4-2010

Search for the Standard Model Higgs Boson in p-pbar Interactions with the decay mode H -> W+ W- -> mu+ nu mu- nu at the D0 Experiment

Dale M. Johnston
University of Nebraska at Lincoln, dalej96@gmail.com

Johnston, Dale M., "Search for the Standard Model Higgs Boson in p-pbar Interactions with the decay mode H -> W+ W- -> mu+ nu mu- nu at the D0 Experiment" (2010). *Theses, Dissertations, and Student Research: Department of Physics and Astronomy*. 9.
http://digitalcommons.unl.edu/physicsdiss/9

SEARCH FOR THE STANDARD MODEL HIGGS BOSON IN $p\bar{p}$ INTERACTIONS WITH THE DECAY MODE $H \rightarrow W^+W^- \rightarrow \mu^+\nu\mu^-\nu$ AT THE DØ EXPERIMENT

by

Dale Morgan Johnston

A DISSERTATION

Presented to the Faculty of
The Graduate College at the University of Nebraska
In Partial Fulfillment of Requirements
For the Degree of Doctor of Philosophy

Major: Physics and Astronomy

Under the Supervision of Professor Aaron Dominguez

Lincoln, Nebraska

April, 2010

SEARCH FOR THE STANDARD MODEL HIGGS BOSON IN $p\bar{p}$ INTERACTIONS WITH THE DECAY MODE $H \to W^+W^- \to \mu^+\nu\mu^-\nu$ AT THE DØ EXPERIMENT

Dale Morgan Johnston, Ph.D.

University of Nebraska, 2010

Advisor: Aaron Dominguez

A search for the standard model Higgs boson in $p\bar{p}$ collisions resulting in two muons and large missing transverse energy is presented. The analysis uses 4.2 fb^{-1}of integrated luminosity at a center-of-mass energy of $\sqrt{s} = 1.96$ TeV collected between April 2002 and December 2008 with the D0 detector at the Fermilab Tevatron collider. No significant excess above the background estimation is observed and limits are derived on Higgs boson production.

For Leah.

ACKNOWLEDGEMENTS

There are many people I would like to thank for their support and guidance in this massive undertaking. First and most especially Leah, without your love, patience, understanding, and support throughout the years I never would have made it this far. Special thanks also to my parents, Dean and Alice, for your love and support and also for the good foundation on which to build my life. I have always had the freedom to explore, grow, and learn whatever and wherever I wanted and that means everything to me. Also, I can never express how much I appreciate the endless support of my whole family and the Varney family, who all believed in me and stood by me through the good times and the hardest ones.

At UNL, thanks to Aaron Dominguez and Ken Bloom. Your help, guidance, and financial support has been instrumental in my making it through the entire research and thesis writing process. Pushing me to move out to Fermilab definitely proved to be the right move and I owe a lot to that push. Thanks for believing in me and sticking with me until I found the right environment in which to do my research. I would also like to thank Dan Claes and Greg Snow for helping to draw me in to high energy physics in the first place and founding the UNLHEP group that I call home. Also thanks to Kayle DeVaughan, my comrade in arms, for his help and friendship throughout the years. Some of the times would likely have been unbearable without it. Thanks to Shawn Langan, another close friend who has stuck with me through the years and given me his ear for a ridiculous number of complaining sessions. Finally, a special thanks to Ed Schmidt, who gave me a start at UNL and an opportunity to do research when I was just a fledgling graduate student.

At D0, I owe a large debt of gratitude to Björn Penning, who helped me shape this analysis channel into what it is and who stood by me through countless all-nighters when many questioned whether success was possible. Joe Haley was also a good friend through it all and helped me keep my spirits up when times were tough. Also thanks to Marco Verzocchi and Wade Fisher who made me learn what it means to do good research and gave me many hours of their time and many (MANY) words of constructive criticism. Without that motivation I would still be spinning my wheels. Thanks especially to Mike Eads for helping me in the trenches and taking some of the pressure off when I needed it most. Thanks also to Andy Haas and Thomas Gadfort, who got me started in this analysis channel and gave me help in the beginning. Finally, there are many wonderful scientists who contributed to my understanding of physics and quality of work at D0, as well as provided me motivation to keep on top of things as best I could. I want to thank them all, especially Stefan Soldner-Rembold, Aurelio Juste, Gavin Davies, and Don Lincoln.

Thanks to the University of Nebraska, Lincoln and the entire D0 Collaboration for their extensive support and for providing wonderful environments to work. To those who I didn't mention, I greatly appreciate all of the support from everyone over the years. I couldn't have done it without each of you. Cheers!

An updated version of the work described in this dissertation, in concert with other analysis channels, has been published in the following journals:

arXiv:1001.4481v3 [hep-ex]

FERMILAB-PUB-10-015-E

Phys.Rev.Lett.:104,061804(2010)

Contents

List of Tables

List of Figures

Chapter 1

Introduction

It is human nature to be curious about the world around us. We strive to create for ourselves a picture of how the world works and where we fit in it. For scientists this is carried one step further with a need to describe and model things well enough to predict the behavior of that world. This quest to understand how everything fits together on a fundamental level is what drives particle physicists to search for, potentially, the last piece of the puzzle in the standard model of particle physics, the Higgs boson. If this particle is found it might give a complete predictive model of all the known fundamental particles and their interactions with each other, excepting gravity.

The theorized Higgs boson is named after Peter Higgs and was dubbed “the god particle” by L. Lederman [1] due to how its interactions with other particles give them their masses. This thesis will attempt to guide you through the steps towards a discovery or exclusion of the Higgs boson. These steps will include motivating the need for the Higgs, describing what sort of detector one needs to find it, outlining an actual search, and discussing the results of this search.

The next chapter will describe the standard model of particle physics, which is a gauge field theory that predicts the existence and behavior of all the fundamental particles. It will then attempt to motivate the need for the Higgs mechanism to complete the model and give the weak gauge bosons and fermions their observed masses. Following this, in Chapter 3, is a brief discussion of particle accelerators and the Tevatron. The D0 detector is also described, which was used for data taking in the Higgs boson search. Chapters 4 and 5 describe the data set collected and the methods used for interpreting it through event reconstruction and modeling. Chapters 6 and 7 contain details about the analysis techniques employed, including the selection of events most likely to contain a Higgs boson and a multivariate technique to help separate those events from the much larger background. Finally, Chapters 8 and 9 focus on the uncertainties in the analysis and the results of the search. Hopefully, as you progress through the reading of this thesis you will gain not only a better understanding of how this specific search was performed, but why we care to do it at all.

Chapter 2

The Standard Model Higgs Boson

2.1 The Standard Model

The standard model of particle physics is a predictive guide describing the known fundamental particles and their interactions. It has been developed over the last 40 years or so as an effective model which aims to explain experimental observations and is not widely believed to be a complete theory valid to all scales. The interactions described in it encompass three of the four basic forces: the electromagnetic, weak, and strong force. 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.

In the standard model, fermions are arranged in families of quarks and leptons. They have spin $\frac{1}{2}$ and interact via the electromagnetic, weak, and strong forces as mentioned above. These forces are carried by the spin $= 1$ gauge bosons. The fermions and bosons are listed in Fig. 2.1. Three types of charged leptons exist: the electron(e), muon(μ), and tau(τ). For each charged lepton there is an uncharged neutrino (ν). There are three quarks bearing an electric charge of $-\frac{1}{3}$ (down(d), strange(s),and bottom(b)) and three with a positive charge of $+\frac{2}{3}$

(up(u), charm(c),and top(t)).

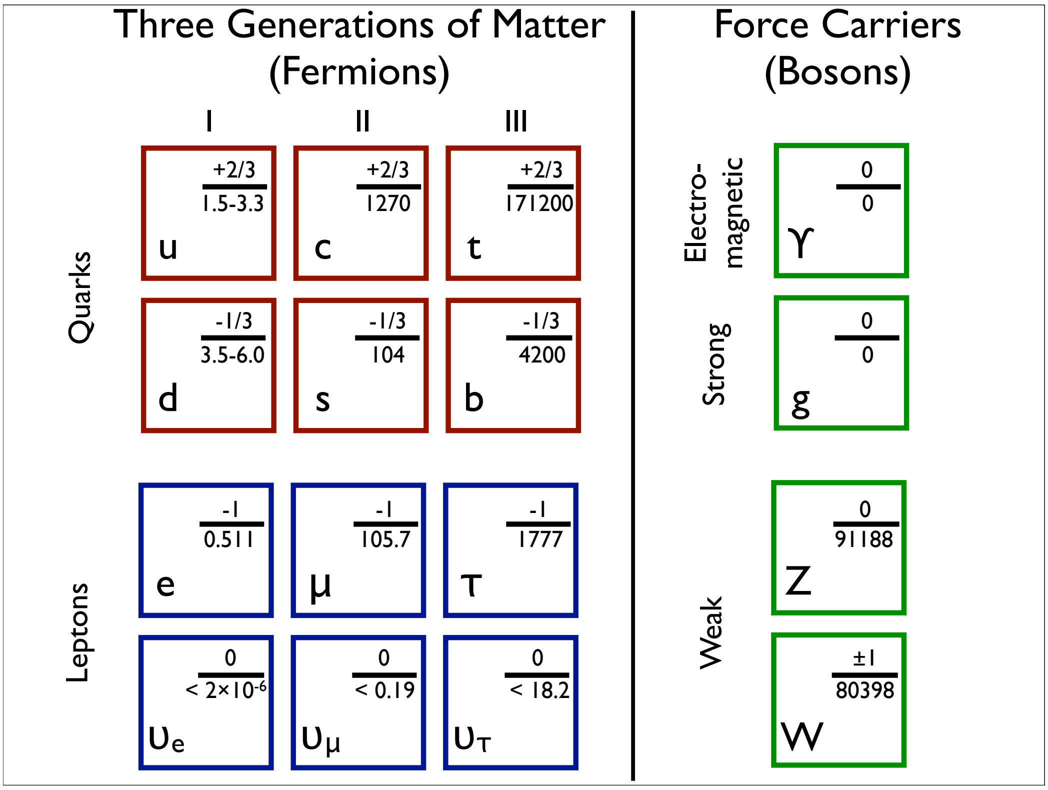

Figure 2.1: The three generations of fermions and the force carrying bosons of the standard model. The number in the top right of each box is the electric charge and the number below that is the mass of the particle in MeV. The masses are from the 2008 version of the Review of Particle Physics [2]

The quarks carry another charge called “color” charge [3]. The three types of color charge are here referred to as blue, green, and red. These color charge types form color triplets. Color charges are carried by quarks and the mediating particles of the strong force, the gluons. The gluon field can turn quarks of one color in the triplet into the same quark with a different color. The gluons form a color-octet and show self-interaction, which leads to an increasing strength of the coupling with increasing distance of the color charges. This makes it impossible to observe free quarks, only colorless states such as mesons ($q\bar{q}$) and baryons (qqq) are observed. This is called *confinement.* For very small distances, the coupling increases and can be calculated using perturbative theory. This is called *asymptotic freedom* [4, 5, 6].

Every fermion has a corresponding anti-fermion with the same mass but with opposite quantum numbers (for example, spin, baryon number, lepton number, etc..). For the neutral neutrinos however, it is not yet clear if neutrinos and anti-neutrinos are different particles or if they are their own anti-particles. In the first case they would be Dirac neutrinos and in the latter, Majorana neutrinos [2]. It appears that all neutrinos are left-handed and all anti-neutrinos show right-handed chirality and this is how they are handled in the standard model. However, the recent observation of neutrino oscillations [7] points to the existence of a neutrino mass. The oscillation parameters have also been measured with man-made neutrino sources [8, 9]. In this case, because chirality and helicity are only the same for massless particles and not for massive ones, massive neutrinos cannot easily be accommodated in the standard model. This is one indication for physics beyond the standard model.

At this point in the discussion of the standard model, it is useful to talk very briefly about group theories and how they relate to symmetries. In general, group theory is a mathematical construct which underlies the treatment of symmetries. A group is a set of objects $a, b, ...$ with a product ab such that

- If a and $b \in \mathrm{G}$, then $ab \in \mathrm{G}$.
- The product is associative: $(ab)c = a(bc)$
- An identity element I exists: $\mathrm{I}a = a\mathrm{I} = a$
- For each element a, an inverse a^{-1} exists: $aa^{-1} = a^{-1}a = \mathrm{I}$

An example of a group is **a set of rotations about an axis**, each described by a matrix $\mathbf{R}$. This group is called SO(2), where the "S" is for "special" (the matrices have determinant = 1), the "O" is for "orthogonal" ($\tilde{\mathbf{R}}\mathbf{R} = 1,\ \tilde{\mathbf{R}} = \mathbf{R}^{-1}$), and the 2 is the dimension n of the $n \times n$ matrix. Since experimental results don't depend on the specific laboratory orientation of the system we are measuring, rotations form a symmetry group of a system.

Another important symmetry group in physics is the group SU(2), which describes **angular momentum** in quantum mechanics. The "U" stands for "unitary", so the matrices satisfy $\mathbf{U}^{\dagger}\mathbf{U} = 1,\ \mathbf{U}^{\dagger} = \mathbf{U}^{-1}$. The 2 comes from the fact that the most fundamental representation is for spin 1/2, which is described by 2×2 matrices (the Pauli spin matrices). These operate in a space with two basis vectors (the spinors $|\uparrow\rangle$, $|\downarrow\rangle$) and all other angular momenta can be constructed from the spin-1/2 case by using the addition rules.

Similar to this, in the standard model the quarks and leptons are grouped into pairs of weak-isospin, T. This is also an SU(2) symmetry, which implies conservation of weak-isospin. The left-handed fermions have a value of $T = \frac{1}{2}$ and can be split into doublets of the axial projection $T_3 = \pm\frac{1}{2}$. For these doublets, the charged weak current can change one member of the pair into the other as long as it has a different value of T_3 (ie. the up-type quarks, u, c, and t with $T_3 = +\frac{1}{2}$ can change into the bottom-type d, s, and b quarks with $T_3 = -\frac{1}{2}$). This current is carried by the $W^{\pm}$ gauge bosons, which have $T = 1$, and only couples to the left-handed fermions. The right-handed fermions have $T = 0$ and are therefore placed in SU(2) singlets that do not interact via the weak force. The left-right asymmetric weak force is said to have $\mathrm{SU(2)_L}$ symmetry and the value of T_3 is conserved in these interactions.

Another symmetry called weak hypercharge(Y) represents the conserved quantum numbers of electric charge, Q, and the axial projection of the weak-isospin, $\mathrm{T_3}$. The weak hypercharge is defined by the equation $\mathrm{Q = T_3 + Y/2}$. This is the $\mathrm{U(1)_Y}$ symmetry.

The electro-weak Lagrangian must be invariant under both of these transformations, $\mathrm{SU(2)_L \times U(1)_Y}$. This leads to an isotriplet of vector fields W^i_ν ($i =$ 1,2,3) coupled with strength g to the isospin. Additionally, a single vector field B_μ couples with strength g' to the weak hypercharge. These two neutral fields mix in such a way that the mass eigenstates we can observe are

$$A_\mu = B_\mu \cos\theta_W + W_\nu^3 \sin\theta_W \qquad \textit{massless} \tag{2.1}$$

and

$$Z_\mu = -B_\mu \sin\theta_W + W_\nu^3 \cos\theta_W \qquad \textit{massive} \tag{2.2}$$

where θ_W is known as the Weinberg or weak mixing angle. It is measured to be $\sin^2\theta_W = 0.231$ [2]. The variables A_μ and Z_μ can be identified with the photon and Z boson, respectively. The two remaining fields, W_μ^1 and W_μ^2, also mix to form the charged gauge boson, $W^\pm$. The decay width of the Z boson to a fermion-antifermion pair is related to the weak mixing angle and provides a method for determining it by measuring the various Z branching ratios.

The strong interaction is invariant under rotations in the color space described earlier. This is an SU(3) symmetry called $\mathrm{SU(3)}_{\mathrm{color}}$. The three color charges form the fundamental representation of the group[1].

The standard model holds up extremely well to experimental tests and is widely used as an example of a successful theoretical structure. Except for the Higgs mass, all the parameters of the SM can be (and have been) determined experimentally. This includes the three gauge coupling constants, the masses of the weak vector bosons and fermions, and the quark mixing angles. The theory, once one calculates the complicated higher-order terms of the perturbative series and the radiative corrections, allows the prediction of any measurable with a high degree of accuracy. These two things combined have lead to the collection of large amounts of electroweak precision data at the e^+e^-, LEP and SLC colliders as well as the $p\bar{p}$ Tevatron collider, which built upon earlier experimental programs. In addition to the precision measurements made, a large theoretical effort has

---

[1] There are 8 generators of this group, denoted λ_i. The first three of these, λ_1, λ_2, λ_3, correspond to the three Pauli-spin matrices and represent an SU(2) subgroup of SU(3). λ_i are known as the Gell-Mann matrices.

been devoted to the calculation of the radiative corrections to the electroweak observables, to match the accuracies which have been or which could be reached experimentally. [10]

There is an impressive list of electroweak observables which have been measured with very good accuracy and can be predicted in the SM with equally good precision. Some of these observables related to the Z boson are its mass m_Z, the total width Γ_Z, and the peak hadronic cross section σ^0_{had}. For the W boson the mass m_W and the total decay width Γ_W are measured. Also included is the effective weak mixing angle, $\sin^2\theta^{lep}_{eff}$, as measured from lepton asymmetries and the mass of the top quark m_t. The experimental values of these electroweak observables along with several others are shown in Fig. 2.2 with their associated errors. The theoretical predictions of the SM are also shown. They were obtained by including all known radiative corrections and using the central values of σ^0_{had}, m_t, α_s, etc..

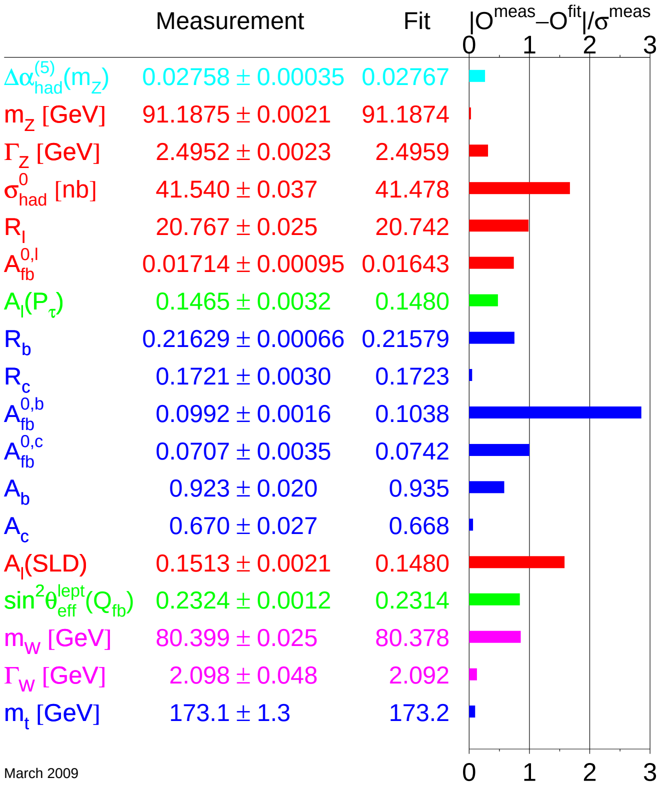

Figure 2.2: Summary of electroweak precision measurements at LEP1, LEP2, SLC and the Tevatron [11]. The SM fit results, which have been derived including all radiative corrections, and the standard deviations are also shown.

As one can see there is remarkable agreement between the theoretical predictions and the actual measurements. Programs like GSM (a Gfitter package [12]) are used to make these comparisons with the SM. A set of input parameters are required, including the mass and decay width of the W boson, the weak mixing angle, the peak hadronic cross section, the strong coupling constant, the mass of the top quark, and many other observables. Then several of these parameters are allowed to float when trying to get a best fit, like the masses of the quarks and the Z boson.

The high level of agreement seen has lead to the common requirement that any new theory must, like the SM, be a renormalizable gauge invariant field theory. To explore this requirement of gauge (phase) invariance imposed on the Lagrangian density (generally just referred to as the Lagrangian) it is useful to consider the effect of a position-dependent phase change of the form

$$\psi(x) \rightarrow \psi'(x) = e^{-i\alpha(x)}\psi(x), \tag{2.3}$$

where ψ is the complex field describing a free Dirac particle and $\alpha(x)$ is a phase with an arbitrary dependence on space and time. The Lagrangian of this particle is written as

$$\mathcal{L} = i\bar{\psi}\gamma^{\mu}\partial_{\mu}\psi - m\bar{\psi}\psi, \tag{2.4}$$

where γ^{μ} are the standard 4×4 Dirac “gamma matrices” and the super/subscript μ’s are an implied summation ($\mu = 0,\ 1,\ 2,\ 3$), then this “local” gauge invariance requirement fails. If the invariance is not upheld, the consequence would be to make $\alpha(x)$ a measurable relation rather than an arbitrary phase, implying that the fields described behaved in certain ways given a specific location in space and time. While the last term is certainly invariant under the transformation in 2.3, the $\partial_{\mu}\psi$ term is not. To fix this problem one can define the “covariant” derivative, D_{μ}, such that additional terms introduced by 2.3 cancel. To do this it is necessary to introduce a vector field, A_μ. The covariant derivative is then

$$D_\mu \equiv \partial_\mu - ieA_\mu. \tag{2.5}$$

This new gauge field is coupled to the Dirac particle (of charge -e) in the same way as the photon field should be and in fact can be interpreted as such, provided we add the necessary kinetic energy term to the Lagrangian. It should be a term which does not interfere with the invariance we started out pursuing, and thus can only depend on the invariant field strength tensor

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu. \tag{2.6}$$

This makes the final Lagrangian look like

$$\mathcal{L} = \bar{\psi}(i\gamma^\mu \partial_\mu - m)\psi + e\bar{\psi}\gamma^\mu A_\mu \psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}, \tag{2.7}$$

which is the Lagrangian of QED [13]. If we try to add an additional mass term($\frac{1}{2}m^2 A_\mu A^\mu$) to this Lagrangian the invariance is not retained. So the gauge particle, in this case the photon, must be massless.

Symmetries of the Lagrangian imply conservation laws, as stated by Noether's theorem [14]. An example of this is the translational invariance of a classical particle which leads to conservation of momentum. The gauge invariance of Eq. 2.7 suggests a conserved current which leads to conservation of charge. The invariance described above is the U(1) symmetry. One can also attempt to continue this technique to get the Lagrangian for the properties and interactions of the rest of the known fundamental particles described earlier.

The full Lagrangian for all of the standard model particles and their interactions must obey the product of all of the above mentioned symmetries

$$\mathrm{SU(3)_{color}} \otimes \mathrm{SU(2)_L} \otimes \mathrm{U(1)_Y}. \tag{2.8}$$

For this to hold true, it can be shown that the weak gauge bosons must be massless. For a more detailed discussion, see for example, Halzen and Martin [15]. Specifically, to get a massive weak vector boson there would have to be a mass term of the form $m^2 B_\mu B^\mu$, where B^μ is the vector field coupled to the weak hypercharge discussed earlier. This term would not remain unchanged under a local $\mathrm{SU(2)_L}$ rotation ($e^{-\frac{i}{2}\tau^i \alpha^i(x)}$). In a similar fashion, to get massive fermions one would need a term like $-m\bar{\ell}_L e_R$, where $\bar{\ell}_L$ is the left-handed doublet and e_R is the right-handed singlet. Again, this would break the $\mathrm{SU(2)_L}$ invariance. This is obviously not the case, since we observe masses for these particles. The method used to help alleviate this problem is called the Higgs Mechanism.

2.2 The Higgs Mechanism

To tackle the problem of massive weak bosons, we need only address the electroweak sector, as the strong force does not enter the picture (that part of the Lagrangian is already gauge invariant and has a massless force carrier). This means we should concern ourselves with maintaining the electroweak symmetry, $\mathrm{SU(2)_L} \otimes \mathrm{U(1)_Y}$. Following the prescription in Halzen and Martin [15], we see that a complex SU(2) doublet of scalar fields can be introduced such that the photon can remain massless while the $W^\pm$ and Z have mass. The doublet has weak hypercharge, Y = +1 and takes the form

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix} \tag{2.9}$$

The new Lagrangian terms for this are a kinetic term involving the interactions of the scalar field with the gauge bosons, fermions, and itself along with a potential energy term like

$$V(\phi^\dagger\phi) = \mu^2(\phi^\dagger\phi) + \lambda(\phi^\dagger\phi)^2 \tag{2.10}$$

The full Higgs Lagrangian takes the form

$$\mathcal{L}_{\text{Higgs}} = \tag{2.11}$$

$$\begin{array}{ll}
(D_\mu\phi)^\dagger(D_\mu\phi) & \{\mathrm{Higgs - gauge - boson\ coupling} \\
 & \\
-\mu^2(\phi^\dagger\phi) - \lambda(\phi^\dagger\phi)^2 & \{\mathrm{Higgs\ potential:} \\
 & \quad \mathrm{Mass\ term\ and\ quartic\ coupling} \\
+k_e(\bar{\ell_L}\phi)e_R + k_e\bar{e_R}(\phi^\dagger\ell_L) & \{\mathrm{Lepton - Higgs\ Yukawa\ terms} \\
 & \\
+k_d(\bar{d^c_L}\phi)d^c_R + k_d\bar{d^c_R}(\phi^\dagger d^c_L) & \{\mathrm{Quark - Higgs} \\
+k_u(\bar{u^c_L}\tilde{\phi})u^c_R + k_u\bar{u^c_R}(\tilde{\phi}^\dagger u^c_L) & \qquad \mathrm{Yukawa\ terms}
\end{array}$$

where ϕ is the Higgs field, μ is the mass term, λ is the Higgs quartic coupling constant, and the arbitrary parameters, k, are the Yukawa coupling constants for each lepton and quark species. A cartoon of the Higgs potential can be seen in Fig. 2.3. It can have non-zero minima if μ^2 is less than zero, effectively hiding a symmetry of the system that was apparent for a minimum at zero. When at one of these non-zero minimum values the system is in a ground state, where the potential is at its lowest energy. There is an infinite number of equivalent ground states described by the minima

$${\phi_1}^2 + {\phi_2}^2 + {\phi_3}^2 + {\phi_4}^2 = \frac{-\mu^2}{2\lambda}, \tag{2.12}$$

and nature picks one at random. We are free to rotate the fields such that $\phi_1 = \phi_2 = \phi_4 = 0$ and ϕ becomes

$$\phi = e^{i\zeta^i \tau^i/2v} \begin{pmatrix} 0 \\ \frac{h(x)+v}{\sqrt{2}} \end{pmatrix} \tag{2.13}$$

where ζ^i are the angles representing the degrees of freedom from ϕ_1, ϕ_2, and ϕ_4. The vacuum expectation value, v, is $v = \sqrt{-\mu^2/\lambda} \approx 250GeV$. This can now be put back in the original Lagrangian (Eq. 2.11) and is in all ways equivalent, but with fewer rotational degrees of freedom. This random choice of a specific direction by nature is known as "spontaneous symmetry breaking".

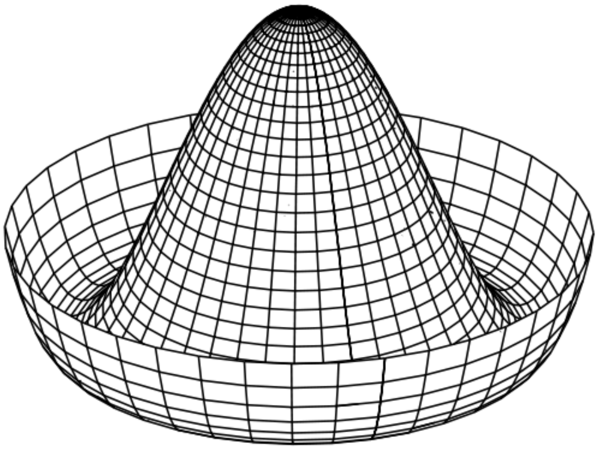

Figure 2.3: Cartoon of the Higgs potential.

The Higgs scalar field we have chosen allows the Lagrangian to have Higgs-fermion and Higgs-weak-boson interactions and since the vaccum expectation value is non-zero, they also have masses. While doing so, it also leaves the photon massless. So, the Higgs mechanism allows us to generate the masses of the weak bosons without spoiling the renormalizability of the electroweak gauge theory.

We can also look back at the weak mixing angle θ_W mentioned in section 2.1, with the relation $g' = g \ \tan\theta_W$, where, from the electroweak gauge group, g' is the weak isoscalar coupling constant and g is the weak isovector coupling constant [16]. This allows us to identify the photon's coupling to charged particles, $gg'/\sqrt{g^2 + g'^2}$, with the electric charge e. The W boson then gets a mass $M_W = gv/2 = ev/2\sin\theta_W$ and the Z boson, $M_Z = M_W/\cos\theta_W$. Since M_W and M_Z can be measured and v determined from M_W, these predictions can be tested. It turns out the agreement is very good (see Fig. 2.2). The masses of the fermions, on the other hand, are not predicted by the Higgs mechanism. These masses stem from the Yukawa terms in the Lagrangian (Eq. 2.11) and the trilinear couplings associated with them are unfortunately completely arbitrary. The only way to determine them would be to actually observe a Higgs boson and measure the effects of the Higgs couplings with each fermion. The masses for the fermions are of the form $k_X v/\sqrt{2}$, where k_X is the arbitrary Yukawa coupling constant representing the coupling strength to the Higgs scalar. Additionally, to reproduce the observed quark and lepton masses, the k_X values must range over many orders of magnitude. At present we do not know what sets these Yukawa coupling values.

While the Higgs boson just described is predicted by the standard model (if we are to maintain renormalizability and gauge invariance), it has not been observed. The search for this particle and a measurement of its mass has become a priority for many scientists in both the theoretical and experimental realms.

2.3 Restrictions on the Higgs Boson Mass

2.3.1 Theoretical Constraints

There are several theoretical constraints on the Higgs boson mass which can be derived from assumptions on the energy range in which the SM is valid before perturbation theory fails and new phenomena are expected to emerge. These include constraints from unitarity in scattering amplitudes, stability of the electroweak vacuum, and finetuning. The following sections briefly describe these constraints. More detailed descriptions can be found in Ref. [17].

(Pertubative) Unitarity

One constraint on the Higgs mass is obtained by studying scattering processes at very high energies. For example, considering the scattering of longitudinal gauge bosons $W_L^+ W_L^- \rightarrow W_L^+ W_L^-$ and applying unitarity requirements for center-of-mass energy $s >> m^2$, one obtains:

$$m_H \lesssim 870 \text{ GeV} \tag{2.14}$$

and if coupled channels like $W_L^+ W_L^- \rightarrow Z_L Z_L$ are considered the bound can be lowered to

$$m_H \lesssim 710 \text{ GeV} \tag{2.15}$$

So, in the SM, if the Higgs boson mass exceeds values of $\sim$ 700 GeV, unitarity will be violated unless new phenomena appear and restore it.

Alternatively, if we assume $m_H >> s >> m_W$ we get a critical scale $\sqrt{s_C}$ at which perturbative unitarity must be violated. For the $W_L^+ W_L^- \rightarrow W_L^+ W_L^-$ only case this scale is

$$\sqrt{s} \lesssim 1.7 \text{ TeV} \tag{2.16}$$

and using more constraining channels this bound can be lowered to

$$\sqrt{s} \lesssim 1.2 \text{ TeV} \tag{2.17}$$

In other words, some new physics beyond the SM should manifest itself at energies in the TeV range to restore unitarity in the scattering amplitudes of longitudinal gauge bosons.

Triviality and Stability

The couplings and masses appearing in the SM Lagrangian can vary with the energy being considered due to quantum corrections. The quartic Higgs coupling is no exception and increases monotonically with the energy scale, $|Q|$. This variation with Q, taking into account only the Higgs self-coupling[2], is described by the Renormalization Group Equation (RGE) [17]:

$$\frac{\mathrm{d}}{\mathrm{d}Q^2}\lambda\ (Q^2) = \frac{3}{4\pi^2}\lambda^2(Q^2) + \text{higher orders} \tag{2.18}$$

Solving this equation, choosing the electroweak symmetry breaking scale $Q_0 = v$ as the reference energy point gives (up to one-loop):

$$\lambda\ (Q^2) = \lambda\ (v^2)\left[1 - \frac{3}{4\pi^2}\lambda(v^2)\log\frac{Q^2}{v^2}\right] \tag{2.19}$$

In the case $Q^2 << v^2$, where the energy is much smaller than the electroweak breaking scale, the quartic coupling becomes extremely small and eventually vanishes. The theory is then deemed "trivial" or non-interacting, since the coupling is zero. Conversely, when $Q^2 >> v^2$ the coupling becomes infinite. This point is called the Landau pole and occurs at [18]:

$$\Lambda_C = v\ \ \exp\left(\frac{4\pi^2}{3\lambda}\right) = v\ \ \exp\left(\frac{4\pi^2 v^2}{M_H^2}\right) \tag{2.20}$$

One can use this to determine the energy cutoff Λ_C below which the self-coupling λ remains finite. Looking at Eq. 2.20, one can see that if Λ_C is large, the Higgs mass must be small to avoid the Landau pole. For example, for $\Lambda_C \sim 10^{16}$ GeV the Higgs boson mass would need to be $M_H \sim 200$ GeV. If Λ_C is small, the mass can be larger. For $\Lambda_C \sim 10^3$ GeV, the Higgs mass is allowed to be of the order of 1 TeV. If the cutoff is set at the Higgs boson mass itself, $\Lambda_C = M_H$, this would

---

[2] This is valid for larger λ.

imply $M_H \sim 700$ GeV. [17]

If one also takes into account fermions and gauge bosons the RGE becomes more complex. Since the Higgs boson couplings are proportional to the particle masses, only the contribution of top quarks and massive gauge bosons need to be considered. Aside from altering the triviality bounds discussed above, the top quark Yukawa coupling in particular can have a large impact for lower values of λ. Its contribution to the solution of the inclusive RGE is negative and becomes dominant in this case, eventually leading to a negative solution and thus a scalar potential $V(Q^2) < V(v)$. The vacuum is not stable anymore since it has no minimum. This is the stability argument and puts strong constraints on the Higgs mass depending on the cut-off Λ_C. The triviality bound and the vacuum stability bound on the Higgs boson mass as a function of the New Physics or cutoff scale Λ_C can be seen in Fig. 2.4.

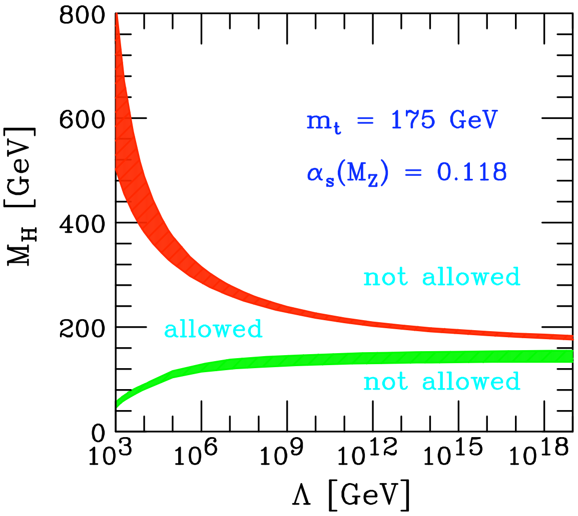

Figure 2.4: The triviality (upper) bound and the vacuum stability (lower) bound on the Higgs boson mass as a function of the New Physics or cutoff scale Λ. The allowed region lies between the bands and the colored/shaded bands illustrate the impact of various uncertainties. [19]

Fine-tuning Constraint

Fine-tuning refers to circumstances when the parameters of a model must be adjusted very precisely in order to agree with observations. Theories requiring fine-tuning are regarded as problematic in the absence of a known mechanism to explain why the parameters happen to have precisely the needed values. In our case, the fine-tuning problem originates from the radiative corrections to the Higgs boson mass. The one-loop radiative corrections involve Higgs boson, massive gauge boson, and fermion loops as seen in Fig. 2.5.

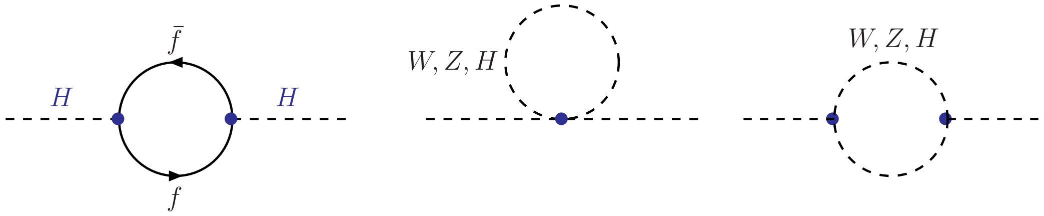

Figure 2.5: Feynman diagrams for the one-loop corrections to the SM Higgs boson mass. [17]

If we cut off the loop integral momenta at a scale Λ that is very large, say 10^{16} GeV [3], one needs a very fine arrangement of the numbers between the Higgs mass and the radiative corrections. This is to have a physical Higgs boson mass in the range of the electroweak symmetry breaking scale, around $M_H \sim 100$ GeV to 1 TeV, which the SM requires for consistency.

For scales, Λ, not much larger than the electroweak scale however, one can consider the Veltman solution [20] which relates the Higgs boson mass to the top quark, W boson, and Z boson masses. Using the relation cancels the quadratic divergences and would even predict a Higgs boson mass. This allows the scale for which fine-tuning is necessary to be raised from something like ~ 2 TeV to 15 TeV in the two-loop, or even 50 TeV in the three-loop case.

The predicted Higgs mass is cut-off dependent. If we define the amount of

---

[3] This is the Grand Unification scale.

fine-tuning as the sensitivity of the electroweak scale to the cutoff point, this leads to the measure Δ_{FT}, where the weak scale is fine-tuned to one part in Δ_{FT}. A higher Δ_{FT} means more fine-tuning and none is necessary for $\Delta_{FT} \leq 1$. In Fig. 2.6, one can see the Higgs boson mass as a function of the cutoff scale Λ. The regions of fine-tuning less than 10 and 100 are shown. The white region corresponds to the one where all constraints are fulfilled and where the Veltman condition is approximately satisfied. It is apparent that for scales $\lesssim$ 1 TeV there is no finetuning problem for any reasonable Higgs boson mass value. At higher scales a Higgs boson mass $\sim$ 200 GeV still allows for an acceptable amount of fine-tuning.

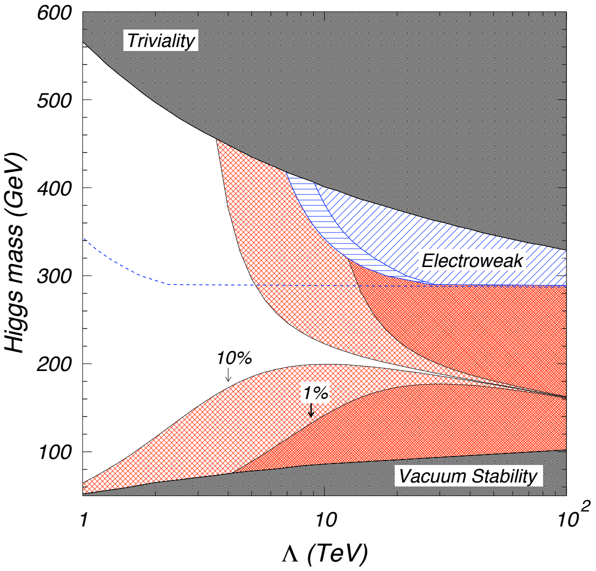

Figure 2.6: The contours for the finetuning parameter Δ_{FT} in the plane $(M_H,\ \Lambda)$. The dark (light) hatched region marked 1% (10%) represents fine-tunings of greater than 1 part in 100 (10). The constraints from triviality, stability and electroweak precision data are also shown. The empty region is consistent with all constraints and has Δ_{FT} less than 10%. [21]

Fig. 2.6 summarizes the theoretical constraints on the Higgs mass as a function of the cut-off scale Λ.

2.3.2 Previous Higgs Boson Searches

Direct Search

The most restrictive limit on the allowed mass of the Higgs boson was set at the LEP collider at CERN [22]. LEP was an electron-positron collider that operated from 1989 to 2000. The production mode for the Higgs boson at LEP was associated production with a Z boson. The primary search channel was the Higgs decaying to 4 jets, where the Z also decayed to jets. Main backgrounds for this channel were multijet events and diboson (WW and ZZ) events.

After a successful detector upgrade, the final LEP2 limit on the Higgs boson mass reached $m_H > 114.4 GeV$ at 95% confidence level (C.L.) [23]. This result stood as the last experimental constraint until the Summer of 2008 when, combining the results from the two Tevatron experiments CDF and D0, it was possible to obtain sensitivity to the standard model Higgs boson in the region not yet excluded by LEP [24].

Indirect Constraints

In addition to doing direct searches for the Higgs boson, one could ask what other observables could be affected by its existence. The Higgs boson would have a significant impact on some electroweak parameters. For example, the Higgs boson would contribute to radiative corrections on the top quark and W boson masses. This means that precision measurements of electroweak parameters can be used to obtain constraints on the Higgs boson mass indirectly. The 1-loop effects of the top quark and Higgs boson on the W boson mass are shown in Fig. 2.7.

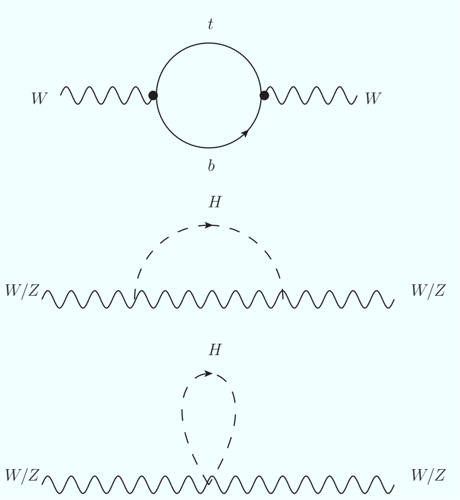

Figure 2.7: The masses of top quark, W and Higgs boson are related via radiative loop corrections. [25]

The three main electroweak observables relevant to the Higgs boson are the electroweak ρ-parameter, the effective leptonic weak mixing angle, and the W boson mass. The ρ-parameter is defined as

$$\rho = \frac{{m_W}^2}{{m_Z}^2}(1 - \sin^2\theta_W). \tag{2.21}$$

One can deduce the Higgs boson mass from precision measurements of the top quark and W boson mass, assuming the standard model is correct. This is achieved by looking at the radiative correction, Δr, which depends on the mass of the top quark and the ratio of the W mass to Higgs mass

$$\Delta\, r = \frac{3G_F}{8\pi^2\sqrt{2}} m_t^2 + \frac{\sqrt{2}G_F}{16\pi^2} m_t^2 \left[\frac{11}{3} \ln\left(\frac{m_H^2}{m_W^2}\right) \; + \; ...\right] \; + \; ... \tag{2.22}$$

If these loop corrections did not exist, as is approximated at the tree-level, then the calculated value for the weak mixing angle would be the one for Δr = 0. One can then do this check and see that indeed there is evidence that these loop corrections must exist to get the value we measure (including the Higgs loop corrections!). The top quark mass and W mass have been measured in several experiments at

LEP, SLC and Tevatron as shown in Fig. 2.8.

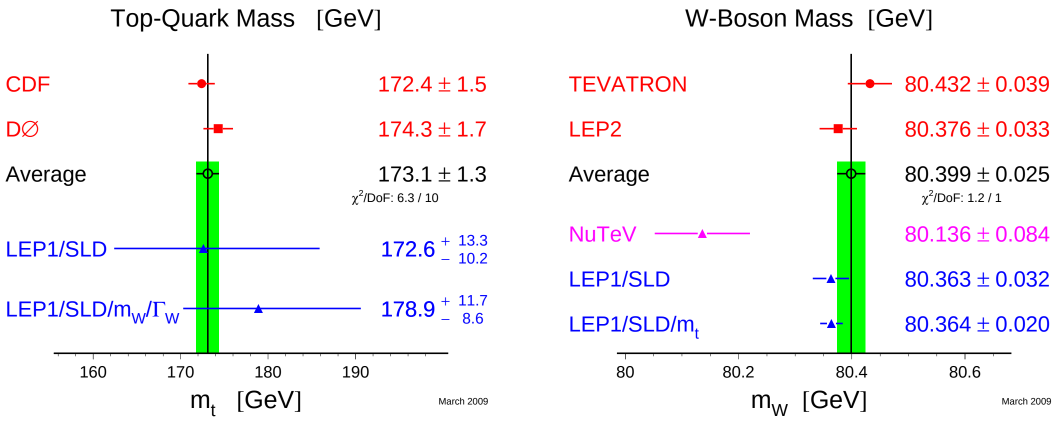

Figure 2.8: World average and individual measurements of the top (left) and W mass (right) including both direct measurements of the top mass as well as indirect measurements via higher order corrections to the W boson mass. [11]

The combined result for both indirect constraints and direct measurements of the W mass and top quark mass is shown in Fig. 2.9. Increased precision in the measured mass of the top quark and W boson have further reduced the uncertainty on these constraints over the last few years. 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.

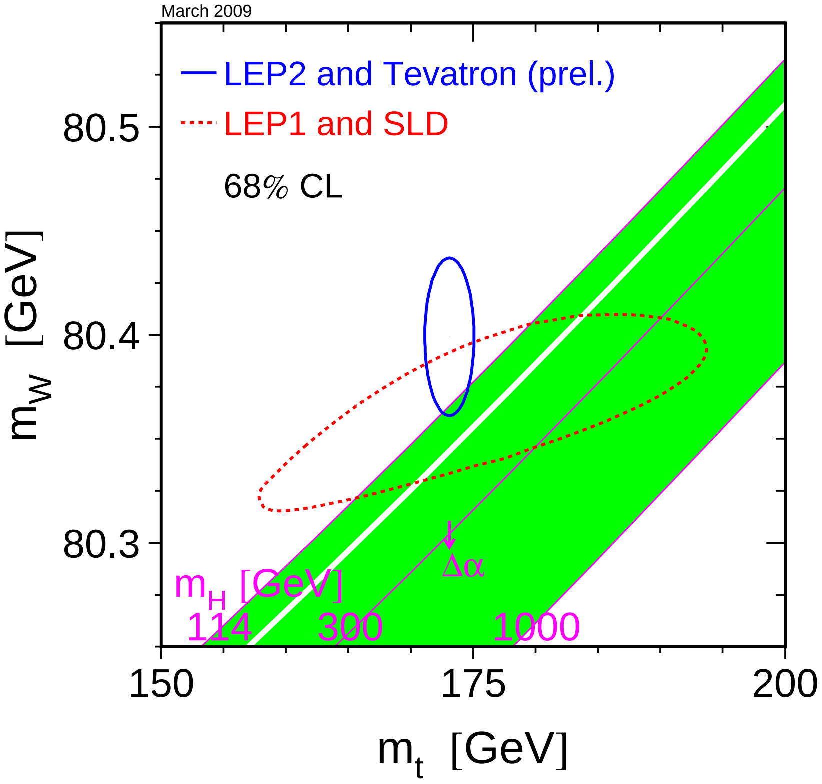

Figure 2.9: The comparison of the indirect constraints on m_W and m_t based on LEP-I/SLD data (dashed contour) and the direct measurements from the LEP-II/Tevatron experiments (solid contour). Also shown is the SM relationship for the masses, as a function of the Higgs mass, in the region favored by theory and allowed by direct searches. The arrow labelled $\Delta\alpha$ shows the variation of this relation if $\alpha(m_Z^2)$ is changed by $\pm 1\sigma$. This variation gives an additional uncertainty to the SM band shown in the figure. [10]

Fig. 2.10 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 LEP-II [23] and the Tevatron [24]. The dotted curve shows the corresponding fits including the low-Q^2 data. The dashed curve is the result obtained using the evaluation of $\Delta\alpha^{(5)}_{had}(m_Z^2)$ from [26].

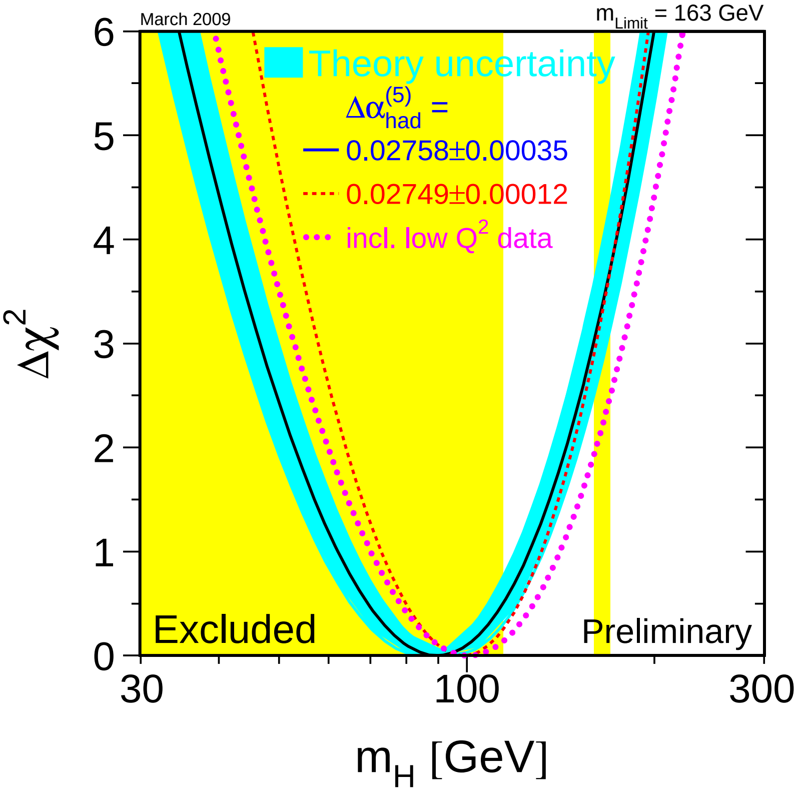

Figure 2.10: The $\Delta\chi^2 = \chi^2 - \chi^2_{min}$ vs. m_H curve. The line is the result of the fit using all high-Q^2 data (last column of Table 2); the band represents an estimate of the theoretical error due to missing higher order corrections. The vertical band shows the 95% CL exclusion limit on m_H from the direct searches at LEP-II and the Tevatron. The dashed curve is the result obtained using the evaluation of $\Delta\alpha^{(5)}_{had}(m_Z^2)$ from [26]. The dotted curve corresponds to a fit that includes the low-Q^2 data. [10]

It is interesting to note that these figures show a prefered lower-mass Higgs in contrast to the higher masses allowed by other constraints. Combining indirect constraints and direct searches, the constraints on the Higgs boson mass summarize to $114.4 < m_H < 191(GeV)$ at 95% C.L.

2.4 Higgs Boson Production and Decay at the Tevatron

2.4.1 Higgs Production

The dominant production processes for Higgs bosons in $p\bar{p}$ collisions at the Tevatron are gluon-gluon fusion, vector-boson fusion, and associated production with either a W or Z vector-boson. The Feynman diagrams of these processes are shown in Fig. 2.11 and 2.12. The cross sections of these processes at the Tevatron are shown in Fig. 2.13. The gluon fusion production (gg → h) 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 (for $m_H = 160$ GeV, $\sigma = 0.431$ pb [27]). This is followed by the WH associated production with a cross section about ten times smaller. The ZH and vector-boson fusion (VBF) production processes follow with even smaller cross sections. All of these production modes are considered in this analysis.

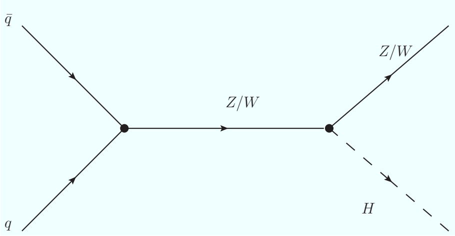

Figure 2.11: Feynman diagram of the associated production of a Higgs boson with a massive vector boson.

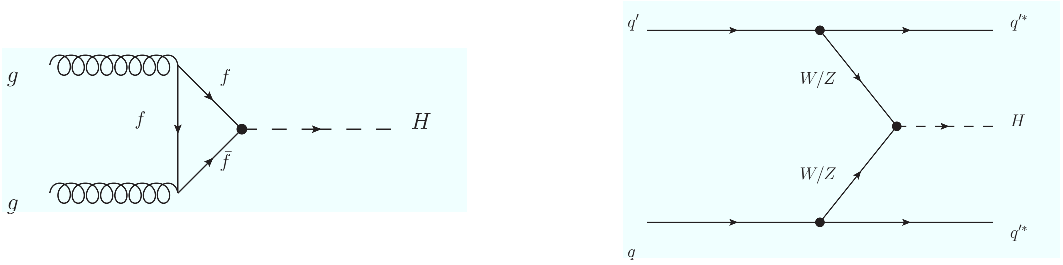

Figure 2.12: Feynman diagrams of Higgs boson production via gluon fusion(left) and vector boson fusion (right).

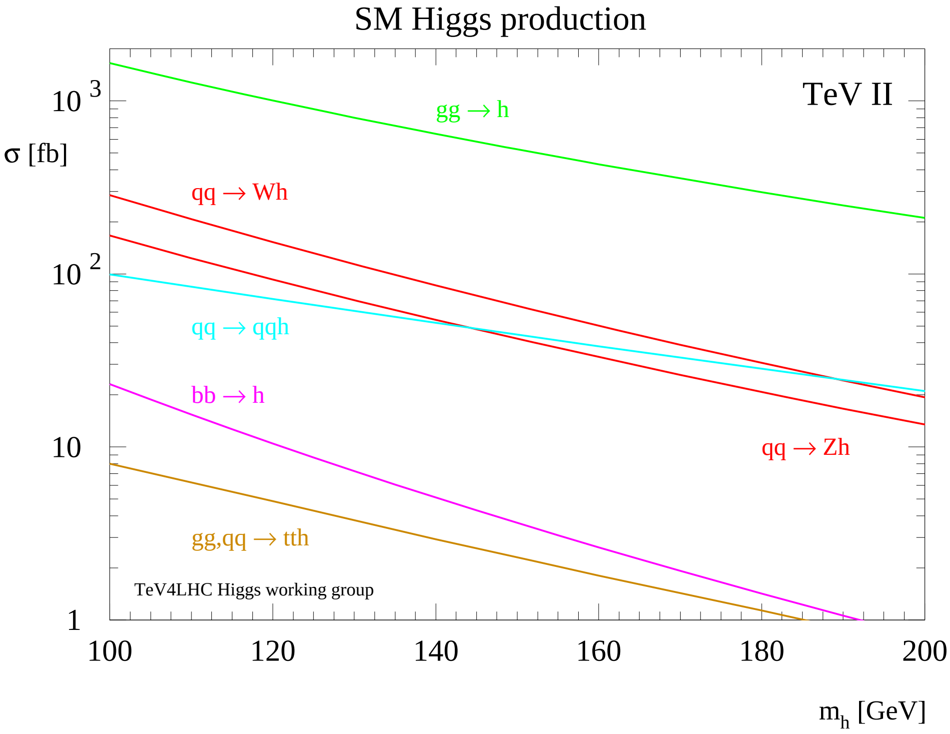

Figure 2.13: The main production modes for a standard model Higgs boson at the Tevatron [27]. The four leading production mechanisms are considered in this analysis. They are gluon fusion, associated production with a massive gauge boson, and vector boson fusion.

2.4.2 Higgs Decay

The branching ratios for various decay modes of the Higgs boson depend on its mass, as shown in Fig. 2.14. Once this mass is set however, all of the couplings to the fermions and weak vector bosons are determined. This mass dependence must be considered along with the production processes described in the previous section when designing an analysis. For small masses, below $m_H \simeq 140 GeV$ the Higgs boson decays primarily into $b\bar{b}$ pairs. For larger masses the decay to W^+W^- becomes dominant. The Higgs boson decay branching ratio predictions are calculated with HDECAY [28].

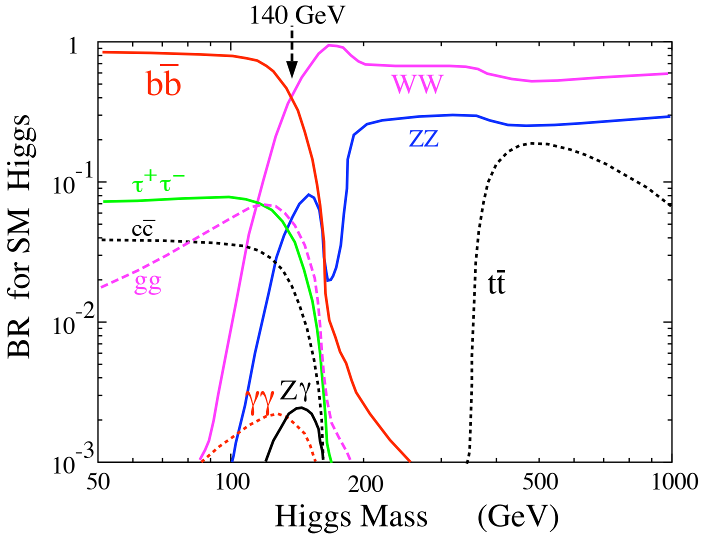

Figure 2.14: Branching ratios of the Higgs boson as a function of the Higgs boson mass. [28]

Given the production cross sections as shown in Fig. 2.13, several thousand Higgs boson events are expected to be produced per detector at the Tevatron with the current luminosity assuming a Higgs mass around 160 GeV. Things like detector acceptance trigger efficiencies, and selection efficiencies lead to the rejection of some of these events. Also, not all channels can be utilized for all possible Higgs boson masses. For example, at low Higgs masses the dominant gluon-fusion channels cannot be easily distinguished from the large ($\sim 10\mu$b) $b\bar{b}$ production. This means that the associated production channels (ZH and WH) become the most useful channels given the much smaller $\sim$ 30 pb production cross section of the W$b\bar{b}$ background. For higher Higgs masses ($m_H > 140 GeV$) the background from multijet production can be reduced by using the leptonic decay $H \rightarrow W^+W^- \rightarrow l^+\nu l^-\nu$, making this the most promising channel. This channel is especially sensitive to Higgs masses $m_H \sim 160$ GeV where the branching ratio to

WW is maximal (~ 0.9011 from Fig. 2.14 at $m_H = 160$ GeV). The most important search channels for $H \rightarrow W^+W^-$ involve decays in the ee$\nu\nu$, e$\mu\nu\nu$ and $\mu\mu\nu\nu$ final states. Contributions of the tau branching modes such as $H \rightarrow W^+W^- \rightarrow$ e$\tau\nu\nu$, $\mu\tau\nu\nu$, and $\tau\tau\nu\nu$ can also be retained to gain some sensitivity. This analysis concentrates on the $\mu\mu\nu\nu$ final state while also considering tau events that make the analysis cuts described in Sec. 6.1. Fig. 2.15 shows the Feynman diagram for the gluon fusion production mode decaying to the $\mu\mu\nu\nu$ final state.

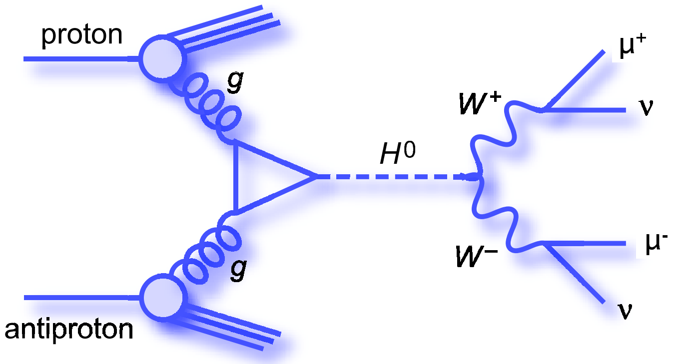

Figure 2.15: Feynman diagram of the gluon fusion production mode decaying to two muons and two neutrinos.

As a quick back-of-the-envelope calculation we can determine N, the number of expected Higgs events from the main production mode (gluon fusion) and our decay mode of interest ($H \rightarrow W^+W^- \rightarrow \mu^+\nu\mu^-\nu$). We can use $N = (\sigma \times BR) \times \epsilon \times L$ for this calculation, where $\sigma \times BR$ is the cross section times branching ratio of the production and decay modes, ϵ is the total efficiency, and L is the total integrated luminosity we analyze. If we assume a perfect world with 100% efficiency (perfect detector, etc..) we can use the numbers mentioned in sections 2.4.1 and 2.4.2 for a Higgs mass of 160 GeV, and an integrated luminosity of 4.2 fb^{-1}(see section 5.1) to get

$$N = (431\ fb \times 0.901 (H \rightarrow WW) \times (0.106)^2 (W \rightarrow \mu\nu) \times 4.2\ \mathrm{fb}^{-1} \simeq 18 \quad (2.23)$$

expected Higgs events. This helps give us some feeling for what we are shooting at in the rest of the analysis. As we will later see (Chapter 6), there are some 55 million events in which to find these < 18 events, a daunting task to be sure.

Chapter 3

Apparatus

The goal of most particle physics experiments is to test the standard model, either by measuring properties of fundamental particles more precisely or by searching for new particles. To do this, one must be able to reproduce the conditions of the very early universe when the average temperature was around 10^{14} K. This amazing feat requires a large experimental setup involving a system to accelerate and store particles for collision (a collider) and one or more detectors which will identify and measure the particles produced in these collisions. It also requires a sufficient number of these high energy collisions to make precision measurements and calculate meaningful probabilities. To create these high energy collisions it is necessary to produce very energetic beams of protons and antiprotons whose kinetic energy can be used to yield the desired energies. The setup would allow for the study of the properties of well established particles like quarks, gluons, and leptons as well as the search for new particles.

Just such an experimental setup exists at Fermi National Accelerator Laboratory(Fermilab) in Batavia, IL [29, 30, 31, 32, 33]. The Tevatron, located on site at Fermilab, is a particle accelerator that collides particles at sufficient energies and rates to probe the standard model. All of the analyzed data considered in this thesis were collected by the D0 detector located on the collision ring of the Tevatron.

Founded in November 1967, Fermilab has been home to many important discoveries in high energy physics. Among them, the discovery of the bottom quark via the Upsilon peak in 1977 [34], the observation of the top quark in 1995 [35, 36], and the tau neutrino in 2000 [37].

3.1 The Tevatron

The Tevatron is a particle accelerator that collides protons and antiprotons at a center of mass energy of 1.96 TeV. Creating these collisions is a long and complicated process which starts with the production of the protons and antiprotons and ends with collisions at the center of the two main detectors, CDF and D0. A schematic of the Tevatron accelerator chain can be seen in Fig. 3.1.

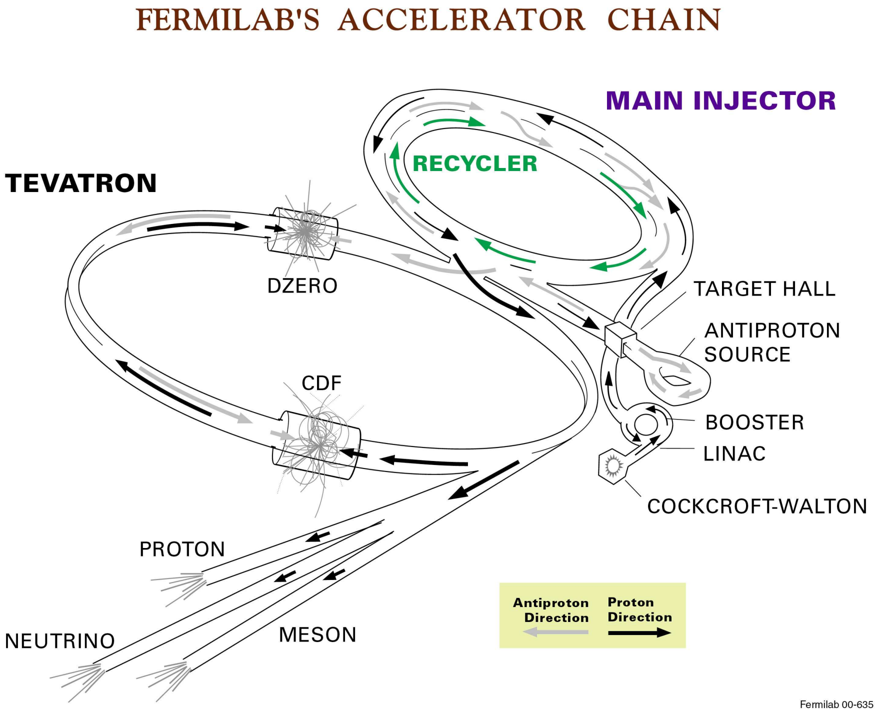

Figure 3.1: Schematic of the Fermilab accelerator chain.

Protons begin their life in the accelerator chain as hydrogen gas (H_2). A magnetron [38] then converts this H_2 into H^- hydrogen ions. The magnetron consists of an oval-shaped cylindrical cathode surrounded by an anode with holes at the bottom and top for introducing the H_2 and extracting the H^- ions. The apparatus is subjected to a magnetic field oriented along the axis of the cylinder. When the hydrogen gas is injected into the bottom hole, a potential difference of a few hundred volts is pulsed in approximately 80 μs pulses at a rate of 15 Hz. During a pulse, the non-varying magnetic and electric fields produce a plasma while confining the electrons to spiral within the anode-cathode gap (about 1 mm in width). The protons bend much less in the magnetic field and hit the cathode, occasionally picking up two electrons and becoming H ions. They are then accelerated towards the anode and the hole in the top of the anode allows some of the H^- ions to escape and accelerate to the extractor electrode.

At the next stage, a Cockcroft-Walton generator is used to accelerate the H^- ions to 750 keV. This device is a simple diode voltage multiplier which generates high voltage by charging capacitors in parallel from an AC voltage source and discharging them in series, facilitated by a number of diodes. From here, the accelerated ions go to the Linac [39] before being fed to the Booster [40]. The Linac is a two-stage linear accelerator consisting of an older drift-tube accelerator responsible for 116 MeV acceleration and a newer (1993) side section which finally brings the beam up to 400 MeV. The Booster is a synchrotron which uses non-superconducting electromagnets to bend the protons around a 75 m radius ring. Using a multi-turn procedure (more than one orbit of the ring) the protons from the H^- ions are merged with the already circulating protons in the Booster ring via charge-exchange injection. In this process the incoming H^- ions from the Linac are steered together with the protons from the ring and allowed to interact electromagnetically. They are then passed through a carbon foil to strip off the electrons and the bare protons are steered back into the ring, allowing other products to pass on to the beam dump. Once these protons are accelerated up to 8 GeV they are ready to be sent to the Main Injector [41]. The Main Injector is another synchrotron of radius 528 m which further accelerates the protons from

8 GeV to 150 GeV when they are finally ready to be put into the Tevatron. This is done in three superbunches separated by 2.64 μs. Each superbunch contains 12 bunches that are 396 ns apart. The Tevatron is also a synchrotron, but it uses superconducting electromagnets to curve and focus the beam of protons to their final energy of 980 GeV where they are ready for collisions with the antiprotons.

The antiprotons are produced starting from protons in the Main Injector. For the sake of antiproton creation, the protons are only accelerated to 120 GeV instead of 150 GeV. These slower protons are then directed to the antiproton source where they collide with a nickel target. The resulting hadronic shower, which includes some antiprotons, is run through an azimuthal magnetic field created by a lithium lens carrying 0.5 MA of current. This lens focuses the negatively charged particles and they are passed through a magnet to steer antiprotons with approximately 8 GeV of energy to the Debuncher. The Debuncher is a synchrotron of radius 505 m. The bunched structure of the protons in the Main Injector give the resulting antiprotons a bunched structure.

These bunches also have an undesirably large spread in energy. To reduce this spread, as the antiprotons move around the Debuncher ring an RF cavity gives them a small boost. The higher energy ones take a longer path than the lower energy ones, meaning that they take a longer time to complete one orbit and the result is a bunch with a smaller spread of energy. It takes about 100 ms for the Debuncher to produce this “cooler” beam of 8 GeV antiprotons, which is then sent to the Accumulator synchrotron. The 474 m circumference of the Accumulator sits just inside the Debuncher, where bunches of antiprotons are sent to collect. Once it becomes inefficient to store more antiprotons in the Accumulator, it is emptied to the Recycler and the Accumulator begins to fill up again. The Recycler is a storage ring below the Main Injector which stores and further cools the beam via electron cooling [42]. In this process a low emittance electron beam is circulated with the antiprotons and they are allowed to come to thermal equilibrium. Finally, this cooled beam of 8 GeV antiprotons is ready to be moved from the Recycler to the Main Injector. As with the protons, the antiprotons are then accelerated up to 150 GeV and passed to the Tevatron in 36 bunches (three superbunches). Once in the Tevatron they are accelerated up to collision energy, 980 GeV. The protons and antiprotons orbit the ring in opposite directions, taking a helical path around each other until they are brought to collision at two points, in the center of the D0 and CDF detectors.

3.2 D0 Detector

The D0 detector [43, 44, 45] was created for detecting particles that result from the proton-antiproton collisions in the Tevatron. It was used to collect data from 1992-1996 after which parts of the detector were upgraded and data taking resumed in March 2001. It is intended that the D0 detector continue this data taking through 2011, when it will have collected several hundreds of millions of events. A few thousand of these events will contain a Higgs boson, if it exists. The detector is made up of many subsystems to meet the challenges of detecting particles that interact in a variety of ways with different materials. In order to properly identify a Higgs-like event , we must be able to account for all of the energy from its parent proton-antiproton collision. We also must measure the momentum of all of the decay products in the collision event to be able to rebuild, or “reconstruct”, the suspected Higgs boson. In Fig. 3.2, the D0 detector is shown with the subsystems labeled. In the following sections each subsystem is described along with some mention of why it is important to Higgs analysis.

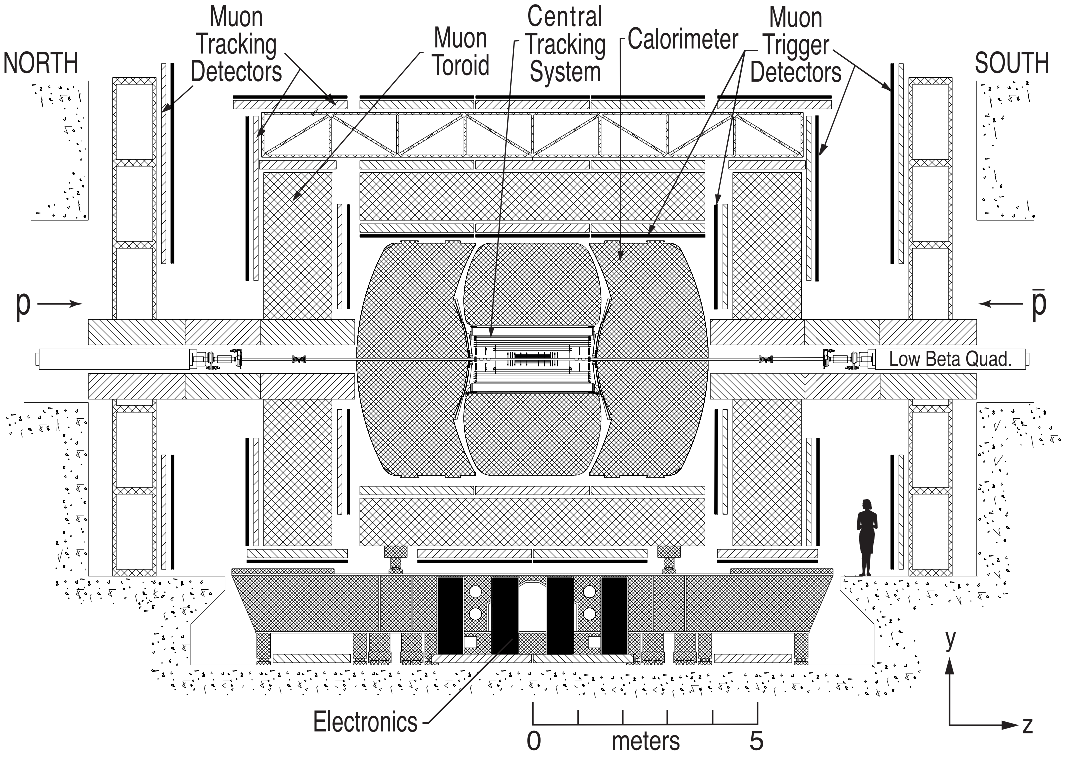

Figure 3.2: Diagram of the D0 detector.

3.2.1 D0 Coordinate System

Before discussing the layout of the D0 detector, it is useful to mention the coordinate system most suited to describing the location of particles in it. The coordinate system of choice is positioned with the origin at the center of the D0 detector and the positive direction of the z-axis pointing in the direction of the proton beam. The positive x-axis points out from the center of the Tevatron ring and the positive y-axis points up (towards the sky). In cylindrical coordinates, the transverse distance is given by $r = \sqrt{x^2 + y^2}$, the azimuthal angle measured from the x-axis is denoted ϕ, and the polar angle measure from the z-axis is denoted θ. It is usually more convenient, however, to use the pseudorapidity defined as

$$\eta = -\ln\left(\tan\frac{\theta}{2}\right) \tag{3.1}$$

instead of the polar angle. In the relativistic limit ($E >> m$), η is a good approximation of the rapidity of the particle

$$y = \frac{1}{2}\ln\left(\frac{E+p_z}{E-p_z}\right), \tag{3.2}$$

which is a Lorentz invariant quantity. Furthermore, particles tend to be produced more or less uniformly in rapidity, which is why the D0 detector is segmented in equal units of pseudorapidity and why η is a more convenient coordinate than θ.

3.2.2 Central Tracking System

When a collision happens in the detector there is a spray of particles that travel outward from that point. Whether there are two resulting particles or one hundred, one of the most important things to measure is the momentum of each of these particles right after the collision. This gives us information about the originating location of the particles, how we expect them to progress through the detector, and a clue about their identities. Applying a known magnetic field immediately after the collision helps us to determine whether the particles are charged and what the momentum of each might be. We also need a way to track the path of these particles as they travel outward through the magnetic field, thus we have the Central Tracking System in the D0 detector.

The central tracking system consists of two tracking layers, an inner Silicon Microstrip Tracker(SMT) and a Central Fiber Tracker(CFT) outside of that. A 2 T superconducting solenoid magnet surrounds both layers. The central tracking system is the first part of the detector that particles encounter after the hard collision in the center of the detector. It is used to determine the location of the collision point by tracing back hits in the tracker to a common point in the beam line. It is also used to measure the paths and momenta of charged particles, which curve in the presence of the magnetic field from the solenoid.

The SMT

The SMT provides both tracking and vertexing over nearly the full η coverage of the calorimeter and muon systems which are described later. The length of the proton-antiproton interaction region that the SMT covers is approximately 25 cm. It is desirable to deploy the silicon detectors such that the tracks are generally perpendicular to their surfaces for all η allowing less ambiguous hits. The design of barrel modules accommodates this by having interspersed disks in the center and assemblies of disks in the forward regions. The barrel detectors primarily measure the r - ϕ coordinate and the disk detectors measure r - z as well as r - ϕ. For particles at high η, the vertices are reconstructed in three dimensions by the disks. The vertices of particles at small values of η are measured in the barrels and central fiber tracker(described in the next section).

The SMT detector has six barrels in the central region. Each barrel has four silicon readout layers. The silicon modules installed in the barrels are called "ladders". Layers 1 and 2 have twelve ladders each; layers 3 and 4 have twenty-four ladders each, for a total of 432 ladders. At the ends of each barrel there is a disk of twelve double-sided wedge detectors, called an "F-disk". Forward (in |z|) of the three disk/barrel assemblies on each side is a unit consisting of three F-disks. At the farthest forward regions there are two larger diameter disks called "H-disks". They provide tracking at high $|\eta|$. Twenty-four full wedges, each consisting of two back-to-back single-sided "half" wedges, are mounted on each H-disk. There are 144 F-wedges and 96 full H-wedges in the tracker. Each side of one of these wedges (upstream and downstream) is read out independently. The SMT is read out by custom-made 128-channel SVXIIe readout chips. In 2006 a new layer, called layer 0, was installed in the SMT [46, 47]. It is now the layer closest to the interaction point and fits inside layer 1. Layer 0 helps to improve the vertex resolution. A computer-generated image of the SMT can be seen in Fig. 3.3.

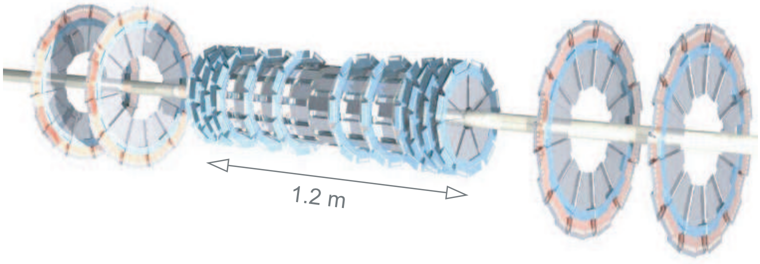

Figure 3.3: A computer generated image of the SMT.

The CFT

The CFT consists of scintillating fibers mounted on eight concentric support cylinders and occupies the radial space from 20 to 52 cm from the center of the beampipe. The outer cylinder provides coverage for $|\eta| < 1.7$. Each cylinder supports one doublet layer of fibers oriented along the beam direction (axial layers) and a second doublet layer at an angle in ϕ of $+3^o$ or -3^o (stereo layers). From the smallest cylinder outward, the fiber doublet orientation alternates between an axial layer (labeled z) and a stereo layer, with the stereo layer alternating $+3^o$ (labeled u) or -3^o(labeled v) and an orientation: zu - zv - zu - zv - zu - zv - zu - zv. The scintillating fibers are coupled to clear fiber waveguides which carry the scintillation light to visible light photon counters (VLPCs) for read out. The small fiber diameter (835 μm) gives the CFT an inherent doublet layer resolution of about 100 μm as long as the location of the individual fibers is known to better than 50 μm [48]. Fig. 3.4 shows a diagram illustrating the layers of scintillating fibers in the CFT.

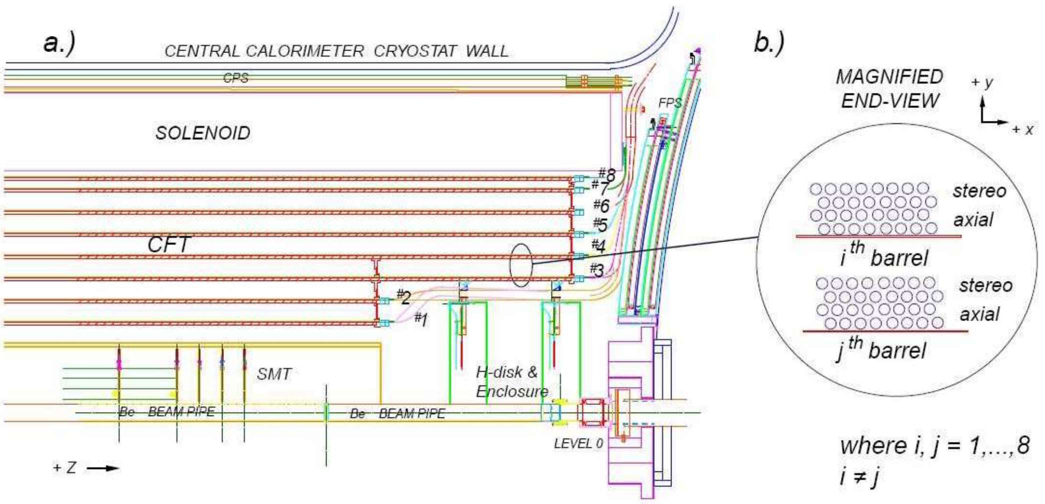

Figure 3.4: Diagrams illustrating the layers of scintillating fibers in the CFT.

The Solenoid Magnet

The superconducting solenoid magnet [49] is 2.73 m in length with an outer radius of 71.0 cm and made from niobium-titanium wire that is cooled to less than 4.7 K using liquid helium. The solenoid produces a 2 T magnet field parallel to the z-axis that causes charged particles to curve as they travel radially out from the interaction point. The radius of curvature in meters is given by

$$R = \frac{p_T}{0.3qB}; \tag{3.3}$$

where p_T is the transverse momentum in units of GeV, q is the charge of the particle in units of electron charge (e), and B is the magnetic field in Tesla; thus providing a measurement of the momenta of charged particles passing through the SMT and CFT.

Besides providing a measurement of momentum, the central tracking system is important for particle identification. Charged particles such as electrons and muons leave tracks, while neutral particles such as photons and neutrinos leave no tracks. Particles produced with color charge such as quarks and gluons will produce a shower or “jet” of particles that result in a group of many tracks.

Sometimes these jets can involve tracks from leptons as well, as is the case in semileptonic b decays where the b hadron decays to a muon inside the tracking system.

3.2.3 Preshower Detectors

As a bridge between the tracking system and the calorimeter, the central and forward preshower detectors are located just outside of the superconducting solenoid and in front of the calorimeter. They are constructed of several layers of scintillator strips which are used to aid in charged particle tracking. They function as both calorimeters and tracking detectors, enhancing the spatial matching between tracks and calorimeter showers. These detectors can also be used off-line to help correct the electromagnetic energy measurement of the calorimeters for losses in the solenoid and upstream material.

3.2.4 Calorimeter

Moving outward past the preshower detectors, the calorimeter measures the position and energies of electromagnetic showers from electrons and photons and hadronic showers resulting from quarks produced in the hard interaction. These energy measurements are most important in the $H \rightarrow W^+W^- \rightarrow \mu^+\nu\mu^-\nu$ analysis in accounting for the energy in background processes like W+jets, where a significant portion of the energy in the event is deposited in the calorimeter via these showers.

To capture all of this energy, the calorimeter system contains three separate calorimeters; a central calorimeter (CC) and two end calorimeters (EC). Each of these sections has its own cryostat and contains electromagnetic (EM) and hadronic layers, where the hadronic layer is further separated into a “fine” hadronic (FH) and “coarse” hadronic (CH) region. The CC covers out to $|\eta| < 1.1$

and the two ECs extend the coverage to around $|\eta| < 4$.

The EM Layer

The calorimeter is made up of many readout layers allowing energy deposits to be sampled at various depths, thus providing additional information for particle identification. The inner four readout layers of the CC and ECs are the EM layers designed to measure electron and photon energies. An electron or photon moving through the calorimeter will produce an EM shower through repeated bremsstrahlung and electron-positron pair production. When the energies of the electrons in the shower fall below a critical value, ionization becomes dominant over bremsstrahlung and the showering stops. The extent of an EM shower in a particular material is characterized by the radiation length, X_0 , which is the average distance an electron or photon will travel before losing 63% $(1 - e^{-1})$ of its energy. The EM layers have a total thickness of approximately 20.5 X_0 near $\eta = 0$. Most of the energy from an electron will be captured by EM layers with just the tail end making it into the first hadronic layer (see Fig. 3.5).

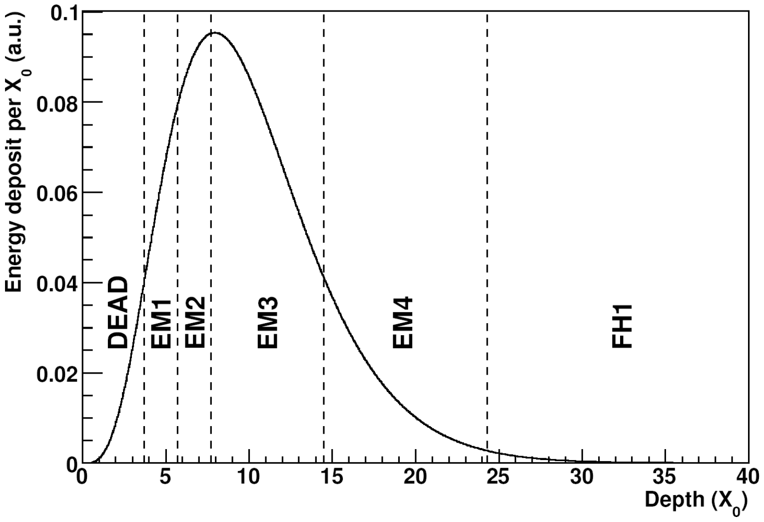

Figure 3.5: Expected energy profile as a function of depth for the shower of a 45 GeV electron. Also shown are the corresponding layers in the calorimeter.

The Hadronic Layer

Jets are composed mostly of pions and kaons that interact via the strong nuclear force. As the hadrons travel through the calorimeter they scatter off nuclei producing more hadrons (mostly pions). The neutral pions will quickly decay to photons resulting in some EM showering while the charged pions will continue to scatter off nuclei and produce more hadrons resulting in a hadronic shower. The extent of a hadronic shower is defined by the nuclear interaction length, λ_I , analogous to the radiation length for EM showers. However, for a given material λ_I is usually much longer than X_0. In particular, the total thickness of the EM layers is only 0.76 λ_I in the CC and 0.97 λ_I in the ECs. This means it takes much more material to contain a hadronic shower. Each readout layer is composed of alternating layers of grounded absorbing plates (uranium in most of the calorimeter) and signal boards at a potential of 2 kV. In the space between the absorbing plates and signal boards is liquid argon. Charged particles from an EM or hadronic shower will ionize the liquid argon resulting in a current between the absorber plates and signal boards. This current indicates the amount of energy lost by the showering particles in that region of the calorimeter. A schematic of the D0 calorimeter can be seen in Fig. 3.6.

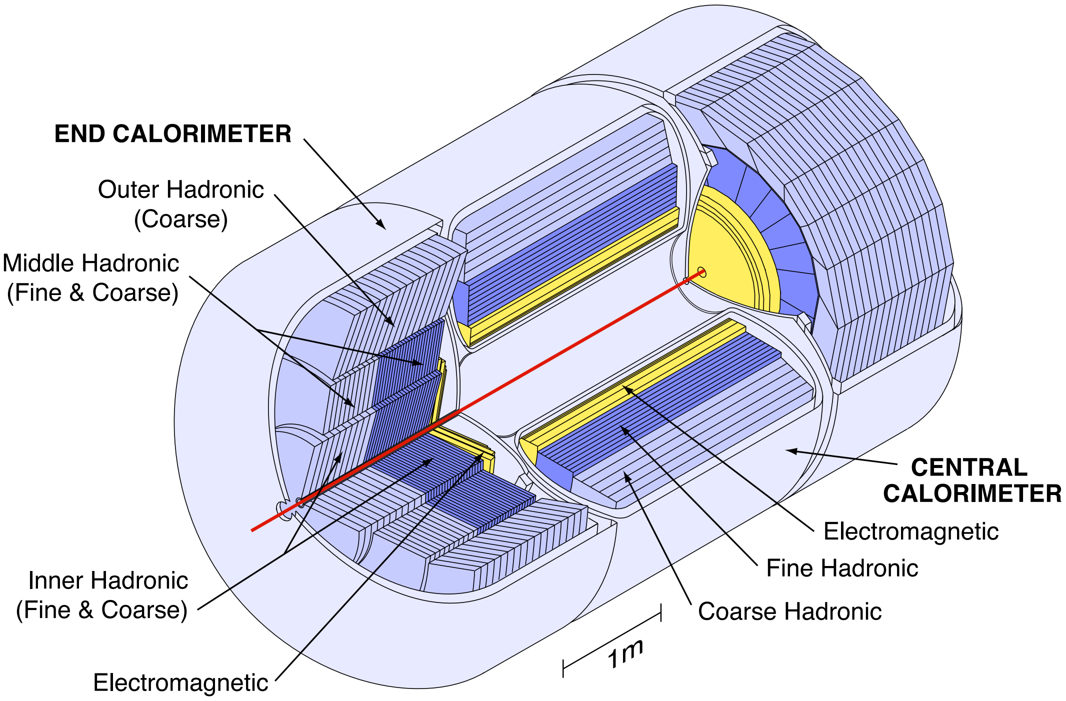

Figure 3.6: Illustration of the D0 calorimeter.

In total, the calorimeter has around 50,000 readout cells. With so many cells there is a significant chance for a false signal from electronics noise and/or uranium decay in the calorimeter. Therefore, before object reconstruction (chapter 4) the so called "T42" zero-suppression algorithm [50] is used to reduce the effects of noise by removing cells that do not measure an energy significantly higher than the noise level. Specifically, cells are removed unless they have an energy at least four standard deviations above the noise ($E^{\text{cell}} \geq 4\sigma^{\text{noise}}$) or have $E^{\text{cell}} \geq 2.5\sigma^{\text{noise}}$ and are adjacent to a cell with $E^{\text{cell}} \geq 4\sigma^{\text{noise}}$.

3.2.5 Muon System

A good muon detector is of obvious importance to any search with muons as a final state particle. In addition to the measurements of position and momentum measured by the tracking system, it is advantageous to get those measurements again at a point much farther away from the interaction region. Not only does this give another one or more measured values to add to the description of the particle, it extends the total path that the muon is tracked over improving the ability to monitor its position. It also creates a way to simply count muons and record , or “trigger” on, events containing them if desired (see section 4.6.1). As muons pass through the calorimeter, they don’t interact as strongly with the electron clouds around the atoms in the material as electrons, causing them to travel a much greater distance relatively unhindered. They also travel farther than hadrons from quarks or gluons because thos particles interact via the strong force with the material. This can be exploited by putting a muon detection system beyond the calorimeter. This way, any hits in the muon system are likely due to actual muons, which greatly reduces the muon fake rate in the system. The primary energy loss of the muons before they reach the muon system is due to ionization.

The muon system has a layer of tracking detectors immediately outside of the calorimeter followed by a layer of scintillation trigger counters and then 1.8 T toroidal magnets. Outside of the toroids are another two sets of detectors and trigger counters. The detectors used in the central region are Proportional Drift Tubes (PDTs) while the detectors on the ends are Mini Drift Tubes (MDTs). The muon system is used to identify and measure the momentum and position of muons.

For muon triggering and measurement, the detector uses the central muon system proportional drift tubes (PDTs) and toroidal magnets, central scintillation counters, and a forward muon system. The central muon system provides coverage for $|\eta| < 1.0$. The forward muon system extends muon detection to $|\eta| \approx 2.0$, uses mini drift tubes (MDTs) rather than PDTs, and includes trigger scintillation counters and beam pipe shielding. A set of scintillation counters known as the cosmic cap is on the top and upper sides of the outer layer of central muon PDTs. This coverage is continued on the lower sides and bottom of the detector, to form the cosmic bottom. These trigger scintillation counters are fast enough to allow association of a muon in a PDT with the appropriate bunch crossing and to reduce the cosmic ray background. Additional scintillation counters, the A-ϕ counters, have been installed on the PDTs mounted between the calorimeter and the toroidal magnet. The A-ϕ counters provide a fast detector for triggering and identifying muons and for rejecting out-of-time background events. The scintillation counters are used for triggering; the wire chambers are used for precise coordinate measurements as well as for triggering. Both types of detectors contribute to background rejection: the scintillator with timing information and the wire chambers with track segments. The layout of the muon system is pictured in Fig. 3.7.

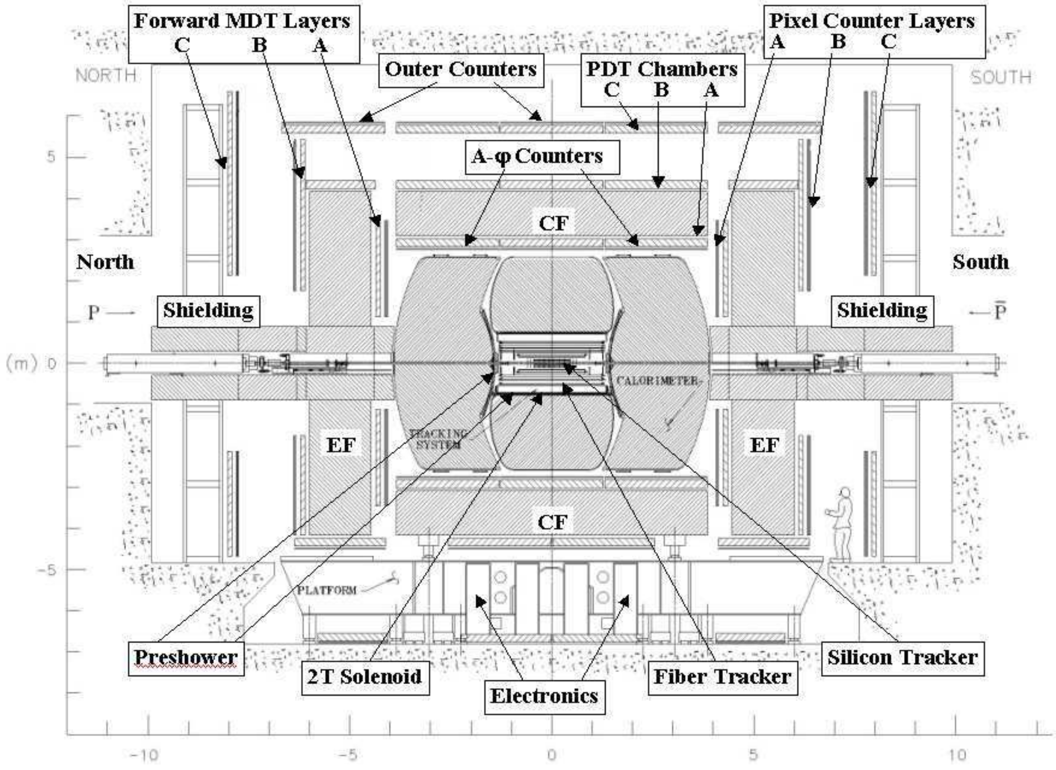

Figure 3.7: Illustration of the D0 Muon System.

3.2.6 Luminosity Monitor

If we want to measure how often a given process occurs in our collisions, we first need to know the total number of collisions we have recorded. We also need to know the total number of collisions that have occurred relative to that. These quantities are measured using a luminosity monitor. The luminosity monitor(LM) at D0 serves as a means to measure the number of collisions taking place at a given moment in the detector (instantaneous luminosity). It also allows the calculation of the total number of interactions over time (integrated luminosity). Luminosity is measured using plastic scintillator arrays located in front of the EC cryostats, covering $2.7 < |\eta| < 4.4$. These circular arrays consist of 24 scintillator wedges that catch remnants from the collisions in the center of the detector. These remnants are in the form of hadrons created by the quarks and gluons that were not directly involved in the hard interaction.

The inelastic proton-antiproton counting rate is used to determine the instantaneous luminosity. The effective cross section is derived from the inelastic cross section, $\sigma_{inelastic}$ (at 1.96 TeV) $= 60.7 \pm 2.4$ mb [51], and 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.

3.3 Trigger

Collecting detector information for all of the $\sim$ 1.7 million bunch crossings in the center of the D0 detector that the Tevatron yields every second would be nearly as pointless as it was difficult. The cross section for inelastic proton-antiproton scattering at the Tevatron is about 50 mb, resulting in an average of a few inelastic collisions per bunch crossing. Of these, only a scant few contain interesting physics. For example, the gluon-fusion production channel for the Higgs boson, as mentioned in section 2.4.1, has a cross section of about 0.431 pb at a Higgs mass of 160 GeV. This is about 10 orders of magnitude smaller than the total inelastic cross section at the Tevatron. Writing all of this information to tape and then sorting through the enormous mountain of useless events to find the interesting ones would be a huge waste of time and space. To help alleviate this difficulty to some degree, D0 employs a multi-stage triggering system [52, 53]. The first level of trigger (L1) uses preliminary information from tracking, calorimetry, and muon systems to lower the accepted event rate to $\approx$ 2 kHz. At the next trigger stage (L2), more detailed information is used to limit the accept rate to $\approx$ 1 kHz. These first two levels of triggering rely mainly on hardware and firmware. The final level of the trigger (L3) has access to all of the event information. It uses software algorithms and a computing farm to reduce the output rate to $\approx$ 50 Hz, which is written to tape.

Level 1 trigger is designed to find potentially interesting patterns of energy deposition indicating the passage of high energy particles and must be very fast (4.2 μs) to keep up with the rate of interactions. It therefore uses a condensed subset of the full detector readout and is implemented entirely in hardware and firmware. Each detector subsystem checks if the event passes preprogrammed trigger conditions (e.g., the calorimeter trigger tests for energy deposits above pre-programmed thresholds), then the results from each detector subsystem are combined to make the final Level 1 decision. After a Level 1 trigger is passed there is 10 μs of dead time required for readout resulting in the maximum accept rate for Level 1 to be set at 5 kHz. After an event passes Level 1 it is sent to Level 2 where it is subjected to more refined tests that may take up to 100 μs.

In addition to firmware, the Level 2 trigger uses microprocessors to take advantage of more precise detector information and spatial correlations to form basic objectssuch as tracks, EM clusters, and jet clusters. Each detector subsystem has a dedicated microprocessor that reduces the data for that subsystem then sends it on to a global processor to make the final Level 2 decision. The rate at which

Level 2 can pass events to Level 3 is limited to 1 kHz by the Level 2 processing time. Only when an event passes the Level 2 trigger is the entire detector read out and sent to the Level 3 CPU farm made up of over 100 Linux computers.

The Level 3 CPUs work together to process each event in under 25 ms. During that time the entire event is reconstructed and sophisticated algorithms, close to those used for offline analyses, are applied to make the final decision of whether to keep the event. The output rate for Level 3 is maintained around 50 Hz, which was chosen to keep the offline event reconstruction from piling up.

Chapter 4

Event Reconstruction

Correctly reconstructing events of interest in the detector is at the heart of any analysis. An "event" is associated with a single bunch crossing in which multiple protons and antiprotons may collide in the detector region and create a spray of particles. This event is considered interesting if it is accepted by a preset suite of triggers (Sec. 4.6.1). One must be able to characterize all of the constituents of an event to be able to work backwards and figure out the processes that took place to spawn them. We must determine each particle's identity and four-momenta from the hits left as it passes through the tracking detector, the energy it deposits in the calorimeter, and information from the muon detector. To help us do this, we create reconstructed "objects" from the detector readout, like tracks which map out the path of a charged particle or missing energy which represents the sum of the unobserved particles needed to conserve momentum.

4.1 Tracks

Tracks are objects we use to calculate the momentum of charged particles passing through the D0 tracking system (the SMT and CFT). They represent the path of a charged particle and from that representation we can determine the particle's momentum (perpendicular to the beam line). This is done by looking at the curvature of these tracks as they travel through the magnetic field permeating the tracking system. This is the most precise way to measure the momentum of the muons in this analysis due to the poor momentum resolution of the muon system.

4.1.1 Track Reconstruction

As the charged particles pass through the SMT and CFT they interact with the various strips and fibers. An interaction in a given layer is designated as a ”hit” in the tracking system. These hits are then used to reconstruct tracks. These reconstructed tracks are curved, with a radius of curvature related to momentum of the particle and the magnetic field in the tracker as was mentioned in section 3.2.2, Eq. 3.3. The resolution for hits in the SMT (CFT) is approximately 10 μm (100 μm) in the azimuthal direction and 35 μm (2 cm) in the z direction. It is often difficult to correctly identify which hits belong to which tracks, or even how many tracks should be reconstructed. There are two methods used to find and reconstruct tracks. This increases the chances that all of the tracks will be identified.

The first method is a “road search” method and works from the inner layers of the tracking system outward. Starting with the hit closest to the beam-line, the algorithm for finding the track looks for a hit further out that has a difference in azimuthal angle from the first hit of $\Delta\phi < 0.08$. It then searches for a third hit consistent with a track that has a radius of curvature $\rho > 30$ cm (i.e., $p_T >$ 180 MeV), a distance of closest approach to the beam line $d_0 < 2.5$ cm, and a χ^2 fit value $\chi^2 < 16$. Track candidates found in this way are used to construct tracks through the entire tracking system using a Kalman filter [54]. The filter tries to extrapolate out from each candidate and include hits that fall into a small window around this extrapolation. It also accounts for multiple scattering in calculating track parameters. The hits must fit the tracks with a $\chi^2 < 16$. If three consecutive layers have no hit in this window then the algorithm stops. Tracks cannot share more than 2/3 of their hits with another track.

The second method used to find tracks is a “histogramming method”. A two-dimensional histogram is created in $\rho \times \phi_0$ space using the transverse projection of all the hits in the SMT. Here, ρ is the curvature of a circle intersecting the origin and the hit and ϕ_0 is the angle of the tangent to that circle at the origin. This histogram is populated by converting the x and y coordinates of each hit to ρ and ϕ_0 using a Hough transformation [55]. Hits from the same track will lie on top of each other in the histogram at the ρ and ϕ_0 of the track. These peaks in the histogram become the track candidates and a two-dimensional Kalman filter is used to more accurately calculate the track parameters and remove tracks with large χ^2. Next, a similar histogramming method is applied to find the z component of the tracks and then a three-dimensional Kalman filter is used to extrapolate these tracks into the CFT. Once all of this is done and the tracks are identified, the entire method is repeated starting from hits in the CFT and working inward.

The tracking system has a momentum resolution of

$$\frac{\delta_{p_T}}{\text{GeV}} = 0.002\frac{p_T^2}{\text{GeV}}. \tag{4.1}$$

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

4.2 Primary Vertex

The precise location of the initial collision between the proton and antiproton is very important so that one can determine the origin for measuring momentum vectors of particles in the various subsystems of the detector. There can be many collision vertices in a single bunch crossing. The number of vertices per bunch crossing is a Poisson distributed random variable. The location of the collision which leads to an event of interest (ie. one that passes a trigger) is referred to as the "primary vertex" (PV). Other vertices get recorded due to their coincidence in time with the interesting event and are called minimum-bias interactions. The PV is the starting point for reconstructing a complete event, but is actually determined by tracing back the tracks in the event to a common point along the beam line.

4.2.1 Primary Vertex Reconstruction

Reconstructing collision vertices is done in three basic steps. First, tracks with $p_T > 0.5$ GeV and at least 2 hits in the SMT are clustered along the beam axis. This allows the separation of possible additional $p\bar{p}$ interactions taking place during the bunch crossing. Then an estimation of the position and width of the beam is performed by fitting all of the tracks in the cluster into a common vertex using a Kalman Filter vertex fitting algorithm. Additionally, a preselection on the tracks corresponding to each cluster is performed based on their distance of closest approach to the beam spot(required to be within 100 standard deviations).

Next, the Adaptive Vertex Fitting algorithm [56], an iterative Kalman Filter fitter, is applied and reweights track errors according to their χ^2 contribution to the vertex by means of the Fermi-like function given in Eq. 4.2.

$$w_i = \frac{1}{1 + e^{(\chi_i^2 - \chi_{cutoff}^2)/2T}} \tag{4.2}$$

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 to identify which ones correspond to the largest momentum transfer. The remaining primary vertices are assigned to additional soft interactions which may have taken place in the same bunch crossing. For each track, a probability of whether it is compatible with the momentum distribution for a soft collision is defined:

$$\mathcal{P}(p_T) = \frac{\int_{p_T}^{\infty} F(q_T)\mathrm{d}q_T}{\int_{0.5}^{\infty} F(q_T)\mathrm{d}q_T} \tag{4.3}$$

where p_T is the measured transverse momentum of the track, $F(q_T)$ is the transverse momentum distribution of tracks produced in soft collisions predicted by Monte Carlo simulations, q_T is a dummy variable, and the 0.5 bound on the integral is due to the $p_T > 0.5$ GeV cut made when determining which tracks to use for vertex reconstruction.

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. 4.4:

$$\mathcal{P}_{soft} = \Pi \sum_{k=0}^{N-1} \frac{-ln\ \mathcal{P}(p_T)}{k!}. \tag{4.4}$$

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.

4.3 Muons

Muons are some of the most important objects in this analysis as mentioned earlier. Identifying muons and reconstructing their location and momentum is accomplished primarily using the tracking detector and the muon system. A muon is identified in the detector starting with a track in the muon system and then matching this with a track in the tracking system. The following sections will describe how these muons are reconstructed in the detector and identify some of the important variables used to classify them.

4.3.1 Muon Reconstruction

As discussed earlier in section 3.2.5, the muon system is made up of an A layer inside the toroid magnet and B and C layers outside. For muon reconstruction the B and C layers are considered together. A track in the muon system begins in the A layer by fitting a straight line to hits in two or more decks of the drift tubes to create a track segment. This is repeated in the BC layer. If more than one track segment is found in a given layer and octant, only the one with the best fit is kept. Once track segments are identified, a fit is performed through the toroid between A segments and BC segments. This fit also accounts for bending from the magnetic field and effects from energy loss and multiple scattering in the material. At this point the muon's momentum can be determined by looking at the deflection due to the toroid, but this is far less precise than using a track in the central tracking system. Therefore, muons that can be matched to a central track use the curvature of that track for momentum determination. The match is done by extrapolating the muon track inward, modeling the calorimeter as a thin scatterer with the appropriate number of radiation lengths, and fitting to a central track. The track with the lowest χ^2 fit is used. The muon identification efficiency is $\sim 92\%$ [57].

4.3.2 Muon Classification

The quality of the muon system measurement, the quality of the central track, and the isolation of the muon from other particles are all important in determining the overall quality of the reconstructed muon. These muon quality variables are used to cut away events with questionable muons as discussed in section 6.1.

Muon Quality

The reconstructed muon candidates are classified using the parameters of muon type and quality [58]. The nseg variable is used to determine the type of muon, where $\mid nseg \mid = 1$, 2, or 3 means the muon has hits in only the A layer, only the B or C layers, or all three layers of the muon system, respectively. A positive value of nseg indicates that the muon candidate is matched to a central track and a negative one indicates it is not. The quality of a muon can be loose, medium or tight depending on the number of hits in the muon chambers and scintillators and the local muon fit. In this analysis, muons are required to have at least loose quality.

In order to further control the muon purity, the matched central track is certified with 3 quality levels, loose, medium and tight. The classification uses the number of SMT hits, the track fit χ^2, and distance of closest approach(dca) to the primary vertex to distinguish levels. A track is defined as loose if $\mid dca_T \mid < 0.2$ cm and there is no SMT hit. This tightens to 0.02 cm if there are SMT hits. A track is considered medium if it passes the loose requirement and the track fit satisfies $\chi^2/ndf < 4$. A tight track fulfills the medium requirement and has SMT hits. Medium quality tracks are used in this analysis to allow slightly higher acceptance than the tight tracks without increasing the fake rate by using a loose track.

Isolation

The isolation quantity is defined as the sum of the transverse momenta of all tracks in a cone of radius $R < 0.5$ around the muon direction plus the energy deposited in the calorimeter in an annular cone of radius $0.1 < R < 0.4$ around the muon direction all divided by the muon transverse momentum. It is an important quantity in this analysis because we are concerned with muons coming from weak boson decays that should have relatively little going on around them in the detector. By restricting the amount of energy around the muon we can be more sure of keeping the muon's energy separate from other objects, like hadronic jets or electrons.

4.4 Jets

A jet is not an individual particle, but represents a collimated shower of energetic particles. While there is no explicit number of jets required for the events in this analysis, they are still important. We need to count all of the energy in an event to properly reconstruct it. Most jets are the result of the hadronization of a high energy quark or gluon produced in the hard interaction. They usually have many tracks associated with them, potentially one for each particle in the jet. After passing through the tracking system they result in showers of energy in the EM and hadronic layers of the calorimeter.

4.4.1 Jet Reconstruction

Jet reconstruction is done in three stages [59, 60]. The first stage is constructing preclusters of towers. A massless four-momentum for each cell in a tower is taken to be the energy of the cell, pointing from the primary vertex to the center of the cell. These cell momenta are summed to get a four-momentum for each tower. All towers with transverse momentum greater than 0.5 GeV are then put into a list of seeds. The highest momentum tower in this list is used first. A cone of $\Delta\mathrm{R} = 0.3$, where $\Delta\mathrm{R} = \sqrt{\Delta\eta^2 \times \Delta\phi^2}$, is placed around the tower and the momentum of all of the towers in the cone are summed to make a precluster. All of the towers used in this precluster are then removed from the list of seeds and the procedure is repeated until the list of seeds is exhausted. The second stage in jet reconstruction is creating protojets using the preclusters as seeds. Preclusters with

$$p_T^{precluster} > 1 \text{ GeV} \tag{4.5}$$

and containing more than one tower make up the list of seeds. A cone of radius ΔR = 0.5 is placed around the highest momentum precluster in the seed list and towers within the cone are summed to form a trial protojet. The cone is then re-centered on the trial protojet and a new protojet is found by summing those towers. The process is repeated until a stable protojet is found. If

$$p_T^{protojet} < 4 \text{ GeV} \tag{4.6}$$

then the protojet is discarded. Any preclusters within ΔR = 0.25 of the stable protojet are removed from the list of seeds and the procedure is then repeated to find the next protojet. The last stage of jet reconstruction is splitting or merging overlapping protojets. A protojet that does not overlap any other protojet is considered a jet. If two protojets overlap and share at least half of the p_T of one of the protojets they are merged; otherwise they are split. This merging is done by summing the two protojets to make a new trial protojet. The procedure described earlier is iterated until a new stable protojet is found. When splitting is required it is done by assigning each of the shared towers to the closer of the two protojets being split. Splitting and merging continues until there are no more overlapping protojets and all the jets are defined.

4.4.2 Jet Energy Scale

In theory, the energy contained in a jet corresponds to the energy of the particles that make up the jet. In practice this is only approximately true due to things like detector effects, dead material, sampling rate and the fixed radius of the cone in which the energy is counted. This means the calorimeter jet energy can differ significantly from the initial energy of the particles forming the jet. To take these effects into account, a correction to the energy of the reconstructed jets is applied [61]. This correction is defined as

$$E_{jet} = \frac{E_{jet}^{raw} - O}{F_\eta \times R \times S} \tag{4.7}$$

where

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 estimation of the corrections is done separately for data and Monte Carlo to account for differences between the detector and detector simulation as well as any other differences between actual jets and simulated jets. First, the offset correction, O, is subtracted from the raw jet energy. The offset energy arises from having multiple $p\bar{p}$ interactions, beam remnants, noise from the electronics and uranium in the calorimeter, or energy from previous collisions ("pile-up"). It is measured from data using minimum-bias events triggered by the luminosity monitor. Then the energy density per tower in these events is measured as a function of the number of reconstructed primary vertices in order to include luminosity dependence (which is related to the number of PV) in the offset energy calculation. The energy deposited within a jet cone using these towers is then defined as the offset energy.

After correcting for the offset, the jet energy is divided by the relative response correction F_η, the absolute response correction R, and the shower correction S. The relative response correction accounts for the fact that the response of the calorimeter is not uniform in rapidity(due to non-uniform materials, etc..). The absolute response correction includes the difference in calorimeter response between hadrons and electrons and also the energy loss in non-instrumented detector regions. Finally, due to the bending of charged particles in the D0 solenoidal field and the arbitrary selection of a fixed cone which does not always fully contain the transverse shape of every jet, some energy that should be included inside the jet cone makes its way outside of the cone (or some that should not be included makes its way in). The showering correction accounts for this effect by adding in, or taking out this energy. This correction does not account for physical showering as one might assume from the name (eg. due to gluon emission).

When considering jet energy corrections it is also important to account for muons that are reconstructed within the jet cones. The energy can be further corrected to consider the energies of this muon. One must be careful however, that this muon actually came from the jet being measured and is not simply an isolated muon that coincidentally passes through the jet cone. If this is the case, the jet energy will be vastly mis-corrected as if that muon were part of the jet and the energy of the muon will also go into the event reconstruction as a muon, thus double counting the energy and incorrectly summing up the total energy in the event. In this analysis, jets are not corrected for muons, but a restriction is placed on events that requires no muon to be closer than a ΔR of 0.1 to the jet cone (thus making sure there is no direct overlap). In addition to this, muons are required to be isolated (see Sec. 6.1) and therefore should not be contained in jets.

4.5 Missing Energy

Neutrinos are near massless particles that travel through the detector without leaving any trace of their passing. They are produced in weak interactions such as the decay of the W bosons in this analysis. While we cannot directly measure the information about the passing neutrino, we can use conservation of momentum to help us determine its kinematic properties. Usually only a single parton (quark or gluon) from each of the colliding hadrons participate in a hard interaction and the remnants of the hadrons continue down the beam pipe undetected. The fact that these remnants are confined to the beam pipe means that they have a negligible transverse component to their momentum. This allows us to require conservation of transverse momentum of the particles resulting from the hard interaction. Any imbalance in the total transverse momentum would then require some amount of missing energy($\vec{\not{E}}_T$) to balance it. This $\vec{\not{E}}_T$ is then the signature of particles escaping the detector undetected such as neutrinos.

4.5.1 $\vec{\not{E}}_T$ Reconstruction

Reconstructing the $\vec{\not{E}}_T$ can be a challenge due to our inability to measure it directly. It requires the accurate accounting of all of the detectable energy in the detector. To find $\vec{\not{E}}_T$ first the uncorrected missing transverse energy is calculated as measured by the calorimeter. This is done by summing the three-momenta in the EM and fine hadronic calorimeter cells. Here the momentum is the energy of the cell E_{cell}, pointing from the primary vertex to the center of the cell along a vector $\vec{\eta}^{cell}$. Then the transverse vector $\vec{\eta}_T$ is associated with the energy E_{cell} so that the transverse energy is $\vec{E}_T = E_{cell} \cdot \vec{\eta}_T$ and the $\vec{\not{E}}_T$ is the negative of the sum of these transverse energies:

$$\vec{\not{E}}_T = -\sum_{i=0}^{N_{cells}} \vec{E}_T^i \tag{4.8}$$

The $\vec{\not{E}}_T$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.

4.6 Data Processing

4.6.1 Trigger Requirements

No explicit trigger requirement is made in the analysis. This means that events which pass any trigger are considered as long as they pass preselection cuts and eventually final analysis cuts (chapter 6). There are many triggers which fire on the presence of one or more high momentum muon with looser requirements than are made at the preselection level and for this reason we assume 100% trigger efficiency and then correct for any deviations from this (see section 6.2.2. Effectively using "all" triggers increases the acceptance for Higgs boson signal events by approximately 10%, as compared to requiring an "OR" of several single muon triggers.

4.6.2 Luminosity and Data Quality

The data collection takes place in rounds of collisions called "stores". Within each store are multiple periods of continuous data taking which typically last 1 to 4 hours with just a few minutes of down time in between. During these "runs", information about the state of the detector is recorded to ensure that all of the various subsystems are operating as expected. A set of standard quality requirements [62] is applied to these data to remove runs marked as bad due to problems with the calorimeter, SMT, CFT, or muon system. In addition to keeping quality information for each run, the data taking is further broken down into intervals of approximately constant integrated luminosity, called 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. These LBNs can also be marked as bad both "online" while the data is being collected or "offline" during a validation step. LBNs marked as bad are removed from this analysis. The same data quality package used to remove the above mentioned events also removes events flagged as *cal_empty_crate, cal_ring_of_fire, cal_noon_noise* and *cal_coherent_noise* [62].

Chapter 5

Samples

To look for a signal in data, one must first have a model predicting what the data would look like with and without this signal. We create simulated events for the background processes and the signal to compare to what we observe in data. The following sections describe the data used, the simulated background and signal samples, and the treatment of events with multiple jets in which a jet fakes a lepton.

5.1 Data

The dataset used for this analysis was collected by the D0 experiment between April 2002 and December 2008. It includes two data collecting epochs corresponding to separate running periods of the Tevatron labeled RunIIa and RunIIb, due to an extended shutdown in spring 2007. During this shutdown, an upgrade to the D0 detector was performed primarily motivated by the projected higher instantaneous luminosities after the shutdown. It consisted of an upgrade to the trigger system and the installation of a new silicon microstrip layer of tracking called “Layer 0”. This layer was installed as the closest layer to the beam-pipe and was designed to compensate for radiation damage in the existing system and to improve impact parameters and lifetime measurements. This upgrade created differences in things like detector geometry and resolutions between the two data epochs and the differences are also propagated to the detector simulations.

At D0, data events are stored as “skims” that have a set of loose requirements imposed. These skims help to categorize events that might be interesting to a particular analysis and cut down on the number of events that must be analyzed. The skim used in this analysis is the 2Muhighpt skim for both data epochs. It requires events to contain at least two loose quality muons (see section 6.1), each with a $p_T > 10$ GeV. For RunIIa, this yields a luminosity of 1.1fb^{-1}. RunIIb data, up to the date when this analysis was frozen, contains a luminosity of 3.1fb^{-1}. These luminosities were obtained from the `vjets_cafe` package which calculates the luminosity by looking at events that passed the `JT125_L3125` trigger. This trigger, which requires at least one jet with $E_T > 125$ GeV, is never prescaled[1] and is present in every data-taking run used for physics analysis [63]. The number of events that pass this trigger are counted and the information is passed to the luminosity system where the appropriate LBNs are retrieved. The luminosity is then calculated for the trigger in each LBN. The integrated luminosity is the sum of LBN luminosities for each trigger with data quality taken into account.

5.2 Simulated Samples

Simulations are used to understand the various processes contributing to the background composition and signal in the analysis. Simulating an event typically requires connecting several descriptions of production and decay processes by com-

---

[1] Prescaled triggers are those that pass events through the trigger system more often than desired (whether they occur too often or only a fraction of the total events are needed) and so the number recorded is scaled down prior to recording by only recording every Nth event passing that trigger. The scaling is usually dependent on the instantaneous luminosity and can be changed at each run transition. An example of a prescaled trigger is the JT65 jet trigger where the prescale ranges from 4000 (1 event in 4000 is kept) at the highest instantaneous luminosities to 5 at the lowest.

bining perturbative theoretical calculations with non-perturbative phenomenological models. To create the simulated events in this analysis we use Monte Carlo (MC) event generators and then process them with a GEANT-based simulation of the D0 detector with zero-bias data events overlaid to produce a detector-level simulation of raw detector data. The zero-bias events used are actual data events collected by the detector, but not based on any physics trigger. Instead, one of these events is recorded every certain amount of time. Overlaying one (or more) of these events on an MC event effectively simulates random interactions that could be occurring at the same time.

The simulation is created in several phases. It starts with the hard interaction of the colliding particles. The hard interaction is calculated with perturbative quantum field theory. The particles interacting in the collision can radiate photons and gluons before the hard interaction takes place, which is called Initial State Radiation(ISR). Particles produced in the hard interaction can also radiate photons and gluons and this is called Final State Radiation(FSR). The FSR is separated from the ISR in the simulation. There are different methods for the simulation of this perturbative step, like parton showering and the color dipole ansatz [64, 65].

Parton showers are based on an improved leading-log approximation. They provide a simple, process-independent method that gives a sensible match to hadronization. Above energies of around 1 GeV, this fragmentation of quarks and gluons into hadrons takes place. To describe the hadronization one must turn to phenomenological models. Current models all have free parameters that need to be measured to properly describe the fragmentation process. The most common Monte Carlo generators used either employ the string fragmentation model as in Pythia [66] or cluster fragmentation as used in the Herwig event generator [67]. The Lund model implemented in Pythia splits gluons into $q\bar{q}$ pairs and turns them into hadrons via the string fragmentation model [68]. 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 muonic and tau final states, diboson, W+jets production, $t\bar{t}$, and multijet QCD events. 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 CTEQ6L parton distribution functions(PDFs) [69]. These PDFs determine the momentum structure of the proton and antiproton where the partons (quarks or gluons) originated and thus, the starting momentum of each parton in the simulated collision. The diboson and signal samples were generated using PYTHIA 6.409 [66]. Alpgen [70] 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 and is refered to as a fixed order matrix element generator(FOME) for this reason. Therefore the Z+jets and W+jets background processes are simulated using ALPGEN 2.11. The topologies simulated using Alpgen for Z+jets are $Z \rightarrow \mu\mu\ +\ Nlp$, $Z \rightarrow \mu\mu\ +\ 2bNlp$, $Z \rightarrow \mu\mu\ +\ 2cNlp$, $Z \rightarrow \tau\tau\ +\ Nlp$, $Z \rightarrow \tau\tau\ +\ 2bNlp$, and $Z \rightarrow \tau\tau\ +\ 2cNlp$. For W+jets they are $W \rightarrow \ell\nu\ +\ Nlp$, $W \rightarrow \ell\nu\ +\ 2bNlp$, and $W \rightarrow \ell\nu\ +\ 2cNlp$. Here, lp is any light flavor quark: u, d, or s. Also, N is 0, 1, 2, or 3 for the heavy flavor b and c-samples, and 0, 1, 2, 3, 4, or 5 for the processes without heavy quark content. Each Alpgen event is passed to Pythia for parton showering and hadronization. The samples then have next-to-next-to-leading order corrections applied as discussed in section 5.2.1 to improve the accuracy of the modeling.

In events with multiple jets, there is a slight probability for one or more jets to be misidentified as leptons. Even though the chances are remote for this to happen to two or more jets in an event, the possibility must be considered due to the large cross section for QCD processes. It would be possible to create a simulated sample of these QCD events and look for ones with two jets that get identified as leptons. This however, would require a prohibitively large number of events be created to get reasonable statistics. Also, it is extremely difficult to create a reliable simulation of these “fake” leptons. For these reasons we use an orthogonal selection of data events instead. In data, we have the advantage of a much larger number of events, where we have already required two leptons (section 5.1). This means that the multijet events we have in our sample already contain at least two “fake” leptons and this gives us much greater statistics. Since the analysis normally requires two muons of opposite charge (see section 6.1), to make an orthogonal selection the multijet background is estimated using like-sign data events. There is no reason to expect that the rate a jet fakes a lepton depends on the charge of the fake, so we can assume the number of events with fakes are equivalent in the like-sign and opposite-sign samples. We also do not expect signal in the like-sign sample (except for possible charge sign flips) so we can simply subtract the like-sign component of our other simulated backgrounds from the like-sign data and whatever events are left are assumed to be multijet(QCD). The procedure for creating the QCD background sample is described in section 6.2.3.

5.2.1 Normalization

Many more simulated events are generated for each process than are expected in data. This means that 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). So, only those Feynman diagrams without loop corrections and radiation processes of gluons or photons are considered. If higher-order corrections are considered, the calculated production cross section naturally becomes more accurate. Second order contributions of this type are called next-to-leading-order(NLO) and third order processes are referred to as next-to-next-to-leading-order(NNLO). Interference terms of virtual gluon loops also contribute to the NLO calculation. Considering as many higher order corrections as possible yields the best result. LO calculations are often very complex and most calculations of production cross sections are not beyond NLO. The ratio of a NLO cross section and the LO cross section is called a k-factor:

$$\mathrm{k} = \frac{\sigma_{NLO}}{\sigma_{LO}} \tag{5.1}$$

The k-factor between NNLO and NLO, k_{NNLO}, is considered for the normalization when available.

The standard model background processes and cross sections times branching ratios can be found in Table 5.2, Table 5.3 and Table 5.4. The cross sections for the diboson Monte Carlo samples are calculated in Ref. [71] with MCFM at NLO, the uncertainties are evaluated at LO. The MC samples are normalized to this NLO SM prediction.

The $t\bar{t}$ samples are normalized to a cross section employing improvements due to soft gluon resummation at NNLO logarithmic accuracy. The resummed results are expanded to analytical cross sections through NNLO. This leads to an approximate NNLO cross section [72, 73], which is used for normalization.

The $Z/\gamma \rightarrow ll$ cross section is calculated in NNLO, using QCD correction in α_s^2 according to Ref. [74]. The $Z + jets$ ($\mu\mu$ and $\tau\tau$) samples are normalized by the cross sections where the additional NNLO k-factor correction is 1.30. The Z + $(heavy\ flavor\ partons)$ receive an additional heavy flavor k-factor correction of 1.67 for a total correction factor of 2.17 (1.30 × 1.67) [75].

The $W \rightarrow l\nu$ cross section is calculated at LO and a k_{NNLO}-factor is applied to all generated W+jets processes. 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 [75]. The $W + jets$ samples get an additional k-factor correction of 1.30 with an added 1.47 correction for heavy flavor samples [76]. The background contribution of multijet production has been estimated from data and is discussed in section 6.2.3.

The signal branching ratios have been calculated for different Higgs boson masses using the HDECAY program [28]. The branching ratios times cross sections are included in Table 5.5 and Table 5.6 and the calculations for each can be found in Ref. [77, 78]. The MC samples are normalized to this NLO prediction.

Process	Mass Range [GeV]	$\sigma \times$ BR [pb]
$Z/\gamma^* \rightarrow \mu\mu$+0lp	15<M<75	338.18
$Z/\gamma^* \rightarrow \mu\mu$+1lp	15<M<75	40.02
$Z/\gamma^* \rightarrow \mu\mu$+2lp	15<M<75	10.04
$Z/\gamma^* \rightarrow \mu\mu$+3lp	15<M<75	2.76
$Z/\gamma^* \rightarrow \mu\mu$+0lp	75<M<130	133.34
$Z/\gamma^* \rightarrow \mu\mu$+1lp	75<M<130	40.29
$Z/\gamma^* \rightarrow \mu\mu$+2lp	75<M<130	9.99
$Z/\gamma^* \rightarrow \mu\mu$+3lp	75<M<130	3.09
$Z/\gamma^* \rightarrow \mu\mu$+0lp	130<M<250	0.86
$Z/\gamma^* \rightarrow \mu\mu$+1lp	130<M<250	0.37
$Z/\gamma^* \rightarrow \mu\mu$+2lp	130<M<250	0.095
$Z/\gamma^* \rightarrow \mu\mu$+3lp	130<M<250	0.032
$Z/\gamma^* \rightarrow \mu\mu$+0lp	250<M<1960	0.069
$Z/\gamma^* \rightarrow \mu\mu$+1lp	250<M<1960	0.034
$Z/\gamma^* \rightarrow \mu\mu$+2lp	250<M<1960	0.012
$Z/\gamma^* \rightarrow \mu\mu$+3lp	250<M<1960	0.0039
$Z/\gamma^* \rightarrow \mu\mu$+2b0lp	15<M<75	0.51
$Z/\gamma^* \rightarrow \mu\mu$+2b1lp	15<M<75	0.20
$Z/\gamma^* \rightarrow \mu\mu$+2b2lp	15<M<75	0.078
$Z/\gamma^* \rightarrow \mu\mu$+2b0lp	75<M<130	0.42
$Z/\gamma^* \rightarrow \mu\mu$+2b1lp	75<M<130	0.20
$Z/\gamma^* \rightarrow \mu\mu$+2b2lp	75<M<130	0.099
$Z/\gamma^* \rightarrow \mu\mu$+2b0lp	130<M<250	0.0034
$Z/\gamma^* \rightarrow \mu\mu$+2b1lp	130<M<250	0.0018
$Z/\gamma^* \rightarrow \mu\mu$+2b2lp	130<M<250	0.00088
$Z/\gamma^* \rightarrow \mu\mu$+2b0lp	250<M<1960	0.00034
$Z/\gamma^* \rightarrow \mu\mu$+2b1lp	250<M<1960	0.00017
$Z/\gamma^* \rightarrow \mu\mu$+2b2lp	250<M<1960	0.00011
$Z/\gamma^* \rightarrow \mu\mu$+2c0lp	15<M<75	4.14
$Z/\gamma^* \rightarrow \mu\mu$+2c1lp	15<M<75	0.95
$Z/\gamma^* \rightarrow \mu\mu$+2c2lp	15<M<75	0.34
$Z/\gamma^* \rightarrow \mu\mu$+2c0lp	75<M<130	0.93
$Z/\gamma^* \rightarrow \mu\mu$+2c1lp	75<M<130	0.55
$Z/\gamma^* \rightarrow \mu\mu$+2c2lp	75<M<130	0.28
$Z/\gamma^* \rightarrow \mu\mu$+2c0lp	130<M<250	0.0076
$Z/\gamma^* \rightarrow \mu\mu$+2c1lp	130<M<250	0.0044
$Z/\gamma^* \rightarrow \mu\mu$+2c2lp	130<M<250	0.0028
$Z/\gamma^* \rightarrow \mu\mu$+2c0lp	250<M<1960	0.00062
$Z/\gamma^* \rightarrow \mu\mu$+2c1lp	250<M<1960	0.00044
$Z/\gamma^* \rightarrow \mu\mu$+2c2lp	250<M<1960	0.00026

Table 5.1: $\sigma \times BR$ of generated events for background processes. The mass ranges listed give the generated invariant mass range of the Z/γ^* in the sample.

Process	Mass Range [GeV]	$\sigma \times$ BR [pb]
Z/γ^* → $\tau\tau$+0lp	15<M<75	336.57
Z/γ^* → $\tau\tau$+1lp	15<M<75	39.90
Z/γ^* → $\tau\tau$+2lp	15<M<75	9.94
Z/γ^* → $\tau\tau$+3lp	15<M<75	2.78
Z/γ^* → $\tau\tau$+0lp	75<M<130	133.17
Z/γ^* → $\tau\tau$+1lp	75<M<130	40.70
Z/γ^* → $\tau\tau$+2lp	75<M<130	10.01
Z/γ^* → $\tau\tau$+3lp	75<M<130	3.29
Z/γ^* → $\tau\tau$+0lp	130<M<250	0.88
Z/γ^* → $\tau\tau$+1lp	130<M<250	0.34
Z/γ^* → $\tau\tau$+2lp	130<M<250	0.099
Z/γ^* → $\tau\tau$+3lp	130<M<250	0.032
Z/γ^* → $\tau\tau$+0lp	250<M<1960	0.034
Z/γ^* → $\tau\tau$+1lp	250<M<1960	0.035
Z/γ^* → $\tau\tau$+2lp	250<M<1960	0.011
Z/γ^* → $\tau\tau$+3lp	250<M<1960	0.0039

Table 5.2: $\sigma \times BR$ of generated events for background processes. The mass ranges listed give the generated invariant mass range of the Z/γ^* in the sample.

Process	$\sigma \times$ BR [pb]
W(incl)+0lp	4597.68
W(incl)+1lp	1234.91
W(incl)+2lp	301.89
W(incl)+3lp	72.62
W(incl)+4lp	16.57X
W(incl)+5lp	5.01
W(incl)+2b0lp	9.49
W(incl)+2b1lp	4.16
W(incl)+2b2lp	1.61
W(incl)+2b3lp	0.75
W(incl)+2c0lp	23.37
W(incl)+2c1lp	13.49
W(incl)+2c2lp	5.53
W(incl)+2c3lp	2.41

Table 5.3: $\sigma \times BR$ of generated events for background processes.

Process	$\sigma \times$ BR [pb]
WW incl	11.66
WZ incl	3.45
ZZ incl	1.37
$t\bar{t}$ incl	7.88

Table 5.4: $\sigma \times BR$ of generated events for background processes.

Process	$\sigma \times$ BR [pb]
$gg \to H(115) \to WW^* \to ll$	0.00927
$gg \to H(120) \to WW^* \to ll$	0.01363
$gg \to H(125) \to WW^* \to ll$	0.01858
$gg \to H(130) \to WW^* \to ll$	0.02359
$gg \to H(135) \to WW^* \to ll$	0.02822
$gg \to H(140) \to WW^* \to ll$	0.03205
$gg \to H(145) \to WW^* \to ll$	0.03485
$gg \to H(150) \to WW^* \to ll$	0.03674
$gg \to H(155) \to WW^* \to ll$	0.03815
$gg \to H(160) \to WW^* \to ll$	0.03987
$gg \to H(165) \to WW^* \to ll$	0.03867
$gg \to H(170) \to WW^* \to ll$	0.03538
$gg \to H(175) \to WW^* \to ll$	0.03201
$gg \to H(180) \to WW^* \to ll$	0.02849
$gg \to H(185) \to WW^* \to ll$	0.02340
$gg \to H(190) \to WW^* \to ll$	0.01984
$gg \to H(195) \to WW^* \to ll$	0.01762
$gg \to H(200) \to WW^* \to ll$	0.01591
$q\bar{q} \to g\bar{g}H(115) \to qqWW^* \to ll$	0.0006
$q\bar{q} \to g\bar{g}H(120) \to qqWW^* \to ll$	0.0010
$q\bar{q} \to g\bar{g}H(125) \to qqWW^* \to ll$	0.0014
$q\bar{q} \to g\bar{g}H(130) \to qqWW^* \to ll$	0.0018
$q\bar{q} \to g\bar{g}H(135) \to qqWW^* \to ll$	0.0022
$q\bar{q} \to g\bar{g}H(140) \to qqWW^* \to ll$	0.0026
$q\bar{q} \to g\bar{g}H(145) \to qqWW^* \to ll$	0.0029
$q\bar{q} \to g\bar{g}H(150) \to qqWW^* \to ll$	0.0032
$q\bar{q} \to g\bar{g}H(155) \to qqWW^* \to ll$	0.0035
$q\bar{q} \to g\bar{g}H(160) \to qqWW^* \to ll$	0.0035
$q\bar{q} \to g\bar{g}H(165) \to qqWW^* \to ll$	0.0035
$q\bar{q} \to g\bar{g}H(170) \to qqWW^* \to ll$	0.0033
$q\bar{q} \to g\bar{g}H(175) \to qqWW^* \to ll$	0.0030
$q\bar{q} \to g\bar{g}H(180) \to qqWW^* \to ll$	0.0027
$q\bar{q} \to g\bar{g}H(185) \to qqWW^* \to ll$	0.0023
$q\bar{q} \to g\bar{g}H(190) \to qqWW^* \to ll$	0.0020
$q\bar{q} \to g\bar{g}H(195) \to qqWW^* \to ll$	0.0018
$q\bar{q} \to g\bar{g}H(200) \to qqWW^* \to ll$	0.0016

Table 5.5: $\sigma \times BR$ of generated events for signal processes.

Process	$\sigma \times$ BR [pb]
ZH(115) → WW, ZZ, $\gamma\gamma$, ee, $\mu\mu$, $\tau\tau$, γZ incl.	0.01053
ZH(120) → WW, ZZ, $\gamma\gamma$, ee, $\mu\mu$, $\tau\tau$, γZ incl.	0.01007
ZH(125) → WW, ZZ, $\gamma\gamma$, ee, $\mu\mu$, $\tau\tau$, γZ incl.	0.00985
ZH(130) → WW, ZZ, $\gamma\gamma$, ee, $\mu\mu$, $\tau\tau$, γZ incl.	0.00998
ZH(135) → WW, ZZ, $\gamma\gamma$, ee, $\mu\mu$, $\tau\tau$, γZ incl.	0.00979
ZH(140) → WW, ZZ, $\gamma\gamma$, ee, $\mu\mu$, $\tau\tau$, γZ incl.	0.00959
ZH(145) → WW, ZZ, $\gamma\gamma$, ee, $\mu\mu$, $\tau\tau$, γZ incl.	0.00934
ZH(150) → WW, ZZ, $\gamma\gamma$, ee, $\mu\mu$, $\tau\tau$, γZ incl.	0.00887
ZH(155) → WW, ZZ, $\gamma\gamma$, ee, $\mu\mu$, $\tau\tau$, γZ incl.	0.00869
ZH(160) → WW, ZZ, $\gamma\gamma$, ee, $\mu\mu$, $\tau\tau$, γZ incl.	0.00816
ZH(165) → WW, ZZ, $\gamma\gamma$, ee, $\mu\mu$, $\tau\tau$, γZ incl.	0.00720
ZH(170) → WW, ZZ, $\gamma\gamma$, ee, $\mu\mu$, $\tau\tau$, γZ incl.	0.00626
ZH(175) → WW, ZZ, $\gamma\gamma$, ee, $\mu\mu$, $\tau\tau$, γZ incl.	0.00587
ZH(180) → WW, ZZ, $\gamma\gamma$, ee, $\mu\mu$, $\tau\tau$, γZ incl.	0.00530
ZH(185) → WW, ZZ, $\gamma\gamma$, ee, $\mu\mu$, $\tau\tau$, γZ incl.	0.00496
ZH(190) → WW, ZZ, $\gamma\gamma$, ee, $\mu\mu$, $\tau\tau$, γZ incl.	0.00432
ZH(195) → WW, ZZ, $\gamma\gamma$, ee, $\mu\mu$, $\tau\tau$, γZ incl.	0.00406
ZH(200) → WW, ZZ, $\gamma\gamma$, ee, $\mu\mu$, $\tau\tau$, γZ incl.	0.00352
WH(115) → WW, ZZ, $\gamma\gamma$, ee, $\mu\mu$, $\tau\tau$, γZ incl.	0.01945
WH(120) → WW, ZZ, $\gamma\gamma$, ee, $\mu\mu$, $\tau\tau$, γZ incl.	0.01746
WH(125) → WW, ZZ, $\gamma\gamma$, ee, $\mu\mu$, $\tau\tau$, γZ incl.	0.01734
WH(130) → WW, ZZ, $\gamma\gamma$, ee, $\mu\mu$, $\tau\tau$, γZ incl.	0.01729
WH(135) → WW, ZZ, $\gamma\gamma$, ee, $\mu\mu$, $\tau\tau$, γZ incl.	0.01741
WH(140) → WW, ZZ, $\gamma\gamma$, ee, $\mu\mu$, $\tau\tau$, γZ incl.	0.01609
WH(145) → WW, ZZ, $\gamma\gamma$, ee, $\mu\mu$, $\tau\tau$, γZ incl.	0.01574
WH(150) → WW, ZZ, $\gamma\gamma$, ee, $\mu\mu$, $\tau\tau$, γZ incl.	0.01550
WH(155) → WW, ZZ, $\gamma\gamma$, ee, $\mu\mu$, $\tau\tau$, γZ incl.	0.01521
WH(160) → WW, ZZ, $\gamma\gamma$, ee, $\mu\mu$, $\tau\tau$, γZ incl.	0.01411
WH(165) → WW, ZZ, $\gamma\gamma$, ee, $\mu\mu$, $\tau\tau$, γZ incl.	0.01285
WH(170) → WW, ZZ, $\gamma\gamma$, ee, $\mu\mu$, $\tau\tau$, γZ incl.	0.01154
WH(175) → WW, ZZ, $\gamma\gamma$, ee, $\mu\mu$, $\tau\tau$, γZ incl.	0.00969
WH(180) → WW, ZZ, $\gamma\gamma$, ee, $\mu\mu$, $\tau\tau$, γZ incl.	0.00850
WH(185) → WW, ZZ, $\gamma\gamma$, ee, $\mu\mu$, $\tau\tau$, γZ incl.	0.00810
WH(190) → WW, ZZ, $\gamma\gamma$, ee, $\mu\mu$, $\tau\tau$, γZ incl.	0.00712
WH(195) → WW, ZZ, $\gamma\gamma$, ee, $\mu\mu$, $\tau\tau$, γZ incl.	0.00570
WH(200) → WW, ZZ, $\gamma\gamma$, ee, $\mu\mu$, $\tau\tau$, γZ incl.	0.00548

Table 5.6: $\sigma \times BR$ of generated events for associated production signal processes.

Chapter 6

Event Selection

Due to the very small signal production cross section relative to the total inelastic cross section (∼12 orders of magnitude smaller) and even to the individual production cross sections of typical background processes, selection requirements have to be applied to suppress the background and enrich the relative signal contribution in the data sample. The goal when choosing selection criteria is 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. To verify this, detailed comparisons of data to the simulated MC samples have been performed with special attention to the level of agreement at each cut-stage. When necessary, corrections to the simulation are applied in such a way as to avoid biasing the signal region, while enhancing the agreement with data. The following sections describe these comparisons and corrections along with the chosen cutflow, choice of cuts and their order, of the analysis.

This analysis is interested in events with a final state consistent with $H \rightarrow W^+W^- \rightarrow \mu^+\nu\mu^-\nu$. Therefore, events are selected with at least two high p_T muons and large missing transverse energy (indicative of a neutrino). Only loose cuts are applied to achieve maximum signal acceptance. After all cuts, a multivariate technique is used to enhance sensitivity(see Chapter 7). Table 6.1 lists all selection requirements. The individual cuts are explained in more detail in the sections following the table.

Cut 0	Preselection	muon ID passed for two leptons two muons with opposite charge $p_T^{\mu\_1} > 20$ GeV $p_T^{\mu\_2} > 10$ GeV isolated muons $M_{\mu\mu} > 15$ GeV min(ΔR(μ_1,any jet)) > 0.1 min(ΔR(μ_2,any jet)) > 0.1 primary vertex z cut of < 60cm
Cut 1	$\Delta\phi(\mu_1, \mu_2)$	< 2.5
Cut 2	Missing Transverse Energy $\vec{\not\!\!E}_T$	> 25.0 GeV
Cut 3	Minimum Transverse Mass M_T^{min}	> 20 GeV

Table 6.1: Summary of the event selection criteria.

6.1 Preselection

A preselection cut level is defined to create a reasonably sized sample where data/MC comparisons can be performed for the background samples. This becomes much more difficult at the final cut stage where statistics for some of the backgrounds are low[1]. Some of the preselection cuts are also designed to limit us to regions of the detector that are best understood to limit our uncertainty due to detector effects. First, we require at least one primary vertex(PV) with a minimum of two tracks attached. A cut of ± 60 cm is placed on the PV's location along the beam axis. Outside of this region it becomes increasingly difficult to determine the location of the PV with any confidence due to things like lower quality tracks for reconstruction.

We also require two loose quality muons with $p_T^{\mu\_1} > 20$ GeV, $p_T^{\mu\_2} > 10$ GeV

---

[1] Of course, this is intentional as we want to get rid of as much background as possible and thus get the highest signal to background ratio we can.

and $|\eta^{det}| < 2$ [57, 58]. These values are chosen to be as low as possible and still have near 100% trigger efficiency for single muons and dimuons. Even though these triggers are not explicitly required, the majority of events come from either dimuon or single muon triggers. Both muons must be matched to central tracks of at least "medium" quality to ensure a good track momentum resolution. The isolation for the leading muon (i.e. highest p_T) must be less than 0.4 to avoid identifying jets as muons and to rule out muons coming from jets (such as jets from b quarks which can eventually decay to muons) without sacrificing acceptance. The second leading muon must have isolation less than 0.5 for similar reasons(see section 4.3.2 for more on these requirements). The muons must have opposite charges. Additionally, the invariant mass of the dilepton system must be greater than 15 GeV. This allows us to further suppress any possible trigger effects, while also allowing us to neglect the Upsilon background which peaks at ~ 9.5 GeV [34] in invariant mass and tapers off quickly. Finally, the minimum distance between one of the two leading muons and a jet is made to be at least 0.1 when there is a jet in the event to ensure that an object is not counted as both a jet and a muon[2].

6.2 Monte Carlo Corrections

6.2.1 Reweightings

Multiple reweightings to correct several issues in the MC modeling of our data are also applied in this analysis as described in the sections below. They are an extension of the simulation process allowing us to create samples which more closely model the data. We apply both default reweightings (as implemented in the standard D0 analysis software) and some additional reweightings for our particular data sample. For each of these additional reweightings, a ratio histogram

---

[2] The isolation requirements effectively do this and the addition of this cut has a very small effect.

is generated by dividing the relevant data distribution by the MC. Then the ratio histogram is smoothed and rebinned into semi-equalized probability bins, where each bin has at least half of the statistics of the bin with the most entries(see Appendix B). This allows us to reweight without having large effects from statistical fluctuations and without relying on an arbitrary function to model the ratio.

Default Reweightings and Improvements

Several “default” reweightings as implemented in the standard `vjets_cafe` package [79] are applied in this analysis. They make small corrections to some distributions not perfectly modeled by the MC generators and detector simulation. These are the z-coordinate of the primary vertex, the instantaneous luminosity, and the Z_{p_T}[3].

In general these default reweightings do a good job of improving the data/MC agreement. For example, the z-coordinate of the PV is well modeled after applying its reweighting. Some of the reweightings are less successful mostly due to being derived on a data sample with vastly different cuts applied and further adjustment is needed. This is evident in the default instantaneous luminosity reweighting, which does not yield adequate agreement, leading to a necessary additional reweighting to data in this variable. Since we use ALPGEN to generate our $Z + jets$ background and the default Z_{p_T} reweighting was derived using it as well, this reweighting does a reasonable job of obtaining good data/MC agreement. The reweighting was derived in the Z-peak region of the invariant mass spectrum, in bins of the number of jets.

The number of SMT hits is also reweighted for each muon. This is done to allow a reweighting of the track χ^2 where significant disagreement is observed. This track reweighting was performed mostly to investigate whether differences

---

[3] The Z_{p_T} referred to here is the transverse momentum of the dimuon system which came from the decay of a Z boson. The reweighting is only applied to events that contain exactly one Z and no other bosons.

between data and MC in the track resolution could propagate via the muon energy loss correction to the missing momentum determination and therefore explain the data/MC disagreement observed for this variable. Since the two variables are highly correlated, the track χ^2 reweighting is done in bins of SMT hits (0, 1-6, 7+).

There are correlations between many of these reweightings which are difficult to take into account. The goal of these reweightings is to correct the major discrepancies and mis-modeled regions and to do a multi-dimensional reweighting would be unreasonably hard compared to the small improvement it would yield. The differences in these distributions after all reweightings are due to the correlations and the serial manner in which the reweightings are applied (one after another). The differences are easily covered by the systematic uncertainties (Chapter 8). In Fig. 6.1 - 6.5 the distributions just discussed can be seen before and after all reweightings. In these figures and all of those that follow, the legend lists the various background processes along with the signal times a factor of 10. In the legends, the light grey is Z+jets, teal is diboson, grey is W+jets, brown is multijet(QCD), the darker blue is $t\bar{t}$, and the red line is the signal times 10.

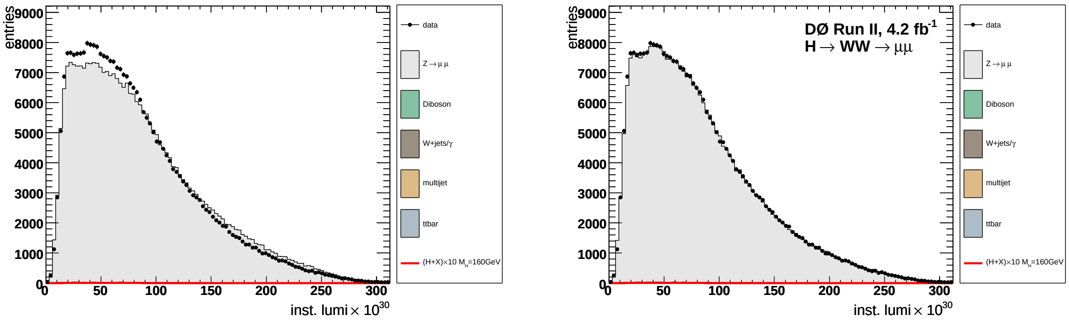

Figure 6.1: Instantaneous luminosity at preselection before(left) and after new reweightings(right).

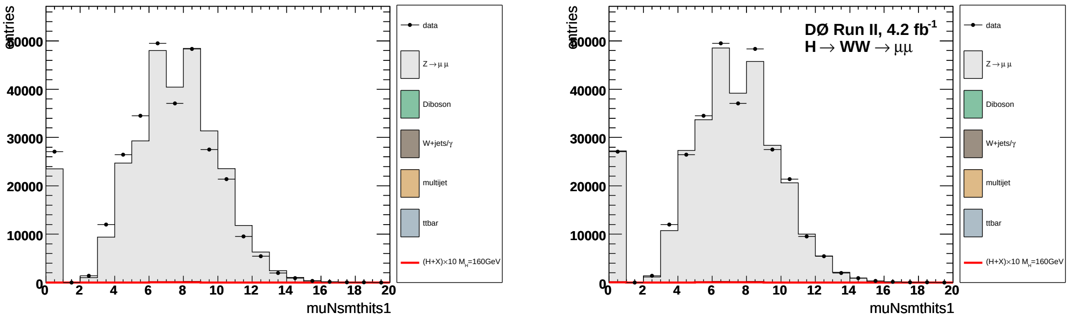

Figure 6.2: Number of SMT hits for the leading muon at preselection before(left) and after new reweightings(right).

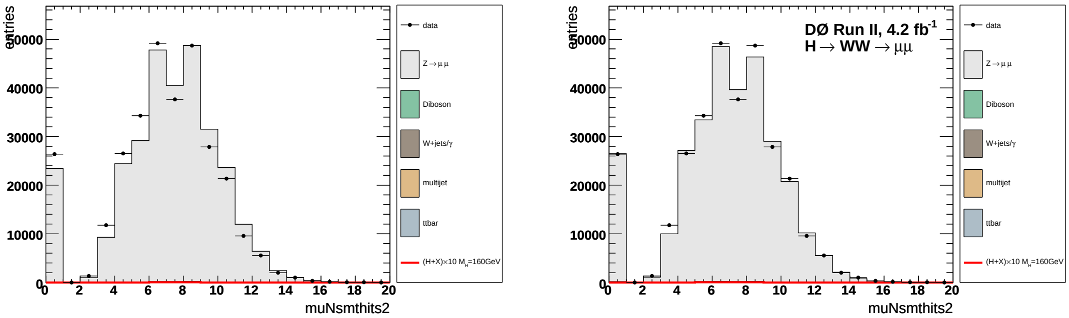

Figure 6.3: Number of SMT hits for the second leading muon at preselection before(left) and after new reweightings(right).

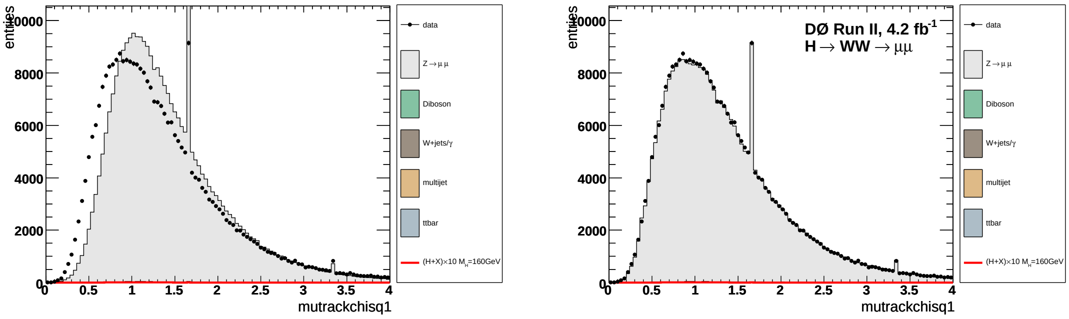

Figure 6.4: Track χ^2 distribution for the leading muon at preselection before(left) and after new reweightings(right). The spikes are a binning effect due to the bit packing in the thumbnail storage file, resulting in a finite numerical granularity of the χ^2 variable.

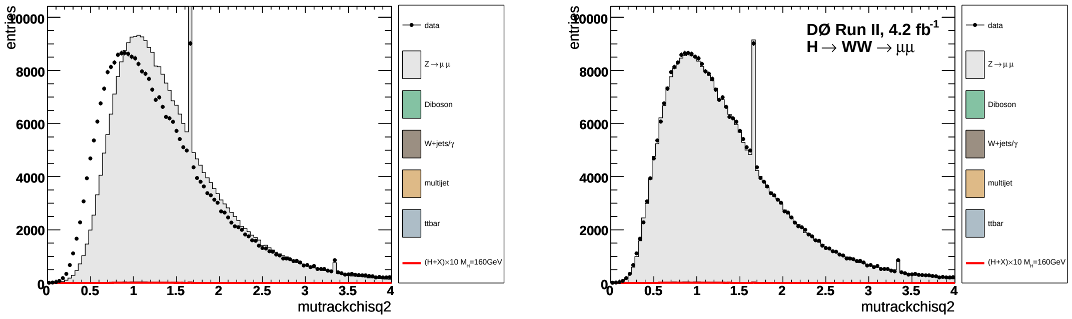

Figure 6.5: Track χ^2 distribution for the second leading muon at preselection before(left) and after new reweightings(right). The spikes are a binning effect due to the bit packing in the thumbnail storage file, resulting in a finite numerical granularity of the χ^2 variable.

Trigger Effect Reweightings

Another issue which affects the overall agreement between data and MC is a trigger bias visible in the η distributions of the muons. Looking at events that come in exclusively on a single muon, dimuon, or muon+jet trigger one can see that the η shapes and consequently the rapidity of the dimuon system are significantly different. By reweighting successively the events based on the rapidity of the two muons and then the rapidity of the dimuon system one can account for this bias. Again, while ideally a multi-dimensional reweighting would be applied to account for correlations, it is significantly more complex and should not change the result within the systematic uncertainties of the analysis. The distributions before and after reweighting can be found in Fig. 6.6 - 6.8.

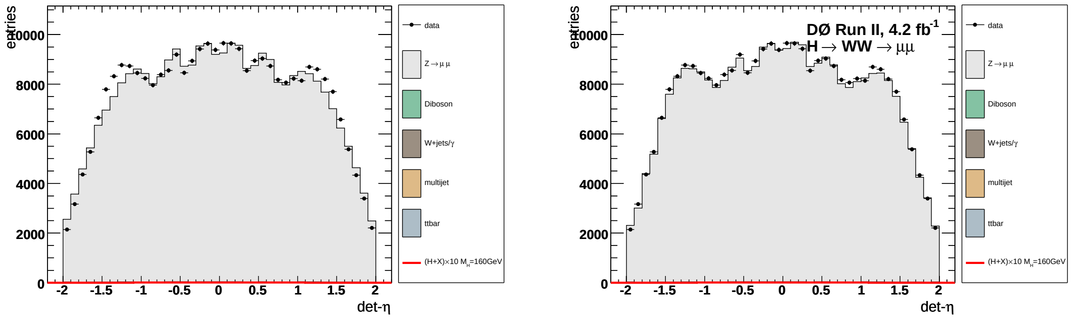

Figure 6.6: Detector η distribution for the leading muon at preselection before reweightings(left) and after reweightings(right).

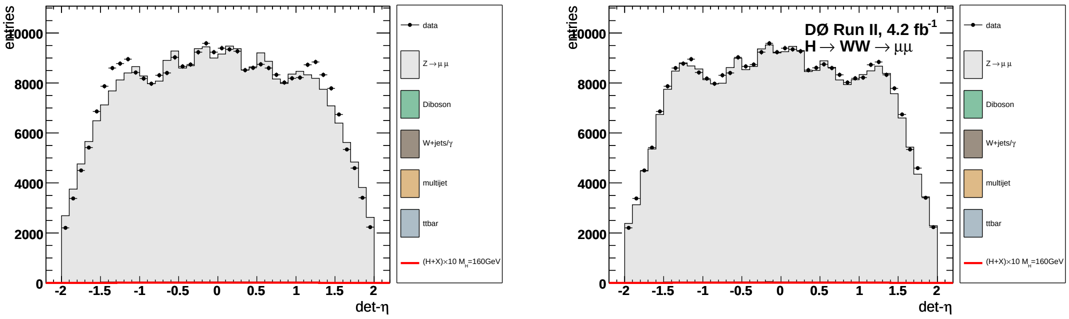

Figure 6.7: Detector η distribution for the second leading muon at preselection before reweightings(left) and after reweightings(right).

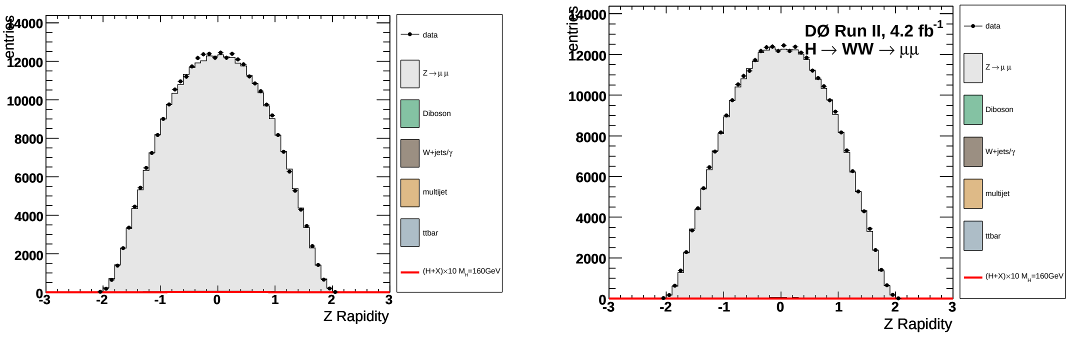

Figure 6.8: Rapidity distribution for the di-muon system at preselection before reweightings(left) and after reweightings(right).

Extra Muon Smearing

The muon momentum resolution from the simulated detector is better than what we actually measure in data(section 4.3). To account for this discrepancy, the muon ID group has created an over-smearing to be applied to the simulated muons. [80, 81] It was observed that the muon MC over-smearing that they provide does not adequately account for this effect. This analysis is particularly sensitive to the muon momentum resolution since the missing transverse energy in events without jets primarily comes from mis-measured muons.

The tuning of the extra muon smearing was performed at the preselection level. It was optimized by comparing (bin-by-bin) the data and MC histograms of the invariant mass between 50 and 130 GeV and to the transverse momentum histograms of the first and second muon. The parameters were chosen by minimizing the χ^2 of this comparison.

The first step was to correct the muon momentum scale in MC. This was done by making the mean of the Z peak agree in data and MC. This is a 0.4% effect that has been ignored in the standard muon over-smearing. After this, a second random gaussian was added to the resolution term used in the standard muon over-smearing. This gaussian was fixed to have the same mean, but be wider than the existing gaussian used in the standard over-smearing. This wider gaussian was added to 1% of muons in the sample. The invariant mass distribution before and after the extra muon smearing can be seen in Fig. 6.9. The transverse momentum distributions of the first and second muons before and after smearing can be found in Fig. 6.10 and Fig. 6.11.

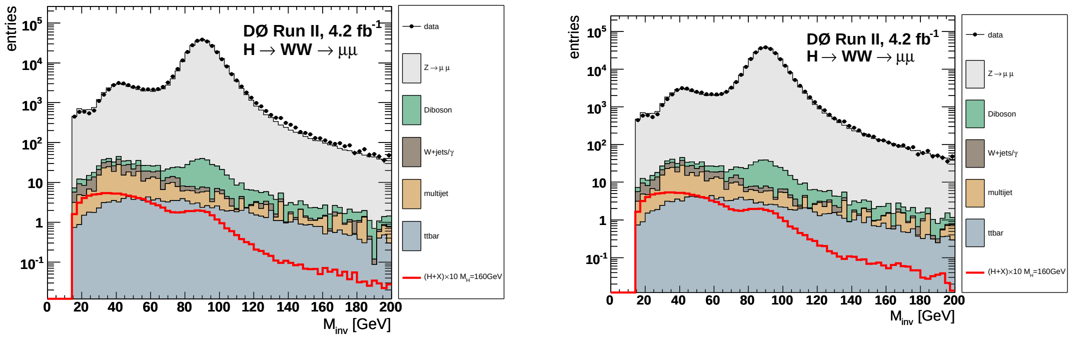

Figure 6.9: The invariant mass distribution after preselection before the additional muon smearing (left) and after the additional muon smearing (right).

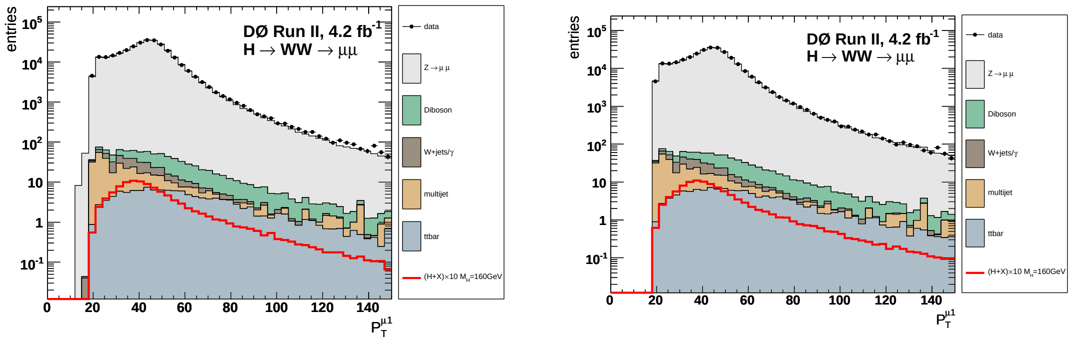

Figure 6.10: The transverse momentum distribution of the leading muon after preselection before the additional muon smearing (left) and after the additional muon smearing (right).

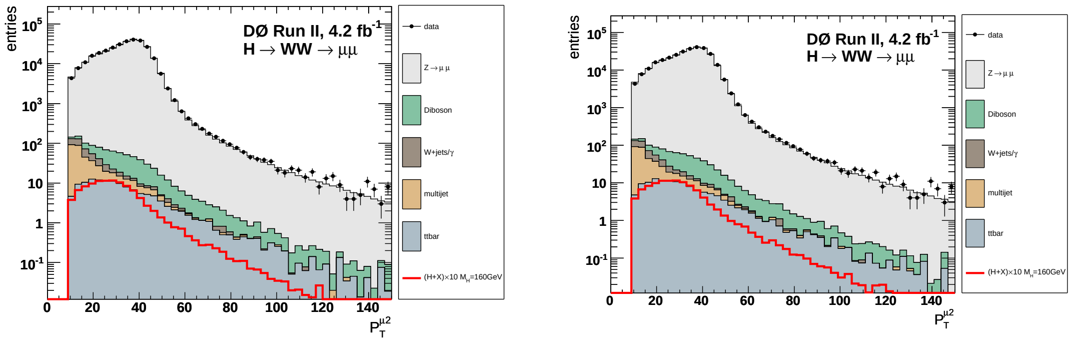

Figure 6.11: The transverse momentum distribution of the second leading muon after preselection before the additional muon smearing (left) and after the additional muon smearing (right).

Generator-Related Reweightings

There are two reweightings stemming from the use of the Pythia event generator in this analysis. In both cases, other event generators are used to improve the description of important kinematic variables. These reweightings are discussed in the following paragraphs.

The WW background is mainly produced via the quark-antiquark annihilation process $q\bar{q} \rightarrow WW$ but the $gg \rightarrow WW$ gluon-induced contribution may contribute significantly although suppressed by two orders of α_s (see [82]). In particular, the $\Delta\phi(\mu_1, \mu_2)$ distribution differs significantly and may even be enhanced by selection cuts relative to the $q\bar{q}$ process. Although more severe at the Large Hadron Collider, due to the high gluon-flux, this effect has been studied for the present analysis. Since the gluon-fusion process is not taken into account in Pythia, the GG2WW [83] event generator is used to derive a correction on the parton level. The correction is applied to the WW background process and an additional systematic uncertainty on the background prediction is included.

A harder $p_T(WW)$ energy spectrum can lead to a boost of the diboson system, which in turn could effect the angular distribution of the leptons decaying from them. This is a very important discriminating variable between the standard model background processes and the Higgs signal. Therefore, we compare the modeling of the p_T of the WW system using the p_T distributions of several simulated event generators, namely, Pythia, Sherpa [84], and MC@NLO [85]. The Sherpa and MC@NLO event generators agree well with each other and generate slightly harder p_T spectra than Pythia, so we reweight the Pythia p_T spectra in the WW background and $H \rightarrow WW$ signal samples to these spectra and include this difference as a systematic uncertainty in the analysis (see section 8.2). The $p_T(WW)$ spectra of all three event generators are shown in Fig. 6.12.

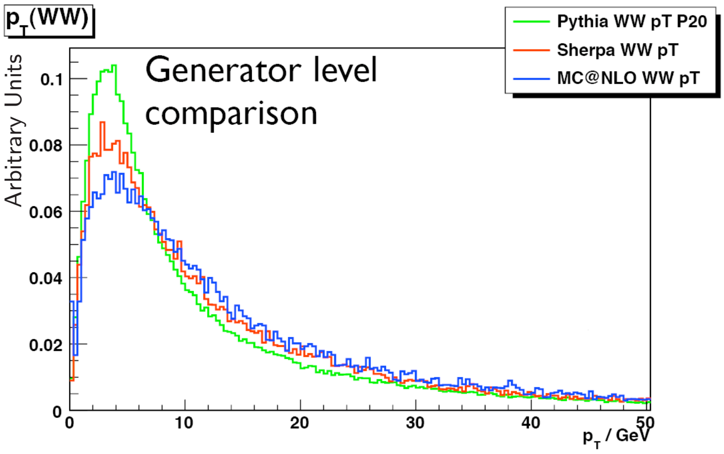

Figure 6.12: Spectra of the momentum of the WW system, $p_T(WW)$, as simulated by Pythia, Sherpa, and MC@NLO. All distributions are normalized to the same area.

6.2.2 DATA/MC Normalization

The best way to get the proper normalization for the MC samples used in the analysis is to see how we reproduce a well understood physical processes. One such process is the decay of the Z boson to two muons. We know the invariant mass of the Z peaks at ~ 91 GeV[4]. This allows us to define a window to require agreement in where we know the major contributing background. Another reason to look at this process is that the Z peak region is heavily dominated by Z boson decays where the resulting two muons have a wide opening angle (are back-to-back) and our signal is an insignificant contribution to this region. The two W bosons in our signal, which would be required to make two muons, are not necessarily back-to-back themselves and thus the $\Delta\phi(\mu_1, \mu_2)$distribution of the muons is roughly flat (not taking into account any preferential configuration of the W bosons). This means that the invariant mass of the two muons from our signal tends to be relatively low, mostly below the Z peak region. Using this region allows us to ensure agreement between data and MC at preselection without biasing our signal.

The global normalization used in this analysis is found by normalizing the sum

---

[4] The latest world average mass measurement is 91.1876 ± 0.0021 GeV according to [2]

of all MC samples to data in the Z peak, as defined by an invariant mass window from 80 GeV to 100 GeV. The normalization has been performed by counting the number of events in this invariant mass range, then applying the ratio as an event weight to MC events. The normalization factor for the RunIIa (RunIIb) sample is found to be 0.92 (0.83)[5]. This means that after normalizing the MC using the luminosity of the respective data samples there is still "too much" MC and thus the samples must be scaled down by this factor. The actual effect is in fact due to inefficiencies in data which make us think we have more luminosity than we do.

There are several sources for these inefficiencies. One source is the rejection of bad calorimeter data event flags. We do not calculate the effect of rejecting these bad calorimeter events in the luminosity, but instead absorb this effect in the normalization factor. The size of this effect has been shown to be about 3% in other analyses. Furthermore, since we use a non-standard isolation definition, we are not applying a data/MC scale factor for the isolation requirement. (Note that we do apply the standard data/MC scale factors for the muon ID [57, 58] and central tracking requirements.) Not applying an isolation scale factor was found to have a 2% effect on the normalization.

The remaining isolation factor is believed to be due to the trigger. Although we do not impose an explicit trigger requirement, the sum of all triggers is not fully efficient for our event selection. This is the primary reason for the large difference in normalization factor between RunIIa and RunIIb data samples. A study was performed that showed imposing the single muon trigger OR requirement in the RunIIb data resulted in a normalization factor that was consistent with unity. Furthermore, a large difference in the normalization factor is observed between data using the v15 trigger list (used in the first ∼2/3 of the data) and data using the v16 trigger list. This can be understood since many triggers (particularly

---

[5] Note that the normalization factors listed here are those found after all the reweightings and additional muon smearing described in the rest of this section. The additional reweightings and muon smearing has an effect of approximately 2% on the normalization factors.

muon triggers) have tighter requirements due to increasing luminosities in v16 as compared to v15 thus decreasing acceptance.

6.2.3 Estimation of Multijet Background

The procedure for creating the QCD background sample is performed after all of the other simulated backgrounds have been properly normalized and reweighted. First, to account for charge flips, the like-sign MC samples are scaled up to the data in the high p_T region of the trailing muon p_T distribution. In this region the QCD contribution is assumed to be very small as the probability for a jet to fake a lepton decreases with increasing p_T due to the decreasing probability that a higher p_T jet would pass the isolation criteria. Then these scaled up backgrounds are subtracted from the like-sign data. Next, the ratio of this background-subtracted like-sign data to the total like-sign data sample is determined as a function of p_T. Finally, all of the like-sign data events are used to increase statistics, and are normalized by this calculated ratio to obtain the proper amount of QCD relative to the other backgrounds. Due to limited statistics we assign a 30% systematic error on this QCD estimation. The overall QCD contribution at preselection level is small due to the muon isolations required at the preselection stage which make a jet less likely and the selection of oppositely charged leptons which essentially cut out half of the fakes.

6.2.4 Control Plots at Preselection

Control plots of the most important distributions at preselection (Cut0) can be found in Figs. 6.13 - 6.25 after all reweightings and normalization. Plots at other cut stages are in the next section.

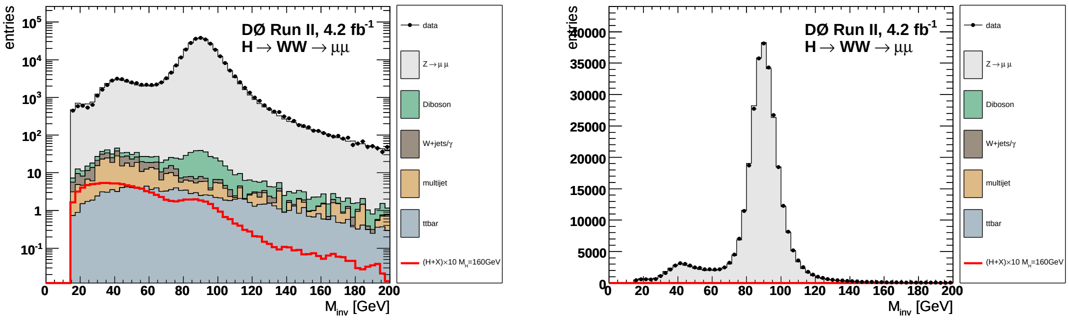

Figure 6.13: Invariant mass distribution at preselection in log scale(left) and linear scale(right).

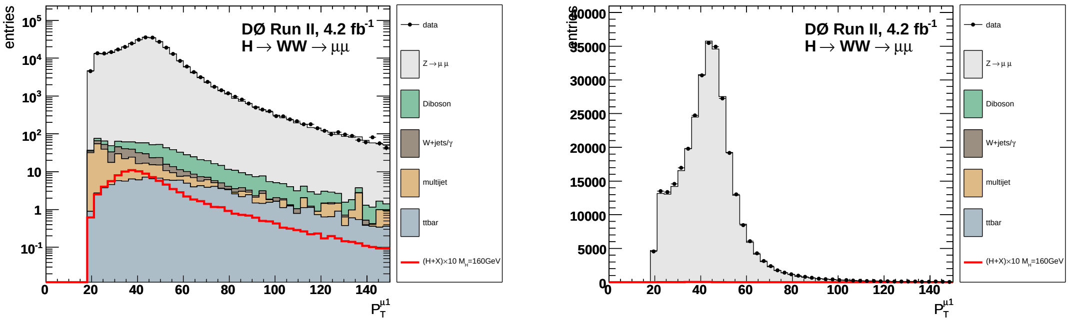

Figure 6.14: Transverse momentum of the leading muon at preselection in log scale(left) and linear scale(right).

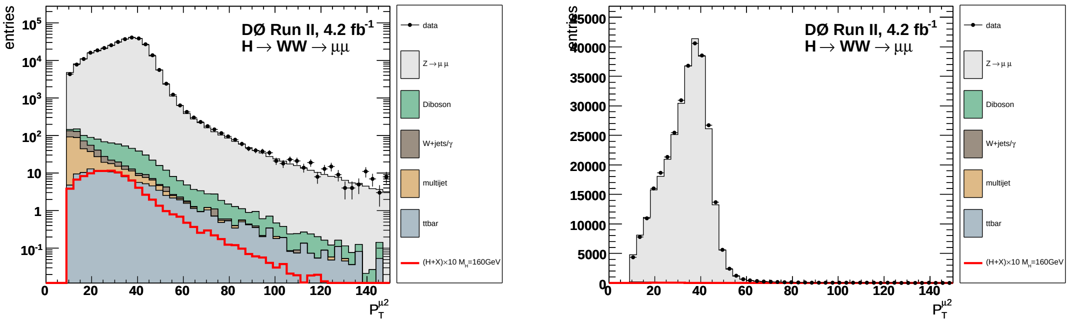

Figure 6.15: Transverse momentum of the second leading muon at preselection in log scale(left) and linear scale(right).

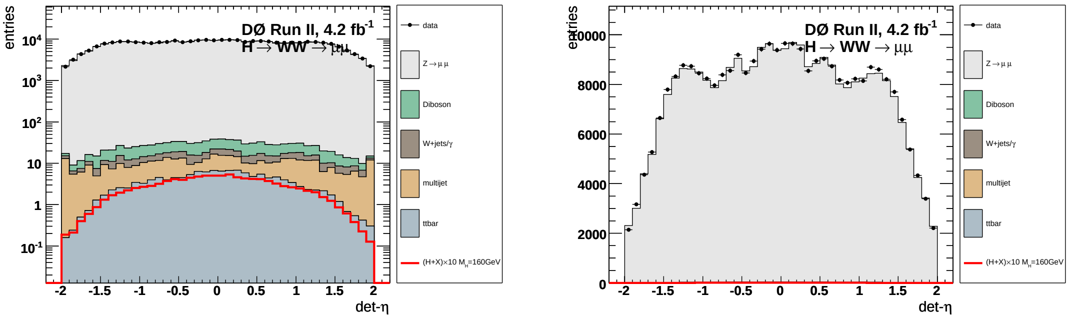

Figure 6.16: Detector η of the leading muon at preselection in log scale(left) and linear scale(right).

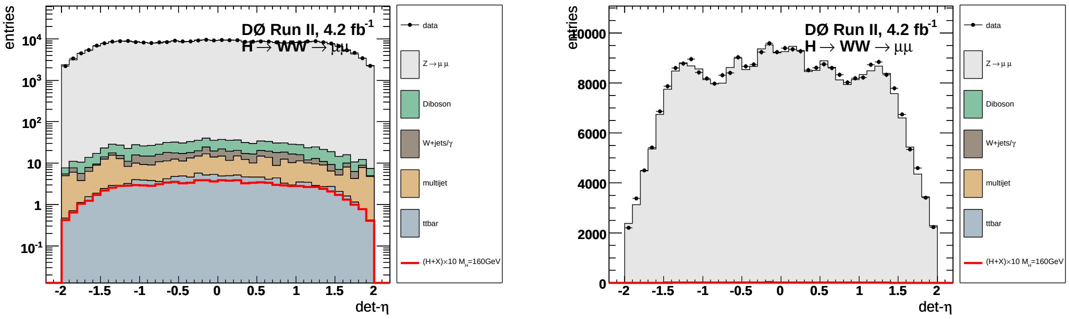

Figure 6.17: Detector η of the second leading muon at preselection in log scale(left) and linear scale(right).

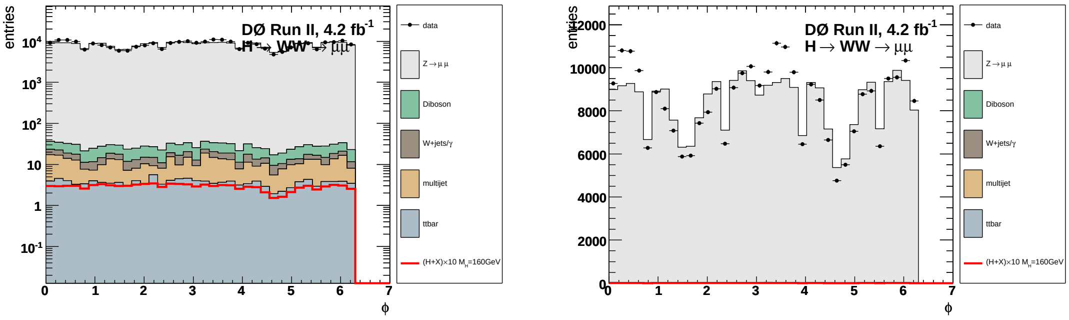

Figure 6.18: ϕ of the leading muon at preselection in log scale(left) and linear scale(right).

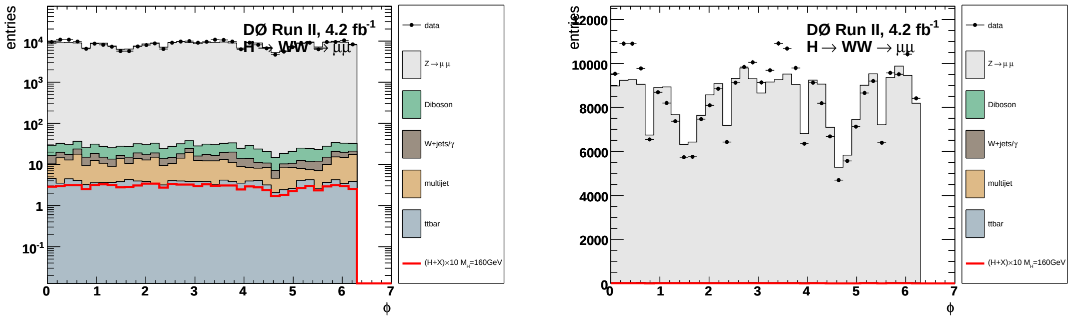

Figure 6.19: ϕ of the second leading muon at preselection in log scale(left) and linear scale(right).

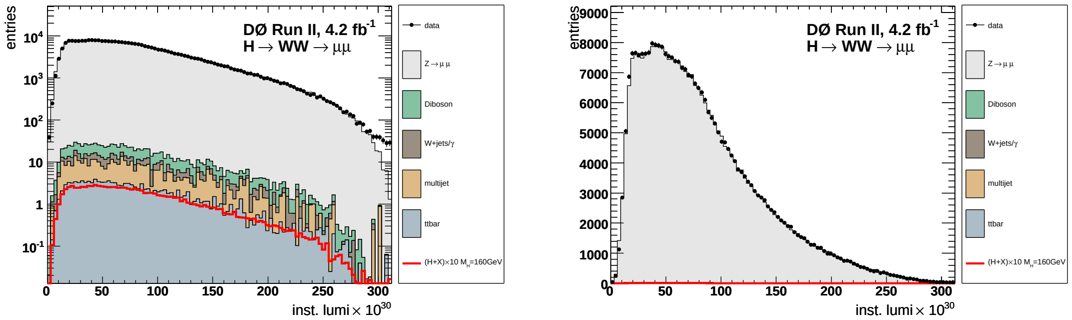

Figure 6.20: Instantaneous luminosity at preselection in log scale(left) and linear scale(right).

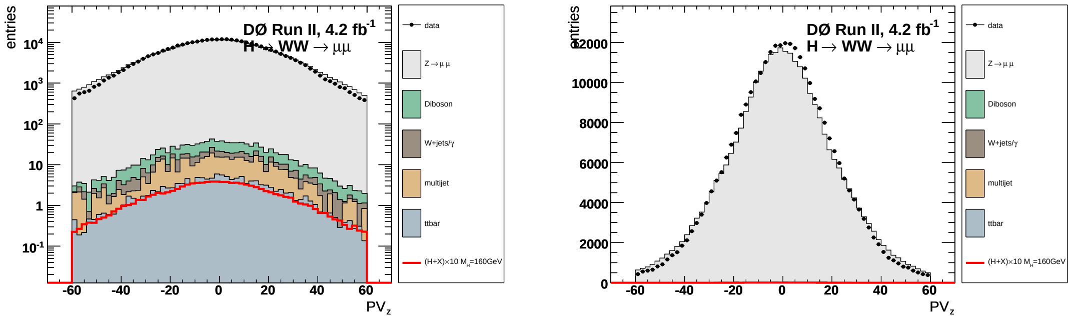

Figure 6.21: Primary vertex z at preselection in log scale(left) and linear scale(right).

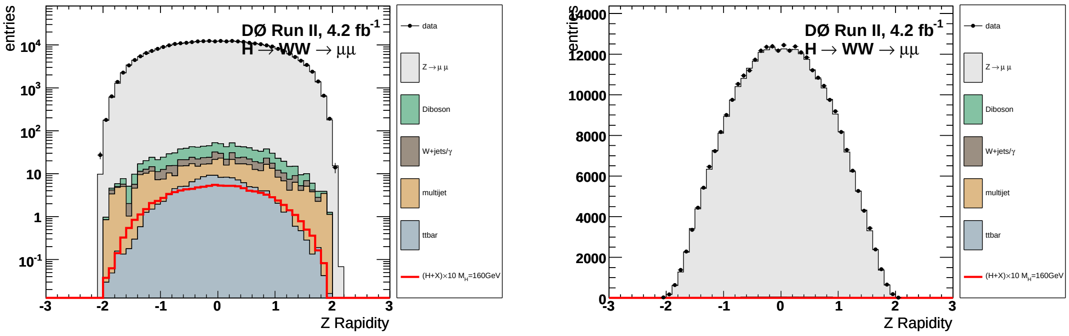

Figure 6.22: Rapidity of the di-muon system at preselection in log scale(left) and linear scale(right).

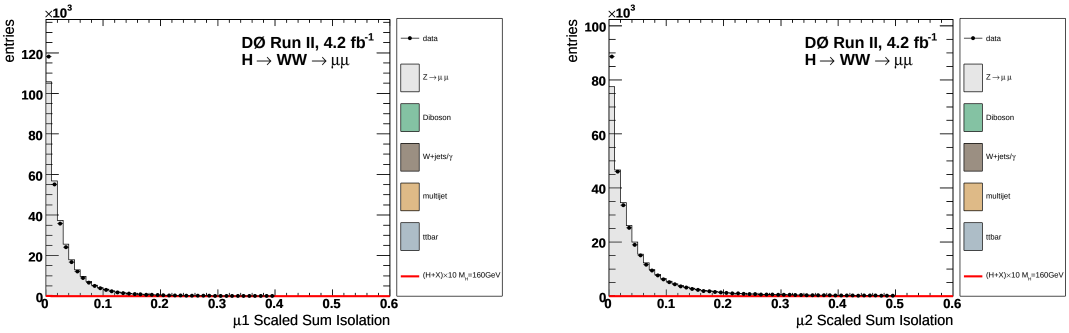

Figure 6.23: Sum of the scaled muon isolations (track and calorimeter) for the leading(left) and second leading(right) muons at preselection.

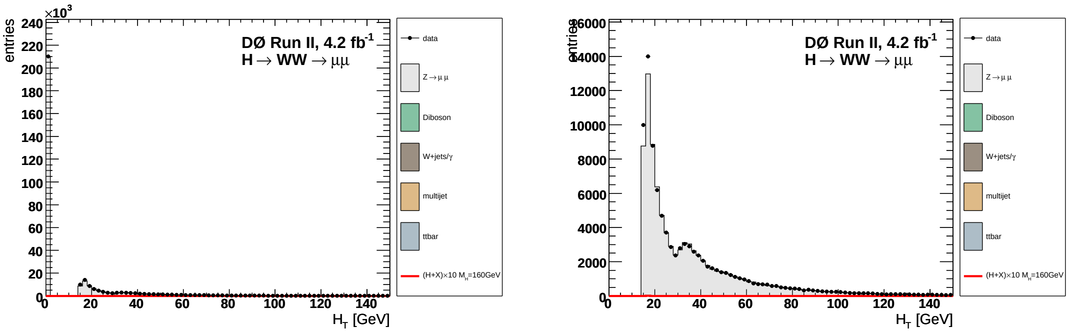

Figure 6.24: htall(left) and htall for events with at least 1 jet(right) at preselection.

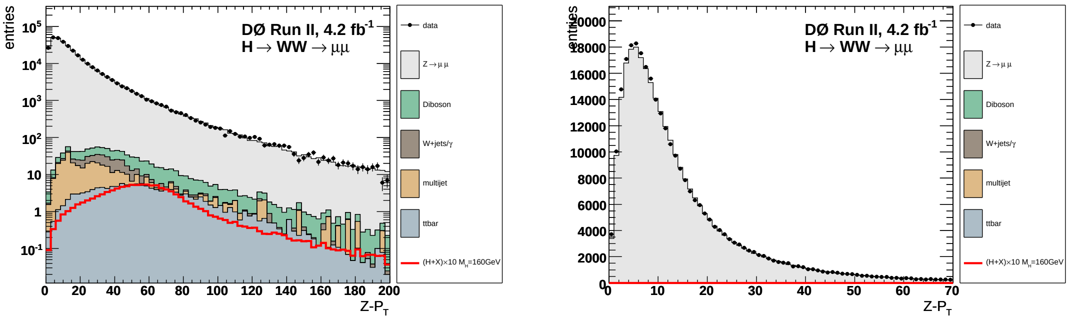

Figure 6.25: Z_{p_T} at preselection in log scale(left) and linear scale(right).

6.3 Final Selection

The goal of any analysis searching for a small signal in a large background is to increase the signal to background ratio as much as possible. This can be achieved by cutting away regions of phase space that contain large background to signal ratios as well as applying multivariate techniques such as the Neural Network(NN)(Chapter 7) in this analysis. Our final selection is composed of three kinematic cuts. The first is a requirement on the opening angle between the two selected muons of $\Delta\phi(\mu_1, \mu_2)$< 2.5. This removes a large amount of the remaining $Z \rightarrow \ell\ell$ events since the leptons decaying from a Z boson tend to be back-to-back, while being largely independent of the particular Higgs mass kinematics. By reducing the dominating $Z + jets$ background prior to the application of the NN, the training for the neural network improves. This is due mainly to the fact that the relative weight of the invariant mass becomes more balanced with respect to the other input variables for the neural network. The $\Delta\phi(\mu_1, \mu_2)$ distribution after preselection and before cutting on $\Delta\phi(\mu_1, \mu_2)$ can be seen in Fig. 6.26.

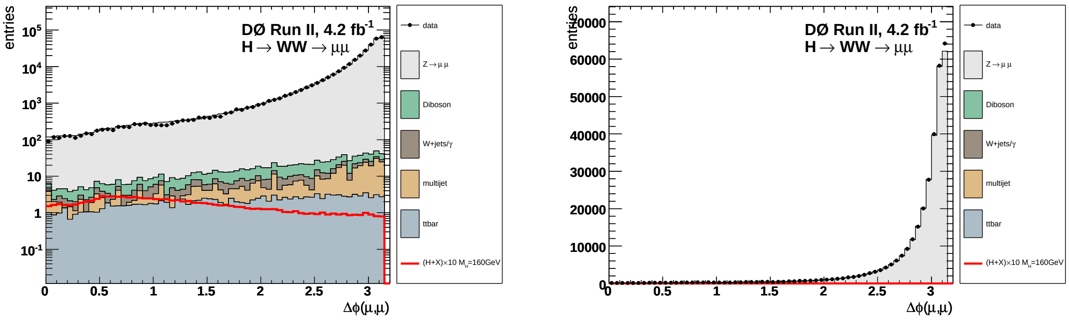

Figure 6.26: $\Delta\phi(\mu_1, \mu_2)$ distribution at preselection in log scale(left) and linear scale(right).

We can determine an optimal place to make this cut on $\Delta\phi(\mu_1, \mu_2)$ from looking at a distribution constructed by integrating the events below successive cut values and calculating the ratio of the signal squared over the sum of the backgrounds. This S^2/B distribution, found in Fig. 6.27, peaks at the optimal cut value for a cut-based analysis. In this analysis we make a looser cut to gain signal acceptance and then exploit the NN, described in the next chapter, to gain additional separation.

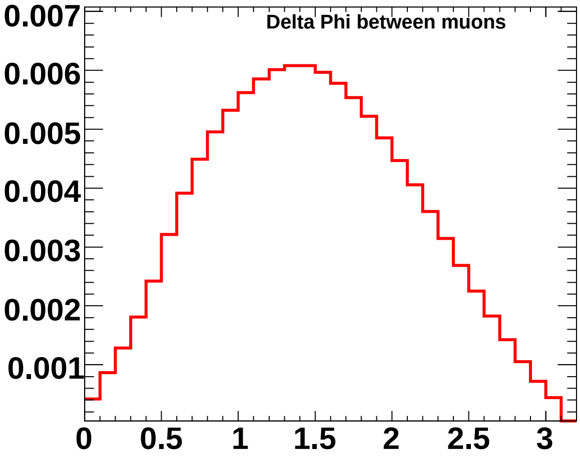

Figure 6.27: The S^2/B distribution constructed by integrating the events below successive cut values and calculating the ratio of the signal squared over the sum of the backgrounds for $\Delta\phi(\mu_1, \mu_2)$ at preselection.

The second cut is on the missing transverse energy in the event. $\vec{\not{E}}_T > 25$ GeV is applied to suppress the large Z/γ^* background and to reject QCD events. As described above, the S^2/B in Fig. 6.28 gives a harder optimal cut then we apply.

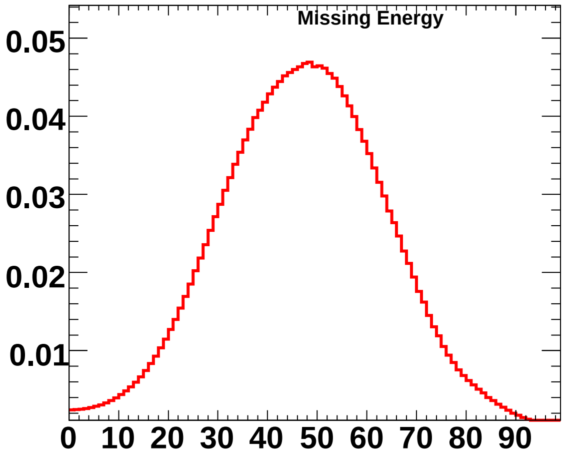

Figure 6.28: The S^2/B distribution constructed by integrating the events below successive cut values and calculating the ratio of the signal squared over the sum of the backgrounds for $\vec{\not\!\!E}_T$ at Cut1.

Finally, there is a cut on the minimum transverse mass (M_T^{min}). This variable is defined by taking the transverse mass of each lepton with the $\vec{\not\!\!E}_T$, then choosing the smaller value. The M_T^{min} is required to be larger than 20 GeV, to suppress backgrounds that do not contain a real W boson. If the transverse mass of a lepton with the $\vec{\not\!\!E}_T$ is small, it means that the $\vec{\not\!\!E}_T$ is either small or almost lying on top of the lepton. In either of these cases it is unlikely to come from a W decay, where one would expect a larger $\vec{\not\!\!E}_T$ and some separation between the lepton and the $\vec{\not\!\!E}_T$ to allow for conservation of momentum. Since the $\vec{\not\!\!E}_T$ is actually coming (at least we assume) from two neutrinos both in magnitude and direction, we don't expect the M_T^{min} to actually be the W mass and thus we cut on a smaller number. Again, this smaller number is below the optimal cut for a cut-based analysis, as can been seen in Fig. 6.29.

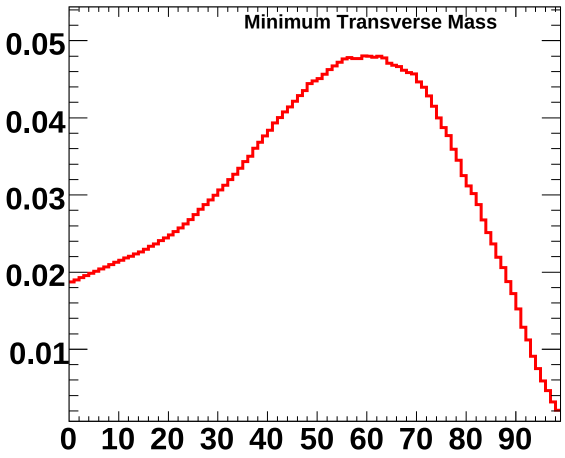

Figure 6.29: The S^2/B distribution constructed by integrating the events below successive cut values and calculating the ratio of the signal squared over the sum of the backgrounds for the minimum transverse mass at Cut2.

6.3.1 Control Plots at Cut1

As described above, Cut1 consists of a cut on the opening angle between the two selected muons of $\Delta\phi(\mu_1, \mu_2) < 2.5$. Figs. 6.30 - 6.33 show control plots at this cut stage. The $\vec{\not\!\!E}_T$ distribution before cutting on it can be seen in Fig. 6.34.

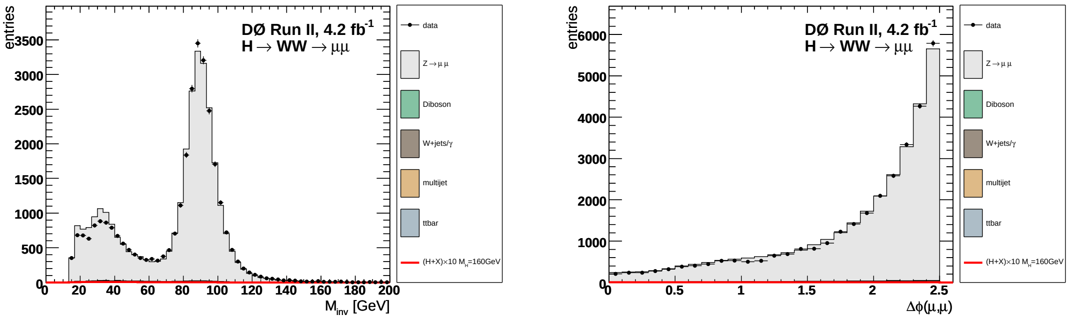

Figure 6.30: Invariant mass distribution(left) and $\Delta\phi_{\mu_1\mu_2}$(right) at Cut1.

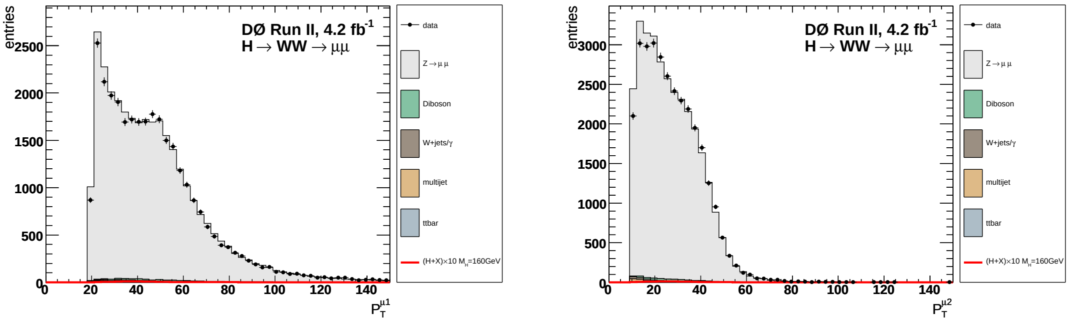

Figure 6.31: p_T distribution for the leading(left) and second leading(left) muons at Cut1.

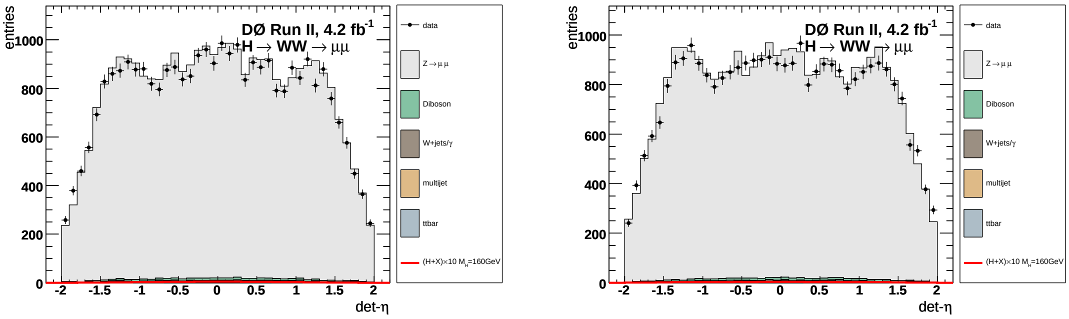

Figure 6.32: Detector η of the leading(left) and second leading(right) muons at Cut1.

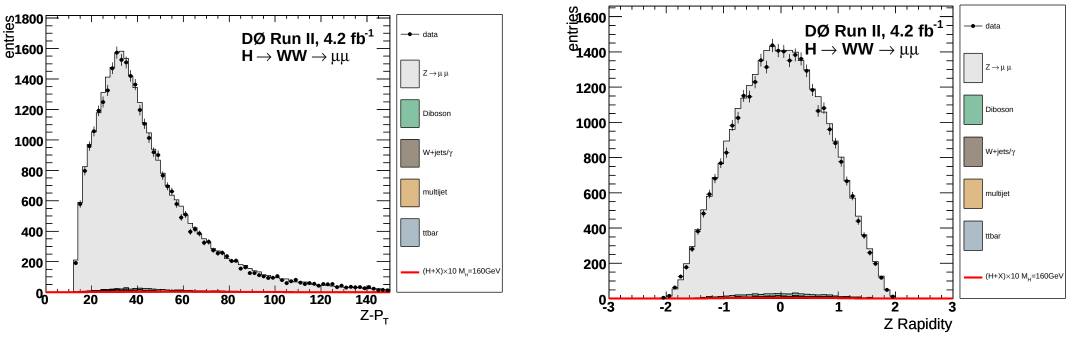

Figure 6.33: Z_{p_T}(left) and rapidity of the dimuon system(right) at Cut1.

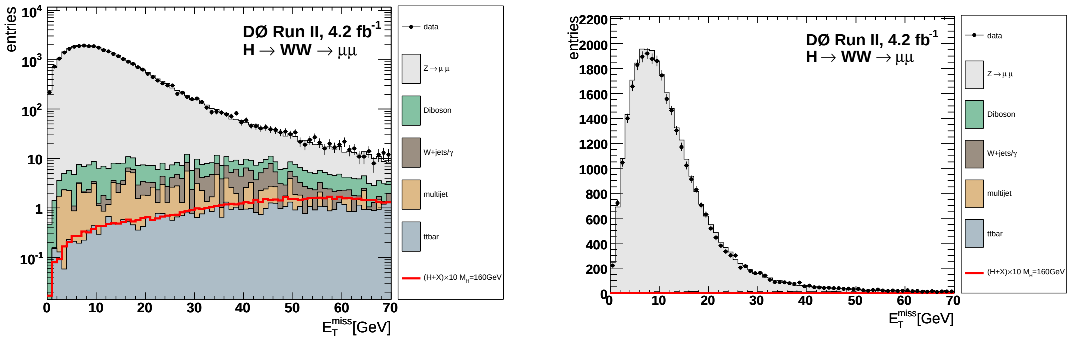

Figure 6.34: Missing energy at Cut1 in log scale(left) and linear scale(right).

6.3.2 Control Plots at Cut2

Cut2 requires that the missing energy in the event be greater than 25 GeV. Figs. 6.35 - 6.41 show control plots at this cut stage. Fig. 6.43 shows the M_T^{min} distribution at Cut2 before cutting on it.

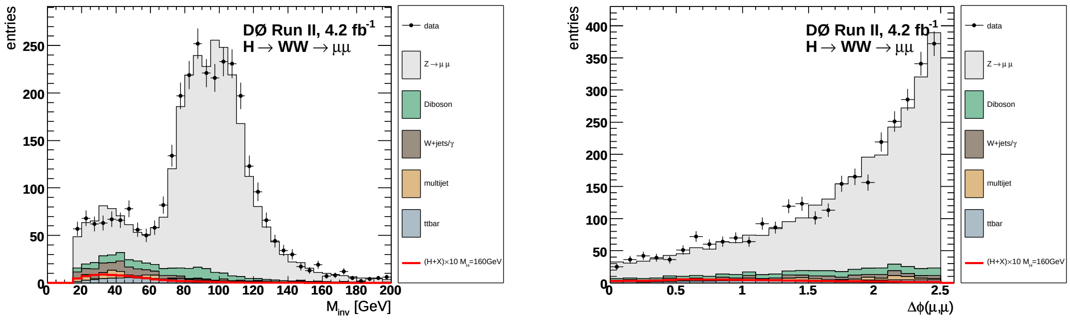

Figure 6.35: Invariant mass distribution(left) and $\Delta\phi(\mu_1, \mu_2)$(right) at Cut2.

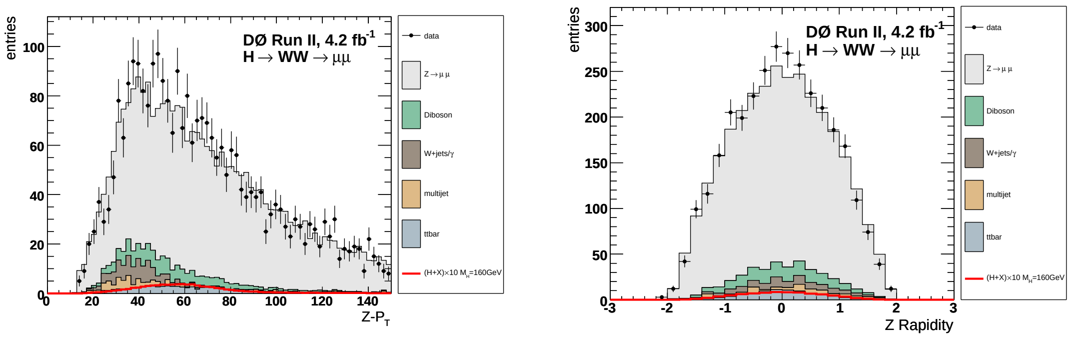

Figure 6.36: Z_{p_T}(left) and rapidity of the dimuon system(right) at Cut2.

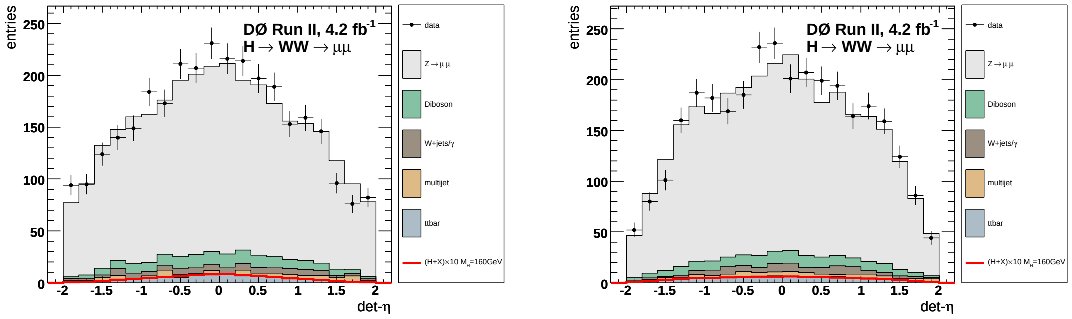

Figure 6.37: Detector η of the leading(left) and second leading(right) muons at Cut2.

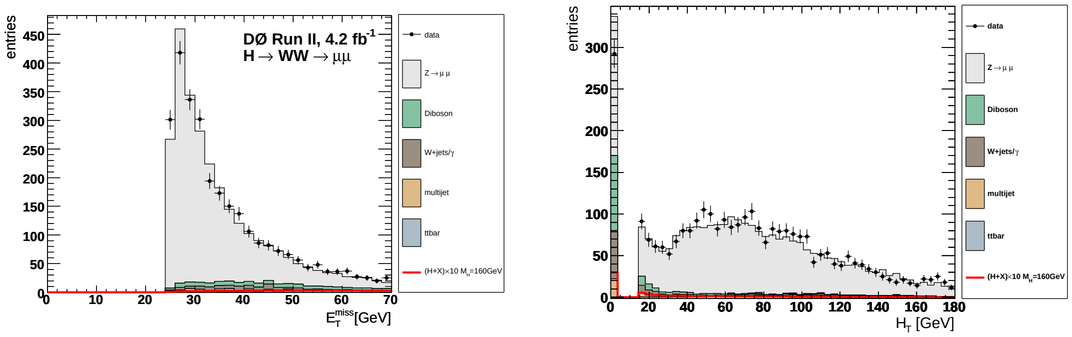

Figure 6.38: Missing energy(left) and ht(right) at Cut2.

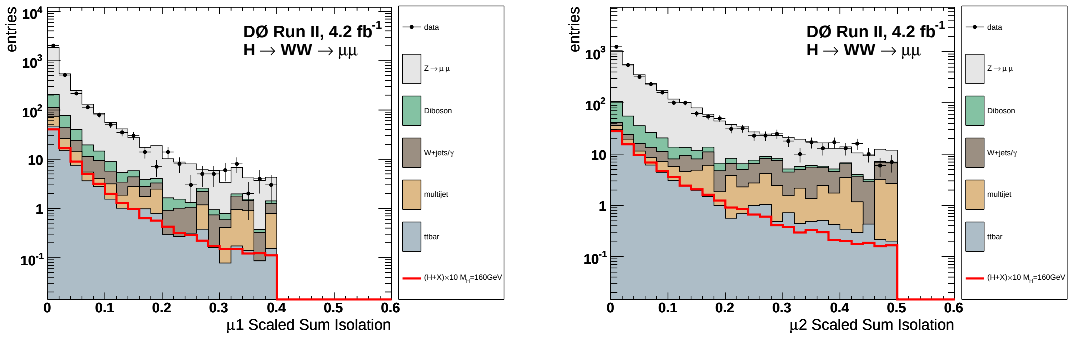

Figure 6.39: Sum of the scaled muon isolations (track and calorimeter) for the leading(left) and second leading(right) muons at Cut2.

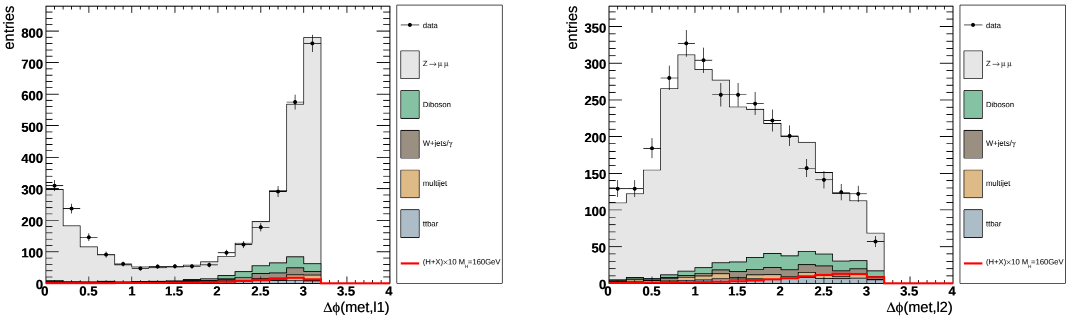

Figure 6.40: $\Delta\phi$ between the first lepton and the missing transverse energy (left), and between the second lepton and the missing transverse energy (right) at Cut2.

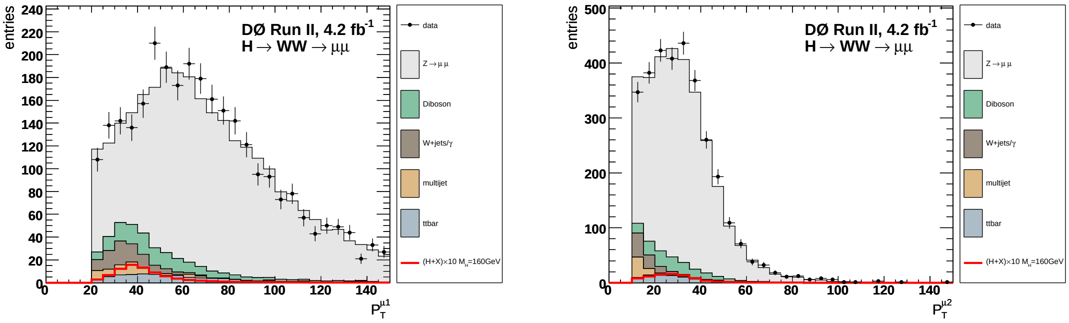

Figure 6.41: p_T distribution for the leading(left) and second leading(left) muons at Cut2.

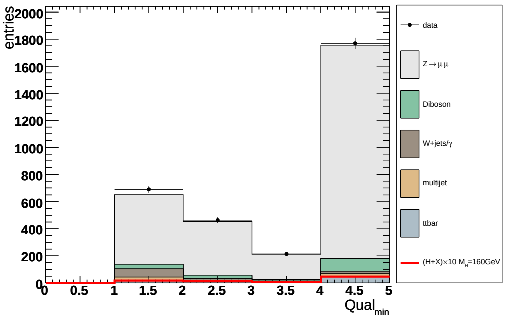

Figure 6.42: Minimum muon quality at Cut2.

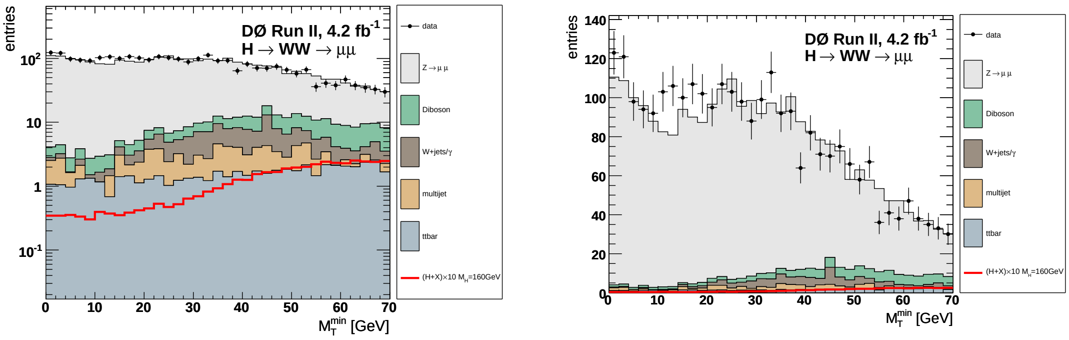

Figure 6.43: M_T^{min} at Cut2 in log scale(left) and linear scale(right).

6.3.3 Control Plots at Cut3

Cut3 requires that the minimum transverse mass of the event be greater than 20 GeV. This is the cut stage used as input to the Neural Network. Figs. 6.44 - 6.50 show control plots at this cut stage. Note that the Higgs signal is now multiplied by a factor of 100 to make the separation power of each variable more apparent.

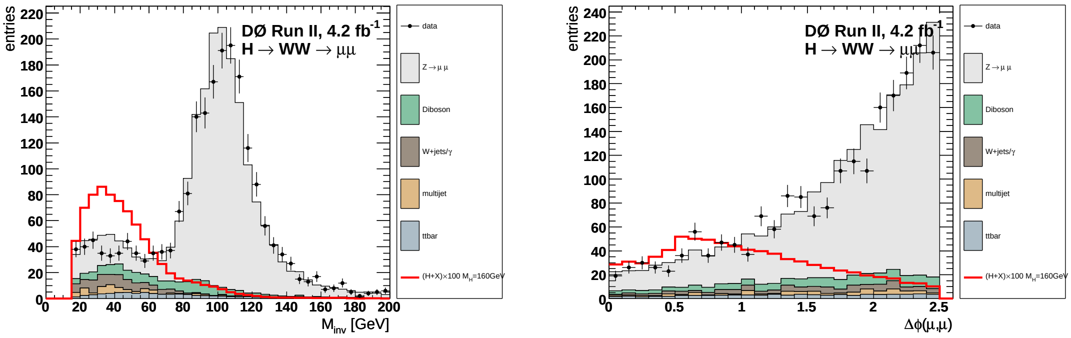

Figure 6.44: Invariant mass distribution(left) and $\Delta\phi(\mu_1, \mu_2)$(right) at Cut3.

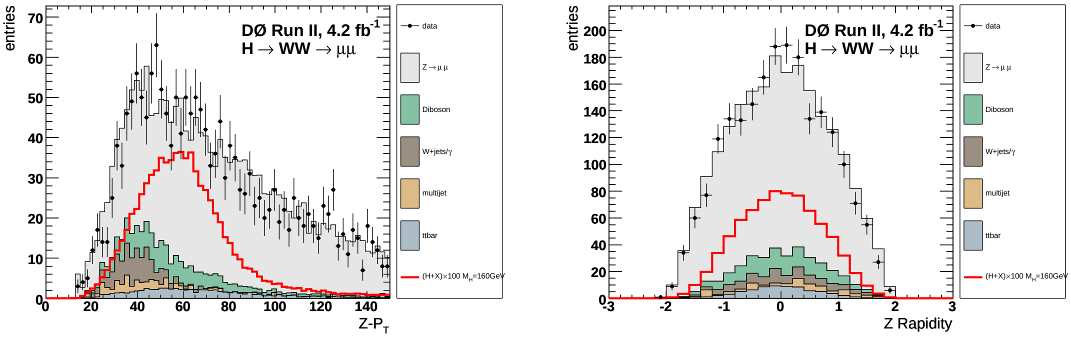

Figure 6.45: Z_{p_T}(left) and rapidity of the dimuon system(right) at Cut3.

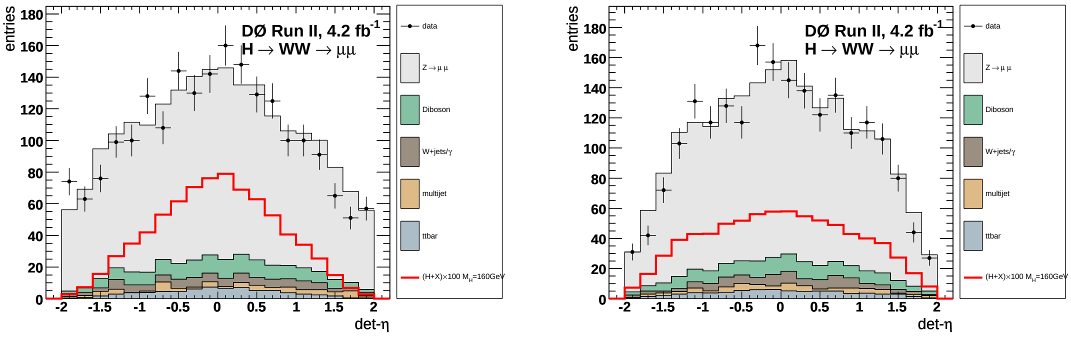

Figure 6.46: Detector η of the leading(left) and second leading(right) muons at Cut3.

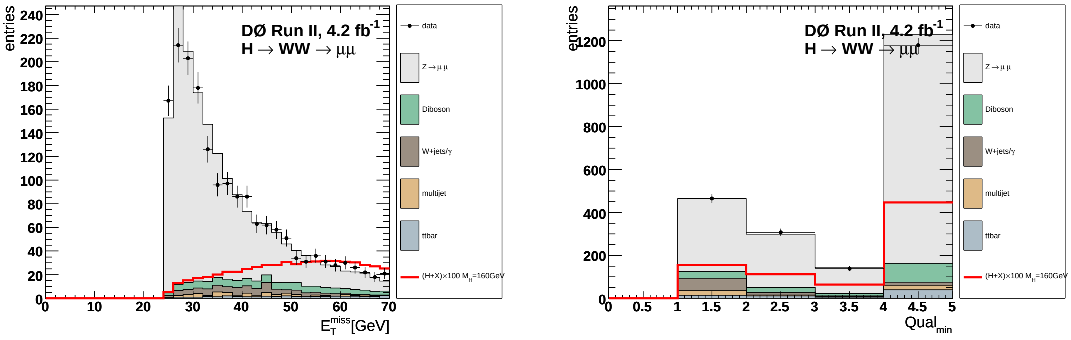

Figure 6.47: Missing energy(left) and minimium muon quality(right) at Cut3.

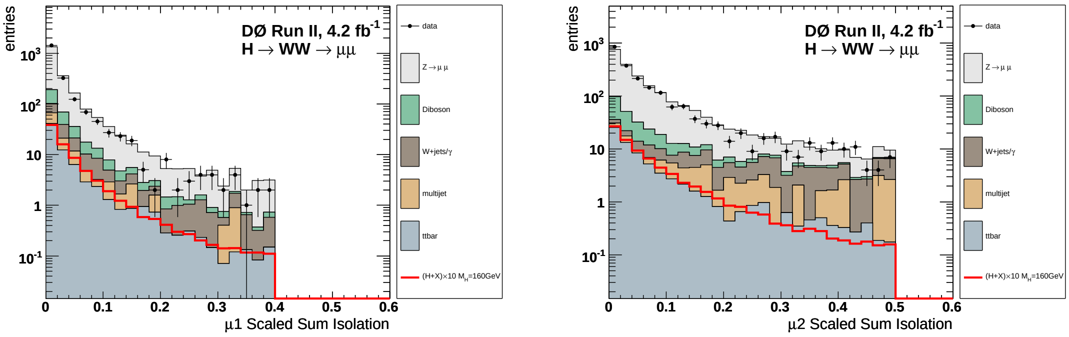

Figure 6.48: Sum of the scaled muon isolations (track and calorimeter) for the leading(left) and second leading(right) muons at Cut3.

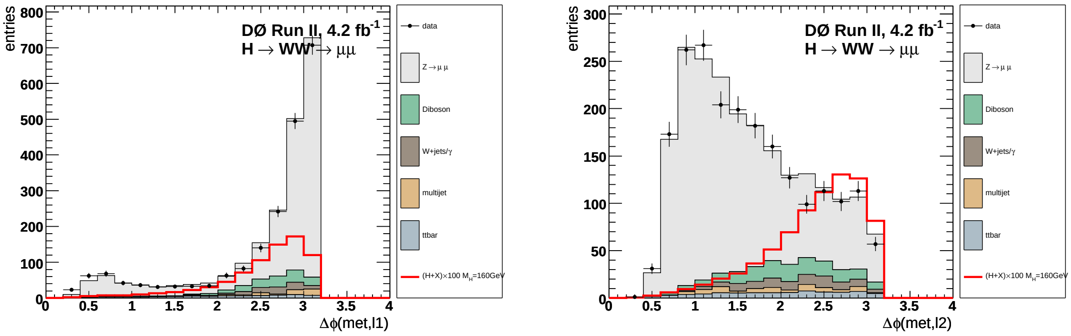

Figure 6.49: $\Delta\phi$ between the first lepton and the missing transverse energy (left), and between the second lepton and the missing transverse energy (right) at Cut3.

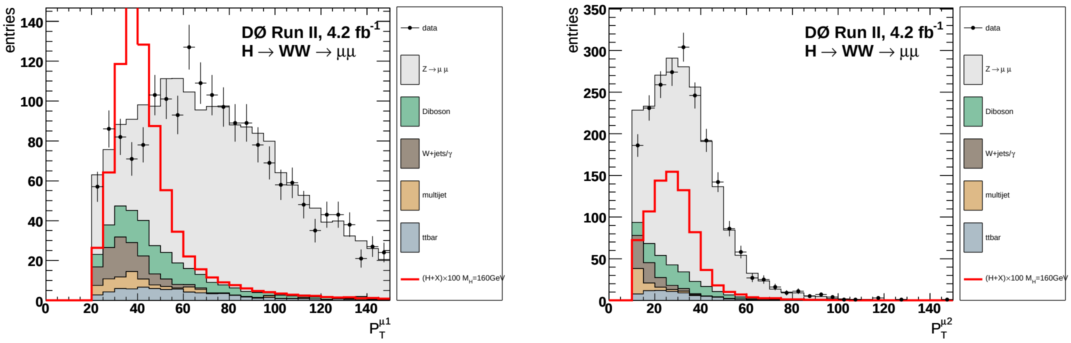

Figure 6.50: p_T distribution for the leading(left) and second leading(left) muons at Cut3.

6.3.4 Selection Cutflow

Cut flows for $m_{\mathrm{H}} = 160$ GeV can be found in Table 6.2. Table 6.3 shows the expected number of signal events remaining at Cut3 for all Higgs masses explored. The over-estimation in the sum of the backgrounds seen in the table is due to the $Z \rightarrow \mu\mu$ sample. There is a known issue with modelling Z/γ* events using Alpgen in the low dilepton invariant mass region ($M_{inv} < 60$ GeV) where too many events are generated relative to the higher mass region for certain event kinematics. This over-estimation does not have a large impact on the final result of the analysis because the analysis cuts and the multivariate technique effectively push these Z events to the non-signal region. In future iterations of this analysis a low-mass reweighting of the Z samples is performed to correct this issue and almost no change to the final result is observed (see section 9.2.1).

Cut	Data	Sum Bkgd	$H_{160} \rightarrow WW$
0	303558	301637.22 ± 179.52	10.37
1	30866	31452.98 ± 32.39	9.20
2	3136	3072.07 ± 10.59	8.15
3	1456	2134.38 ± 9.20	7.79

Cut	$Z \rightarrow \mu\mu$	$Z \rightarrow \tau\tau$	$W + jet/\gamma$	$t\bar{t}$
0	298725.75 ± 178.05	1897.58 ± 15.53	164.03 ± 2.26	131.59 ± 0.38
1	30740.79 ± 30.66	147.20 ± 2.03	106.64 ± 1.87	93.05 ± 0.32
2	2599.11 ± 7.01	73.50 ± 1.31	91.56 ± 1.73	82.89 ± 0.31
3	1761.50 ± 5.69	12.11 ± 0.48	86.00 ± 1.69	71.36 ± 0.28

Cut	ZZ	WZ	WW	QCD
0	114.85 ± 0.03	128.36 ± 0.06	194.69 ± 0.30	280.39 ± 16.74
1	65.23 ± 0.02	75.62 ± 0.05	123.34 ± 0.23	101.12 ± 10.06
2	26.21 ± 0.01	30.47 ± 0.03	110.14 ± 0.22	58.19 ± 7.63
3	22.02 ± 0.01	25.34 ± 0.03	106.87 ± 0.22	49.17 ± 7.01

Table 6.2: Number of expected background events and number of observed events, after successive selections with respect to $m_{\mathrm{H}} = 160$ GeV. Only statistical uncertainty is given for all backgrounds.

H_{115}	H_{120}	H_{125}	H_{130}	H_{135}	H_{140}	H_{145}	H_{150}	H_{155}
1.44	2.05	2.93	3.83	4.79	5.56	6.25	6.79	7.29

H_{160}	H_{165}	H_{170}	H_{175}	H_{180}	H_{185}	H_{190}	H_{195}	H_{200}
7.79	7.65	7.09	6.47	5.79	4.94	4.19	3.71	3.37

Table 6.3: Expected number of signal events remaining at Cut3 for all Higgs masses from 115 GeV to 200 GeV. The number listed is the combination of all Higgs signals considered (gluon fusion, vector boson fusion, and associated production).

Chapter 7

Multivariate Discriminant

In order to increase the sensitivity of an analysis, it is advantageous to combine the information from several variables using a multivariate classifier. There are many multivariate classifiers to choose from, each with varying degrees of robustness and predictive power. This analysis makes use of an artificial neural network(ANN or just NN). [86]

7.1 Neural Networks

Neural networks are very sophisticated multivariate techniques commonly used in high energy physics. They are capable of modeling extremely complex functions. In particular, NNs are non-linear statistical data modeling tools. In other words, they can tackle more than just linearly-separable problems where a line (or more generally in higher dimensions, a hyperplane) can be drawn which separates the points in pattern space.

Neural networks learn by example. Once training samples representative of the data have been gathered, one can use training algorithms to automatically learn the structure of the data. NNs are applicable in virtually every situation in which a relationship between the predictor variables (the inputs) and predicted variables (the outputs) exists, even when that relationship is very complex and not easy to articulate in the usual terms of “correlations”.

NNs are made up of artificial neurons or “nodes”. A neuron receives a number of inputs (either the initial inputs from the user, or from the output of other neurons in the NN). Each input comes via a connection that has a strength (or weight). Each neuron also has a single threshold value. The weighted sum of the inputs is formed, and the threshold subtracted, to compose the “activation” of the neuron. The activation signal is then passed through a transfer function to produce the output of the neuron.

Any network must have inputs (which carry the values of variables of interest in the outside world) and outputs (which form predictions, or control signals). Along with these, there also can be hidden neurons that play an internal role in the network. The input, hidden and output neurons need to be connected together for the NN to function. A simple network has a feedforward structure, which means signals flow from the inputs, forwards through any hidden nodes, eventually reaching the output nodes. This gives the structure a stable behavior. A typical feedforward network has neurons arranged in a distinct layered topology. The input layer serves to introduce the values of the input variables. The hidden and output layer neurons are each connected to all of the nodes in the preceding layer.

When the network is executed (used), the input variable values are placed in the input nodes, and then the hidden and output layer nodes are progressively executed. Each of them calculates its activation value by taking the weighted sum of the outputs of the nodes in the preceding layer, and subtracting the threshold. The activation value is passed through the transfer function to produce the output of the neuron. When the entire network has been executed, the outputs of the output layer act as the output of the entire network. A cartoon of a neural network can be seen in Fig. 7.1.

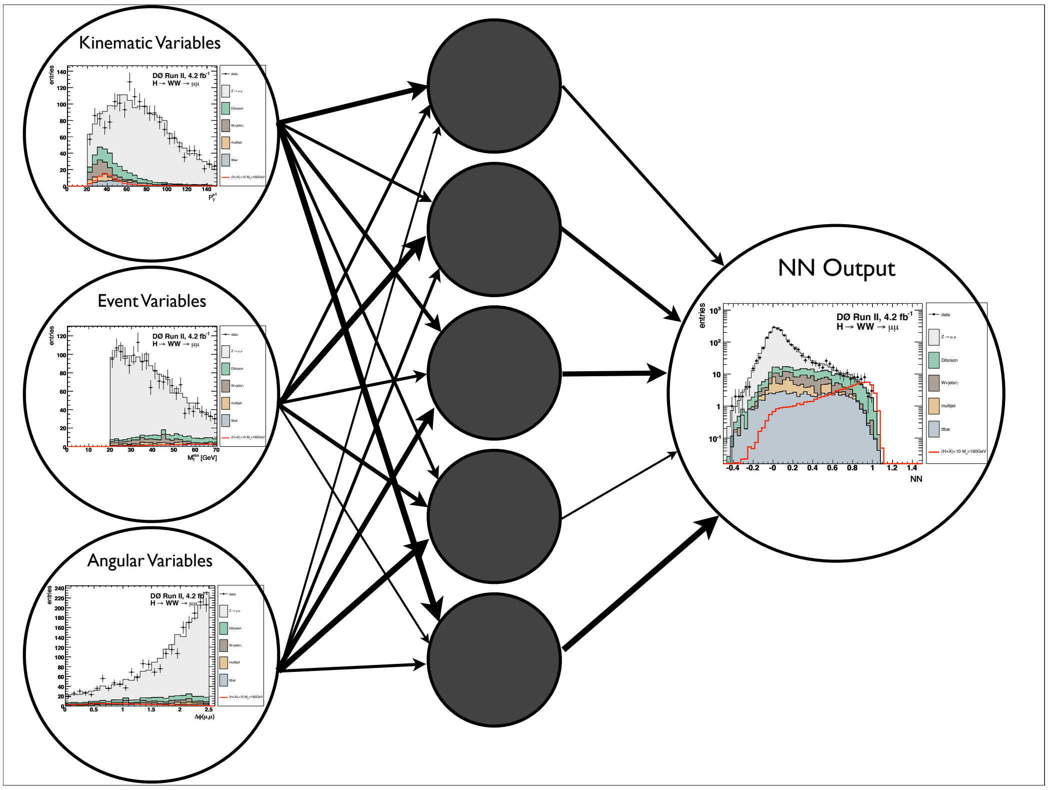

Figure 7.1: NN Cartoon.

The transfer function of a node is typically chosen so that it can accept input in any range, and produces output in a strictly limited range. In general, while the input can be in any range, there is a saturation effect so that the node is only sensitive to inputs within a fairly limited range. A transfer function commonly used is the sigmoid function. An example of a sigmoid function is the logistic function:

$$P(t) = \frac{1}{1+e^{-t}} \tag{7.1}$$

Sigmoid functions are smooth and easily differentiable, facts that are critical in allowing the network training algorithms to operate. (This is the reason why , for example, a step function is not used.)

To summarize the operation of the NN, neurons each perform a biased weighted sum of their inputs and pass this activation level through a transfer function to produce their output, and the neurons are arranged in a layered feedforward topology. The network thus has a simple interpretation as a form of input-output model, with the weights and thresholds (biases) being the free parameters of the model. Such networks can model functions of almost arbitrary complexity, with the number of layers, and the number of nodes in each layer, determining the function complexity.

Once the number of layers, and number of nodes in each layer, has been selected, the network's weights and thresholds must be set so as to minimize the prediction error made by the network. This is the role of the training algorithms. The training samples are used to automatically adjust the weights and thresholds in order to minimize this error. This process is equivalent to fitting the model represented by the network to the training data available. The error of a particular configuration of the network can be determined by running all the training cases through the network, comparing the actual output generated with the desired or target outputs. The differences are combined together by an error function to give the network error. A cycle of running through the entire training sample is called an "epoch". Many training epochs are run to better train the NN and thus get closer to the desired output (better discrimination between signal and background in our case) when using it on the test sample and data.

It is not possible to analytically determine where the global minimum of the error is, and so neural network training is essentially an exploration of the error space. From an initially random configuration of weights and thresholds (a random point in error space), the training algorithms incrementally search for the global minimum. Eventually, the algorithm stops in a low point, which may be a local minimum (but hopefully is the global minimum) and the weights and threshholds are fixed for use on the actual data to produce the output (a linear sum of the sigmoid functions produced at each node in the layer preceeding it). The output in this analysis is designed such that more signal-like events get a higher output value while background events are moved to lower output values. NNs of this type have been applied in RunI and RunII analysis and are used for b-tagging and object identification techniques [87].

7.2 $H \to W^+W^-$ Neural Network

Every network used in this analysis consists of three layers of nodes: an input layer, a hidden layer and an output layer. There are 12 input variables, 8 hidden nodes and one output variable. All of the training in this analysis is done for 600 epochs. The details of the inputs, training, and output are described in the next few sections.

7.2.1 Training

The NN was trained using `TMultiLayerPerceptron` [88] which is part of the `ROOT` package [89]. Separate networks were trained for each Higgs mass. All background samples and the $H \to W^+W^- \to \mu^+\nu\mu^-\nu$ signal sample are used for the training. The NN is trained at the final selection stage - *Cut3*(section 6.3). Events in the background and signal samples go into the training weighted by their appropriate event weights. Only half of the events (every other event, to avoid possible biases from the event generation) are used for training the NN and the other half are reserved as a test sample. After training, only the events in the test sample are used in the rest of the analysis. If all of the events were used for training and in the analysis, one would run the risk of teaching the NN to separate the exact events given to it rather than how to use the connections between the different variables for discriminantion. Additionally, due to the very large number of $Z + jets$ background events at the final selection, these events are given a reduced weight for training purposes. Even with this reduced weight the NN does a very good job of distinguishing them from signal and this allows it to focus on separating the harder to reduce backgrounds, such as diboson and

$W+jets$. $Z+jets$ events with a dimuon invariant mass greater(less) than 70 GeV receive an additional weight, on top of their normal event weight, of 0.1(0.5). These additional weights were somewhat optimized by trying various weights in different mass regions for the $Z + jets$ background and picking the configuration that yielded the best overall separation (highest signal to background ratio in the most "signal-like" output bins).

7.2.2 Input Variables

A list of input variables for the NN has been created based on the separation power between signal and background of the various distributions. These variables can be divided into three classes: object kinematics, event kinematics and angular variables. They are listed in in Table 7.1. The separation power of each input variable can be seen in Sec. 6.3.3 where each of the input variables is plotted for all backgrounds and overlayed with the signal (multiplied by a factor of 10). The relative importance of each variable can change significantly over the analyzed mass range due to mass dependent kinematical characteristics of the expected signal. The relative importance of each input variable can be seen in Table 7.2 in arbitrary units.

NN Analysis Variables	
Object kinematics	
p_T of leading muon	$p_T(\mu_1)$
p_T of trailing muon	$p_T(\mu_2)$
sum of the transverse momenta of the leptons(zpt)	$\vec{p}_T(\mu_1) + \vec{p}_T(\mu_2)$
sum of the momenta of all jets:	$H_T = \sum_i \text{jet}_i$
minimal quality of one of the two leptons:	$\mathcal{Q}_{min}(\mu_1, \mu_2)$
Event Kinematics	
invariant mass of both leptons	$M_{\text{inv}}(\mu_1, \mu_2)$
missing transverse energy	$\vec{\not{E}}_T$
Topological Variables	
angle between selected muons	$\Delta\phi(\mu_1, \mu_2)$
angle between the first muon and $\vec{\not{E}}_T$	$\Delta\phi(\vec{\not{E}}_T, \mu_1)$
angle between the second muon and $\vec{\not{E}}_T$	$\Delta\phi(\vec{\not{E}}_T, \mu_2)$
minimum transverse mass	M_T^{min}
log of the sum of the scaled isolations for the two muons	$log_{10}(Iso_1 + Iso_2)$

Table 7.1: Input variables for the Neural Network.

Input Variable	Relative Importance
$M_{\text{inv}}(\mu_1, \mu_2)$	0.071
M_T^{min}	0.054
$p_T(\mu_1)$	0.039
$\vec{\not{E}}_T$	0.034
$\vec{p}_T(\mu_1) + \vec{p}_T(\mu_2)$	0.029
$p_T(\mu_2)$	0.028
$\Delta\phi(\vec{\not{E}}_T, \mu_2)$	0.018
$H_T = \sum_i \text{jet}_i$	0.017
$\Delta\phi(\vec{\not{E}}_T, \mu_1)$	0.014
$log_{10}(Iso_1 + Iso_2)$	0.011
$\Delta\phi(\mu_1, \mu_2)$	0.009
$\mathcal{Q}_{min}(\mu_1, \mu_2)$	0.007

Table 7.2: Relative importance of the input variables for the Neural Network.

7.2.3 NN Output

After training the neural network, the events from the test sample (the half of the simulated events not used for training) and the data are run through it. The output is a distribution in which more signal-like events are pushed to higher values. While the NN is trained at *Cut3*, all of the events that pass the preselection requirements are run through it to produce plots for checking data/MC agreement in the NN output variable. This serves to reassure us that we are not treating the data significantly differently than the simulated events in the test sample when passing them through the NN. The NN output distributions at preselection are shown in Fig. 7.2 and at the final selection stage in Fig. 7.3 in both linear and log scale. As can be seen in Fig. 7.3, no evidence for a signal is apparent.

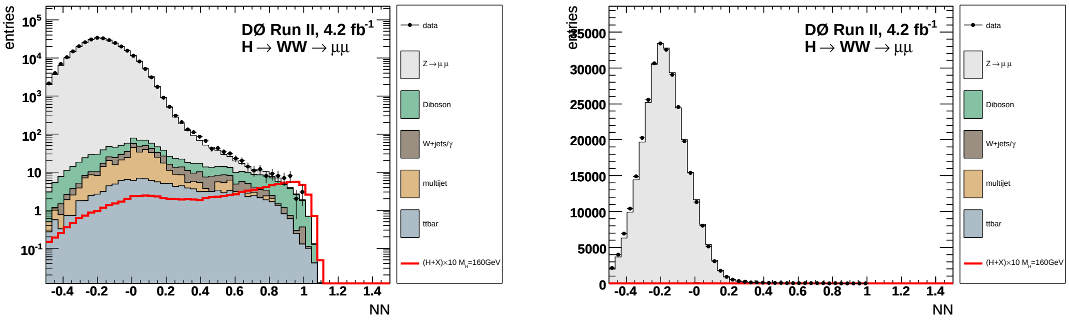

Figure 7.2: NN distribution at preselection for m_H=160 GeV in log scale(left) and linear scale(right).

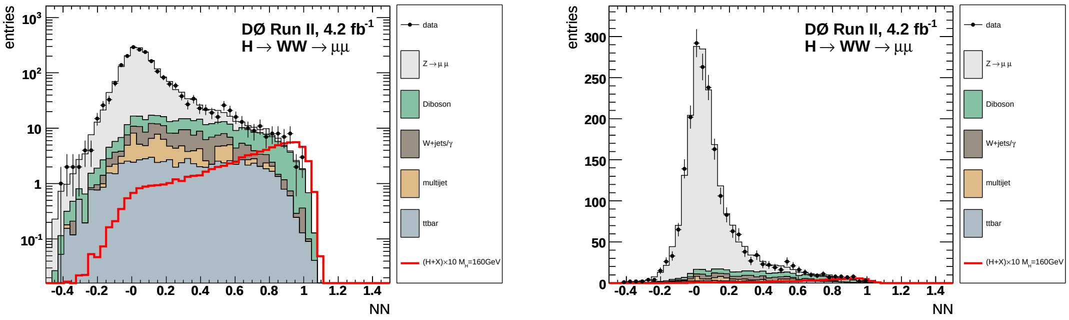

Figure 7.3: NN distribution at Cut3 for m_H=160 GeV in log scale(left) and linear scale(right).

Chapter 8

Systematic Uncertainties

In general, the uncertainty of a measurement is defined by a range of values which encompass the “true” value. There are two main classes of uncertainties. The first are *statistical uncertainties* which arise from the fact that a measurement is based on a finite set of observations. This type of uncertainty is easily treated using standard statistics theory. The second class of uncertainties are known as *systematic uncertainties.* These uncertainties are due to things like the nature of the detector used for the measurements, the methods used for event readout or reconstruction, or assumptions made when creating Monte Carlo simulation. The estimation of systematic uncertainties must be handled with some care as it can impact the result greatly. The estimation and treatment of “systematics” is the focus of the following chapter.

Systematic uncertainties in this analysis are placed in two categories. The first category consists of uncertainties that only change the overall normalization or “rate” of the various physics processes. These will be called “flat” systematics because they are not expected to change the shape of the final variable distribution. They are determined by propagating the systematic variation in question through the selection requirements (Chapter 6) and calculating the relative uncertainty (relative difference).

The second are those uncertainties which do change the shape of the final variable distribution as well as possibly changing the normalization. Any such uncertainties are called shape-changing or just “shape” uncertainties. These variations are propagated through both the selection requirements and the neural network. The shape differences are derived by comparing the non-modified (“nominal”) and modified shape of the NN output and are then put in the form of fractional uncertainties. The integral of the fractional uncertainty gives the relative uncertainty in percent. Because the Neural Network is separately trained for each analyzed Higgs boson mass, the shape uncertainties depend on which NN is used.

Each systematic uncertainty is assumed to be 100% correlated among backgrounds and signal except in a few special cases which will be mentioned in their descriptions later. This is because most of the uncertainties are related to either object identification and reconstruction, or kinematic modeling which are both assumed to be consistent throughout the samples. All sources of systematic uncertainty are assumed to be mutually independent, and no inter-correlation is propagated. While this may be slightly conservative, it is largely the case and the correlations would not be easily determined.

8.1 Flat Systematics

The largest flat uncertainties are the cross section factors used to normalize the background processes. These are related to the accuracy of the theoretical cross section calculation and the uncertainty on the normalization of the multijet estimation. Each flat systematic uncertainty is considered to arise from a Gaussian parent distribution when propagated into the final limit calculation (section 9.1.2). Below is a list of systematic uncertainties considered to be independent of NN shape and therefore flat:

- The muon ID uncertainty is taken to be 4%. The muon ID parameters, including the systematic uncertainties associated with varying these parameters, were determined by looking at a sample of Z bosons which decay to two muons. A tag-and-probe method was used to look at ID efficiencies when changing various muon requirements and the uncertainties were estimated by looking at these changes.

- The PDF related uncertainty on the theoretical cross sections are for diboson pair production 7%, top pair production $t\bar{t}$ 10%, W+jets 20%, Z+jets 6% and, signal processes 10% [91, 92, 93, 94, 95, 96]. The PDF is the probability distribution used to determine the momentum each parton has in a simulated collision. These PDF uncertainties are calculated by varying the distribution used and recalculating the cross sections.

- QCD normalization (30 %) from the statistical uncertainty and the uncertainty in the like-sign to opposite-sign ratio.

- The systematic error on the luminosity is assumed to be 6.1% error as measured by the D0 luminosity group [90]. The uncertainty of the Z peak normalization (described in section 6.2.2) was studied by varying the boundaries of the peak fitting function and by comparing fit vs. bin-wise counting of the events. This results in an uncertainty of 5% which agrees well with the luminosity uncertainty. The 6.1% has been chosen for combination purposes with other D0 analysis.

- The charge flip uncertainty has been estimated by calculating the ratio of the events in the Z peak in the like-sign sample with respect to the number of Z peak events in the opposite-sign sample. Since there are also real physics events causing like-sign events, this is only an upper limit on the charge mis-identification rate and therefore a conservative estimate for the uncertainty. The charge misidentification rate is found to be 1%.

An overview of all flat systematics used in the analysis can be found in Table 8.1.

	Σ Bkgd	Signal	$Z+jets$	$W+jets$	$t\bar{t}$	ZZ	WZ	WW	QCD
σ	-	10	6	20	10	6	6	6	-
QCD	-	-	-	-	-	-	-	-	30
Luminosity	6.1	6.1	-	-	-	-	-	-	-
Lepton ID	4	4	-	-	-	-	-	-	-
Charge flips	-	1	1	-	-	-	-	-	-

Table 8.1: Flat systematic uncertainties in percent.

8.2 Shape Dependent Systematics

As mentioned above, the shape uncertainties are those that, when propagated through the analysis selection, impact the shape of the distribution used to perform the limit calculation. The dependence of the NN output distribution on these uncertainties is determined by varying each parameter by its associated uncertainty ($\pm 1\sigma$) and re-evaluating the shape of the distribution(without retraining). Again, the magnitude of the resulting shape dependence is considered to arise from a Gaussian parent distribution. Some of the uncertainties are "symmetrized". This means that, in the case where only a variation in one direction is possible (i.e. turning off a reweighting which was applied in the nominal case), the uncertainty is taken to be both a positive and negative fluctuation in normalization with the same shape. The dominant shape uncertainties are the jet correction uncertainties, which have an impact on the $\vec{\not{E}}_T$ calculation, and the additional muon momentum smearing which can cause muons to miss or make cuts that they normally would not. Even these most important shape uncertainties are quite small. The current set of shape dependent systematics along with a description of how each was calculated can be found below:

- The uncertainty on the jet energy scale(section 4.4.2), calculated by varying the jet energy scale correction by $\pm 1\sigma$.
- The uncertainty on the $Z - p_T$ reweighting is calculated by not applying the reweighting and symmetrizing 20 percent of the resulting uncertainty. Here, the 20% was estimated by putting the reweighted and non-reweighted samples along with the data into a cross section fitter and determinig what a 1σ variation would be relative to the non-reweighted sample.
- Jet Resolution, which has been calculated by varying the jet resolution by $+1\sigma$ and symmetrizing the resulting uncertainty.
- The uncertainty on the JetID by varying the JetID by -1σ and symmetrizing the resulting uncertainty.
- The uncertainty on the muon momentum smearing determined by varying the smearing by $\pm 1\sigma$.
- The $\Delta\phi(l, l)$ distributions of the production processes $qq \rightarrow WW$ and $gg \rightarrow WW$ differ significantly. Since this difference is not considered by Pythia, an appropriate reweighting at the parton level has been derived and applied to the WW background sample (see section 6.2.1). The uncertainty is estimated by not applying the reweighting and symmetrizing 30% of the resulting uncertainty. While this 30% is almost certainly overkill, the full difference is very small and a conservative estimate was chosen.
- In order to estimate the uncertainty due to the difference between the p_T momenta of the WW system in the Pythia and Sherpa event generators for WW background and signal, the $p_T(WW)$ spectrum reweighting (see section 6.2.1) is removed. Then the uncertainty is calculated by taking the difference with the reweighted samples and symmetrizing the resulting uncertainty.

- The uncertainty on the pseudo-rapidity reweighting for the leading muon is calculated by not applying the reweighting and symmetrizing the resulting uncertainty.
- The uncertainty on the pseudo-rapidity reweighting for the second leading muon is calculated by not applying the reweighting and symmetrizing the resulting uncertainty.
- The uncertainty on the track quantity related reweightings (nSMT hits and track chisq) for the two leading muons is calculated by not applying the reweightings and symmetrizing the resulting uncertainty.
- The uncertainty on the reweighting of the rapidity of the dimuon system is calculated by not applying the reweighting and symmetrizing the resulting uncertainty.
- The $+1\sigma$ and -1σ variations for all four of the extra parameters involved in the extra muon smearing. These parameters were then varied (individually) by this amount and the shape of the result was used for the systematic uncertainty.

In the following figures (Fig. 8.1 - 8.7, the shapes of the fractional uncertainties of the NN output are shown for the signal and background contributions. Only the systematics for a Higgs mass of $m_H = 160$ GeV are shown. To avoid possible biases in low statistics tails, the shapes have been smoothed using the method of “equalized binning”. This is when the individual bins of the initial systematic and nominal histograms are merged into bins of equal statistics such that the bins have roughly equal uncertainties, thus removing statistical fluctuations while preserving the prominent features of the fractional shapes. If a given background has insufficient statistics to yield a meaningful shape for a particular systematic, that systematic is flattened. These flattened uncertainties still differ from what we refer to as flat systematics in that they still vary depending on NN training (and therefore tested Higgs mass). It is possible for a systematic to be flattened for one Higgs mass and not for another if the fluctuations in the NN output are deemed purely statistical. For symmetric systematics, both the positive and negative systematic are shown as a positive fractional uncertainty in the figures below. The last figure (Fig. 8.8) shows all of the systematics added in quadrature for each sample.

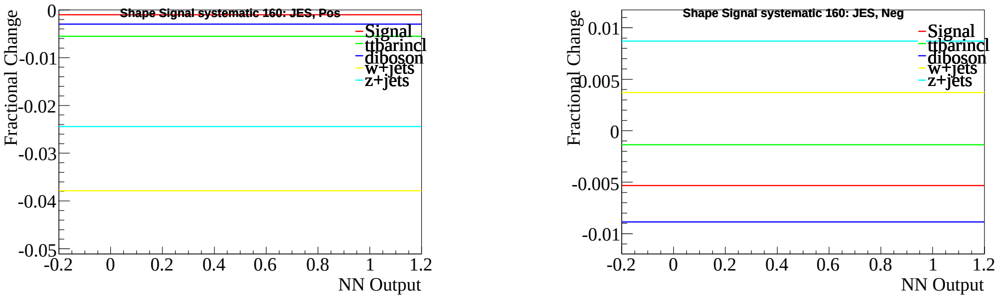

Figure 8.1: Fractional shape systematics for varying the Jet Energy Scale $+1\sigma$ (left) and -1σ (right) for the various backgrounds, assuming a Higgs mass of $m_H = 160$ GeV.

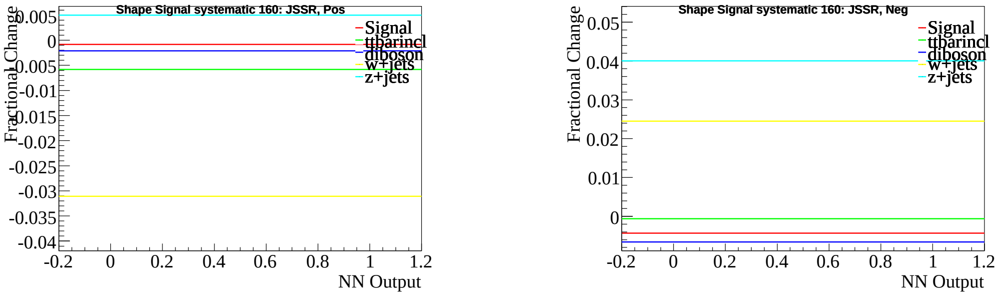

Figure 8.2: Fractional shape systematics for varying the Jet Energy Resolution $+1\sigma$ (left) and -1σ (right) for the various backgrounds, assuming a Higgs mass of $m_H = 160$ GeV.

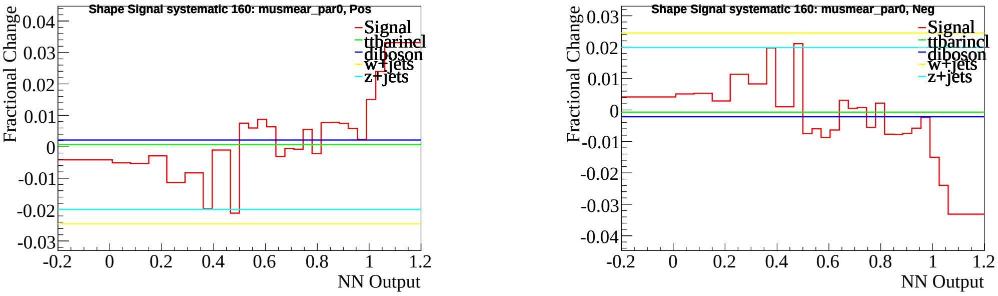

Figure 8.3: Fractional shape systematics for varying the muon momentum smearing $+1\sigma$ (left) and -1σ (right) for the various backgrounds, assuming a Higgs mass of $m_H = 160$ GeV.

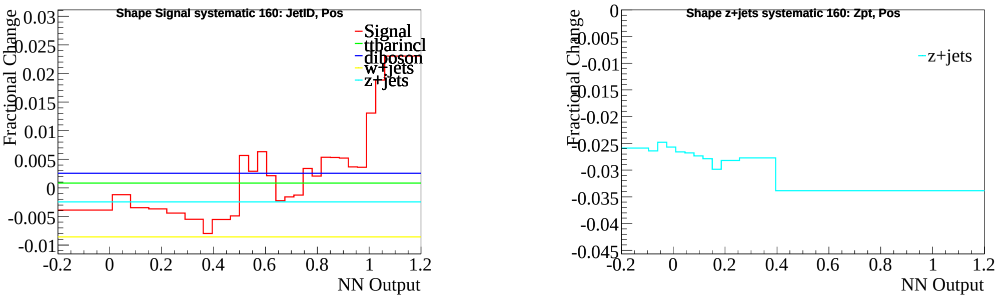

Figure 8.4: Fractional shape systematics for varying the Jet ID (left) and $Z - p_T$ reweighting (right) by 1σ for the various backgrounds, assuming a Higgs mass of $m_H = 160$ GeV.

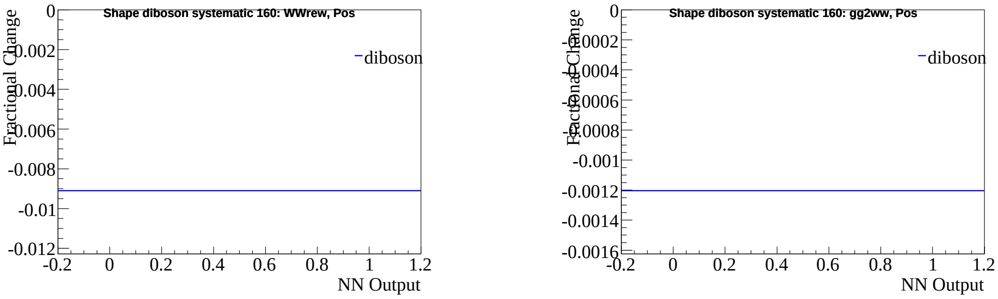

Figure 8.5: Fractional shape systematics for varying the WW reweighting (left) and $gg2WW$ (right) by 1σ for the various backgrounds, assuming a Higgs mass of $m_H = 160$ GeV.

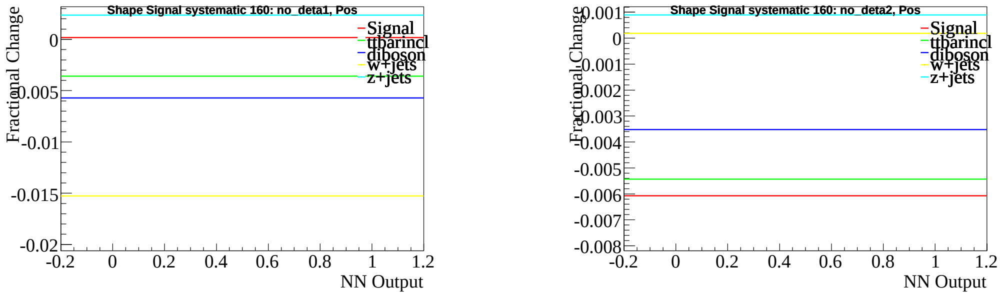

Figure 8.6: Fractional shape systematics for varying the detector pseudo-rapidity reweighting for the leading (left) and second leading (right) muon by 1σ for the various backgrounds, assuming a Higgs mass of $m_H = 160$ GeV.

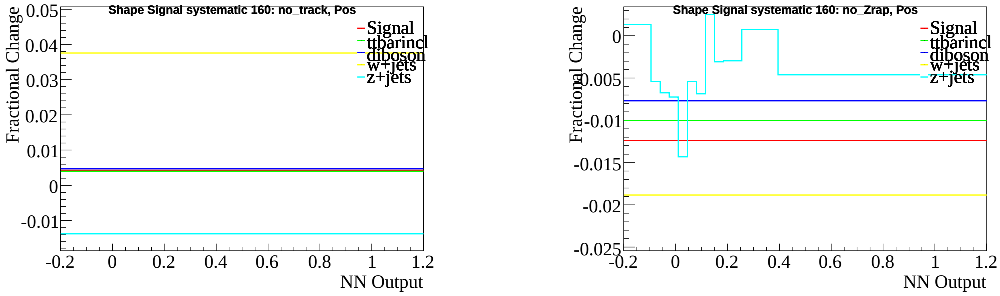

Figure 8.7: Fractional shape systematics for varying the reweighting of the track related variables (left) and the rapidity of the di-muon system (right) by 1σ for the various backgrounds, assuming a Higgs mass of $m_H = 160$ GeV.

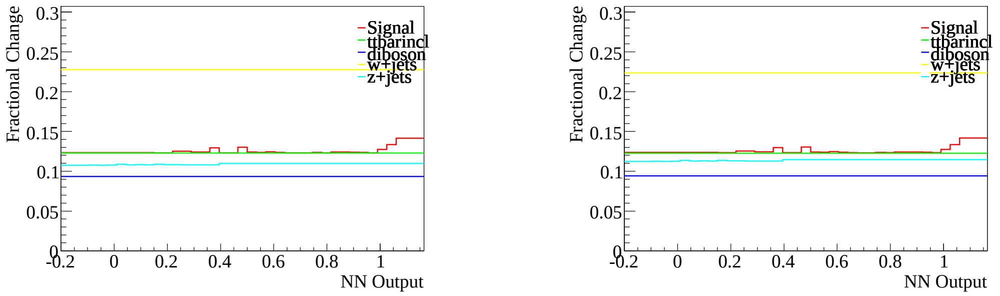

Figure 8.8: Total fractional uncertainty, adding all of the systematic uncertainties in quadrature, for a Higgs mass of $m_H = 160$ GeV. Both the positive(left) and negative(right) uncertainties are shown as positive fractional uncretainties.

Chapter 9

Limits and Conclusions

When searching for a small signal in a sea of background events, one must be able to quantify the outcome of an analysis. In the case where no signal is observed, it is important to be able to describe to what degree you are certain that there is no signal. To do this we set an upper limit on the SM Higgs production cross section times the branching ratio at the point where, if there were a signal, we would observe it in 95% of identical experiments to this one and in 5% we would not. This is a 95% confidence level(C.L.). The limit is quoted as a ratio with the current theoretical SM Higgs cross section, such that a limit of 1.0 would exclude the existence of the predicted Higgs boson at that particular mass with 95% confidence. The current analysis is performed for 18 different Higgs mass points using a mass range of $m_{\mathrm{H}} = 115 - 200$ GeV in steps of 5 GeV.

9.1 Mass Limits

Since no evidence of SM Higgs production is observed, we proceed to set upper limits on its cross section. We do this using a modified $\mathrm{CL_s}$ method [97, 98, 99], as implemented in Ref. [100]. Version V00-03-17 of the `collie` package has been used. Appendix A further describes the $\mathrm{CL_s}$ method and limit setting procedures.

9.1.1 Limits Without Systematics

Limits are first calculated without systematic uncertainties using the `CLFast` feature of the `collie` package[1]. The test-variable used in the limit setting procedure is the NN output discussed in Chapter 7. These limits are listed in Table 9.1. A plot of the expected and observed CLFast limits as a function of Higgs mass can be found in Fig. 9.1. The Log Likelihood Ratio (LLR) as a function of Higgs mass can be found in Fig. 9.2. The limit setting procedure and LLR test statistic are discussed in Appendix A in more detail. In the LLR distribution, the $\mathrm{LLR_b}$ line represents the background-only hypothesis and the green and yellow bands are the one and two σ bands on this line, respectively. Thus, the closer the observed data, $\mathrm{LLR_{obs}}$, are to this line the more background-like they are. Similarly, the closer the data are to the signal+background hypothesis line, $\mathrm{LLR_{s+b}}$, the more evidence there is for a signal. Some measure of the sensitivity of the analysis can be obtained by looking at the separation between $\mathrm{LLR_b}$ and $\mathrm{LLR_{s+b}}$. This means that if the $\mathrm{LLR_{s+b}}$ line is too close to the $\mathrm{LLR_b}$ line, there is not sufficient sensitivity for exclusion.

---

[1] Statistical uncertainties are also ignored at this stage and included in the next step. They have a small effect, on the order of 1-2%

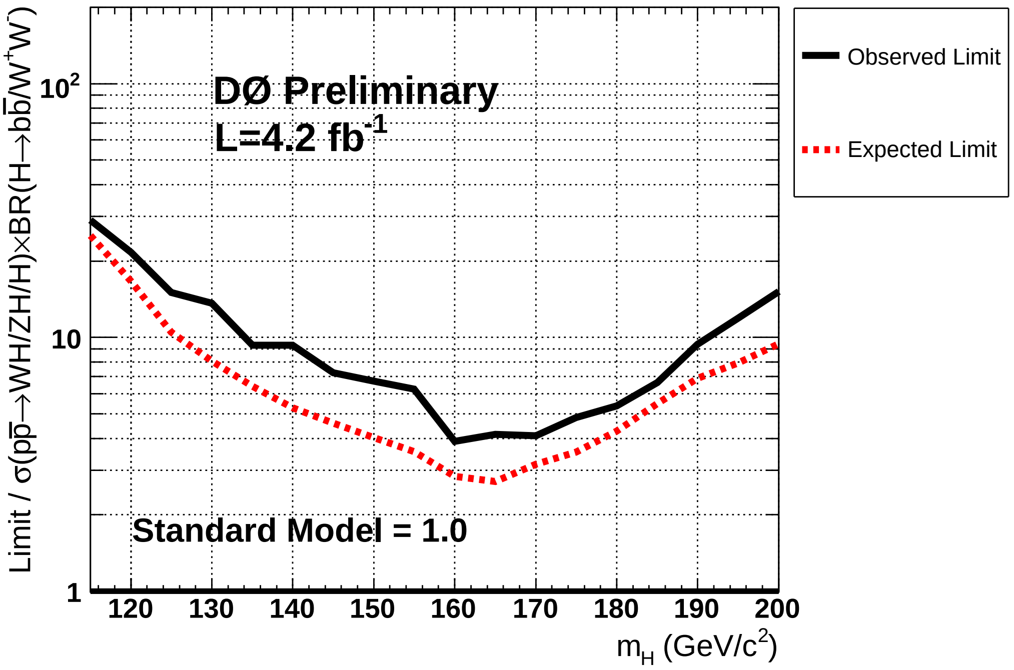

Figure 9.1: Expected and observed limits in terms of $\sigma \times BR(H \to WW^*) \times SM$ for the $\mu\mu$ final state. These are CLFast limits and do not take into account systematic uncertainties.

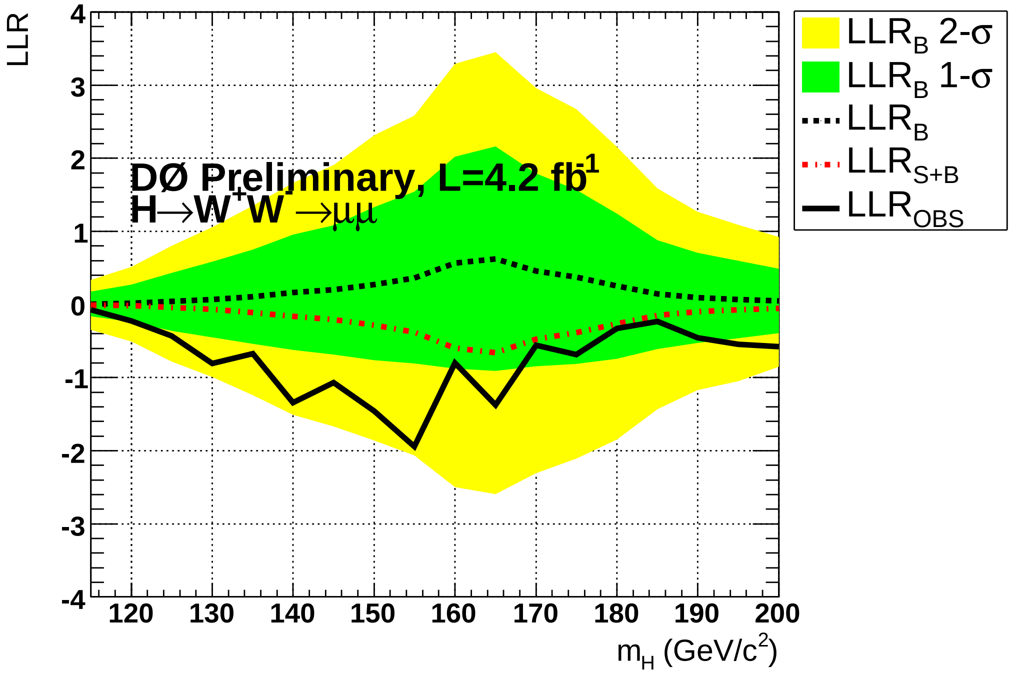

Figure 9.2: Log Likelihood Ratio for the expected and observed limit. The green and yellow shaded areas represent the ±1σ and ±2σ confidence levels for the background only hypothesis and the solid black graph represents the observed limit. The 'background-only' and 'signal + background' hypothesis are given by the black and red dashed lines, respectively. These are CLFast limits and do not take into account systematic uncertainties.

m_h [GeV]	115	120	125	130	135	140	145	150	155	160	165	170	175	180	185	190	195	200
	Observed limit $\sigma \times BR(H \to WW^* \to \mu\mu)$ [$(\sigma \times$ BR)/SM]																	
$\mu\mu$	28.97	21.64	15.05	13.63	9.32	9.30	7.26	6.73	6.25	3.90	4.15	4.10	4.84	5.37	6.64	9.40	11.89	15.14
	Expected limit $\sigma \times BR(H \to WW^* \to \mu\mu)$ [$(\sigma \times$ BR)/SM]																	
$\mu\mu$	25.21	16.65	10.50	8.07	6.40	5.27	4.60	4.02	3.54	2.83	2.70	3.16	3.53	4.27	5.50	6.91	7.94	9.44

Table 9.1: Expected and observed upper limits at 95% C.L. on the cross section times branching ratio for $\sigma \times BR(H \to WW^*)/SM$ for the $\mu\mu$ final states with respect to the standard model expectation. These are CLFast limits and do not take into account systematic uncertainties.

9.1.2 Limits Including Systematic Uncertainties

Systematic uncertainties as described in Chapter 8 are taken into account using the `CLFit2` feature of the `collie` package. The final limits can be found in Fig. 9.3 and Table 9.2. The Log Likelihood Ratio can be found in Fig. 9.4.

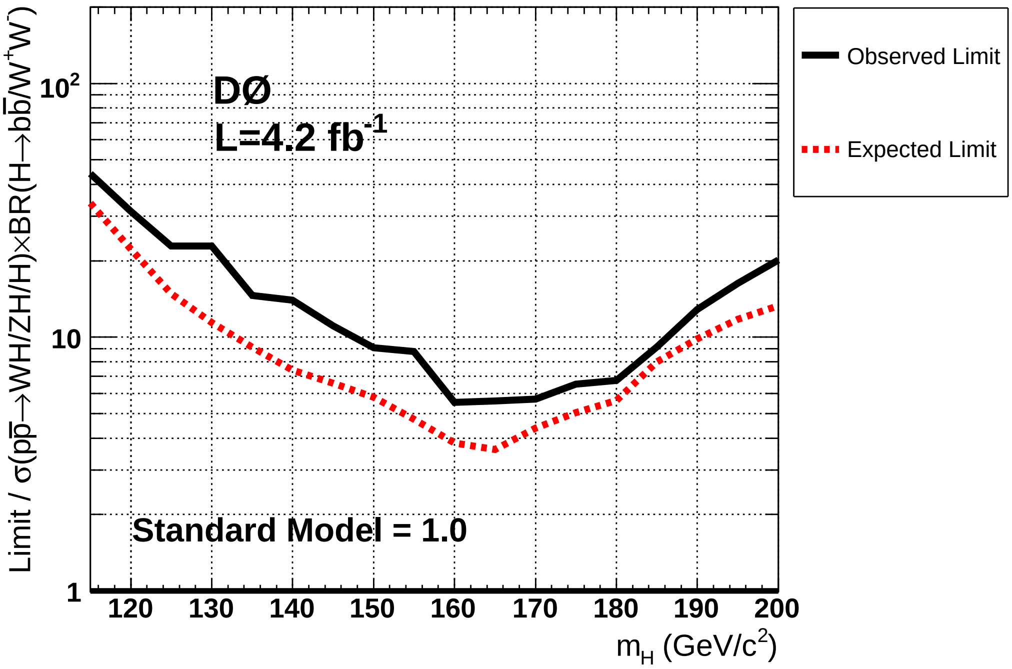

Figure 9.3: Expected and observed limits in terms of $\sigma \times BR(H \to WW^*) \times SM$ for the $\mu\mu$ final state.

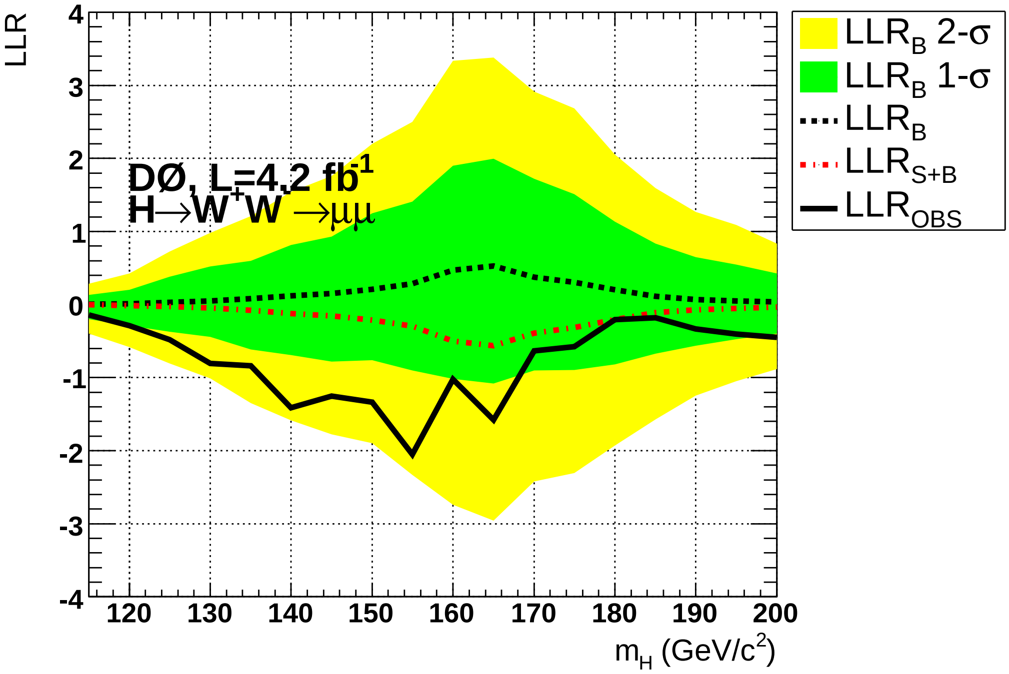

Figure 9.4: Log Likelihood Ratio for the expected and observed limit. The green and yellow shaded areas represent the $\pm 1\sigma$ and $\pm 2\sigma$ confidence levels for the background only hypothesis and the solid black graph represents the observed limit. The 'background-only' and 'signal + background' hypothesis are given by the black and red dashed lines, respectively.

m_h [GeV]	115	120	125	130	135	140	145	150	155
	Observed limit $\sigma \times BR(h \rightarrow ww^{*} \rightarrow \mu\mu)$ [($\sigma \times$ BR)/SM]								
$\mu\mu$	44.14	31.22	22.86	XXXX	14.59	13.96	11.09	9.09	8.78
	Expected limit $\sigma \times BR(H \rightarrow WW^{*} \rightarrow \mu\mu)$ [($\sigma \times$ BR)/SM]								
$\mu\mu$	33.81	22.21	14.85	11.43	9.12	7.41	6.59	5.80	4.74

m_h [GeV]	160	165	170	175	180	185	190	195	200
	Observed limit $\sigma \times BR(h \rightarrow ww^{*} \rightarrow \mu\mu)$ [($\sigma \times$ BR)/SM]								
$\mu\mu$	5.53	5.60	5.70	6.53	6.75	9.16	12.86	16.32	20.15
	Expected limit $\sigma \times BR(H \rightarrow WW^{*} \rightarrow \mu\mu)$ [($\sigma \times$ BR)/SM]								
$\mu\mu$	3.84	3.60	4.39	5.03	5.62	7.98	9.84	11.77	13.30

Table 9.2: Expected and observed upper limits at 95% C.L. on the cross section times branching ratio for $\sigma \times BR(H \rightarrow WW^{*})/SM$ for the $\mu\mu$ final states with respect to the standard model expectation.

9.2 Conclusion

The search for the standard model Higgs boson is an attempt to complete our picture of how fundamental particles interact and where they get their mass. Assuming the standard model of particle physics is the whole story, the Higgs mechanism is necessary to fully explain our observations. This analysis describes one small piece of this search when a Higgs boson produced from the collision of a proton and an antiproton, at center of mass energy s = $\sqrt{1.96}$ TeV, decays to two W bosons which then decay to two muons and missing energy. We find no evidence for the SM Higgs boson in this search channel at D0 and a limit was set on the mass, shown in the previous section. While the final result does not succeed in finding or excluding the SM Higgs boson alone, it is a vital part of a combined effort. This analysis also represents one of the most sensitive single analysis to date in the mass range studied.

In addition to doing a search on a specific decay mode of the Higgs boson, this analysis was performed with the intent of combining the result with other "high mass" [2] SM Higgs searches in the D0 Collaboration as well inclusive searches at both D0 and her sister experiment at the Tevatron, CDF. These combinations were performed by special working groups assembled for this purpose. The combined CDF/D0 SM Higgs search result is the first published collaboration between the two experiments in the RunII era of the Tevatron. This is a significant achievement in itself, aside from the impressive results. Both the D0 only and combined Higgs searches are published in Physical Review Letters as of the writing of this thesis, [101] and [24] respectively. Fig. 9.5, 9.6, and 9.7 show the combined LLR, limit ratio, and $1 - CL_s$ distributions respectively. This is the first published exclusion of the SM Higgs boson for any mass range outside of the LEPII result [23]. The quoted exclusion at 95% C.L. for a SM Higgs is $163 < m_H < 166$ GeV, with an expected exclusion of $159 < m_H < 168$ GeV.

---

[2] In this case "high mass" refers to other $H \rightarrow WW$ analyses, which are more sensitive to the higher mass range $m_H > 140$ GeV due to the production and decay mode cross sections.

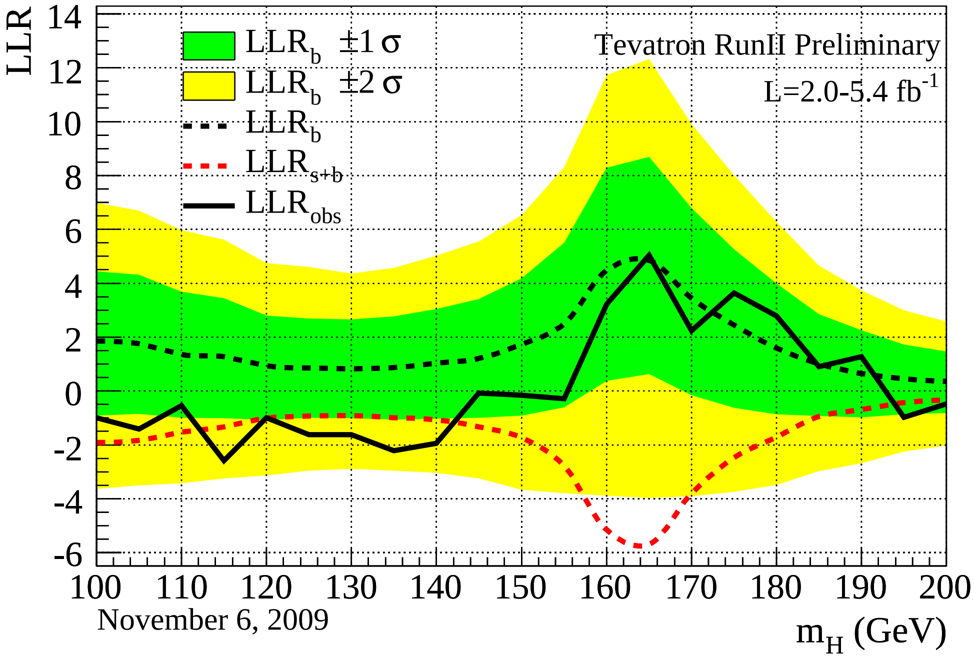

Figure 9.5: Log Likelihood Ratio for the expected and observed limit for the combined CDF and D0 analyses. The green and yellow shaded areas represent the $\pm 1\sigma$ and $\pm 2\sigma$ confidence levels for the background only hypothesis and the solid black graph represents the observed limit. The 'background-only' and 'signal + background' hypothesis are given by the black and red dashed lines, respectively.

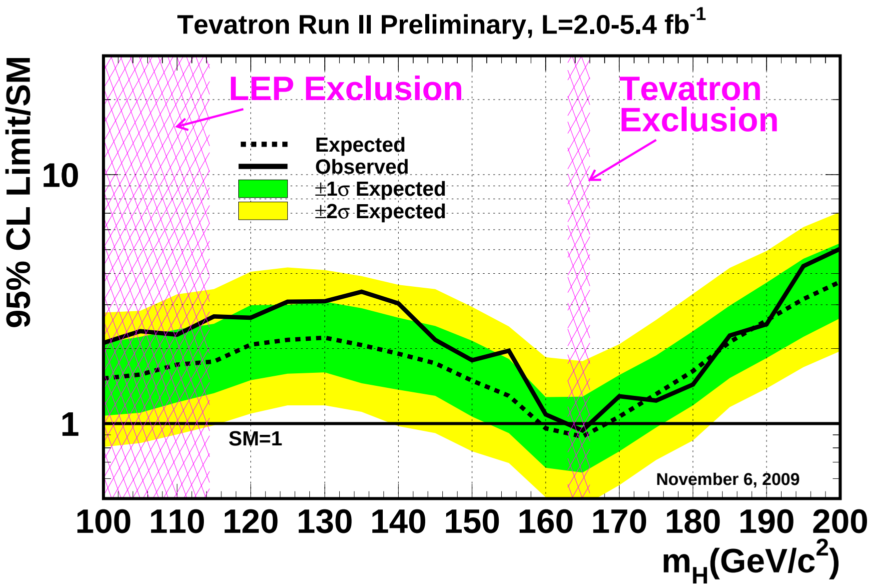

Figure 9.6: Expected and observed upper limits on the ratios to the SM cross section as functions of the Higgs boson mass for the combined CDF and D0 analyses.

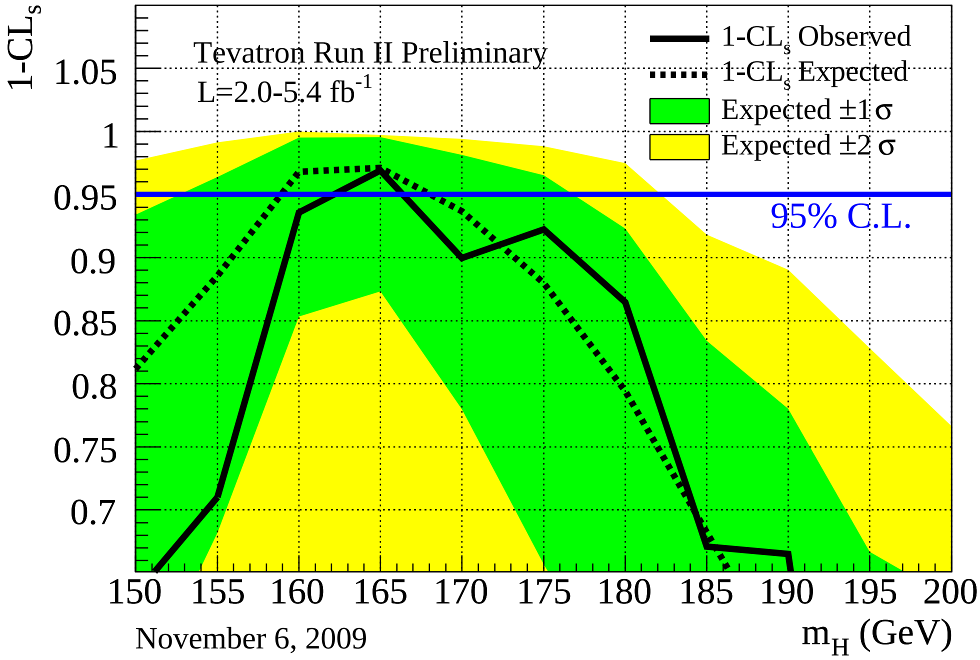

Figure 9.7: The exclusion strength $1 - CL_s$ as a function of the Higgs boson mass, as obtained with CL_s method. for the combination of the CDF and D0 analyses.

9.2.1 Improvements

While this analysis was done with the utmost care and attention to detail, there are many things that could be improved given more time. The most obvious "improvements" to be made are increased luminosity, better efficiencies, and better signal acceptance. The luminosity increase will simply come with time and the next section describes the prospects in that regard. The various efficiencies in the analysis, such as object ID efficiencies and trigger efficiencies are being improved by various working groups at D0. Also, by improving our understanding of the detector and our simulation of it we can loosen our analysis cuts to include more events and thus increase our acceptance. From there we can rely on more advanced multivariate techniques to separate the increased background.

This brings us to another possible place for improvement, the implementation of a more complex multivariate technique. Changing the inputs or training parameters of the Neural Network could afford some small improvements, but there are no major ($>$ few percent) improvements to be had there. One way to gain some sensitivity is to employ a technique that uses leading order matrix elements for the signal and important backgrounds to make a new discriminating variable. This Matrix Element(ME) technique essentially looks at the kinematic variables of the leptons and missing energy in each event and calculates a probability that it is signal or one of the backgrounds. This probability distribution can then be used as a discriminating variable by itself, or fed into something like the NN used in this analysis. The ME technique has been tried with some success in this analysis channel. While it has not been used in an official result yet, up to a 10% improvement might be possible using the ME in addition to the NN. Future iterations of this analysis will likely include a ME input variable.

Other multivariate techniques, such as the Random Forest method [102], are also currently being studied to replace the NN. One advantage to such techniques is the lack of a training procedure. With no training, it is unnecessary to throw away half of the simulated events (the training sample). This greatly increases the statistics of the background samples after analysis cuts where some of them are rather sparse (like the W+jets sample). Work is in progress to see how much there is to gain from such techniques.

Aside from improving the multivariate methods used for limit setting, we continue to work on better modeling of our backgrounds. A good example of this is the known issue with modelling Z/γ^* events using Alpgen in the low dilepton invariant mass region ($M_{inv} < 60$ GeV). This issue can be seen in Table 6.2 where there are too many simulated events as compared to data. While this does not affect the high signal region in the NN output distribution, it does impact our overall agreement between simulation and data and in later versions of this analysis an additional reweighting to the dilepton mass for Z events in the low mass region is applied to correct this issue [103].

9.2.2 Outlook

Future Higgs Searches at the Tevatron

At both the D0 and CDF experiments, the data recorded by the Summer of 2009 corresponds to about 6 $\text{fb}^{-1}$of integrated luminosity. By the end of 2011, 10 $\text{fb}^{-1}$of integrated luminosity should be attainable. With this additional luminosity and considering anticipated improvements in things like mass resolution, b-tagging, additional acceptance, and new analysis channels a sensitivity to the SM Higgs in a mass range of $m_H = 145 - 190$ GeV is likely. In the low mass region, the luminosities achievable in the most optimistic scenarios would just be sufficient to claim evidence but not an observation. While discovery is not likely at the Tevatron, The Large Hadron Collider(LHC) at CERN may have a significantly better chance if the SM Higgs exists.

Future 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. Fig. 9.8 shows the discovery significance for 10 fb^{-1}of data at the Atlas detector [104] and the integrated luminosity needed for a 5 σ discovery at CMS [105]. 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 pb^{-1} of data.

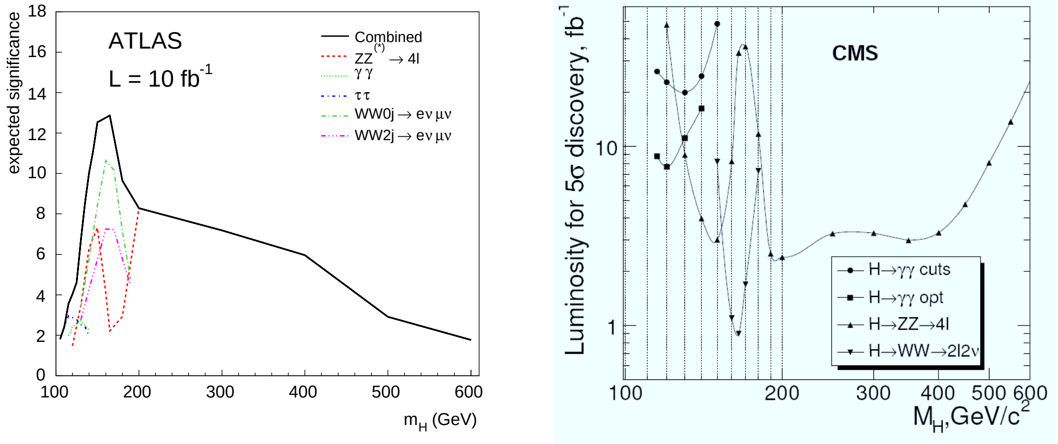

Figure 9.8: The median discovery significance of the Atlas detector for the various channels and the combination with an integrated luminosity of 10 fb^{-1}for for masses up to 600 GeV [104] (left). The integrated luminosity needed for the 5 σ discovery of the inclusive Higgs boson production $pp \to H + X$ with the Higgs boson decay modes $H \to \gamma\gamma$, $H \to ZZ \to 4l$ and $H \to WW \to 2l2\nu$ in the CMS experiment [105] (right).

While there has still been no direct observation of the Higgs boson despite the best efforts of particle physicists the world over, an important milestone has been reached. By achieving sensitivity for the standard model Higgs boson for the first time at a hadron collider, the stage is set to nail down this elusive prey. If it exists, evidence for the Higgs boson could be found at the Tevatron in the near future and a discovery is likely at the LHC. This discovery would bring us one step closer to seeing the structure of our universe on the smallest scales and hopefully an understanding of how it all fits together.

Appendix A

Limit Setting

The limit setting procedures in this analysis use the so called “$\mathrm{CL_s}$ Method” to determine a 95% confidence level(CL) limit on the cross section times branching ratio for various possible Higgs boson masses. The following sections will describe this method and how it is applied.

A.1 $\mathrm{CL_s}$ Method

There are two possibilities considered when calculating a limit, the background only hypothesis and the signal+background hypothesis. In the former, only the Standard Model(SM) background processes contribute to the accepted event rate. In the latter, the signal from some form of new physics (e.g., the Higgs boson) adds to the background. Confidence levels are then computed by comparing the observed data configuration to the expectations for these two hypotheses. The $\mathrm{CL_s}$ method uses the estimated signal, s_i , background, b_i , and the number of observed data events, n_i , in each bin in the calculation of confidence levels. In this analysis, the output of a Neural Network(NN) is the test-variable used in the calculation (Chapter 7).

We want to define a test statistic such that the two hypothesis mentioned above are as different as possible. This variable will then be used to compare them with the data. An optimal choice [98] is the likelihood ratio of Poisson probabilities:

$$Q = \frac{P_{\mathrm{poiss}}(data|signal + background)}{P_{\mathrm{poiss}}(data|background)} \tag{A.1}$$

where

$$P_{\mathrm{poiss}}(data|signal + background) = \prod_{i=1}^{n_{\mathrm{bins}}} \frac{(s_i + b_i)^{n_i} e^{-(s_i+b_i)}}{n_i!} \tag{A.2}$$

and

$$P_{\mathrm{poiss}}(data|background) = \prod_{i=1}^{n_{\mathrm{bins}}} \frac{(b_i)^{n_i} e^{-b_i}}{n_i!}. \tag{A.3}$$

The products run over all bins of all distributions to be combined (if there is more than one). Here the signal events in each bin, s_i, get all of the appropriate weights such as the expected signal cross-section and decay branching ratios, the integrated luminosity, and the detection efficiency for the signal. Likewise, the background estimation in each bin, b_i, depends on the SM background cross-sections, the integrated luminosity, and selection efficiencies. Finally, the number of observed events in bin i is just n_i. The test statistic can also be expressed in the logarithmic form (called the log-likelihood ratio or LLR):

$$\chi_n = -2\ln Q = 2\sum_{i=1}^{n_{\mathrm{bins}}} s_i - 2\sum_{i=1}^{n_{\mathrm{bins}}} n_i \ln(1 + s_i/b_i) \tag{A.4}$$

This LLR test statistic (seen in Fig. 9.2 and 9.4), χ_n, has the desirable benefit of being monotonically increasing in the number of candidates. In each bin, the ratio s_i/b_i must be kept finite, either by generating enough Monte Carlo statistics for signal and background or by rebinning or smoothing.

Next, to test the consistency of the data with the background-only hypothesis, the confidence level $1 - \mathrm{CL_b}$ is defined as

$$1 - \mathrm{CL_b} = P(\chi_N \geq \chi_{\mathrm{obs}}|background). \tag{A.5}$$

This is the fraction of experiments in a large ensemble of N background-only experiments which would produce results at least as background-like as the observed data. In practice, this means generating many pseudo-experiments to calculate this probability.

Similarly, to test the consistency of the data with the signal+background hypothesis, the confidence level $\mathrm{CL_{s+b}}$ is defined as

$$\mathrm{CL_{s+b}} = P(\chi_N \leq \chi_{\mathrm{obs}} | \mathit{signal} + \mathit{background}). \tag{A.6}$$

This is the fraction of experiments in a large ensemble of N signal+background experiments which would produce results less signal-like than the observed data. Again, many pseudo experiments are generated to create the ensemble. When defined in this way, a signal+background hypothesis is excluded at the 95% confidence level if $\mathrm{CL_{s+b}} < 0.05$.

One could use $\mathrm{CL_{s+b}}$ for exclusion, however it is possible for statistical downward fluctuations to lead to deficits of observed events that are not consistent with the expected background. This, in turn, can lead to an exclusion of the signal+background hypothesis even if the expected signal is too small to actually achieve experimental sensitivity along with other unphysical results [100]. To help alleviate the problem, one can define a confidence level $\mathrm{CL_s}$ as:

$$\mathrm{CL_s} = \mathrm{CL_{s+b}} / \mathrm{CL_b} \tag{A.7}$$

While using $\mathrm{CL_s}$ rather than $\mathrm{CL_{s+b}}$ does come with some sensitivity loss, it helps regulate the undesirable behavior and in no case is the limit more restrictive than from using $\mathrm{CL_{s+b}}$. We can therefore consider a signal hypothesis to be excluded at the 95% CL if $\mathrm{CL_s} < 0.05$. This is sometimes referred to as the "Modified Frequentist" or "semi-frequentist" model, since it reduces the dependence on the signal distribution.

Because all of the s_i , b_i, and observed data in each bin depend on the NN output test-variable, $\mathrm{CL_b}$, $\mathrm{CL_{s+b}}$, and $\mathrm{CL_s}$ all depend on it as well. For this analysis, it means the limit on the Higgs cross section times branching ratio at a given test mass is the smallest value such that $\mathrm{CL_s} \geq 0.05$. Once this limit is calculated, it is reported as a ratio to the theoretical SM cross section times branching ratio for the Higgs boson at that mass (see for example, Fig. 9.3). This means that a value of ≤ 1 represents exclusion of the theorized SM Higgs boson for that mass at the 95% CL. The calculation is performed using the software package `collie` [100] which uses a technique called profile maximization discussed in the next section.

A.2 Propagation of Uncertainties

The estimation of systematic uncertainties is always done cautiously and conservatively. This potential overestimation of the uncertainty naturally leads to the degradation of calculated limits. To help alleviate this, the magnitude of systematic uncertainties can be constrained by comparing the predicted shape and rate of the final variable, in this analysis the NN, to what is observed in data. One can measure this level of agreement by calculating a χ^2 for the final variable distribution, which is a function of the difference between the observed and predicted values and the uncertainties on the predicted values.

When setting limits, a technique called profile maximization or “profiling” can be used, where the best fit of the predicted model to data is determined. This is done by minimizing the χ^2 function over the possible values of parameters describing the systematic uncertainties. 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. During this fitting process all systematic uncertainties are varied separately. If a shape uncertainty (see section 8.2) is used, each bin is individually varied according to the corresponding fractional uncertainty.

Appendix B

Reweighting

Several factors contribute to the inability of Monte Carlo(MC) generated events to model observed data perfectly. Things like limitations in the knowledge of underlying theory, the precision of perturbative calculations, and the complex modeling of the D0 detector response contribute to this failure. To help correct these short-comings of our simulation, correction factors are derived from data and applied to the simulation in the form of event reweightings. The method for deriving and applying these reweightings is described in the following sections.

B.1 Deriving Reweighting Histograms

To derive appropriate reweighting histograms for each of the reweighted variables described in section 6.2 several steps are necessary. First, the various MC backgrounds are normalized by their event weights and summed to get a distribution representative of the total MC background. The QCD background is left out of the sum because it is derived from data (section 6.2.3). This also means that it should be subtracted from the data distribution[1]. Next, both the data and MC distributions for a given variable are normalized to unity by dividing each by its

---

[1] Here we are assuming that the QCD events form the like-sign data sample have the same approximate size and kinematics in the opposite sign data that we use in the analysis. This is the same assumption we make when deriving the QCD in the first place (see section 6.2.3).

own integral.

Once the distributions have been normalized so that only shape differences remain, the two histograms are smoothed and rebinned into semi-equalized probability bins, where each bin has at least half of the statistics of the bin with the most entries. Finally, the histograms are divided to get a ratio histogram which can be used to reweight the simulated background distributions. An example of the data/MC ratio with and without this rebinning technique can be found in Fig. B.1. The advantage of this rebinning is to reduce statistical fluctuations in the tails of distributions that might adversely affect the reweightings without relying on an arbitrary function to model the ratio.

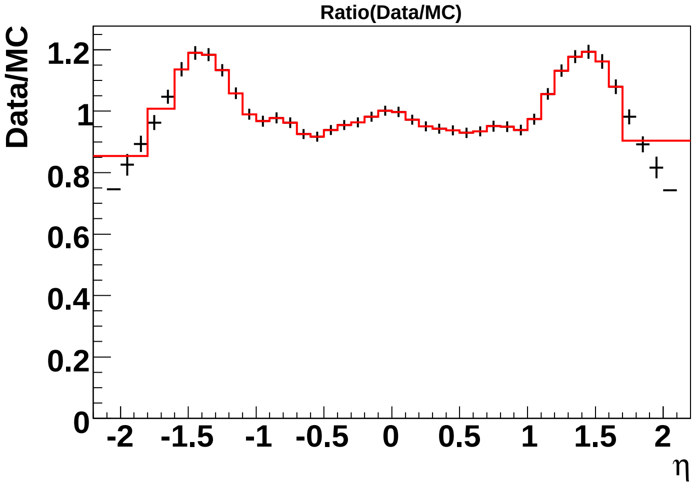

Figure B.1: An example comparison of the data/MC ratio with and without rebinning. The distribution used for the example is the pseudo-rapidity of the second highest momentum muon in each event. The solid red line is the rebinned ratio using semi-equalized probability bins and the black points are the ratio without rebinning.

B.2 Applying Reweightings

Applying the reweightings derived as in the previous section is relatively straightforward. The rebinned-ratio histograms are simply multiplied to the corresponding MC distributions for each background as an event weight. For each event, the ratio histogram is used to “look-up” the value of the weight to be applied. This is done for all events in all simulated samples. Any events that fall outside of the range of the ratio histogram simply get an event weight of 1. This should rarely happen, since the binning for the ratio histograms was chosen to include most, if not all, of the events from every background.

The order in which the reweightings are derived and applied is important and chosen in such a way that reweightings of correlated quantities are applied in consecutive steps. An example of this are the detector η and dimuon rapidity distributions as seen in Fig. 6.6, 6.7, and 6.8. It should be remembered that the reweighted distributions shown in those figures, along with all of the ones in Sec. 6.2 have all reweightings applied in the order specified and not just the reweighting of the distribution shown. Thus, correlation effects are present. These correlation effects are small compared to the systematics assigned to the reweightings and the multi-dimensional reweighting procedure necessary to account for these correlations was dismissed as impractical and unnecessary.

Bibliography

[1] L. Lederman. *The God Particle: If the Universe Is the Answer, What Is the Question?*. Dell Publishing, (1993).

[2] C. Amsler *et al.*, (Particle Data Group), Physics Letters B **667**, 1 (2008).

[3] O. W. Greenberg, “Spin and Unitary-Spin Independence in a Paraquark Model of Baryons and Mesons”, Phys. Rev. Lett. **13**, 598-602 (1964).

[4] D. J. Gross and F. Wilczek, “Asymptotically Free Gauge Theories.I”, Phys. Rev. D **8**, 3633 (1973).

[5] D. J. Gross and F. Wilczek, “Asymptotically Free Gauge Theories.II”, Phys. Rev. D **9**, 980 (1974).

[6] H. Georgi and H. D. Politzer, “Electroproduction scaling in an asymptotically free theory of strong interactions”, Phys. Rev. D **9**, 416 (1974).

[7] Y. Fukuda *et al.*, (Super-Kamiokande), Phys. Rev. Lett. **81**, 1562 (1998).

[8] M. H. Ahn *et al.*, (K2K), Phys. Rev. Lett. **90**, 041801 (2003).

[9] T. Araki *et al.*, (KamLAND), Phys. Rev. Lett. **94**, 081801 (2005).

[10] ALEPH, CDF, D0, DELPHI, L3, OPAL and SLD Collaborations, LEP and Tevatron Electroweak Working Groups, SLD Electroweak and Heavy Flavour Groups, *Precision Electroweak Measurements and Constraints on the Standard Model*, arXiv:0911.2604 [hep-ex].

[11] *The LEP Electroweak Working Group*, (2009). http://lepewwg.web.cern.ch/LEPEWWG/plots/winter2009/.

[12] **GSM**, “A Gfitter Package for the Global Electroweak Fit” 10 Jan. 2010. http://gfitter.desy.de/GSM/

[13] D. Griffiths. *Introduction to Elementary Particles.* Wiley, (1987).

[14] H. Goldstein, C. Poole, and J. Safko. *Classical Mechanics.* pages 589-598. Addison Wesley, (2002).

[15] F. Halzen, A. D. Martin. *Quarks and Leptons.* Wiley, (1984).

[16] C. Quigg, "Unanswered Questions in the Electroweak Theory," Ann. Rev. Nucl. Part. Sci. **59**, 505 (2009).

[17] A. Djouadi, "The Anatomy of electro-weak symmetry breaking. I: The Higgs boson in the standard model," Phys. Rept. **457**, 1 (2008).

[18] S. Dawson, "Introduction to Electroweak Symmetry Breaking," AIP Conf. Proc. **1116**, 11 (2009).

[19] T. Hambye and K. Riesselmann, "Matching conditions and Higgs boson mass upper bounds reexamined," Phys. Rev. D **55**, 7255 (1997).

[20] M. Veltman, "The Infrared - Ultraviolet Connection," Acta Phys. Polon. B **12**, 437 (1981).

[21] C.F. Kolda and H. Murayama, JHEP **0007** 035, (2000).

[22] CERN: European Organization for Nuclear Research. www.cern.ch

[23] LEP Working Group for Higgs boson searches, R. Barate *et al.*, Phys. Lett. B bf 565, 6175 (2003)

[24] [CDF Collaboration and D0 Collaboration], "Combined CDF and D0 Upper Limits on Standard Model Higgs-Boson Production with 2.1 - 5.4 $\text{fb}^{-1}$of Data," Phys. Rev. Lett. **104**, 061802 (2010).

[25] R. Boughezal and B. Tausk, "Three-loop electroweak corrections to the rho-parameter, sin$\hat{2}$ theta_eff and the W-boson mass in the large Higgs mass limit", Acta Phys.Polon. B **36** 3275 (2005).

[26] J. F. de Troconiz and F. J. Yndurain, Phys. Rev. D **71** 073008 (2005).

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

[28] 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**:5674, (1998).

[29] L. M. Lederman, "The Tevatron", Scientific American, **264(3)**, 48 (1991).

[30] H. T. Edwards, "The Tevatron Energy Doubler: A Superconducting Accelerator", Ann. Rev. Nucl. Part. Sci., **35**, 605 (1985).

[31] TeVI Group, "Design Report Tevatron I projects", Fermilab Technical Report FERMILAB-DESIGN-1984-01 (1984).

[32] F. T. Cole *et al.*, "A Report on the Design of the Fermi National Accelerator Laboratory Superconducting Accelerator", Beams Document 1888-v1 (1979).

[33] J. Thompson, "Introduction to Colliding Beams at Fermilab", Fermilab Technical Report FermiLab TM-1909 (1994).

[34] S. W. Herb *et al.*, “Observation of a Dimuon Resonance at 9.5 GeV in 400-GeV Proton-Nucleus Collisions”, Phys. Rev. Lett. **39**, 252255 (1977).

[35] S. Abachi *et al.*, (D0 Collaboration), “Observation of the Top Quark”, Phys. Rev. Lett. **74**, 26322637 (1995).

[36] F. Abe *et al.*, (CDF Collaboration), “Observation of Top Quark Production in $p\bar{p}$ Collisions with the Collider Detector at Fermilab”, Phys. Rev. Lett. **74**, 26262631 (1995).

[37] K. Kodama, *et.al.*, (DONUT Collaboration), “Observation of tau neutrino interactions”, Physics Letters B **504** 218-224 (2001).

[38] C. W. Schmidt and C. D. Curtis, “A 50 mA Negative Hydrogen-Ion Source”, IEEE Transactions on Nuclear Science, NS-26, 4120 (1979).

[39] C. D. Curtis *et al.*, “Linac H-Beam Operation and Uses at FermiLab”, IEEE Transactions on Nuclear Science, NS-26, 3760 (1979).

[40] E. L. Hubbard *et al.*, “Booster synchrotron”, Fermilab Technical Report Fermilab TM-405 (1973).

[41] C. S. Mishra, “The Fermilab Main Injector”, Fermilab Technical Report Fermilab Conf-92/372 (1992).

[42] S. Nagaitsev *et al.*, “Antiproton Cooling in the Fermilab Recycler Ring”, Fermilab Technical Report Fermilab-Conf-05-550-AD (2005).

[43] S. Abachi *et al.* (D0 Collaboration), “The D0 Detector”, Nucl. Inst. Methods **A338**, 185 (1994).

[44] S. Abachi *et al.* (D0 Collaboration), “The D0 Upgrade: The Detector and its Physics”, Fermilab Technical Report Fermilab-Pub-96/357-E (1996).

[45] V. M. Abazov *et al.* (D0 Collaboration), “The Upgraded D0 Detector”, Nucl. Inst. Methods **A565**, 463 (2006).

[46] M. Roco, “The Silicon Microstrip Tracker for the D0 Upgrade”, Internal D0 Note 3553 (1998).

[47] D0 Upgrade Collaboration, “D0 Silicon Tracker Technical Design Report” (1994).

[48] F. Borcherding *et al.* (D0 Collaboration), “Technical Design Report for the Upgrade L1/L2 Tracking Trigger - TDR for the CTT”, Internal D0 Note 3551 (1998).

[49] J. Brzezniak (D0 Collaboration), “Conceptual Design of a 2 Tesla Superconducting Solenoid for the Fermilab D0 Detector Upgrade”, Fermilab Technical Report FERMILAB-TM-1886 (1994).

[50] G. Bernardi *et al.*, “Improvements from the T42 Algorithm on Calorimeter Objects Reconstruction”, Internal D0 Note 4335 (2004).

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

[52] M. Abolins *et al.*, “The Run IIb Trigger Upgrade for the D0 Experiment” IEEE Trans. Nucl. Sci. **51**, 340344 (2004).

[53] V. M. Abazov *et al.* (D0 Collaboration), “Run IIb Upgrade Technical Design Report”, Fermilab Pub-02-327-E (2002).

[54] R. Fruhwirth, “Application of Kalman Filtering to Track and Vertex Fitting”, Nucl. Inst. Methods **A262**, 444 (1987).

[55] P. V. C. Hough, “Machine Analysis of Bubble Chamber Pictures”, Proc. Int. Conf. High Energy Accelerators and Instrumentation (1959).

[56] C. Tully, A. Schwartzman. “Primary Vertex Reconstruction by Means of Adaptive Vertex Fitting.”, 2005. D Note 4918.

[57] S. W. Cho *et al.*, “Muon ID Certification for p20 data”, Internal D0 Note 5824 (2008).

[58] P. Calfayan *et al.*, “Muon Identification Certification for p17 data”, Internal D0 Note 5157 (2007).

[59] G. C. Blazey *et al.*, “Run II Jet Physics”, arXiv:hep-ex/0005012 (2000).

[60] E. Busato and B. Andrieu, “Jet Algorithms in the D0 Run II Software: Description and User’s Guide”, Internal D0 Note 4457 (2004).

[61] Jet Energy Scale Working Group, “Jet Energy Scale Determination at D0 Run II (final p17 version)”, Internal D0 Note 5382 (2008).

[62] DQ Group. *D0 Data Quality Coordination.* http://www-d0.fnal.gov/computing/data_quality/.

[63] T. Andeen *et al.*, “The D0 Experiment’s Integrated Luminosity for Tevatron Run IIa”, FERMILAB-TM-2365 (2007).

[64] T. Sjostrand *et al.*, “High-Energy-Physics Event Generation with PYTHIA 6.1”, Comp. Phys. Commun., **135**:238259, (2001).

[65] F. Krauss and J.-C. Winter, “Initial-state showering based on colour dipoles connected to incoming parton lines.”, JHEP **0807**, 040 (2008).

[66] T. Sjöstrand, *et al.*, “High-Energy-Physics Event Generation with PYTHIA 6.1”, Computer Phys. Commun. **135** 238 (2001).

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

[68] B. Andersson, "The Lund Model", Cambridge University Press, (1998).

[69] J. Pumplin, *et al.*, "New Generation of Parton Distributions with Uncertainties from Global QCD Analysis", JHEP **0207**, 012 (2002).

[70] M. L. Mangano *et al.*, "ALPGEN, a generator for hard multiparton processes in hadronic collisions", JHEP **0307**, 001 (2003).

[71] T. Nunnemann. *MCFM Cross Sections.* http://www-clued0.fnal.gov/ñunne/cross-sections/mcfm cross-sections.html.

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

[73] Kidonakis *et al.*, "Top quark production at the Tevatron at NNLO", Eur. Phys. J. **C33**, s466-s468 (2004).

[74] W. L. van Neerven, R. Hamberg, T. Matsuura. Nucl. Phys., B **359**:343, (2002). Erratum-ibid. B **644**:403, (2002).

[75] T. Nunnemann. "NNLO Cross-Sections for Drell-Yan, Z and W Production using Modern Parton Distribution Functions", Internal D0 Note 4476 (2004).

[76] Y. Peters *et al.*, "Study of the W +jets heavy flavor scale factor in p17." Internal D0 Note 5406 (2007).

[77] D0 Higgs Physics Group. *Compilation of Higgs cross sections and branching provided by the D0 Higgs group.* http://www-d0.fnal.gov/Run2Physics/higgs/higgs_xs_br.html.

[78] M. Spira, "HIGLU: A Program for the Calculation of the Total Higgs Production Cross Section at Hadron Colliders via Gluon Fusion including QCD Corrections," arXiv:hep-ph/9510347 (1995).

[79] V+Jets Working Group, "`vjets_cafe` package". https://plone4.fnal.gov/P1/D0Wiki/physics/VplusJets/

[80] M. Arthaud *et al.*, "Muon Momemtum Oversmearing for p17 Data", Internal D0 Note 5444 (2007).

[81] M. Arthaud *et al.*, "Muon Momemtum Oversmearing for p20 Data", Internal D0 Note 5449 (2007).

[82] T. Binoth *et al.*, "Gluon induced W W background to Higgs boson searches at the LHC", arXiv:hep-ph/0503094.

[83] N. Kauer. "GG2WW parton-level integrator and event generator", http://hepsource.sourceforge.net/GG2WW/.

[84] T. Gleisberg *et al.*, JHEP **07**, 001 (2003).

[85] B. R. Webber, S. Frixione, "The MC@NLO 3.3 Event Generator", arXiv:hep-ph/0612272v1 (2006).

[86] StatSoft, "Electronic Statistics Textbook", StatSoft Inc. (2009). http://www.statsoft.com/textbook/neural-networks/.

[87] T. Gadfort *et al.*, "Performance of the Dzero NN b-tagging Tool on p20 Data", Internal D0 Note 5554 (2007).

[88] A. Hocker *et al.*, "TMVA (Toolkit for Multivariate Data Analysis with ROOT) Users Guide", CERN-OPEN-2007-007, (2008).

[89] ROOT. http://root.cern.ch

[90] T. Andeen *et al.*, "The D0 Experiment's Integrated Luminosity for Tevatron Run IIa", Internal D0 Note 5398 (2007).

[91] S. Catani *et al.*, JHEP, **028**:0307 (2003).

[92] K. A. Assamagan *et al.*, "The Higgs Working Group: Summary Report", arxiv:hep-ph/0406152 (2003).

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

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

[95] Luminosity Working Group. "Adjustments to the Measured Integrated Luminosity in Run IIa", Internal D0 Note 5139 (2007).

[96] T. Gleisberg *et al.*, JHEP **07**:001 (2003).

[97] T. Junk, "Confidence Level Computation for Combining Searches with Small Statistics", Nucl. Instrum. Meth. A **434**, 435 (1999).

[98] V. Buescher *et al.*, "Recommendation of the Ad-Hoc Committee on Limit-Setting Procedures to be Used by D0 in Run II", Internal D0 Note 4629 (2004).

[99] W. Fisher, "Calculating Limits for Combined Analyses", Internal D0 Note 4975 (2006).

[100] W. Fisher, "Systematics and Limit Calculations", FERMILAB-TM-2386-E (2006).

[101] V. M. Abazov *et al.* [The D0 Collaboration], "Search for Higgs boson production in dilepton and missing energy final states with 5.4 $\mathbf{fb^{-1}}$ of $\boldsymbol{p\bar{p}}$ collisions at $\boldsymbol{\sqrt{s} = 1.96}$ TeV", arXiv:1001.4481 [hep-ex].

[102] L. Breiman, “Random Forests”, Machine Learning 45 (2001). http://www.stat.berkeley.edu/ breiman/RandomForests/

[103] A. Dominguez *et al.*, “Search for the Higgs Boson in Di-muon plus Missing Transverse Energy Final States with 5.4 $\text{fb}^{-1}$of Data”, Internal D0 Note 5963 (2009).

[104] Atlas Collaboration. “Expected Performance of the ATLAS Experiment, Detector, Trigger and Physics”, CERN-OPEN-2008-020 (2008).

[105] CDF Collaboration. “Technical Design Report, Volume II: Physics Performance”, CERN/LHCC 2006-021.