity block number (LBN). After each run or store transition or after 60 seconds the LBN monotonically increases. The time period is chosen such that for each LBN the integrated luminosity is approximately constant.

3.2.7 The Trigger and DAQ-System

Due to the bunch crossing rate of 369 ns there are 1.7 million collisions per second. The events from this crossings are largely filled with soft scattering processes and noise. Only a small fraction of events has the large momentum transfer necessary to create events interesting for analysis. The DAQ system and the offline reconstruction capabilities limit the rate of events which can be recorded to an average rate of 100 Hz. In order to match these rates a filtering procedure is performed in a three tiered fashion (see also Ref. [26]). An overview over the DØ trigger and DAQ system realizing the data filtering is shown in Fig. 3.16.

The trigger system is very closely integrated with the data read out. Each event passing the first two trigger stages gets fully digitized before being sent to the L3 farm. Here a single farm nodes processes all data block associated with one data event.

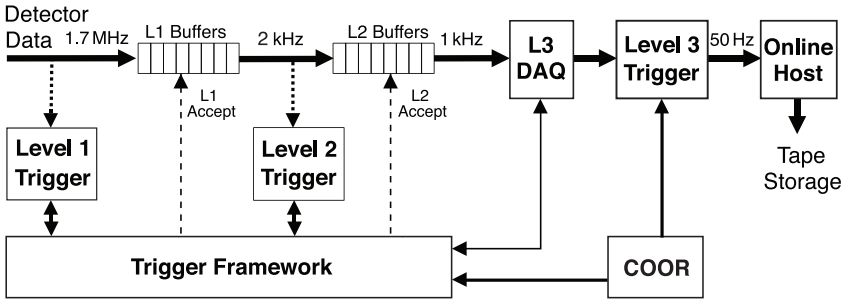


Figure 3.16: The DØ trigger and data acquisition system [26].

The L1 Trigger

The L1 trigger is based on specialized hardware. The most important components are:

- **Calorimeter Trigger** `L1Cal`: Searching for calorimeter trigger exceeding designated values of the transverse energy.
- **Central Track Trigger** `L1CTT`: Searching for tracks exceeding designated momentum thresholds, partly in combination with the
- **Muon System Trigger** `L1Muon`: Searching for muon tracks with high transverse momentum
- **Forward Proton Trigger** `L1FPD`: Selects diffractive produced events by triggering on protons or antiprotons scattered at very small angles

The rate of the L1 trigger accepts is limited by the maximum readout rates of the participating subsystems and by a desire to minimize the dead time associated with the readout.

The L2 Trigger

The L2 trigger consists of two stages. In the preprocessor stage the information from the subsystems are collected and the data are analyzed to form physics objects. The preprocessor stage consists of the Level 2 calorimeter trigger (`L2CAL`), the preshower detector trigger (`L2PS`), the muon system trigger (`L2MUO`), the SMT trigger (`L2STT`) and the Level 2 central track trigger (`L2CTT`). These various preprocessors collect data from the front-ends and L1 trigger system and analyze these data to form physics objects. In the global stage (`L2Global`) the data from across the subsystems is combined and physics objects are formed. The decision whether an event is kept or not is made within 100μs at Level 2. The L2 trigger reduces the data rate to about 1 kHz.

The L3 Trigger

The final L3 trigger enriches further the physics samples while maintaining an acceptable output which can be written to tape. Using a fully programmable software trigger the L3 farm performs a limited reconstruction of events and considering as well the relations among physical objects (e.g. the azimuthal angle or invariant mass between two objects).

3.2.8 Data Acquisition

Due to decreasing instantaneous luminosity during a store the amount of interactions is not constant but decreases as well. In order to maintain a constant readout rate of 50Hz so called prescales are used. A prescale allows only a fraction of $1/n$ of a particular type of event to pass the L1 trigger. A store, typically lasting about 8-12 hours, is divided into individual runs of 2 hours at high luminosity and about 4 hours towards the middle and end of the store. At each run transition the prescales are adjusted according to the present luminosity. Events passing all 3 trigger stages are getting transferred from the primary readout crates to the farm nodes by the primary data acquisition system (L3DAQ). Both, triggering and data acquisition are controlled by a coordination program (COOR).

3.2.9 Detector Upgrade

An upgrade of the DØ detector was performed during a long shut-down of the Tevatron in 2007. The upgrade was primarily motivated by the higher instantaneous luminosities after the shutdown and consisted of an upgrade of the trigger system and the installation of an additional silicon microstrip layer of tracking ("Layer 0") close to the beam-pipe to compensate for radiation damages in the existing system to improve impact parameters and lifetime measurements.

4
Object Identification

The events recorded by the DØ sub-detectors consist mainly of digitized electronic signals resulting from the collected charge of calorimeter cells, light yield of scintillators, hits of the tracking system etc. This format is not convenient for data analysis. Instead the raw data are converted into another format, more suitable for analysis. During the reconstruction raw data are converted into energy and position information, applying the relevant calibrations/alignments. Then different particles generated in the collisions are identified and their kinematic properties are measured. Additional quantities used for the identification of different types of particles, like the missing transverse momentum or the isolation of particles relative to other objects in the calorimeter or in the tracker, are also calculated.

This analysis relies mostly on the selection of leptons - electrons, muons and taus - from the data sample and also on the reconstruction of the missing transverse momentum due to the neutrinos escaping the detector. The following section will summarize briefly the identification of tracks, vertices, leptons, jets and missing transverse momentum ($\not\!\!E_{\mathrm{T}}$).

4.1 The RECO Program

The DØ offline reconstruction program [33] can process both, raw data written by the data acquisition system after a successful L3 trigger and also the result of Monte Carlo (MC) simulations of the collisions for different types of physics processes. The output of the reconstruction program [34] is a standardized format with well defined reconstructed objects. In the first step of the reconstruction chain, the reconstruction program unpacks the digitized signals from the detector, associates the electronic channels with physical detector elements and applies detector specific calibration constants. The second step requires most computing power and consists of reconstructing tracks using the hits in the SMT and CFT detectors. These tracks are stored and used later as input to the third level in the reconstruction, vertexing. In this third step, primary vertices (associated with $p\bar{p}$ collisions) and secondary vertices (stemming from the decay of long-lived particles) are identified and stored. The last and final step uses the information of dedicated sub-detectors to reconstruct and store high-level object such as electromagnetic particles (e.g. electrons, photons) muons, taus, neutrinos ($\not\!\!E_{\mathrm{T}}$) and jet candidates.

4.2 Track Reconstruction

Any charged particle moving in a solenoidal magnetic field describes a curved trajectory called *track*. When transversing through detector elements the particles interact with matter by ionization and leave hits in the CFT and SMT. These hits are spatially clustered. These clusters are provided to the reconstruction algorithms [35], the Histogram Track Finder (HTF) [36] and the Alternative Algorithm [37] which run sequentially. The AA algorithm shows generally a better performance and lower fake rate for low p_T and high impact parameters. It constructs a large pool of track hypotheses by extending seed clusters of tracking hits from the SMT to the rest of the tracking system. It filters down the number of track candidates based on well defined criteria and eliminates all overlapping hypotheses until no more tracks remain in the pool. The HTF algorithm on the other side is more efficient for high p_T tracks. This algorithm performs the track-finding by filling a histogram in the track parameter space (track curvature and azimuthal angle) with values consistent with each hit in the CFT and the SMT. Hits from the same particle will produce a peak in the histogram, contrary to random hits.

The AA algorithm

The Alternative Algorithm (AA) forms an initial track hypothesis by using combinations of hits in the SMT barrel or disk. Starting from any combination of three hits in the SMT barrels or disks, the algorithm extrapolates the sequence of hits moving outwards to the next SMT or CFT layer. Hits found within the expected region are associated to the track hypothesis if they match certain conditions such as the axial angle between the different hits or the curvature. If a hit is the quality of the fit of the track candidate has to satisfy a given χ^2 threshold. In case of multiple tracks passing these conditions the hypothesis is split and a new track candidate is formed for each valid hit combination. A "miss" is recorded when no hit is found in the layer. Tracks with less than three hits in the SMT are reconstructed by using the primary vertex candidate which was found using reconstructed tracks of at least three hits in the SMT. Any three hits in the CFT are then required to fulfill the additional condition of the track hypothesis passing near a reconstructed primary vertex.

The HTF algorithm

The trajectory of a particle in a magnetic field can be characterized in a plane perpendicular to the direction of the field by three parameters: ρ, the radius of the curvature, $d0$, the distance of closest approach (DCA) with respect to $(0,0)$, and ϕ, the azimuthal angle of the track at the point of closest approach to $(0,0)$. For track candidates with small impact parameters, every pair of hits in x and y that belongs to the same track corresponds to a single point in the $\rho - \phi$ plane. Filling each pair of hits into the 2-dimensional $\rho - \phi$ histogram, a peak in the histogram would correspond to a track candidate.

A final list of tracks is generated by merging the output of both algorithms and removing duplicates. The final track list is sorted by the number of hits, fewest misses and lowest χ^2 value.

4.3 Vertex Reconstruction

Interaction and decay of particles results in a set of tracks coming from one vertex. The primary vertex (PV) represents the point of the hard interaction. A precise reconstruction of the PV is essential to distinguish the objects from the hard interaction from overlapping events, to calculate the missing transverse energy and to distinguish tracks from the primary and possible secondary vertex (SV). Proper identification and precise measurement of secondary vertices is particularly important for identifying the flavor of jets and selecting the contribution of jets originating from heavy quarks. The algorithm can be distinguished in three basic steps:

At first tracks identified with $p_T > 0.5$ GeV and at least 2 hits in the fiducial SMT region are clustered along the beam axis. This allows the separation of possible additional $p\bar{p}$ interactions taking place during the bunch crossing. Then a two step approach is applied for each of the track clusters. First an estimation of position and width of the beam is performed by fitting all the tracks in the cluster into a common vertex using a Kalman Filter vertex fitting algorithm. Then a preselection on the tracks corresponding to each cluster is performed based on their distance of closest approach to the beam spot. Subsequently the Adaptive Vertex Fitting algorithm is applied [38]. This technique is an iterative Kalman Filter fitter that re-weights track errors according to their χ^2 contribution to the vertex by means of the Fermi-like function given in Eq. 4.1.

$$w_i = \frac{1}{1 + e^{(\chi_i^2 - \chi_{cutoff}^2)/2T}} \tag{4.1}$$

Here, χ_i^2 is the χ^2 contribution of the i-th track to the primary vertex, χ_{cutoff}^2 is the distance where the function drops to 0.5, and T is a parameter controlling the sharpness of the function. The weight is re-computed with respect to the newly fitted vertex at each iteration until convergence is achieved.

In the last step the primary vertices are selected with the goal to identify the ones corresponding to the largest momentum transfer (high Q^2 processes). The remaining primary vertices are assigned to additional soft interactions with may take place in the same bunch crossing. A probability if the track is compatible with the momentum distribution for a soft collision is defined for each track:

$$\mathcal{P}(p_T) = \frac{\int_{p_T}^{\infty} F(q_T)\mathrm{d}q_T}{\int_{0.5}^{\infty} F(q_T)\mathrm{d}q_T} \tag{4.2}$$

here p_T is the measured transverse momentum of the track and $F(q_T)$ is the transverse momentum distribution of tracks produced in soft collisions respectively predicted by Monte Carlo simulations.

Finally the joint probability that all the tracks coming from a vertex are consistent with being created in a soft collision is calculated by using Eq. 4.3:

$$\mathcal{P}_{soft} = \Pi \sum_{k=0}^{N-1} \frac{-\ln \mathcal{P}(p_T)}{k!} \tag{4.3}$$

The vertex with the smallest value of $\mathcal{P}_{soft}$ is identified as the hard scatter interaction vertex. It is the vertex with the largest momentum transfer and used to define the kinematic quantities for all the reconstructed objects in the event.

4.4 Reconstruction of EM Objects and Electron

Electrons and photons deposit most of their energy in the electromagnetic (EM) part of the calorimeter in the form of EM showers. The dominant processes for the shower development are bremsstrahlung for electrons and positrons and pair production for photons. The cross-section of these processes becomes almost independent of energy above 1 MeV. After the first step of the process the energy of the initial particle (electron or photon) is shared between two particles ($e\gamma$ for electrons, e^+e^- for photons). These particles then interact further, leading at each step to the creation of more particles each carrying a fraction of the initial energy. The shower maximum with the largest number of particles is reached when the average energy per particle becomes low enough to stop further multiplication. From this point on the shower dissolves slowly through ionization losses for electrons or by Compton scattering for photons.
The reconstruction of these EM showers starts with the clustering of EM towers [39] . The initial clustering is performed using the *Simple Cone Algorithm*, looking for EM towers exceeding a minimum transverse energy of 0.5 GeV and declaring them as seeds. In the next step a sum over all calorimeter towers within $\Delta R < 0.4$ is performed around these seeds to construct calorimeter clusters. If necessary the $\eta - \phi$ position of the cluster-center gets recalculated when adding towers in the cone. The object is accepted if the energy of the simple cone cluster formed this way is larger than $E_T > 1.5$ GeV and more than 90% of the total energy is deposited in the EM part of the calorimeter. The cluster must also be well-isolated within the detector, which translates into a dearth of energy in the area surrounding the cluster. The isolation ratio is defined by creating a list of towers within a 8 tower wide circle around the highest E_T tower. Then the total energy $E_{tot} = E_{EM} + E_{had}$ within a cone of $\Delta R = 0.4$ around the initial cluster position is calculated using this list. Calculating the energy E_{core} deposited in a smaller circle with $\Delta R = 0.2$ one can compare the ratio of the energies deposited in both cones. If the ratio $(E_{tot} - E_{core})/E_{core}$ is found to be less than 0.2 the cluster is isolated (see Fig. 4.1).

Isolated clusters are more likely to be of electromagnetic origin apposed to clusters caused by hadronic showers. Another criterion for the electron is the longitudinal shower development of the cluster in the calorimeter which differs for electromagnetic and hadronic objects. To provide a measure how similar the shower development of the cluster is to the expectation of an electron a 7×7 covariance matrix (*Hmatrix* or $hmx7$) is calculated using seven correlated variables. These seven variables are the energy deposited in each of the four electromagnetic layers of the calorimeter, shower energy in the EM calorimeter, z position of the primary vertex divided by its uncertainty and the width of the shower in $r - \phi$ in the third EM layer.

The χ^2 is defined as:

$$\chi^2_{hxm7} = \sum_{ij}^{N} (x_i - \mu_i) H_{ij} (x_j - \mu_j) \tag{4.4}$$

with the observed values x_i and the means μ_i for the shower shape observables, which are derived from Monte Carlo simulations of electrons. For electrons this $hmx7$-variables should have small values, the requirement applied is $\chi^2_{hmx7} < 50$.

Electrons are charged and will produce a track in the tracker. Therefore their identification requires - in contrast to an EM cluster originating from photons - an associated (matched) track

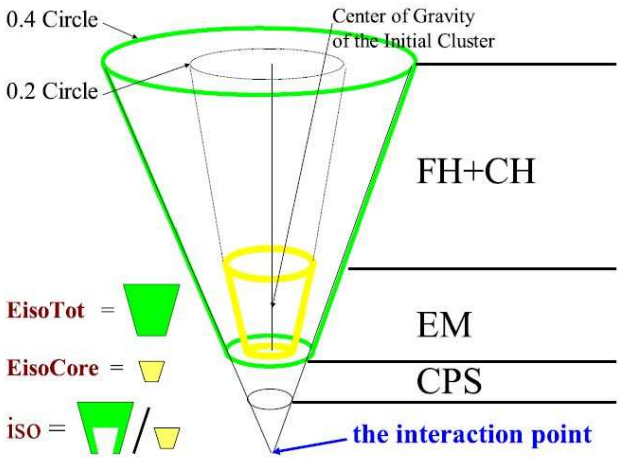


Figure 4.1: Definition of calorimeter isolation. The green cone gives the outer and the yellow one the inner isolation cone. The isolation variable is defined as the ratio of the inner and outer core energies.

in the tracking detectors. A track is matched if it points to the cluster in the third layer within a cone of $\Delta\eta \times \Delta\phi < 0.05 \times 0.05$. The track-matching is quantified by calculating χ^2_{EMtrk}, which is defined from the difference in ϕ and z in EM3 layer and the squared significance of the resolution of the transverse energy of the cluster over the transverse momentum of the track E_T/p_T:

$$\chi^2_{EMtrk} = \left(\frac{\Delta\phi}{\sigma\phi}\right)^2 + \left(\frac{\Delta z}{\sigma z}\right)^2 + \left(\frac{E_T/p_T - 1}{\sigma_{E_T/p_T}}\right)^2 = \chi^2_{spatial} + \left(\frac{E_T/p_T - 1}{\sigma_{E_T/p_T}}\right)^2 \tag{4.5}$$

However, this method poses shortcomings in finding low p_T electrons. An alternative algorithm is also run which basically reverses the recipe given. First tracks are identified, then preshower clusters are associated with these tracks and finally the information is matched to a calorimeter tower. This method is described in detail in Ref. [40]. Now isolated electrons within jets are reconstructed using the Road Method (RM) [41]. The RM extrapolates tracks of charged particles into the calorimeter.

The criteria described so far define "loose isolated" electrons. A "tight isolated" electron is defined as a subset of the "loose" isolated electrons which has to match the same criteria plus one additional requirement on the output of the electron likelihood *em-llhood8* [43] which is required to be at least 0.85. The purpose of the electron likelihood is to discriminate electrons from background processes, e.g. jets from multijet production faking electrons. Eight variables (hence the name) are used as input for the electron likelihood: The electromagnetic fraction, the fraction of energy deposited in the EM part of the calorimeter divided by the total cluster energy, two covariance matrices using a different number of variables, E_T/p_T, the probability $\mathcal{P}(\chi^2_{spatial})$ [44], the z position of the closest approach of the matched track to the primary vertex, the number of tracks within a cone of $\Delta R = 0.05$ around the matched track and the sum of all track momenta within a cone of $\Delta R = 0.4$ around the track, excluding the candidate track itself. Figure 4.2 shows the performance of the electron likelihood for the endcaps and the central part of the calorimeter.

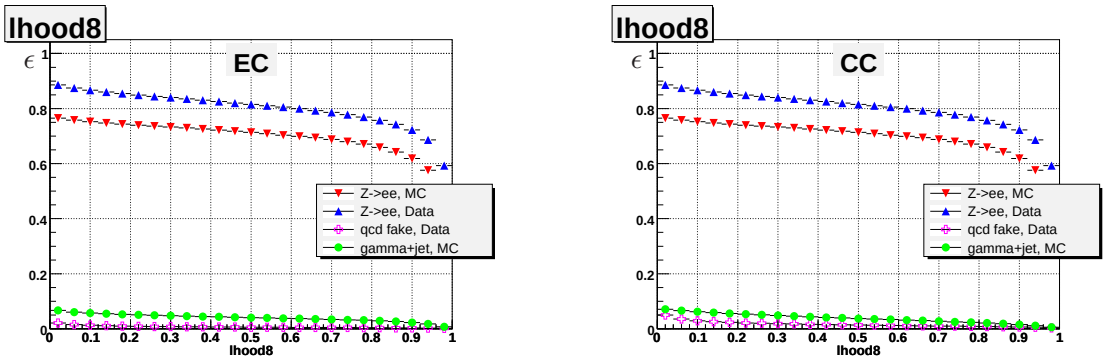


Figure 4.2: The performance of the electron likelihood for $Z \rightarrow ee$ events in data and MC and for background processes of multijet and $\gamma + jet$ production [42]. The left plot shows the efficiency for the endcaps of the calorimeter, the right plot the efficiencies for the central part.

Although the efficiencies in both parts are comparable the contribution of instrumental background in the endcaps of the detector is higher than in the central part. Figure 4.3 shows the signal efficiency vs. the background rejection of both, the endcap and the central part of the detector. The electron ID used in the present analysis is `tight_trk`.

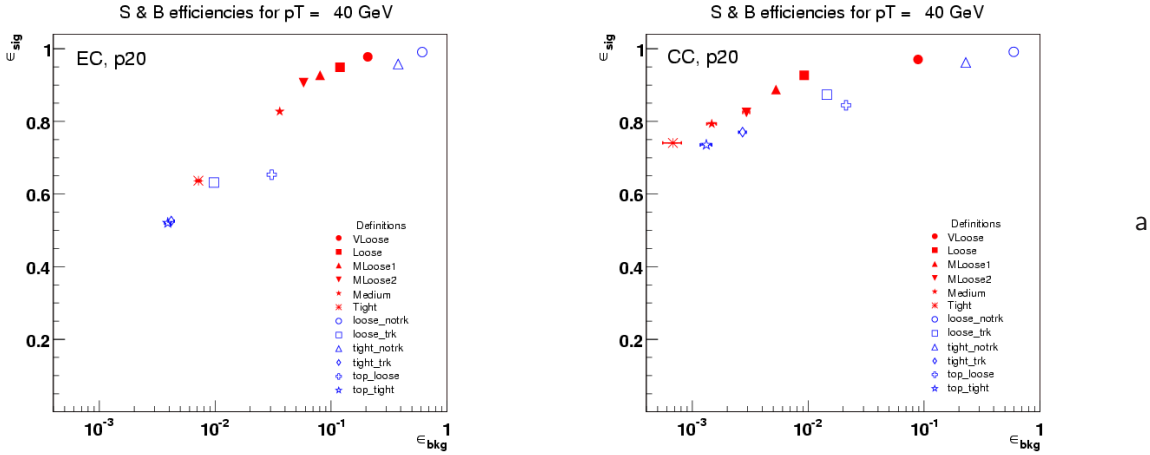


Figure 4.3: The rejection of various electron ID definitions for a 40 GeV electron. The comparison is done using Run IIb data. The electron ID used in the present analysis is `loose_trk` [45]. The left plot shows the rejection for the endcaps of the calorimeter, the right plot the rejection for the central part.

4.5 Muon Reconstruction

The reconstruction of muons uses information from the muon system and the central tracking detectors [46]. The former delivers unambiguous muon identification, the latter provides precise momentum resolution and a high efficiency of finding tracks in the entire angular acceptance region of the muon system. In each of the three layers of the muon system track segments are reconstructed from the layer wire and scintillator hits. The separately reconstructed segments of the three layers are matched with each other to form a muon track candidate. The muon track candidates get matched with a central track, creating a "central track-matched muon". The efficiency of the muon reconstruction could be increased by extrapolating central tracks to the calorimeter and looking for the signature of a minimum ionizing particle. But this algorithm shows still a low efficiency of $\sim 50\%$, much less efficient than all other types of muon signature. According to the number of hits and layers in the muon system and the quality of the matched track of the reconstructed muon, each muon is classified into several muon quality and track quality categories [47]. In the present analysis the muons have to fulfill the `loose` muon quality and the `medium` track quality. The criteria for the muons can be summarized as follows.

- The transverse momentum of the muon p_T^{μ} has to exceed 12 GeV
- The muon has to fulfill the requirements of a loose muon candidate
  - Track isolation: $\mathcal{I}_{trk} = \sum_{r_i>0.1}^{r_i>0.4} p_T^i < 2.5 \ \ \mathrm{GeV}$, here r_i gives the distance from the track in $\eta \times \phi$.
  - Calorimeter isolation: $\mathcal{I}_{cal} = \sum_{r_i>0.1}^{r_i>0.4} E_T^i < 2.5 \ \ \mathrm{GeV}$
  - At least one SMT hit: $N_{Hits}^{SMT} \geq 1$
  - $\chi^2_{trk} < 4$
  - Distance of closes approach to the beam spot in the transverse plane < 0.02
- The whole muon coverage is used: $|\eta| < 2$
- Distance from primary vertex: $\Delta \mathrm{Z}(\ell, \ell') < 1.5 \ \ \mathrm{cm}$
- A veto on cosmic muons is applied in order to suppress muons from cosmic radiation. The standard cosmic ray veto requires scintillator hits in all three layers in less than $|t| < 10$ ns.

4.6 Tau Reconstruction

In contrast to electrons and muons tau-leptons decay very rapidly, their lifetime is about 290×10^{-15} sec. Therefore the decay takes place within the beam pipe and the tau leptons have to be identified using their decay products. In $\sim 66\%$ of the cases taus decay into one or more charged hadrons plus possibly in neutral pions. The taus can be identified by a combination of tracks in the central tracking system and clusters of energy in the calorimeters. The presence of neutral pions can be inferred by significant energy deposited in the EM layers of the calorimeter. The reconstruction begins by finding calorimeter clusters, using a cone algorithm with a cone size of $R = 0.3$. Electromagnetic subclusters are then found by using a nearest neighbor algorithm in the

third layer of the EM calorimeter. If a cluster is found, EM cells in the other layers and preshower hits are added to the cluster. The subclusters are designed to identify π^0 particles originating for example from $\pi^{\pm} \rightarrow \rho^{\pm}\nu \rightarrow \pi^{\pm}\pi^0\nu$ decays. All tracks with $p_T > 1.5$ GeV within a cone size of $R = 0.5$ around the cluster center are ordered in terms of their momentum. The highest p_T track is associated with the cluster and up to two more tracks are associated if they are within 2 cm of the first track at the production vertex. A second track is added if the mass of the first and second track is less than 1.1 GeV and a third is added if the mass of the first two tracks is less than 1.7 GeV. A detailed description of the reconstruction algorithm can be found in Ref. [48]. At this stage in the reconstruction the tau candidates are split into three types, defined by their detector signature:

Type 1: Calorimeter cluster with one track associated and no EM subcluster, corresponding mostly to the decay $\tau^{\pm} \rightarrow \pi^{\pm}\nu$.

Type 2: Calorimeter cluster with one track associated and at least one EM subcluster, corresponding mostly to the decay $\tau^{\pm} \rightarrow \pi^{\pm}\pi^0\nu$.

Type 3: Calorimeter Cluster, with two or three associated tracks, with or without EM subcluster. This corresponds mostly to the decays $\tau^{\pm} \rightarrow \pi^{\pm}\pi^{\pm}\pi^{\mp}(\pi^0)\nu$.

These decay channels reflect back on the classification in the detector. However, the classification is not strict. If for example a decay as $\tau^{\pm} \rightarrow \pi^{\pm}\nu$ starts showering in one of the first layers of the calorimeter and depositing energy in the EM layer it may be reconstructed as type 2 tau. Or if $\tau^{\pm} \rightarrow \rho^{\pm}\nu \rightarrow \pi^{\pm}\pi^0\nu$ falls into the ICD region with a poor EM calorimeter coverage the reconstruction will classify this event as type I.

Jets misidentified as taus contribute significantly to the sample selected at this stage. In order to separate the tau events from jets faking hadronic taus a set of neural networks (NN) have been constructed [49]. These neural networks are trained using a set of variables discriminating between jets and taus. The output of the neural network will be a value in the range of $[0, 1]$. Lower values correspond to background-like events and higher values of the NN output correspond events likely to be a real tau candidate event. Using this NN gives a much better separation power than each single variable. The variables used in the neural network are:

- $(E^{EM1} + E^{EM2})/E^{\tau}$, where E^{EM1} and E^{EM2} correspond to the energies deposited in the first and second layer of the EM calorimeter and E^{τ} is the energy of the cluster, within a cone of $R < 0.5$. This variable offers the best discrimination power between τ-type 1 and 3.
- $\Sigma p_T^{trk}/\Sigma p_T^{\tau_{trk}}$, where p_T^{trk} is the p_T of a track within a cone of $R < 0.5$. $p_T^{\tau_{trk}}$ corresponds to the sum of the transverse momenta of the tracks, associated with the tau candidate.
- Fine hadronic fraction, fraction of E_T in the hadronic part of the calorimeter.
- $E_T^{\tau}/\left(E_T^{\tau} + \Sigma p_T^{trk}\right)$, energy of the cluster divided by the sum of the cluster energy and the total track momentum.
- $\sqrt{(\Delta\phi/sin\Theta)^2 + (\Delta\eta)^2}/\pi$ where the differences are between the vector sum of τ-track directions and the vector sum of the EM cluster. Used for τ-types 2 and 3.

- Transverse energy of the leading EM subcluster divided by the transverse energy in the layer 3 of the calorimeter in a cone of $R < 0.5$. Only used for tau-type 2.
- $(E_T^1 + E_T^2)/E_T^\tau$ where E_T^1 and E_T^2 are the transverse energies of the two most energetic calorimeter towers. Used for all tau types. Also a modified shower profile is used $profile2 = profile/(0.67 + 0.22 \cdot |\eta_d|)$ for $|\eta_{det} > 1.5|$ to remove η dependencies of the $profile$ variable.
- E_T^{em}/E_T^τ, where E_T^{em} corresponds to the transverse energy of the EM cluster. This variable is used for tau type I and II only.
- Transverse energy of the leading τ track divided by the transverse energy of the τ. This variable is considered in all three tau types.
- Calorimeter isolation, $\mathcal{I}_{cal} = (E_T^\tau - E_{core}^\tau/E_{core}^\tau)$, here E_T^τ is the cluster energy in the $R < 0.5$ cone and E_{core}^τ the energy in the $R < 0.3$ cone. Again a modified isolation parameter is defined: $\mathcal{I}_{cal}^2 = \mathcal{I}_{cal}/(1.5 \cdot |\eta_d| - 0.5)$ for the region $\eta_d > 1$. This variable is considered in all three tau types.
- RMS of the shower, $\sqrt{\Sigma_{i=1}^n [(\Delta\phi_i)^2 + (\Delta\eta_i)^2]\, E_{T_i}/E_T}$, represents the width of the calorimeter cluster of the tau. Again the variable is used for all three tau types.

Figure 4.4 shows the efficiency of the tau reconstruction algorithm as a function of the visible transverse momentum of a tau lepton $p_T^{tau,vis}$, which excludes neutrinos, as function of η_{det}. For a tau efficiency of 95% the typical rejection efficiency is about $\sim 85\% - \sim 90\%$, depending on the tau type.

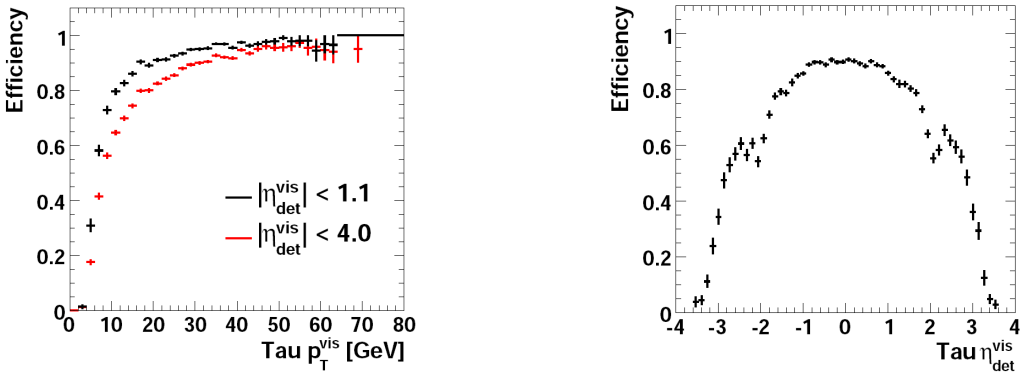


Figure 4.4: Reconstruction efficiency of hadronically decaying tau leptons as a function of $p_T^{\tau,vis}$ and η_{det} [50].

4.7 Jet Reconstruction

Jets are identified in the detector as large amounts of hadronic energy in small angular regions in the calorimeter. These are the experimental signature of quarks or gluons caused by the hadronisation

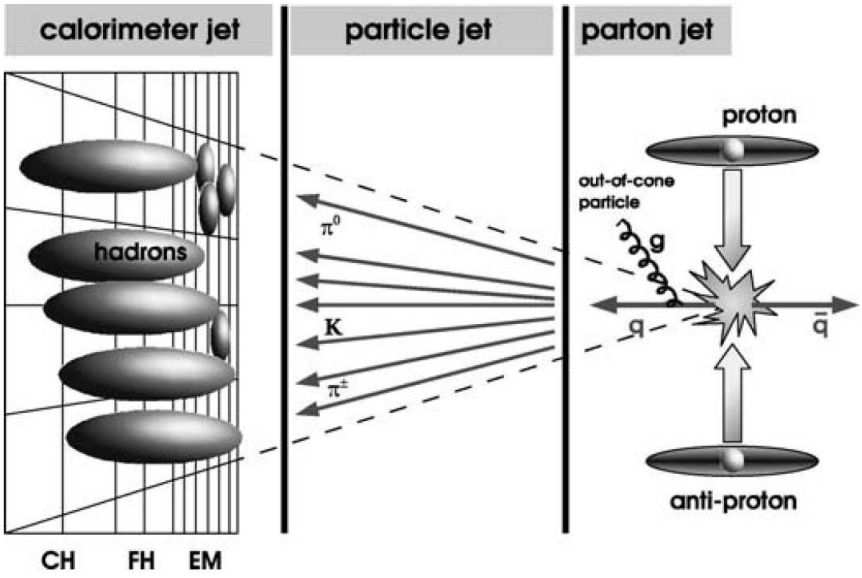


Figure 4.5: Schematic of the jet development illustrating the idea of jet energy scale corrections, taking energy measurements on the calorimeter level to the particle or the parton level [9].

process. For $p\bar{p}$ collision jets are formed by partons from the hard interaction, from initial state radiation (ISR) and from beam remnants (beam jets). Jet algorithms enable the identification of jets and to categorize them according to their quality. Study their resolution and correct the jet energies such that one obtains the initial parton energy. A schematic of the jet development is shown in Fig. 4.5. The details of these procedures are described below.

4.7.1 Jet Reconstruction and Identification

Two different jet reconstruction algorithms are available at DØ , a *cluster algorithm* [51] and the *cone algorithm*. The cone algorithms principle is based on the idea of associating all particles within a cone of radius $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$ in $\eta \times \phi$. The cone algorithm used at DØ is called "*Improved Legacy Cone Algorithm*" (ILCA or Run II Cone Algorithm) [52, 53]. For improving the jet energy scale (JES) and the $\not\!\!E_{\mathrm{T}}$ resolution noisy calorimeter cells are suppressed by the $T42$ algorithm [54].

The cone algorithm consists out of three stages. The first stage resembles the Simple Cone Algorithm used for the reconstruction of electrons as described in Sec. 4.4 by forming preclusters using calorimeter towers. After being sorted in E_T these towers are clustered to higher-order E_T towers using a cone of $\Delta R = 0.3$. All towers with a transverse energy of more than $E_T > 0.5$ GeV can act as seeds for the pre-clustering. Preclusters with $E_T > 1$ GeV and more than one tower are considered as input to the ILCA algorithm. The precluster with the highest E_T is used as seed for the formation of "proto-jets". Then all preclusters within a cone of $\Delta R = 0.5$ are assigned iteratively to the "proto-jet" until the E_T-weighted cone center is found and stable. In order to avoid sensitivity to soft radiation, stable cones around midpoints of any combination of two "proto-

jets" are searched for. Finally a list of "proto-jets" from preclusters and midpoints is considered in the last step of merging and splitting. So far the condition that each precluster is only present in one jet is not fulfilled. Therefore in the next step all pairs of "proto-jets" within a distance larger than the cone size but smaller than twice the cone size are considered. These cones get either split or merged, according to their overlapping energies. If the overlap exceeds 50% of the lower-energy cone of the two jets the jets are merged. If the overlap is smaller the preclusters get assigned to the cone with the cone axis closer in $\eta \times \phi$. All jets with $E_T > 6$ GeV are kept for the next step of jet identification.

In this analysis only 0.5-cone jets with a transverse energy larger than 15 GeV and $|\eta_{det}| < 2.5$ are used. To ensure that they are well separated from electromagnetic objects, it is required that the fraction of energy deposited in the EM calorimeter is smaller than 0.9. To remove jets which are faked by noise appearing in the coarse hadronic calorimeter, the fraction of energy in this region has to be smaller than 0.4. Another quality criterion is the L1 confirmation which compares the energy of the reconstructed jet to the energy reconstructed by the L1 trigger readout. Jets are accepted only if the ratio of the energy measurements is above a certain threshold, which depends on $|\eta_{det}|$ and p_T. Further details about reconstruction and identification of jets are given in Ref. [55].

4.8 Jet Energy Scale

Due to detector effects, dead material, sampling rate and the fixed cone radius the calorimeter jet energy can differ significantly from the initial energy of the particles forming the jet. Therefore a correction to the energy of the reconstructed jets is applied. This correction is defined [56, 57] as:

$$E_{jet} = \frac{E_{jet}^{raw} - O}{F_\eta \times R \times S} \tag{4.6}$$

with

E_{jet}	:	corrected jet energy at particle level,
E_{jet}^{raw}	:	uncorrected jet energy,
O	:	offset energy correction
F_η	:	relative response correction,
R	:	absolute response correction,
S	:	showering correction,

The individual components of the jet energy scale correction are derived and applied sequentially in the order as given in Eq. 4.6. The estimation of the corrections is done separately for data and Monte Carlo. Starting from the raw jet energy the offset correction O is subtracted. The offset energy arises from multiple $p\bar{p}$ interactions, beam remnants, electronics and uranium noise in the calorimeter or energy from previous collisions ("pile-up"). The offset correction is measured from data using minimum bias events which are triggered by the luminosity monitor. The energy density per tower is measured depending on the number of reconstructed primary vertices in order

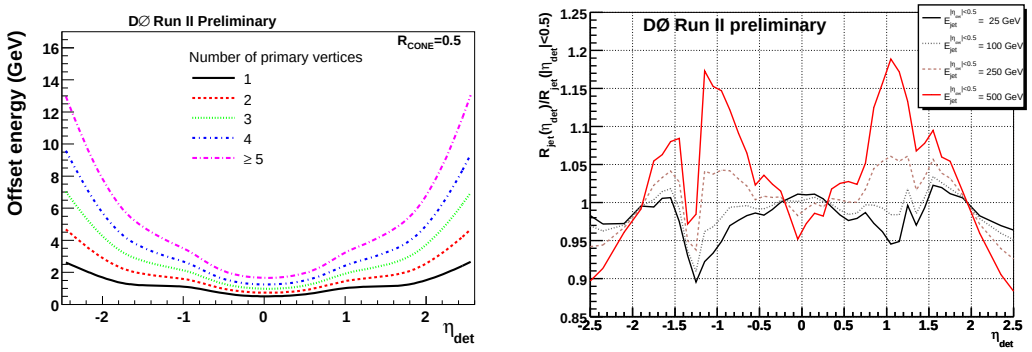


Figure 4.6: Example plots for offset and relative response corrections [57]. Left: Offset energy as a function of η_{det} for different primary vertex multiplicities. Right: Example of relative response correction in data.

to include the luminosity dependence in the offset energy calculation. The energy deposited within a jet cone is then defined as the offset energy. Figure 4.6 (left) shows the offset energy as function of η_{det} for different primary vertex multiplicities.

After correcting for the offset the jet energy is divided by the relative response correction F_η. This relative response correction calibrates the jet energy for the fact that the response of the calorimeter is not uniform in rapidity. Particularly the ICR and the massless gaps (explained in detail in Sec. 3.2.2) show a different response compared to the regular calorimeter cells in the central and endcap parts of the calorimeter. For the measurement of F_η the *Missing Transverse Energy Projection Fraction* (MPF) method is applied. Here dijet or photon+jet events are used with the tag object always in the central detector region and the probe object in the region to be considered. The difference in transverse energy between tag and probe object can then be used to extract the relative response. Due to differences between the physical detector and its simulation this response determination is performed separately in data and Monte Carlo. Figure 4.6 (right) shows the relative response correction in data. Significant spikes in the ICD region are visible.

After the determination of the relative response the absolute response correction $\mathcal{R}$ can be measured and applied to the jet energy. $\mathcal{R}$ corrects for example the difference in calorimeter response of hadrons and electrons and energy loss in non-instrumented detector regions. Figure 4.7 (left) shows the absolute response correction as function of jet energy for jets with cone $R = 0.7$.

The last step is the determination and application of the shower correction $\mathcal{S}$. Due to showering in the calorimeter or the bending of charged particles in the DØ solenoidal field, energy belonging inside (outside) the jet can be missing from (added to) the jet energy. The correction does not account for physical showering e.g. due to gluon emission. $\mathcal{S}$ is measured in photon+jet events with exactly one primary vertex. The ratio between the jet energies at the particle level and the reconstruction level yields the showering correction. Figure 4.7 (right) shows the showering correction in data.

When muons are reconstructed within the jet cones, the energy scale has to be further corrected to consider the energies of the muon and the neutrino.

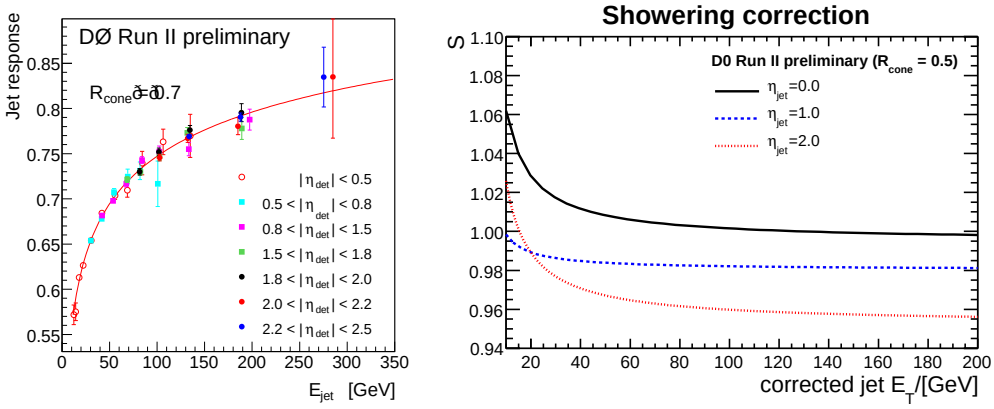


Figure 4.7: Example plots for absolute response and showering correction [57]. Left: Example of absolute response for $R = 0.7$ in data versus the offset and relative response corrected jet energy. Right: Example of showering correction in data versus corrected transverse jet energy.

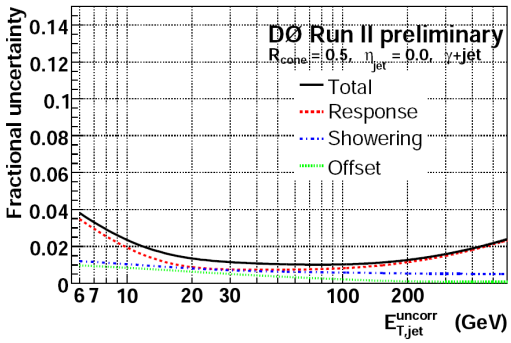


Figure 4.8: Relative uncertainty of the p17 JES correction in data as function of uncorrected transverse jet energy for $R = 0.5$ and $\eta = 0.0$ [57, 56].

The uncertainty on the Jet Energy Scale is for many measurements one of the largest uncertainties. The uncertainty in dependency of the jet momentum is shown in Fig. 4.8.

4.8.1 Jet Resolution in Data and MC

The jet energy resolution (JER) is also derived using data [58]. The resolutions can be correlated to the calorimeter resolution as given in Eq. 3.4. Due to limitations of the modeling of the DØ detector it is necessary to modify the default MC simulation and match it to the performance observed in data. For jet resolution, jet reconstruction efficiencies and identification efficiencies, a method called *"Jet shifting, smearing and removal"* (JSSR) is introduced [59, 60]. The method uses the transverse momentum imbalance in photon+jets and $Z + \text{jet}/\gamma$ events. The JSSR method is applied on jets after η dependent corrections. Simulated jets are first oversmeared according to their resolution function and then shifted according to a correction function of the resolution. This resolution correction is derived as the relative difference between data and Monte Carlo.

4.9 Missing Transverse Energy

The missing transverse energy ($\not\!\!E_{\mathrm{T}}$) is a crucial quantity for the present analysis. $\not\!\!E_{\mathrm{T}}$ is the signature of particles escaping the detector undetected such as neutrinos or weakly interacting particles predicted by non-SM theories like for example some supersymmetric models. Partons involved in the original hard scattering process have large longitudinal momenta but their total transverse momentum at the interaction point is approximately 0. Therefore no net transverse momentum of the particles produced in the $p\bar{p}$ interaction should be produced. Thus the magnitude of the vectorial sum of the x- and y-components are expected to be zero within the detector resolution. However, when a particle escapes the detector without interacting it produces an imbalance of transverse momentum. $\not\!\!E_{\mathrm{T}}$ is the amount of energy needed to restore the balance. The location of each cell in the detector is used and combined with the primary vertex position information to yield a direction for the energy $\vec{\eta}^{\text{cell}}$. The transverse direction $\vec{\eta}_T$ is associated to the energy measured in the cell E^{cell} such that $\vec{E}_T^{\text{cell}} = E^{\text{cell}} \cdot \vec{\eta}^{\text{cell}}$. The missing energy in an event is simply the negative sum of all transverse energy of all cell's.

$$\vec{E}_T^{\text{miss}} = -\sum_{i=0}^{N_{cells}} \vec{E}_T^i \tag{4.7}$$

Splitting up the missing energy into its transverse components can be done from the following relation:

$$\left(E_T^{\text{miss}}\right)^2 = \left(E_{T_x}^{\text{miss}}\right)^2 + \left(E_{T_y}^{\text{miss}}\right)^2 \tag{4.8}$$

$$E_{T_x} = \sum_{i=0}^{N_{cell}} E_i \cos\theta_i = -E_{T_x}^{\text{miss}} \tag{4.9}$$

$$E_{T_y} = \sum_{i=0}^{N_{cell}} E_i \sin\theta_i = -E_{T_y}^{\text{miss}} \tag{4.10}$$

It is important to note that only energetic cells are included in this calculations. The $\not{E}_{\mathrm{T}}$ and calorimeter E_T are corrected for the presence of reconstructed muons in the event, noisy cells from the coarse hadronic section of the calorimeter, jets and corrections to their energy scale. Due to the large noise the energy of the coarse hadronic calorimeter is not taken into account.

5
Calorimeter Calibration

The energy measurement is sensitive to the final mechanical precision of the calorimeter and physical and electronic non-uniformities smear out the detector resolution and limit the ability to perform precision measurements. Particularly this analysis depends almost exclusively on calorimeter objects and their energy measurement. Moreover, a precise knowledge of the calorimeter enables one to decrease some statistical uncertainties like the Jet Energy Scale. The calibration has been performed as part of this thesis to improve the acceptance and the energy measurements.

5.1 Data Selection

The calibration has been performed using data taken after the shutdown in July 2006 and has been repeated regularly to monitor the long term stability and behavior of the calorimeter. Care has been taken to cover the entire range of instantaneous luminosities in each individual calibration. A possible bias of the calibration by the choice of triggers has been studied but no dependency was found. Therefore not explicit trigger requirements were applied to maximize the available statistics. Runs affected by operational issues or susceptible to noise have been discarded. The calorimeter calibration is performed on raw data, e.g. data not yet converted into a more convenient format as explained in Sec. 4 without applying any additional calibration constants like corrections for jets or the corrections from timing studies.

5.2 Patterns of Noise

When studying the energy spectra in the calorimeter cells three erroneous patters can be found distinguished.

- *Noisy cell*: Noise induced in the cells may lead to energy spectra which are many orders of magnitude lager than the expectation from the average energy distribution. Additionally the shape can be distorted.
- *Low occupancy cell*: The energy deposition is significantly smaller than the average energy.
- *Dead cell*: The cell is showing no energy deposition.

Each region of the detector of fixed η ("η-ring") has 64 cells in ϕ for each layer. Figures 5.1-5.4 show the energy spectrum of a given cell plotted in black compared to the average energy

distribution of a given η-ring, plotted in red. The calculation of the average energy spectra can be compromised if spectra of cells affected by any of the erroneous patterns are considered, hence these are excluded from the calculation. The calibration itself is performed for all cells. To create a list of faulty cells for calculating a correct average spectra one has to manually inspect all energy profiles as shown in Fig. 5.1 and classify them accordingly. There are altogether about 55.000 cells in the DØ calorimeter, whereas 396 are located in the Inter Cryostat Detectors (ICD) region and have to be treated separately due to luminosity dependence. Although the ICD cells represent just a small fraction of the calorimeter they are they only layer in this region, hence covering up to 20% of the possible event yield. Additionally their design based on photomultiplier tubes rather than liquid argon makes these cells more susceptible to effects like aging, radiation damages, electronic drifts and luminosity effects.

Figure 5.1 shows an example of a good cell. The energy deposited in the cell itself is given by the black distribution. One sees that both distributions agree well. Figure 5.2 shows the energy distributions of a cell affected by noise patterns. The measured spectrum of the cell exceeds the average expected energy by several orders of magnitude and the shape of the distributions differ significantly. An entirely unresponsive cell is displayed in Fig. 5.3, only very few events of same the same energy have been detected. A cell with lower response is shown in Fig. 5.4, one sees that the energy yield is significantly lower as expected from the average energy spectrum and both distributions shows different shapes.

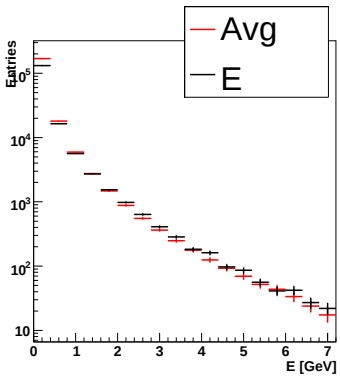


Figure 5.1: Example of a good cell with well agreeing energy distributions. The average energy per cell expected by integrating over all cells per ϕ ring and normalizing the spectra is given in red. The energy deposited in the cell itself is given by the black distribution.

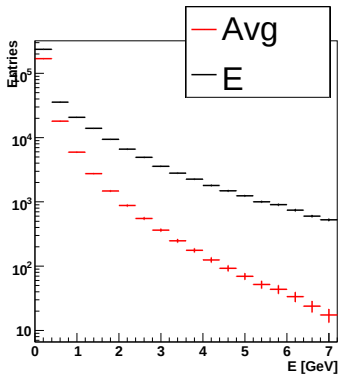


Figure 5.2: Example of cell affected by noise patterns. The average energy per cell expected by integrating over all cells per ϕ ring and normalizing the spectra is given in red. The energy deposited in the cell itself is given by the black distribution.

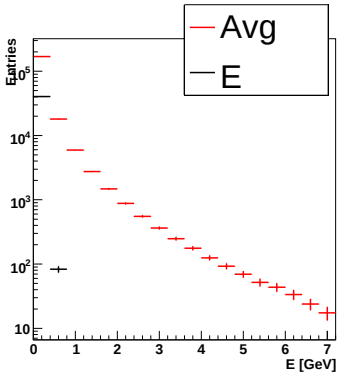


Figure 5.3: Example of a dead cell where no energy spectrum can be measured. The average energy per cell expected by integrating over all cells per ϕ ring and normalizing the spectra is given in red. The energy deposited in the cell itself is given by the black distribution.

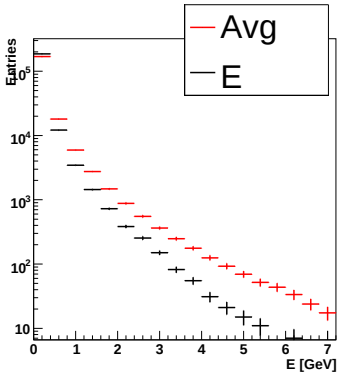


Figure 5.4: Example for a cell with a low energy yield. The average energy per cell expected by integrating over all cells per ϕ ring and normalizing the spectra is given in red. The energy deposited in the cell itself is given by the black distribution.

5.3 Calibration Procedure

The calibration procedure for the DØ calorimeter consists of two steps: Calibration of the readout electronics using pulser data and correction of non-uniformities due to mechanical variations in the detector using collider data. This calibration method is called ϕ *intercalibration*. The method is based on equalizing the response for each cell in a given ϕ ring of the calorimeter. Since the beams of the Tevatron are not polarized, the energy flux from the interaction region is ϕ-independent an uniform calorimeter occupancy in ϕ is expected. Hence any ϕ-non-uniformity in the measured occupancy is caused by differences in the calorimeter response. In order to correct for these deformations the effective gains of the cells in rings of fixed η are corrected until the response is uniform. This corrections are usually a correction factor applied in the reconstruction software to adjust the gain. For the ICD cells it may be necessary to adjust the high voltage of the photomultiplier tube as well. This corrections are then migrated to a data base which applies the corrections to each cell during reconstruction of the events.

In principle almost every physical process with calorimeter related quantities can be chosen to equalize the responses of the cells in ϕ at given η. Here events in a precision readout tower with a transverse energy above an E_T threshold of E_{tr} =3 GeV are counted. The sum of events counted in all cells of an η ring is used to calculate the average amount of events per cell. Again erroneous cells are not taken into account.

$$\bar{N} = \frac{\sum_i N_i}{64 - n_e} \qquad i = 1...64 - n_e \tag{5.1}$$

Here n_e represents the number of excluded defective cells and N_i the number of hits above the E_T

threshold. In the next step the energy threshold E_{tr} is adapted such that the number of events passing the modified E_T-cut $E'_{tr,i}$ is equals $\bar{N}$. The ratio $\alpha_i = \frac{E_{tr}}{E'_{tr,i}}$ gives the calibration constant for a given cell i.

Summarizing the calibration procedure:

- Obtaining the numbers of events per ϕ-tower above a given E_T threshold E_{tr}
- Calculating the average number of events per ϕ-tower for the entire η-ring
- For each tower in ϕ the lower cut E'_{tr} gets adjusted such that the occupancy of the tower matches the average occupancy
- The calibration constants for a particular ϕ-tower is given by

$$\alpha = \frac{E_{tr}}{E'_{tr}} \tag{5.2}$$

Figure 5.5 shows the energy spectra of a low occupancy cell before and after applying the correction. After calibration the average and the individual energy spectra agree well.

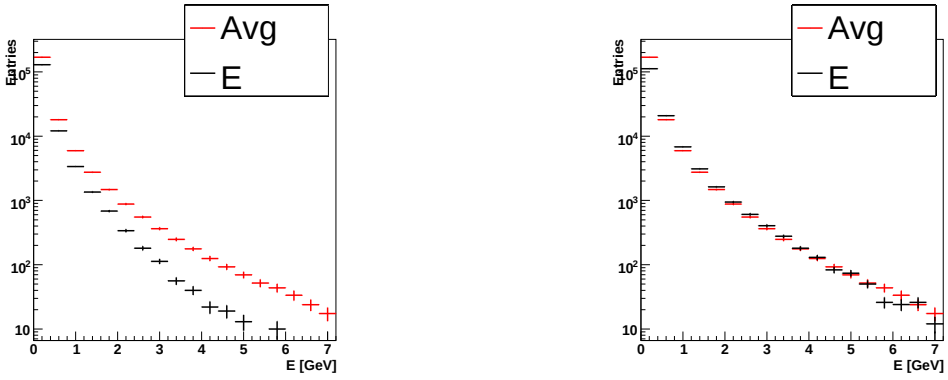


Figure 5.5: Energy profiles for a cell at $\eta = -1.3,\ \phi = 9$. Left one sees a significant deviations between the average energy spectrum in red and the energy spectrum of the cells represented by the black distribution. After applying the calibration the energy spectra agree well (right).

The calibration is particularly crucial for the ICD detectors, a region susceptible to energy mismeasurements during Run IIa. The photomultiplier tubes of the individual cells show partly strong luminosity dependences, electronic drifts and aging effects. A successful re-commissioning effort requires a calibration procedure able to obtain luminosity dependent calibration constants. It has to be sufficiently fast to allow frequent monitoring of hardware changes like replacement of photomultiplier tubes or changes of high voltage of the photomultiplier tubes. During the re-commissioning the calibration has been repeated several times to select faulty cells and adjust the photomultiplier tubes. An example for cell with a luminosity dependent gain is shown in Fig. 5.6.

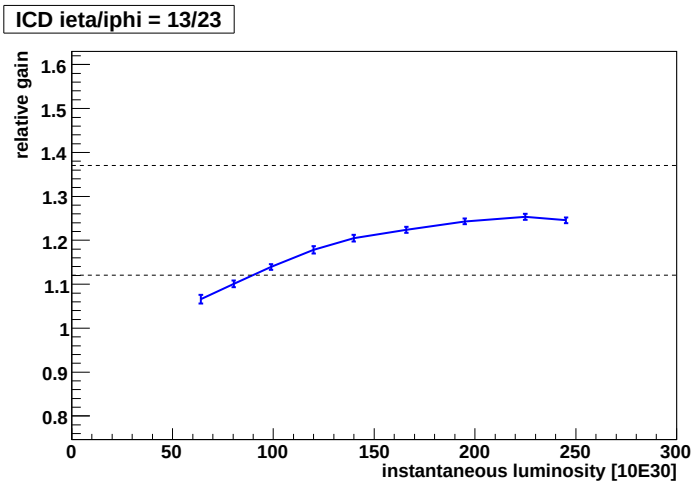


Figure 5.6: Example of a luminosity dependent gain for a calorimeter cell located in the intercryostat region.

North-South-Calibration

When comparing the average energy profiles per $|\eta|$ ring for the north and south part of the detector differences in the absolute scale of the measured average energy profiles are revealed. The north region corresponds to negative η values and the south side has positive η values. These differences vary, depending on the layer and η region which are compared. The north side shows less cells erroneous cells and therefore has been chosen as reference. This North-South-intercalibration is performed after applying the ϕ-intercalibration.

5.4 Result of the Calorimeter Calibration

By establishing an equal response throughout the calorimeter the energy resolution and misidentification rate of EM object have significantly decreased. Particularly the successful commissioning of the ICD detectors led to an 20% increased signal acceptance for some analysis like the $ZH \rightarrow \nu\nu b\bar{b}$ search. Figure 5.7 shows the ϕ distribution of jets, given in units of ϕ slices, in the ICD from jet+$\not\!\!E_{\mathrm{T}}$ triggers with transverse energies exceeding 15 GeV.

Figure 5.8 shows that luminosity dependence and gain are stable for a given cell. The calibration data is recorded with a time difference of five months,the The first one, represented by the blue line, is performed on data from October 2008. The second calibration is given by the red line and corresponds to data taken in March 2009.

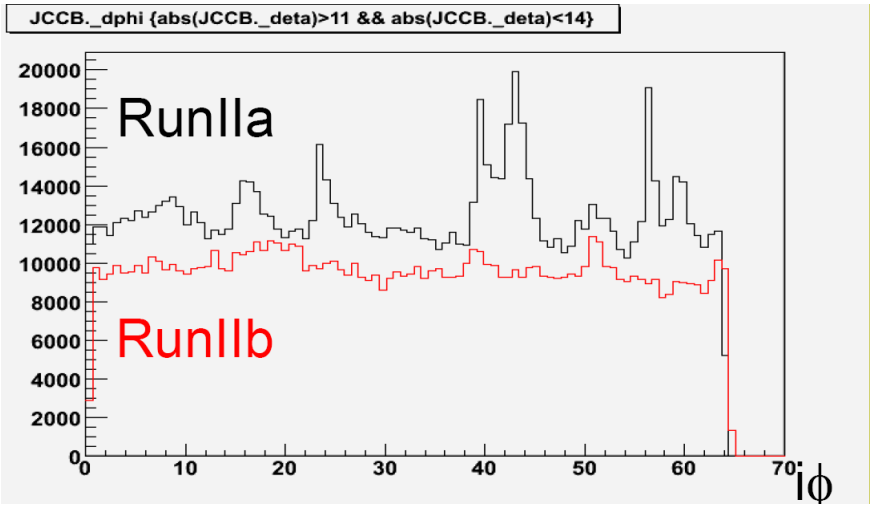


Figure 5.7: ϕ distribution of jets in the Inter Cryostat Detectors (ICD) from jet+$\not\!\!E_{\mathrm{T}}$ triggers with transverse energies exceeding 15 GeV. The x-axises gives the ϕ coordinate of the η ring, ranging from ϕ =1-64. The distributions is acquired by integrating over all layers of the ICD. The black solid line gives the distribution for Run IIa, the red line the distribution for Run IIb.

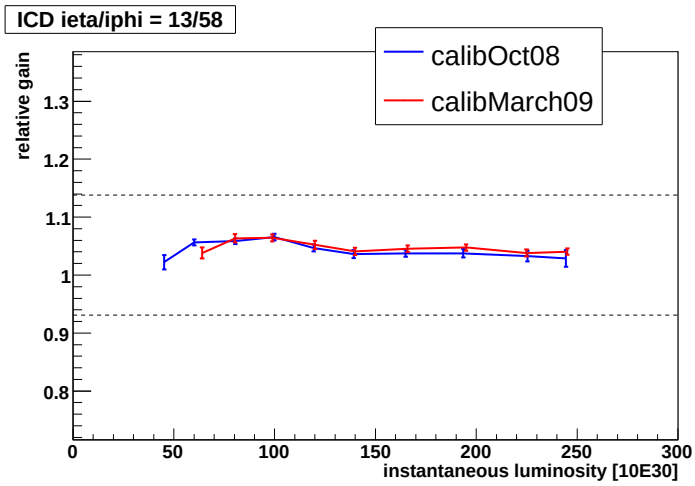


Figure 5.8: Comparison of two calibrations performed with a time difference of five months. Luminosity dependence and gain are stable.

6
Data and Monte Carlo Samples

The understanding of the experimental observations in the study of $p\bar{p}$ interactions and the search for the Higgs boson benefits from a proper modeling of all possible background processes and Higgs signals. For this purpose simulations of the background and signal samples, including all the effects of the interaction of particles with the detector and of data reconstruction, are used. In samples dominated by known Standard Model processes one can check that the data being analyzed are not affected by hardware failures and readout problems during data taking. A careful understanding of these processes is crucial to develop methods for enriching the signal events in the data sample. In the following section the data samples and the Monte Carlo (MC) simulations used in this analysis are described.

6.1 Data Sample

The analysis is based on data collected between April 2002 and December 2008. This covers the entire Run II data. The DØ detector recorded in this time 4.2 fb^{-1} of integrated luminosity delivered by the TEVATRON. The data is ordered in units called a *store, run* and *luminosity block*. A store is the data taking period between putting a beam into the TEVATRON in colliding mode and the time the beam is discarded ('dumped') and the Tevatron is prepared to host a new store. A store lasts typically about 8 hours but can be significantly longer or shorter, depending on the instantaneous luminosity of the particular store and decisions made by the main control room crew of the Tevatron. Each store is divided into several runs, a run corresponds to about 2 hours of data taking and therefore has about the same instantaneous luminosity and the same trigger prescales. Each run is again divided in luminosity blocks, the smallest unit used to measure the quantity of the data, lasting typically about one minute. Finally each single event recorded gets an unique number assigned. Runs or individual luminosity blocks can be marked as affected by hardware failures or noise for any of the sub-detectors [61]. The state of each single subdetector is continuously monitored and recorded and quality checks are applied to the data after recording. Events suffering from known issues during data taking are removed from the analysis. Also single events containing specific characteristic patterns of noise are removed [62]. The integrated luminosity can be calculated depending on the triggers used for the analysis, list of bad luminosity blocks and events which are discarded due to possible problems with noise. This calculation is based on a non-prescaled trigger and cross checked by normalizing the total number of $Z/\gamma^* \rightarrow e^{\pm}e^{\mp}$ events in the invariant mass distribution in data and Monte Carlo at preselection level. The integrated luminosity of the data sample for the $e^{\pm}e^{\mp}$ final state was found to be approximately $4.2\ \mathrm{fb}^{-1}$.

6.2 Background Processes

A precise knowledge of the various background contributions is necessary to achieve an effective background reduction while retaining most of the expected signal events. The following paragraphs briefly explain the main backgrounds in order of their severity.

6.2.1 Di-Boson Production

Pair-production of gauge bosons is the most important background process in this analysis, although its production cross section is small compared to the other background processes. The boson-pair production consists out of the three sub-processes WW, WZ and ZZ. Among these the pair production of W bosons is the one with the highest production cross section and most difficult to separate from the signal. The corresponding Feynman diagram, Fig. 6.1, shows that the decay of the WW system is identical to the WW decays in $H \to WW$ processes. This means each event contains two high quality leptons and real $\not\!\!E_{\mathrm{T}}$ due to the neutrinos escaping the detector. The only difference is caused by the fact that the Higgs is a spin-0 particle, resulting in different angular correlations between the two leptons in the final state. Therefore the most powerful quantity which can be used to separate between the Higgs boson decay and the W pair production is the opening angle in the transverse plane, $\Delta\phi(\ell,\ell)$, although only a partial background suppression is possible using this distribution only.

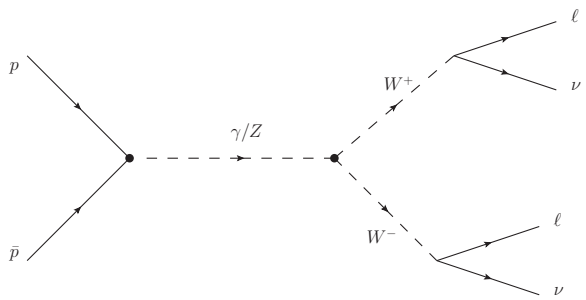


Figure 6.1: Feynman diagram for WW production.

Since a complete or almost complete separation of the expected Higgs signal and the WW production is not possible this process is an *irreducible background*.

6.2.2 $W+$jets Production

The second most important background is the $W+$jets production where the W decays to a lepton and a neutrino. If a jet gets mis-identified as lepton this event can pass the preselection. The kinematics differ significantly from the signal processes, $\not\!\!E_{\mathrm{T}}$ will be in average lower than in a Higgs decay. Due to the large production cross section of about 2600 pb^{-1}, a significant amount of these events pass the various selection requirements. Hence this process posses the second largest contribution to the sample composition at the end of the selection, even if it is suppressed very efficiently by the selection requirements used in the analysis. In addition the theoretical cross

section is known only at the NLO level, introducing thereby a sizable systematic uncertainty. The Feynman diagrams of these processes are displayed in Fig. 6.2.

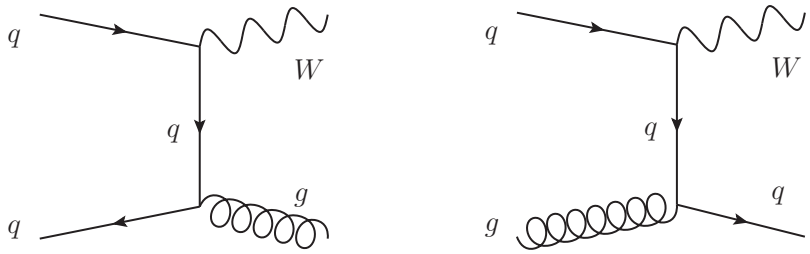


Figure 6.2: Leading order Feynman diagrams for W + jets production.

6.2.3 Top Pair Production

Figure 6.3 shows the decay of a single top quark. Since the dominant top quark production channel at the Tevatron is pair-production the final state depicted appears twice. As one sees the decay of the W boson leads here as well to high values of $\not\!\!E_{\mathrm{T}}$ and two good quality leptons, therefore mimicking the expected signal closely. Fortunately the t quark always decays in association with a b quark, leading to high jet multiplicities with a rather hard spectrum. The absence of these jets can be used to reject this background efficiently.

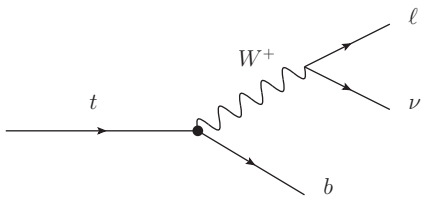


Figure 6.3: Feynman diagram of the decay of the top quark.

In March 2009 both, the DØ and CDF collaborations reported observation of the single top production process [63, 64]. Although the decay of single top production processes with only one lepton in the final state and high jet multiplicities are not expected to fake the signal dedicated studies have been performed. No single top event has been found to pass to final selection requirements as described in Sec. 7.1.

6.2.4 Z + jet/γ Production

The main contribution of the data at the initial selection stage is made of Drell-Yan (DY) processes like $q\bar{q} \to Z/\gamma \to \ell\ell$, leading to the prominent peak of the Z boson in the invariant mass spectrum of the di-electron system as shown in Fig. 6.9. These final states features two high quality leptons but little or none $\not\!\!E_{\mathrm{T}}$. Also many kinematic and topological distributions differ significantly from the corresponding signal contributions. Therefore these background processes are easily separated.

Drell-Yan events passing the final selection (see Sec. 7.1) requirements show mis-measured lepton momenta, leading to fake $\not\!\!E_{\mathrm{T}}$ in the event. The Feynman diagrams of these processes are displayed in Fig. 6.4.

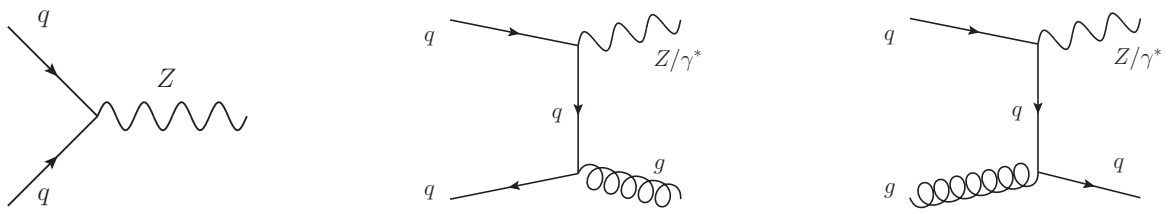


Figure 6.4: Leading order Feynman diagrams for $Z + \text{jet}/\gamma$ production.

6.2.5 Multijet production

The multijet background consist mainly out of dijet events in which jets or low-p_T and low quality leptons are faking the desired signal. Although no adequate Monte Carlo description of this background is available the shape and contribution can be estimated using data. Multijet background can be easily separated from the signal processes and plays only a minor role at the final selection stage.

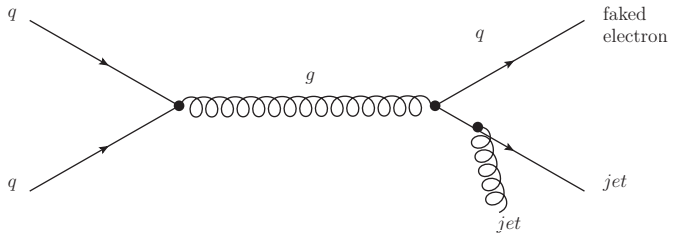


Figure 6.5: Feynman diagram for multijet production.

6.3 Monte Carlo Samples for Signal and Background

Simulations are needed to study the various processes contributing to the background composition and the signal. Usually the process of describing an event by simulations requires to connect several descriptions of production and decay of a process by combining perturbative theoretical calculations with non-perturbative phenomenological models. This is done by Monte Carlo generators. Furthermore a reliable simulation of detector effects is needed.

6.3.1 Event simulation

Physical processes can be of different nature such as the three most studied at collider experiment, e^+e^-, $e^{\pm}p$ and $p\bar{p}$ collisions. The latter is illustrated in Fig. 6.6. The key in these simulations is

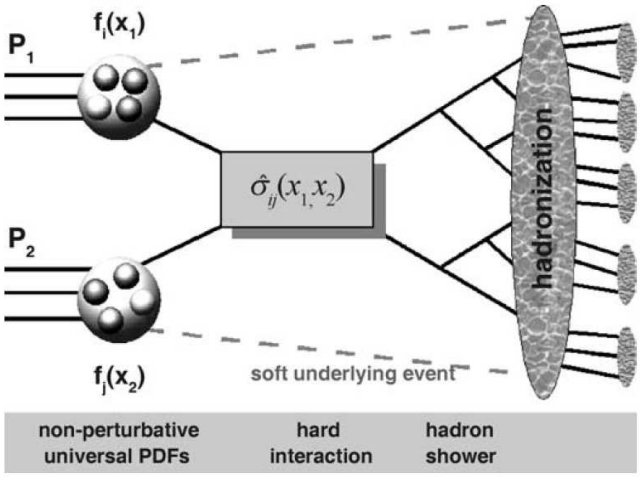


Figure 6.6: Schematic drawing of an event in $p\bar{p}$ collisions [9].

to break up the event into different steps.

The simulation starts with the hard interaction of the colliding particles. The hard interaction is calculated with perturbative quantum field theory. The particles participating in the collision can radiate photons and gluons before the hard interaction takes place, called Initial State Radiation (ISR). The radiation of photons and gluons from particles produced in the hard interaction is called Final State Radiation (FSR). As seen in Fig. 6.6, this radiation from the hard interacting particles and the color charged initial state leads to complex event structure in collisions. In the event simulation the radiation of particles from final particles of the hard interaction is separated from the radiation before the hard interaction. There are different methods for the simulation of this perturbative step, like the parton shower and the color dipole ansatz [65, 66]. Parton showers are based on an improved leading-log approximations. Hence they cannot be accurate for well separated partons but they offer a simple, process-independent machinery that gives a sensible match to hadronization. Once energies of about 1 GeV are reached, the fragmentation of quarks and gluons into hadrons takes place. Phenomenological models are necessary for the description of the hadronisation. All models have free parameters that need to be measured in order to properly describe the fragmentation process. The most commonly used Monte Carlo generators use either the string fragmentation model as in PYTHIA [67], or cluster fragmentation as realized in the HERWIG event generator [68]. The Lund model implemented in PYTHIA splits gluons into $q\bar{q}$ pairs and turns them into hadrons via the string fragmentation model [69]. In the last step the unstable hadrons decay into stable final state particles that can be identified in the detector.

The main background processes for the analysis are Z decays in electronic and tau final states, diboson, $W+$jets production, $t\bar{t}$ and multijet production. The Standard Model background processes and cross sections used in the present analysis are listed in Tables 6.3, 6.4 and 6.5. All Monte Carlo samples were generated at a center of mass energy of $\sqrt{s} = 1.96$ TeV assuming a top quark mass of $m_t = 170.0$ GeV and using the CTEQ6L1 parton distribution functions (PDFs) [70]. The diboson, $Z+$jet/γ and signal samples were generated using PYTHIA [69]. ALPGEN [71] is

designed for the generation of the Standard Model processes in hadronic collisions with emphasis on final states with large jet multiplicities. It is based on the exact LO evaluation of partonic matrix elements. Therefore the $W+$ jets background processes are simulated using both generators, ALPGEN and PYTHIA. The ALPGEN samples are utilized by the $H \rightarrow WW^* \rightarrow e^+\nu e^-\nu$ analysis whereas the PYTHIA samples are employed by the $H \rightarrow WW^* \rightarrow \mu + \tau_{had} + X$ analyses. Both analyses are discussed in the following chapters. The topologies simulated using ALPGEN are $W + \ell\nu + Nlp$, $W + \ell\nu + 2b + Nlp, W + \ell\nu + 2c + Nlp$. Here lp is any flavor of u, d, s, $N = 0, 1, 2, 3$ for the heavy flavor b and c-samples and $N = 0, 1, 2, 3, 4, 5$ for the processes without heavy quark content. A detailed overview of these samples is given in Table 6.5. Each ALPGEN event is passed to PYTHIA for parton showering and hadronization.

6.3.2 Cross Section Calculation and Normalization

For all background processes an amount of Monte Carlo events has been generated which exceeds the amount of events expected in the data for this particular process. For processes with small production cross section as the diboson samples, $t\bar{t}$ or signal processes this corresponds to hundred or even thousands of events generated per expected data event. Hence the Monte Carlo samples have to be normalized to the expected number of data events based on their cross section. Often the production cross sections are calculated in leading-order (LO). This means only Feynman diagrams without loop corrections and radiation processes of gluons or photons are considered. When considering higher-order corrections the calculated production cross section becomes more accurate. Second order contributions are called next-to-leading-order (NLO) and third order processes next-to-next-to-leading-order (NNLO). Interference terms of virtual gluon loops as well contribute to the NLO calculation. The best value for the production process is obtained when considering as many higher order corrections as possible. Often the calculations of LO processes are already very complex so most calculations of production cross sections are NLO. The ratio of a NLO cross section and the LO cross section is called *k factor*

$$k = \frac{\sigma_{NLO}}{\sigma_{LO}} \tag{6.1}$$

The k-factor of the ration between NNLO and NLO is called k_{NNLO}. These k_{NNLO} factors, if available, have to be considered for the normalization as well.

Signal Monte Carlo Samples

All signal processes are generated in NLO using PYTHIA for a mass range of $115 - 200$ GeV in steps of 5 GeV. The dominant $gg \rightarrow H$ production cross sections are calculated at next-to-next-to-leading logarithmic (NNLL) in QCD. The calculation includes the most advanced theoretical information available at present for this observable: soft-gluon resummation up to NNLL accuracy, the exact treatment of the bottom-quark contribution up to next-to-leading order, and two-loop electroweak effects [72]. The most recent parametrization of parton distribution functions at next-to-next-to-leading order, called MSTW 2008, are adapted and the corresponding uncertainties are evaluated [73]. Additional higher order corrections of 3-loop $\mathcal{O}(\alpha\alpha_s)$ correction to the Higgs boson production cross section arising from light quarks. These calculations probe the factorization of

QCD and electroweak perturbative corrections to this process. The second largest contribution to Higgs boson production is the vectorboson-fusion (VBF) $q\bar{q} \rightarrow q\bar{q}H$. This process is known at NLO in QCD. Although being the second largest production cross section the relative fraction to the selected signal events is just about 10%. The Higgs decay via two W bosons in a dileptonic final state is simulated, $H \rightarrow WW^* \rightarrow \ell^+\nu\ell^-\nu$. The production cross sections multiplied with their branching ratio ($\sigma \times BR$) of these processes is summarized in Table 6.1. Minor additional contributions are coming from ZH and WH production. These processes are known incorporating NNLO QCD corrections and NLO electroweak corrections [21]. The simulation of these processes has been performed considering the decay into pairs of heavy gauge bosons except WZ, leptons, photons and γZ. The $\sigma \times BR$ of these processes are summarized in Table 6.2. Although almost negligible in the high mass range these contributions add about 5%-10% sensitivity in the low mass domain. All signal cross sections have been calculated using the HIGLU program [74], further details can be found in Ref. [75].

Process	$\sigma_{\mathrm{NNLL}} \times \mathrm{BR}$ [pb]
$gg \to H(115) \to WW^* \to \ell^+\nu\ell^-\nu$	0.00927
$gg \to H(120) \to WW^* \to \ell^+\nu\ell^-\nu$	0.01363
$gg \to H(125) \to WW^* \to \ell^+\nu\ell^-\nu$	0.01858
$gg \to H(130) \to WW^* \to \ell^+\nu\ell^-\nu$	0.02359
$gg \to H(135) \to WW^* \to \ell^+\nu\ell^-\nu$	0.02822
$gg \to H(140) \to WW^* \to \ell^+\nu\ell^-\nu$	0.03205
$gg \to H(145) \to WW^* \to \ell^+\nu\ell^-\nu$	0.03485
$gg \to H(150) \to WW^* \to \ell^+\nu\ell^-\nu$	0.03674
$gg \to H(155) \to WW^* \to \ell^+\nu\ell^-\nu$	0.03815
$gg \to H(160) \to WW^* \to \ell^+\nu\ell^-\nu$	0.03987
$gg \to H(165) \to WW^* \to \ell^+\nu\ell^-\nu$	0.03867
$gg \to H(170) \to WW^* \to \ell^+\nu\ell^-\nu$	0.03538
$gg \to H(175) \to WW^* \to \ell^+\nu\ell^-\nu$	0.03201
$gg \to H(180) \to WW^* \to \ell^+\nu\ell^-\nu$	0.02849
$gg \to H(185) \to WW^* \to \ell^+\nu\ell^-\nu$	0.02340
$gg \to H(190) \to WW^* \to \ell^+\nu\ell^-\nu$	0.01984
$gg \to H(195) \to WW^* \to \ell^+\nu\ell^-\nu$	0.01762
$gg \to H(200) \to WW^* \to \ell^+\nu\ell^-\nu$	0.01591

Process	$\sigma_{\mathrm{NLO}} \times \mathrm{BR}$ [pb]
$q\bar{q} \to q\bar{q}H(115) \to qqWW^* \to \ell^+\nu\ell^-\nu$	0.0006
$q\bar{q} \to q\bar{q}H(120) \to qqWW^* \to \ell^+\nu\ell^-\nu$	0.0010
$q\bar{q} \to q\bar{q}H(125) \to qqWW^* \to \ell^+\nu\ell^-\nu$	0.0014
$q\bar{q} \to q\bar{q}H(130) \to qqWW^* \to \ell^+\nu\ell^-\nu$	0.0018
$q\bar{q} \to q\bar{q}H(135) \to qqWW^* \to \ell^+\nu\ell^-\nu$	0.0022
$q\bar{q} \to q\bar{q}H(140) \to qqWW^* \to \ell^+\nu\ell^-\nu$	0.0026
$q\bar{q} \to q\bar{q}H(145) \to qqWW^* \to \ell^+\nu\ell^-\nu$	0.0029
$q\bar{q} \to q\bar{q}H(150) \to qqWW^* \to \ell^+\nu\ell^-\nu$	0.0032
$q\bar{q} \to q\bar{q}H(155) \to qqWW^* \to \ell^+\nu\ell^-\nu$	0.0035
$q\bar{q} \to q\bar{q}H(160) \to qqWW^* \to \ell^+\nu\ell^-\nu$	0.0035
$q\bar{q} \to q\bar{q}H(165) \to qqWW^* \to \ell^+\nu\ell^-\nu$	0.0035
$q\bar{q} \to q\bar{q}H(170) \to qqWW^* \to \ell^+\nu\ell^-\nu$	0.0033
$q\bar{q} \to q\bar{q}H(175) \to qqWW^* \to \ell^+\nu\ell^-\nu$	0.0030
$q\bar{q} \to q\bar{q}H(180) \to qqWW^* \to \ell^+\nu\ell^-\nu$	0.0027
$q\bar{q} \to q\bar{q}H(185) \to qqWW^* \to \ell^+\nu\ell^-\nu$	0.0023
$q\bar{q} \to q\bar{q}H(190) \to qqWW^* \to \ell^+\nu\ell^-\nu$	0.0020
$q\bar{q} \to q\bar{q}H(195) \to qqWW^* \to \ell^+\nu\ell^-\nu$	0.0018
$q\bar{q} \to q\bar{q}H(200) \to qqWW^* \to \ell^+\nu\ell^-\nu$	0.0016

Table 6.1: Higgs gluon-fusion and vectorboson-fusion processes generated using PYTHIA along with their cross sections from the NNLL and NLO theory prediction. This are the two dominant Higgs production processes for the present analysis.

Process	$\sigma_{\mathrm{NLO}} \times \mathrm{BR}$ [pb]
$ZH(115) \to Z(H \to WW,\ ZZ,\ \gamma\gamma,\ ee,\ \mu\mu,\ \tau\tau,\ \gamma Z) \to Z + \text{inclusive}$	0.01775
$ZH(120) \to Z(H \to WW,\ ZZ,\ \gamma\gamma,\ ee,\ \mu\mu,\ \tau\tau,\ \gamma Z) \to Z + \text{inclusive}$	0.02056
$ZH(125) \to Z(H \to WW,\ ZZ,\ \gamma\gamma,\ ee,\ \mu\mu,\ \tau\tau,\ \gamma Z) \to Z + \text{inclusive}$	0.02397
$ZH(130) \to Z(H \to WW,\ ZZ,\ \gamma\gamma,\ ee,\ \mu\mu,\ \tau\tau,\ \gamma Z) \to Z + \text{inclusive}$	0.02738
$ZH(135) \to Z(H \to WW,\ ZZ,\ \gamma\gamma,\ ee,\ \mu\mu,\ \tau\tau,\ \gamma Z) \to Z + \text{inclusive}$	0.03029
$ZH(140) \to Z(H \to WW,\ ZZ,\ \gamma\gamma,\ ee,\ \mu\mu,\ \tau\tau,\ \gamma Z) \to Z + \text{inclusive}$	0.03230
$ZH(145) \to Z(H \to WW,\ ZZ,\ \gamma\gamma,\ ee,\ \mu\mu,\ \tau\tau,\ \gamma Z) \to Z + \text{inclusive}$	0.03331
$ZH(150) \to Z(H \to WW,\ ZZ,\ \gamma\gamma,\ ee,\ \mu\mu,\ \tau\tau,\ \gamma Z) \to Z + \text{inclusive}$	0.03342
$ZH(155) \to Z(H \to WW,\ ZZ,\ \gamma\gamma,\ ee,\ \mu\mu,\ \tau\tau,\ \gamma Z) \to Z + \text{inclusive}$	0.03273
$ZH(160) \to Z(H \to WW,\ ZZ,\ \gamma\gamma,\ ee,\ \mu\mu,\ \tau\tau,\ \gamma Z) \to Z + \text{inclusive}$	0.03164
$ZH(165) \to Z(H \to WW,\ ZZ,\ \gamma\gamma,\ ee,\ \mu\mu,\ \tau\tau,\ \gamma Z) \to Z + \text{inclusive}$	0.02903
$ZH(170) \to Z(H \to WW,\ ZZ,\ \gamma\gamma,\ ee,\ \mu\mu,\ \tau\tau,\ \gamma Z) \to Z + \text{inclusive}$	0.02595
$ZH(175) \to Z(H \to WW,\ ZZ,\ \gamma\gamma,\ ee,\ \mu\mu,\ \tau\tau,\ \gamma Z) \to Z + \text{inclusive}$	0.02326
$ZH(180) \to Z(H \to WW,\ ZZ,\ \gamma\gamma,\ ee,\ \mu\mu,\ \tau\tau,\ \gamma Z) \to Z + \text{inclusive}$	0.02077
$ZH(185) \to Z(H \to WW,\ ZZ,\ \gamma\gamma,\ ee,\ \mu\mu,\ \tau\tau,\ \gamma Z) \to Z + \text{inclusive}$	0.01858
$ZH(190) \to Z(H \to WW,\ ZZ,\ \gamma\gamma,\ ee,\ \mu\mu,\ \tau\tau,\ \gamma Z) \to Z + \text{inclusive}$	0.01669
$ZH(195) \to Z(H \to WW,\ ZZ,\ \gamma\gamma,\ ee,\ \mu\mu,\ \tau\tau,\ \gamma Z) \to Z + \text{inclusive}$	0.01490
$ZH(200) \to Z(H \to WW,\ ZZ,\ \gamma\gamma,\ ee,\ \mu\mu,\ \tau\tau,\ \gamma Z) \to Z + \text{inclusive}$	0.01341
$WH(115) \to W(H \to WW,\ ZZ,\ \gamma\gamma,\ ee,\ \mu\mu,\ \tau\tau,\ \gamma Z) \to W + \text{inclusive}$	0.0305
$WH(120) \to W(H \to WW,\ ZZ,\ \gamma\gamma,\ ee,\ \mu\mu,\ \tau\tau,\ \gamma Z) \to W + \text{inclusive}$	0.0350
$WH(125) \to W(H \to WW,\ ZZ,\ \gamma\gamma,\ ee,\ \mu\mu,\ \tau\tau,\ \gamma Z) \to W + \text{inclusive}$	0.0404
$WH(130) \to W(H \to WW,\ ZZ,\ \gamma\gamma,\ ee,\ \mu\mu,\ \tau\tau,\ \gamma Z) \to W + \text{inclusive}$	0.0458
$WH(135) \to W(H \to WW,\ ZZ,\ \gamma\gamma,\ ee,\ \mu\mu,\ \tau\tau,\ \gamma Z) \to W + \text{inclusive}$	0.0503
$WH(140) \to W(H \to WW,\ ZZ,\ \gamma\gamma,\ ee,\ \mu\mu,\ \tau\tau,\ \gamma Z) \to W + \text{inclusive}$	0.0533
$WH(145) \to W(H \to WW,\ ZZ,\ \gamma\gamma,\ ee,\ \mu\mu,\ \tau\tau,\ \gamma Z) \to W + \text{inclusive}$	0.0546
$WH(150) \to W(H \to WW,\ ZZ,\ \gamma\gamma,\ ee,\ \mu\mu,\ \tau\tau,\ \gamma Z) \to W + \text{inclusive}$	0.0542
$WH(155) \to W(H \to WW,\ ZZ,\ \gamma\gamma,\ ee,\ \mu\mu,\ \tau\tau,\ \gamma Z) \to W + \text{inclusive}$	0.0527
$WH(160) \to W(H \to WW,\ ZZ,\ \gamma\gamma,\ ee,\ \mu\mu,\ \tau\tau,\ \gamma Z) \to W + \text{inclusive}$	0.0505
$WH(165) \to W(H \to WW,\ ZZ,\ \gamma\gamma,\ ee,\ \mu\mu,\ \tau\tau,\ \gamma Z) \to W + \text{inclusive}$	0.0461
$WH(170) \to W(H \to WW,\ ZZ,\ \gamma\gamma,\ ee,\ \mu\mu,\ \tau\tau,\ \gamma Z) \to W + \text{inclusive}$	0.0409
$WH(175) \to W(H \to WW,\ ZZ,\ \gamma\gamma,\ ee,\ \mu\mu,\ \tau\tau,\ \gamma Z) \to W + \text{inclusive}$	0.0363
$WH(180) \to W(H \to WW,\ ZZ,\ \gamma\gamma,\ ee,\ \mu\mu,\ \tau\tau,\ \gamma Z) \to W + \text{inclusive}$	0.0322
$WH(185) \to W(H \to WW,\ ZZ,\ \gamma\gamma,\ ee,\ \mu\mu,\ \tau\tau,\ \gamma Z) \to W + \text{inclusive}$	0.0286
$WH(190) \to W(H \to WW,\ ZZ,\ \gamma\gamma,\ ee,\ \mu\mu,\ \tau\tau,\ \gamma Z) \to W + \text{inclusive}$	0.0255
$WH(195) \to W(H \to WW,\ ZZ,\ \gamma\gamma,\ ee,\ \mu\mu,\ \tau\tau,\ \gamma Z) \to W + \text{inclusive}$	0.0227
$WH(200) \to W(H \to WW,\ ZZ,\ \gamma\gamma,\ ee,\ \mu\mu,\ \tau\tau,\ \gamma Z) \to W + \text{inclusive}$	0.0202

Table 6.2: Higgs processes from associated WH and ZH production generated using PYTHIA along with their cross sections from the NLO theory prediction. The Higgs is simulated to decay into pairs of heavy gauge bosons except WZ, leptons, photons and γZ. All allowed final states are considered as indicated by the label inclusive.

Background Monte Carlo Samples

A summary of all Standard Model processes that are simulated using Monte Carlo is given in Tables 6.3, 6.4 and 6.5. The diboson Monte Carlo samples are calculated in Ref. [76] with MCFM at NLO, the uncertainties are evaluated at LO. The $t\bar{t}$ samples are normalized to a cross section employing improvements due to soft gluon resummation at next-to-next-to-leading logarithmic accuracy. The resummed results are expanded to analytical cross sections through next-to-next-to-leading order. This leads to an approximate NNLO cross section [77].
The $Z/\gamma \rightarrow \ell\ell$ cross section is calculated in NNLO, using QCD correction in α_s^2 according to Ref. [78, 79]. The $W \rightarrow \ell\nu$ cross section is calculated at LO and a k_{NNLO}-factor is applied to all generated $W+$jets processes. The complete set of $W+$jets events consists of samples with an additional gluon splitting into light partons (Wlp+jets), c quarks (Wcc+jets) and b quarks (Wbb+jets). In the NLO and higher perturbative calculations, the relative contributions of the three samples are different from leading order calculations. In particular the Wcc+jets and Wbb+jets contributions are expected to be higher in NLO. As no reliable NLO calculations exist for these samples the heavy flavor scale factor, defined as the relative fraction of W plus heavy flavor (Wcc+jets and Wbb+jets) to Wlp+jets, has to be measured from data. This has been done by selecting an $\ell + jets$ final state and employing flavor tagging methods, a technique to identify the heavy flavor content of a jet. Details can be found in Ref. [80]. The uncertainties on the cross sections are due to uncertainties of the parton distribution functions as well as variations in the renormalization and factorization scale [79]. They are also listed in Tables 6.3, 6.4 and 6.5, if available. The background contribution of multijet production has been estimated from data and is discussed in Sec. 6.6.

Process	Mass Range [GeV]	$\sigma_{\mathrm{NNLO}} \times \mathrm{BR}$ [pb]
$Z/\gamma^* \rightarrow \tau\tau$	15<M<60	455.00 ± 15.10
$Z/\gamma^* \rightarrow \tau\tau$	60<M<130	241.60 ± 8.70
$Z/\gamma^* \rightarrow \tau\tau$	130<M<250	1.96 ± 0.60
$Z/\gamma^* \rightarrow \tau\tau$	250<M<500	0.16 ± 0.02
$Z/\gamma^* \rightarrow ee$	5<M<15	5856.00
$Z/\gamma^* \rightarrow ee$	15<M<60	455.00 ± 15.10
$Z/\gamma^* \rightarrow ee$	60<M<130	241.60 ± 8.70
$Z/\gamma^* \rightarrow ee$	130<M<250	1.96 ± 0.60
$Z/\gamma^* \rightarrow ee$	250<M<500	0.16 ± 0.02
$Z/\gamma^* \rightarrow ee$	500<M	0.004584
$Z/\gamma^* \rightarrow \mu\mu$	15<M<60	455.00 ± 15.10
$Z/\gamma^* \rightarrow \mu\mu$	60<M<130	241.60 ± 8.70
$Z/\gamma^* \rightarrow \mu\mu$	130<M<250	1.96 ± 0.60
$Z/\gamma^* \rightarrow \mu\mu$	250<M<500	0.16 ± 0.02

Table 6.3: $Z/\gamma^* \rightarrow \ell\ell$ processes generated with PYTHIA along with their cross sections from the NNLO theory prediction.

Process	σ_{NLO} [pb]
$WW \rightarrow$ inclusive	11.66 ± 0.60
$WZ \rightarrow$ inclusive	3.45 ± 0.22
$ZZ \rightarrow$ inclusive	1.37 ± 0.06

Process	$\sigma_{\mathrm{NNLO(approx.)}}$ [pb]
$t\bar{t} \rightarrow$ inclusive	7.88 ± 0.20

Table 6.4: Diboson and $t\bar{t}$ processes generated with PYTHIA along with their cross sections from the NLO and NNLO theory prediction. All allowed final states are considered in the simulation.

Process	$\sigma_{\mathrm{LO}} \times \mathrm{BR}$ [pb]
ALPGEN	
$W + \ell\nu + 0lp$	4527.41
$W + \ell\nu + 1lp$	1287.67
$W + \ell\nu + 2lp$	308.52
$W + \ell\nu + 3lp$	74.42
$W + \ell\nu + 4lp$	16.98
$W + \ell\nu + 5lp$	5.18
$W + \ell\nu + 2b + 1lp$	9.35
$W + \ell\nu + 2b + 1lp$	4.31
$W + \ell\nu + 2b + 2lp$	1.53
$W + \ell\nu + 2b + 3lp$	0.73
$W + \ell\nu + 2c + 0lp$	24.40
$W + \ell\nu + 2c + 1lp$	13.49
$W + \ell\nu + 2c + 2lp$	5.46
$W + \ell\nu + 2c + 3lp$	2.55
PYTHIA	
$W \rightarrow \mu\nu$	2583 $\pm$ 93
$W \rightarrow \tau\nu$	2583 $\pm$ 93

Table 6.5: $W +$ jets processes generated with ALPGEN or PYTHIA along with their cross sections from the LO theory prediction. Here lp is any flavor of u, d, s.

6.4 Triggering

A combination of various triggers is used to select events containing two electrons in the data sample. These are mostly single or di-electron triggers but as well triggers sensitive to high values of $\not\!\!E_{\mathrm{T}}$ in combination with an electron. The triggers used in the analysis are listed in Tables A.1, A.2 in Appendix A. At Level 1 and Level 2 two major groups can be identified, the calorimeter only triggers (called E1 and E2) and the triggers which also require a track (called TE1TE5). On Level 1 these triggers usually require one electromagnetic tower to exceed 10 GeV or 15 GeV. Additionally an electron trigger accepts as well events with two neighboring towers exceeding 5 GeV and if required a track in the central tracking system matched with the calorimeter tower. On Level 2 additional selections are applied such as transverse momentum of the electron candidate and the track, isolation and a dedicated likelihood for the second stage of triggering. Finally on Level 3 tighter criteria are applied like electron likelihood and shower shape requirements. Details regarding the electron trigger menus and the trigger likelihood can be found in Ref. [81].

The trigger efficiencies are measured using the tag-and-probe method. Events are selected in a mass window around the Z mass and the tag electron is required to pass an unprescaled single electron trigger at all three levels. Then the other electron from the Z decay, the probe electron, is unbiased with respect to the trigger decision and can be used for studies. The tag electron has to fulfill very tight conditions to ensure that the data sample has a very low contribution of fake electrons. The requirements are as follows:

- Electromagnetic fraction > 0.9
- Isolation < 0.15
- HMatrix7 < 50
- Electron Likelihood > 0.85
- Electron transverse momentum > 15 GeV
- Track transverse momentum > 10 GeV
- Match to an unprescaled single electron trigger within $\Delta R < 0.4$

These variables are explained in Sec. 4.4. For the probe electron, different electron qualities are considered. The electron qualities used in this analysis require for the tag-electron:

- Electromagnetic fraction > 0.9
- Isolation < 0.2
- Electron Likelihood > 0.85

A logical OR of the efficiencies of the triggers combined for this analysis is calculated. The effective luminosity used in the analysis is a function of the triggers. The requirement of two electrons per event and the loose electron ID criteria leads to an almost $\sim 100\%$ selection efficiency.

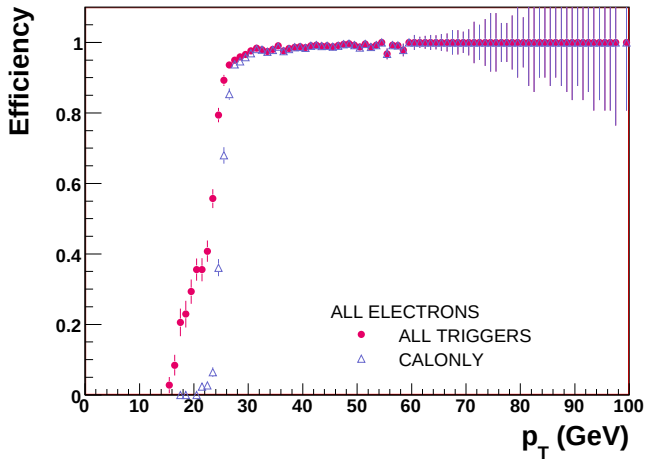


Figure 6.7: Trigger efficiencies as function of the transverse momentum p_T for a combination of all electron triggers including tracking information and triggers based on calorimeter information only. The efficiency at the plateau is close to 100% [81].

The trigger requirements and algorithms were modified throughout the entire data taking period to adapt to changes in the hardware, increase in the instantaneous luminosity and the optimization of the trigger system. The different trigger conditions applied during a given period of data taking are put in *trigger lists*. Figure 6.7 shows the efficiency for the OR of all electron triggers including tracking information and triggers based on calorimeter information only [81]. More details on the triggers used in this thesis can be found in Ref. [82].

6.5 Preselection

The total inelastic cross section σ_{tot} is about 12 orders of magnitude larger than the signal production cross section. Even individual production cross sections of typical background processes like $t\bar{t}$ or Z boson production are factors or even orders of magnitude larger than the signal production cross section. Therefore selection requirements have to be applied to suppress the background and enrich the relative signal contribution in the data sample. The most basic selection aims just for an optimal acceptance of the final state of interest while maintaining a background composition for which the various kinematic, topological and event quantities are well described.

The specific requirements placed on the electron identification criteria such as isolation, electromagnetic fraction and electron likelihood have been carefully chosen to be as loose as possible while still controlling the various background contributions. For obtaining optimal results new electron identification definitions (electron ID) have been co-developed. The electron ID used in the present analysis is defined by:

- *Isolation*: The calorimeter isolation, the ratio of the energy deposited in a cone around the electron seed with $\Delta R = 0.2$ and the energy deposited in a wider cone with $\Delta R = 0.4$ should

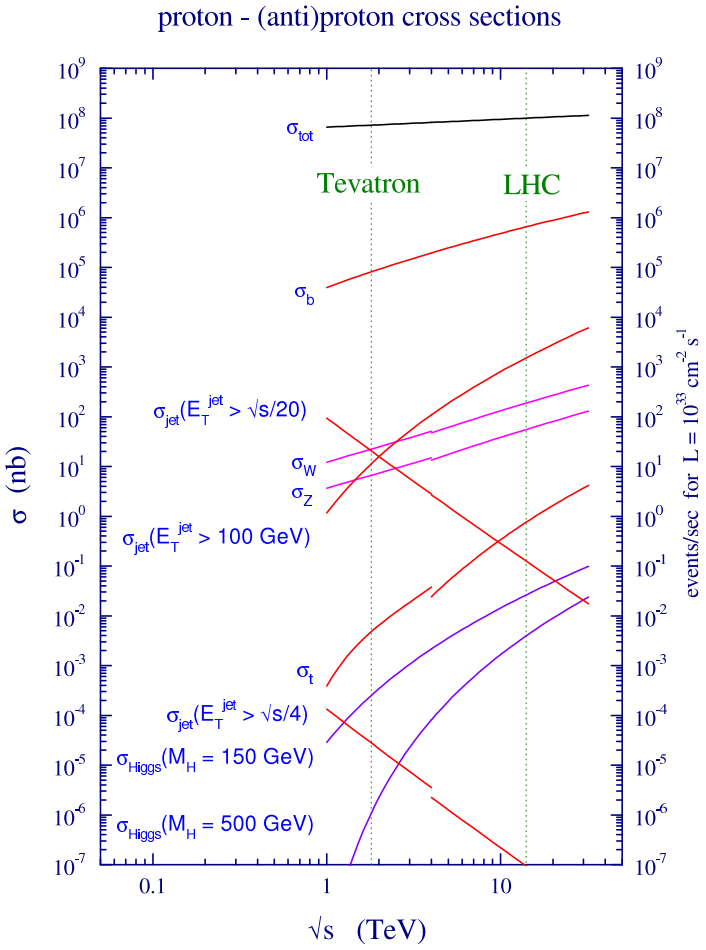


Figure 6.8: Comparison of the inclusive production cross section for various processes depending on the center of mass energy. The dashed green lines indicate the center of mass energies at the Tevatron and at the Large Hadron Collider [83].

be less than 0.2.

- *EM-Fraction*: The electromagnetic fraction, the fraction of energy deposited in the EM part of the calorimeter divided by the total cluster energy has to exceed 0.9.
- p_T: The transverse momentum of the cluster has to be at least 3 GeV.
- *Likelihood*: The value of the EM-ID likelihood has to be at least 0.2.
- *Trigger*: Trigger requirements discussed in Sec. 6.4 are used

The two electrons are required to originate from the same vertex and to be of opposite charge. The transverse momentum of each lepton has to exceed $p_T > 15$ GeV. In addition the invariant mass M_{ee} of the di-electron system is required to exceed 15 GeV. The detector regions between

$|\eta_{\text{det}}| > 1.1$ and $|\eta_{\text{det}}| < 1.4$, the inter cryostat region (see also Sec. 3.2.2), are neglected in the present analysis. This stage is referred to as *preselection stage*. The distribution of the invariant mass of the di-electron system is shown in Fig. 6.9. The exact sample composition is given in Table 7.5. The data sample is dominated by $Z + \text{jet}/\gamma$ event at the preselection stage with less than 0.01% of possible signal contributions. Therefore corrections of the Monte Carlo simulations and if necessary adaptions are evaluated at this stage and adapted if necessary. The various background contributions are discussed in Sec. 6.2.

Table 6.6 shows the production cross section of the major background contributions at TEVATRON energies, the event produced in $4.2\ \text{fb}^{-1}$ of data and the events selected after triggering and applying the preselection requirements. Typical efficiencies for heavy boson processes are at the percent level. The W production is efficiently suppressed.

	Z	W	$t\bar{t}$	ZZ	WZ	WW
σ [pb]	6690	23000	7.88	1.37	3.45	1.66
Events produced	28098000	96600000	33096	5754	14490	48972
Events selected	219830	241	132	98	112	172
Efficiency [%]	0.8	$2.5{\cdot}10^{-4}$	0.4	1.7	0.8	0.4

Table 6.6: Events produced in $4.2\ \text{fb}^{-1}$ of data according to their total production cross section and events selected after applying triggers and the preselection criteria. The preselection requirements lead to typical selection efficiencies at the percent level. The large W cross section is efficiently suppressed.

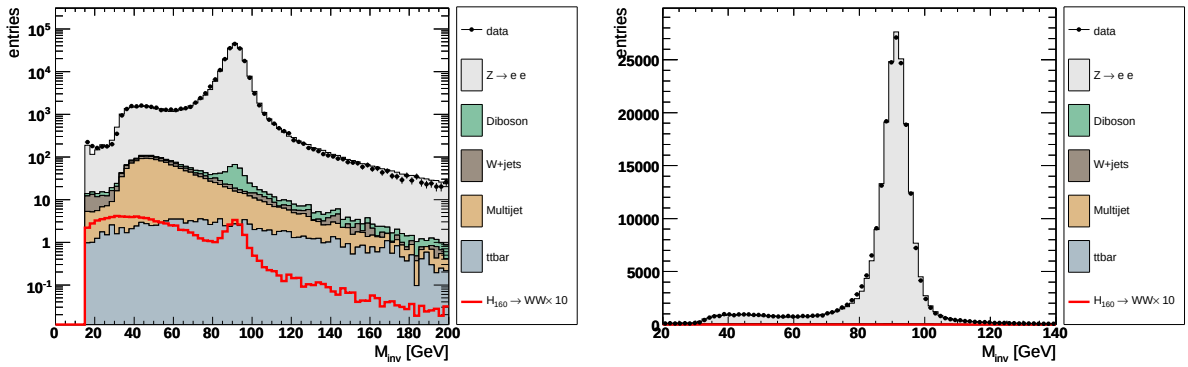


Figure 6.9: Distribution of the invariant mass M_{inv} in logarithmic scale (left) and linear scale (right) at preselection level for data (points) and sum of all backgrounds (filled histograms). The entire Run II data set is used.

In all plots comparing data and Monte Carlo the possible Higgs boson contributions are plotted with respect to a Higgs boson mass of $m_H = 160$ GeV and uses all production processes mentioned in Sec. 6.3. Figures 6.10 and 6.11 show the p_T spectra of the leading and trailing p_T electron at the

preselection level whereas Figs. 6.12, 6.13 respectively show the corresponding η_{det} distributions. The background at the preselection level is dominated by $Z/\gamma \rightarrow \ell\ell$ and multijet events. More control distributions can be found in Appendix B.1, namely the $\not\!\!E_{\mathrm{T}}$ distribution, the scalar sum of the jet p_T, the angles between $\not\!\!E_{\mathrm{T}}$ and either of the leptons and the direction of the ϕ coordinate for either lepton.

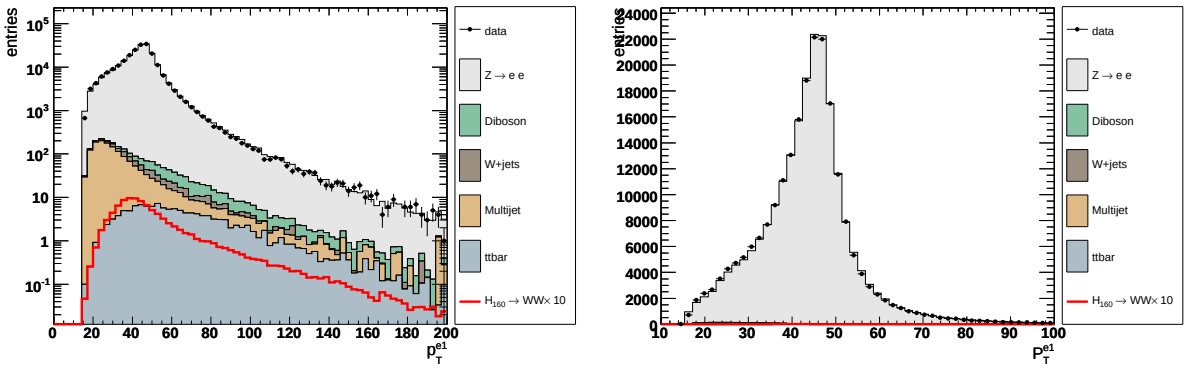


Figure 6.10: Distribution of the transverse momentum for the leading p_T electron in logarithmic scale (left) and linear scale (right) at preselection level for data (points) and sum of all backgrounds (filled histograms). The expected signal, multiplied by a factor of 10, for a 160 GeV/c^2 Standard Model Higgs is also shown. The entire Run II data set is used.

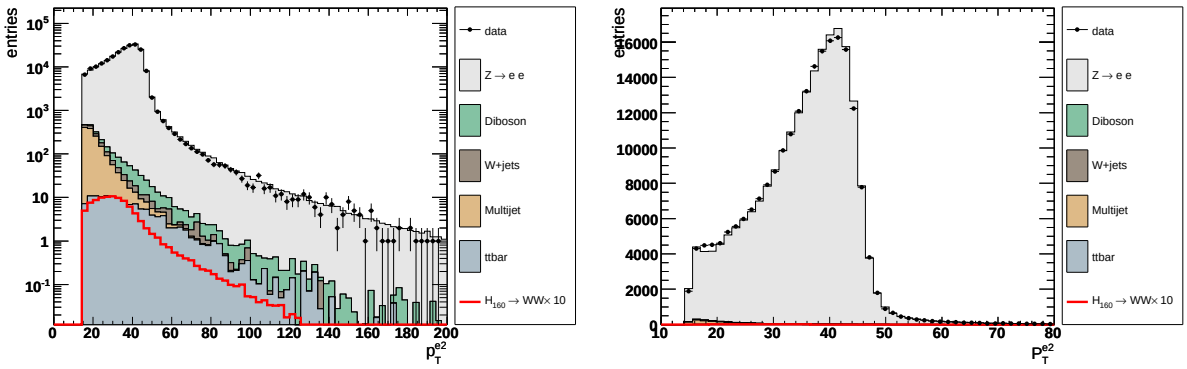


Figure 6.11: Distribution of the transverse momentum for the trailing p_T electron in logarithmic scale (left) and linear scale (right) at preselection level for data (points) and sum of all backgrounds (filled histograms). The expected signal, multiplied by a factor of 10, for a 160 GeV/c^2 Standard Model Higgs is also shown. The entire Run II data set is used.

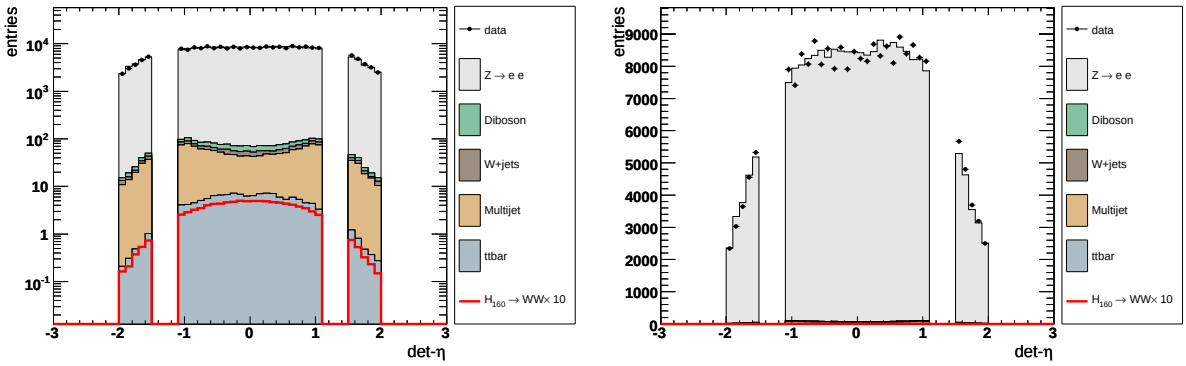


Figure 6.12: Distribution of the detector η for the leading p_T electron in logarithmic scale (left) and linear scale (right) at preselection level for data (points) and sum of all backgrounds (filled histograms). The expected signal, multiplied by a factor of 10, for a 160 GeV/c^2 Standard Model Higgs is also shown. The entire Run II data set is used.

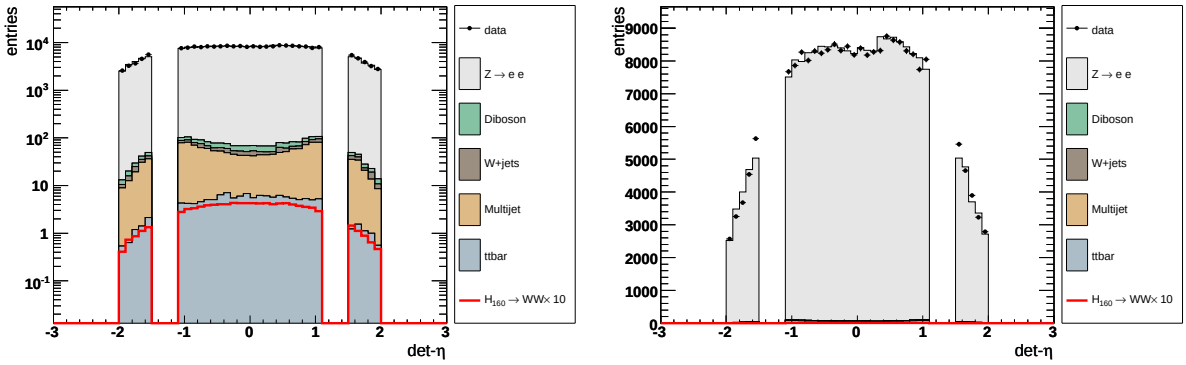


Figure 6.13: Distribution of the detector η for the trailing p_T electron in logarithmic scale (left) and linear scale (right) at preselection level for data (points) and sum of all backgrounds (filled histograms). The expected signal, multiplied by a factor of 10, for a 160 GeV/c^2 Standard Model Higgs is also shown. The entire Run II data set is used.

6.6 Multijet Background in the e^+e^- final state

The main source of instrumental background arises from QCD multijet events in which jets are misidentified as electrons. This happens mostly when a jet is formed with most of its energy being carried by an isolated π^0 or η which then decays into a pair of spatially close photons. Such a photon pair might be mis-identified as electron or photon in the calorimeter and therefore be reconstructed as a single photon and pass the loose electron quality identification requirements. In addition tracks can be associated with the EM object formed by the photon pair, either by charged hadrons from the jet forming tracks that overlap with the EM object, particularly for loose lepton quality criteria, or by one of the photons converting to e^+e^- that will form tracks that can be

associated with the EM object. Such events can possibly pass the track-matched electron quality identification requirements.

The background contribution of multijet production was determined directly from data. Events containing at least one loose electron are used for estimating the multijet background. To get an estimation of the multijet contribution the electron fakes are obtained by requiring EM-objects with HMatrix7 $\chi^2 > 35$ and no likelihood requirement for the electron candidate. Track matching for the selected electron candidates is a crucial requirement of the EM-likelihood, thus by discarding the likelihood requirement track-matching is discarded as well. Besides that all standard selection requirements and p_T thresholds are applied. A sample containing a very small signal contribution and enriched in QCD multijet and W + jets is selected by requiring the two lepton candidates to have the same charge. This sample is referred to as fake sample. This fake sample is used to normalize the multijet contribution in the signal sample, a method utilized and described also in Ref. [84]. The number of like-sign events $N^{\pm\pm}_{QCD}$, corresponding to the multi-jet contribution in the like-sign signal sample, is computed from the excess of like-sign events $N^{\pm\pm}_{data}$ above the expected contribution $N^{\pm\pm}_{MC}$ from all Standard Model backgrounds:

$$N^{\pm\pm}_{QCD} = N^{\pm\pm}_{data} - N^{\pm\pm}_{MC}.$$

Because the kinematic turn-ons of the p_T spectra differ between like-sign and opposite-sign a p_T dependent correction is derived. The normalization factor for the QCD sample is defined as

$$f^i_{QCD}(p_T) = \frac{N^{\pm\pm}_{data}(p_T) - N^{\pm\pm}_{MC}(p_T)}{N^{\pm\pm}_{fake}(p_T)},$$

Where $N^{\pm\pm}_{fake}$ corresponds to the number of the like-sign events in the fake sample, the i corresponds to the electron which is used for deriving the correction. The number of multi-jet events in the unlike-sign signal sample $N^{\pm\mp}_{QCD}$ is determined from the number of unlike-sign fake events $N^{\pm\mp}_{fake}(p_T)$ using the formula

$$N^{\pm\mp}_{QCD} = f^i_{QCD}(p_T) \cdot N^{\pm\mp}_{fake}(p_T).$$

These corrections are applied first for the leading p_T lepton and then for the trailing p_T electron. The correction function for the trailing lepton is derived after applying the leading lepton corrections. The correction curves can be found in Figs. 6.14 and 6.15. One sees a significant difference of the scale factors for the leading p_T lepton. This is caused by the upgrade of the Layer 0 in the tracker and improved trigger, therefore an decreasing the fake rate.

These $f^i_{QCD}(p_T)$ ratios are determined at preselection stage (see Sec. 6.5) and kept constant throughout the analysis. It is assumed that the ratio of like-sign and unlike-sign events in the fake sample and in the multijet contribution to the search sample is the same [84]. The like sign data sample shows a significant Z-peak, caused by charge-misidentification. This Z peak has been subtracted from the like-sign sample prior to performing the fit. The correction factor needed to match the MC to data for $Z \to ee$ is 2.7.

The overall multijet contribution at preselection stage is at the few percent level due to the requirement of a track-matched electron candidates, stringent isolation requirements and the selection of oppositely charged leptons. One sees in Fig. 6.15 that the multijet contribution in the Run IIb data set has decreased significantly due to the improved track matching.

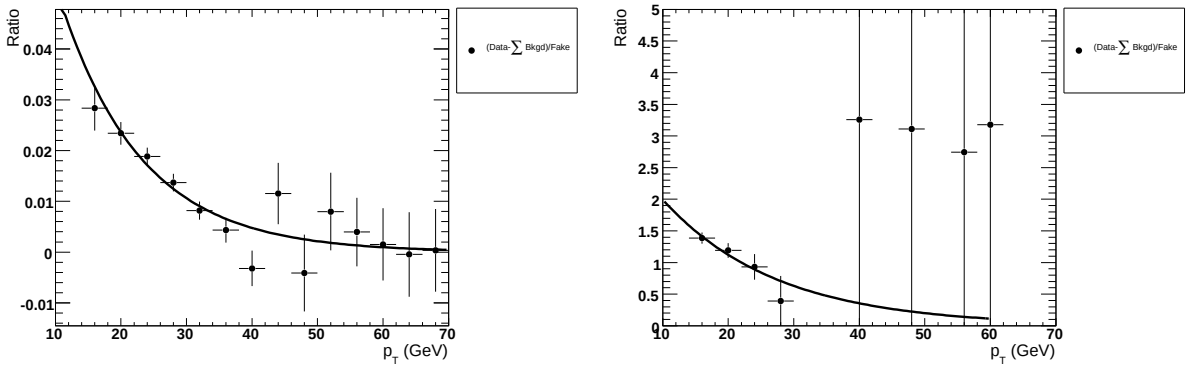


Figure 6.14: Fit of the exponential normalization function for QCD background for Run IIa as a function of the transverse momentum of the leading p_T lepton (left) and the trailing p_T lepton (right). Each fit is performed in the like-sign sample, correcting the leading p_T electron first and then the trailing.

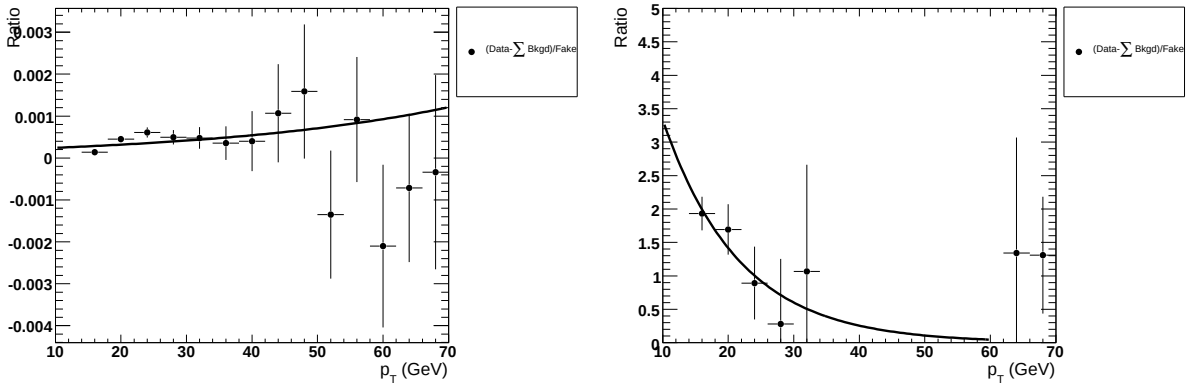


Figure 6.15: Fit of the exponential normalization function for QCD background for Run IIb as a function of the transverse momenta of the leading p_T lepton (left) and the trailing p_T lepton (right). Each fit is performed in the like sign sample, correcting the leading p_T electron first and then the trailing.

6.7 Electroweak Reference Sample

The production cross section for the $W+$jets process at the Tevatron is about 2600 pb^{-1}. Although the stringent selection requirements of two leptons and the following event selection (see Sec. 7.1) are very efficient in suppressing background processes, the $W+$jets production remains one of the more challenging backgrounds for the present analysis. Therefore several cross-checks have been performed to ensure a proper description of this crucial background. An independent sample has been selected in which the overall contribution of $W \to \ell\nu$ events is significantly enriched. In order to increase the limited statistics the opposite sign requirement is abandoned. Additional selection requirements are applied to suppress all other background processes apart from $W+$jets . By doing so the normalization and description of the $W+$jets sample can cross checked and adapted. These cuts are:

- By requiring the minimum lepton quality of either one of the leptons, $\mathcal{Q}_{min}(e_1, e_2)$, to be less than 0.9, diboson, $Z+$jet/γ and signal samples are suppressed.
- By requiring 70 GeV $< M_{inv}$ and $M_{inv} > 100$ GeV further $Z+$jet/γ events are removed.
- By requiring $M^T_{min}(\ell, \not\!\!E_{\rm T}) > 30$ GeV, $\not\!\!E_{\rm T} > 20$ GeV and $\not\!\!E_T^{Sig} > 5$ one removes mismeasured events from multijet production.

The minimum transverse mass $M^T_{min}(\ell, \not\!\!E_{\rm T})$ and missing transverse energy $\not\!\!E_{\rm T}$ are defined in Sec. 7.1. The resulting sample is dominated by $W+$jets events and diboson WW production. This selection reveals a significant difference between the number of expected and observed events. The correction factor needed to match the data with Monte Carlo has been estimated by scaling the Monte Carlo expectation to data and was found to be 1.3. This factor is applied to correct the discussed cross-section corrections to all ALPGEN $W+$jets background processes. Applying this scale factor leads to consistent behavior in a subset of the data with previous versions of this analysis using PYTHIA generated Monte Carlo events. Recent studies of these background samples have confirmed this correction factor and are discussed in Ref. [85]. As shown in Figs. 6.16 to 6.19, shape and scaling of W-like backgrounds are well described by the Monte Carlo simulation after application of the corrections. All variables entering the neural network (see Sec. 7.2) are shown as well. The yields of the various background contributions are shown in Table 6.7.

Data	Sum Bkgd	$H_{160} \to WW$	$Z \to ee$	$Z \to \tau\tau$	
189.0 ± 13.8	191.1 ± 11.9	1.3 ± 0.0	41.4 ± 10.8	0.6 ± 0.3	
$W +$ jets	$t\bar{t}$	ZZ	WZ	WW	QCD
105.6 ± 3.3	14.7 ± 0.5	0.2 ± 0.0	1.2 ± 0.0	21.5 ± 0.9	5.9 ± 3.7

Table 6.7: Number of expected background events and number of observed events for the $W +$ jets enriched sample. Only statistical uncertainty is given for all backgrounds. The entire Run II data set is used.

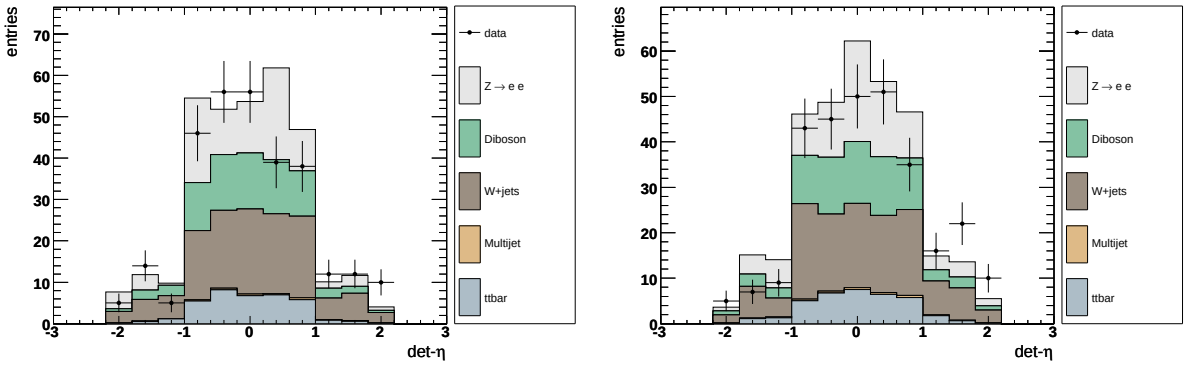


Figure 6.16: Distribution of η_{det} for both electrons for data (points) and sum of all backgrounds (filled histograms) for a $W +$ jets enriched sample. The entire Run II data set is used.

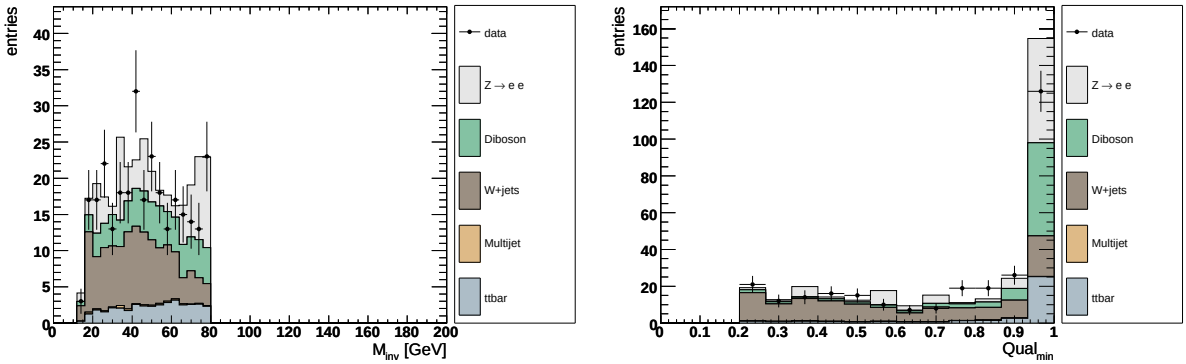


Figure 6.17: Distribution of the invariant mass M_{inv} (left) and the minimum quality of either of the leptons (right) for data (points) and sum of all backgrounds (filled histograms) for a $W +$ jets enriched sample. The entire Run II data set is used.

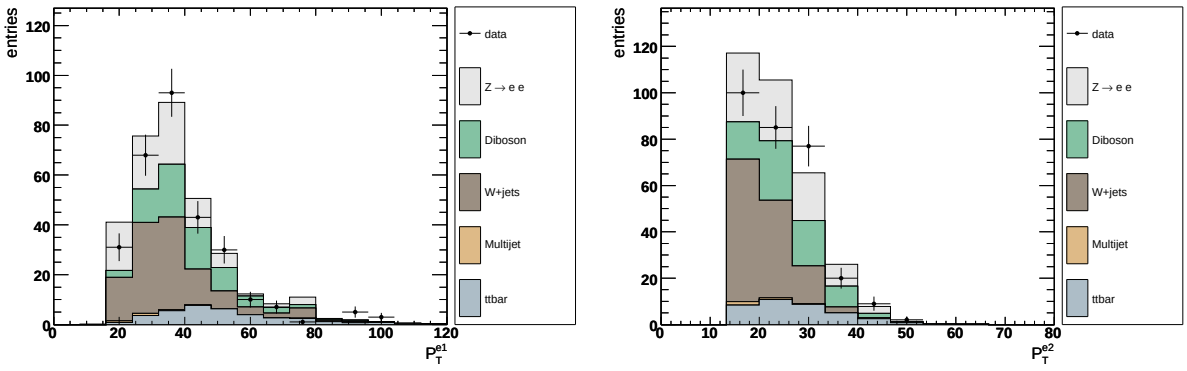


Figure 6.18: Distribution of the transverse momentum for the leading p_T electron and trailing p_T electron for data (points) and sum of all backgrounds (filled histograms) for a $W +$ jets enriched sample. The entire Run II data set is used.

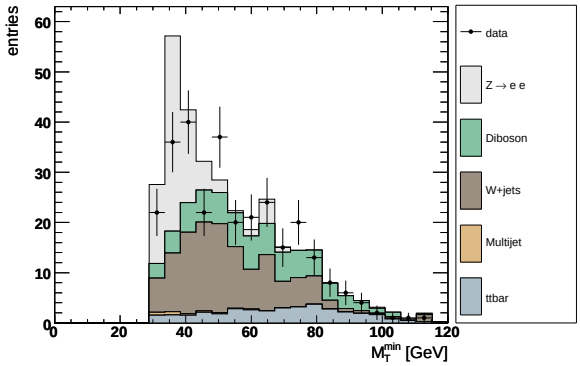


Figure 6.19: Distribution of the minimal transverse mass for data (points) and sum of all backgrounds (filled histograms) for a $W +$ jets enriched sample. The entire Run II data set is used.

6.8 Monte Carlo Corrections

Due to limitations in the knowledge of the underlying theory, the precision of perturbative calculations and the complex modeling of the DØ detector response, Monte Carlo generated events do not reproduce the data perfectly. Therefore correction factors are derived from data and applied to the simulation. Generally two types of corrections can be distinguished:

- *Modifications of the shape of distributions*: Monte Carlo generators use often leading order matrix elements for the hard interactions. Higher order corrections can result in differences in the cross section but also in the shape of the resulting distributions. In addition various tunable parameters in the event can alter the shape of the distributions of physical quantities. Important here is that the Monte Carlo samples are overlayed with zero bias events with the intent of making the simulated events more like the detector data. Moreover, the zero bias overlay, used to improve the modeling of energy offsets due to pile-up events, minimum bias interactions and luminosity effects is recorded in distinct time periods. However, the zero bias events were not collected with the same luminosity spectrum as the data, which could result in small difference between the data and MC. Therefore each of the MC samples has been re-weighted such that is shows the same luminosity profile as the data set. This class of reweightings preserves the normalization of samples.

- *Modifications affecting the event yield*: Inefficiencies in detecting and triggering particles are not modeled perfectly in Monte Carlo. Trigger efficiencies are even not modeled at all. This efficiencies usually show a strong dependency of the η-region in the detector, the particle momenta, z-position of the hard interaction or other possible detector related quantities. The simulation is corrected for these efficiency functions which are measured in data. These corrections are not preserving the normalization.

Lepton Identification Corrections

The efficiency to reconstruct, identify and select leptons is higher in Monte Carlo events than in data. This behavior is caused mostly by tracking inefficiencies, increasing in size as more stringent lepton quality criteria are applied. These efficiencies are measured individually for any given lepton type and quality [46, 86]. Monte Carlo to data correction factors are then parametrized as function of the pseudorapidity η and ϕ and the simulated leptons are corrected accordingly.

Lepton Resolution Corrections

The muon and electron energy resolution are not the same in data and Monte Carlo. This is caused mostly by poorly simulated dead material and deviations in alignment and geometry. In addition the effect of temporarily dead readout channels is not properly handled. The energy resolution in Monte Carlo is matched to the one in data by applying an additional smearing. The amount of smearing necessary is tuned according to the observed width of the Z boson.

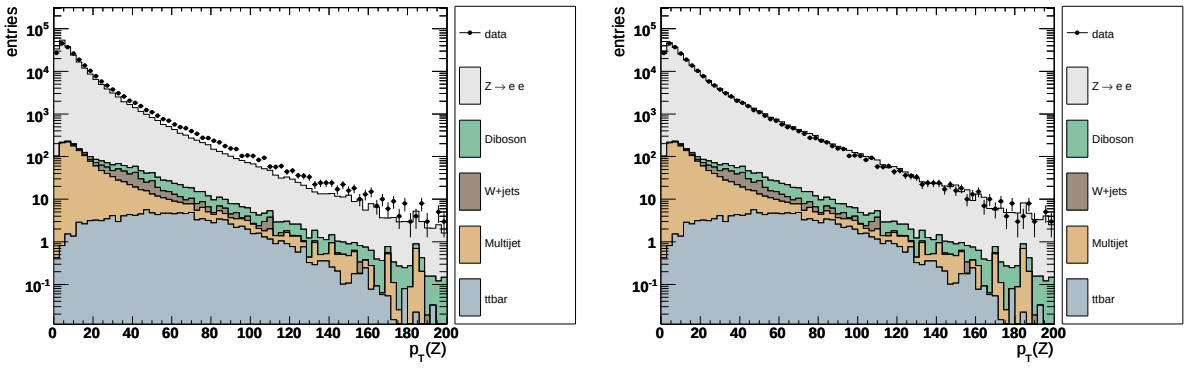


Figure 6.20: Distribution of $p_{\mathrm{T}}(Z)$ before reweighting (left) and after reweighting (right). The entire Run II data set is used.

$p_T(Z)$ re-weighting

Because the p_T distribution of the Z boson is not perfectly described in PYTHIA, the $Z/\gamma^* \to \ell\ell$ Monte Carlo simulations are corrected for this effect by applying a weight to each Monte Carlo event according to the results given in Ref. [87], based on the measurement of the differential Z boson production cross section as a function of p_T as obtained from $Z \to ee$ events in Run IIa and Run IIb data. The ratio of the tuned versus default Monte Carlo is fitted using a modified Fermi function. The fits are performed in three different mass bins (15 GeV$< M_{ee} <$60 GeV, 60 GeV$< M_{ee} <$130 GeV and 130 GeV$< M_{ee} <$250 GeV). The effect of the $p_T(Z)$ reweighting is shown in Fig. 6.20.

Instantaneous Luminosity re-weighting

The Monte Carlo is overlaid with zero bias (ZB) events. These events are recorded using a special minimum bias trigger associated with non-diffractive interactions. The Monte Carlo simulation uses zero bias events to reproduce the effects of additional interactions in the same bunch crossing which may affect for example the determination of quantities like the lepton isolation. However, the zero bias events used for the overlay were recorded in five distinguished data sets. Due to the continuously changing and increasing instantaneous luminosity delivered by the Tevatron the luminosity profiles of these five data sets do not correspond accurately to the luminosity profile of the data. Therefore the distribution of the instantaneous luminosity of the MC is reweighted accordingly. As the zero bias events are used to reproduce detector and luminosity effects, a good description of the instantaneous luminosity in Monte Carlo is crucial. The procedure contains three steps, measuring the luminosity profile for any given MC set, measuring the luminosity profile of the data set and finally reweighting the MC according to the data/MC profiles ratio. The luminosity profile before and after correction can be found in Figs. 6.21 and 6.22. In Fig. 6.21 one sees a wavy structure of the instantaneous luminosity distributions caused by the limited zero bias events available.

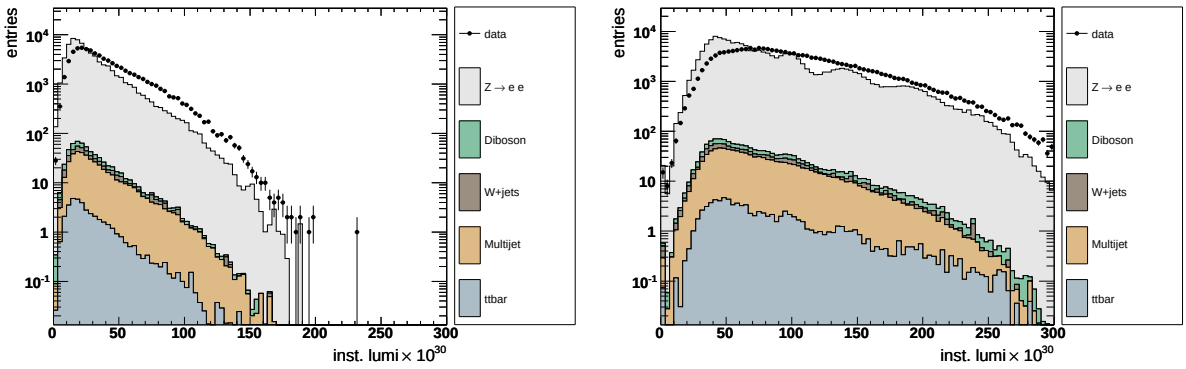


Figure 6.21: Instantaneous luminosity profile of Run IIa (left) and Run IIb (right) before corrections.

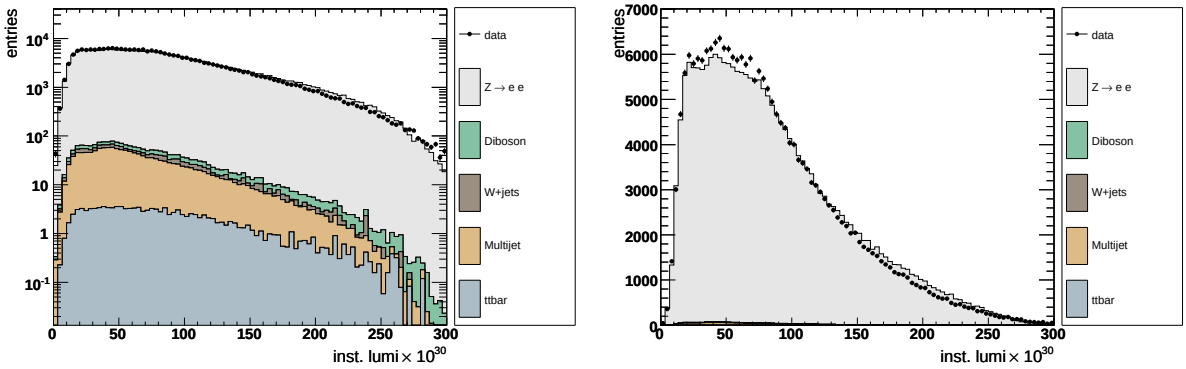


Figure 6.22: Distribution of the instantaneous luminosity profile for the entire Run II data set (Run IIa + Run IIb) after application of the Monte Carlo correction. The left distributions is plotted in logarithmic scale whereas the right one gives the luminosity profile in linear scale. The remaining inaccuracy has no impact on the analysis.

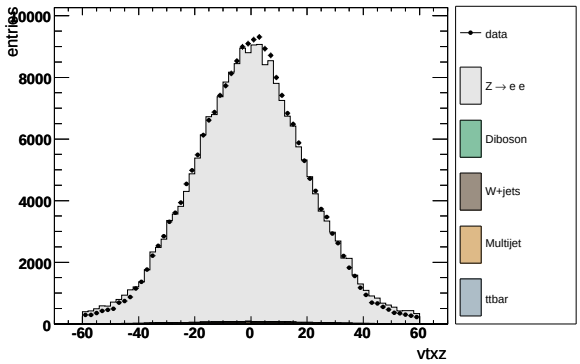


Figure 6.23: The z distribution of the vertex for the entire Run II data set after applying corrections. The entire Run II data set is used.

Vertex z Reweighting

The shape of the interaction region at DØ depends on shapes of the p and $\bar{p}$ bunches and on the beta parameter of the interaction point. In simulated events the z coordinate of the primary vertex is distributed according to a Gaussian shape and a width of 25 cm and centered around $z = 0$. Whereas in data the shape of the luminosity region and the primary vertex differ, partly depending on the instantaneous luminosity, but as well show a wider z range with a loss in efficiency for $|z| > 40$ cm. The characteristics and the shape of the luminous region have been measured. In order to improve the agreement between the z vertex distribution of the Monte Carlo and the data an appropriate correction is applied [88]. The distribution of vertex z after applying the correction is shown in Fig. 6.23.

7
The $H \rightarrow WW^* \rightarrow e^+\nu e^-\nu$ Channel

At the stage of the preselection as explained in Sec. 6.5 about 220.000 background events are expected with a signal expectation of less than 10 events. Obviously the signal-to-background ratio S/B is insufficient to make any statement about the existence of the Higgs boson in the selected sample. Therefore appropriate selection requirements to suppress the various background contributions and enrich the signal content have to be applied. This analysis follows a two folded approach. First "cuts", requirements on distributions of variables are placed, then a multivariate method is trained and applied as described in Chapter 7.2.

7.1 Event Selection for the $H \rightarrow WW^* \rightarrow e^+\nu e^-\nu$ final state

In order to gain maximal sensitivity only loose kinematic cuts are applied. These selection requirement are:

- **Cut1:** By rejecting electron pairs with a wide opening angle the major part of the $Z \rightarrow \ell\ell$ background is removed.
- **Cut2:** Exploiting the presence of the two neutrinos in the signal sample a missing transverse momentum of more than $\not\!\!E_{\mathrm{T}} > 20$ GeV is required. This suppresses further the $Z + \text{jet}/\gamma$ background and rejects a large fraction of the multijet events.
- **Cut3:** Another very useful quantity is the $\not\!\!E_{\mathrm{T}}$ significance ($\not\!\!E_T^{Sig}$) which is a measure of the energy mis-measurement of the jets. It can be determined by projecting the probable energy fluctuations of a given jet onto the direction of the missing transverse energy and performing the sum over all jets in an event:

$$\not\!\!E_T^{Sig} = \frac{\not\!\!E_{\mathrm{T}}}{\sqrt{\sum_{jets} \sigma^2_{E_T(jets)||\not\!\!E_{\mathrm{T}}}}} \tag{7.1}$$

  here σ corresponds to an experimentally determined resolution for jets [28]. This provides a measure to which extend the mis-measurements of jet-energies contribute to the $\not\!\!E_{\mathrm{T}}$.
- **Cut4:** The last selection cut requires a minimal transverse mass of the system formed out of either of the leptons and the $\not\!\!E_{\mathrm{T}}$: $M^T_{min}(\ell, \not\!\!E_{\mathrm{T}}) > 30$ GeV. This reduces further contributions from from Z/γ^* and $W \rightarrow \ell\nu$ events.

After application of the last of these requirement one obtains a relatively clean sample dominated by background processes involving W decays. Tab. 7.1 summarizes all selection requirements.

Cut 0	Preselection	lepton ID, leptons with opposite charge and $p_T^{e_{1,2}} > 15$ GeV Single EM Trigger OR, $M_{ee} > 15$ GeV
Cut 1	$\Delta\phi(\ell,\ell)$	< 2.0
Cut 2	Missing Transverse Energy $\not\!\!E_{\mathrm{T}}$	> 20.0 GeV
Cut 3	$\not\!\!E_T^{Sig}$	> 6.0
Cut 4	$M_{min}^T(\ell, \not\!\!E_{\mathrm{T}})$	> 30.0 GeV

Table 7.1: Summary of selection criteria applied in the analysis.

This selection reduced the overall expected background by three orders of magnitude with respect to the preselection stage while showing an efficiency of about $\sim 75\%$ for signal events as seen in Table 7.5 and Tables 7.2, 7.3 and 7.4. The number of expected and observed events agree well. The remaining sample consists mainly of $W+$jets , $t\bar{t}$ and WW production. These processes cannot be easily reduced by applying selection criteria on typical distributions without affecting strongly the signal yield. Therefore as second approach an Artificial Neural Network (ANN or NN), described in Sec. 7.2, is used to further separate the signal and backgrounds processes.

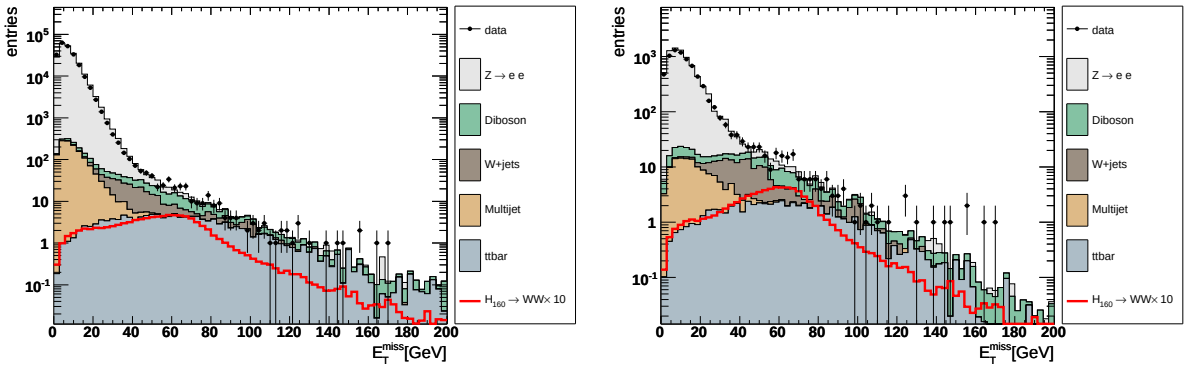


Figure 7.1: Distribution of the missing transverse energy $\not\!\!E_T^{Sig}$ at the preselection level (left) and after the $\Delta\phi(\ell,\ell)$ cut (right). The expected signal, multiplied by a factor of 10, for a 160 GeV/c^2 Standard Model Higgs is also shown. The entire Run II data set is used.

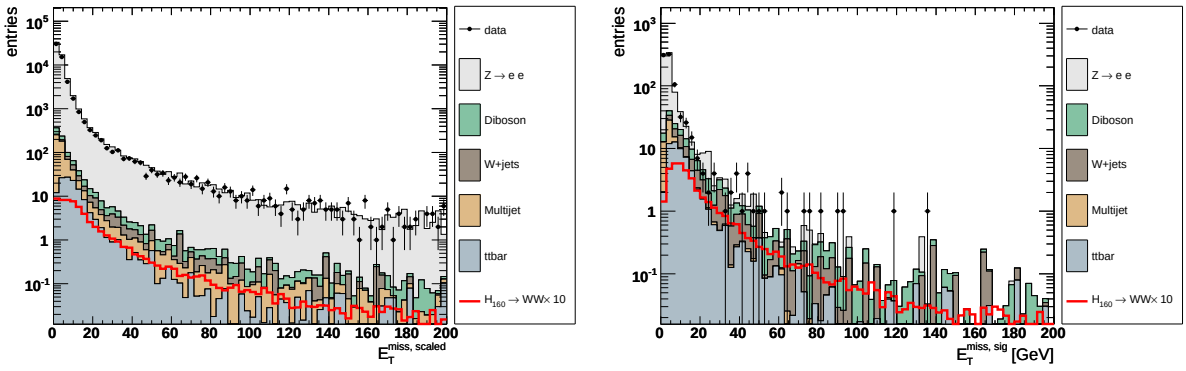


Figure 7.2: Distribution of the missing transverse energy significance after preselection (left) and after the $\not\!\!E_T$ cut is applied (right). The expected signal, multiplied by a factor of 10, for a 160 GeV/c^2 Standard Model Higgs is also shown. The entire Run II data set is used.

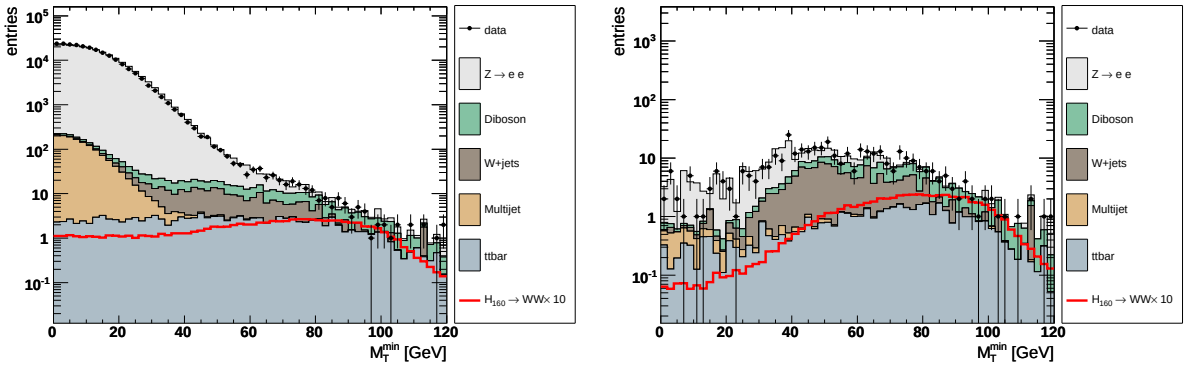


Figure 7.3: Distribution of the minimal transverse mass after preselection (left) and after the $\not\!\!E_T^{Sig}$ cut is applied (right). The expected signal, multiplied by a factor of 10, for a 160 GeV/c^2 Standard Model Higgs is also shown. The entire Run II data set is used.

7.1.1 Marginal Distributions

This section presents supplementary plots, the marginal distributions. The marginal distribution of a quantity is defined as the distribution obtained when applying all selection requirements except the one involving the quantity of interest. These distribution shows the complementary set of events which are removed by the event selection. They are shown Figs. 7.4, 7.5, 7.6 and Fig. 7.7 for each of the selection requirements.

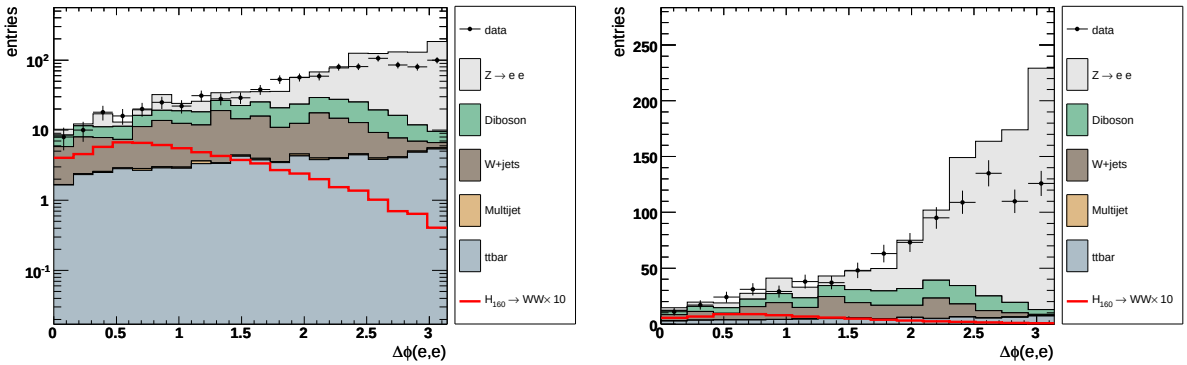


Figure 7.4: Distribution of the angle $\Delta\phi(\ell,\ell)$ after applying all cuts except of the requirement of $\Delta\phi(\ell,\ell)$ itself in logarithmic scale (left) and linear scale (right). The expected signal, multiplied by a factor of 10, for a 160 GeV/c^2 Standard Model Higgs is also shown. The entire Run II data set is used.

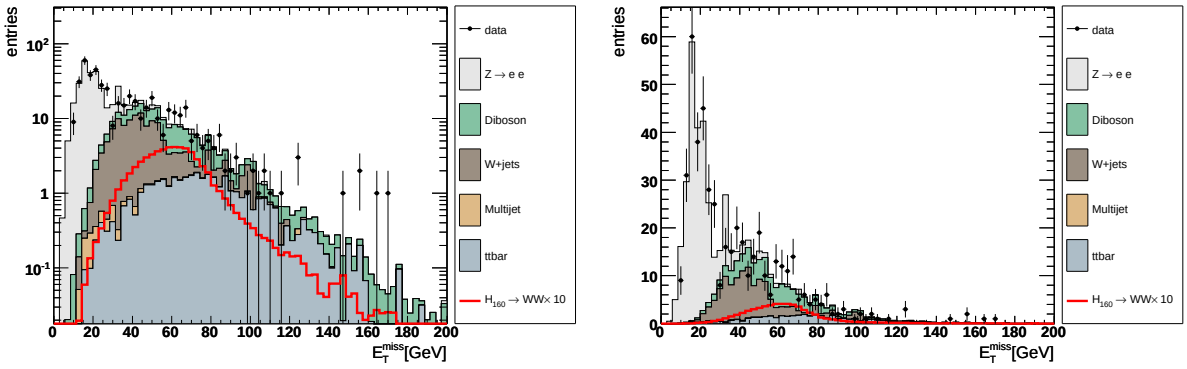


Figure 7.5: Distribution of the $\not\!\!E_{\mathrm{T}}$ after applying all cuts except of the requirement of $\not\!\!E_{\mathrm{T}}$ itself in logarithmic scale (left) and linear scale (right). The expected signal, multiplied by a factor of 10, for a 160 GeV/c^2 Standard Model Higgs is also shown. The entire Run II data set is used.

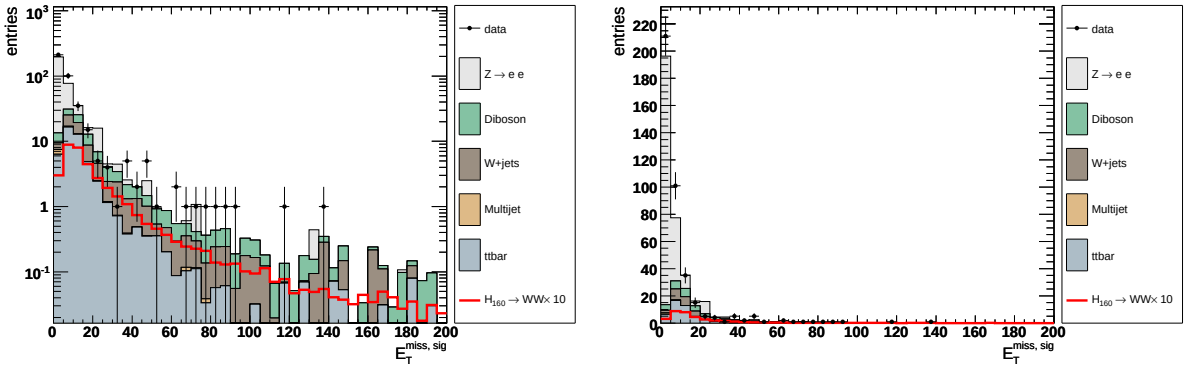


Figure 7.6: Distribution of the $\not\!\!E_T^{Sig}$ after applying all cuts except of the requirement of $\not\!\!E_T^{Sig}$ itself in logarithmic scale (left) and linear scale (right). The expected signal, multiplied by a factor of 10, for a 160 GeV/c^2 Standard Model Higgs is also shown. The entire Run II data set is used.

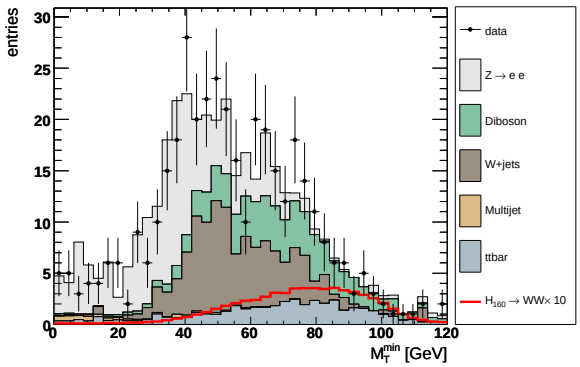


Figure 7.7: Distribution of the minimal transverse mass after applying all cuts except of the requirement of $M_{min}^T(\ell, \not\!\!E_T)$ itself in linear scale. The expected signal, multiplied by a factor of 10, for a 160 GeV/c^2 Standard Model Higgs is also shown. The entire Run II data set is used.

One observes a significant discrepancy, shown in Fig. 7.4. By definition these marginal distributions enhance contributions which are removed before entering the final selection. Nevertheless it is important to understand the sources of discrepancies. It should be noted that the discrepancy is not visible on preselection stage as shown in Fig. 7.8.

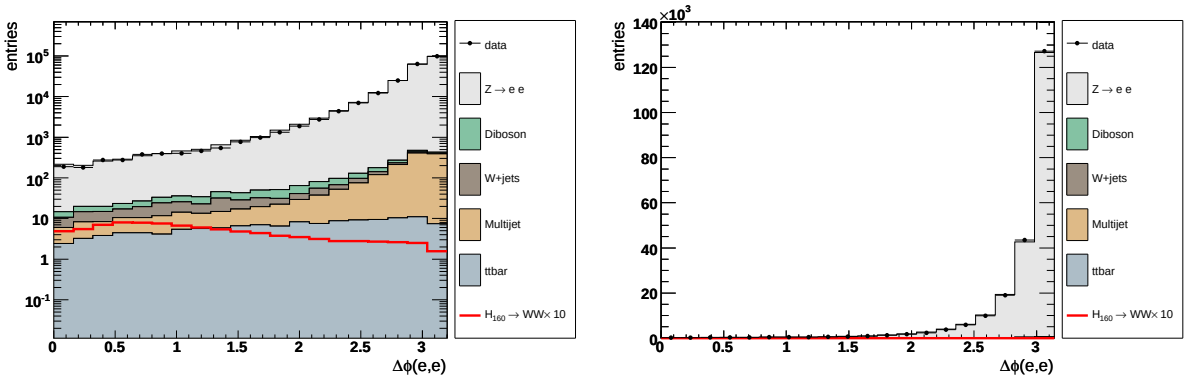


Figure 7.8: Distribution of $\Delta\phi(\ell,\ell)$ at preselection stage in logarithmic (left) and linear (right) scale. One sees that Monte Carlo simulations describe the data well. The entire Run II data set is used.

Further studies have shown that these events correspond to events with low vectorial sums $\vec{p}_T^{\,1} + \vec{p}_T^{\,2}$ and high invariant masses around 90 GeV. Furthermore this mis-description is only visible in samples corresponding to Run IIb. Figure 7.9 shows the distribution of the invariant mass $M_{\ell\ell}$ and the $\vec{p}_T^{\,1} + \vec{p}_T^{\,2}$ when applying all analysis selection requirements except of the cut on $\Delta\phi(\ell,\ell)$. It is evident that the excess is most likely due to a mis-description of a small fraction of the $Z \rightarrow \ell\ell$ events. The excess is affecting less than one percent of the events selected at the preselection stage in the Z peak. Because these events do not pass the preselection the final result is not affected.

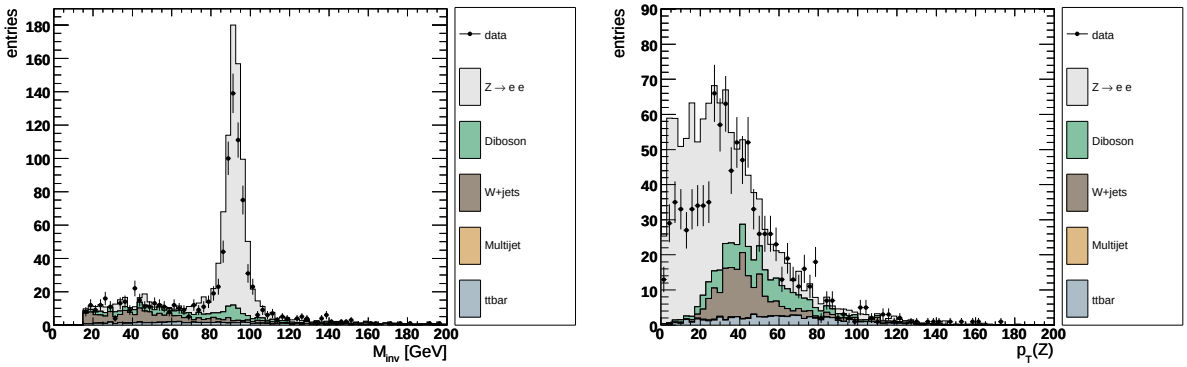


Figure 7.9: Distribution of the $M_{\ell\ell}$ (left) and the vectorial sum $\vec{p}_T^{\,1} + \vec{p}_T^{\,2}$ of the electron momenta (right) when applying all cuts except of the requirement on $\Delta\phi(\ell,\ell)$ itself. The entire Run II data set is used.

Selection	$m_H = 115$	$m_H = 120$	$m_H = 125$	$m_H = 130$	$m_H = 135$	$m_H = 140$
Preselection	2.3	2.9	4.0	5.1	6.1	6.9
$\Delta\phi(\ell,\ell)$	1.3	1.8	2.5	3.4	4.2	4.8
$\not\!\!E_{\mathrm{T}}$	1.1	1.5	2.1	2.9	3.6	4.2
$\not\!\!E_T^{Sig}$	0.8	1.2	1.7	2.4	3.1	3.6
$M_{min}^T(\ell, \not\!\!E_{\mathrm{T}})$	0.7	1.1	1.6	2.3	2.9	3.4

Table 7.2: Overview of expected signal events for an integrated luminosity of 4.2 fb^{-1} for m_H = 115-140 GeV in 5 GeV steps. Due to the high Monte Carlo statistic no statistical uncertainties are given.

Selection	$m_H = 145$	$m_H = 150$	$m_H = 155$	$m_H = 160$	$m_H = 165$	$m_H = 170$
Preselection	8.1	8.5	9.2	9.5	9.2	8.4
$\Delta\phi(\ell,\ell)$	5.8	6.1	6.8	7.4	7.4	6.7
$\not\!\!E_{\mathrm{T}}$	5.1	5.4	6.2	6.9	6.9	6.3
$\not\!\!E_T^{Sig}$	4.5	4.7	5.4	6.2	6.3	5.7
$M_{min}^T(\ell, \not\!\!E_{\mathrm{T}})$	4.3	4.6	5.3	6.1	6.2	5.6

Table 7.3: Overview of expected signal signal events for an integrated luminosity of 4.2 fb^{-1} for m_H = 145-170 GeV in 5 GeV steps. Due to the high Monte Carlo statistic no statistical uncertainties are given.

7.1.2 Event Yield

The number of candidate events observed and number of background events expected at each step of the application of the selection requirements are shown in Table 7.5 for a Higgs mass of $m_{\mathrm{H}} = 160$ GeV. After applying all selection criteria the data are in agreement with the SM expectation. Event yields for $m_{\mathrm{H}} = 160$ GeV are be found in Table 7.5. An overview of the signal yields for all studied Higgs masses are given Tables 7.2, 7.3 and Table 7.4.

Selection	$m_H = 175$	$m_H = 180$	$m_H = 185$	$m_H = 190$	$m_H = 195$	$m_H = 200$
Preselection	7.8	7.1	6.4	5.7	5.0	4.6
$\Delta\phi(\ell,\ell)$	6.1	5.4	4.6	3.9	3.4	3.0
$\not\!\!E_{\mathrm{T}}$	5.7	5.0	4.2	3.6	3.1	2.7
$\not\!\!E_{\mathrm{T}}^{Sig}$	5.1	4.5	3.8	3.2	2.7	2.4
$M_{min}^{T}(\ell, \not\!\!E_{\mathrm{T}})$	5.0	4.4	3.7	3.1	2.6	2.3

Table 7.4: Overview of expected signal events for an integrated luminosity of 4.2 fb^{-1} for m_H = 175-200 GeV in 5 GeV steps. Due to the high Monte Carlo statistic no statistical uncertainties are given.

Cut	Data	Sum Bkgd	$Z \rightarrow ee$	$Z \rightarrow \tau\tau$
Preselection	221530.0 ± 470.7	221936.6 ± 706.6	218695.7 ± 704.3	1134.9 ± 15.9
$\Delta\phi(\ell,\ell)$	7042.0 ± 83.9	7733.8 ± 454.8	7268.3 ± 454.5	41.2 ± 3.3
$\not\!\!E_{\mathrm{T}}$	1034.0 ± 32.2	1053.7 ± 112.2	723.4 ± 111.8	32.6 ± 2.7
$\not\!\!E_{\mathrm{T}}^{Sig}$	386.0 ± 19.7	389.1 ± 17.8	140.4 ± 17.2	12.1 ± 1.6
$M_{min}^{T}(\ell, \not\!\!E_{\mathrm{T}})$	336.0 ± 18.3	332.5 ± 14.5	107.7 ± 14.0	1.4 ± 0.5

Cut	W + jets	$t\bar{t}$	ZZ	WZ	WW	QCD
Preselection	240.8 ± 4.7	131.4 ± 1.4	98.2 ± 0.2	112.5 ± 0.2	172.2 ± 2.6	1351.1 ± 55.1
$\Delta\phi(\ell,\ell)$	112.0 ± 3.2	65.2 ± 1.0	31.8 ± 0.1	36.4 ± 0.1	71.6 ± 1.7	107.3 ± 15.5
$\not\!\!E_{\mathrm{T}}$	106.3 ± 3.2	61.2 ± 1.0	12.4 ± 0.1	15.4 ± 0.1	70.1 ± 1.6	32.3 ± 8.5
$\not\!\!E_{\mathrm{T}}^{Sig}$	100.5 ± 3.2	45.8 ± 0.8	8.00 ± 0.1	10.2 ± 0.1	67.6 ± 1.6	4.6 ± 3.2
$M_{min}^{T}(\ell, \not\!\!E_{\mathrm{T}})$	97.7 ± 3.2	39.9 ± 0.8	7.7 ± 0.1	9.7 ± 0.1	66.8 ± 1.6	1.7 ± 2.0

Table 7.5: Number of expected background events and number of observed events, after successive selection requirements. Only statistical uncertainties are given for all backgrounds. The entire Run II data set is used, corresponding to an integrated luminosity of 4.2 fb^{-1}

7.2 Neural Network

In order to further separate the remaining backgrounds and signal are Neural Network approach is used. The following sections discuss structure and training of the neural network and the choice of input variables.

7.2.1 Structure of Neural Networks

The requirements discussed in Sec. 7.1 remove 99.8 % of the total sum of expected background events while preserving 75 % of signal events. Nevertheless the sensitivity to an eventual Higgs boson would be greatly enhanced by further improved background rejection. However, the characteristics of the background prevent the use of additional selection requirements on the distributions of individual kinematic quantities. Therefore a multivariate method approach is used. Multivariate statistical methods are designed to evaluate more than one variable at a time. After comparing the perfor)ecision Trees,

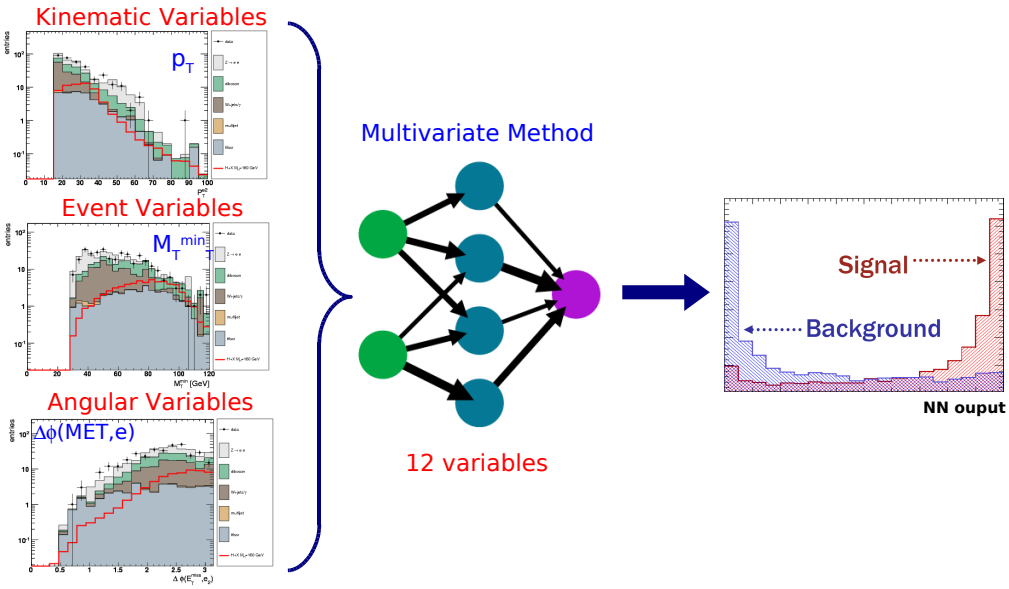


Figure 7.10: Principle of a Neural Network (NN). A given set of input variables offering separation power between background and signal is used to train the network. The resulting output distribution after successful training should assign low values to background like events and high NN output values to signal like events.

Neural networks are a multivariate technique commonly used in high energy physics. They have been applied in LEP analysis and both Run I and Run II analyses at the Tevatron. They are also often used for object identification techniques such as electron identification or b-tagging ([89], [64], [63]). They can be described as parametrized nonlinear functions for regression or classification modeling. Inputs to a neural network are variables which provide discrimination power

between signal and background. Every network used in this analysis consists of three layers of nodes, an input layer, a hidden layer and an output layer. A sigmoid function from the sum of the weighted input variables is calculated at each hidden node. The linear sum of these sigmoid functions appears at the output node. A neural network is trained using samples of simulated signal and background events. During the training process, weights are adjusted at each node such that the signal is moved towards one and the background towards zero. A complete cycle of running through the entire training sample is called an "epoch". The complete training cycle used in this analysis has 600 epochs.
The Artificial Neural Network (ANN) used is the `TMultiLayerPerceptron` implementation, part of the `ROOT` package [90]. For training and analysis the available MC samples are divided into orthogonal samples of equal size. The training samples are divided again, using one half for the training itself and the other one for testing of the network training. For each analyzed Higgs mass a separate neural network is trained, using all background processes along with their weights for normalization and efficiency corrections. The training is performed using slightly looser selection requirements to maintain high statistics for the training and control samples. Therefore the requirement on $\not\!\!E_T^{Sig}$ was lowered from 6 to 5.

7.2.2 Neural Network Variable Selection

A set of input quantities has been selected for the NN which show separation power between the signal and at least one of the backgrounds types. These quantities can be categorized into three different classes: object kinematics, event kinematics and angular variables. The input variables of the NN are listed in Table 7.6.

The relative importance of individual input quantities changes significantly over the analyzed mass range as the background composition and the kinematic properties of the signal processes change. This effect is mostly driven by change of importance of the $W+$ jets and WW production which are the most important background types at low and high values of m_H respectively. For lower Higgs masses the $W+$ jets background becomes more dominant whereas WW production is more relevant at high and intermediate masses.

There are altogether 12 different input variables used. The input variables are well described by the Monte Carlo and the MC/Data comparisons are shown in Fig. 7.11 - 7.16.

NN Analysis Variables	
Object kinematics	
p_T of leading electron	$p_T(e_1)$
p_T of trailing electron	$p_T(e_2)$
sum of the transverse momenta of the leptons:	$\lvert\vec{p}_T(e_1) + \vec{p}_T(e_2)\rvert$
scalar sum of the momenta of all jets:	$H_T = \sum_i \text{jet}_i$
minimal quality of one of the two leptons:	$\min\{\mathcal{Q}(e_1), \mathcal{Q}(e_2)\}$
Event Kinematics	
invariant mass of both leptons	$M_{\text{inv}}(e_1, e_2)$
minimal transverse mass of either lepton and $\not\!\!E_{\text{T}}$	M_T^{min}
missing transverse energy	$\not\!\!E_{\text{T}}$
met significance	$\not\!\!E_T^{Sig}$
Topological Variables	
azimuthal angle between selected electrons	$\Delta\phi(e_1, e_2)$
azimuthal angle between leading electron and $\not\!\!E_{\text{T}}$	$\Delta\phi(\not\!\!E_{\text{T}}, e_1)$
azimuthal angle between trailing electron and $\not\!\!E_{\text{T}}$	$\Delta\phi(\not\!\!E_{\text{T}}, e_2)$

Table 7.6: Input variables for the Neural Network.

7.2.3 Training and Testing of the Neural Network

In order to avoid over-training and create an artificially increased sensitivity the signal and background samples used for testing and training are orthogonal. The individual events are evenly distributed throughout the entire sample, e.g each generated Monte Carlo sample is represented in training and testing. The training samples are divided again, using one half for the training itself and the other one for testing during the network training. The training is performed using very similar selection requirements as discussed in Sec. 7.1. The only exception is the cut on $\not\!\!E_T^{Sig}$ which was lowered from 6 to 5 to maintain high statistics for the training and control samples. The selection requirements as in Sec. 7.1 have been developed considering the following neural network approach. For example assuming the neural network would be trained and applied using a sample with none or looser selection requirements placed upon such that the data sample is still dominated by $Z \to \ell\ell$ events. In this case the training emphasizes the separation power of variables separating Drell-Yan processes, leading to an non-optimal neural network structure for the $W +$ jets and diboson background within the number of training cycles possible with the given Monte Carlo statistics. The signal training was done using the $H \to WW^* \to \ell^+\nu\ell^-\nu$ and the $q\bar{q} \to g\bar{g}H \to qqWW^* \to \ell^+\nu\ell^-\nu$ samples. Associated Higgs production processes WH and ZH are not included in the training due to significant kinematic difference. The Neural Network has been re-trained for each Higgs mass.

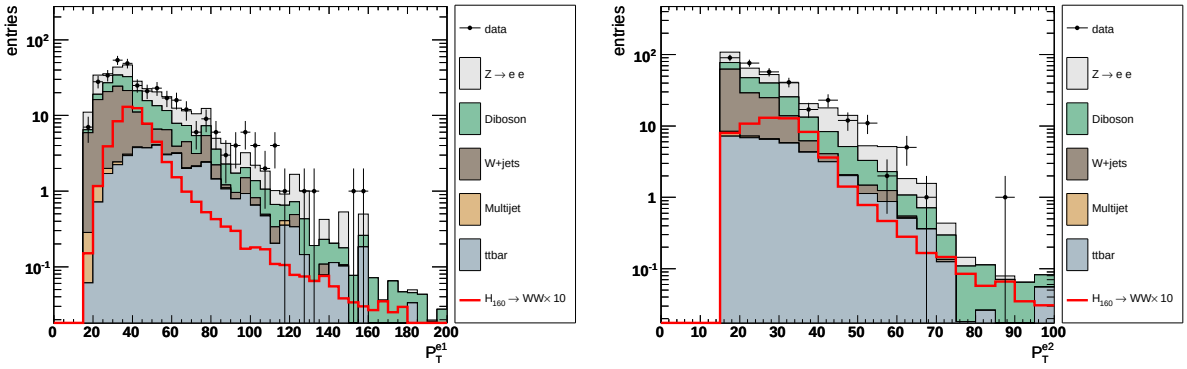


Figure 7.11: Distribution of the leading lepton (left) and the trailing lepton p_T (right) used in the Neural Network. The signal indicated by the red graph is representative of a Higgs boson mass of $m_{\mathrm{H}} = 160$ GeV. The entire Run II data set is used.

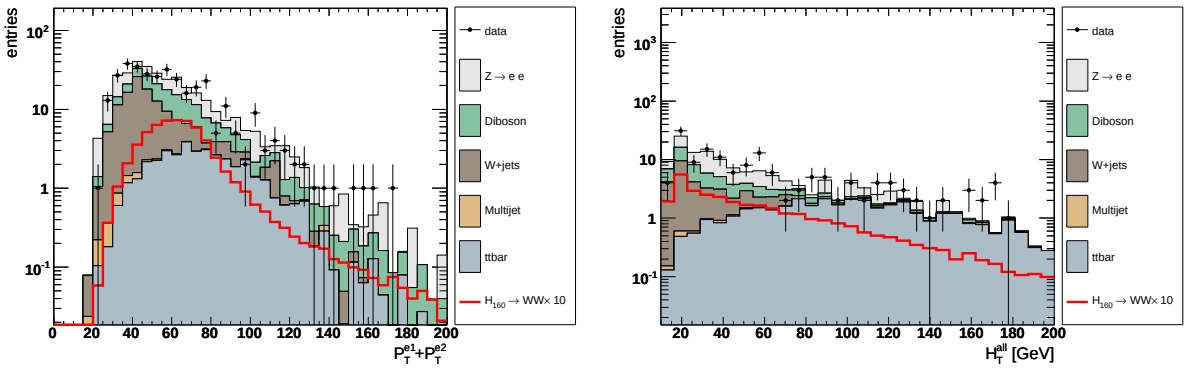


Figure 7.12: Distribution of the scalar sum of the lepton momenta $|\vec{p}_T(e1) + \vec{p}_T(e2)|$ (left) and the scalar sum of the jet energies (right) as used in the Neural Network. The signal indicated by the red graph is representative of a Higgs boson mass of $m_{\mathrm{H}} = 160$ GeV. The entire Run II data set is used.

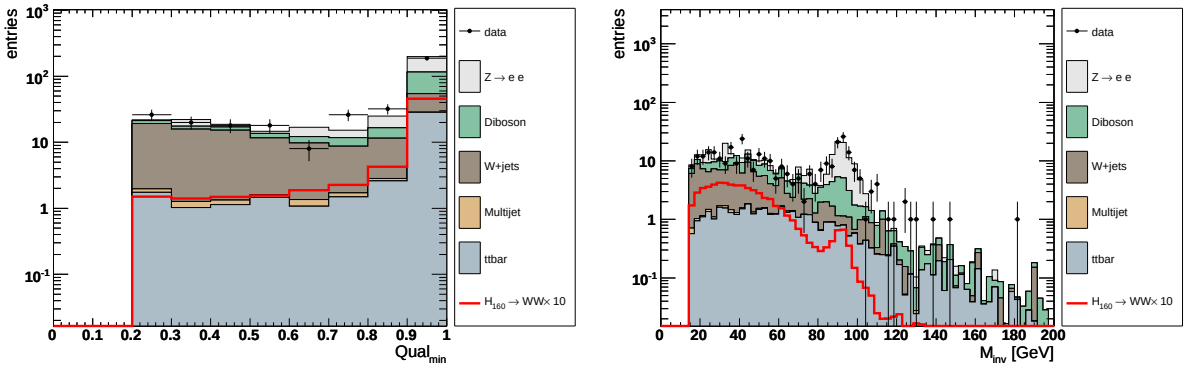


Figure 7.13: Distribution of the minimum lepton quality of either of the two leptons (left) and the invariant mass of the di-electron system M_{ee} (right) used in the Neural Network. The signal indicated by the red graph is representative of a Higgs boson mass of $m_{\mathrm{H}} = 160$ GeV. The entire Run II data set is used.

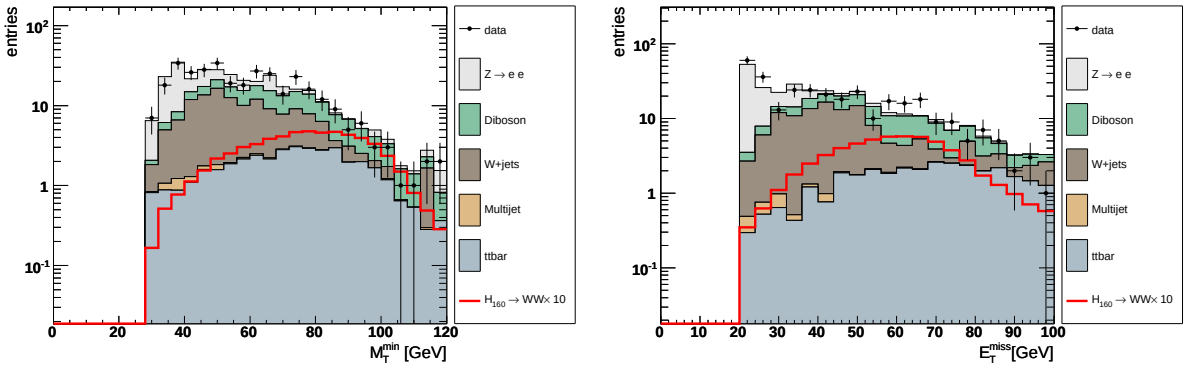


Figure 7.14: Distribution of the minimal transverse mass of the leptons (left) and of the missing transverse energy $M^{T}_{min}(\ell, \not\!\!E_{\mathrm{T}})$ (right) used in the Neural Network. The signal indicated by the red graph is representative of a Higgs boson mass of $m_{\mathrm{H}} = 160$ GeV. The entire Run II data set is used.

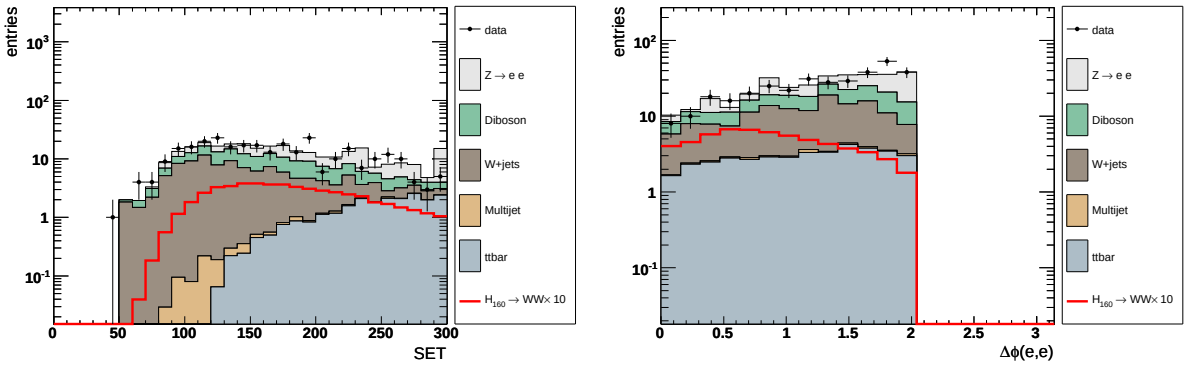


Figure 7.15: Distribution of the scalar transverse energy of the event (left) and the opening angle between the two selected electrons $\Delta\phi(e1, e2)$ (right) used in the Neural Network. The signal indicated by the red graph is representative of a Higgs boson mass of $m_{\mathrm{H}} = 160$ GeV. The entire Run II data set is used.

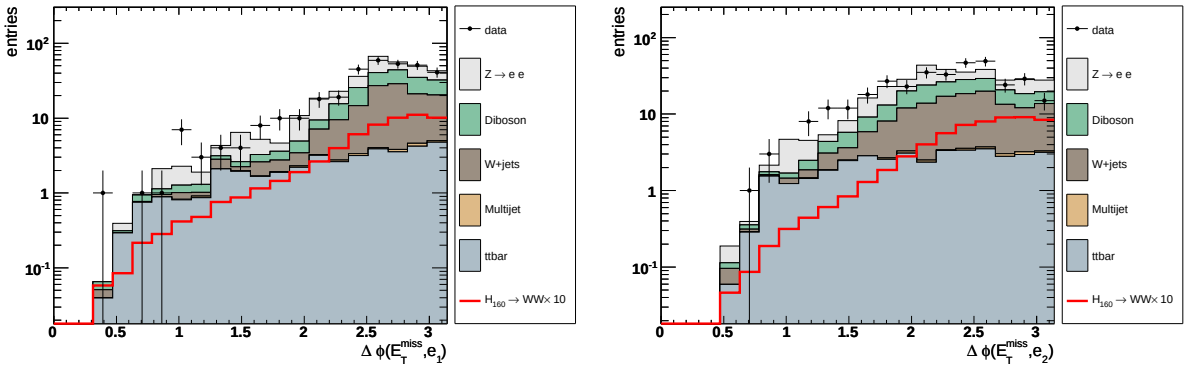


Figure 7.16: Distribution the opening angle between the leading (left) and trailing lepton (right) and $\not\!\!E_{\mathrm{T}}$ used in the Neural Network. The signal indicated by the red graph is representative of a Higgs boson mass of $m_{\mathrm{H}} = 160$ GeV. The entire Run II data set is used.

7.2.4 Neural Network Output for selected Higgs boson masses

The neural network output distributions after applying each of the selection requirement as listed in Table 7.1 are shown in Figs. 7.17, 7.18 and 7.19 for the most sensitive Higgs mass, $m_H = 165$ GeV.

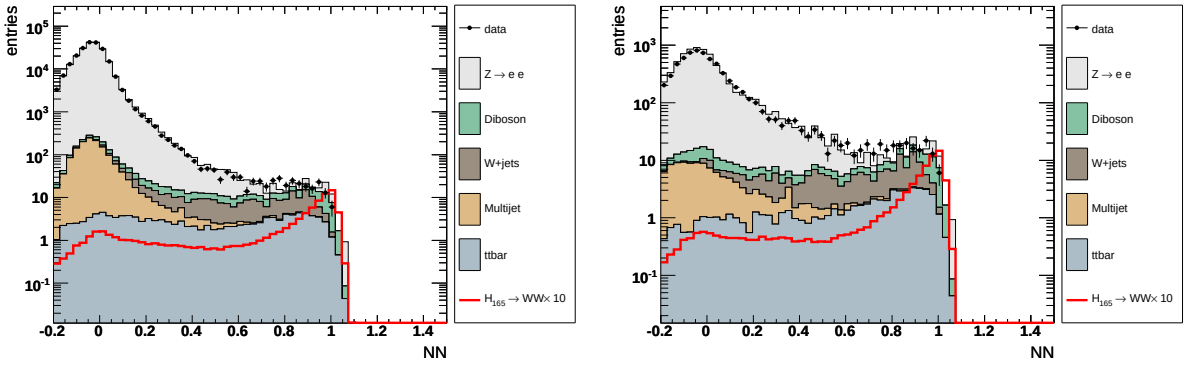


Figure 7.17: NN output at preselection level (left) and after the $\Delta\phi(\ell,\ell)$ selection cut (right) for a Higgs boson mass of $m_{\mathrm{H}} = 165$ GeV. The entire Run II data set is used.

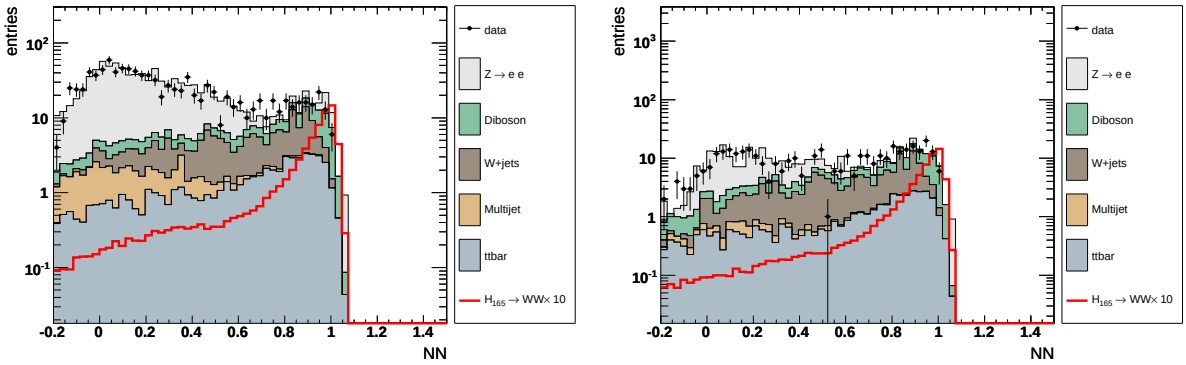


Figure 7.18: NN output after the $\not\!\!E_{\mathrm{T}}$ (left) and after the $\not\!\!E_T^{Sig}$ selection cut (right) for a Higgs boson mass of $m_{\mathrm{H}} = 165$ GeV. The entire Run II data set is used.

The neural network output distributions at preselection level and after applying all selection requirements for Higgs boson masses of $115,\ 130,\ 150$ and 170 GeV are displayed in Figs. 7.20 - 7.23. The Higgs masses displayed are $m_H = 115,\ 130,\ 150$ and 170 GeV. The neural network output for the remaining analyzed Higgs masses are found in Appendix B.2. All distributions show a good agreement between data and simulation.

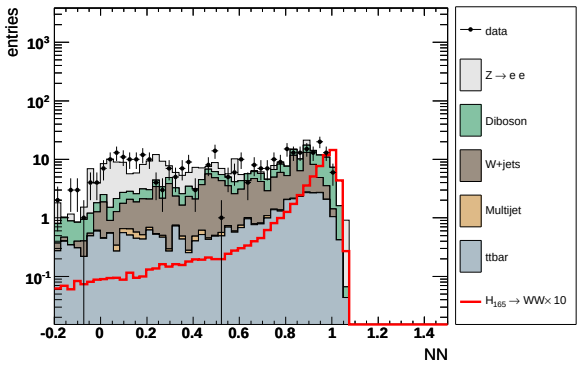


Figure 7.19: NN output after the $M^T_{min}(\ell, \not\!\!E_T)$ selection cut for a Higgs boson mass of $m_H = 165$ GeV. The entire Run II data set is used.

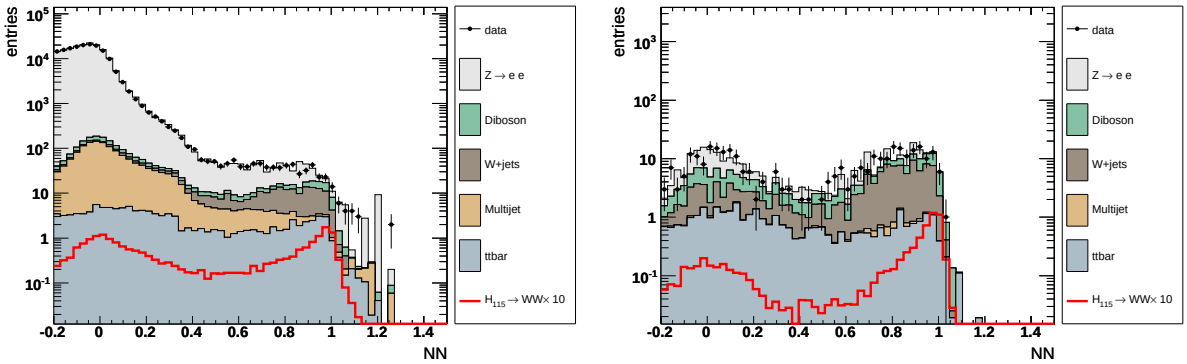


Figure 7.20: NN output at the preselection level (left) and after all selection requirements (right) for a Higgs boson mass of $m_H = 115$ GeV. The entire Run II data set is used.

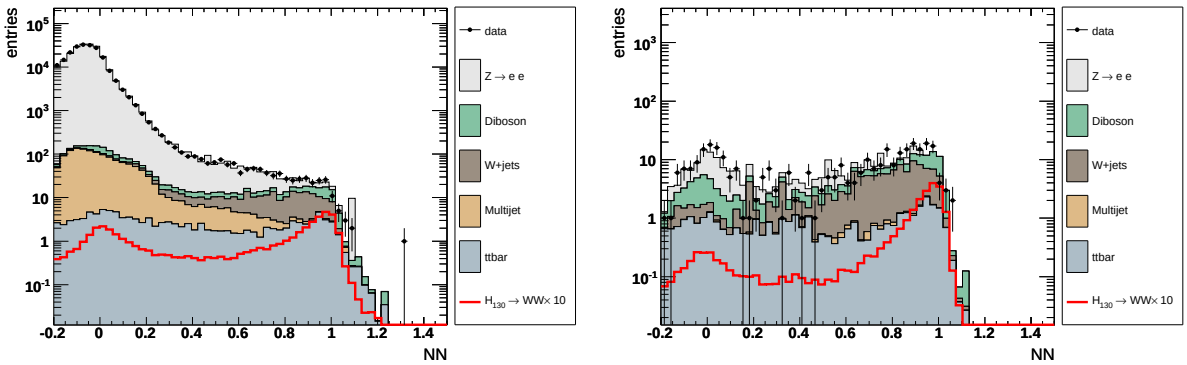


Figure 7.21: NN output at the preselection level (left) and after all selection requirements (right) for a Higgs boson mass of $m_H = 130$ GeV. The entire Run II data set is used.

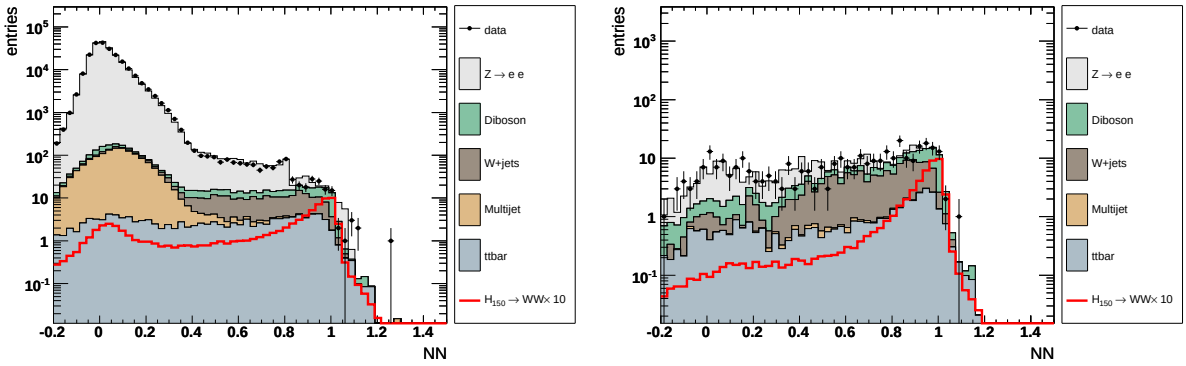


Figure 7.22: NN output at the preselection level (left) and after all selection requirements (right) for a Higgs boson mass of $m_{\mathrm{H}} = 150$ GeV. The entire Run II data set is used.

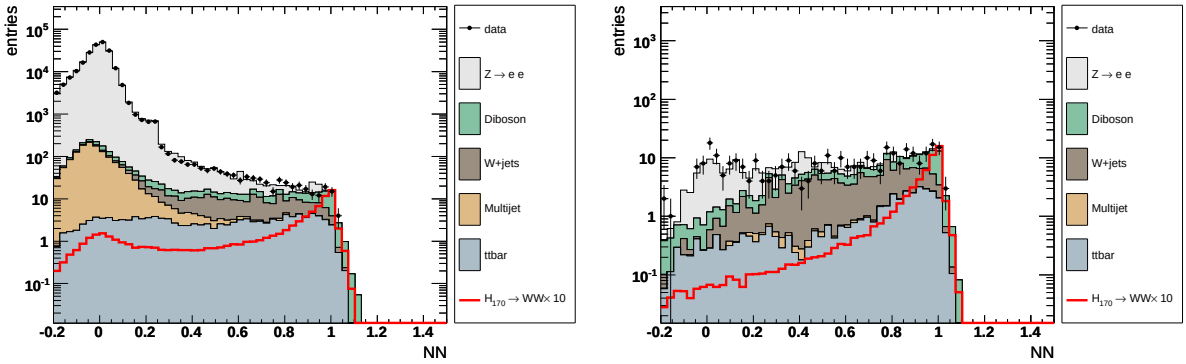


Figure 7.23: NN output at the preselection level (left) and after all selection requirements (right) for a Higgs boson mass of $m_{\mathrm{H}} = 170$ GeV. The entire Run II data set is used.

7.2.5 Event Display

Figure 7.24 shows a typical event selected by the Neural Network. The Higgs event candidate features an NN value of $NN_{160} = 1.003$. The $\not\!\!E_{\mathrm{T}}$ of this event is 72 GeV and the lepton momenta $40,\ 30$ GeV respectively.

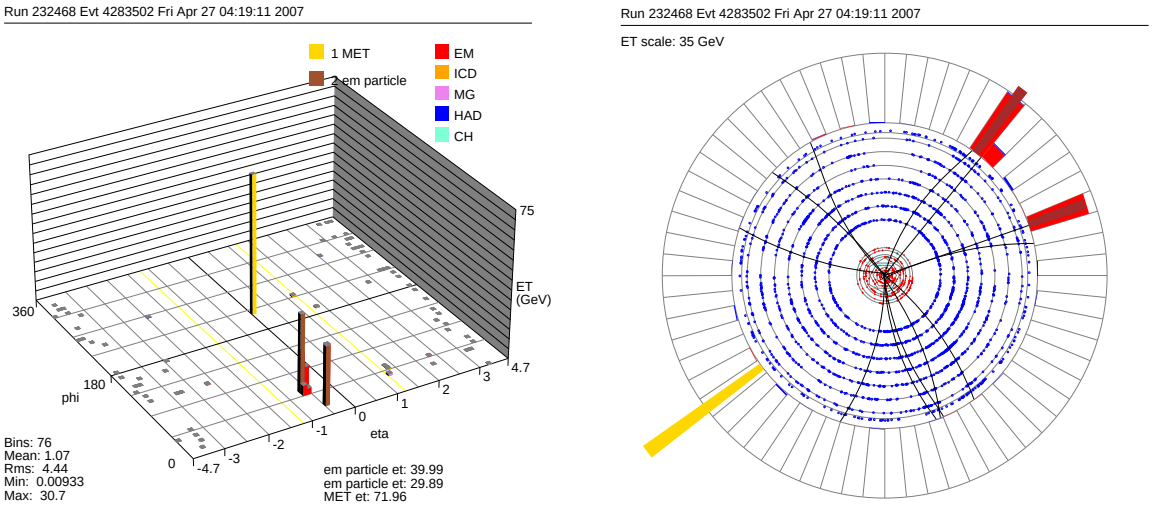


Figure 7.24: Event selected using the Neural Network with a high NN output of 1.0, $\not\!\!E_{\mathrm{T}} = 72$ GeV and p_T of the leptons of 40 GeV, respectively 30 GeV. The left plots gives the lego view of the event, the right plot the X-Y plane. $\not\!\!E_{\mathrm{T}}$ is represented by the yellow and the two electrons by the red bar.

One sees in Fig. 7.24 that the two electrons show a small opening angle, the desired events get selected by the neural network.

7.3 Systematic Uncertainties of the $H \rightarrow WW^* \rightarrow e^+\nu e^-\nu$ channel

The uncertainty or margin of error of a measurement is given by a range of values which are likely to enclose the true value. There are two classes of uncertainties which may impact the final result. The first sort of uncertainties are *statistical uncertainties* which are a result of stochastic fluctuations. They arise from the fact that a measurement is based on a finite set of observations and can be handled using the well developed mathematical theory of statistics. The second class of uncertainties are the *systematic uncertainties*. They arise from uncertainties associated with the nature of the measurement apparatus, methods used for the readout or the particle reconstruction, model-dependent uncertainties as the choice and parametrization of MC simulations and further assumptions made. The evaluation of systematic uncertainties is generally not trivial and will be discussed in the following.

7.3.1 Classes of Systematic Uncertainties

The uncertainties on various experimental quantities used in the analysis may have a sizable effect on the result and have to be investigated carefully. Therefore each source of a potential systematical uncertainty is considered in two ways:

- **Rate Uncertainties:** Uncertainties related to efficiencies and overall normalization of the contributing physical processes. This class of uncertainties is not changing the shape of the NN output classifier but the ratio of the various contributions to the entire data set.

- **Shape Uncertainties:** If the variation of a source of systematic uncertainty changes the NN values of a given MC event it impacts the multivariate classification of events. This uncertainties are referred to as shape or shape-changing uncertainties. Although these may also influence efficiencies or normalization, any uncertainty shown to impact the shape of the multivariate classifier final variable (NN) is treated as shape uncertainty.

The rate uncertainties are determined by propagating the systematic variation through the selection requirements as described in Sec. 7.1 and calculating the relative uncertainty. For the shape uncertainties the variations are propagated through the selection requirements and the neural network. By comparing the non-modified ("nominal") and modified shape of the neural network output the shape differences are derived, called *fractional uncertainty* as displayed in Fig. 7.26. The integral of the fraction uncertainty gives the relative uncertainty in percent. Because the Neural Network is separately trained for each analyzed Higgs boson mass the shape uncertainty depend on the neural network used. The plots in the following area all generated assuming a Higgs boson mass of m_H =165 GeV.

7.3.2 Sources of Statistical Uncertainties

The following sources of systematic uncertainties are considered:

Electron-ID

The uncertainty used for the electron ID is composed of statistical uncertainties on the scale factors in the parametrizations, systematics due to background subtraction method which increase the smaller the transverse momentum, uncertainties due to the dependence of $\Delta R(e, jet)$ and material distribution used in the Monte Carlo. Dependencies regarding the jet multiplicity of the event studied as well but were found to be negligible. Details can be found in Ref. [86]. A systematic uncertainty of 2.5% is used. Details regarding the estimation of the uncertainties can be found in Ref. [42]

$p_T(Z)$ reweighting

In order to evaluate the uncertainty introduced by the correction of the $p_T(Z)$ distribution the Monte Carlo samples are re-processed, discarding the correction. Then a side band of the analysis is selected by requiring $\Delta\phi(\ell,\ell) > 2.0$ is used. The selected sample offers a very high statistic and purity of Drell-Yan events. The scale of the uncertainty is evaluated by performing a fit of the corrected MC samples to the data using the $\Delta\phi(\ell,\ell)$ distribution. The difference in shape between the corrected and uncorrected MC samples is used as nuisance parameter of the fit. Nuisance parameters are intermediate variables used to derived properties such as the mean, variance and covariances of a quantity. In this examples the nuisance parameter is the estimation of the confidence interval for the fit. Figure 7.25 left shows the impact of the $p_T(Z)$ correction on the data sample. Propagating this correction through the neural network results in an output distribution with distorted shape. Figure 7.25 right shows the NN output distribution for the corrected and not corrected Drell-Yan samples.

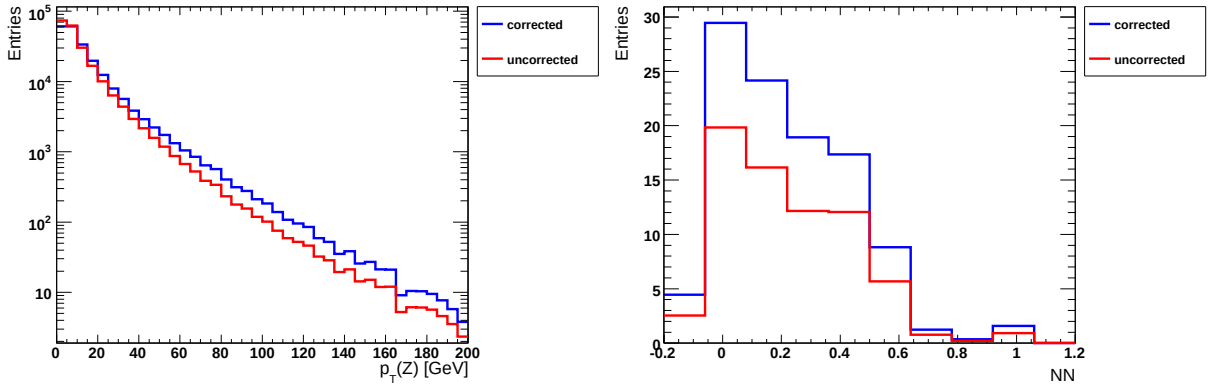


Figure 7.25: The left plot shows the difference in shape of the Drell-Yan processes with and without applying $p_T(Z)$ corrections. The samples size corresponds to preselection level in order to avoid statistical fluctuations. The right plots shows the difference of the neural network output shapes at the last cut stage with and without corrections. One sees that the shape is changed by applying the correction.

Dividing this two neural network output distributions on a bin by bin base leads to the *fractional shape uncertainties*, e.g. the relative change of the output distribution per bin. This fractional

shape for the given distributions is shown in Fig. 7.26. One sees clearly that the relative changes in Fig. 7.26 correspond to the region with modified shapes in Fig. 7.25. The distribution of the fractional shape systematic has been smoothed to avoid statistical fluctuations in the tails.

As listed in Table 7.7 the uncertainty is at the order of $\sim 4\%$ for the $Z \rightarrow \ell\ell$ sample, the total uncertainty introduced is $\sim 1.5\%$. The fractional uncertainty is shown in Fig. 7.26.

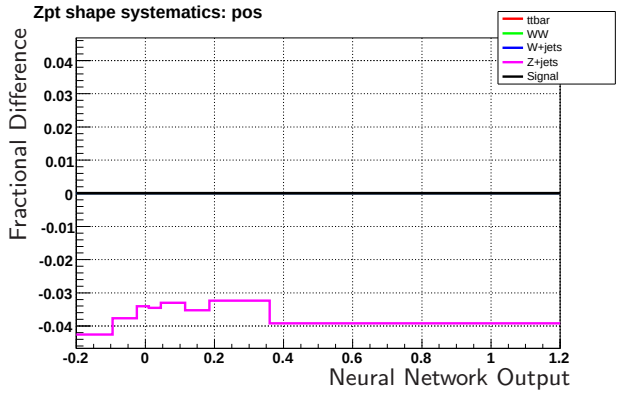


Figure 7.26: Fractional shape systematics for varying the $Z - p_T$ correction. The uncertainty is symmetrized to obtain the -1σ value.

Lepton Resolution

To calculate the uncertainty of the lepton momentum correction, discussed in Sec. 6.8, the analysis is repeated using the identical signal and background Monte Carlo processes but not applying the lepton smearing. The relative difference of the event yield between corrected and not corrected sample is assumed to be the 1σ deviation. Because this approach results in a single-sided systematic uncertainty the results has been symmetrized. Using the maximal difference is an conservative approach, still the effect on the overall uncertainty is negligible and less than $< 1\%$

Jet Energy Scale, Resolution and Identification

The uncertainties are evaluated by varying the jet energy scale (JES) by $\pm 1\sigma$ using the uncertainty parametrizations as measured by the JES group. Figure 4.8 shows the combined jet energy scale uncertainty. Over a wide kinematic range the uncertainties are at the level of 12%. All subcorrections are know to the same level of precision [56]. Due to the size of the response correction it becomes the dominant source of uncertainty on the combined jet energy scale. At low p_T the uncertainty is dominated by the purity of the γ+jet sample [91]. At high p_T the uncertainties are affected by decreasing statistics in the di-jet sample. Overall the limiting factors are the photon purity and the photon energy calibration. The high-precision jet energy calibration is directly visible in physics results, the uncertainty for the present thesis is 0.5% for signal and up to 8% for background events. The overall effect of the JES uncertainty is at the percent level even if some backgrounds, in particular the Drell-Yan (DY) processes, may show much larger variations. Jet resolution effects are estimated by shifting all Monte Carlo jets (in p_T) by the relative data-to-MC

jet energy scale [60] by varying the jet resolution by $\pm 1\sigma$. This assumes that the full jet energy scale difference is due to not understood resolution and efficiency effects in the Monte Carlo. This systematic uncertainty is small because the background rejection shows only a small dependency on jet multiplicities and jet energies. The overall systematic uncertainty is about 8%. This can be partially constrained on data because the events mostly affected show small NN values. Both uncertainties, JES and Jet-ID, show large fluctuations because DY events should contain none or little $\not\!\!E_{\mathrm{T}}$, all variables related to the missing transverse energy are more affected by variations of these quantities. The fractional uncertainties for varying the jet energy scale, the jet resolution and the jet identification are shown in Figs. 7.27-7.29.

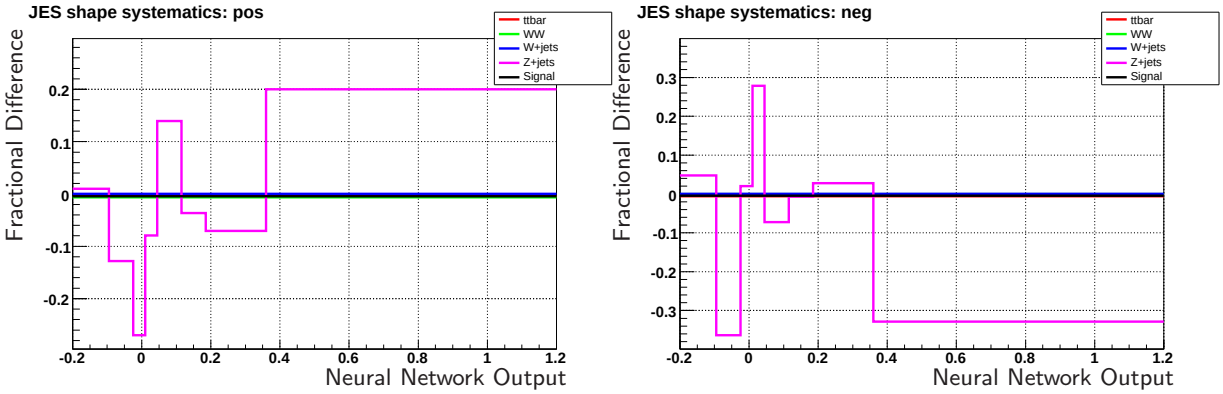


Figure 7.27: Fractional shape systematics for varying the Jet Energy Scale $+1\sigma$ (left) and -1σ (right) for the various backgrounds.

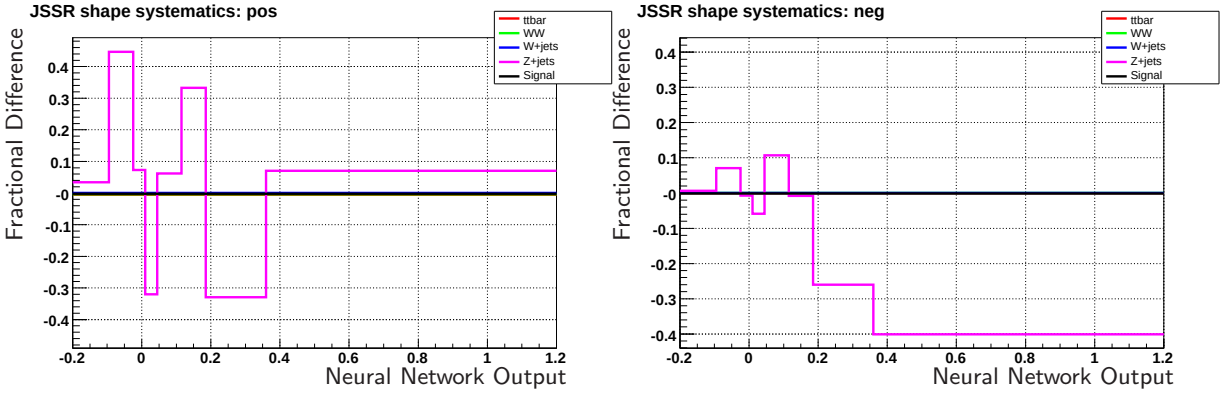


Figure 7.28: Fractional shape systematics for varying the Jet Energy Resolution $+1\sigma$ (left) and -1σ (right) for the various backgrounds.

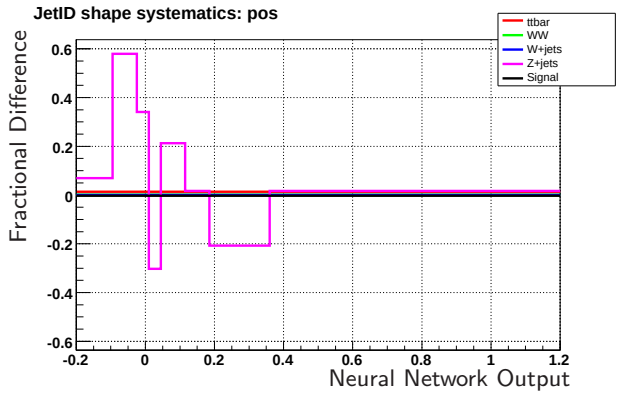


Figure 7.29: Fractional shape systematics for varying the Jet ID $+1\sigma$. The uncertainty is symmetrized to obtain the -1σ value.

Cross section

The PDF related uncertainty on the theoretical cross-sections are for the diboson pair production 7%, top pair production $t\bar{t}$ 10%, $W+$jets 20%, $Z+$jet/γ 6% and Higgs boson signal processes 10% [92, 93, 94, 95].

Multijet Normalization

The uncertainty on the estimation of the multijet background is calculated to be 2% of the final event yield by using the statistical uncertainty and the uncertainty in the like-sign / unlike-sign ratio. The uncertainty has been doubled in size in order to take into account possible $W+$jets fake rates which are not studied separately. The effect of larger uncertainties have been studied. Because the multijet background is well separated by the neural network and shows rather large statistical fluctuations after applying all selection requirements the effect on the final result is negligible.

Luminosity Uncertainty

The systematic error on the luminosity is 6.1% as measured by the luminosity group [96]. Studies performed evaluating the uncertainty of the Z peak normalization by varying the boundaries of the fitted function and comparing fit vs. bin-wise counting of the events result in an uncertainty of 5%, therefore agreeing well with the value cited by the luminosity group. The luminosity uncertainty has a sizable effect on the final result and is correlated with Higgs boson searches at DØ and CDF for the TEVATRON combination.

WW system momentum

A higher energetic $p_T(WW)$ spectrum may lead to a boost of the diboson system, therefore the angular distribution $\Delta\phi(\ell,\ell)$ is possibly modified. This is the best discriminating variable between the standard model background processes and the Higgs signal and entering the training of the

neural network. Therefore a systematic uncertainty on the modeling of the p_T of the WW system is determined by comparing the p_T distributions of PYTHIA, SHERPA [97], and MC@NLO [98]. The SHERPA and MC@NLO event generators agree well with each other and generate harder p_T spectra than PYTHIA (see also Ref [99]). The $p_T(WW)$ spectra of all three event generators are shown in Fig. 7.30.

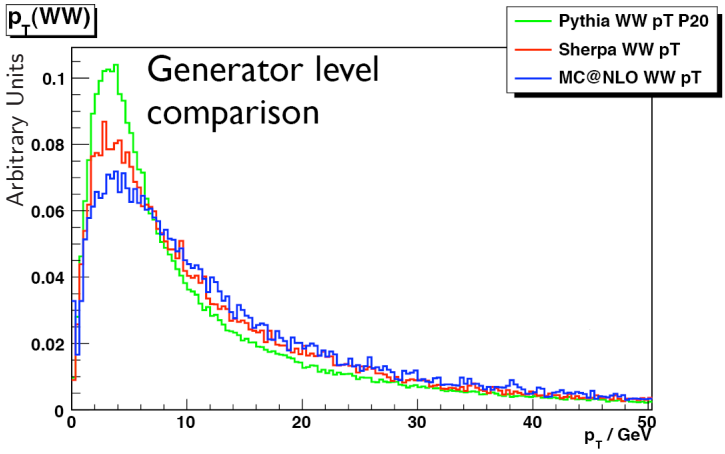


Figure 7.30: Spectra of the momentum of the WW system, $p_T(WW)$, as simulated by PYTHIA, given in green, SHERPA in red and MC@NLO in blue. All distributions are normalized to the same area. The $p_T(WW)$ spectrum simulated by PYTHIA is softer than the spectra generated by SHERPA respectively MC@NLO which are well in agreement.

Therefore a corresponding correction for the $p_T(WW)$ system is applied to the WW and $H \rightarrow WW$ samples. The fractional shape uncertainty between the sample with and without reweighting applied is taken as 1σ deviation. The fractional uncertainty is shown in Fig. 7.31 for a Higgs boson mass of $m_H = 165$ GeV.

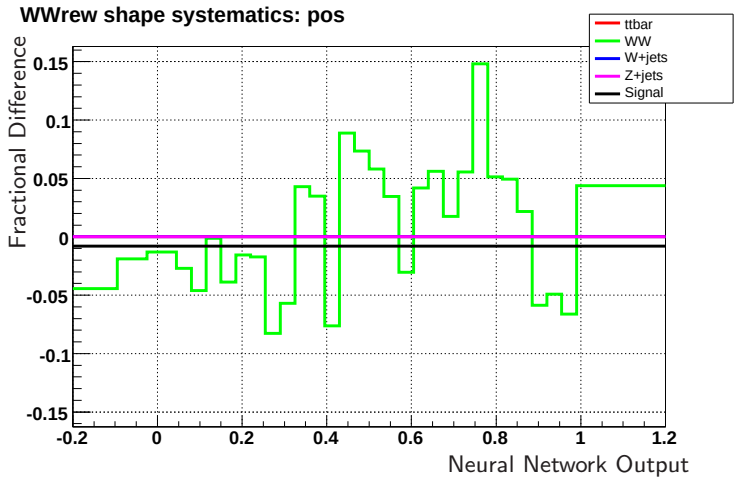


Figure 7.31: Fractional shape systematics for varying the WW reweighting $+1\sigma$. The uncertainty is symmetrized to obtain the -1σ value.

Charge Mis-ID

The charge flip uncertainty has been estimated by comparing the number of events within the Z peak in the like-sign sample and to the number of Z peak events in the opposite sign sample. All Standard Model processes have been taken into account. The charge mis-ID rate is found to be at the order of 1%. The sign of particles charges is measured by using the curvature of the track in the magnetic field of the tracker. The uncertainty on the tracks increases with high values of η_{det} and small curvatures, thus the charge-mis-ID rate is higher for evens with one or both lepton in EC with respect to CC events. The η_{det} dependency has been studied but is not further considered for the final systematic uncertainty due to the small size of the uncertainty of less than 1.0%. Figure 7.32 shows the Z-peak observed in the like-sign data sample.

VBF Modeling

The irreducible WW background is mainly produced via the quark-antiquark annihilation process $q\bar{q} \rightarrow WW$ but the gluon-induced contribution may contribute significantly $gg \rightarrow WW$ although suppressed by two orders of α_s (see Ref. [100] for details). Particularly the $\Delta\phi(\ell,\ell)$ distribution of this latter production channel differs significantly and may be enhanced by selection cuts relative to the quark-antiquark annihilation process. Although this effect is more severe at the Large Hadron Collider due to the high gluon-flux, this effect has been studied for the present analysis and taken as additional systematic uncertainty on the background prediction. The gluon-fusion process is not considered in PYTHIA. Therefore the GG2WW [101] event generator is used to derive an appropriate correction on parton level. The correction is applied to the WW background process. The uncertainty is found to be at the order of 2%. The fractional uncertainty is shown in Fig. 7.33 for a Higgs boson mass of $m_H = 165$ GeV.

One has to be careful not to bias the assessment of systematical uncertainties by statistical

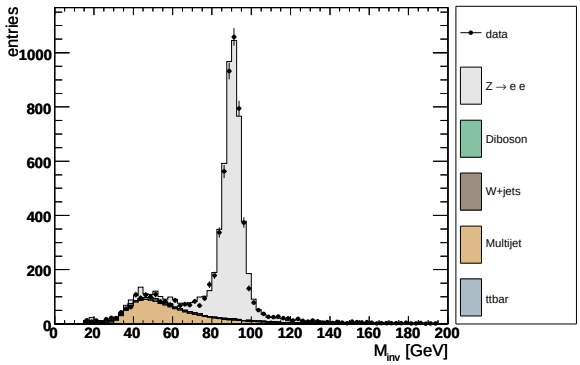


Figure 7.32: Spectrum of the invariant mass of the di-electron system for electrons with same charge, the Z peak is caused by electrons with charge-mis-identification. The data set corresponds to the entire Run II data set.

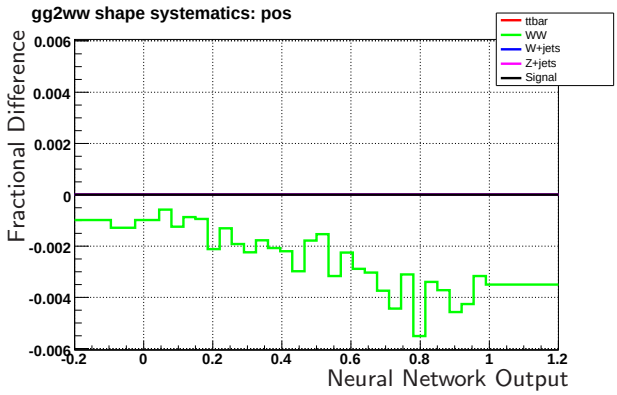


Figure 7.33: Fractional shape systematics for varying the VBF modeling $+1\sigma$. The uncertainty is symmetrized to obtain the -1σ value.

fluctuations due to limited available MC statistics. The low statistic tails of some background contributions such as $W+$jets and $Z+$jet/γ are particularly susceptible to potential biases. This is even more important because this analysis does not warrant a truly background free region. Techniques such as smoothing can be used to address fluctuations due to limited statistics in simulated events. The Collie package provides a histogram smoothing algorithm based on Gaussian kernel estimation as described in Ref. [102]. This algorithm is a more robust alternative to that provided in the `Root` histogramming software package and the actual implementation has been partially developed in the course of this analysis. The smoothing algorithm is designed such that it preserves all statistically important features of the histogram. Neither the mean nor the variance of the distributions are changing during the smoothing procedure.

Smoothing is applied only to the $W+$jets , $Z+$jet/γ and the $t\bar{t}$ samples. The signal and main backgrounds remain therefore unchanged. An examples for the effect of the smoothing algorithm on the $W+$jets background contribution is shown in Fig. 7.34.

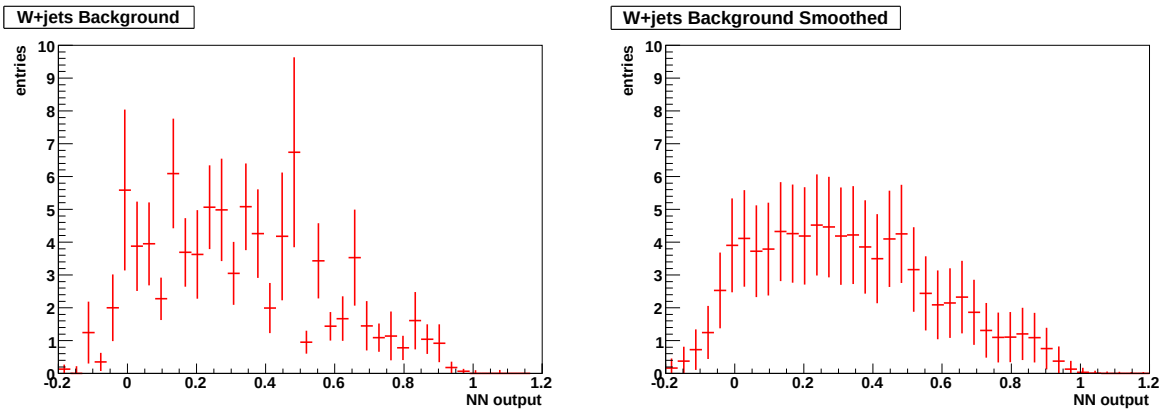


Figure 7.34: Distribution of the Neural Network output discriminator for the W + jets background only for a Higgs boson mass of $m_H = 165$ GeV. The left histogram shows the not smeared, the right figure the smoothed histogram.

	Σ Bkgd	Signal	$Z \rightarrow ee$	$Z \rightarrow \tau\tau$	W + jets	$t\bar{t}$	ZZ	WZ	WW	QCD
JES +	-2.66	-0.49	-7.15	0.00	-0.18	-0.43	-0.78	-0.10	-0.67	0.00
JES -	0.80	0.33	1.93	0.00	0.38	0.38	0.65	-0.21	0.00	0.00
Jet Reso. Up	-0.22	-0.33	-0.52	27.94	0.02	-0.58	-0.78	-1.34	-0.16	0.00
Jet Reso. Down	0.24	0.00	0.37	0.00	0.51	0.03	-0.26	0.10	-0.07	0.00
$Z - p_T$ Rew.	-1.46	0.00	-4.30	-3.62	0.00	0.00	0.00	0.00	0.00	0.00
Jet ID	7.70	-0.33	22.50	23.53	0.04	0.50	1.95	-0.62	-0.16	0.00
WW NLO	-0.01	-0.98	0.00	0.00	0.00	0.00	0.00	0.00	-0.04	0.00
Lep. mom.	0.72	0.00	2.15	-36.05	3.14	-1.01	-2.36	-3.54	-1.80	0.00
VBF model	-0.22	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-1.05	0.00
σ	-	10.0	6.0	6.0	20.0	10.0	6.0	6.0	6.0	-
QCD	2	0	-	-	-	-	-	-	-	-
Lep. ID.	2.5	2.5	-	-	-	-	-	-	-	-
Charge flips	-	1.0	1.0	-	-	-	-	-	-	-

Table 7.7: Systematic uncertainties in percent

However, after all cuts the systematic uncertainties are small compared to the statistical uncertainties due to limited statistics.

7.3.3 Fitting of Systematic Uncertainties

During the limit setting process a technique is used called profile maximization (profiling) which refers to the practice of determining the "best fit" of the predicted model to data by minimizing a χ^2 function (i.e. maximizing the likelihood) over the possible values of parameters describing the systematic uncertainties [103]. Given a set of predictions, observations, and systematic uncertainties, one can define a model which represents the best fit to the data observation within the constraints of the systematic uncertainties. A reliable mean of performing this fit is achieved by minimizing the χ^2 function. The input to the problem are the best estimates for each background source model and systematic uncertainty. During this fitting process all systematic uncertainties are varied separately. If a shape uncertainty is used each bin is individually varied according to

the corresponding fractional uncertainty. The following sources of uncertainties are considered as shape uncertainties when calculating the limit.

- Jet Energy Scale
- $Z - p_T$ reweighting
- Jet Resolution
- Jet-ID
- WW reweighting
- VBF modeling

The results for the fit of the systematic uncertainties, using a Higgs boson signal of $m_H = 165$ GeV, are shown in Fig. 7.35. This figure shows the shift in each systematic uncertainty from nominal, in terms of number of standard deviations for that systematic uncertainty.

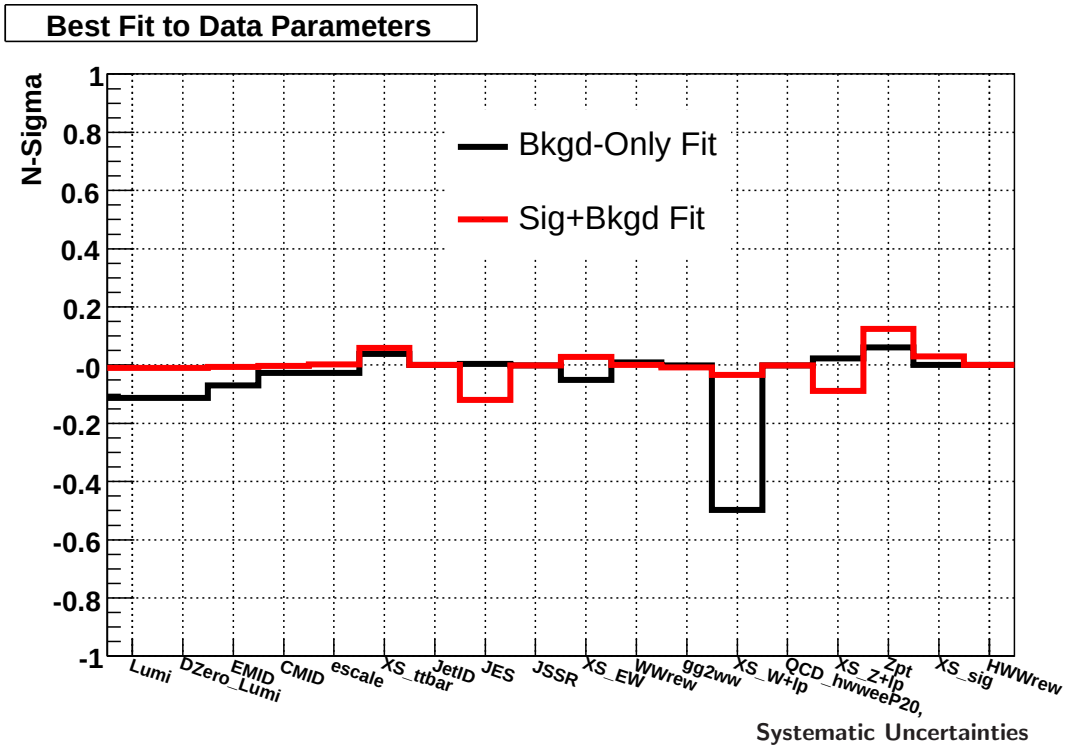


Figure 7.35: Fitted values for the systematic uncertainties in the background-only hypothesis (black) and signal-plus-background hypothesis (red). The signal hypothesis is an Higgs boson signal for $m_H = 165$ GeV normalized. The y-axis shows the shift in each systematic uncertainty from nominal, in terms of number of standard deviation for that systematic uncertainty. The downward fluctuation one sees corresponds to less than 1σ and is caused by the large uncertainty of the W + jets cross section.

7.4 Limit Calculation and Result

Figure 7.36 shows the NN output for the most sensitive Higgs boson mass of $m_H = 165$ GeV at the last cut stage. The right plot gives the distribution one obtains after subtracting the expected background from the data. Because the data and MC prediction is well in agreement the distributions fluctuates around 0 given the uncertainties.

The data appears to be consistent with a background only hypothesis. To evaluate the possible presence of a signal in the NN output and to derive limits on the production cross section times branching ratio $\sigma \times BR(H \to WW^* \to ee)$ a detailed statistical analysis has been performed.

7.4.1 Statistical Methods of the Limit Calculation

The methods follows a semi-frequentist approach. This will be described briefly in the following.

Assuming that Q represents a statistical estimator chosen to evaluate the difference between the Test and Null hypotheses. A common choice is the logarithm of the ratio of Poisson likelihoods

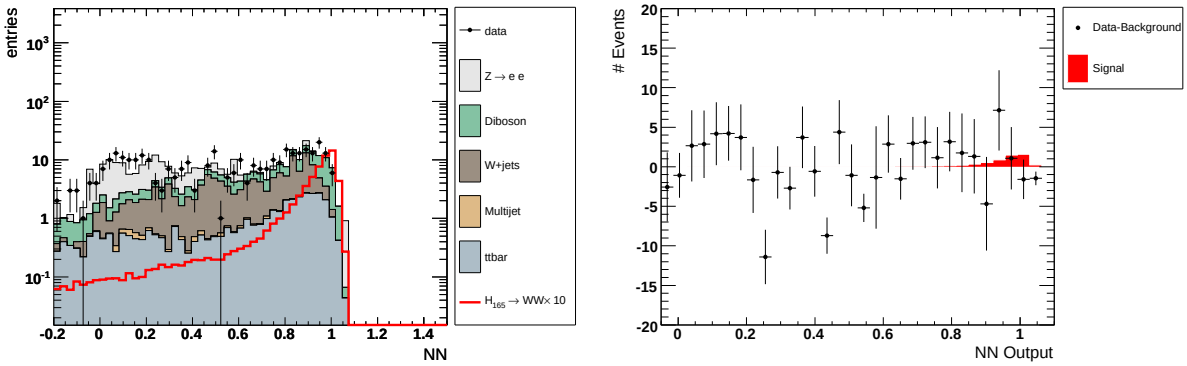


Figure 7.36: Comparison between Data and Monte Carlo of the NN output for m_H=165 GeV at the final cut stage (left). The Higgs signal is scaled up by a factor of ten. The right plot shows the same NN output distribution but the background expectation is subtracted from the data and the signal contribution is normalized to its theoretical cross section. Only statistical uncertainties are given.

(log-likelihood ratio or LLR):

$$Q = -2\log\frac{P(s+b)}{P(b)} \tag{7.2}$$

The definition of the semi-frequentist confidence level (CL_s) is then given by:

$$CL_s = CL_{S+B}/CL_b \tag{7.3}$$

with the values CL_{s+b} and CL_b defined as the probability for the signal-plus-background (Test) and background-only (Null) hypotheses to produce outcomes less signal-like than the one observed in data [104]:

$$CL_{s+b} = P_{s+b}(Q \geq Q_d) \tag{7.4}$$

$$CL_b = P_b(Q \geq Q_d) \tag{7.5}$$

Q_d represents the data observation for the estimator. The probability distribution functions for the Test and Null hypotheses are populated via random Poisson trials with mean values given by the expected numbers of events in the Test and Null hypotheses respectively. Systematic uncertainties are incorporated by varying the expected numbers of events in each hypothesis according to the size and correlations of each parameter describing them.

The calculation is performed using the software package COLLIE [105]. This package uses a technique called profile maximization (profiling) which refers to the practice of determining the "best fit" of the predicted model to data by minimizing a χ^2 function (i.e. maximizing the likelihood) over the possible values of parameters describing the systematic uncertainties. The fit is described in Sec. 7.3.3.

The implementation of the limit calculation uses two fits to a data observation, one each for the Test (background + signal) and Null (background-only) hypotheses. Therefore a new likelihood ratio is defined by the ratio of χ^2 values minimized for each scenario fit independently:

$$Q' = -2 \log \frac{\chi^2_{min}(\text{Test}|\vec{D})}{\chi^2_{min}(\text{Null}|\vec{D})} \tag{7.6}$$

where $\vec{D}$ represents the data or pseudo-data used for the fit.

This approach ensures that uncertainties and their correlations are propagated to the outcome with their proper weights. The profiling approach utilizes binned final-variable distributions rather than a single-bin (fully integrated) value. Detailed informations on `Collie` and the profiling technique can be found in Ref. [103].

7.4.2 Upper Limits for the Di-Electron Channel

The limits are derived using systematic uncertainties as listed in Tab. 7.7 and the CL_s method as explained above. Therefore this method incorporates systematic uncertainties as uncertainties on the expected numbers of signal and background events. The estimator used for the limit setting is the shape of the NN output after the last selection requirement for data, expected backgrounds and signal. This NN output is shown in Fig. 7.19 for m_H=165 GeV. Two limits are obtained, one is the limit predicted by the simulation of the signal and the various background contributions, called the *expected* limit. The limit resulting from the observation in data versus the simulation of the signal is the *observed limit*. In case of well understood and modeled physics processes one expects both limits to agree within the given uncertainties assuming no presence of new physics. In the case of new physical processes like a Higgs boson signal one would expect a difference between observed and expected limit with increasing significance as the sensitivity of the analysis increases.
The limits are calculated at 95% C.L. and presented as $\sigma(p\bar{p} \to H \to W^+W^-)$ relative to the SM expectation, the final observed and expected limits are given in Tables 7.8 and 7.9. The graphical representation of the limit is shown in Fig. 7.37. Here the dashed red line represents the expected limit whereas the solid black line represents the observed limits.

m_H=	115	120	125	130	135	140	145	150	155
ee (exp.)	44.2	28.4	18.2	12.5	9.9	8.0	6.1	5.9	4.7
ee (obs.)	30.1	18.7	16.4	12.6	9.1	8.0	6.9	5.9	3.2

Table 7.8: Expected and observed upper limits at 95% C.L. on the cross section times branching ratio for $\sigma \times BR(H \to WW^*)/\sigma_{SM}$ for the ee final states with respect to the Standard Model expectation. The limits are givens for Higgs boson mass ranges of $m_H = 115 - 155$ GeV

The Log Likelihood Ratio (LLR) plot, shown in Fig. 7.38 allows a more quantitative assessment how well expected and observed limit agree. Here the green and yellow shaded areas represent the $\pm 1\sigma$ and $\pm 2\sigma$ confidence levels for the background only hypothesis and the black solid graph represent the observed limit. The "background-only" and "background + signal" hypothesis are

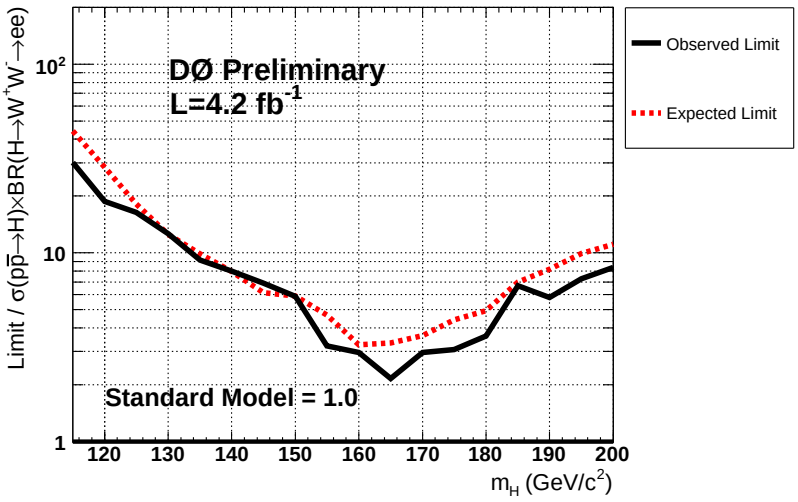


Figure 7.37: Expected and observed limits in terms of $\sigma \times BR(H \rightarrow WW^*)/\sigma_{SM}$ for the ee final state.

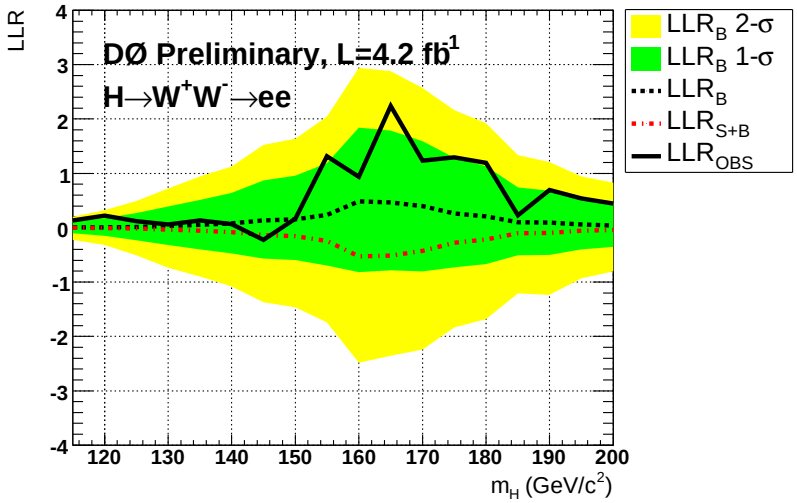


Figure 7.38: Log Likelihood Ratio for the expected and observed limit. The green and yellow shaded areas represent the $\pm 1\sigma$ and $\pm 2\sigma$ confidence levels for the background only hypothesis and the solid black graph represents the observed limit. The "background-only" and "background + signal" hypothesis are given by the black respectively the red dashed line.

$m_H=$	160	165	170	175	180	185	190	195	200
ee (exp.)	3.3	3.3	3.6	4.42	4.9	7.0	8.2	9.9	11.1
ee (obs.)	3.0	2.2	3.0	3.1	3.6	6.7	5.8	7.3	8.4

Table 7.9: Expected and observed upper limits at 95% C.L. on the cross section times branching ratio for $\sigma \times BR(H \to WW^*)/\sigma_{SM}$ for the ee final states with respect to the Standard Model expectation. The limits are givens for Higgs boson mass ranges of $m_H = 160 - 200$ GeV

given by the black and red dashed line respectively. Positive values on the LLR plot correspond to background like regions and negative values are signal like. One notices the bulge between the expect limits given the signal only and signal+background hypothesis, showing that the present analysis is gaining sensitivity for the SM Higgs boson.

7.5 Combined DØ Limits

To obtain maximal sensitivity the present analysis is combined with the orthogonal final states $H \to WW^* \to e^+\nu e^-\nu$ and $H \to WW^* \to \mu^+\nu\mu^-\nu$. The integrated luminosity of the $\mu^+\mu^-$ final state corresponds to 3.0 fb^{-1} and 4.2 fb^{-1} for the $e^\pm\mu^\mp$ final states.

The other channel follow the analysis as discussed in this thesis, therefore the treatment of simulations, analysis techniques and systematic uncertainties are almost identical. Again the various signal and Standard Model background processes are simulated with PYTHIA and ALPGEN, the MC corrections and the normalizations are treated as discussed in Sec. 6. The main background processes remain the same.

Muon tracks are reconstructed from hits in the wire chambers and scintillators in the muon system and must match a track in the central tracker (see also Sec 4.5). To select isolated muons, the scalar sum of the transverse momentum of all tracks, other than that of the muon, in a cone of $\mathcal{R} = 0.5$ around the muon track is calculated, where $\mathcal{R} = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2}$ and ϕ is the azimuthal angle. The transverse energy deposited in the calorimeter in a hollow cone of $0.1 < \mathcal{R} < 0.4$ around the muon is also measured. In the $e\mu$ final state, both quantities are required to be $< 0.15 \times p_T^\mu$, where p_T^μ is the transverse momentum of the muon. In the $\mu\mu$ final state, the sum of the variables is required to be $< 0.4\ (0.5)\ \times p_T^\mu$ for the leading (trailing) muon. Muons are restricted to the fiducial coverage of the muon system $|\eta| < 2.0$. Muons from cosmic rays are rejected by requiring a timing criterion on the hits in the scintillator layers as well as applying restrictions on the position of the muon track with respect to the selected primary vertex.

At preselection stage muons must have $p_T^\mu > 10$ GeV whereas the electron in the $e\mu$ final state is as well required to exceed $p_T^e > 15$ GeV. In the $\mu\mu$ final state one of the two muons is required to have $p_T^\mu > 15$ GeV. In all final states, the two leptons originating from the same primary vertex are required to be of opposite charge and again the dilepton invariant mass is required to exceed 15 GeV. For the dimuon final state the number of jets with transverse energy $E_T > 15$ GeV is required to be $n_{\text{jet}} < 2$ where jets are reconstructed in the calorimeter with a cone of radius $\mathcal{R} = 0.5$. Both muons must be separated from the nearest jet by $\Delta\mathcal{R} > 0.1$. Trigger requirements are not applied for the $\mu\mu$ and $e\mu$ channels.

Some selections are final-state dependent and optimized to further suppress contributions from $Z +$ jet/γ , diboson (WW, WZ, ZZ), $W +$ jets , and multijet backgrounds. Table 7.10 shows the selection criteria used for the three different channels. Table 7.11 shows the number of expected and observed events after pre-selection and final selections for all three channels. Only statistical uncertainties are given except for the multijet background where systematical uncertainties are given. The difference observed in the $\mu\mu$ channel corresponds to less than 3 σ considering the systematical uncertainties.

Final state		$e\mu$	ee	$\mu\mu$
Cut 0	Pre-selection	lepton ID, leptons with opposite charge and $p_T^\mu > 10$ GeV and $p_T^e > 15$ GeV invariant mass $M_{\ell\ell} > 15$ GeV $\mu\mu$: $n_{\rm jet} < 2$ for $p_T^{\rm jet} > 15$ GeV, $\Delta\mathcal{R}(\mu, {\rm jet}) > 0.1$ and $p_T^\mu > 15$ GeV for the leading μ		
Cut 1	Missing Transverse Energy $\not\!\!E_{\rm T}$ (GeV)	> 20	> 20	
Cut 2	$\not\!\!E_T^{Sig}$	> 6	> 6	
Cut 3	M_T^{min} $(\ell, \not\!\!E_{\rm T})$ (GeV)	> 20	> 30	
Cut 4	p_T^μ (GeV) for $n_{\rm jet} = 0$			> 20
	$\not\!\!E_{\rm T}$ (GeV) for $n_{\rm jet} = 1$			> 20
Cut 5	$\Delta\phi(\ell, \ell)$	< 2.0	< 2.0	< 2.5

Table 7.10: Summary of the selection criteria for the three final states (blank entries indicate that no requirement is applied on that quantity for the specific channel).

As well the $e\mu$ and $\mu\mu$ channels are improving the separation of signal from background by using an artificial neural network. As discussed in Sec. 7.2 a separate NN is trained for each Higgs boson mass tested and again the weighted sum of all background events was used during training. The systematic uncertainties are in a comparable order of magnitude as in Sec. 7.3, the total uncertainty on the background level is approximately 13% and for the signal efficiency about 10%.

After all selection cuts the NN output distributions in data agree within uncertainties with the expected backgrounds and the NN output distributions are used to set limits on the Higgs boson. Figures 7.39 and 7.40 show the result of the fit constraining the systematics uncertainties for the combination of all three channels before and after background subtraction. Tables 7.12 and 7.13 present expected and observed upper limits at 95% C.L. for $\sigma(p\overline{p} \rightarrow H \rightarrow W^+W^-)$ relative to that expected in the Standard Model expectation for each of the three final states and for their combination for each Higgs boson mass considered.

Figure 7.41 shows the expected and observed limits for $\sigma(p\overline{p} \rightarrow H \rightarrow W^+W^-)$ relative to the Standard Model for the different Higgs boson masses and the LLR distribution for the 3.0–4.2 fb^{-1} of Run II data. Figure 7.42 shows the confidence level for the exclusion of the cross section $\sigma(p\overline{p} \rightarrow H \rightarrow W^+W^-)$ in units of the Standard Model cross section for all the three channels combined as a function of the Higgs boson mass. So far, no region of the Standard Model

	$e\mu$ pre-selection	$e\mu$ final	$\mu\mu$ pre-selection	$\mu\mu$ final
$Z \to ee$	280.6 ± 3.3	$0.0^{+0.1}_{-0.0}$	–	–
$Z \to \mu\mu$	274.6 ± 0.9	5.8 ± 0.1	235670 ± 158	3921 ± 22
$Z \to \tau\tau$	3260 ± 3	7.3 ± 0.1	1735 ± 10	66 ± 2
$t\bar{t}$	272.0 ± 0.3	82.5 ± 0.2	19.93 ± 0.05	12.55 ± 0.04
W+jets	183 ± 4	78.6 ± 2.8	214 ± 7	134 ± 5
WW	421.2 ± 0.1	154.7 ± 0.1	159.0 ± 0.3	92.8 ± 0.3
WZ	20.5 ± 0.1	6.6 ± 0.1	47.3 ± 0.5	19.4 ± 0.3
ZZ	5.3 ± 0.1	0.60 ± 0.01	40.5 ± 0.2	15.1 ± 0.1
Multijet	279 ± 168	$1.1^{+9.6}_{-1.1}$	386 ± 20	64 ± 8
Signal ($m_H = 165$ GeV)	17.1 ± 0.01	12.2 ± 0.1	5.43 ± 0.01	4.85 ± 0.01
Total Background	4995 ± 168	337 ± 10	238272 ± 159	4325 ± 24
Data	4995	329	239923	4084

Table 7.11: Expected and observed number of events in each channel after pre-selection and final selections (the NN input stage). Statistical uncertainties in the expected yields are shown for all backgrounds whereas the systematic uncertainty is shown for the multijet background.

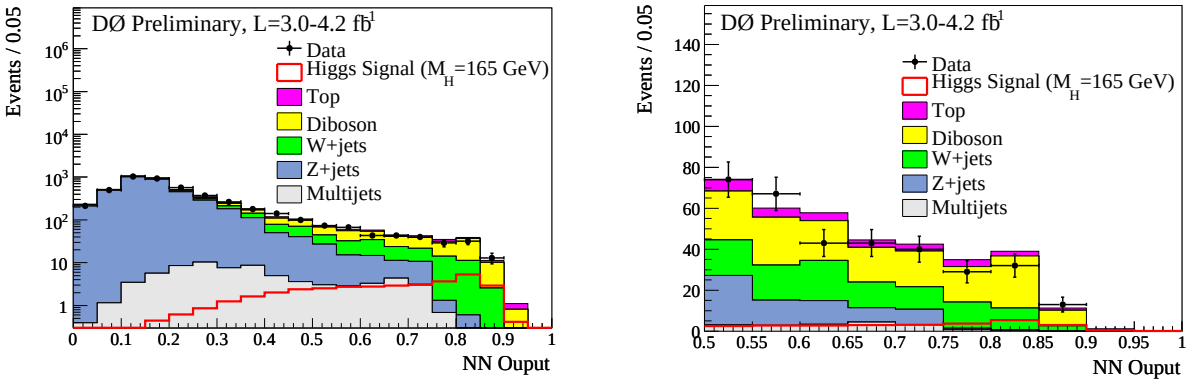


Figure 7.39: Data, Standard Model signal expectation and backgrounds as a function of the neural net output variable. The expected Standard Model signal for a Higgs boson mass of $m_H = 165$ GeV is shown by the red histogram. The neural net output is shown in logarithmic scale (left) and in linear scale for the high NN region (right).

Higgs boson mass range can be excluded and no significant excess of events is observed in data. However using the data of DØ alone the sensitivity of the current analysis has reached an expected confidence level for the exclusion of the Standard Model cross section for an Higgs boson with $m_H \approx 160$ GeV close to 80%. With increased integrated luminosity it will be possible to exclude the presence of a Higgs boson with masses in this range.

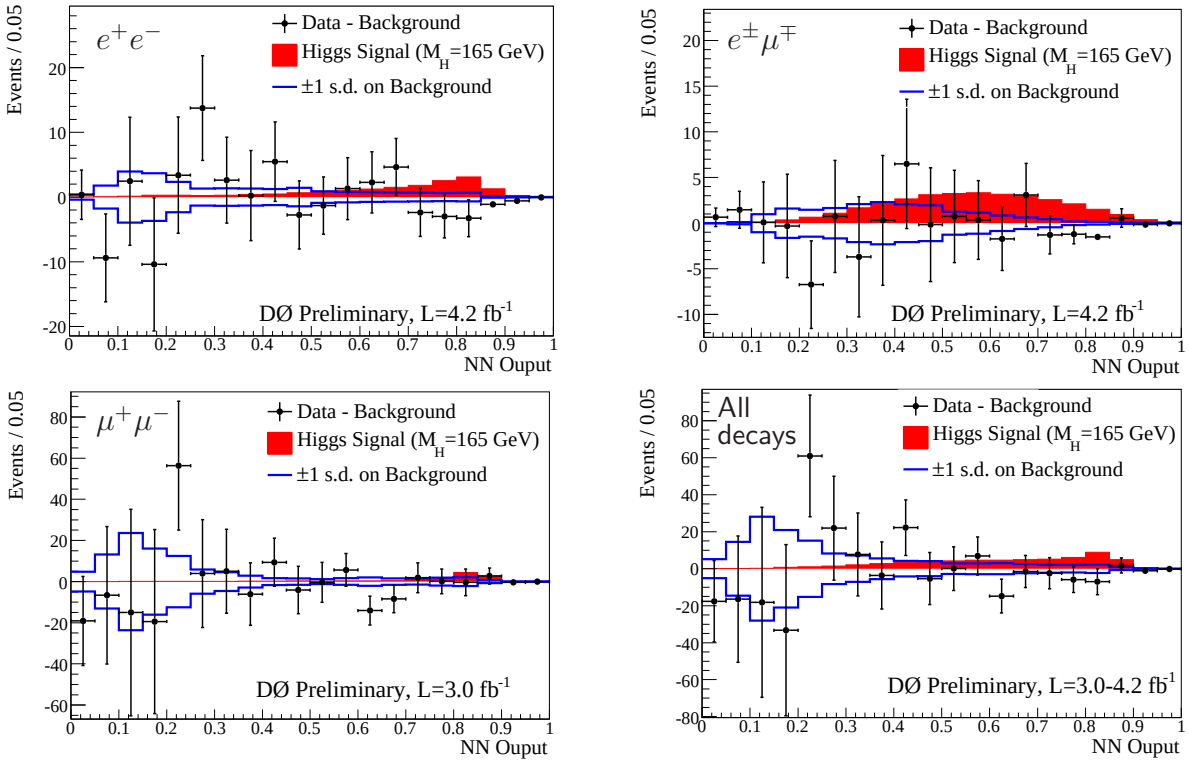


Figure 7.40: Data and Standard Model signal expectation after background subtraction as a function of the neural net output variable. The background resulting from the constrained fit using the signal plus background hypothesis is subtracted. The Standard Model signal expectation is shown by the red histogram. The constrained total systematic uncertainty is shown by the blue line.

$m_H=$	115	120	125	130	135	140	145	150	155
$e\mu$ (exp.)	40	23	19	11	11	8.5	7.7	4.3	3.5
$e\mu$ (obs.)	48	34	24	15	11	7.6	10	5.0	3.8
$\mu\mu$ (exp.)	94	36	23	16	16	11	10	8.2	7.6
$\mu\mu$ (obs.)	104	25	20	15	11	14	10	10	7.5
Run II (exp.)	23	13	9.6	6.3	5.6	4.7	4.0	2.9	2.4
Run II (obs.)	28	13	13	7.8	5.5	6.2	5.6	4.0	2.3

Table 7.12: Expected and observed upper limits at 95% C.L. relative to the Standard Model for $e\mu$, and $\mu\mu$ final states in Run II and their combination for different Higgs boson masses (m_H). The combined Run II limit incorporates the results for the ee channel given in Table 7.8.

m_H=	160	165	170	175	180	185	190	195	200
$e\mu$ (exp.)	2.6	2.3	2.7	3.2	3.8	5.1	6.4	7.6	8.4
$e\mu$ (obs.)	2.3	1.8	1.9	2.2	2.5	3.6	4.0	4.5	5.6
$\mu\mu$ (exp.)	6.4	5.7	5.9	7.7	9.2	12	14	16	19
$\mu\mu$ (obs.)	7.2	6.4	4.3	6.9	6.9	9.0	14	10	14
Run II (exp.)	1.8	1.7	2.0	2.4	2.9	4.0	4.8	5.8	6.7
Run II (obs.)	1.7	1.3	1.3	1.7	1.6	2.9	2.8	3.1	3.7

Table 7.13: Expected and observed upper limits at 95% C.L. relative to the Standard Model for $e\mu$, and $\mu\mu$ final states in Run II and their combination for different Higgs boson masses (m_H). The combined Run II limit incorporates the results for the ee channel given in Table 7.9.

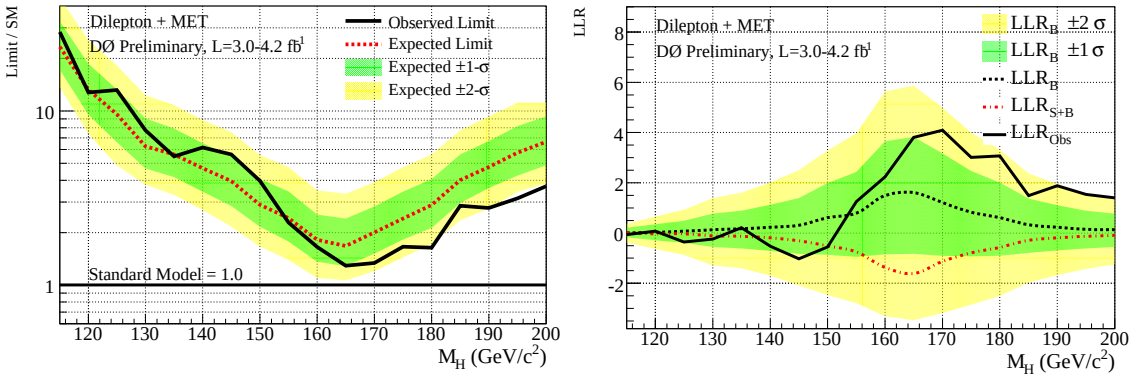


Figure 7.41: Excluded cross section ($\sigma(p\overline{p} \rightarrow H \rightarrow W^+W^-)$) at 95% C.L. in units of the Standard Model cross section (left) and LLR (right) for all three channels combined as a function of the Higgs boson mass, using 3.0–4.2 fb$^{-}1$ of Run II data.

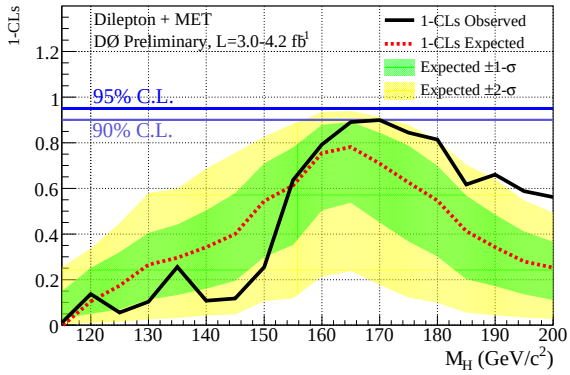


Figure 7.42: Confidence level for the exclusion of the cross section $\sigma(p\overline{p} \rightarrow H \rightarrow W^+W^-)$ in units of the Standard Model cross section for all three channels combined as a function of the Higgs boson mass, using 3.0–4.2 fb$^{-}1$ of Run II data.

8

The $H \rightarrow WW^* \rightarrow \mu + \tau_{had} + X$ Channel

8.1 The $\mu + \tau_{had} + X$ final state

Final states containing taus were neglected in initial Higgs sensitivity studies for DØ and the TEVATRON. In order to achieve maximal sensitivity it is desirable to make also use of these final states. The following chapter will discuss the data and Monte Carlo samples used along with corrections necessary for performing a High mass search in the $\mu + \tau_{had} + X$ final state. This analysis, although presented at the end of this thesis, was performed as one of the first projects while performing post-graduate research. Therefore this projects utilizes a smaller data sample and employs different multivariate methods as the di-electron analyses. As explained in detail in Sec. 4.6 three different tau types are distinguished:

Type I: Calorimeter cluster with one track associated and no EM subcluster, corresponding mostly to the decay $\tau^{\pm} \rightarrow \pi^{\pm}\nu$.

Type II: Calorimeter cluster with one track associated and at least one EM subcluster, corresponding mostly to the decay $\tau^{\pm} \rightarrow \pi^{\pm}\pi^{0}\nu$.

Type III: Calorimeter Cluster, with two or three associated tracks, with or without EM subcluster. This corresponds mostly to the decays $\tau^{\pm} \rightarrow \pi^{\pm}\pi^{\pm}\pi^{\mp}(\pi^{0})\nu$.

This analysis uses tau-type I and tau-type II. Tau-type III is neglected due to the large background contamination.

8.1.1 Introduction to the $\mu + \tau_{had} + X$ Channel

So far final states containing τ and muons have been neglected. If in the $H \rightarrow WW^*$ decay one or each of the W bosons decays into τ leptons, the subsequent tau decay can either take place via an electron, muon or hadrons. A significant part of the tau leptons decays further into electrons or muons and are mostly selected by the $H \rightarrow WW^* \rightarrow e^{\pm}\nu\mu^{\mp}\nu$ analysis. Among the remaining final states the $\mu + \tau_{had} + X$ final state, e.g. the μ coming from one of the W boson decays and the other decaying via an hadronic tau, offers the highest efficiency. Another advantage of this particular decay channel is the fact, that the τ identification algorithms are sensitive to low quality electrons. These low quality electrons are likely to be missed in the $H \rightarrow WW^* \rightarrow e^{\pm}\nu\mu^{\mp}\nu$ analysis and can therefore be recovered. The use of hadronic tau final states τ_{had} requires an advanced detector understanding and challenging background-suppression techniques but leads to an optimal use of

the data. Many of the principles and methods discussed in previous chapters remain unchanged for the $\mu + \tau_{had} + X$ analysis. Therefore next sections will focus on differences in methods and techniques.

8.1.2 Data Sample

The data used for the $\mu + \tau_{had} + X$ analysis was collected between April 2002 and November 2005, covering the Run IIa data sample. This data sample corresponds to an integrated luminosity of $1\ \mathrm{fb}^{-1}$. The analysis is based on data samples with at least one muon per event.

8.1.3 Monte Carlo Signal Samples And Backgrounds

The signal and Standard Model background processes have been generated in a similar fashion as discussed in Sec. 6.3.2. One of the differences is the use of PYTHIA 6.319 [69] along with CTEQ6L1 parton distribution functions. The calculation of the cross sections is based on the same PDFs. All Standard Model background processes and cross sections used in the present analysis are listed in Tables 6.1, 6.3 and 6.4. The main differences are the use of PYTHIA instead of ALPGEN to simulate the W + jets background processes and only 4 different Higgs boson masses are considered, $m_H = 120, 140, 160$ and 180 GeV. The gluon-fusion production mode is the only production mode considered. The signal samples used are given in Table 6.1. Additional Monte Carlo signal samples were neither available nor necessary given the sensitivity in 2007. The background contribution of multijet production is estimated from data and explained in Sec. 8.1.4.

8.1.4 Multijet Background in the $\mu + \tau_{had} + X$ Final State

The background contribution from multijet (QCD) production is determined directly from data using a method similar to the one described in Sec. 6.6. The same data set as described in Sec. 8.1.2 data is used to obtain a sample of tau-like jets. All muon and tau preselection criteria as defined in Sec. 4.5 and Sec. 4.6 are applied except of the isolation criteria. In order to obtain a sample of tau like jets the isolation criteria have been modified. The track isolation of the muon track in a hollow cone between $0.1 - 0.4$ in $\eta \times \phi$ is now required to be less than 7 GeV rather than 2.5 GeV. The calorimeter isolation around the muon candidate in a cone of the same size is required to be less than 2.5 GeV. To select the enriched sample the energy deposited in the hollow cone has to be in the range from 2.5 GeV to 7 GeV. Failing the calorimeter isolation leads to an increased contribution of multijet events in the selected sample. Upper limits are set to avoid a kinematic bias which could be introduced when the soft events entering the sample show a different trigger response. For the τ identification at DØ a neural network is used as discussed in Sec. 4.6. In a similar fashion as for the muons the neural network output of the τ candidate is required to be within the range $0.3 < \mathrm{NN} < 0.7$ opposed to $\mathrm{NN} > 0.9$ as required for the default tau selection. Again the initial selection is inverted while still preserving a lower cut on the neural network to avoid events in the sample which introduce a kinematic bias. Finally the leptons are required to be of same charge. These requirements lead to an enriched sample of multijet and W + jets events with almost no Z + jet/γ contributions. This sample is referred to as fake sample in the following. To obtain the contribution from the multijet background in the $\mu + \tau_{had} + X$ final

state a normalization function is calculated by performing two subsequent fits to the p_T distribution of both leptons using the multijet enriched sample. The correction function for the muon is derived from the multijet enriched fake sample as follows:

$$f^{\mu}_{QCD}(p^{\mu}_T) = \frac{N_{data}(p^{\mu}_T)^{\pm\pm} - N^{\pm\pm}_{MC}(p^{\mu}_T)}{N^{\pm\pm}_{fake}(p^{\mu}_T)},$$

Here $N^{\pm\pm}_{fake}(p^{\mu}_T)$ represents the number of like-sign events in the fake sample. These corrections as shown in Fig. 8.1 are applied to the multijet sample. This corrected sample QCD^{μ}_{corr} is then used to perform a second fit as function of p^{τ}_T as follows:

$$f^{\tau}_{QCD}(p^{\tau}_T) = \frac{N_{data}(p^{\tau}_T)^{\pm\pm} - N^{\pm\pm}_{MC}(p^{\tau}_T)}{N^{\pm\pm}_{QCD^{\mu}_{corr}}(p^{\tau}_T)}.$$

The result of the second fit is displayed in Fig. 8.2. The number of multi-jet events in the unlike-sign signal sample $N^{\pm\mp}_{QCD}(p^{\mu}_T)$ is determined from the number of unlike-sign fake events $N^{\pm\mp}_{fake}(p^{\mu}_T)$ using the following formula:

$$N^{\pm\mp}_{QCD}(p^{\mu}_T) = f^{\mu}_{QCD}(p^{\mu}_T) \cdot f^{\tau}_{QCD}(p^{\tau}_T) \cdot N^{\pm\mp}_{fake}(p^{\mu}_T, p^{\tau}_T).$$

The normalization functions $f_{QCD}(p^{\mu}_T)$ and $f_{QCD}(p^{\tau}_T)$ is determined once after the preselection and then kept constant. It is assumed that the ratio of like-sign and unlike-sign events in the fake sample and the data sample is equal [84].

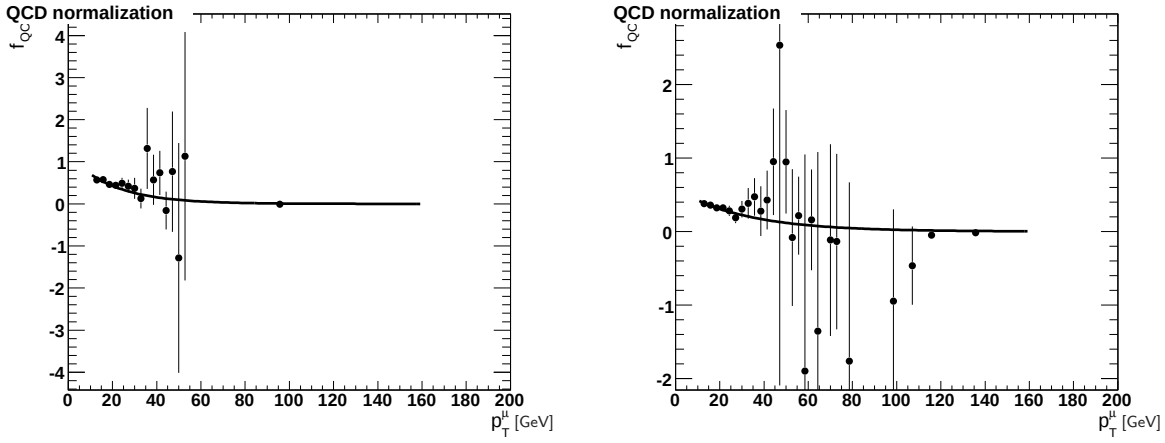


Figure 8.1: Fit of the normalization function for them multijet background as a function of p^{μ}_T. The fit for τ-type I (left) and τ-type II (right) performed in the like sign sample is applied to the corresponding τ-type in the opposite sign sample. The data are fitted using an exponential function.

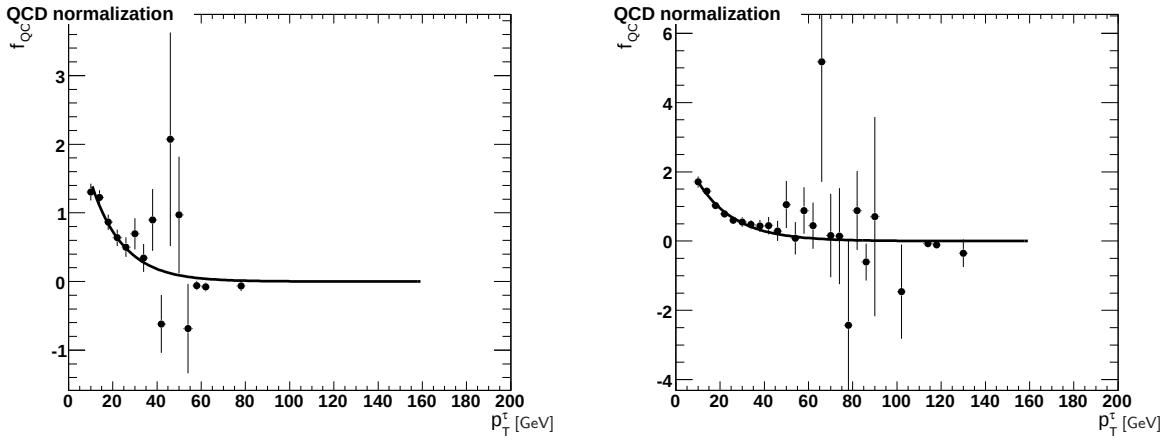


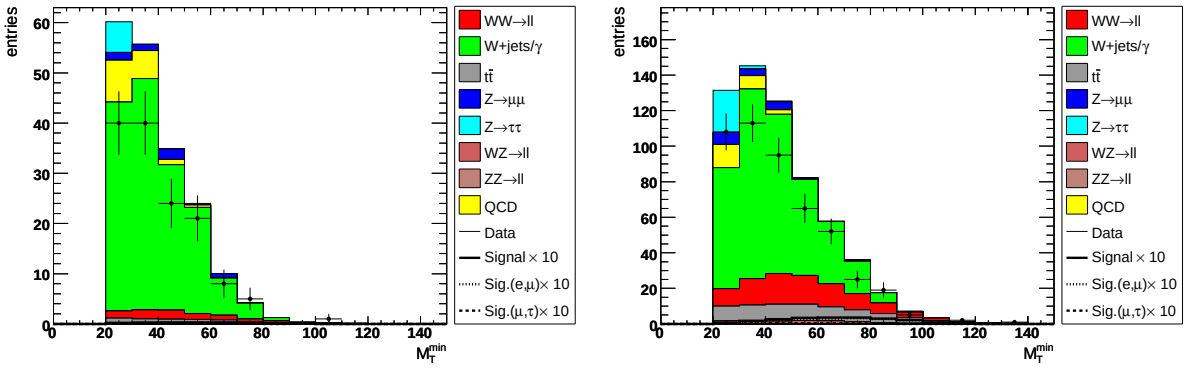
Figure 8.2: Fit of the normalization function for multijet background as a function of p_T^τ. The fit for τ-type I (left) and τ-type II (right) performed in the like sign sample is applied to the corresponding τ-type in the opposite sign sample. The fit for p_T^τ is performed using an exponential function after applying the correction functions obtained using the p_T^μ fit for the multijet background.

8.1.5 $W +$ jets Background

As already discussed in Sec. 6.7 the Monte Carlo cross section of the $W +$ jets samples needs to be corrected to improve the Monte Carlo description of the data. Particularly any mis-description of the $W +$ jets cross section is more visible in the $\mu + \tau_{had} + X$ final state because these background processes show a larger relative contribution than in the $H \rightarrow WW^* \rightarrow e^+\nu e^-\nu$ analysis. Since the $W +$ jets processes are characterized by a muon with high p_T from the W boson decay and $\not\!\!E_{\mathrm{T}}$ from the neutrino the $M_{min}^T(\ell, \not\!\!E_{\mathrm{T}})$ distribution tends to be shifted towards higher values. Therefore a sample enriched with $W +$ jets events is selected by requiring an $\not\!\!E_{\mathrm{T}}$ greater than 20 GeV, M_T greater than 20 GeV and leptons of opposite charge. All other selection requirements correspond to ones described in Sec. 4.5 and Sec. 4.6. This $W +$ jets enriched samples, shown in Fig. 8.3 are used to derive a correction factor for the $W +$ jets cross section by scaling the cross section to the data. This is individually performed for τ-type I and τ-type II. The scale factors obtained are 0.76 for τ type I and 0.82 for τ type II. Figures 8.3 and 8.4 show the $M_{min}^T(\ell, \not\!\!E_{\mathrm{T}})$ distribution with and without applying the scale factor, respectively. After correction the data are well described by the Monte Carlo simulations.

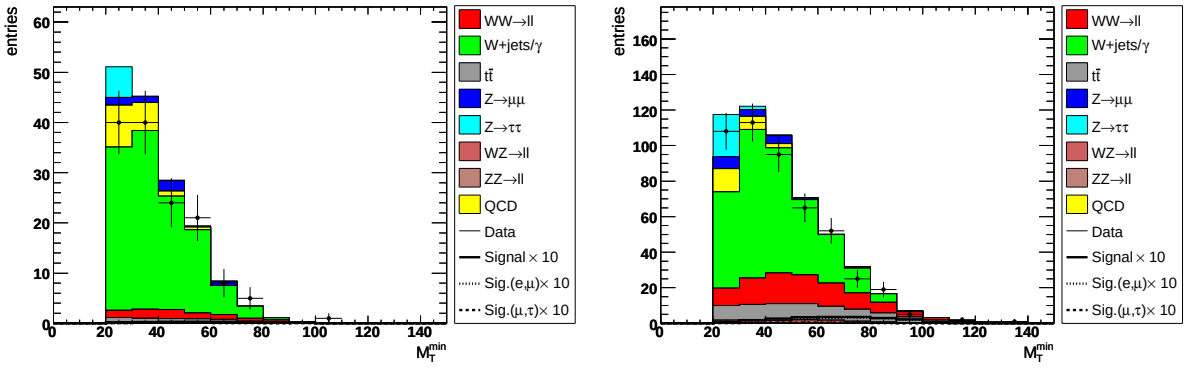
One of the reasons that the $H \rightarrow WW^* \rightarrow \mu + \tau_{had} + X$ channel shows a higher relative contribution of $W +$ jets background is the fact, that jets may be mis-identified either as tau leptons or - although less likely - as muons. In order to ensure that both major background contributions, $W +$ jets and multijet production are well described the following control samples are selected:

- To study jets mis-identified as taus the valid range of the tau NN output is restricted to $0.3 \leq NN \leq 0.7$. For minimizing the impact of multijet events the transverse mass between muon and $\not\!\!E_{\mathrm{T}}$ is required to be greater then 40 GeV. The selection requirements of the muon remain



τ-type I, no W + jets corrections τ-type II, no W + jets corrections

Figure 8.3: M_T^{min} distributions for tau type I (left) and tau type II (right) before applying corrections for the Monte Carlo cross section. The Run IIa data sample is used.

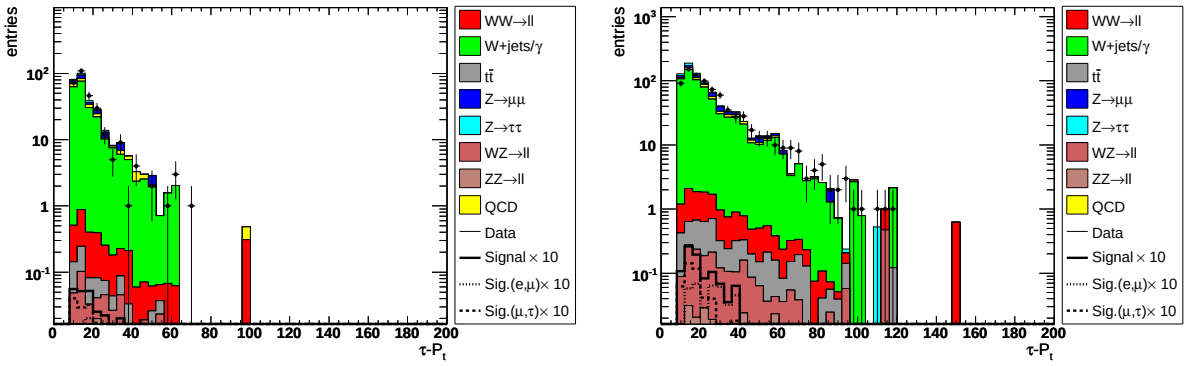


τ-type I, W + jets corrections applied τ-type II, W + jets corrections applied

Figure 8.4: M_T^{min} distributions for tau type I (left) and tau type II (right) after applying corrections for the Monte Carlo cross section. The Run IIa data sample is used.

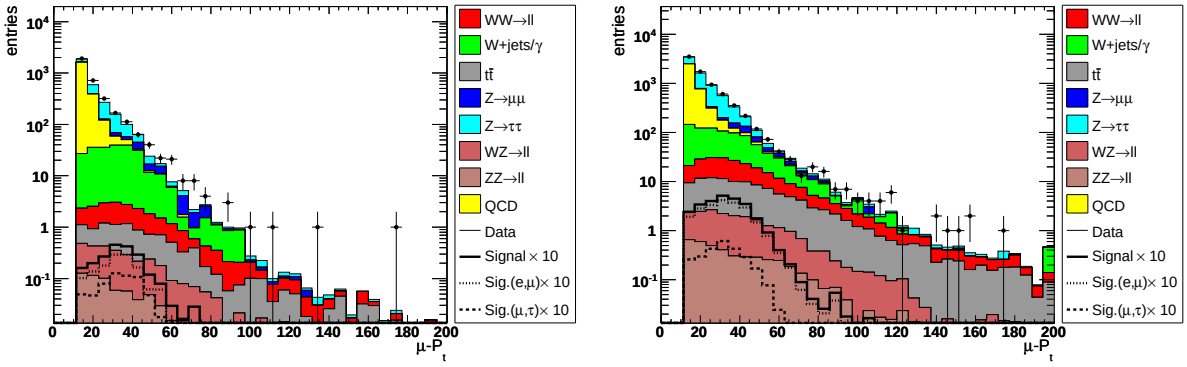
unchanged. By loosening the selection requirement on the tau candidate while preserving good quality muons the contribution of tau-candidates originating from mis-identified jet-candidates increases. Figure 8.5 shows the distribution of the transverse mass for both tau types. One sees that a good description of the data is given and the dominant contribution are W + jets events.

- To study jets mis-identified as muons the calorimeter isolation criteria is discarded while the tau requirement remains unchanged to the description in Sec. 4.6. Therefore the amount of jets misidentified as muons is increased. The resulting p_T distributions for τ-type I and τ-type II are shown in Fig. 8.6 and show a reasonable description of the data by the simulation.



τ-type I τ-type II

Figure 8.5: Tau p_T distributions for τ-type I (left) and τ-type II (right) with loose NN cut and calorimeter isolation required. In order to minimize multijet contamination the transverse mass between muon and $\not\!\!E_{\mathrm{T}}$ is required to be greater than 40 GeV. The Run IIa data sample is used.



τ-type I τ-type II

Figure 8.6: Muon p_T distributions for τ-type I (left) and τ-type II (right) with a tight NN cut but no calorimeter isolation required. The Run IIa data sample is used.

8.2 Monte Carlo Corrections

Most corrections are applied as discussed in Sec. 6.8. However, differences exists in the evaluation of efficiency corrections for the leptons. The reweighting of the instantaneous luminosity is not applied because the effect is negligible in Run IIa.

8.2.1 Muon Efficiency Corrections

Default corrections for muon and track identification are applied to Monte Carlo. The measurement of the reconstruction efficiency for a loose muon with a track match proceeds in a similar way as

the measurement of the trigger efficiencies in Sec. 6.4. A high-statistics sample of $Z \rightarrow \mu\mu$ events is selected using tight identification requirements for the first muon ('tag') while the second muon ('probe') is identified using only the tracking or the muon detector. The second muons are then used to determine the tracking and trigger efficiencies.

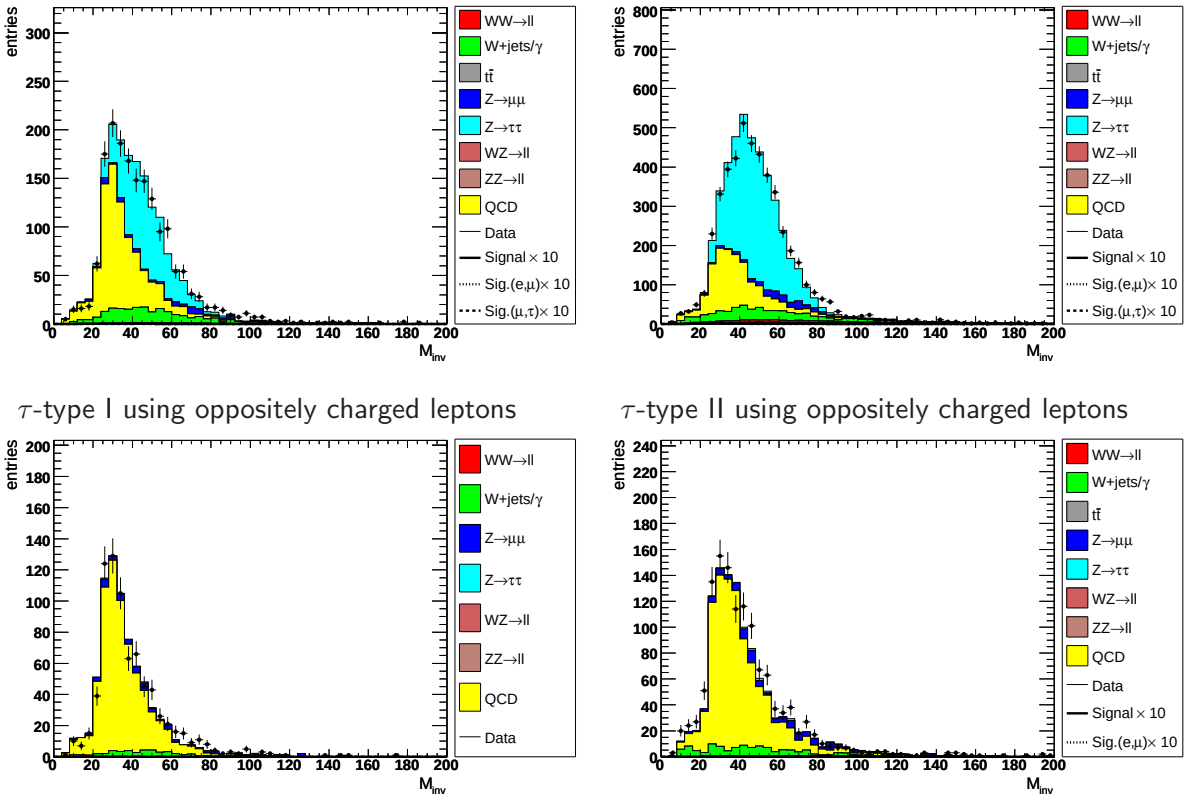
In $\mathrm{Z} \rightarrow \mu\mu$ events the fraction of tracks is calculated to which a loose muon can be matched. These efficiencies have been measured for data and Monte Carlo and their ratio provides a correction factor for Monte Carlo. Both corrections, muon ID and muon track are applied to all background and signal processes.

8.2.2 Tau Track SMT Efficiency Correction

The probability for tracks to have associated SMT hits is overestimated in Monte Carlo simulations. An appropriate correction is applied. The correction was derived by measuring the probabilities for obtaining at least one SMT hit in the CFT detector in terms of pseudorapidity and track z [106]. The average correction when applied to the $Z \rightarrow \tau^+\tau^-$ Monte Carlo is 97%.

8.2.3 Monte Carlo Normalization

For the $\mu\tau_{had}$ final state the normalization is the NLO $Z/\gamma^* \rightarrow \tau\tau$ cross section scaled to the data in the mass region 35 GeV$< M_{\tau\tau} <$75 GeV. Applying this normalization leads to a total luminosity $\times$ trigger efficiency of $\sim$ 1000 pb^{-1} for the used data sample. Monte Carlo corrections have been applied before normalizing to $Z/\gamma^* \rightarrow \tau\tau$. The systematic uncertainties on the normalization factor include the $Z +$ jet$/\gamma$ cross sections, the PDF uncertainties and the statistical uncertainties on the data/Monte Carlo normalization factor. By using un-prescaled running trigger the recorded luminosity of the data is measured to be 1050 pb^{-1} . The normalization of the multijet background follows an iterative approach when applying the method described in Sec. 8.1.4. At the preselection level the multijet background is subtracted from the data distribution and the sum of all Monte Carlo backgrounds is normalized to the Z boson mass peak in the invariant mass region 35 GeV$<$ $M_{\tau\tau} <$75 GeV. The luminosity corresponding to the Monte Carlo normalization is then used to scale the multijet background accordingly and the fit of the scale factor is re-iterated until the results does not change any more. As initial value 1050 pb^{-1} is used, the maximal recorded luminosity neglecting any trigger inefficiencies of the Run IIa data sample. This iterative method has the advantage that by subtracting the expected Monte Carlo background processes scaled to the luminosity both, the contribution of multijet processes and a Monte Carlo scale factor are simultaneously estimated. The effective luminosity, differs slightly for the two tau-types. For tau-type I the luminosity estimate corresponds to 1002 pb^{-1} and for tau-type II to 914 pb^{-1} . Figure 8.7 shows the Z mass peak in linear scale for both, like-sign and opposite sign charge requirements, using the background normalization and shape determination as explained.



τ-type I using oppositely charged leptons

τ-type II using oppositely charged leptons

τ-type I using leptons with the same charge

τ-type II using leptons with the same charge

Figure 8.7: Spectrum of the invariant mass for tau-type I and II in the like-sign and opposite sign sample. The Z boson peak is well described in either sample although the sample composition changes significantly. The Run IIa data sample is used.

Although the individual background processes feature varying shapes and relative contributions a good description is achieved in either case.

8.3 Event Selection for the $H \rightarrow WW^* \rightarrow \mu + \tau_{had} + X$ final state

The event selection of the $\mu + \tau_{had} + X$ follows a similar approach as presented in the di-electron final state. However, using muons and taus rather than electrons leads to different requirements for the leptons and to a different background composition. These differences will be discussed before presenting data and Monte Carlo comparisons in the $\mu + \tau_{had} + X$ final state and developing appropriate selection requirements for reducing the various background sources.

8.3.1 Triggers

The data for the $\mu + \tau_{had} + X$ final state is only triggered via muon triggers since tau trigger became only later available. The L1 term for muon trigger at Level 1 is based on information from the scintillator and wire chamber, showing an average efficiency of 78% for the wire term and 95% for a tight timing criteria in the scintillator term. At Level 2 the triggers are classified as either loose, medium or tight, depending on the number of hits in various layers. At Level 2 the transverse momentum p_T, measured with the muon system only, can be required to be above a given threshold. Almost all muon triggers use medium muons at Level 2, with a p_T cut of 0, 3 or 5 GeV. The average efficiency for a L2 medium muon with p_T >3 GeV, with respect to muons that have triggered the L1 tight scintillator and L1 loose wire trigger term is in average 96%. At Level 3 information of the local muon system is used and a p_T cut is applied. Evidently this leads to a decrease of the overall efficiency, depending on the threshold of the p_T cut. For a muon with p_T >20 GeV this efficiency with respect to the previous trigger terms is about 72%. However, to obtain the best efficiency possible it is desirable to use the logical OR of all the available single muon triggers. This gives a significant increase in the efficiency when compared to any individual single muon trigger. This behavior as function of η_{det} is shown in Fig. 8.8.

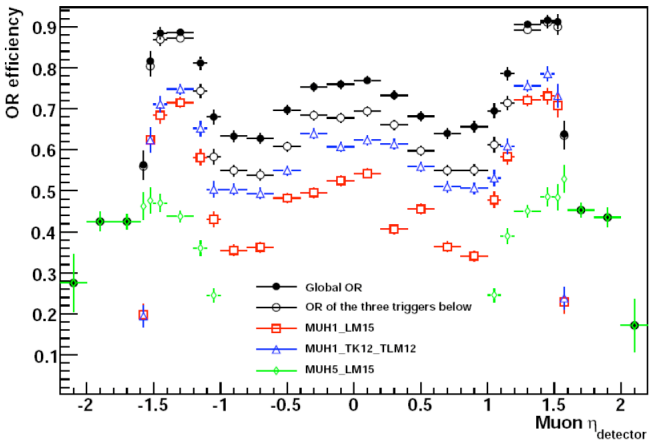


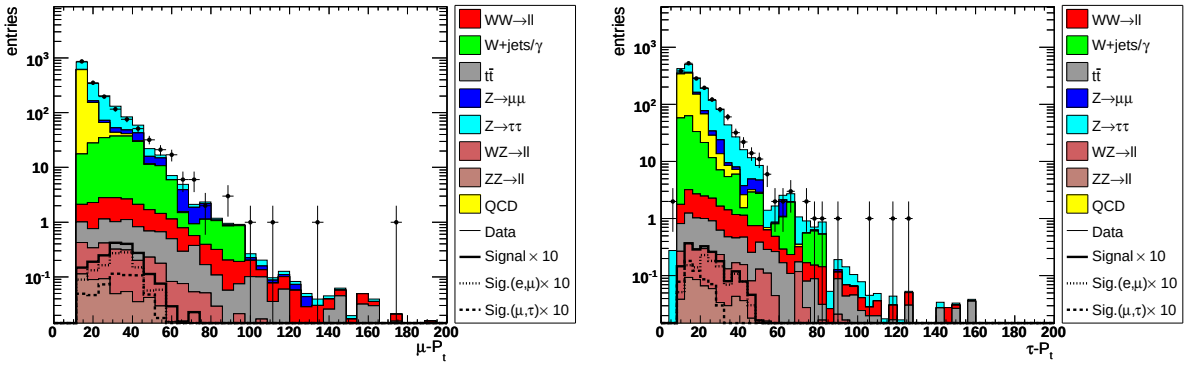
Figure 8.8: Comparison of the efficiency for the OR of the single muon triggers with three different single muon triggers as a function of η_{det} for the Run IIa data sample. The logical OR of all trigger, represented by the black markers, gives a much higher efficiency than any single trigger [107].

The data selection implicitly features a very high efficiency and no differences are observed when requiring the muon OR. Therefore no further trigger requirements are applied. It may be noted that in the soon available update of the $H \rightarrow WW^* \rightarrow \mu + \tau_{had} + X$ analysis dedicated τ triggers are used, leading to a higher purity of the sample at the preselection stage. Those triggers were not available at the time of performing the presented analysis.

8.3.2 Event Selection

The muon identification is based on the "loose" muon criteria defined by the Muon ID Group [108]. Muons are required to have a transverse momentum larger than 12 GeV. Details of the

muon identification criteria and the requirements for a "loose" muon are discussed in Sec. 4.5. The tau has to be selected as τ-type I or τ-type II and to pass a tau-identification neural network cut of $NN > 0.9$, see Sec. 4.6 for details. Tau type III is neglected due to the high background contamination. The tau candidates are required to exceed a transverse momentum of $p_T > 10$ GeV. Leptons are required to originate from the same primary vertex ($\Delta z(\ell$, primary vertex$) < 1.5$ cm), not being matched to each other ($\Delta \mathrm{R}(\mu, \tau) > 0.4$) and to be of opposite charge. All remaining objects such as jets and $\not\!\!\mathrm{E}_{\mathrm{T}}$ are treated as described in Chapter 4. Comparisons between data and MC are shown in Figs. 8.11 and 8.12 for the pseudorapidity of muons and taus. The p_T distributions for the selected muon and tau at preselection level are shown in Fig. 8.9 for τ -type I and Fig. 8.10 for τ-type II.

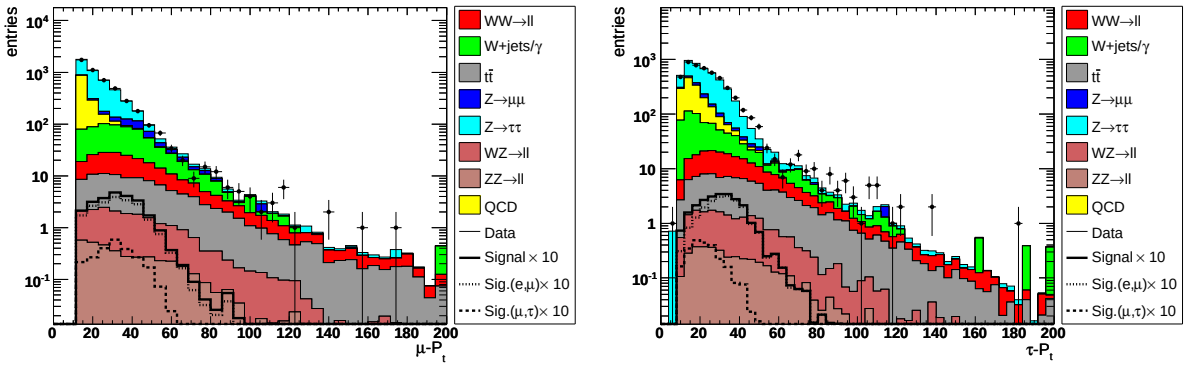


$p_T(\mu)$ $p_T(\tau)$

Figure 8.9: Distributions of the muon and tau p_T at preselection level for the $\mu + \tau_{\mathrm{type\ I}}$ final state for data (points) and sum of all backgrounds (filled histograms). The expected signal, multiplied by a factor of 10, for the Standard Model Higgs of mass 160 GeV is shown as well.

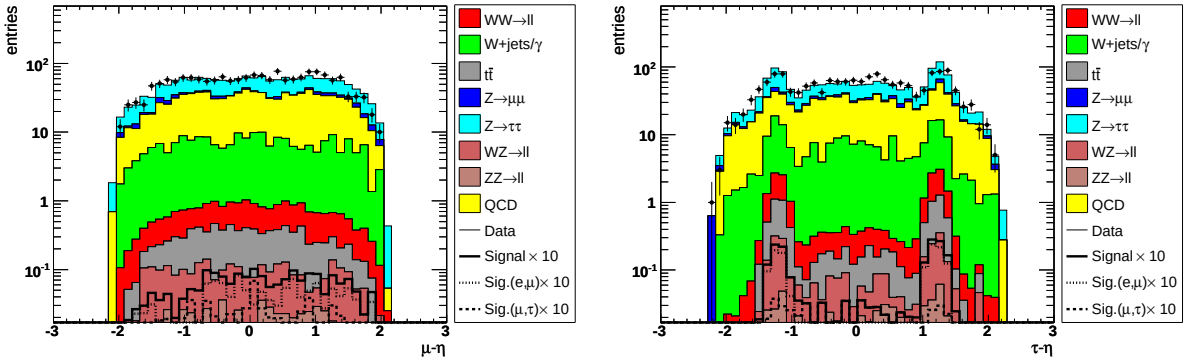
The $\mu + \tau_{had} + X$ analysis uses also a two-folded approach for background suppression. First a cut-based analysis followed by application of a likelihood discriminant. Because the kinematic differences of the various analyzed Higgs masses are distinguished and the composition of the background processes change significantly depending on the possible Higgs boson mass, the cuts are adapted for each analyzed Higgs boson mass. A comparison of the cuts for the various Higgs boson masses is given in Table 8.1.

- The selected leptons are required to have opposite charge and must have at least p_T >12 GeV for the muon and p_T >10 GeV for the tau candidate.
- A cut on the missing transverse energy in the event $\not\!\!\mathrm{E}_{\mathrm{T}} > 20$ GeV is applied to suppress QCD and $Z +$ jet/γ background.
- To remove most of the multijet background the requirement $\not\!\!\mathrm{E}_{T}^{Sig} > 7$ is applied.
- A cut on the minimal transverse energy requiring $M_{min}^{T} > 45$ GeV is applied.



$p_T(\mu)$ $p_T(\tau)$

Figure 8.10: Distributions of the muon and tau p_T at preselection level for the $\mu + \tau_{\text{type II}}$ final state for data (points) and sum of all backgrounds (filled histograms). The expected signal, multiplied by a factor of 10, for the Standard Model Higgs of mass 160 GeV is shown as well.



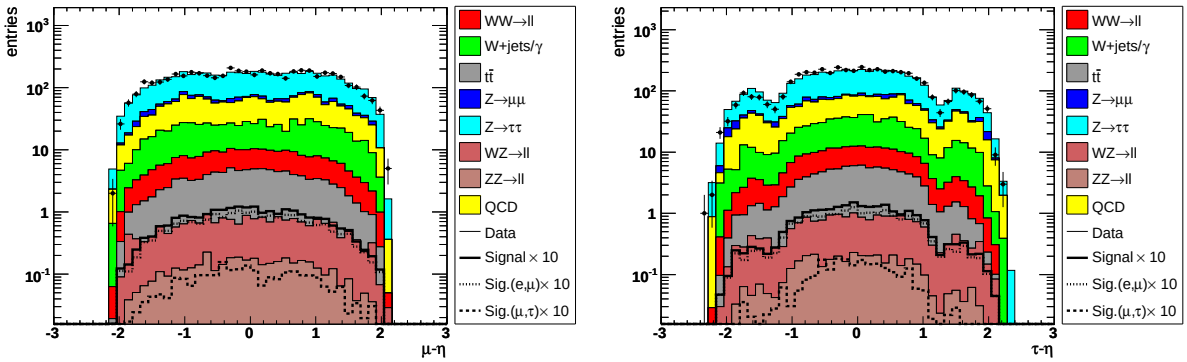
$\mu - \eta_{det}$ $\tau - \eta_{det}$

Figure 8.11: Distributions of the muon and tau η_{det} at preselection level for $\mu + \tau_{\text{type I}}$ final state for data (points) and sum of all backgrounds (filled histograms). The expected signal, multiplied by a factor of 10, for the Standard Model Higgs of mass 160 GeV is shown as well.

- The requirements on the sum of the transverse momenta $p_T^{\mu}+p_T^{\tau}+\not{E}_{\mathrm{T}}$, the minimal transverse mass $M_{min}^{T}(\ell, \not{E}_{\mathrm{T}})$ and the invariant mass $M_{\mu\tau}$ are optimized to reject contributions from $Z + \text{jet}/\gamma$, WW, WZ, ZZ and $W +$ jets backgrounds processes.

- Top pair production shows a high multiplicity of jets with large transverse momentum. By requiring the scalar sum of the p_T of jets (H_T) to be less than 70 GeV ("Cut 7") most $t\bar{t}$ are removed.

- Taking advantage of the narrow angular distributions of the leptons from spin 0 decays the



$\mu - \eta_{det}$ $\tau - \eta_{det}$

Figure 8.12: Distributions of the muon and tau η_{det} at preselection level for $\mu + \tau_{\text{type II}}$ final state for data (points) and sum of all backgrounds (filled histograms). The expected signal, multiplied by a factor of 10, for the Standard Model Higgs of mass 160 GeV is shown as well.

opening angle between the leptons in the transverse plane $\Delta\phi(\ell, \ell)$ is required to be smaller than 2.0.

At the final cut stage after applying selection requirements a likelihood discriminant is used to further suppress the remaining background processes. The discriminant is discussed in detail in Sec. 8.4. This approach exploits the differences of the event kinematics for Higgs boson decays and the quantities used to identify tau candidates.

Selection criteria		$m_{\text{H}} = 120$	$m_{\text{H}} = 140$	$m_{\text{H}} = 160$	$m_{\text{H}} = 180$
Cut 1	Missing Transverse Energy $\not\!\!E_{\text{T}}$	> 20	> 20	> 20	> 20
Cut 2	$\not\!\!E_T^{Sig}$	> 7	> 7	> 7	> 7
Cut 3	$M_{min}^T(\ell, \not\!\!E_{\text{T}})$	> 35	> 40	> 45	> 45
Cut 4	Sum of $p_T^l + p_T^{l'} + \not\!\!E_{\text{T}}$	50-140	60-150	70-160	80-180
Cut 5	Invariant mass $M_{\mu\tau}$	< 50	< 60	< 60	< 80
Cut 6	H_T	< 70	< 70	< 70	< 70
Cut 7	$\Delta\phi(\ell, \ell)$	< 2	< 2	< 2	< 2

Table 8.1: Summary of the selection criteria for the analyzed Higgs masses $m_{\text{H}} = 120$ GeV, $m_{\text{H}} = 140$ GeV, $m_{\text{H}} = 160$ GeV and $m_{\text{H}} = 180$ GeV .

There are two main contributions to the Higgs signal, $H \rightarrow WW \rightarrow \mu\tau_{had}\nu\nu$ and $H \rightarrow WW \rightarrow \mu e\nu\nu$ where the electron is mis-identified but satisfies the τ selection criteria. In the following plots the contribution of the $e\mu\nu\nu$ final state is indicated by a small dashed line, while the $\mu + \tau_{had} + X$ signal is represented by a bold dashed line. The sum of the two signal contributions

is given by a solid black line. The analysis is optimized to simultaneously maximize both signal contributions.
In this and the following sections distributions and cuts corresponding to a Higgs boson mass of $m_{\mathrm{H}} = 160$ GeV are shown. The data sets corresponds to the Run IIa sample. More details regarding the remaining Higgs boson masses of $m_{\mathrm{H}} = 120$ GeV, $m_{\mathrm{H}} = 140$ GeV and $m_{\mathrm{H}} = 180$ GeV are given in Ref. [109]. Figure 8.13 shows the missing transverse energy distribution at preselection stage for the $\mu + \tau_{had} + X$ final state for tau-type I and tau-type II. Figures 8.14 and 8.15 shows all quantities which are used for applying selection criteria in the tau-type I final state. The corresponding distributions for tau-type II are displayed in Figs 8.16 and 8.17.

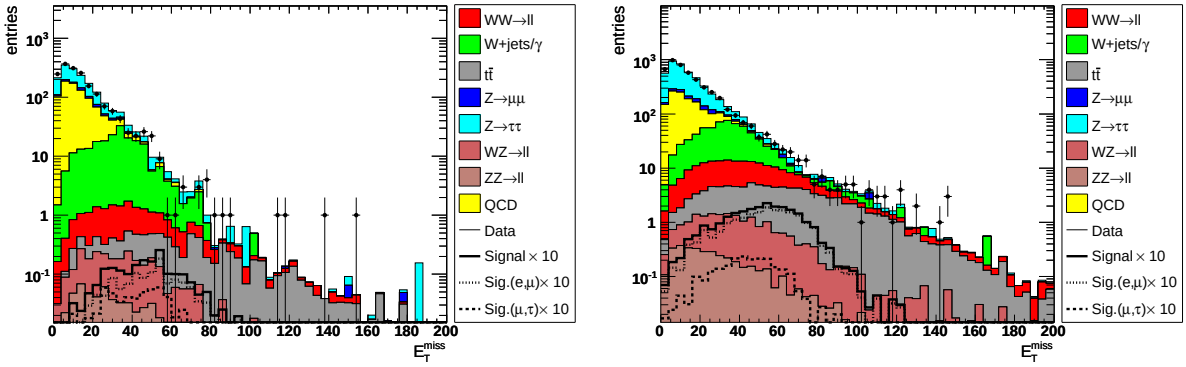


Figure 8.13: Distribution of the missing transverse energy $\not\!\!E_{\mathrm{T}}$ at preselection level for the $\mu + \tau_{had} + X$ final state for tau-type I (left) and tau-type II (right).The data is represented by points, the background by filled histograms. The expected signal, multiplied by a factor of 10, for the Standard Model Higgs with a mass of 160 GeV is also shown.

8.3.3 Veto on Events selected by the $H \rightarrow WW^* \rightarrow e^{\pm}\nu\mu^{\mp}\nu$ Analysis

To avoid double counting events already in the $H \rightarrow WW^* \rightarrow e^{\pm}\nu\mu^{\mp}\nu$ analysis are rejected. This veto is applied at the last cut stage, leading to a significant drop in sensitivity. The signal yield before and after applying the veto are also listed in Tables 8.4, 8.5.

8.4 Likelihood Discriminant

After applying the selection requirements as discussed in Sec. 8.3 no further reduction of the background can easily be achieved any more by imposing requirements on individual distributions. Therefore a likelihoods discriminant is applied to further discriminate between signal and background processes.

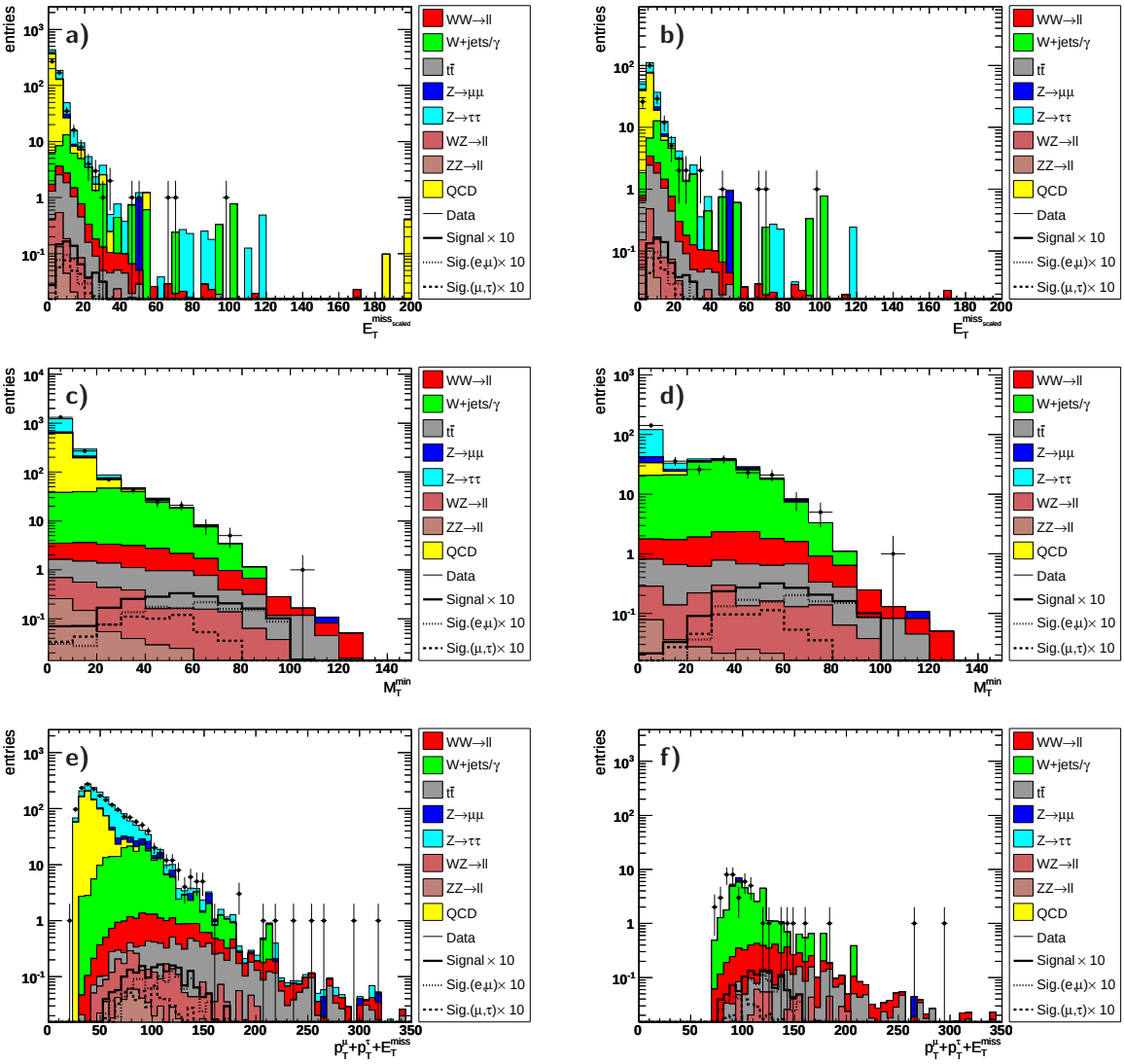


Figure 8.14: Distributions of kinematic quantities which are used in the event selection for tau-type I. Each distribution is shown at preselection stage and before the cut on the quantity is placed according to Table 8.1. The expected signal, multiplied by a factor of 10 is also shown for a the Standard Model Higgs boson assuming a mass of 160 GeV.
a) shows the $\not\!\!E_T^{Sig}$ distribution at preselection stage and b) after applying the selection requirement on $\not\!\!E_{\mathrm{T}}$,
c) shows the $M_{min}^{T}(\ell, \not\!\!E_{\mathrm{T}})$ distribution at preselection stage and d) after applying the selection requirement on $\not\!\!E_T^{Sig}$,
e) shows the $\sum p_T$ distribution at preselection stage and f) after applying the selection requirement on $M_{min}^{T}(\ell, \not\!\!E_{\mathrm{T}})$.

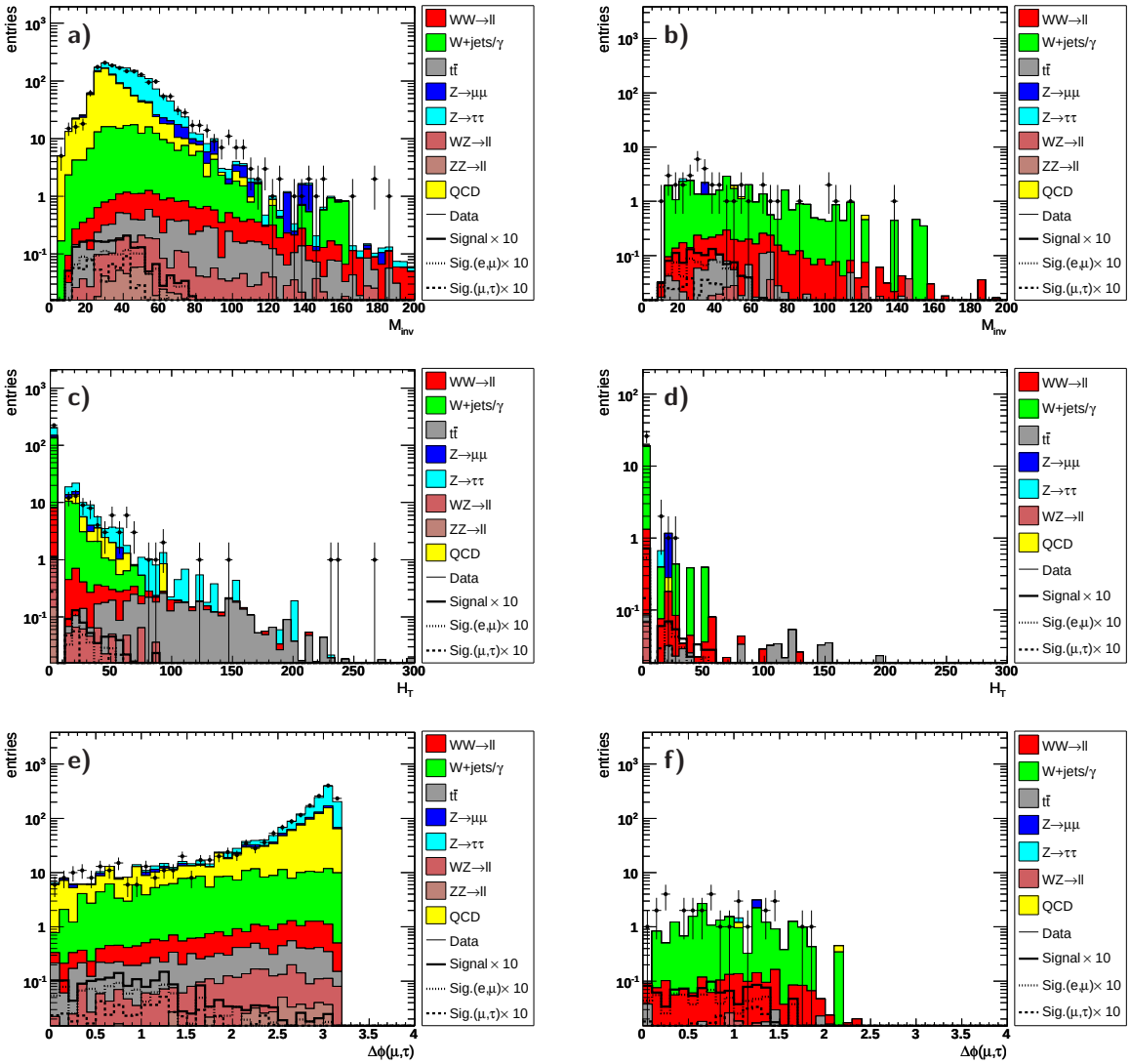


Figure 8.15: Distributions of kinematic quantities which are used in the event selection for tau-type I. Each distribution is shown at preselection stage and before the cut on the quantity is placed according to Table 8.1. The expected signal, multiplied by a factor of 10 is also shown for a the Standard Model Higgs boson assuming a mass of 160 GeV.
a) shows the $M_{\ell\ell}$ distribution at preselection stage and b) after applying the selection requirement on $\sum p_T$,
c) shows the H_T distribution at preselection stage and d) after applying the selection requirement on $M_{\ell\ell}$,
e) shows the $\Delta\phi(\ell,\ell)$ distribution at preselection stage and f) after applying the selection requirement on H_T.

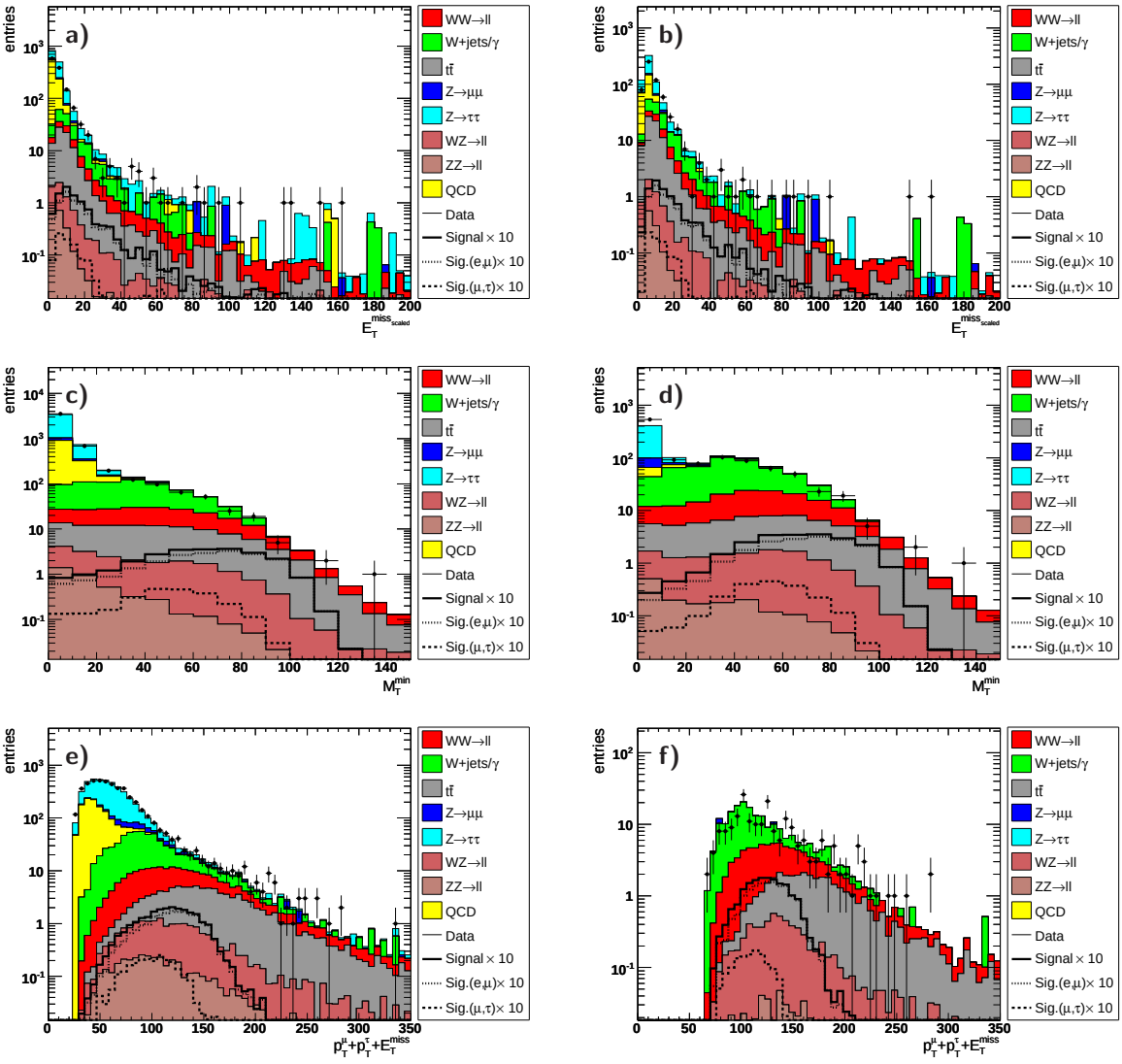


Figure 8.16: Distributions of kinematic quantities which are used in the event selection for tau-type II. Each distribution is shown at preselection stage and before the cut on the quantity is placed according to Table 8.1. The expected signal, multiplied by a factor of 10 is also shown for a the Standard Model Higgs boson assuming a mass of 160 GeV.
a) shows the $\not\!\!E_T^{Sig}$ distribution at preselection stage and b) after applying the selection requirement on $\not\!\!E_{\mathrm{T}}$,
c) shows the $M_{min}^{T}(\ell, \not\!\!E_{\mathrm{T}})$ distribution at preselection stage and d) after applying the selection requirement on $\not\!\!E_T^{Sig}$,
e) shows the $\sum p_T$ distribution at preselection stage and f) after applying the selection requirement on $M_{min}^{T}(\ell, \not\!\!E_{\mathrm{T}})$.

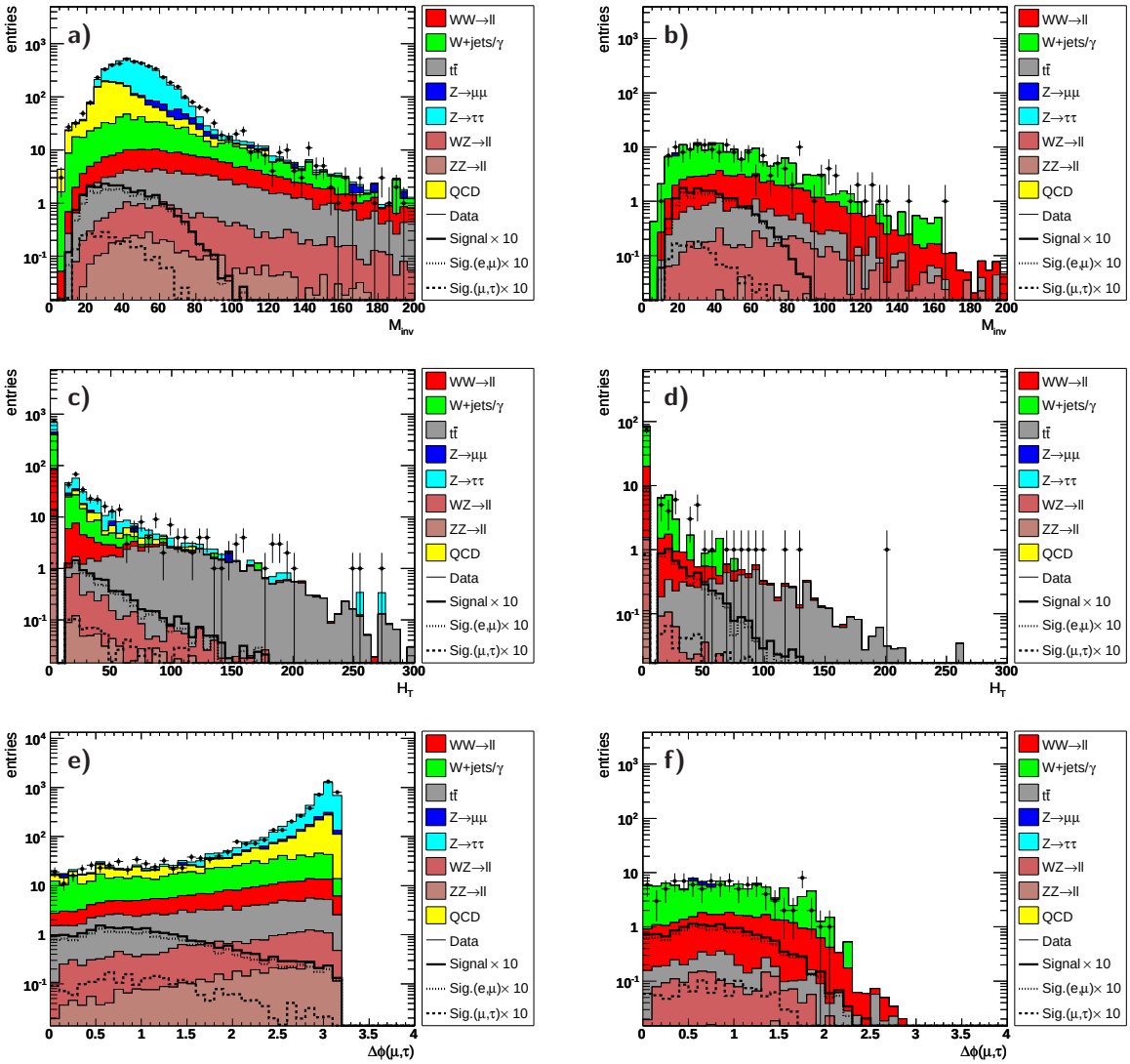


Figure 8.17: Distributions of kinematic quantities which are used in the event selection for tau-type II. Each distribution is shown at preselection stage and before the cut on the quantity is placed according to Table 8.1. The expected signal, multiplied by a factor of 10 is also shown for a the Standard Model Higgs boson assuming a mass of 160 GeV.
a) shows the $M_{\ell\ell}$ distribution at preselection stage and b) after applying the selection requirement on $\sum p_T$,
c) shows the H_T distribution at preselection stage and d) after applying the selection requirement on $M_{\ell\ell}$,
e) shows the $\Delta\phi(\ell,\ell)$ distribution at preselection stage and f) after applying the selection requirement on H_T.

8.4.1 Likelihood Construction

The analysis uses two separate likelihoods, one to enhance the contribution of $e\mu$ events in which the electron is not well isolated and mis-identified as tau, the second likelihoods is designed to enrich real $\mu + \tau_{had} + X$ events. The best sensitivity is achieved by combining both likelihoods. Each likelihood is constructed using a separate set of quantities which emphasize different signal properties, the first set is related to variables used in the τ identification and the second is based on kinematic properties, they will be referred to as *tau likelihood* or *kinematic likelihood*, respectively. The following variables are used for construction of the τ likelihood:

- **EMF**: Fraction of the calorimeter energy deposited in the EM subcluster
- $\mathbf{p_T^{\tau} + p_T^{trk}}$: Sum of the transverse momenta of the track assigned to the tau and of the energy of the tau-candidate deposited in the calorimeter
- τ **ID NN**: Output of the tau-identification neural network
- $\mathbf{E_T/p_T^{trk1}}$ Ratio of Energy of calorimeter cluster and p_T of the leading track
- $\sum \mathbf{p_T^{trk}}(\mathbf{\Delta}\mathrm{R}(\tau, \mathrm{trk}) < \mathbf{0.4})$: Sum p_T of all tracks in a Cone of 0.4 around the tau candidate

The second likelihood is based on kinematic quantities of the event:

- $\mathbf{p_T^{\mu}}$: p_T of the muon
- $\mathbf{M_T}(\mu, \not{\mathbb{E}}_{\mathrm{T}}$**)** Transverse mass of μ and $\not{\mathbb{E}}_{\mathrm{T}}$
- $\mathbf{M_T^{min}}(\ell, \not{\mathbb{E}}_{\mathrm{T}}$**)** Minimal transverse mass of either of the leptons and $\not{\mathbb{E}}_{\mathrm{T}}$
- $\mathbf{M_c} = \sqrt{\mathbf{p_T^2}(\ell\ell) + \mathbf{m^2}(\ell\ell)} + \not{\mathbb{E}}_{\mathrm{T}}$: cluster mass, approximation for the $H \rightarrow WW$ mass
- $\mathbf{\Delta}\phi(\mu, \tau)$: Angle between μ and τ candidate in the transverse plane
- $\mathbf{\Delta\Theta}(\mu, \tau)$: Solid angle between μ and τ candidate

At the final cut stage the background is dominated by W + jets production (Table 8.4, 8.5). The available Monte Carlo statistics is divided into two sets of equal size, one is used for construction and the other for analysis purpose. The reference distributions for the construction of the kinematic likelihood are obtained after applying all selection requirements. To minimize fluctuations of the reference distributions for the τ based likelihood the histograms used to construct the likelihood are selected after the $\not{\mathbb{E}}_{\mathrm{T}}$ requirement because the following selection requirements do not change the shape of the tau candidate related quantities whereas the available statistic is at this stage greatly enhanced.

The likelihoods are constructed according to

$$\mathcal{L} = \frac{\mathcal{P}_{Sig}(x_1, x_2, \ldots)}{\mathcal{P}_{Sig}(x_1, x_2, \ldots) + \mathcal{P}_{Bkgd}(x_1, x_2, \ldots)} \tag{8.1}$$

$$= \frac{\prod_i \mathcal{P}^i_{Sig}}{\prod_i \mathcal{P}^i_{Sig} + \prod_i \mathcal{P}^i_{Bkgd}} = \frac{\prod_i \mathcal{P}^i_{Sig}/\mathcal{P}^i_{Bkgd}}{\prod_i \mathcal{P}^i_{Sig}/\mathcal{P}^i_{Bkgd} + 1} \tag{8.2}$$

Here $\mathcal{P}^i_{Sig}$ represents the number of expected signal events and $\mathcal{P}^i_{Bkgd}$ the number of expected background events for bin i. The kinematic distribution on which the likelihood is based are given by x_i. The input distributions are shown in Figs. C.1 and C.2 in Appendix C.2.

In general the events selected by mis-identifying an electron as tau resemble more closely the $WW \rightarrow \ell^+\nu\ell^-\nu$ background whereas the shape of the hadronic events is similar to $W+$jets production. Therefore two classes are selected where the classification is based on the Monte Carlo true information. Using the corresponding input distributions for these classes two distinguished likelihoods for either case are constructed and applied to the selected samples.

- $H \rightarrow WW^* \rightarrow e^{\pm}\nu\mu^{\mp}\nu$ events where one electron fails the electron criteria required in the $e\mu$ analysis. Because this sample contains the mis-identified electrons it will be referred to as **inclusive sample** in the following.

- $H \rightarrow WW \rightarrow \mu\tau_{had} + X$: Events containing a 'real' tau decaying hadronically, therefore excluding electron pick-ups on Monte Carlo true level. This will be referred to as **exclusive sample** in the following.

The two likelihood discriminants based on these two samples are called correspondingly *inclusive* and *exclusive* likelihoods. To achieve optimal separation power these likelihoods are applied separately to each tau type and for each Higgs mass studied. The final combination is performed such that an event selected by both, the *inclusive* and the *exclusive* likelihood, is not double counted. The signal contribution from final states such as $H \rightarrow WW \rightarrow \mu\tau \rightarrow \mu + e + \nu$ has been studied but is negligible. The low statistics of tau-type I compared to tau type II leads to large fluctuations of the input distributions, therefore impairing the separation power of the resulting likelihoods to the point where the application of a dedicated, tau-type I based likelihood is ineffective. Therefore the input distributions of tau-type II are used for all tau types, leading to an improved sensitivity. The likelihood distributions at the final cut stage are shown in Figs. 8.19, 8.18, 8.21 and 8.20. Comparing Figs. 8.18 and 8.19 shows that the kinematic likelihood offers a better separation power than the tau likelihood. Therefore only the kinematic likelihood is used for tau-type I.

In contrast to the kinematic likelihoods the shape of the two signal contributions differ significantly for the tau-likelihoods. When using $e+\mu$ events for the likelihood construction the $\mu+\tau_{had}$ contribution is shifted significantly towards the background-like region (Fig. 8.20 right). Therefore the separation power of true hadronic tau events is degraded, centering the distribution of the signal events around 0.5 (Fig. 8.20 left). This behavior is due to the case that $W+$jets production and hadronic $\tau+\mu$ events show similar shapes in many distributions (see also Fig. C.2 in Appendix C.2).

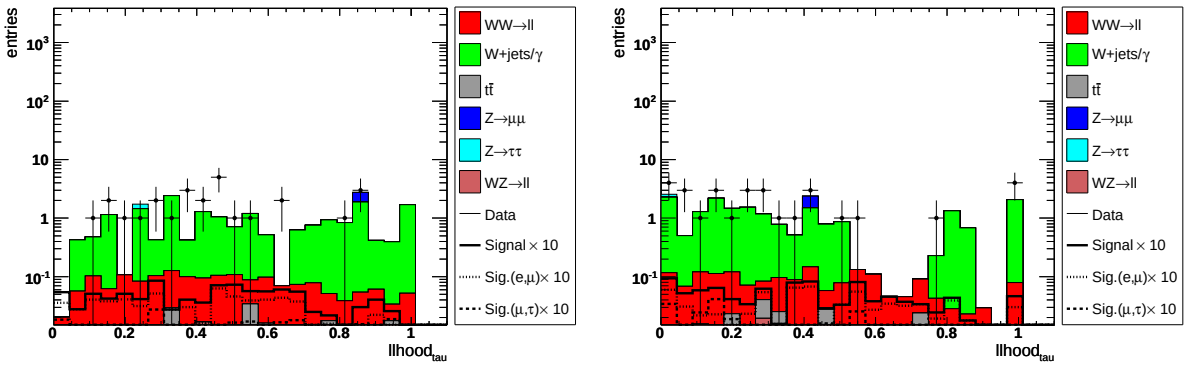


Figure 8.18: Tau likelihood for $\mu + \tau_{had}$ events with taus of type I (left) (*exclusive*) and $\mu + e$ (right) final states where an electron is reconstructed as tau (*inclusive*).

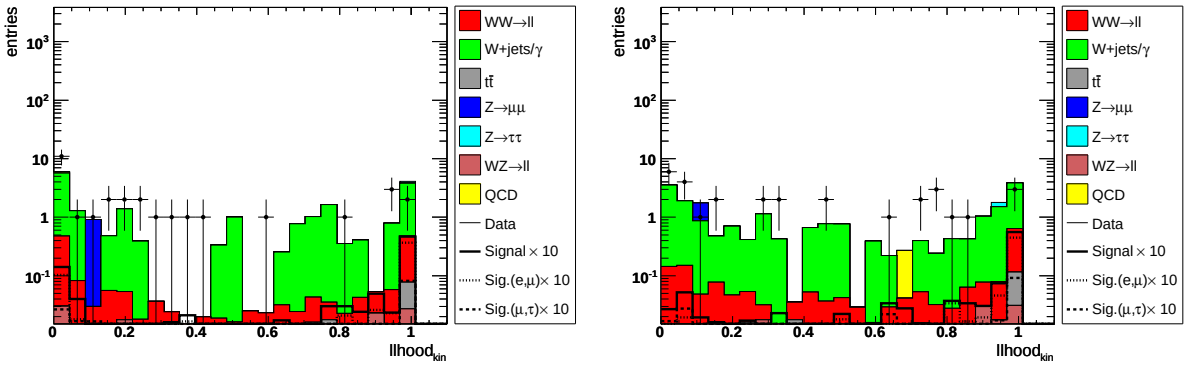


Figure 8.19: Kinematic likelihood for $\mu + \tau_{had}$ events with taus of type I (left) (*exclusive*) and $\mu + e$ (right) final states where an electron is reconstructed as tau (*inclusive*).

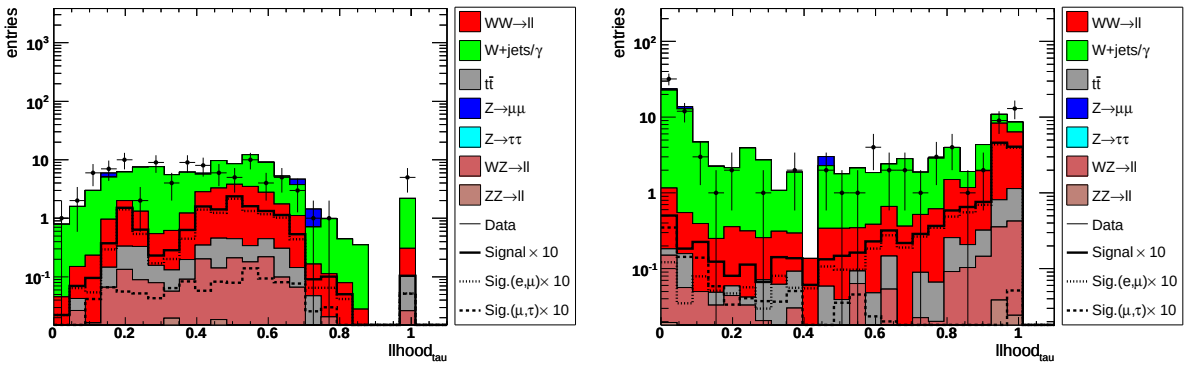


Figure 8.20: Tau likelihood for $\mu + \tau_{had}$ events with taus of type II (left) (*exclusive*) and $\mu + e$ (right) final states where an electron is reconstructed as tau (*inclusive*).

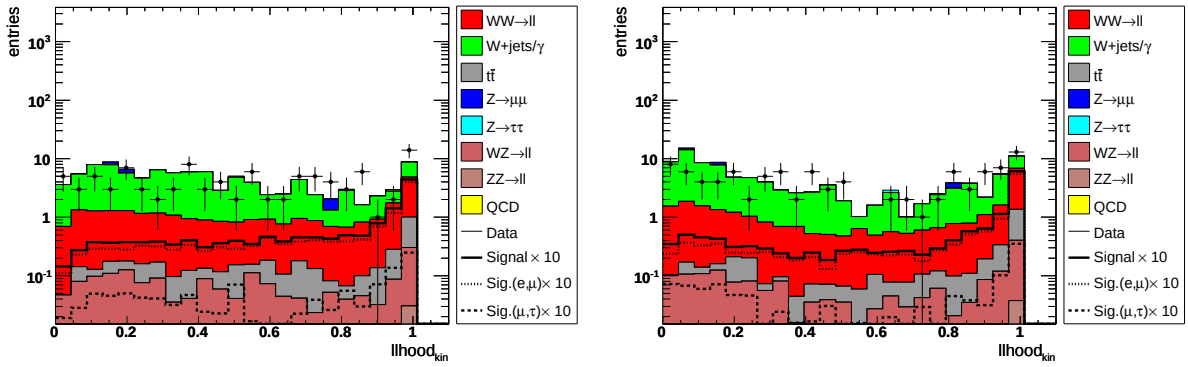


Figure 8.21: Kinematic likelihood for $\mu + \tau_{had}$ events with taus of type II (left) (*exclusive*) and $\mu + e$ (right) final states where an electron is reconstructed as tau (*inclusive*).

8.4.2 Likelihood Cuts

The likelihood output probabilities of the kinematic ($\mathcal{L}_{kin}$) and tau likelihoods ($\mathcal{L}_{\tau}$) are combined to construct a 2 dimensional discriminant. The 2D discriminant for τ-type I and τ-type II are shown in Fig. 8.22 and Fig. 8.23. The background contribution can be reduced relative to the signal contribution by placing appropriate selection requirements on this discriminant. These cuts have been developed by scanning a variety of possible selections maximizing the $S/\sqrt{S+B}$ ratio where S is the expected number of signal events and B is the expected number of background events. The optimal cuts are listed in Tables 8.2 and 8.3.

training sample	τ-type	$m_{\mathrm{H}} = 160$	$m_{\mathrm{H}} = 180$
inclusive	I	$0.4 < \mathcal{L}_{tau} < 0.8$	$\mathcal{L}_{kin} > 09$
inclusive	II	$\mathcal{L}_{tau} < 0.85$	$\mathcal{L}_{tau} > 0.9$
exclusive.	I	$\mathcal{L}_{tau} < 0.65$	$0.3 < \mathcal{L}_{tau} < 0.85$ and $\mathcal{L}_{kin} < 0.1$
exclusive	II	$0.65 < \mathcal{L}_{tau} < 0.8$	$0.6 < \mathcal{L}_{tau} < 0.8$ and $\mathcal{L}_{kin} < 0.8$

Table 8.2: List of the selection requirements applied on the 2D likelihood planes for both likelihood training classes and both selections.

training sample	τ-type	$m_{\mathrm{H}} = 120$	$m_{\mathrm{H}} = 140$
inclusive	I	$0.4 < \mathcal{L}_{tau} < 0.8$	$\mathcal{L}_{kin} > 09$
inclusive	II	$\mathcal{L}_{tau} < 0.85$	$\mathcal{L}_{tau} > 0.9$
exclusive.	I	$\mathcal{L}_{tau} < 0.65$	$0.3 < \mathcal{L}_{tau} < 0.85$ and $\mathcal{L}_{kin} < 0.1$
exclusive	II	$0.65 < \mathcal{L}_{tau} < 0.8$	$0.6 < \mathcal{L}_{tau} < 0.8$ and $\mathcal{L}_{kin} < 0.8$

Table 8.3: List of the selection requirements applied on the 2D likelihood planes for both likelihood training classes and both selections.

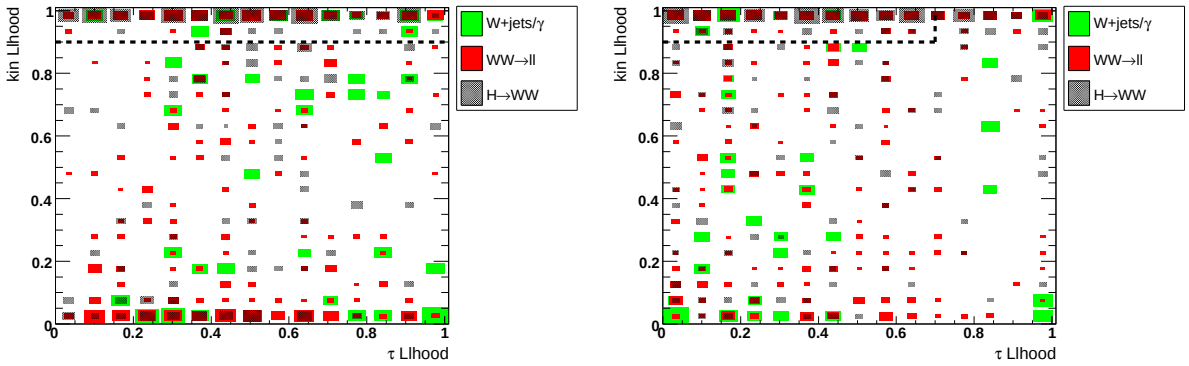
The selections applied are indicated by a dashed black line in Figs. 8.22 and 8.23.

8.5 Results of the $\mu + \tau_{had} + X$ analysis

Applying these selection cuts no excess in data above the background expectation is observed and upper limits on the $\sigma \times BR(H \rightarrow WW^* \rightarrow \mu + \tau_{had} + X)$ are set, including various sources of systematic uncertainties.

8.5.1 Event yield

After application of all selection requirements as listed in Sec. 8.4.2 the likelihoods are applied separately for the sample including electrons mis-identified as taus and the exclusive samples. A veto is applied, rejecting events selected by $\mu + \tau_{had}$ events if those are as well selected by the $\mu + e$



τ likelihood versus kinematic likelihood for the exclusive $H \to WW \to \mu\tau_{had} + \nu\nu$ sample, τ-type I.

τ likelihood versus kinematic likelihood for the inclusive $H \to WW \to e\mu + \nu\nu$ sample, τ-type I.

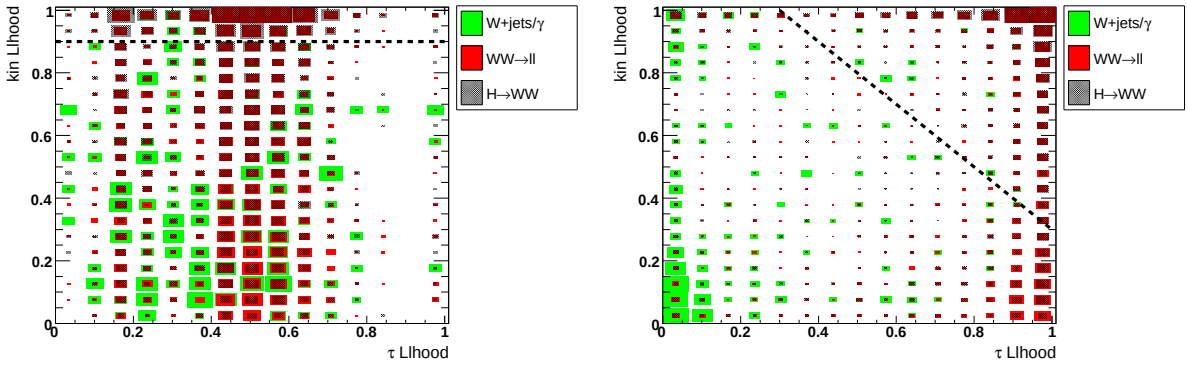
Figure 8.22: τ likelihood versus kinematic likelihood after Cut 8 ($\Delta\phi(\mu, \tau)$) for the inclusive Likelihoods and τ-type I. The applied cut is listed in Table 8.2 and Table 8.3 and is indicated by the black dashed line. The distributions are normalized to their integral.

likelihood. This veto is based on unique identifier assigned to each individual event, for example run- and event number for the data. The correlation of the events selected by one of the two likelihoods is small.

The number of events observed in data and the expected signal and background contributions are shown in Table 8.4 for tau type I at the various selection stages and in Table 8.5 for tau-type II.

Cut	Data	Sum Bkgd	Signal	WW	$W \rightarrow \mu\nu$	$W \rightarrow e\nu$	$W \rightarrow \tau\nu$	$t\bar{t}$	$Z \rightarrow \mu\mu$	$Z \rightarrow \tau\tau$	WZ	ZZ	QCD
Preselection	1749.00 $\pm$ 41.82	1719.19 $\pm$ 33.58	0.20	13.06	176.35	4.43	16.83	6.50	64.36	671.07	2.24	0.58	763.78
$\not\!\!E_{\mathrm{T}}$	408.00 $\pm$ 20.20	441.58 $\pm$ 15.30	0.18	10.63	141.54	3.85	10.73	6.04	18.65	128.52	1.75	0.39	119.50
$\not\!\!E_{\mathrm{T}}^{Sig}$	301.00 $\pm$ 17.35	292.75 $\pm$ 10.71	0.17	9.84	138.36	3.85	10.73	4.10	16.05	88.69	1.40	0.24	19.51
M_T^{min}	44.00 $\pm$ 6.63	39.77 $\pm$ 3.53	0.12	4.07	28.91	1.82	0.50	1.80	0.92	0.28	0.56	0.06	0.84
$\sum p_T$	40.00 $\pm$ 6.32	35.89 $\pm$ 3.39	0.12	3.29	27.44	1.82	0.50	0.84	0.88	0.28	0.35	0.04	0.43
M_{inv}	30.00 $\pm$ 5.48	22.62 $\pm$ 2.78	0.11	1.70	16.45	1.82	0.50	0.52	0.88	0.28	0.09	0.02	0.33
H_T	30.00 $\pm$ 5.48	22.18 $\pm$ 2.78	0.11	1.64	16.45	1.82	0.50	0.14	0.88	0.28	0.09	0.02	0.33
$\Delta\phi(\mu,\tau)$	30.00 $\pm$ 5.48	21.66 $\pm$ 2.74	0.11	1.59	16.12	1.82	0.50	0.12	0.88	0.28	0.09	0.02	0.23
Llhood incl.	2.00 $\pm$ 1.41	4.63 $\pm$ 1.22	0.05	0.45	3.12	0.66	0.00	0.07	0.00	0.28	0.05	0.01	0.00
Veto $e\mu$.	2.00 $\pm$ 1.41	4.63 $\pm$ 1.22	0.05	0.45	3.12	0.66	0.00	0.07	0.00	0.28	0.05	0.01	0.00
Llhood excl.	5.00 $\pm$ 2.24	4.91 $\pm$ 1.19	0.05	0.46	4.08	0.00	0.00	0.05	0.00	0.28	0.04	0.01	0.00
Veto $e\mu$.	3.00 $\pm$ 1.73	1.78 $\pm$ 0.68	0.01	0.18	1.22	0.00	0.00	0.05	0.00	0.28	0.04	0.01	0.00

Table 8.4: Number of expected background events and number of observed data events, after successive selections for τ-type I. The expected signal content is given with respect to $m_{\mathrm{H}} = 160$ GeV. After the $\Delta\phi(\mu,\tau)$ the likelihood cuts for the inclusive ($H \rightarrow WW \rightarrow \mu e\nu\nu$) sample and the exclusive ($H \rightarrow WW \rightarrow \mu\tau_{had}\nu\nu$) sample are applied individually and a veto against the $H \rightarrow WW^* \rightarrow e^{\pm}\nu\mu^{\mp}\nu$ analysis is applied. The data set corresponds to 1.0 fb^{-1} in integrated luminosity. Only statistical uncertainties are given.



τ likelihood versus kinematic likelihood for the exclusive $H \to WW \to \mu\tau_{had} + \nu\nu$ sample, τ-type II.

τ likelihood versus kinematic likelihood for the inclusive $H \to WW \to e\mu + \nu\nu$ sample, τ-type II.

Figure 8.23: τ likelihood versus kinematic likelihood after Cut 8 ($\Delta\phi(\mu, \tau)$) for the inclusive Likelihoods and τ-type II. The applied cut is listed in Table 8.2 and Table 8.3 and is indicated by the black dashed line. The distributions are normalized to their integral.

Cut	Data	Sum Bkgd	Signal	WW	$W \rightarrow \mu\nu$	$W \rightarrow e\nu$	$W \rightarrow \tau\nu$	$t\bar{t}$	$Z \rightarrow \mu\mu$	$Z \rightarrow \tau\tau$	WZ	ZZ	QCD
Preselection	4786.00 ± 69.18	4735.08 ± 47.95	2.47	129.94	401.26	28.40	42.14	83.35	176.41	2778.40	15.96	3.88	1075.35
$\not\!\!E_{\mathrm{T}}$	1332.00 ± 36.50	1348.85 ± 22.76	2.31	106.77	335.08	23.87	24.51	77.77	62.86	524.67	13.14	2.40	177.77
$\not\!\!E_T^{Sig}$	1058.00 ± 32.53	956.66 ± 17.06	2.20	100.96	321.06	21.18	21.90	51.58	53.99	335.10	11.52	1.64	37.73
M_T^{min}	208.00 ± 14.42	215.37 ± 6.38	1.80	57.12	103.39	7.99	5.28	31.30	2.63	0.29	6.24	0.51	0.62
$\sum p_T$	162.00 ± 12.73	168.43 ± 6.03	1.69	43.86	92.42	7.47	4.87	11.69	2.51	0.28	4.53	0.35	0.47
M_{inv}	107.00 ± 10.34	109.79 ± 5.10	1.50	24.45	64.26	5.60	3.64	6.96	2.51	0.28	1.69	0.15	0.25
H_T	99.00 ± 9.95	104.50 ± 5.09	1.44	23.86	64.01	5.60	3.64	2.51	2.51	0.28	1.69	0.14	0.25
$\Delta\phi(\mu,\tau)$	98.00 ± 9.90	101.90 ± 5.05	1.42	22.90	62.66	5.60	3.64	2.32	2.51	0.28	1.61	0.14	0.25
Llhood incl.	25.00 ± 5.00	23.41 ± 1.86	1.02	11.88	5.99	2.59	0.48	1.63	0.01	0.00	0.77	0.05	0.00
Veto $e\mu$.	15.00 ± 3.87	14.11 ± 1.76	0.27	3.67	4.90	2.59	0.48	1.63	0.01	0.00	0.77	0.05	0.00
Llhood excl.	14.00 ± 3.74	9.26 ± 1.44	0.20	2.13	4.91	0.71	0.00	0.94	0.02	0.00	0.36	0.03	0.16
Veto $e\mu$.	8.00 ± 2.83	7.08 ± 1.37	0.05	0.58	4.28	0.71	0.00	0.94	0.02	0.00	0.36	0.03	0.16

Table 8.5: Number of expected background events and number of observed data events, after successive selections for τ-type II. The expected signal content is given with respect to $m_{\mathrm{H}} = 160$ GeV. After the $\Delta\phi(\mu,\tau)$ the likelihood cuts for the inclusive ($H \rightarrow WW \rightarrow \mu e\nu\nu$) sample and the exclusive ($H \rightarrow WW \rightarrow \mu\tau_{had}\nu\nu$) sample are applied individually and a veto against the $H \rightarrow WW^* \rightarrow e^{\pm}\nu\mu^{\mp}\nu$ analysis is applied. The data set corresponds to 1.0 fb^{-1} of integrated luminosity. Only statistical uncertainties are given.

8.5.2 Systematics

The statistical uncertainty in the $H \rightarrow WW^* \rightarrow \mu + \tau_{had} + X$ analysis is much larger than for the $H \rightarrow WW^* \rightarrow e^+\nu e^-\nu$ analysis. Nevertheless studies are performed to identify sources of systematic uncertainty affecting both, the signal efficiency and background estimations. Shape uncertainties are neglected due to the larger statistical uncertainties compared to the systematic uncertainties. Details about the evaluation of systematics are given in Sec. 7.3. The sources of systematic uncertainties evaluated are:

- The muon track matching efficiency is 2.3% according to Ref. [108].
- The muon ID uncertainty is 0.5% according to Ref. [108].
- The tau ID uncertainty is 3.6% according to Ref. [89].
- The uncertainty on the tau and muon track reconstruction efficiency is with 4% the same according to Ref. [108].
- The multijet normalization uncertainty has been estimated to 5%. The uncertainties on the multijet normalization is estimated by applying the multijet fit using the opposite sign lepton sample and redoing the analysis.
- PDF uncertainties are $< 4\%$.
- Jet Energy Scale uncertainties, Jet ID uncertainties and JET resolution uncertainties vary between 0 and 2% for signal and background Monte Carlo for the various selections, tau types and variations. A detailed overview is given in Table 8.6.
- The systematic uncertainties on the tau energy scale has been evaluated by comparing the distributions of E_T/p_T^{trk} in data and Monte Carlo, where E_T is the tau energy measured by the calorimeter. The difference of the mean value of this distributions are used for rescaling the tau energies and re-running the analysis. This systematic uncertainties vary between $3-5\%$ for signal and $1-8\%$ for background Monte Carlo. Details are given in Ref. [106].
- The systematic uncertainty on the normalization factor is conservatively taken to be 10 %. It results from the NNLO $Z/\gamma \rightarrow \ell\ell$ cross section uncertainty (3.6%), the uncertainty on the multijet background (5%) and the uncertainty on the Z peak mass window (6%) added in quadrature.
- The systematic uncertainty on the description of the $W+$jets background is conservatively estimated and introduces an systematic uncertainty up to 17.5%.

An overview of all systematic uncertainties is given in Table 8.6.

The biggest systematic uncertainty is observed for τ-type I and therefore does not affect the main contribution of the final result. The systematics on the $W+$jets description is derived by comparing the event yields with and without applying the corrections. All systematic uncertainties are derived assuming a Higgs boson mass of $m_{\mathrm{H}} = 160$ GeV but the uncertainties are applied to

	τ-type I $e+\mu$		τ-type I $\mu+\tau$		τ-type I-type I $e+\mu$		τ-type II $\mu+\tau$	
	MC	Signal	MC	Signal	MC	Signal	MC	Signal
Track ID	2.3	2.3	2.3	2.3	2.3	2.3	2.3	2.3
Mu ID	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5
Tau ID	3.6	3.6	3.6	3.6	3.6	3.6	3.6	3.6
Mu Trk Rec	4	4	4	4	4	4	4	4
Tau Trk Rec	4	4	4	4	4	4	4	4
PDF	4	4	4	4	4	4	4	4
$W+$ jets xsec	17.5		2.5		6.8		12.1	
Tau Energy Scale	7.88	5.53	1.23	5.43	2.66	3.71	4.49	3.28
JES	0.56	0.6	0.51	0.0	0.71	0.88	0.38	0.56
Resolution	0.56	0.97	0.51	0.13	0.33	0.63	1.85	1.07
Eff.	0.56	1.91	0.51	1.88	0.50	0.73	1.65	0.47
$\sqrt{\sum_i \sigma_i^2}$	20.88	10.08	8.66	9.98	10.99	9.05	15.47	8.88

Table 8.6: Systematic uncertainties in % for all selected classes and tau types.

all Higgs mass points for the final limit calculation. The systematic uncertainties are dominated by the energy scales and the $W+$ jets background. The systematic uncertainties in this channel are much less relevant than the statistical uncertainties. Details of the systematic uncertainties for individual signal and background processes can be found in Table C.7 in Appendix C.3

	$m_{\mathrm{H}} = 120$ GeV		$m_{\mathrm{H}} = 140$ GeV		$m_{\mathrm{H}} = 160$ GeV		$m_{\mathrm{H}} = 180$ GeV	
	expected	observed	expected	observed	expected	observed	expected	observed
$\mu + e$	15.4 pb	15.4 pb	12.8 pb	10.7 pb	10.1 pb	16.2 pb	9.3 pb	9.7 pb
$\mu + \tau_{had}$	147.8 pb	130.3 pb	65.2 pb	41.2 pb	31.9 pb	20.9 pb	34.4 pb	44.4 pb

Table 8.7: Expected and observed cross section limits at 95% C.L. for events selected as $\mu + \tau_{had}$ and $\mu + e$ for the Higgs masses $m_{\mathrm{H}} = 120, 140, 160, 180$ GeV. The data sample used corresponds to 1.0 fb^{-1} of integrated luminosity.

8.5.3 Results

No evidence for a SM Higgs boson decaying to $\mu + \tau_{had} + X$ has been found in the analyzed sample and a limit on the production cross section times branching ratio has been set. The limit is calculated using the CL_s method with a log likelihood test statistics. Expected and observed limits at 95% C.L. for both event classes are presented in Table 8.7. The limits for each τ type and the combination of both can be found in Table 8.8 and in Fig. 8.24.

	$m_{\mathrm{H}} = 120$ GeV		$m_{\mathrm{H}} = 140$ GeV		$m_{\mathrm{H}} = 160$ GeV		$m_{\mathrm{H}} = 180$ GeV	
	expected	observed	expected	observed	expected	observed	expected	observed
τ-type I	90.2 pb	90.4 pb	68.7 pb	51.7 pb	28.2 pb	22.6 pb	29.4 pb	20.7 pb
τ-type II	14.8 pb	14.8 pb	13.6 pb	10.5 pb	9.9 pb	10.9 pb	10.3 pb	17.0 pb
comb.	15.1 pb	14.6 pb	13.2 pb	9.9 pb	9.2 pb	9.7 pb	9.7 pb	13.0 pb

Table 8.8: Cross section limits obtained for the combination of all contributing channels and for $m_{\mathrm{H}} = 120, 140, 160, 180$ GeV 95% C.L. The data sample used corresponds to 1.0 fb^{-1} of integrated luminosity.

8.5.4 Discussion

This analysis presented was the first analysis utilizing a hadronically decaying tau in a search for the Standard Model Higgs boson. By identifying the tau lepton explicitly through its decay to a hadronic jet and vetoing on the decays modes involving only electrons and muons a completely orthogonal channel is considered. The analysis has been presented at the Summer conferences 2007. Being the first attempt to use hadronic decaying taus in the Standard Model Higgs boson search it was discovered that adding this channel to the DØ Higgs combination leads to only negligible sensitivity improvement. Considering the improvements in tau-identification, background separation techniques and tau-trigger an update using very similar methods as the $H \rightarrow WW^* \rightarrow e^+\nu e^-\nu$ channel is currently carried out, based on 3.9 fb^{-1} of integrated luminosity. Preliminary combinations performed show the possibility of improving the DØ Higgs boson limit for about ∼5%. The update of the analysis will be made public in Fall 2009.

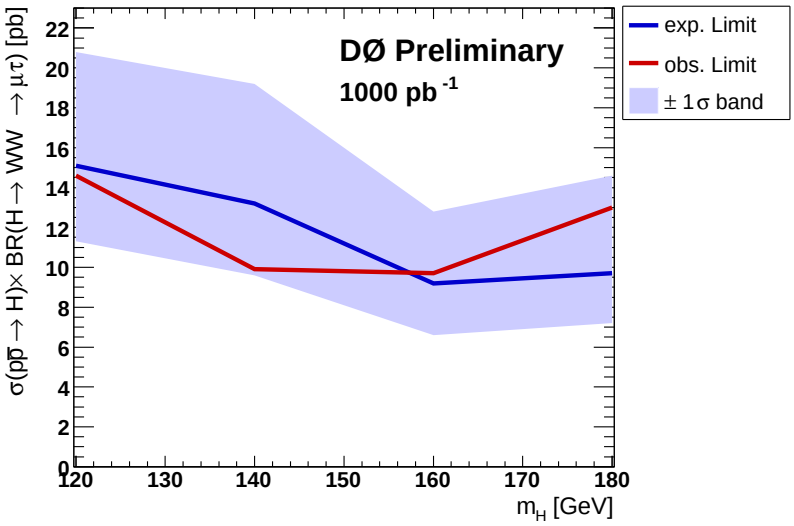


Figure 8.24: Expected and observed limits at 95% C.L. for $m_{\mathrm{H}} = 120, 140, 160$ and 180 GeV. The red graph represent the observed limit, the blue graph the expected limit. The uncertainty for the expected limits corresponds to $\pm 1\sigma$ for a C.L. of 95%.

9
Future Prospects

Along with DØ the CDF [110] experiment, another high p_T detector at the TEVATRON, performs searches for the Standard Model Higgs boson. The corresponding $H \rightarrow WW^*$ searches at CDF result in very comparable sensitivities as the DØ analyses as presented in Sec. 7.5. The expected limits of the CDF experiment since 2004 are shown in Fig. 9.1.

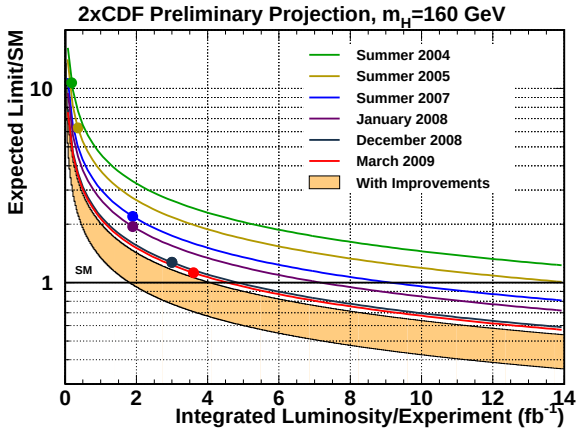


Figure 9.1: Achieved and projected expected upper limits on the Standard Model Higgs boson cross section from March 2009. The solid lines are $1/\sqrt{\int \mathscr{L} dt}$ projections. The top of the orange band corresponds to the Summer 2007 performance expected limit scaled such that it corresponds to twice the integrated luminosity. The bottom of the orange band to theoretical achievable limits incorporating anticipated improvements by the CDF collaboration. This plots is shown for $m_H = 160$ GeV [111].

The Higgs boson results of both detectors are combined using various mutually exclusive analysis channels, yielding the first Standard Model Higgs exclusion in a mass range from $m_H = 160 - 170$ GeV. The exclusion, in combination with other updated results of Spring 2009 from the TEVATRON like the W and top mass measurements entered the latest update of the indirect constraints on the Higgs mass as seen in Fig. 9.2. This leads to an upper limit from the electroweak precision measurements 163 GeV [20] for the Higgs boson mass at 95% C.L..

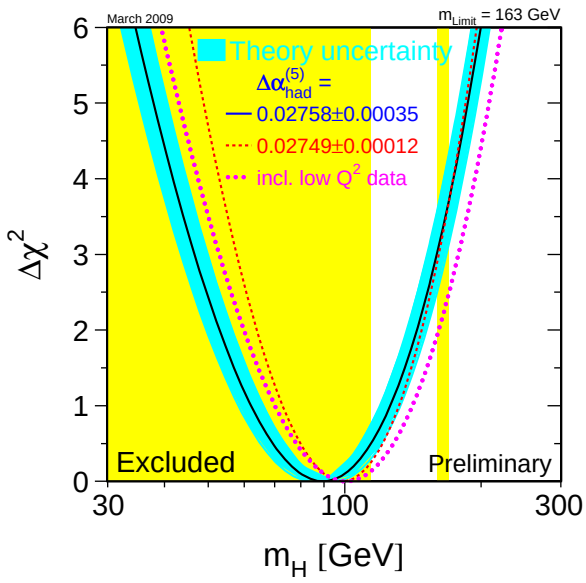


Figure 9.2: Constraint on the Higgs boson mass as $\Delta\chi^2 = \chi^2 - \chi^2_{min}$ vs m_H curve. The line is the results of the fit using high-Q^2 data whereas the band represents an estimate of the theoretical error due to missing higher order corrections. The vertical yellow bands show the direct exclusion limits by LEP2 and Tevatron. The dotted curves shows the corresponding fits using low-Q^2 data and the dashed respectively solid curves various evaluations of the hadronic correction from light quark contributions to $\alpha(m_Z^2)$ [19].

9.1 Future Prospects for Higgs Searches at the Tevatron

Both experiments, CDF and DØ, have analyzed different data sets and various analyses are based on different data samples. Therefore the effective analyzed luminosity in the TEVATRON combination corresponds to 2.6 fb^{-1} of data for a Higgs boson mass of $m_H = 115$ GeV and 3.8 fb^{-1} for a Higgs boson mass of $m_H = 165$ GeV. At both experiments the recorded data corresponds to about $\int \mathscr{L} dt$=6 fb^{-1} at the time of writing this thesis (July 2009). Considering the performance of the TEVATRON and the data taking efficiency an integrated luminosity of 10 fb^{-1} of Run II seems to be well in the reach of DØ and CDF until end of 2011. Considering possible improvements in mass resolution, b-tagging, additional acceptance and new analysis channels, sensitivity in a mass range of $m_H = 145 - 190$ GeV may be reached. In the low mass region sensitivity for $m_H = 115$ GeV seems feasible but requires continuous improvements of the analysis methods and high data taking efficiency.
However, the luminosities achievable even in the most optimistic scenarios would at the most be sufficient to claim evidence, no observation. Fig. 9.3 shows projections for the integrated luminosity necessary for a 95% C.L. exclusion or a 3σ evidence respectively. Hence an observation of

the Standard Model Higgs boson is expected to be made in the near future at the Large Hadron Collided (LHC) at CERN.

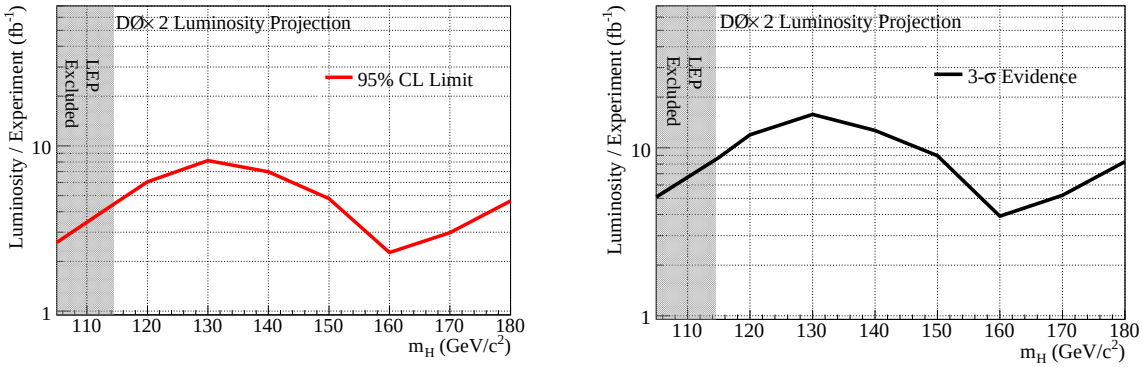


Figure 9.3: Projections of limits on the Standard Model Higgs boson at the TEVATRON, assuming that the doubled DØ sensitivity corresponds to the sensitivity of the TEVATRON combination. The left plots shows the projection for the luminosity needed to achieve a 95%C.L. exclusion. The right plots gives a projection for the integrated luminosity necessary for a 3 σ evidence. [112]

9.2 Future Prospects for Higgs Searches at the LHC

The LHC collider at CERN, Geneva will provide a center of mass energy of $\sqrt{s} = 14$ TeV, offering an unprecedented physics reach. One of the design goals for both detectors was to achieve sensitivity for the Higgs boson across the entire Higgs boson mass range. Figure 9.4 shows the discovery significance for 10 fb^{-1} of data at the ATLAS detector respectively the integrated luminosity needed for a 5 σ discovery at CMS [113, 114].

One sees that the sensitivity is excellent over a wide mass range. The LHC experiments will be able to confirm the presented exclusion with as little as 200 $\mathrm{pb}^{-1}$of data.

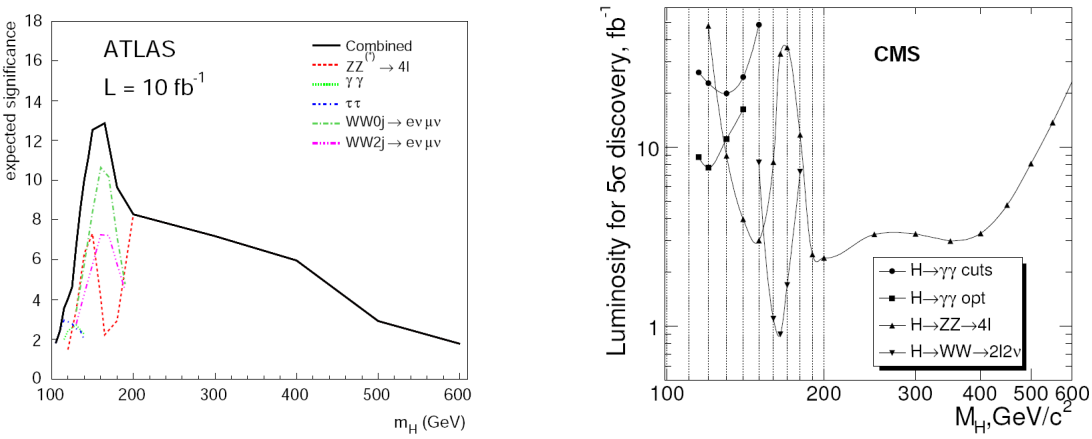


Figure 9.4: left: The median discovery significance of the ATLAS detector for the various channels and the combination with an integrated luminosity of 10 fb^{-1} for for masses up to 600 GeV [113].
right: The integrated luminosity needed for the 5 σ discovery of the inclusive Higgs boson production $pp \to H + X$ with the Higgs boson decay modes $H \to \gamma\gamma$, $H \to ZZ \to 4\ell$ and $H \to WW \to 2\ell 2\nu$ in the CMS experiment[114] .

10
Summary

The presented thesis uses data recorded between April 2002 and December 2008 with the DØ detector at the TEVATRON collider. The proton-antiproton center of mass energy is $\sqrt{s} = 1.96$ TeV and the data corresponds to an integrated luminosity of $\int \mathscr{L} dt = 1.0\text{-}4.2$ fb^{-1}. Several searches for a high mass Higgs boson decaying into final states with two leptons and high transverse momentum have been conducted, asking for either two electrons or a muon and a tau lepton in the final state. The main background sources for all analyses are $Z \to \ell\ell$, $W +$ jets , di-bosons and $t\bar{t}$ production. By applying selection requirements the major part of the background is removed, leaving in each case $W +$ jets and the irreducible di-boson production as main background processes. These remaining background processes are separated by applying multivariate methods, a neural network in case of the $e^{\pm}e^{\mp}$ analysis and a likelihood in the $\mu + \tau_{had} + X$ analysis. No excess in data above the expectation from the Standard Model processes is observed in any of the analyses. Considering various sources of systematic uncertainties upper limits on the Higgs production cross section $\sigma \times BR(H \to WW)$ at 95% confidence level are set. The result obtained for the $e^{\pm}e^{\mp}$ channel is $\sigma(m_H = 165 \text{ GeV}) \times BR(H \to WW^*)/\sigma_{SM} < 2.2$. The limit set in the $\mu + \tau_{had} + X$ final state was measured to be $\sigma(m_H = 160 \text{ GeV}) \times BR(H \to WW^*) < 9.7$ pb^{-1} for the $\mu + \tau_{had} + X$ channel.

The result is combined with other Higgs boson searches at DØ. The improvement achieved over the years led to an almost linear development of the sensitivity, considerably better than the sensitivity improvements expected by increasing the integrated luminosity which would scale like $1/\sqrt{\int \mathscr{L} dt}$. The combination of the various $H \to WW^*$ analyses at DØ is shown in Fig. 7.41, the observed upper limit is $\sigma \times BR(H \to WW^*)/\sigma_{SM} = 1.3$, the expected upper limit is $\sigma \times BR(H \to WW^*)/\sigma_{SM} = 1.7$.

Ultimately the Higgs boson results of both detectors are combined using a range of various mutually exclusive analysis channels. During the combination sources of systematic uncertainties are considered and correlated across the experiments [15]. The combination is dominated by the $H \to WW^* \to \mu^+\nu\mu^-\nu$, $H \to WW^* \to e^{\pm}\nu\mu^{\mp}\nu$ and $H \to WW^* \to e^+\nu e^-\nu$ searches in the high Standard Model Higgs boson mass region. In the low mass region search channels like $W/ZH \to \ell\ell b\bar{b}$ dominate. This combination leads to the exclusion of the Standard Model Higgs boson in a mass range from $m_H = 160 - 170$ GeV. It is the first experimental exclusion of a Standard Model Higgs boson mass range since the direct searches at LEP, discussed in Sec. 2.7.1. The research conducted for this thesis plays a crucial role in the exclusion. At a Higgs boson mass of $m_H = 165$ GeV the limit yielded is:

observed limit: $\sigma \times BR(H \to WW)/\sigma_{SM} = 0.8$
expected limit: $\sigma \times BR(H \to WW)/\sigma_{SM} = 1.2$

For a low mass Higgs boson of $m_H = 115$ GeV the combination leads to observed and expected limits of $\sigma \times BR(H \to WW)/\sigma_{SM}$ =2.6, respectively $\sigma \times BR(H \to WW)/\sigma_{SM}$ =2.4. Figure 10.1 shows the expected and observed limits at 95% C.L. for the combination of various Higgs boson searches relative to the Standard Model based on 0.9-4.2 fb^{-1} of Run II data. The green and yellow shaded areas represent the $\pm 1\sigma$ and $\pm 2\sigma$ confidence levels for the background only hypothesis and the solid black graph represents the observed limit. The expected limit is given by the dashed black graph.

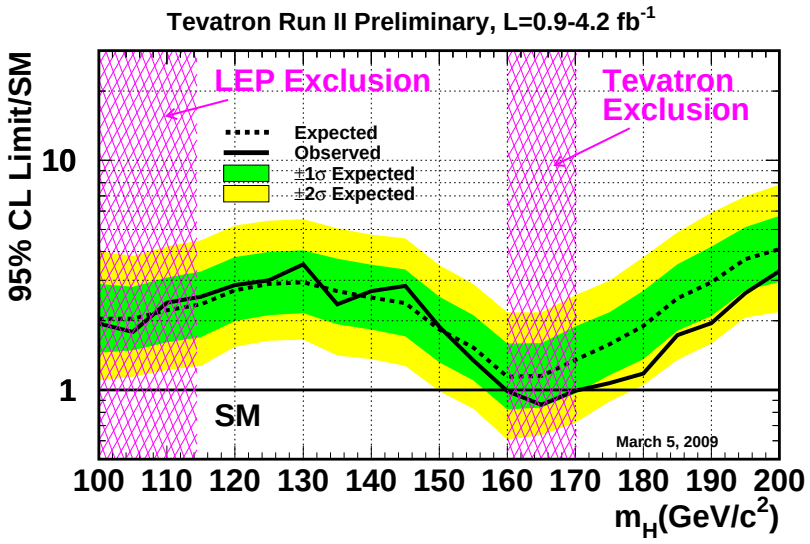


Figure 10.1: Observed and expected (median, for the background-only hypothesis) 95% C.L. upper limits on the ratios to the Standard Model Higgs boson cross section, as functions of the Higgs boson mass for the analyses. The limits are expressed as a multiple of the Standard Model prediction. The green and yellow shaded areas represent the $\pm 1\sigma$ and $\pm 2\sigma$ confidence levels.[15]

Although searched for more than twenty years the Higgs boson still evades direct observation. An important milestone has been reached within the scope of this thesis, achieving sensitivity for the Standard Model Higgs boson for the first time at hadron colliders. If it exists, evidence for the Standard Model Higgs boson could be found at the Tevatron in the near future and a discovery is very likely at the LHC. A discovery of the Higgs boson would guide and improve our understanding of the physics of the smallest particles along with other upcoming observations.

Bibliography

[1] DEPHI L3 ALEPH, OPAL Collaborations. *Search for the Standard Model Higgs Boson at LEP*. hep-ex/0306033.

[2] Fragment 117, Diogenes Laertius IX, 72.

[3] A. D. Dolgov. *Neutrinos in Cosmology*. arXiv:hep-ph/0202122.

[4] C. Walter. *Experimental Neutrino Physics*. 2009. arXiv:hep-ex/0810.3937.

[5] Lewin, Smith. *Astropart. Phys.*, 6:87–112, 1996.

[6] S. Weinberg. *Rev Mod Phys*, 61:1, 1989.

[7] A. Albrecht et al. *Report of the Dark Energy Task Force*. arXiv:astro-ph/0609591.

[8] Particle Data Group. *The Review of Particle Physics*. `http://pdg.lbl.gov/`.

[9] A. Quadt. *Top quark physics at Hadron Colliders*. 2006. Eur. Phys.J. **C48**, 835-1000.

[10] F. Halzen, A. D. Martin. *Quarks and Leptons*. Wiley, 1984.

[11] S. Dawson. *Introduction to Electroweak Symmetry Breaking*. *AIP Conf.Proc.*, pages 1116:11–3, 2009. arXiv:hep-ph/0812.2190v1.

[12] K. Riesselmann T. Hambye. *Matching conditions and Higgs mass upper bounds revisited*. *Phys.Rev. D*, 55:7255–7262, 1997. arXiv:hep-ph/9610272v1.

[13] Marc Sher. *Phys. Lett. B*, 331:448, 1994.

[14] C. Tully M. Kado. *The Searches for Higgs Bosons at LEP*. Annu. Rev. Nucl. Part. Sci. 2002. 52:65-113.

[15] Higgs working group Tevatron New Phenomena. *Combined CDF and DZero Upper Limits on Standard Model Higgs-Boson Production with up to 4.2 fb^{-1} of Data*. arXiv:hep-ex/0903.4001.

[16] The LEP working group for Higgs Boson searches. *Search for the Standard Model Higgs Boson at LEP*. 2003. CERN-EP/2003-011.

[17] R. Boughezal. *Three loop corrections to electroweak precision observables in the large Higgs mass limit*. 2003. `http://www.freidok.uni-freiburg.de/volltexte/2004/`.

[18] *The LEP Electroweak Working Group*. 2009. `http://lepewwg.web.cern.ch/LEPEWWG/plots/winter2009/`.

[19] The LEP working group for Higgs Boson searches. *Precision Electroweak Measurements and Constraints on the Standard Model*. 2008. CERN-PH-EP/2008-020.

[20] *The LEP Electroweak Working Group*. `http://lepewwg.web.cern.ch/LEPEWWG/`.

[21] Tev4LHC Higgs working group. *Standard Model Higgs cross sections at hadron colliders*. `http://maltoni.web.cern.ch/maltoni/TeV4LHC/SM.html`.

[22] J. Kalinowski A. Djouadi, M. Spira. *HDECAY: a Program for Higgs Boson Decays in the Standard Model and its Supersymmetric Extension. Comput. Phys. Commun.*, 108:56–74, 1998. arXiv:hep-ph/9704448.

[23] M. Carena et al. *Report of the Tevatron Higgs working group*. hep-ph/0010338.

[24] DØ Collaboration. `http://www-d0.fnal.gov/Run2Physics/displays/presentations/`.

[25] M. Weber. `http://d0server1.fnal.gov/projects/Silicon/www/SMT_files/talks/`.

[26] V. M. Abazov et al. *Nucl. Instrum. Meth. A*, 565:463, 2006.

[27] D. Chaping et al. *Measurement of $Z \to ee$ and $W \to e\nu$ Production Cross Sections using one tight central Electron*. 2005. DØ Note 4897.

[28] C. Royon M. Voutilainen. *Jet p_T resolution using* `JetCorr`. 2007. DØ Note 5381.

[29] DØ Collaboration. `http://www-d0.fnal.gov/phys_id/luminosity/`.

[30] T. Andeen et. al. *The D Experiments Integrated Luminosity for Tevatron Run IIa. 2007a*. 2007. FERMILAB-TM-2365.

[31] T. Edwards et. al. *Determination of the effective inelastic p anti-p cross-section for the DØ*. 2004. FERMILAB-TM-2278-E.

[32] T. M. Liss. . 2003. FERMILAB-FN-0741. S. Klimenko, J. Konigsberg. *Averaging of the inelastic cross sections measured by the CDF and the E811 experiments*. 2003. FERMILAB-FN-0741.

[33] Algorithms Group. *How to Run RECO*. `http://www-d0.fnal.gov/computing/algorithms/howto/howtoreco.html`.

[34] Algorithms Group. *DØ Run II Event Data Model*. `http://www-d0.fnal.gov/software/data_model/data_model.html`.

[35] R. van Kooten et al. *Tracking Algorithm Recommendation Committee (TARC) Report II*. 2003. http://d0server1.fnal.gov/www/Computing/Web/Meeting/Reviews/TARC/TARC2Recom.pdf.

[36] A. Khanov. *HTF: histogramming method for finding tracks. The algorithm description.* 2000. DØ note 3778.

[37] G. Borissov. *Ordering a Chaos or... Technical Details of AA Tracking*. 2003. `http://wwwd0.fnal.gov/atwork/adm/d0_private/2003-02-28/adm_talk.ps`.

[38] C. Tully A. Schwartzman. *Primary Vertex Reconstruction by Means of Adaptive Vertex Fitting*. 2005. DØ Note 4918.

[39] F. Fleuret. *The DØ Electron/Photon Analysis Package EMAnlyze*. 2001. DØ Note 3888.

[40] S. Protopopescu F. Fleuret, F. Beaudette. *The DØ Soft Electron Reconstruction Package SEMReco*. 2001. DØ Note 3872.

[41] J.-F. Grivaz F. Beaudette. *The Road Method*. 2002. DØ Note 3976.

[42] O. Atramentov et al. *Electron and Photon Identification with p20 data*. DØ Note 5761 (v2).

[43] L. Wand. *Electron Likelihood Efficiency in p17*. 2006. DØ Note 5114.

[44] L. Wang et al. *Electron Likelihood in p17*. DØ Note 5114.

[45] *New vs Old Electron ID*. `https://plone4.fnal.gov/P1/D0Wiki/object-id/emid/NewVsOld/`.

[46] P. Calfayan et al. *Muon Identification Certification for p17 data*. 2008. DØ Note 5157.

[47] Philippe Calfayan et al. *Muon Identification Certification for p17 data*. 2006.

[48] D. Chakraborty et al. *Reconstruction of tau leptons in hadronic final states at D0 in Run 2*. 2003. DØ Note 4210.

[49] P. Svoisky S. Protopopescu. *Tau identification with neural networks for p17 data*. 2006. DØ Note 5094.

[50] I. Torchiani. *Search for Higgs Bosons and Supersymmetric Particles in Tau Final States*. 2008. http://www-d0.fnal.gov/results/publications_talks/thesis/torchiani/thesis.pdf.

[51] B. R. Webber S. Catani, Y. L. Dokshitzer. *The $k_\perp$-Clustering Algorithm for Jets in Deep Inelastic Scattering and Hadron Collisions*. 1992. Phys.Lett. B285 291-299.

[52] J. R. Dittmann et al. G. C. Blazey. *Run II Jet Physics*. 2000. arXicv:hep-ex/0005012.

[53] B. Andrieu E. Busato. *JetAlgorithms in the DØ Run II Software: Description and Uses's Guide*. 2004. DØ Note 4457.

[54] G. Bernardi, E. Busato, J.-R. Vlimant. *Improvements from the T42 Algorithm on Calorimeter Object Reconstruction*. 2004. DØ note 4335.

[55] A. Harel. *Jet ID Optimization*. 2005. DØ Note 4919.

[56] Jet Energy Scale Working Group. *Jet Energy Scale Determination at DØ Run II (final p17 version)*. 2007. DØ Note 5382.

[57] DØ Preliminary Jet Energy Scale. `http://www-d0.fnal.gov/phys_id/jes/public/plots_v7.1/`.

[58] M. Voutelainen, C. Royon. *Jet p_T resolution using JetCorr v7.1*. 2006. DØ Note 5381.

[59] N. Makovec, J.-F. Grivaz. *Shifting, Smearing and Removing Simulated Jets*. 2005. DØ Note 4914.

[60] C. Orchando, J.-F. Grivaz. *SSR for p17*. 2008. DØ Note 5609.

[61] DQ Group. *DØ Data Quality Co-ordination*. `http://www-d0.fnal.gov/computing/data_quality/`.

[62] Sophie Trincaz-Duvoid Gregorio Bernardi. *Improvement of the NADA Algorithm: Hot Cell Killing in DØ Run II Data*. 2002. DØ Note 4057.

[63] CDF Collaboration. *First Observation of Electroweak Single Top Quark Production*. 2009. arxiv:hep-ph/0903.0885.

[64] DØ Collaboration. *Observation of Single Top Quark Production*. 2009. arxiv:hep-ph/0903.0850.

[65] T. Sjostrand et al. *High-Energy-Physics Event Generation with PYTHIA 6.1. Comp. Phys. Commun.*, 135:238–259, 2001. hep-ph/0010017.

[66] F. Krauss J.-C. Winter. *Initial-state showering based on colour dipoles connected to incoming parton lines*. *JHEP ,2008*, page 0807:040, 2008. arXiv:hep-ph/0712.3913v1.

[67] B. Andersson. *The Lund Model*. Cambridge University Press, 1998.

[68] G. Corcella, I. G. Knowles, G. Marchesini, S. Moretti, K. Odagiri, P. Richardson, M. H. Seymor, B. R. Webber. *HERWIG 6.5*.

[69] T. Sjostrand. *Comp. Phys. Commun.*, 82:74, 1994. `CERN-TH 7112/93` (1993).

[70] J. Pumplin et al. *JHEP 0207*, 012, 2002.

[71] M.L. Mangano et al. *JHEP 07*, 001, 2003.

[72] D. de Florian, M. Grazzini. *Higgs production through gluon fusion: updated cross sections at the Tevatron and the LHC*. 2009. arXiv:hep-ph/0901.2427v2.

[73] R. S. Thorne A. D. Martin, W. J. Stirling, G. Watt. *Parton distributions for the LHC*. 2009. arXiv:hep-ph/0901.0002.

[74] M. Spira. arxiv:hep-ph/9510347.

[75] DØ Higgs Physics Group. *Compilation of Higgs cross sections and branching provided by the DØ Higgs group*. `http://www-d0.fnal.gov/Run2Physics/higgs/higgs_xs_br.html`.

[76] T. Nunnemann. *MCFM Cross Sections*. `http://www-clued0.fnal.gov/ nunne/cross-sections/mcfm_cross-sections.html`.

[77] S. Moch, P. Uwer. *Theoretical status and prospectes for top-quark pair production at hadron colliders*. 2008. hep-ph/0804.1476.

[78] W.L. van Neerven R. Hamberg, T. Matsuura. *Nucl. Phys.*, B359:343, 2002. Erratum-ibid. **B644**, 403 (2002).

[79] T. Nunnemann. DØ Note 4476.

[80] Y. Peters et al. *Study of the W+jets heavy flavor scale factor in p17*. 2007. DØ Note 5406.

[81] M. Hohlfeld. *Measurement of Single Electron Trigger Efficiencies for RunIIB*. 2008. DØ Note 5783.

[82] *DØ Trigger Group*. http://www-d0.fnal.gov/computing/algorithms/tsg/tsg.html.

[83] Markus Warsinsky. *Studies of b-associated production and muonic decays of neutral Higgs bosons at the ATLAS experiment within the Minimal Supersymmetric Standard Model*.

[84] M. Hohlfeld J. Elmsheuser. *Phys. Rev. Lett.*, 96:011801, 2006. DØ Note 4759.

[85] G. Hesketh. *W/Z p_T reweighting for* ALPGEN *and* PYHTIA. 2008. DØ Note 5786.

[86] *Single Electron Efficiencies in p17 Data and Monte-Carlo Using p18.05.00 d0correct*. 2006. DØ Note 5105.

[87] T. Bolton M. Shamim. *Generator Level Reweighting of pT of Z Boson*. DØ Note 5565.

[88] H. Schellmann. *The longitudinal shape of the luminous region at D0*. DØ Note 5142.

[89] P. Svoisky S. Protopopescu. *τ identification with neural networks for $p17$ data*. DØ Note 5094.

[90] ROOT. `http://root.cern.ch`.

[91] Jeroen Hegeman. *Measurement of the Top quark pair production cross section in proton-antiproton collisions at $\sqrt{s}$ =1.96 TeV, hadronic Top decays with the DØ detector*. 2008. `http://lss.fnal.gov/archive/thesis/fermilab-thesis-2009-07.pdf`.

[92] S. Catani et al. *JHEP*, 028:0307, 2003.

[93] K.A. Assamagan et al. *The Higgs Working Group: Summary Report*. 2003. arxiv:hep-ph/0406152.

[94] F. Petriello C. Anastasiou, R. Boughezal. *Mixed QCD-electroweak corrections to Higgs boson production in gluon fusion*. arXiv:0811.3458v2 [hep-ph].

[95] U. Aglietti et al. Two-loop electroweak corrections to Higgs production in proton-proton collisions. arXiv:hep-ph/0610033.

[96] Luminosity Working Group. *Adjustments to the Measured Integrated Luminosity in Run IIa*. 2007. DØ note 5139.

[97] T. Gleisberg et al. *JHEP 07*, 001, 2003.

[98] B. R. Webber S. Frixione. *The MC@NLO 3.3 Event Generator*. 2006. arXiv:hep-ph/0612272v1.

[99] T. Gleisberg et al. *Phys. Rev.*, 072:034028, 2005.

[100] T. Binoth *et al. Gluon induced WW background to Higgs boson searches at the LHC*. arXiv:hep-ph/0503094.

[101] N. Kauer. *GG2WW parton-level integrator and event generator*. `http://hepsource.sourceforge.net/GG2WW/`.

[102] K. S. Cranmer. *Comput. Phys. Commun., 136, 198*, 2001.

[103] W. Fisher. *Tests of the P17-P20 Combined Higgs Limits for the 2007 Lepton-Photon Conference*. DØ Note 5491.

[104] T. Junk. *Nucl. Instr. and Meth*, A434:435, 1999.

[105] W. Fisher. *Systematics and Limit Calculation*. 2006. DØ note 5309.

[106] M. Owen et al. *Search for MSSM Higgs Boson Production in the Decay $h \rightarrow \tau_\mu \tau$ with the DØ Detector at $\sqrt{s} = 1.96$ TeV*. 2007. DØ note 5331.

[107] P.Calfayan. *ORing single muon triggers in p17 data*. 2007. DØ note 5329.

[108] P. Calfayan et al. *Muon Identification Certification for p17 data*. 2007. DØ note 5157.

[109] B. Penning et al. *Search for the Higgs boson in $H \rightarrow WW^* \rightarrow \mu, \tau_{had}$ decays with 1150 pb^{-1} at DO in Run II*. 2007. DØ note 5332.

[110] CDF Collaboration. *The Collider Detector at Fermilab*. `http://www-cdf.fnal.gov/`.

[111] CDF Collaboration. *Combined Upper Limit on Standard Model Higgs Boson Production at CDF for Winter 2009*. `http://www-cdf.fnal.gov/physics/new/hdg/results/combcdf_mar09/`.

[112] DØ Collaboration. *Results and Publications*. `http://www-d0.fnal.gov/Run2Physics/WWW/results.htm`.

[113] Atlas Collaboration. *Expected Performance of the ATLAS Experiment, Detector, Trigger and Physics*. 2008. CERN-OPEN-2008-020.

[114] CDF Collaboration. *Technical Design Report, Volume II: Physics Performance*. 2006. CERN/LHCC 2006-021.

So if you're lonely
You know I'm here waiting for you.
Franz Ferdinand

11
Acknowledgments - Danksagung

Time to come to the probably most-read part of this thesis. Without the help and support of many wonderful people during the last years this thesis and and a lot of other things in my life would be fundamentally different. First of all I want to thank Prof. Dr. Karl Jakobs and Dr. Ralf Bernhard for their assistance, endurance and support during the years of this thesis.
Spending so many years at DØ itself I owe a lot to many people in both, science and social life (often deeply connected). I want to express my gratitude particularly to Yvonne, Schwani, Krisztian, Maiko, Alexander, Greg, Emanuel, Rainer, Dale, Ruchika, Marc, Su-Jung and Olav (thanks for the Brunch) for for their continuous friendship, being great flat-mates (quite some of them have been at some point) and many nice hours. I thank all this great scientist and good friends I was able to learn from at DØ: Stefan, Aurelio, Wade, Darien, Dmitry, George, Jan. I want to point out Marco, thank you for making me suffer... looking forward to another couple of years working together.
I thank Evelyn, Susanne and Jochen for the last minute proof-reading. Hope I can help out at times as well.
Often on another continent but never far away there are Hase, Hanna, Max and Lisa. The best friends one can ask for, offering me their unconditional friendship and hospitality since I know them. Thank you! Not to forget to mention Kerstin who helped to ease the burden at times and Jörg who was always helpful looking at things from a different point of view.
Of course all of this would not have been possible in the first place without my loving, caring and supportive parents Irma and Pieter. There are no words to describe their support of thirty years. Kai and Gilles, thank you too for being always there. Particularly Gilles! The responsibilities you will face soon such are so much more challenging than any PhD in physics.
Last but not least Marcelle. A truly impressive person and physicist: Marcelle, you are brilliant!

Zeit zu dem wohl am häufigsten gelesenen Teil dieser Arbeit zu kommen. Ohne die Hilfe und Unterstützung vieler toller Menschen in den letzen Jahren sähen sicher diese Arbeit aber auch viele andere Aspekte meines Leben gänzlich anders aus. Zuerst möchte ich Prof. Dr. Karl Jakobs und Dr. Ralf Bernhard für ihre Hilfe, Geduld und Unterstützung über die Jahre danken.
Nach so langer Zeit an DØ verdanke ich vieles sowohl in Wissenschaft als auch im Privatem guten Freunden. Für ihre Freunschaft, viele schöne Stunden und das gute Zusammenwohnen möchte ich meinen besonderen Dank an Yvonne, Schwani, Krisztian, Maiko, Alexander, Greg, Emanuel, Rainer, Dale, Ruchika, Marc, Su-Jung und Olav ausdrücken. Und natürlich all den guten Freunden und Wissenschaftlern, von denen ich an DØ lernen durfte: Stefan, Aurelio, Wade, Darien, Dmitry,

George, Jan. Besonders möchte ich Marco hervorheben und ihm danken mich leiden zu lassen. Ich freue mich auf die kommenden Jahre der Zusammenarbeit.
Vielen Dank an Evelyn, Susanne und Jochen für das Korrekturlesen in letzter Minute.
Oft auf einem anderen Kontinent aber trotzdem immer nah sind Hase, Hanna, Max an Lisa. Die allerbesten Freunde die man sich nur wünschen kann, ihre bedingungslose Freundschaft und Gastfreundschaft hat mir vieles erleichtert. Vielen Dank! Nicht zu vergessen Kerstin die mir einen Teil des Weges die Bürde ebenfalls erleichtert hat und Jörg der immer half die Perspektive zu wechseln. Natürlich wurde alles erst ermöglicht durch die liebevolle Unterstützung meiner Eltern Irma und Pieter und meiner Großmutter Mathilde. Mir fehlen die Worte zu beschreiben wie dankbar ich für alles der letzen dreissig Jahre bin. Kai und Gilles, vielen Dank, dass ihr immer dagewesen seid. Besonders Gilles! Die Verantwortung die du bald übernimmst stellt eine weit größere Herausforderung dar als jede Dissertation. Und schließlich noch Marcelle. Eine beeindruckender Mensch und tolle Physikerin. Marcelle, Du bist brilliant!

A
Trigger Requirements for Run II

A.1 Triggers for Run IIa

triggers for Run IIa
EM_MX_SH EM_MX_EMFR8 EM_HI_EMFR8 EM_HI_F0 EM_MX_F0 EM_HI_2EM5 EM_HI_2EM5_SH EM_HI_2EM5_EMFR8 EM_HI_2EM5_F0 E1_SHT20 E2_SHT20 E3_SHT20 E1_SH30 E2_SH30 E3_SH30 E1_L50 E1_VL70 E4_SH30 E4_SHT20 E1_NC90 E1_SHT22 E2_SHT22 E3_SHT22 E4_SHT22 E1_L70 E1_SHT25 E3_SHT25 E4_SHT25 E1_SH35 E3_SH35 E4_SH35 E1_ISH30 E3_ISH30 E4_ISH30 E1_ISHT22 E3_ISHT22 E4_ISHT22 EM_MX_TR EM_MX_SH_TR EM_HI_TR EM_HI_SH_TR EM_HI_2EM5_TR EM_HI_2EM5_SH_TR E1_T13L15 E1_T25VL30 E1_SHT15_TK13 E2_T13L15 E2_T25VL30 E2_SHT15_TK13 E3_T13L15 E3_T25VL30 E3_SHT15_TK13 E4_T13L15 E4_T25VL30 E4_SHT15_TK13 E5_SHT15_TK13 E6_T13L15 E6_SHT15_TK13 E5_SHT20 E5_SH30 E6_SHT20 E6_SH30 E7_SHT20 E7_SH30 E5_T13L15 E7_SHT15_TK13 E7_T13L15 E8_SHT20 E8_SH30 E9_SHT20 E9_SH30 E8_SHT15_TK13 E8_T13L15 E9_SHT15_TK13 E9_T13L15 E9_IT7SHT8 E9_SHT8_ITK10 E5_SHT22 E6_SHT22 E7_SHT22 E1_T15L20 E1_T13SH15 E2_T15L20 E2_T13SH15 E3_T15L20 E3_T13SH15 E4_T15L20 E4_T13SH15 E5_T15L20 E5_T13SH15 E6_T15L20 E6_T13SH15 E7_T15L20 E7_T13SH15 E9_T15L20 E8_T13SH15 E8_IT10SHT10 E8_SHT10_ITK10 E8_T15L20 E9_T13SH15 E9_IT10SHT10 E9_SHT10_ITK10 E1_T13SHT15 E1_T15SH20 E1_ISHT15_TK13 E3_T13SHT15 E3_T15SH20 E3_ISHT15_TK13 E4_T13SHT15 E4_T15SH20 E4_ISHT15_TK13 E13_SHT25 E13_SH35 E13_T13SHT15 E13_T15SH20 E13_ISHT15_TK13 E13_ISHT22 E13_ISH30 E17_ISHT22 E17_SHT25 E17_ISH30 E17_SH35 E18_ISHT22 E18_SHT25 E18_ISH30 E18_SH35 E19_ISHT22 E19_SHT25 E19_ISH30 E19_SH35 E20_ISHT22 E20_SHT25 E20_ISH30 E20_SH35 E21_ISHT22 E21_SHT25 E21_ISH30 E20_SH35 E17_T13SHT15 E17_T15SH20 E17_ISHT15_TK13 E17_IT10SHT10 E17_SHT12_ITK10 E18_T13SHT15 E18_T15SH20 E18_ISHT15_TK13 E19_T13SHT15 E19_T15SH20 E19_ISHT15_TK13 E20_T13SHT15 E20_T15SH20 E20_ISHT15_TK13 E20_IT10SHT10 E20_SHT12_ITK10 E21_T13SHT15 E21_T15SH20 E21_ISHT15_TK13 E21_IT10SHT10 E21_SHT12_ITK10

Table A.1: Trigger selection used for Run IIa

A.2 Triggers for Run IIb

triggers for Run IIb
E1_ISH30 E1_ISHT15_TK13 E1_ISHT22 E1_L70 E1_L80 E1_LH2ISH24 E1_LH2ISHT17T14 E1_LH2L70 E1_LH2SH27 E1_LH3ISH25 E1_LH3SH27 E1_SH35 E1_SH60 E1_SHT25 E1_SHT27 E1_SHT27_NOLUM E1_SHT50 E1_T13SHT15 E1_T14LH2SH17 E1_T15SH20 E2_ISH30 E2_ISHT15_TK13 E2_ISHT22 E2_L70 E2_L80 E2_LH2ISH24 E2_LH2ISHT17T14 E2_LH2L70 E2_LH2SH27 E2_LH3ISH25 E2_LH3SH27 E2_SH35 E2_SH60 E2_SHT25 E2_SHT27 E2_SHT50 E2_T13SHT15 E2_T14LH2SH17 E2_T15SH20 TE1_ISH30 TE1_ISHT15_TK13 TE1_ISHT22 TE1_L70 TE1_L80 TE1_L80_NOLUM TE1_LH2ISH24 TE1_LH2ISH24_NOLUM TE1_LH2ISHT17T14 TE1_LH2ISHT17T14_NOLUM TE1_LH2L70 TE1_LH2L70_NOLUM TE1_LH2SH27 TE1_LH2SH27_NOLUM TE1_SH35 TE1_SH60 TE1_SH60_NOLUM TE1_SHT25 TE1_SHT25_NOLUM TE1_SHT50 TE1_SHT50_NOLUM TE1_T13SHT15 TE1_T14LH2SH17 TE1_T14LH2SH17_NOLUM TE1_T15SH20 TE2_ISH30 TE2_ISHT15_TK13 TE2_ISHT22 TE2_L70 TE2_L80 TE2_LH2ISH24 TE2_LH2ISHT17T14 TE2_LH2L70 TE2_LH2SH27 TE2_SH35 TE2_SH60 TE2_SHT25 TE2_SHT50 TE2_T13SHT15 TE2_T14LH2SH17 TE2_T15SH20 TE3_ISH30 TE3_ISHT15_TK13 TE3_ISHT22 TE3_L70 TE3_L80 TE3_L80_NOLUM TE3_LH2ISH24 TE3_LH2ISH24_NOLUM TE3_LH2ISHT17T14 TE3_LH2ISHT17T14_NOLUM TE3_LH2L70 TE3_LH2L70_NOLUM TE3_LH2SH27 TE3_LH2SH27_NOLUM TE3_SH35 TE3_SH60 TE3_SHT25 TE3_SHT25_NOLUM TE3_SHT50 TE3_SHT50_NOLUM TE3_T13SHT15 TE3_T14LH2SH17 TE3_T14LH2SH17_NOLUM TE3_T15SH20 TE4_ISH30 TE4_ISHT15_TK13 TE4_ISHT22 TE4_L70 TE4_L80 TE4_L80_NOLUM TE4_LH2ISH24 TE4_LH2ISH24_NOLUM TE4_LH2ISHT17T14 TE4_LH2ISHT17T14_NOLUM TE4_LH2L70 TE4_LH2L70_NOLUM TE4_LH2SH27 TE4_LH2SH27_NOLUM TE4_SH35 TE4_SH60 TE4_SHT25 TE4_SHT25_NOLUM TE4_SHT50 TE4_SHT50_NOLUM TE4_T13SHT15 TE4_T14LH2SH17 TE4_T14LH2SH17_NOLUM TE4_T15SH20 TE5_ISH30 TE5_ISHT15_TK13 TE5_ISHT22 TE5_L70 TE5_L80 TE5_LH2ISH24 TE5_LH2ISHT17T14 TE5_LH2L70 TE5_LH2SH27 TE5_SH35 TE5_SH60 TE5_SHT25 TE5_SHT50 TE5_T13SHT15 TE5_T14LH2SH17 TE5_T15SH20

Table A.2: Trigger selection used for Run IIa

B
Appendix for the $H \rightarrow WW^* \rightarrow e^+\nu e^-\nu$ selection

B.1 Control Distributions for the $H \rightarrow WW^* \rightarrow e^+\nu e^-\nu$ selection

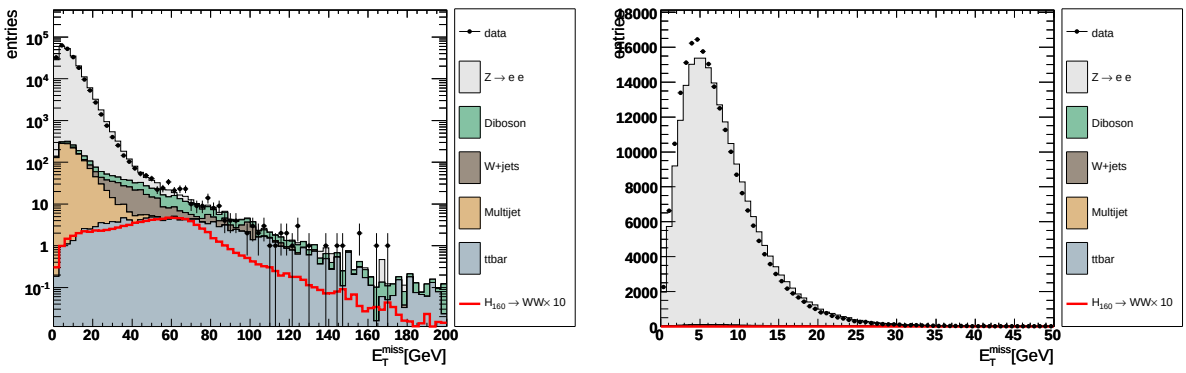


Figure B.1: Distribution of the missing transverse energy $\not\!\!E_{\mathrm{T}}$ in logarithmic scale (left) and linear scale (right) at preselection level for data (points) and sum of all backgrounds (filled histograms). The entire Run II data set is used. The expected signal, multiplied by a factor of 10, for a 160 GeV/c^2 Standard Model Higgs is also shown.

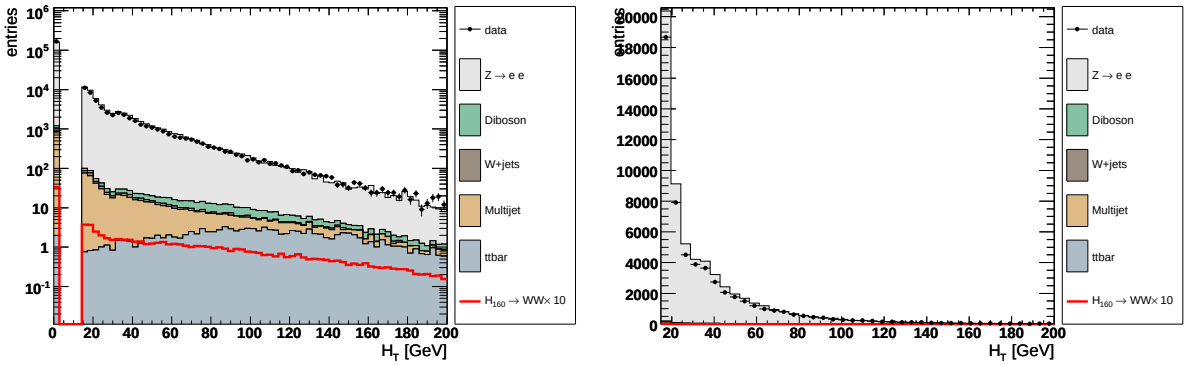


Figure B.2: Distribution of the sum of the transverse jet momenta H_T (right) in logarithmic scale (left) and linear scale (right) at preselection level for data (points) and sum of all backgrounds (filled histograms). The entire Run II data set is used. The expected signal, multiplied by a factor of 10, for a 160 GeV/c^2 Standard Model Higgs is also shown.

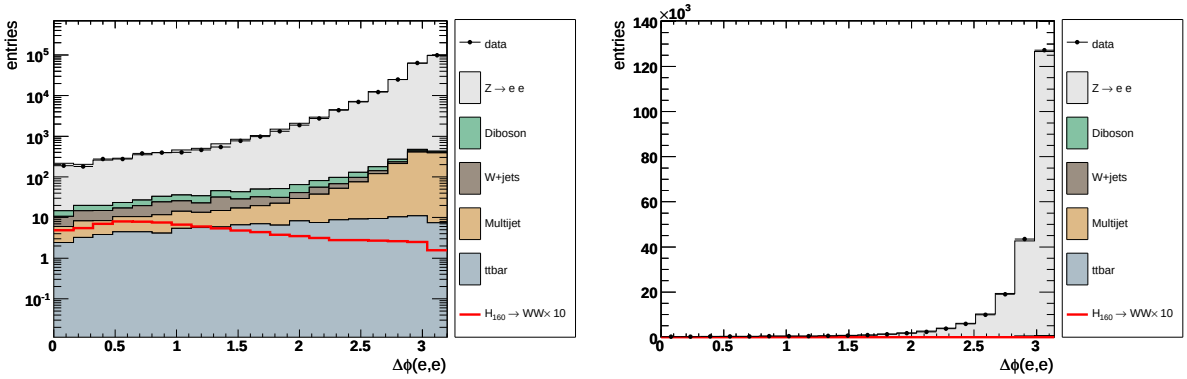


Figure B.3: Distribution of the lepton opening angle in logarithmic scale (left) and linear scale (right) at preselection level for data (points) and sum of all backgrounds (filled histograms). The entire Run II data set is used. The expected signal, multiplied by a factor of 10, for a 160 GeV/c^2 Standard Model Higgs is also shown.

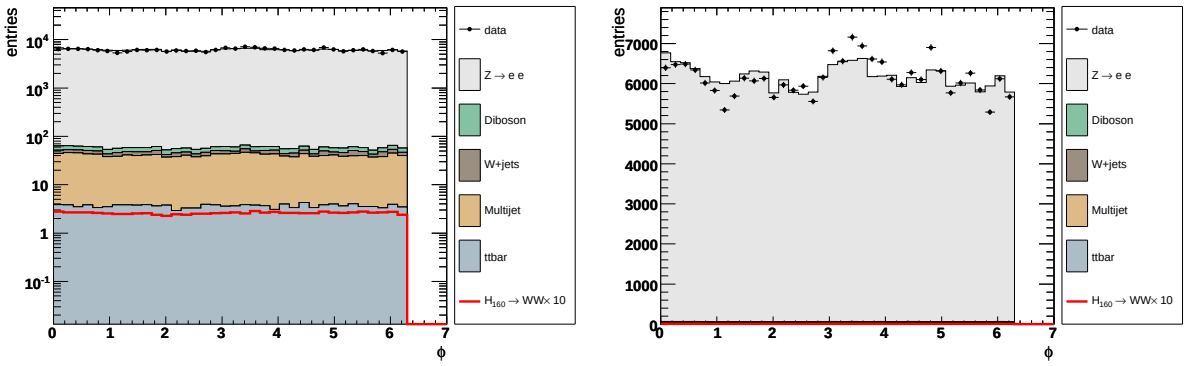


Figure B.4: ϕ distribution of the leading lepton at preselection level in logarithmic scale (left) and linear scale (right) at preselection level for data (points) and sum of all backgrounds (filled histograms). The entire Run II data set is used. The expected signal, multiplied by a factor of 10, for a 160 GeV/c^2 Standard Model Higgs is also shown.

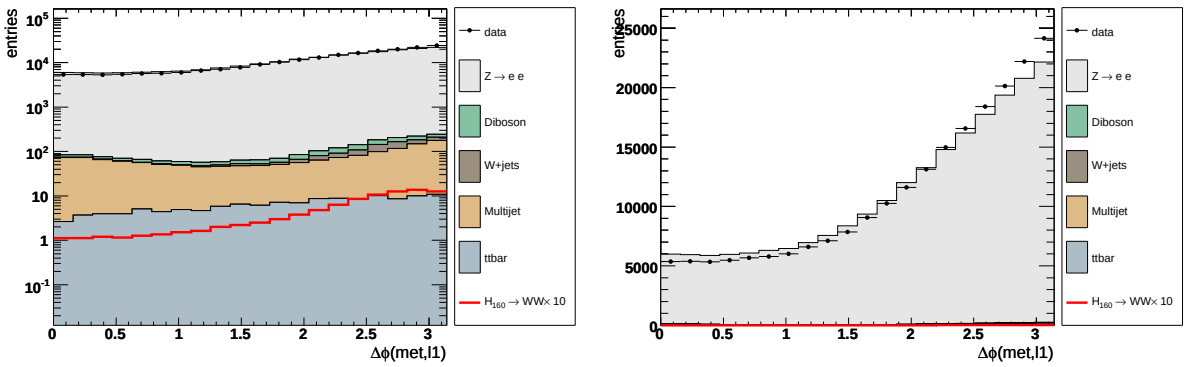


Figure B.5: Distribution of the opening angle between for the leading lepton and $\not\!\!E_T$ in logarithmic scale (left) and linear scale (right) at preselection level for data (points) and sum of all backgrounds (filled histograms). The entire Run II data set is used. The expected signal, multiplied by a factor of 10, for a 160 GeV/c^2 Standard Model Higgs is also shown.

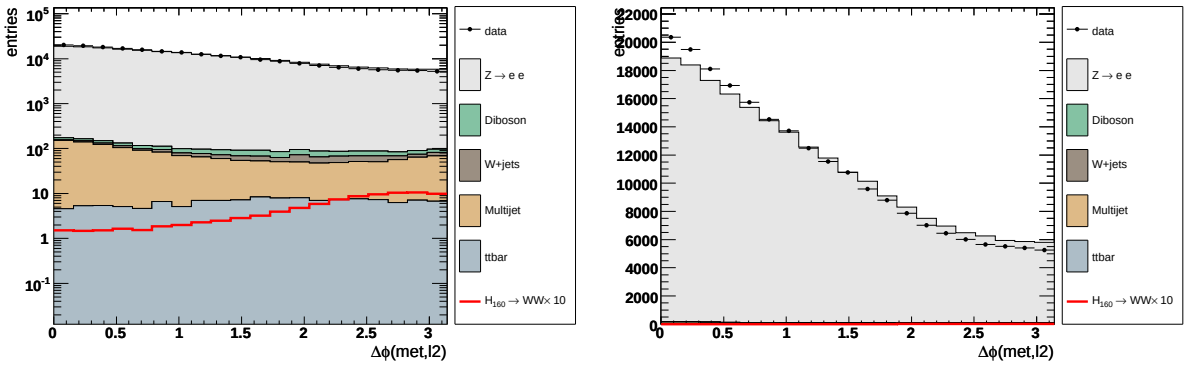


Figure B.6: Distribution of the opening angle between the trailing lepton and $\not\!\!E_{\mathrm{T}}$ in logarithmic scale (left) and linear scale (right) at preselection level for data (points) and sum of all backgrounds (filled histograms). The entire Run II data set is used. The expected signal, multiplied by a factor of 10, for a 160 GeV/c^2 Standard Model Higgs is also shown.

B.2 NN control distributions

Distributions for the NN output on preselection stage and after the $\Delta\phi(\ell,\ell)$ cut for m_H=120,160 and 200 GeV can be found in Sec. 7.2, all remaining Higgs boson mass dependent output distributions are to be found in this section:

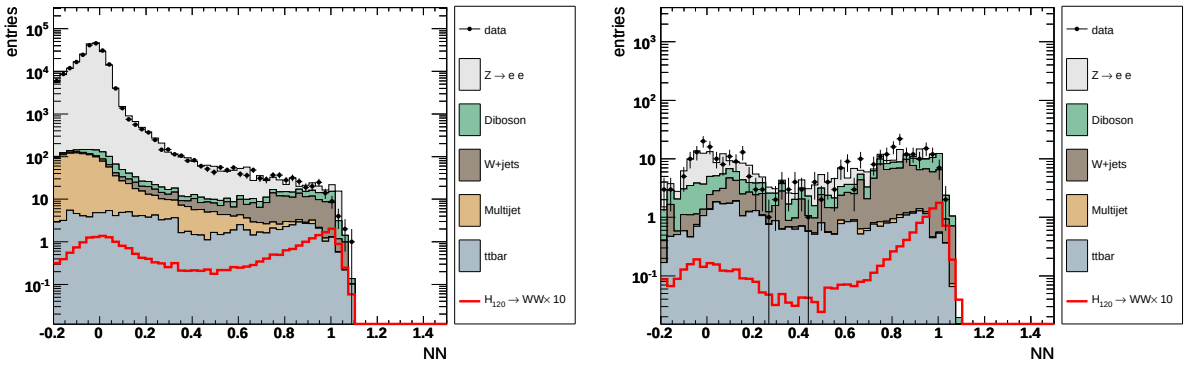


Figure B.7: NN output for $m_{\mathrm{H}} = 120$ GeV at preselection level (left) and (right) NN output for $m_{\mathrm{H}} = 120$ GeV after the $\Delta\phi(\ell,\ell)$ selection cut (right). The signal indicated by the red graph is representative of a Higgs boson mass of $m_{\mathrm{H}} = 160$ GeV. The entire Run II data set is used.

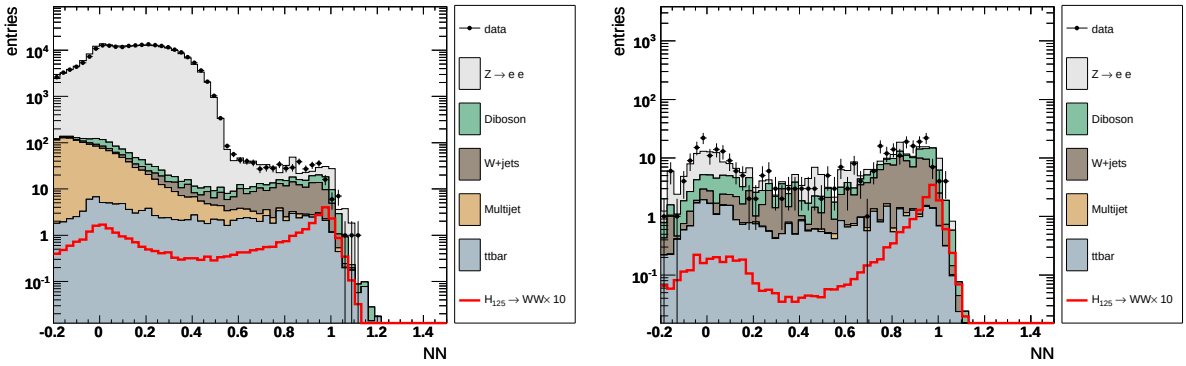


Figure B.8: NN output for $m_{\mathrm{H}} = 125$ GeV at preselection level (left) and (right) NN output for $m_{\mathrm{H}} = 125$ GeV after the $\Delta\phi(\ell,\ell)$ selection cut (right). The signal indicated by the red graph is representative of a Higgs boson mass of $m_{\mathrm{H}} = 160$ GeV. The entire Run II data set is used.

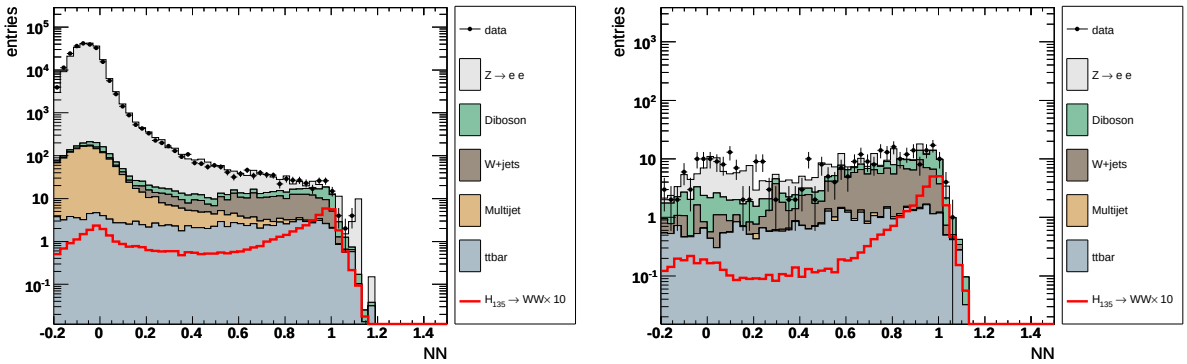


Figure B.9: NN output for $m_{\mathrm{H}} = 135$ GeV at preselection level (left) and (right) NN output for $m_{\mathrm{H}} = 135$ GeV after the $\Delta\phi(\ell,\ell)$ selection cut (right). The signal indicated by the red graph is representative of a Higgs boson mass of $m_{\mathrm{H}} = 160$ GeV. The entire Run II data set is used.

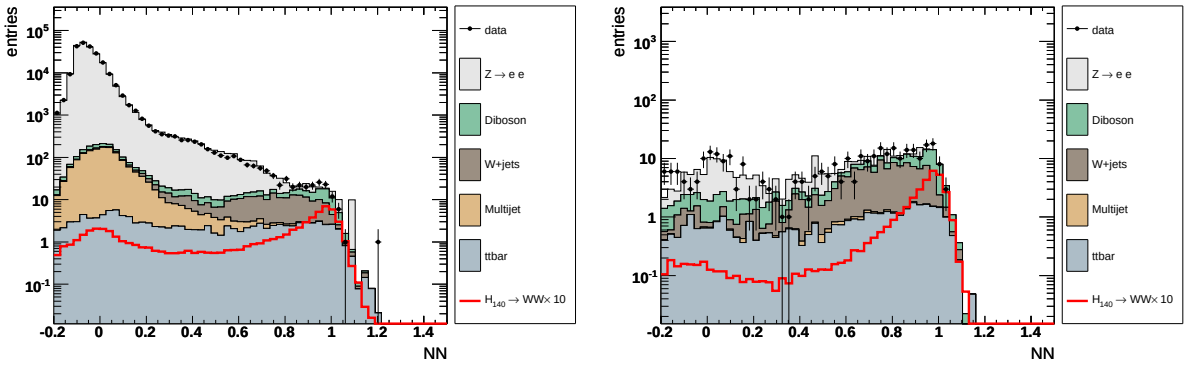


Figure B.10: NN output for $m_{\mathrm{H}} = 140$ GeV at preselection level (left) and (right) NN output for $m_{\mathrm{H}} = 140$ GeV after the $\Delta\phi(\ell,\ell)$ selection cut (right). The signal indicated by the red graph is representative of a Higgs boson mass of $m_{\mathrm{H}} = 160$ GeV. The entire Run II data set is used.

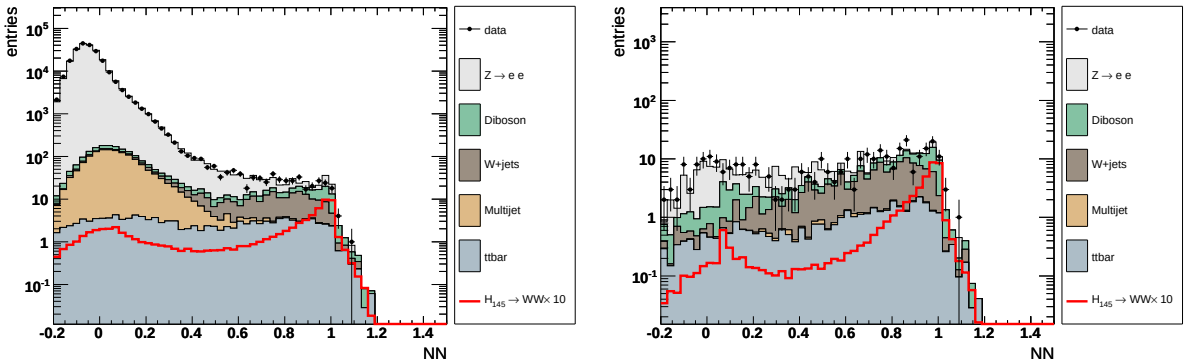


Figure B.11: NN output for $m_{\mathrm{H}} = 145$ GeV at preselection level (left) and (right) NN output for $m_{\mathrm{H}} = 145$ GeV after the $\Delta\phi(\ell,\ell)$ selection cut (right). The signal indicated by the red graph is representative of a Higgs boson mass of $m_{\mathrm{H}} = 160$ GeV. The entire Run II data set is used.

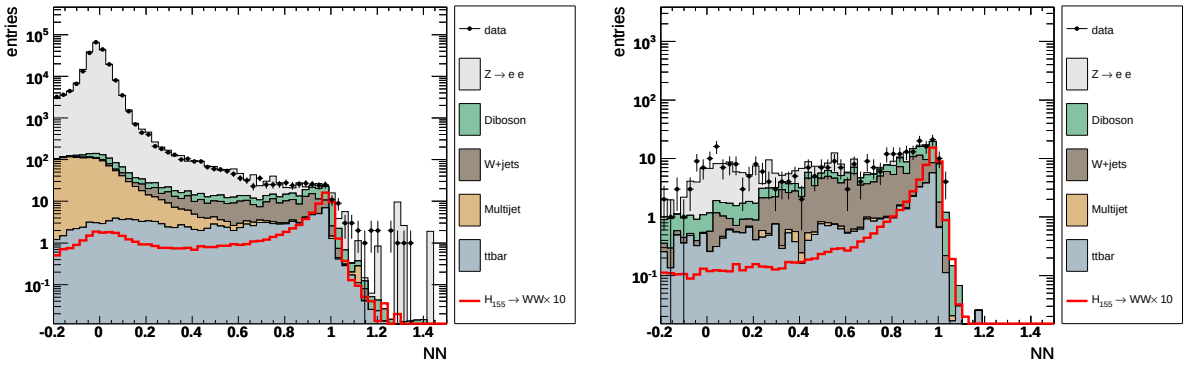


Figure B.12: NN output for $m_{\mathrm{H}} = 155$ GeV at preselection level (left) and (right) NN output for $m_{\mathrm{H}} = 155$ GeV after the $\Delta\phi(\ell,\ell)$ selection cut (right). The signal indicated by the red graph is representative of a Higgs boson mass of $m_{\mathrm{H}} = 160$ GeV. The entire Run II data set is used.

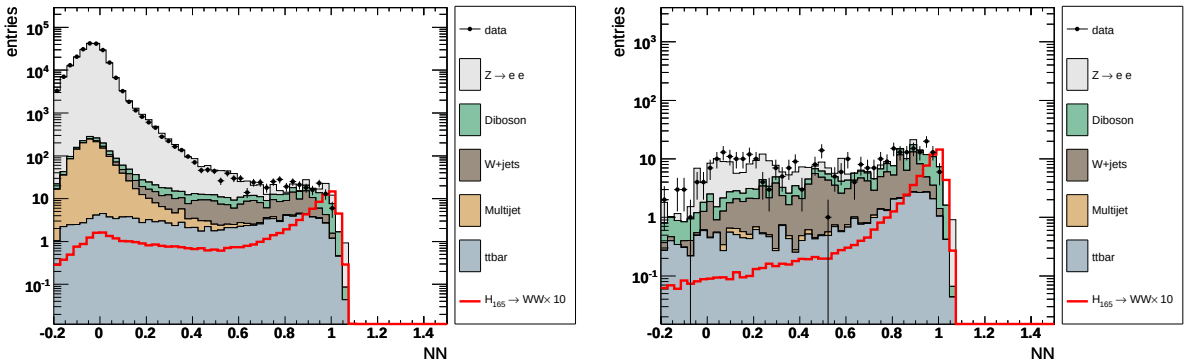


Figure B.13: NN output for $m_{\mathrm{H}} = 165$ GeV at preselection level (left) and (right) NN output for $m_{\mathrm{H}} = 165$ GeV after the $\Delta\phi(\ell,\ell)$ selection cut (right). The signal indicated by the red graph is representative of a Higgs boson mass of $m_{\mathrm{H}} = 160$ GeV. The entire Run II data set is used.

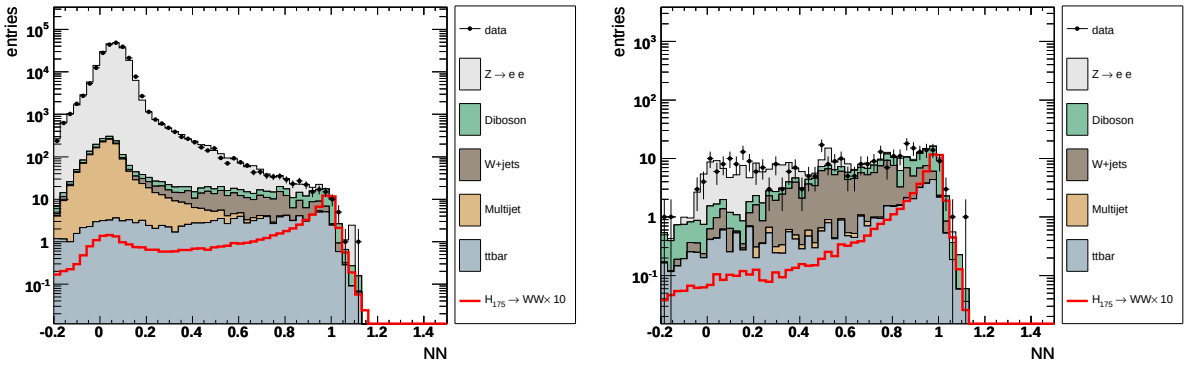


Figure B.14: NN output for $m_{\mathrm{H}} = 175$ GeV at preselection level (left) and (right) NN output for $m_{\mathrm{H}} = 175$ GeV after the $\Delta\phi(\ell,\ell)$ selection cut (right). The signal indicated by the red graph is representative of a Higgs boson mass of $m_{\mathrm{H}} = 160$ GeV. The entire Run II data set is used.

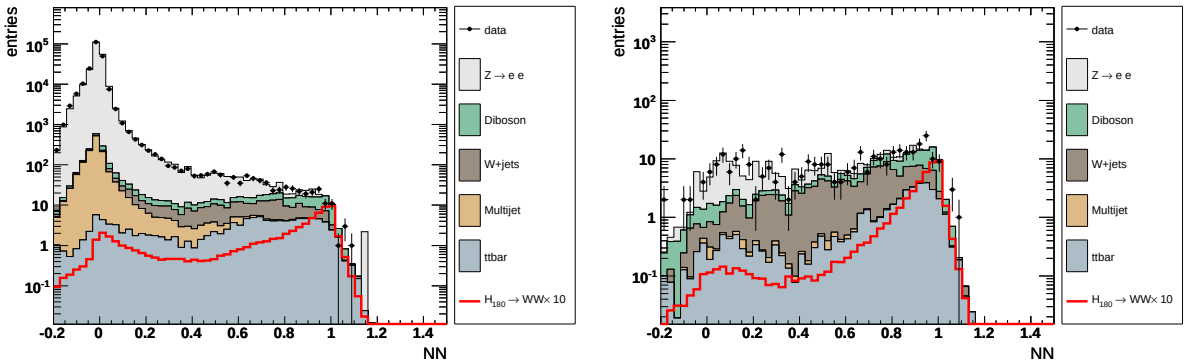


Figure B.15: NN output for $m_{\mathrm{H}} = 180$ GeV at preselection level (left) and (right) NN output for $m_{\mathrm{H}} = 180$ GeV after the $\Delta\phi(\ell,\ell)$ selection cut (right). The signal indicated by the red graph is representative of a Higgs boson mass of $m_{\mathrm{H}} = 160$ GeV. The entire Run II data set is used.

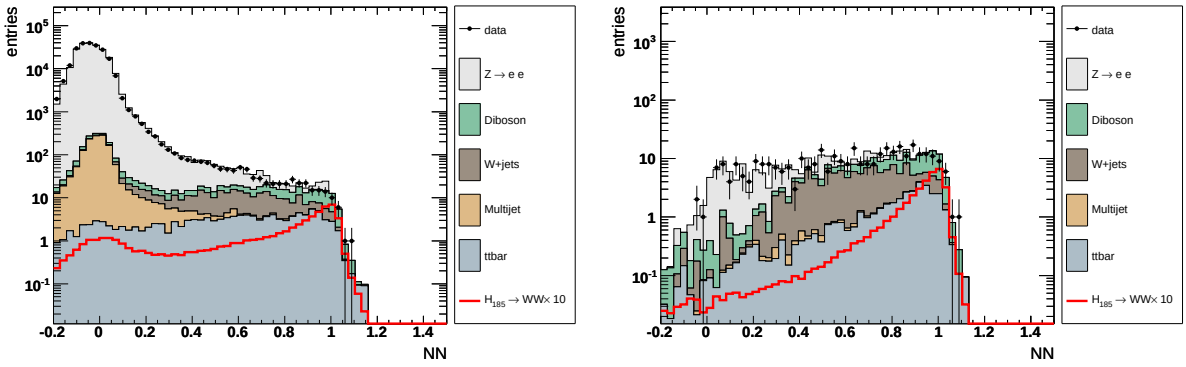


Figure B.16: NN output for $m_{\mathrm{H}} = 185$ GeV at preselection level (left) and (right) NN output for $m_{\mathrm{H}} = 185$ GeV after the $\Delta\phi(\ell,\ell)$ selection cut (right). The signal indicated by the red graph is representative of a Higgs boson mass of $m_{\mathrm{H}} = 160$ GeV. The entire Run II data set is used.

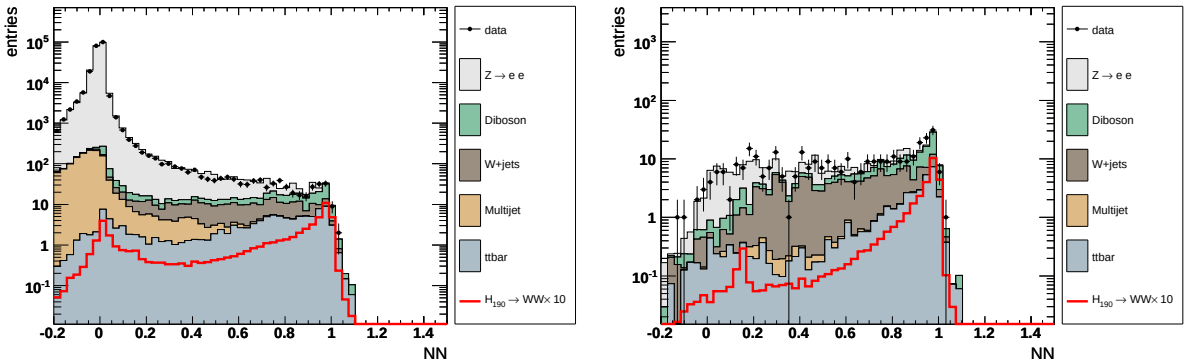


Figure B.17: NN output for $m_{\mathrm{H}} = 190$ GeV at preselection level (left) and (right) NN output for $m_{\mathrm{H}} = 190$ GeV after the $\Delta\phi(\ell,\ell)$ selection cut (right). The signal indicated by the red graph is representative of a Higgs boson mass of $m_{\mathrm{H}} = 160$ GeV. The entire Run II data set is used.

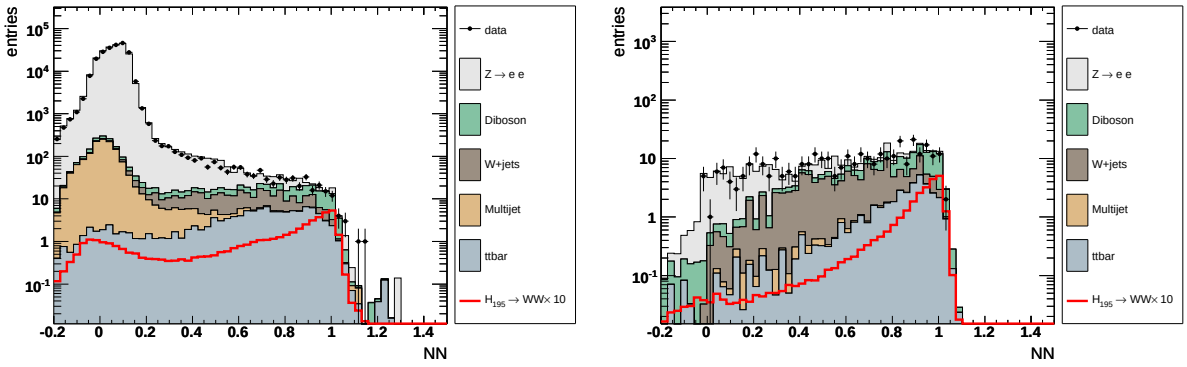


Figure B.18: NN output for $m_{\mathrm{H}} = 195$ GeV at preselection level (left) and (right) NN output for $m_{\mathrm{H}} = 195$ GeV after the $\Delta\phi(\ell,\ell)$ selection cut (right). The signal indicated by the red graph is representative of a Higgs boson mass of $m_{\mathrm{H}} = 160$ GeV. The entire Run II data set is used.

C
Appendix for the $H \rightarrow WW^* \rightarrow \mu + \tau_{had} + X$ selection

C.1 Event Yields $H \rightarrow WW^* \rightarrow \mu + \tau_{had} + X$

Cut	Data	Sum Bkgd	Signal	WW	$W \rightarrow \mu\nu$	$W \rightarrow e\nu$	$W \rightarrow \tau\nu$	$t\bar{t}$	$Z \rightarrow \mu\mu$	$Z \rightarrow \tau\tau$	WZ	ZZ	QCD
Preselection	1749.00 ± 41.82	1719.19 ± 33.58	0.05	13.06	176.35	4.43	16.83	6.50	64.36	671.07	2.24	0.58	763.78
$\not\!\!E_{\mathrm{T}}$	408.00 ± 20.20	440.16 ± 15.29	0.04	10.61	141.54	3.85	10.73	6.05	18.65	127.10	1.75	0.39	119.50
$\not\!\!E_{\mathrm{T}}^{Sig}$	301.00 ± 17.35	292.92 ± 10.71	0.04	9.83	137.97	3.85	10.73	4.10	16.05	89.24	1.41	0.24	19.51
M_T^{min}	80.00 ± 8.94	76.83 ± 5.09	0.03	5.86	56.32	2.63	3.09	2.42	3.75	0.30	0.79	0.09	1.58
$\sum p_T$	72.00 ± 8.49	70.66 ± 4.96	0.03	4.28	54.18	2.63	3.09	0.78	3.72	0.30	0.47	0.05	1.16
M_{inv}	47.00 ± 6.86	43.93 ± 4.00	0.02	1.98	33.16	2.12	1.99	0.42	2.95	0.28	0.18	0.03	0.83
H_T	47.00 ± 6.86	43.57 ± 4.00	0.02	1.92	33.16	2.12	1.99	0.11	2.95	0.28	0.18	0.03	0.83
$\Delta\phi(\mu,\tau)$	47.00 ± 6.86	42.36 ± 3.96	0.02	1.85	32.07	2.12	1.99	0.10	2.95	0.28	0.14	0.03	0.83
Llhood incl.	19.00 ± 4.36	20.36 ± 2.71	0.02	1.06	15.77	0.66	0.50	0.07	2.16	0.00	0.11	0.02	0.00
Veto incl.	19.00 ± 4.36	20.36 ± 2.71	0.02	1.06	15.77	0.66	0.50	0.07	2.16	0.00	0.11	0.02	0.00
Llhood excl.	30.00 ± 5.48	23.07 ± 2.97	0.01	0.91	16.68	1.32	1.07	0.08	2.95	0.00	0.04	0.02	0.00
Veto excl.	17.00 ± 4.12	12.04 ± 2.31	0.00	0.34	6.23	1.32	1.07	0.08	2.95	0.00	0.04	0.02	0.00

Table C.1: Number of expected background events and number of observed data events, after successive selections for τ-type I. The expected signal content is given with respect to $m_{\mathrm{H}} = 120$ GeV. After the $\Delta\phi(\mu,\tau)$ the likelihood cuts for the inclusive ($H \rightarrow WW \rightarrow \mu e\nu\nu$) sample and the exclusive ($H \rightarrow WW \rightarrow \mu\tau_{had}\nu\nu$) sample are applied individually and a veto against the $H \rightarrow WW^* \rightarrow e^{\pm}\nu\mu^{\mp}\nu$ analysis is applied. The data set corresponds to 1.0 fb^{-1} in integrated luminosity. Only statistical uncertainties are given.

Cut	Data	Sum Bkgd	Signal	WW	$W \rightarrow \mu\nu$	$W \rightarrow e\nu$	$W \rightarrow \tau\nu$	$t\bar{t}$	$Z \rightarrow \mu\mu$	$Z \rightarrow \tau\tau$	WZ	ZZ	QCD
Preselection	4786.00 ± 69.18	4737.06 ± 47.96	0.65	129.94	401.26	30.37	42.14	83.35	176.41	2778.40	15.96	3.88	1075.35
$\not{E}_T$	1332.00 ± 36.50	1351.17 ± 22.79	0.57	106.78	335.08	25.83	24.51	77.75	62.86	525.05	13.15	2.40	177.77
$\not{E}_T^{Sig}$	1058.00 ± 32.53	957.56 ± 17.08	0.53	100.96	321.06	23.13	21.90	51.51	53.99	334.12	11.52	1.64	37.73
M_T^{min}	313.00 ± 17.69	325.35 ± 8.50	0.40	72.76	171.74	14.64	10.21	36.95	7.80	0.72	7.54	0.72	2.26
$\sum p_T$	230.00 ± 15.17	246.32 ± 8.02	0.40	47.93	152.31	13.41	9.80	8.54	6.94	0.68	4.21	0.38	2.12
M_{inv}	148.00 ± 12.17	149.18 ± 6.52	0.37	22.61	94.59	10.90	7.61	4.46	5.59	0.66	1.53	0.14	1.09
H_T	143.00 ± 11.96	146.20 ± 6.52	0.36	22.22	94.59	10.90	7.61	1.87	5.59	0.66	1.53	0.13	1.09
$\Delta\phi(\mu,\tau)$	143.00 ± 11.96	140.17 ± 6.36	0.35	21.13	92.59	8.69	7.14	1.71	5.59	0.66	1.45	0.13	1.09
Llhood incl.	56.00 ± 7.48	50.01 ± 3.51	0.28	16.18	19.82	8.69	1.59	1.23	1.49	0.00	0.93	0.08	0.00
Veto incl.	42.00 ± 6.48	42.35 ± 3.47	0.14	9.27	19.06	8.69	1.59	1.23	1.49	0.00	0.93	0.08	0.00
Llhood excl.	42.00 ± 6.48	41.05 ± 3.39	0.10	5.97	27.29	3.06	2.66	0.58	0.83	0.00	0.63	0.04	0.00
Veto excl.	21.00 ± 4.58	29.19 ± 3.06	0.01	1.01	20.38	3.06	2.66	0.58	0.83	0.00	0.63	0.04	0.00

Table C.2: Number of expected background events and number of observed data events, after successive selections for τ-type II. The expected signal content is given with respect to $m_{\mathrm{H}} = 120$ GeV. After the $\Delta\phi(\mu,\tau)$ the likelihood cuts for the inclusive ($H \rightarrow WW \rightarrow \mu e\nu\nu$) sample and the exclusive ($H \rightarrow WW \rightarrow \mu\tau_{had}\nu\nu$) sample are applied individually and a veto against the $H \rightarrow WW^* \rightarrow e^{\pm}\nu\mu^{\mp}\nu$ analysis is applied. The data set corresponds to 1.0 fb^{-1} in integrated luminosity. Only statistical uncertainties are given.

Cut	Data	Sum Bkgd	Signal	WW	$W \rightarrow \mu\nu$	$W \rightarrow e\nu$	$W \rightarrow \tau\nu$	$t\bar{t}$	$Z \rightarrow \mu\mu$	$Z \rightarrow \tau\tau$	WZ	ZZ	QCD
Preselection	1749.00 ± 41.82	1719.19 ± 33.58	0.16	13.06	176.35	4.43	16.83	6.50	64.36	671.07	2.24	0.58	763.78
$\not\!\!E_{\mathrm{T}}$	408.00 ± 20.20	441.27 ± 15.30	0.14	10.62	141.20	3.85	10.73	6.04	18.65	128.56	1.75	0.39	119.50
$\not\!\!E_{\mathrm{T}}^{Sig}$	301.00 ± 17.35	292.26 ± 10.70	0.13	9.82	138.02	3.85	10.73	4.13	16.05	88.51	1.40	0.24	19.51
M_T^{min}	58.00 ± 7.62	60.38 ± 4.52	0.08	4.96	42.57	2.63	2.61	2.19	2.98	0.30	0.62	0.08	1.45
$\sum p_T$	54.00 ± 7.35	55.45 ± 4.40	0.08	3.89	40.75	2.63	2.61	0.89	2.95	0.30	0.36	0.05	1.03
M_{inv}	39.00 ± 6.24	40.79 ± 3.88	0.07	2.20	29.02	2.63	2.04	0.60	2.95	0.28	0.12	0.03	0.94
H_T	39.00 ± 6.24	40.28 ± 3.88	0.07	2.14	29.02	2.63	2.04	0.15	2.95	0.28	0.12	0.03	0.94
$\Delta\phi(\mu,\tau)$	39.00 ± 6.24	38.60 ± 3.81	0.07	2.03	27.59	2.63	2.04	0.10	2.95	0.28	0.12	0.03	0.83
Llhood incl.	3.00 ± 1.73	7.36 ± 1.57	0.02	0.65	4.97	1.32	0.00	0.07	0.00	0.28	0.06	0.01	0.00
Veto incl.	3.00 ± 1.73	7.36 ± 1.57	0.02	0.65	4.97	1.32	0.00	0.07	0.00	0.28	0.06	0.01	0.00
Llhood excl.	7.00 ± 2.65	6.37 ± 1.50	0.02	0.46	4.93	0.00	0.00	0.05	0.88	0.00	0.03	0.00	0.00
Veto excl.	6.00 ± 2.45	4.94 ± 1.36	0.01	0.28	3.68	0.00	0.00	0.05	0.88	0.00	0.03	0.00	0.00

Table C.3: Number of expected background events and number of observed data events, after successive selections for τ-type I. The expected signal content is given with respect to $m_{\mathrm{H}} = 140$ GeV. After the $\Delta\phi(\mu,\tau)$ the likelihood cuts for the inclusive ($H \rightarrow WW \rightarrow \mu e\nu\nu$) sample and the exclusive ($H \rightarrow WW \rightarrow \mu\tau_{had}\nu\nu$) sample are applied individually and a veto against the $H \rightarrow WW^* \rightarrow e^{\pm}\nu\mu^{\mp}\nu$ analysis is applied. The data set corresponds to 1.0 fb^{-1} in integrated luminosity. Only statistical uncertainties are given.

Cut	Data	Sum Bkgd	Sig.	WW	$W \rightarrow \mu\nu$	$W \rightarrow e\nu$	$W \rightarrow \tau\nu$	$t\bar{t}$	$Z \rightarrow \mu\mu$	$Z \rightarrow \tau\tau$	WZ	ZZ	QCD
Preselection	4786.00 ± 69.18	4737.06 ± 47.96	1.96	129.94	401.26	30.37	42.14	83.35	176.41	2778.40	15.96	3.88	1075.35
$\not{E}_T$	1332.00 ± 36.50	1350.43 ± 22.78	1.76	106.78	335.45	24.60	24.51	77.77	62.86	525.14	13.16	2.40	177.77
$\not{E}_T^{Sig}$	1058.00 ± 32.53	957.75 ± 17.08	1.64	100.95	321.03	22.53	21.90	51.34	53.99	335.10	11.53	1.64	37.73
M_T^{min}	250.00 ± 15.81	269.35 ± 7.48	1.25	65.24	136.20	11.04	7.50	34.03	6.29	0.70	6.87	0.60	0.88
$\sum p_T$	195.00 ± 13.96	208.90 ± 7.07	1.21	47.03	123.04	9.81	7.08	10.21	5.42	0.68	4.53	0.36	0.74
M_{inv}	136.00 ± 11.66	144.11 ± 6.11	1.10	27.34	88.66	7.94	5.44	6.40	5.42	0.66	1.82	0.17	0.25
H_T	128.00 ± 11.31	139.53 ± 6.10	1.06	26.76	88.66	7.94	5.44	2.41	5.42	0.66	1.82	0.17	0.25
$\Delta\phi(\mu,\tau)$	124.00 ± 11.14	133.15 ± 5.99	1.03	24.97	84.91	7.48	5.44	2.18	5.42	0.66	1.69	0.16	0.25
Llhood incl.	42.00 ± 6.48	39.48 ± 2.84	0.77	17.75	9.65	6.30	1.74	1.43	1.50	0.00	1.01	0.09	0.00
Veto incl.	25.00 ± 5.00	28.82 ± 2.77	0.29	8.18	8.56	6.30	1.74	1.43	1.50	0.00	1.01	0.09	0.00
Llhood excl.	41.00 ± 6.40	47.75 ± 3.59	0.26	7.02	33.96	1.95	1.51	0.88	1.49	0.09	0.76	0.09	0.00
Veto excl.	27.00 ± 5.20	39.97 ± 3.47	0.07	2.06	31.14	1.95	1.51	0.88	1.49	0.09	0.76	0.09	0.00

Table C.4: Number of expected background events and number of observed data events, after successive selections for τ-type II. The expected signal content is given with respect to $m_{\mathrm{H}} = 140$ GeV. After the $\Delta\phi(\mu,\tau)$ the likelihood cuts for the inclusive ($H \rightarrow WW \rightarrow \mu e\nu\nu$) sample and the exclusive ($H \rightarrow WW \rightarrow \mu\tau_{had}\nu\nu$) sample are applied individually and a veto against the $H \rightarrow WW^* \rightarrow e^{\pm}\nu\mu^{\mp}\nu$ analysis is applied. The data set corresponds to 1.0 fb^{-1} in integrated luminosity. Only statistical uncertainties are given.

Cut	Data	Sum Bkgd	Signal	WW	$W \rightarrow \mu\nu$	$W \rightarrow e\nu$	$W \rightarrow \tau\nu$	$t\bar{t}$	$Z \rightarrow \mu\mu$	$Z \rightarrow \tau\tau$	WZ	ZZ	QCD
Preselection	1749.00 ± 41.82	1719.19 ± 33.58	0.15	13.06	176.35	4.43	16.83	6.50	64.36	671.07	2.24	0.58	763.78
$\not\!\!E_{\mathrm{T}}$	408.00 ± 20.20	440.26 ± 15.29	0.14	10.63	141.54	3.85	10.73	6.04	18.65	127.20	1.75	0.39	119.50
$\not\!\!E_{\mathrm{T}}^{Sig}$	301.00 ± 17.35	293.12 ± 10.71	0.13	9.84	137.97	3.85	10.73	4.11	16.05	89.42	1.40	0.24	19.51
M_T^{min}	44.00 ± 6.63	40.68 ± 3.58	0.09	4.07	29.23	1.82	1.07	1.84	0.92	0.28	0.56	0.06	0.84
$\sum p_T$	38.00 ± 6.16	37.62 ± 3.46	0.08	3.43	28.09	1.82	1.07	1.14	0.88	0.28	0.42	0.05	0.43
M_{inv}	31.00 ± 5.57	28.17 ± 3.05	0.08	2.28	20.34	1.82	1.07	0.97	0.88	0.28	0.15	0.04	0.33
H_T	31.00 ± 5.57	27.32 ± 3.05	0.07	2.19	20.34	1.82	1.07	0.21	0.88	0.28	0.15	0.04	0.33
$\Delta\phi(\mu,\tau)$	31.00 ± 5.57	24.26 ± 2.87	0.07	1.99	18.21	1.82	0.50	0.17	0.88	0.28	0.14	0.03	0.23
Llhood incl.	1.00 ± 1.00	1.25 ± 0.60	0.01	0.14	1.08	0.00	0.00	0.00	0.00	0.00	0.02	0.01	0.00
Veto incl.	1.00 ± 1.00	1.25 ± 0.60	0.01	0.14	1.08	0.00	0.00	0.00	0.00	0.00	0.02	0.01	0.00
Llhood excl.	3.00 ± 1.73	6.98 ± 1.40	0.04	0.75	5.77	0.00	0.00	0.08	0.00	0.28	0.09	0.01	0.00
Veto excl.	3.00 ± 1.73	5.79 ± 1.27	0.03	0.64	4.69	0.00	0.00	0.08	0.00	0.28	0.09	0.01	0.00

Table C.5: Number of expected background events and number of observed data events, after successive selections for τ-type I. The expected signal content is given with respect to $m_{\mathrm{H}} = 180$ GeV. After the $\Delta\phi(\mu,\tau)$ the likelihood cuts for the inclusive ($H \rightarrow WW \rightarrow \mu e\nu\nu$) sample and the exclusive ($H \rightarrow WW \rightarrow \mu\tau_{had}\nu\nu$) sample are applied individually and a veto against the $H \rightarrow WW^* \rightarrow e^{\pm}\nu\mu^{\mp}\nu$ analysis is applied. The data set corresponds to 1.0 fb^{-1} in integrated luminosity. Only statistical uncertainties are given.

Cut	Data	Sum Bkgd	Signal	WW	$W \rightarrow \mu\nu$	$W \rightarrow e\nu$	$W \rightarrow \tau\nu$	$t\bar{t}$	$Z \rightarrow \mu\mu$	$Z \rightarrow \tau\tau$	WZ	ZZ	QCD
Preselection	4786.00 ± 69.18	4737.06 ± 47.96	1.87	129.94	401.26	30.37	42.14	83.35	176.41	2778.40	15.96	3.88	1075.35
$\not{E}_T$	1332.00 ± 36.50	1349.29 ± 22.77	1.76	106.79	335.37	25.22	24.51	77.73	62.86	523.50	13.14	2.40	177.77
$\not{E}_T^{Sig}$	1058.00 ± 32.53	956.07 ± 17.07	1.66	100.98	320.80	22.53	21.90	51.50	53.99	333.50	11.51	1.65	37.73
M_T^{min}	208.00 ± 14.42	216.08 ± 6.41	1.36	57.14	103.35	8.69	5.28	31.33	2.63	0.29	6.23	0.51	0.62
$\sum p_T$	170.00 ± 13.04	171.54 ± 5.81	1.28	48.21	86.78	8.04	4.08	17.38	0.82	0.28	5.05	0.39	0.51
M_{inv}	127.00 ± 11.27	133.78 ± 5.34	1.17	33.59	70.37	8.04	3.86	13.37	0.78	0.28	2.75	0.26	0.47
H_T	115.00 ± 10.72	122.05 ± 5.27	1.10	32.73	69.60	7.44	3.86	3.97	0.78	0.28	2.70	0.26	0.44
$\Delta\phi(\mu,\tau)$	108.00 ± 10.39	111.55 ± 5.08	1.04	28.98	65.47	6.82	2.86	3.25	0.78	0.28	2.43	0.23	0.44
Llhood incl.	27.00 ± 5.20	20.81 ± 1.73	0.60	10.02	5.92	1.76	0.00	1.88	0.04	0.00	0.94	0.09	0.16
Veto incl.	21.00 ± 4.58	14.12 ± 1.69	0.18	3.68	5.57	1.76	0.00	1.88	0.04	0.00	0.94	0.09	0.16
Llhood excl.	16.00 ± 4.00	12.31 ± 1.58	0.16	3.18	5.34	1.23	0.00	1.50	0.02	0.00	0.80	0.07	0.16
Veto excl.	2.00 ± 1.41	5.10 ± 1.10	0.01	0.19	1.11	1.23	0.00	1.50	0.02	0.00	0.80	0.07	0.16

Table C.6: Number of expected background events and number of observed data events, after successive selections for τ-type II. The expected signal content is given with respect to $m_{\mathrm{H}} = 180$ GeV. After the $\Delta\phi(\mu,\tau)$ the likelihood cuts for the inclusive ($H \rightarrow WW \rightarrow \mu e\nu\nu$) sample and the exclusive ($H \rightarrow WW \rightarrow \mu\tau_{had}\nu\nu$) sample are applied individually and a veto against the $H \rightarrow WW^* \rightarrow e^{\pm}\nu\mu^{\mp}\nu$ analysis is applied. The data set corresponds to 1.0 fb^{-1} in integrated luminosity. Only statistical uncertainties are given.

C.2 Likelihood Input Distributions

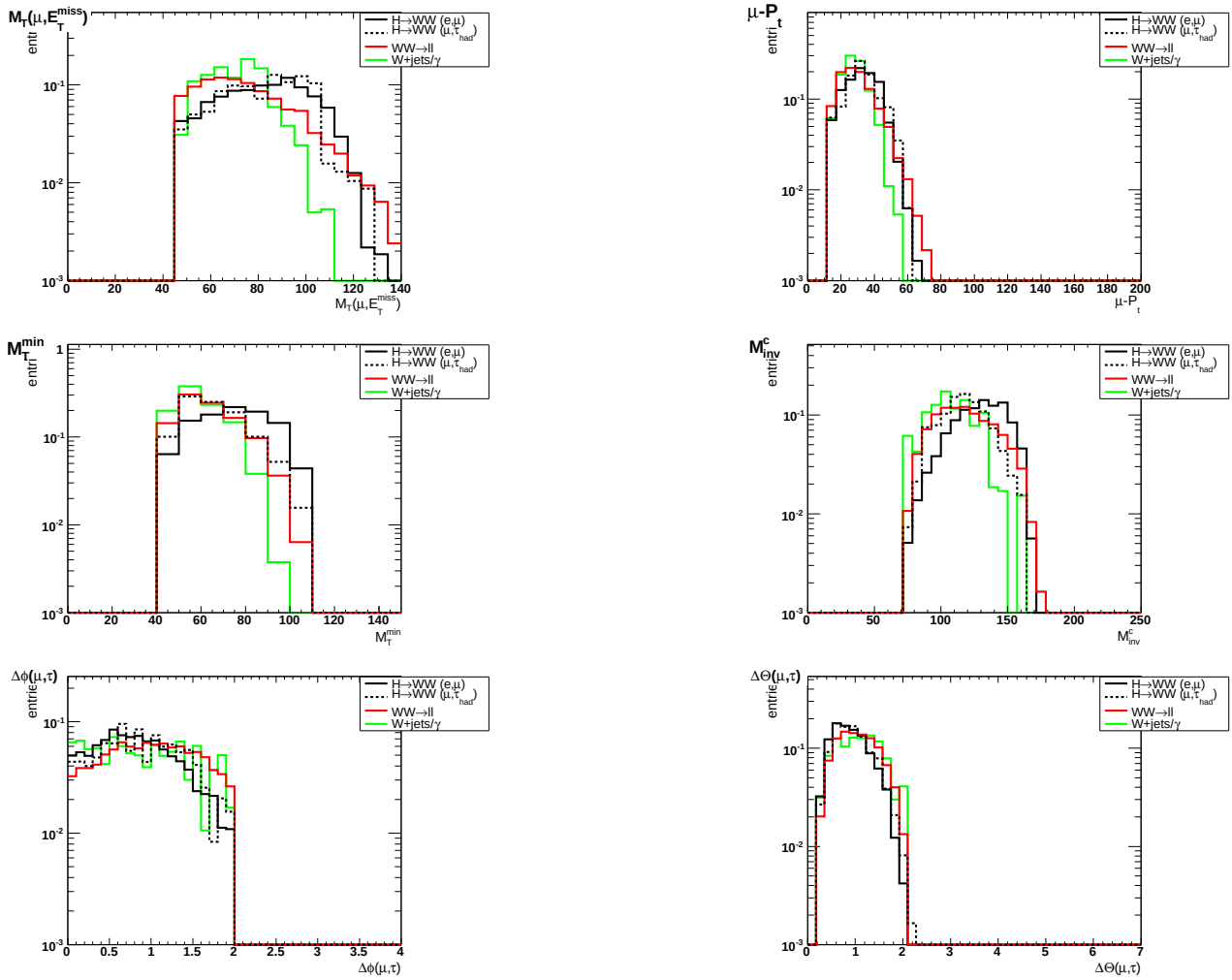


Figure C.1: Reference Histograms used for construction of the kinematic likelihood. The input distributions are divided in two distinct samples, one with $H \rightarrow WW^* \rightarrow e^{\pm}\nu\mu^{\mp}\nu$ where the e is faking the tau and the second using $H \rightarrow WW^* \rightarrow \mu + \tau_{had} + X$ events. The red line represents the shape of the $WW \rightarrow \ell^+\nu\ell^-\nu$ sample, the green histogram the $W+$jets background and the black lines the two signal samples. $e\mu$ events as solid and $\mu\tau$ events as dashed graph. One sees that the $WW \rightarrow \ell^+\nu\ell^-\nu$ background is generally very similar in shape to the $H \rightarrow WW^* \rightarrow e^{\pm}\nu\mu^{\mp}\nu$ events whereas the $W+$jets background resembles the $H \rightarrow WW^* \rightarrow \mu + \tau_{had} + X$ signature.

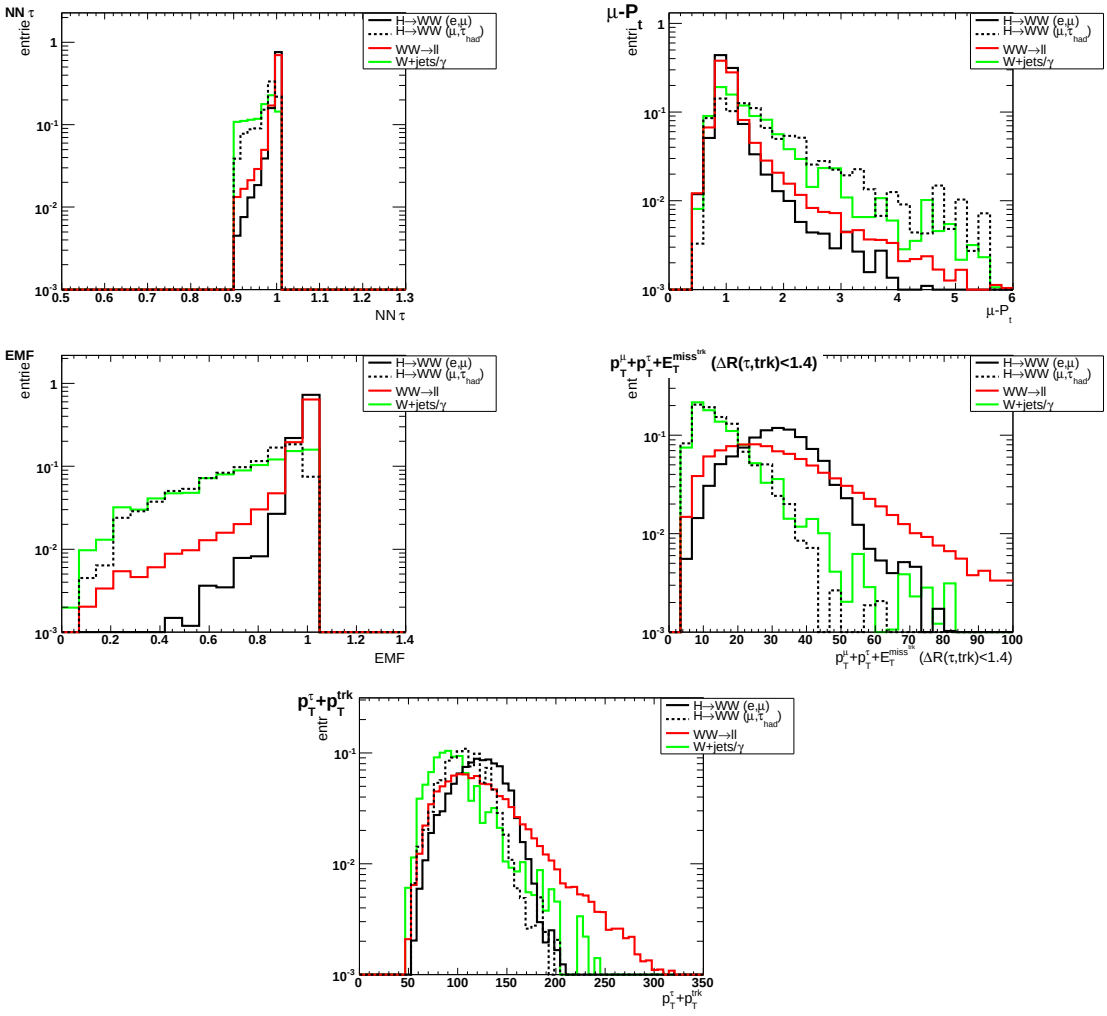


Figure C.2: Reference Histograms used for construction of the tau likelihood. The input distributions are divided in two distinct samples, one with $H \rightarrow WW^* \rightarrow \mu + \tau_{had} + X$ where the e is faking the tau and the second using $H \rightarrow WW^* \rightarrow \mu + \tau_{had} + X$ events. The red line gives the shape of the $WW \rightarrow \ell^+\nu\ell^-\nu$ sample, the green histogram represents the $W+$jets background and the black lines the two signal samples - $e\mu$ events (solid) and $\mu\tau$ events (dashed). One sees that the $WW \rightarrow \ell^+\nu\ell^-\nu$ background is generally very similar in shape to the $H \rightarrow WW^* \rightarrow e^{\pm}\nu\mu^{\mp}\nu$ events whereas the $W+$jets background resembles the $H \rightarrow WW^* \rightarrow \mu + \tau_{had} + X$ signature.

C.3 Detailed Systematics for $H \rightarrow WW^* \rightarrow \mu + \tau_{had} + X$ analysis

Variation	Data	Sum Bkgd	$H \to WW$	WW	$W \to \mu\nu$	$W \to e\nu$	$W \to \tau\nu$	$t\bar{t}$	$Z \to \mu\mu$	$Z \to \tau\tau$	WZ	ZZ	QCD
τ I incl jssr up	0.00	-0.56	0.00	0.00	0.00	0.00	—	-18.91	—	0.00	-19.94	0.00	—
τ I incl jssr down	0.00	-0.44	0.06	-0.97	0.00	0.00	—	-19.74	—	0.00	0.00	0.00	—
τ I incl reso up	—	-0.56	0.00	0.00	0.00	0.00	—	-18.91	—	0.00	-19.94	0.00	—
τ I incl reso down	—	-0.44	0.97	-0.97	0.00	0.00	—	-19.74	—	0.00	0.00	0.00	—
τ I incl eff up	—	-0.56	0.00	0.00	0.00	0.00	—	-18.91	—	0.00	-19.94	0.00	—
τ I incl eff down	—	-0.44	-1.91	-0.97	0.00	0.00	—	-19.74	—	0.00	0.00	0.00	—
τ II incl jssr up	-4.00	-0.71	0.33	-0.21	0.00	0.00	0.00	-8.83	0.00	—	0.00	0.00	—
τ II incl jssr down	0.00	-0.59	-0.88	-0.81	0.00	0.00	0.00	-1.77	0.00	—	-1.29	-8.61	—
τ II incl reso up	—	-0.33	0.13	-0.35	0.00	0.00	0.00	-2.24	0.00	—	0.00	0.00	—
τ II incl reso down	—	-0.27	-0.63	-0.63	0.00	0.00	0.00	1.51	0.00	—	-1.29	-8.61	—
τ II incl eff up	—	-0.15	0.05	-0.06	0.00	0.00	0.00	-1.84	0.00	—	0.00	0.00	—
τ II incl eff down	—	-0.50	-0.73	-0.85	0.00	0.00	0.00	-0.12	0.00	—	-1.29	-8.61	—
τ I excl jssr up	0.00	-0.32	0.00	0.00	0.00	—	—	-23.36	—	0.00	0.00	0.00	—
τ I excl jssr down	0.00	-0.51	0.00	-1.88	0.00	—	—	-24.38	—	0.00	0.00	0.00	—
τ I excl reso up	—	-0.32	0.00	0.00	0.00	—	—	-23.36	—	0.00	0.00	0.00	—
τ I excl reso down	—	-0.51	-0.13	-1.88	0.00	—	—	-24.38	—	0.00	0.00	0.00	—
τ I excl eff up	—	-0.32	1.59	0.00	0.00	—	—	-23.36	—	0.00	0.00	0.00	—
τ I excl eff down	—	-0.51	-1.16	-1.88	0.00	—	—	-24.38	—	0.00	0.00	0.00	—
τ II excl jssr up	0.00	-0.25	0.84	0.56	0.00	0.00	—	-4.01	0.00	—	0.00	0.00	0.00
τ II excl jssr down	0.00	-0.38	-0.03	-0.09	0.00	0.00	—	-2.29	0.00	—	-2.76	-12.77	0.00
τ II excl reso up	—	-1.69	1.07	-0.01	0.00	0.00	—	2.27	0.00	—	0.00	0.00	—
τ II excl reso down	—	-1.85	0.69	0.20	0.00	0.00	—	1.69	0.00	—	-2.76	-12.77	—
τ II excl eff up	—	-1.32	0.27	0.15	0.00	0.00	—	5.83	0.00	—	0.00	0.00	—
τ II excl eff down	—	-1.65	0.47	-0.01	0.00	0.00	—	4.32	0.00	—	-2.76	-12.77	—

Table C.7: Systematic error on Jet Energy Scale, Jet Resolution and Jet ID in % for the various final states and selections. All numbers given are obtained by varying the associated quantity one sigma up respectively down.