

## Abstract

# A Search for the Standard Model Higgs Boson in CDF II Data

Sarah E. Lockwitz

2012

This dissertation presents a search for the standard model Higgs boson in the associated production process  $p\bar{p} \rightarrow ZH \rightarrow e^+e^-b\bar{b}$ . Data amounting to an integrated luminosity of  $7.5 \text{ fb}^{-1}$  at  $\sqrt{s} = 1.96 \text{ TeV}$  collected at the Collider Detector at Fermilab (CDF) at the Tevatron are analyzed. Two objectives are pursued in the methods applied: maximize acceptance, and distinguish the signal from background. The first aim is met by applying a neural-network-based electron identification and considering multiple electron triggers in an effort to improve  $Z$  acceptance. In an attempt to maximize the Higgs acceptance, three  $b$  quark identification schemes are used allowing for varying event conditions. The latter goal is met by employing more multivariate techniques. First, the dijet mass resolution is improved by a neural network. Then, both single variables and boosted decision tree outputs are fed into a segmented final discriminant simultaneously isolating the signal-like events from the  $Z$  with additional jets background and the kinematically different  $t\bar{t}$  background. Good agreement is seen with the null hypothesis and upper production cross section ( $\sigma_{ZH}$ ) times branching ratio ( $BR(H \rightarrow b\bar{b})$ ) limits are set for 11 mass hypotheses between 100 and 150  $\text{GeV}/c^2$  at the 95% confidence level. For a Higgs boson mass of 115  $\text{GeV}/c^2$ , this channel sets an observed (expected) upper limit of 3.9 (5.8) times the standard model value of  $\sigma_{ZH} \times BR(H \rightarrow b\bar{b})$ . The inclusion of this channel within the combined CDF and Tevatron limits is discussed.

**A Search for the Standard Model Higgs  
Boson in CDF II Data**

A Dissertation  
Presented to the Faculty of the Graduate School  
of  
Yale University  
in Candidacy for the Degree of  
Doctor of Philosophy

by  
Sarah E. Lockwitz

Dissertation Director: Paul L. Tipton

May 2012

Copyright © 2012 by Sarah E. Lockwitz

All rights reserved.

To graduating...

# Contents

<b>1</b>	<b>Introduction</b>	<b>5</b>
<b>2</b>	<b>Motivation: The Standard Model and the Higgs Mechanism</b>	<b>8</b>
2.1	Standard Model Introduction . . . . .	8
2.1.1	Fundamental Particles and Forces . . . . .	8
2.1.1.1	Fermions . . . . .	9
2.1.1.2	Bosons: Force Mediators . . . . .	10
2.2	Standard Model Formalism . . . . .	12
2.3	Electroweak Symmetry Breaking . . . . .	15
2.3.1	Illustration of Principles . . . . .	15
2.4	Glashow-Weinberg-Salam Theory of Weak Interactions . . . . .	17
2.5	Limits on the Higgs Boson Mass . . . . .	20
2.5.1	Theory . . . . .	20
2.5.2	Experiment . . . . .	21
2.6	Higgs Phenomenology . . . . .	23
2.6.1	Higgs Production and Decay . . . . .	23
2.6.2	Backgrounds to Search . . . . .	25
<b>3</b>	<b>The Tevatron and the Collider Detector at Fermilab</b>	<b>27</b>
3.1	The Tevatron . . . . .	27

3.2	The Collider Detector at Fermilab: CDF . . . . .	30
3.2.1	Tracking . . . . .	31
3.2.1.1	Solenoid . . . . .	31
3.2.1.2	Silicon . . . . .	32
3.2.1.3	Central Outer Tracker (COT) . . . . .	33
3.2.2	Calorimetry . . . . .	34
3.2.2.1	Central Electromagnetic Calorimetry (CEM) . . . . .	35
3.2.2.2	Plug Electromagnetic Calorimetry (PEM) . . . . .	35
3.2.2.3	Central Hadronic Calorimetry . . . . .	36
3.2.2.4	Plug Hadronic Calorimeter . . . . .	37
3.3	Muon System . . . . .	37
3.4	Cherenkov Luminosity Counter . . . . .	40
3.5	CDF's Data Acquisition . . . . .	41
<b>4</b>	<b>Event Selection and Reconstruction</b>	<b>43</b>
4.1	Data . . . . .	43
4.1.1	Triggers Used . . . . .	43
4.1.2	Event Credibility Checks . . . . .	45
4.1.2.1	Cosmic Muon Filter . . . . .	45
4.1.2.2	Good Run Check . . . . .	45
4.2	Event Reconstruction . . . . .	46
4.2.1	Electron Selection & $Z$ Reconstruction . . . . .	46
4.2.1.1	Electron Identification . . . . .	46
4.2.1.2	Reconstructing a $Z$ Boson . . . . .	49
4.2.2	Jet Selection & Higgs Candidate Reconstruction . . . . .	54
4.2.2.1	$b$ Tagging and Final Analysis Channels . . . . .	55

4.2.2.2	Neural Network Jet-Energy Corrections . . . . .	56
4.3	Model . . . . .	58
4.3.1	Monte Carlo . . . . .	58
4.3.2	Trigger Modeling . . . . .	59
4.3.3	Electron Corrections . . . . .	62
4.3.3.1	Electron Efficiency Scale Factors . . . . .	62
4.3.3.2	Data Energy Scale Factors . . . . .	65
4.3.3.3	Plug-Energy Smearing . . . . .	67
4.3.4	Misidentified Electrons . . . . .	68
4.3.5	Mistagged $b$ Jets . . . . .	70
4.3.6	$b$ -Tag Scale Factors . . . . .	71
4.3.7	Model Validation . . . . .	72
4.3.7.1	Pretag Model Validation . . . . .	72
4.3.7.2	Final Analysis Channels: Tag Level Model . . . . .	74
<b>5</b>	<b>Systematic Uncertainties</b>	<b>78</b>
5.1	Rate Uncertainties . . . . .	78
5.2	Shape Uncertainties . . . . .	80
<b>6</b>	<b>Final Analysis Discriminant</b>	<b>83</b>
6.1	Super Bins . . . . .	83
6.2	Discriminant . . . . .	84
<b>7</b>	<b>Results</b>	<b>91</b>
7.1	Limit Calculation . . . . .	91
7.2	Combination with $ZH \rightarrow \mu^+\mu^-b\bar{b}$ . . . . .	95
7.3	CDF and Tevatron Combinations . . . . .	96

<b>8 Conclusion</b>	<b>98</b>
8.1 Notes on Future Work . . . . .	99
<b>A Introduction to Neural Networks</b>	<b>102</b>
<b>B Pretag Distributions</b>	<b>110</b>
<b>C Final Analysis Channel Distributions</b>	<b>115</b>
<b>D Final Analysis Discriminants</b>	<b>121</b>
<b>E Individual Channel Limits</b>	<b>125</b>

# Chapter 1

## Introduction

The Greeks proposed that everything was made of four elements: earth, fire, water, and air. While this was eventually rejected, it does illustrate an aspect of human nature: the desire to know what makes up the world around us. This feeds into a desire to understand how the universe works. This is what the study of physics attempts to answer.

The Greeks are also responsible for supplying the word “atom” to represent the smallest constituent of matter. This notion was later rejected as well with the discovery of the atomic structure. However, the desire to know the fundamental building blocks of nature and understand how they interact lives on. These are the topics addressed in particle physics.

Theories in particle physics have evolved over the years with ever-increasing knowledge from experiment. The sum total of this effort developed the standard model. This model attempts to encapsulate all that is known about fundamental particles and their interactions. It has been wildly successful predicting various phenomena and explaining others. It has some shortfalls, however. One, that serves as the inspiration for this dissertation, is that it does not explain why certain force

carriers have mass. A very simple mechanism for this was proposed in the 1960s, and, since it was so simple and produced the desired results on paper, evidence of this mechanism has been pursued for decades. The evidence would be a particle: the Higgs boson. Indirect evidence and previous exclusions suggest the Higgs boson's mass ( $m_H$ ) is between 100-200 GeV/ $c^2$ . This energy regime is within the reach of the Tevatron at Fermilab, however, theory also predicts the frequency of background processes to be large.

This dissertation describes a search for the standard model Higgs boson in the associated production process  $p\bar{p} \rightarrow ZH \rightarrow e^+e^-b\bar{b}$ . The frequency of this process is only relevant between masses of 100-150 GeV/ $c^2$ , however, the analysis is combined with other searches to evaluate the larger mass window. Again, the backgrounds are very large for this signature ( $S/B \sim 0.0045$ ). This necessitates the use of sophisticated techniques to make a useful statement on the existence of a Higgs boson. The two goals pursued in this regard are to maximize acceptance, and to distinguish the signal.

The signal consists of two  $b$  jets and two electrons. To allow for a variety of  $b$ -identification scenarios, three  $b$ -identification schemes are allowed and treated as separate analysis channels due to their differing  $S/B$ . To improve electron identification over previous cut-based efforts, a novel neural-network identification scheme is employed. To distinguish the signal, first a neural network is applied to improve the dijet mass ( $\sim m_H$ ) resolution. Then, boosted decision tree outputs along with single variables are used as inputs to a final neural-network discriminant that distinguishes the  $ZH$  signal from varying backgrounds. If there is no observed signal-like event excess, upper bounds can be set on the production cross section times branching ratio ( $\sigma_{ZH} \times BR(H \rightarrow b\bar{b})$ ) using these final discriminant shapes. This is done for 11 mass hypotheses between 100 and 150 GeV/ $c^2$ .

The following Chapter gives the theoretical motivation for this search and includes previous bounds on the Higgs boson mass. Chapter 3 describes the experimental apparatus. Both the Tevatron, which produced the analyzed data, and the Collider Detector at Fermilab (CDF), which collected the data, are described here. Chapter 4 is considerable; it describes the event selection and reconstruction including the model description. The treatment of systematic effects is explained in Chapter 5. Then, Chapter 6 describes the final analysis discriminant. The results of the search along with their inclusion in larger search efforts is reported in Chapter 7. The final Chapter has the conclusions of the analysis and a brief discussion on the future of this search channel.

# Chapter 2

## Motivation: The Standard Model and the Higgs Mechanism

### 2.1 Standard Model Introduction

Despite its unassuming name, the standard model (SM) of particle physics is the primary theory we have today describing all of the fundamental particles and their interactions (except for gravity). It was developed in the 1970s with the unification of the electromagnetic and weak forces by Glashow, Salam, and Weinberg. It has been developed and tested extensively with much success; roughly ten Nobel prizes have been awarded for work on the SM since 1976 [1]. While it is not without its flaws, this is our working model of the world at its most fundamental level. It is the model experimentalists today test and is being examined in this thesis.

#### 2.1.1 Fundamental Particles and Forces

The standard model attempts to incorporate all of the known fundamental particles and all of the forces with the exception of gravity. The fundamental particles are

typically classified into two types: fermions and bosons.

### 2.1.1.1 Fermions

Fermions are matter particles. They have spin  $1/2$  and follow Fermi-Dirac statistics meaning that they obey the Pauli exclusion principle. Fermions are further classified into leptons and quarks.

The most well-known lepton is the electron. It is charged, stable, and can easily be a free particle. It is low-lying fruit in the particle physics world and is therefore the most familiar particle. However, in the 1930s, it was discovered that there is a heavier version, or flavor, of the electron. This particle, called the muon ( $\mu$ ) [2], seemed extraneous leading Nobel laureate I. I. Rabi to famously quip, “Who ordered that?[3]” Additionally, there is an even heavier flavor called the tau ( $\tau$ ) [4]. These, together with their anti-particles, are the charged leptons.

To explain the energy spectra of  $\beta$  decay (to conserve energy), Wolfgang Pauli proposed in the 1930s a particle that was simply escaping detection and carrying away energy. This particle was later discovered and aptly named neutrino, literally meaning “little neutral one.” Neutrinos are flavored ( $\nu_e, \nu_\mu, \nu_\tau$ ), electrically neutral particles with very little mass ( $\bar{\nu}_e < 2 \text{ eV}/c^2$  [5]). While the study of neutrinos, namely flavor conservation, flavor oscillation and their mass, is currently a very rich and active field in particle physics, in the study covered in this thesis, they simply escape the detector with no trace. The three flavors of neutrinos, along with their anti-particles, make up the neutral leptons.

The other class of fermions is called quarks. Among other particles, combinations of quarks make up familiar composite particles like the proton and neutron. Quarks have non-integer electric charge ( $\pm\frac{2}{3}e, \pm\frac{1}{3}e$ ), a color charge, and also have different flavor generations. The color charge is a construct to describe a charge with three

degrees of freedom instead of one. The common way to denote color charge is  $r$  (red),  $g$  (green),  $b$  (blue), and their anti-colors ( $\bar{r}$ ,  $\bar{g}$ ,  $\bar{b}$ ). There are six flavors of quarks: up ( $u$ ), down ( $d$ ), strange ( $s$ ), charm ( $c$ ), top ( $t$ ), and bottom ( $b$ ). Quark properties, along with the other fermions, are summarized in Table 2.1.

Leptons			Quarks		
Lepton	Electric Charge	Mass	Quark	Electric Charge	Mass
$e$	1	$\sim 0.511 \text{ MeV}/c^2$	$u$	$\frac{2}{3}$	$1.7\text{-}3.1 \text{ MeV}/c^2$
$\nu_e$	0	$< 2 \text{ eV}/c^2$	$d$	$-\frac{1}{3}$	$4.1\text{-}5.7 \text{ MeV}/c^2$
$\mu$	1	$\sim 105.66 \text{ MeV}/c^2$	$c$	$\frac{2}{3}$	$1.29^{+0.05}_{-0.11} \text{ GeV}/c^2$
$\nu_\mu$	0	$< 0.19 \text{ MeV}/c^2$	$s$	$-\frac{1}{3}$	$100^{+30}_{-20} \text{ MeV}/c^2$
$\tau$	1	$\sim 1777 \text{ MeV}/c^2$	$t$	$\frac{2}{3}$	$172.9 \pm 0.6 \pm 0.9 \text{ GeV}/c^2$
$\nu_\tau$	0	$< 18.2 \text{ MeV}/c^2$	$b$	$-\frac{1}{3}$	$4.19^{+0.18}_{-0.06} \text{ MeV}/c^2$

Table 2.1: A summary of fermion values. The horizontal lines designate the three generations of fermions and quarks. The electric charge is given in units of the electron’s charge. For each of the particles, there is also an antiparticle. The quoted upper limit for  $\nu_e$  is actually the  $\bar{\nu}_e$  value. The values are from [5].

### 2.1.1.2 Bosons: Force Mediators

Bosons are the other class of particles. These are the force mediator particles. They have integral spin and obey Bose-Einstein statistics. The most well-known boson is the photon ( $\gamma$ ). We can see evidence of this all around us (you are reading this page due to scattered photons). The other bosons are the  $W^\pm$ ,  $Z$ , gluons, and the theoretical graviton (not included in the SM) and their properties are summarized in Table 2.2. The four known forces are the gravitational, electromagnetic, strong, and weak forces. Again, even though gravity is the most familiar force to humans, it is roughly 32-39 orders of magnitude weaker than the other forces at the length scale in particle physics [6], and is therefore neglected in the standard model.

The electromagnetic force is probably the next most familiar. It is mediated by

Boson	Force Mediated	Mass (GeV/ $c^2$ )
$\gamma$	Electromagnetic	0
$Z$	Electroweak	91.2
$W^\pm$	Weak	80.4
$g$	Strong	0

Table 2.2: A summary of boson values. The masses are quoted to only one decimal place and are from [5].

the photon, and besides keeping magnets on one’s fridge, it is responsible for, among other effects, binding atoms and molecules together, and keeping one from passing through his or her chair while reading this.

The strong force is mediated by gluons ( $g$ ) and keeps the nucleus together and creates bound states of quarks called hadrons. Hadrons come in two varieties: baryons (groups of three valence quarks) and mesons (groups of two valence quarks — one quark and one antiquark). Baryons are particles like protons and neutrons while meson examples are pions and kaons. The term “valence quarks” is used because the quarks in the particle can be thought of as existing in a sea of gluons with many quarks and anti-quarks. In this case, only the net content is what distinguishes a particle. Nature only seems to allow colorless states so baryons have combinations like  $rbg$  or  $\bar{r}\bar{b}\bar{g}$  whereas mesons could have a color state of  $r\bar{r}$  (or any color, anti-color). This colorless state requirement leads to an interesting effect. No one has ever seen a bare quark. An experimentalist once claimed it would be possible to observe bare quarks with a setup involving two niobium spheres. This was never reproducible and led Nobel laureate Leon Lederman to write of the t-shirts students would wear at the research university reading, “You have to have niobium balls if you want to trap quarks [7].” This colorless state requirement is formulated by describing quarks as having asymptotic freedom [8, 9]. This means, as one attempts to pull quarks apart in a bound state, the strong force increases. When this is taken to the extreme, the

energy from the force becomes so great that particles are created out of the vacuum preserving a colorless state. (When this happens in a collision, we call the resultant spray of particles a jet (see Section 4.2.2)).

The strong force was named for keeping the nucleus together, so when  $\beta$  decay was discovered, the theme continued and the responsible force was named the weak force. This force does not form bound states; it allows particles to change flavor. It is mediated by the  $W^\pm$  boson for electromagnetically charged weak interactions and the  $Z$  boson for electromagnetically neutral weak interactions.

## 2.2 Standard Model Formalism

The standard model is a gauge field theory meaning that local symmetries (or conserved quantities by Noether's theorem) allow for phase transformations that leave the Lagrangian, and therefore measurable quantities, invariant. The symmetry groups used to describe the standard model are

$$SU(3)_C \times SU(2)_L \times U(1)_Y \tag{2.1}$$

The  $SU(3)$  group represents the color interactions, and the  $SU(2) \times U(1)$  group represents the electroweak interactions. Each of these groups imply certain symmetries must be maintained in the Lagrangian; transformations of the field  $\phi(x) \rightarrow U(x)\phi(x)$  for

$$\begin{aligned} U(1) : \phi(x) &\rightarrow e^{-i\alpha(x)}\phi(x) \\ SU(N) : \phi(x) &\rightarrow e^{-ig\alpha^a(x)T^a}\phi(x) \end{aligned} \tag{2.2}$$

Where  $T^a$  is a generator matrix of the group must leave the Lagrangian invariant. With these symmetries in mind, we can con-

struct Lagrangians to describe interactions. The Lagrangian for electrodynamics looks like

$$\mathcal{L} = -\frac{1}{4}(F_{\mu\nu})^2 + |D_\mu\phi|^2 - V(x) \quad (2.3)$$

where  $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$  is the electromagnetic field tensor,  $D_\mu = \partial_\mu + ieA_\mu$ ,  $A_\mu$  is a component of the electromagnetic field, and  $V(x)$  is a potential.

From the Lagrangian, the action is found and path integrals can be calculated to find matrix elements of measurable quantities of interest. To aid in these calculations, Richard Feynman developed pictorial aids now called Feynman diagrams [10]. The diagrams represent the process of interest and each line segment and interaction point translate to factors or terms in the integral. They are so useful and widespread in particle physics, Julian Schwinger remarked, “Like the silicon chip of more recent years, the Feynman diagram was bringing computation to the masses [11].” For a primer on Feynman diagrams, please see Chapter 5 of [12], or nearly any quantum field theory textbook. Examples are shown in Figures 2.5 and 2.6.

Using this model for the  $U(1)$  electrodynamics component has been quite successful. Values like  $g-2$  of the muon and electron, the Lamb energy shift, the decay rates of positronium, and hydrogen hyperfine splitting energies and transition energies all show precise agreement between theory and experiment. However, when extended to  $SU(2)$ , the model has a glaring weakness. It describes massless force carriers. This is appropriate for the photon but not for the massive  $W$  and  $Z$  bosons. Since a force’s range goes like  $\sim 1/M$  and the weak force’s range is small, it was known that the  $W$  and  $Z$  bosons must have a non-zero mass even before they were discovered. Simply injecting a mass term ( $\sim m_A^2 A_\mu A^\mu$ ) into the Lagrangian does not preserve the required symmetry as can be seen by:

$$\mathcal{L} = -\frac{1}{4}(F_{\mu\nu})^2 + |D_\mu\phi|^2 + \frac{1}{2}m_A^2 A_\mu A^\mu \quad (2.4)$$

Then these transformations should leave the Lagrangian invariant

$$\begin{aligned} \phi(x) &\rightarrow e^{i\alpha(x)}\phi(x) \\ A_\mu(x) &\rightarrow A_\mu(x) - \frac{1}{e}\partial_\mu\alpha(x) \end{aligned} \quad (2.5)$$

The  $F_{\mu\nu}$  term goes like

$$\begin{aligned} -\frac{1}{4}(F_{\mu\nu})^2 &= -\frac{1}{4}(\partial_\mu A_\nu - \frac{1}{e}\partial_\mu\partial_\nu\alpha(x) - \partial_\nu A_\mu + \frac{1}{e}\partial_\nu\partial_\mu\alpha(x))^2 \\ &= -\frac{1}{4}(\partial_\mu A_\nu - \partial_\nu A_\mu - \frac{1}{e}\partial_\mu\partial_\nu\alpha(x) + \frac{1}{e}\partial_\nu\partial_\mu\alpha(x))^2 \\ &= -\frac{1}{4}(\partial_\mu A_\nu - \partial_\nu A_\mu)^2 \end{aligned} \quad (2.6)$$

The second term looks like

$$\begin{aligned} |D_\mu\phi|^2 &= |\partial_\mu e^{i\alpha(x)}\phi + ie(A_\mu - \frac{1}{e}\partial_\mu\alpha(x))e^{i\alpha(x)}\phi|^2 \\ &= |i\phi e^{i\alpha(x)}\partial_\mu\alpha(x) + e^{i\alpha(x)}\partial_\mu\phi + ieA_\mu e^{i\alpha(x)}\phi - i(\partial_\mu\alpha(x))e^{i\alpha(x)}\phi|^2 \\ &= |e^{i\alpha(x)}(\partial_\mu\phi + ieA_\mu\phi)|^2 = |\partial_\mu\phi + ieA_\mu\phi|^2 \end{aligned} \quad (2.7)$$

However, the mass term is not invariant

$$\begin{aligned} A_\mu A^\mu &= (A_\mu - \frac{1}{e}\partial_\mu\alpha(x))(A^\mu - \frac{1}{e}\partial^\mu\alpha(x)) \\ &= A_\mu A^\mu - \frac{1}{e}(\partial_\mu\alpha(x))A^\mu - \frac{1}{e}A_\mu\partial^\mu\alpha(x) + \frac{1}{e^2}(\partial_\mu\alpha(x))(\partial^\mu\alpha(x)) \\ &= A_\mu A^\mu + \text{Not invariant} \end{aligned} \quad (2.8)$$

This derivation explains the gauge problems that occur with a massive photon. The problems that arise for  $W$  and  $Z$  bosons are analogous but more difficult. A popular

approach to address this is described in the next Section.

## 2.3 Electroweak Symmetry Breaking

To preserve the standard model symmetries, a different approach is needed to include the masses of the weak force carrier bosons. One popular approach is to introduce a symmetric potential that breaks the symmetry when expanded about the minima. This method was developed into electroweak symmetry breaking (EWSB) in the 1960s by Brout, Englert, Higgs, Kibble, Guralnik, and Hagen [13], [14], [15]. However, it is often simply called the Higgs mechanism, perhaps to the dismay of some of the others. The method was developed earlier for the theory of superconductivity by Landau and Ginzburg and has similarities to an even earlier description of the Debye cloud in electrolyte solutions [16].

First, an example of an Abelian case will illustrate the principles of EWSB. This will then be followed by how EWSB fits into the standard model. All examples follow from the notation and motivation in Chapter 20 of [12] and Chapter 2 of [17].

### 2.3.1 Illustration of Principles

Let us again consider the Lagrangian

$$\mathcal{L} = -\frac{1}{4}(F_{\mu\nu})^2 + |D_\mu\phi|^2 - V(\phi) \quad (2.9)$$

for  $\phi$  a complex scalar field. This time, take the potential to be

$$V(\phi) = -\mu^2\phi^*\phi + \frac{\lambda}{2}(\phi^*\phi)^2. \quad (2.10)$$

This potential is illustrated in Figure 2.1 and when rotated in three dimensions,

resembles the bottom of a wine bottle, a punt. The minimum of this potential is at

$$\phi_0 = \sqrt{\frac{\mu^2}{\lambda}}. \quad (2.11)$$

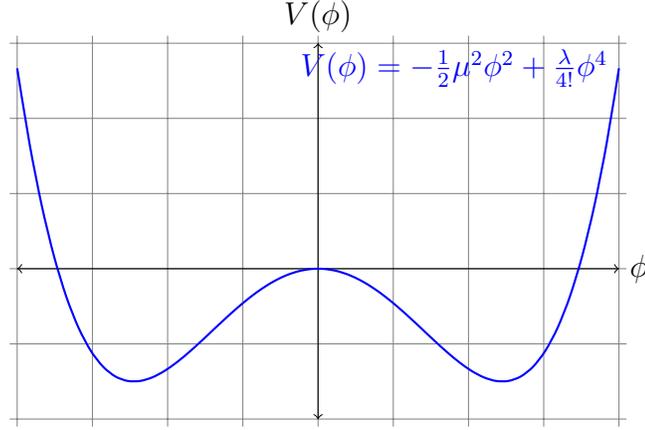


Figure 2.1: The symmetry-breaking potential.

Rewriting the field in terms of  $\phi_0$  and the real and complex field components yields

$$\phi(x) = \phi_0 + \frac{1}{\sqrt{2}} (\phi_1(x) + i\phi_2(x)). \quad (2.12)$$

Putting this into the potential, Equation 2.10, and dropping terms higher than field squared gives

$$V(\phi) = -\frac{1}{2} \frac{\mu^4}{\lambda} + \frac{1}{2} (\sqrt{2}\mu)^2 \phi_1^2 + \text{higher orders}. \quad (2.13)$$

Additionally, the kinetic energy term  $|D_\mu\phi|^2$ ,

$$\begin{aligned}
|D_\mu\phi|^2 &= \left[ \frac{1}{\sqrt{2}}\partial_\mu(\phi_1 + i\phi_2) + ieA_\mu(\phi_0 + \frac{1}{\sqrt{2}}(\phi_1 + i\phi_2)) \right] \\
&\quad \times \left[ \frac{1}{\sqrt{2}}\partial^\mu(\phi_1 - i\phi_2) - ieA^\mu(\phi_0 + \frac{1}{\sqrt{2}}(\phi_1 - i\phi_2)) \right] \\
&= \frac{1}{2} [(\partial_\mu\phi_1)^2 + (\partial_\mu\phi_2)^2] + eA^\mu\phi_2\partial_\mu\phi_1 - \sqrt{2}eA^\mu\phi_0\partial_\mu\phi_2 - eA^\mu\phi_1\partial_\mu\phi_2 \\
&\quad + e^2A_\mu A^\mu[\phi_0^2 + \frac{1}{2}(\phi_1^2 + \phi_2^2) + \sqrt{2}\phi_0\phi_1].
\end{aligned} \tag{2.14}$$

From Equation 2.13, it can be seen  $\phi_1$  has a mass term  $\frac{1}{2}m_1^2\phi_1^2$  with  $m_1 = \sqrt{2}\mu$  and from Equation 2.14,  $m_A = \sqrt{2}e\phi_0$ . The formerly massless field  $A$  now has a non-zero mass.

## 2.4 Glashow-Weinberg-Salam Theory of Weak Interactions

Now, to apply the method described earlier to Lagrangians invariant under  $SU(2) \times U(1)$  transformations, the Lagrangian is

$$\begin{aligned}
\mathcal{L} &= (D_\mu\phi)^\dagger(D_\mu\phi) - \mu^2\phi^\dagger\phi - \lambda(\phi^\dagger\phi)^2 - \frac{1}{4}E_{\mu\nu}E^{\mu\nu} - \frac{1}{4}f_{\mu\nu}f^{\mu\nu} \\
E_{\mu\nu} &= \partial_\mu A_\nu - \partial_\nu A_\mu - g(A_\mu \times A_\nu) \\
f_{\mu\nu} &= \partial_\mu B_\nu - \partial_\nu B_\mu.
\end{aligned} \tag{2.15}$$

The covariant derivative becomes

$$D_\mu\phi = \partial_\mu\phi + \frac{ig}{2}\mathbf{A}_\mu \cdot \vec{\tau}\phi + \frac{ig'}{2}B_\mu\phi. \tag{2.16}$$

Where  $\tau$  are the Pauli matrices

$$\tau_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \tau_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \tau_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (2.17)$$

The field is now a complex isodouble spinor, and the field at the minimum potential can now be written as

$$\phi_0 = \begin{pmatrix} 0 \\ \eta \end{pmatrix}; \quad \eta = \frac{\mu}{\sqrt{2\lambda}}. \quad (2.18)$$

Changing to the frame of this minimum and adding a small perturbation  $\sigma(x)/\sqrt{2}$  yields

$$\phi(x) = \begin{pmatrix} 0 \\ \eta + \sigma(x)/\sqrt{2} \end{pmatrix} \quad (2.19)$$

$$\begin{aligned} V &= \mu^2 \phi^\dagger \phi - \lambda (\phi^\dagger \phi)^2 \\ &= -\left( \frac{\mu^2}{2} \eta^2 - \mu^2 \sigma^2 - \frac{\mu^2 \sigma^3}{\sqrt{2}\eta} - \frac{\mu^2 \sigma^4}{8\eta^2} \right) \end{aligned} \quad (2.20)$$

$$D_\mu \phi = \begin{bmatrix} \partial_\mu + \frac{ig}{2} A_{\mu 3} + \frac{ig'}{2} B_\mu & \frac{ig}{2} A_{\mu 1} + \frac{g}{2} A_{\mu 2} \\ \frac{ig}{2} (A_{\mu 1} + iA_{\mu 2}) & \partial_\mu - \frac{ig}{2} A_{\mu 3} + \frac{ig'}{2} B_\mu \end{bmatrix} \begin{pmatrix} 0 \\ \eta + \sigma(x)/\sqrt{2} \end{pmatrix}. \quad (2.21)$$

Then

$$|D_\mu \phi|^2 = \frac{1}{2} \partial_\mu \sigma \partial^\mu \sigma + \left[ \frac{1}{4} g^2 (A_{\mu 1} A_1^\mu + A_{\mu 2} A_2^\mu) + \frac{1}{4} g^2 (A_{\mu 3} - \frac{g'}{g} B_\mu)^2 \right] \left( \eta + \frac{\sigma}{\sqrt{2}} \right)^2. \quad (2.22)$$

To relate the gauge fields to bosons, it is common to write the terms in brackets in terms of

$$\begin{aligned}
W_\mu^\pm &= \frac{1}{\sqrt{2}}(A_{\mu 1} \mp iA_{\mu 2}) \\
Z_\mu &= \cos\theta_w A_{\mu 3} - \sin\theta_w B_\mu = \frac{1}{\sqrt{g^2 + g'^2}}(gA_{\mu 3} - g'B_\mu) \\
A_\mu &= \sin\theta_w A_{\mu 3} + \cos\theta_w B_\mu = \frac{1}{\sqrt{g^2 + g'^2}}(g'A_{\mu 3} + gB_\mu).
\end{aligned} \tag{2.23}$$

The angle,  $\theta_w$ , is the Weinberg angle or weak mixing angle defined by  $\tan \theta_w \equiv g'/g$ .

It relates the gauge fields by

$$\begin{pmatrix} Z \\ A \end{pmatrix} = \begin{pmatrix} \cos\theta_w & -\sin\theta_w \\ \sin\theta_w & \cos\theta_w \end{pmatrix} \begin{pmatrix} A_3 \\ B \end{pmatrix}. \tag{2.24}$$

The Lagrangian, without constant terms, is then

$$\begin{aligned}
\mathcal{L} &= \frac{1}{2}\partial_\mu\sigma\partial_\mu\sigma + \mu^2\sigma^2 - \frac{1}{4}F_{\mu\nu}^2 - \frac{1}{4}(W_{\mu\nu}^{+\dagger}W_{\mu\nu}^+ + W_{\mu\nu}^{-\dagger}W_{\mu\nu}^-) \\
&+ \frac{g^2\eta^2}{4}(W_\mu^{+\dagger}W^{+\mu} + W_\mu^{-\dagger}W^{-\mu}) - \frac{1}{4}(Z_{\mu\nu}Z^{\mu\nu}) + \frac{g^2\eta^2}{4\cos^2\theta_w}Z_\mu Z^\mu \\
&+ \frac{g}{4}[(\mathbf{A}_\mu \times \mathbf{A}_\nu)(\partial^\mu \mathbf{A}^\nu \times \partial^\nu \mathbf{A}^\mu) + (\partial_\mu \mathbf{A}_\nu \times \partial_\nu \mathbf{A}_\mu)(\mathbf{A}^\mu \times \mathbf{A}^\nu)] \\
&- \frac{g^2}{4}(\mathbf{A}_\mu \times \mathbf{A}_\nu) \cdot (\mathbf{A}^\mu \times \mathbf{A}^\nu) \\
&+ \left(\frac{g^2}{2\sqrt{2}}\eta\sigma + \frac{g^2\sigma^2}{8}\right)(W_\mu^{+\dagger}W^{+\mu} + W_\mu^{-\dagger}W^{-\mu}) + \frac{1}{\cos^2\theta_w}Z_\mu Z^\mu + \frac{\mu^2}{\eta\sqrt{2}}\sigma^3 + \frac{\mu^2}{8\eta^2}\sigma^4
\end{aligned} \tag{2.25}$$

where the proposed origin of the  $W^\pm$  and  $Z$  boson masses can be seen in the sixth, seventh, and ninth terms to be

$$\begin{aligned}
m_W &= \frac{g\eta}{\sqrt{2}} \\
m_Z &= \frac{g\eta}{\sqrt{2}\cos\theta_w}.
\end{aligned} \tag{2.26}$$

The  $\mathbf{A}_\mu$  field is interpreted as the electrodynamic field. Recognizing  $(I + \tau_3)/2$  as

the charge operator, the electromagnetic charge is determined to be

$$|e| = g \sin \theta_w. \tag{2.27}$$

The  $\sigma$  field is the proposed Higgs field, and, if correct, this should lead to a physically realizeable particle. Its kinetic energy is given by the first term, and self-interaction terms are the last terms. Its mass is related by the second term and it is worth noting that this formalism does not predict the Higgs mass. We must search for it. This dissertation is such a search.

## 2.5 Limits on the Higgs Boson Mass

### 2.5.1 Theory

The Glashow-Weinberg-Salam theory of EWSB does not predict a value of the Higgs boson mass; it only gives an outline of how the boson interacts. This has led the search across orders of magnitude in energy through the decades as technology has progressed. We will consider two estimates of the upper bound of the Higgs mass from theory. The first involves unitarity arguments from  $WW$  scattering [18]. It sets an upper limit on  $m_H$  of about 1 TeV ( $m_H \leq \sqrt{8\pi\sqrt{2}/3G_F}$ , where  $G_F$  is the Fermi constant) where perturbation theory is still valid. Beyond this, tree level diagrams violate unitarity and phenomena beyond our current understanding would have to come into play.

The second upper bound is set in theory using a triviality argument. This means that, between two energy scales,  $\mu$  and  $\Lambda$ , the theory can be reduced to a trivial, non-interacting theory by taking the limit of  $\Lambda$  to infinity and seeing that  $\lambda(\mu)$  goes to zero [19]. This argument [20] uses one-loop corrections and the resulting relation

between the running coupling constants,  $\lambda$ , is

$$\frac{1}{\lambda(\mu)} = \frac{1}{\lambda(\Lambda)} + \frac{3}{2\pi^2} \log\left(\frac{\Lambda}{\mu}\right) \quad (2.28)$$

where  $\mu$  and  $\Lambda$  are energy scales ( $\mu < \Lambda$ ). For a stable potential,  $\lambda$  must always be positive. To make an inequality, the value  $1/\lambda(\Lambda)$  can be dropped, and considering the Higgs scale,  $\mu = m_H$ , using the relation  $m_H = \sqrt{2\lambda(m_H)\eta^2}$  results in

$$\Lambda \geq m_H e^{\frac{4\pi^2\eta^2}{3m_H^2}} \quad (2.29)$$

The vacuum expectation value is well-known from muon decay experiments [21] [22] and  $\eta \approx 246$  GeV. This gives a relation between the energy scale and the upper limit of Higgs mass beyond which the SM does not describe. Figure 2.2 shows that this upper limit is at most 800 GeV.

A theoretical lower bound on the Higgs mass can also be set in terms of  $\Lambda$  [23, 24, 25], beyond which the Higgs potential becomes unstable due to another minimum [5]. This, along with the upper bound is plotted in Figure 2.2 in terms of  $\Lambda$ .

## 2.5.2 Experiment

More restrictive limits on the Higgs boson mass come from precision electroweak data, and from exclusions resulting from previous searches. The electroweak data help predict the Higgs mass by exploiting relations between the Higgs boson and massive bosons and fermions it interacts with. The masses of the top quark and  $W$  boson are especially sensitive to interactions with the Higgs boson due to radiative corrections. Radiative corrections of this type are due to a Higgs boson decaying to two top quarks or  $W$  bosons in a loop that then recombine to form a Higgs boson. Figure 2.3a shows the relationship between the top quark and  $W$  masses for a number

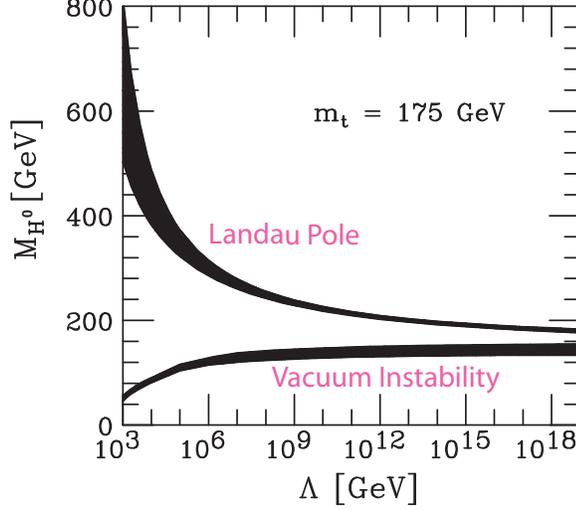
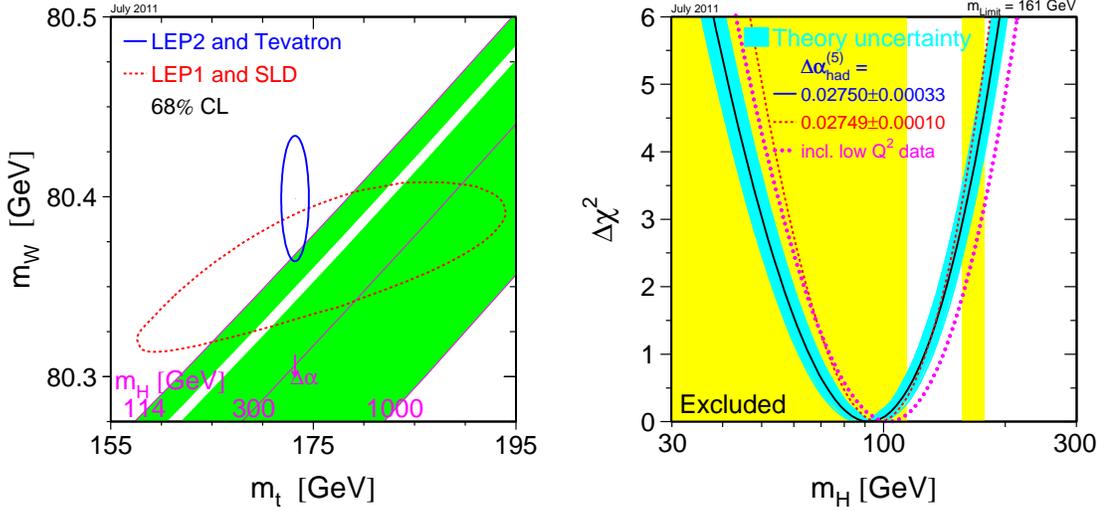


Figure 2.2: Theory limits on the Higgs boson mass as a function of energy scale. The top region (above the shaded error band) is excluded by triviality arguments and the bottom region (below the shaded error band) is ruled out by stability of the Higgs potential. Image from [26].

of Higgs boson masses. Figure 2.3b shows the  $\chi^2$  fit across electroweak measurements versus Higgs boson mass. The data were collected around the world at LEP, SLD, CDF, D0, and the pink dashed line includes NuTeV data [27]. The results predict a Higgs boson mass of  $92_{-26}^{+34}$  GeV/ $c^2$  at one standard deviation and lower than 161 GeV/ $c^2$  at the 95% confidence level.

Direct searches for the Higgs boson have further shaped the potential Higgs boson mass landscape. In 2003, the combined final result from the LEP experiments ruled out the SM Higgs boson below 114.4 GeV/ $c^2$  at the 95% confidence level [29]. The results from the 2010 Tevatron combination excluded mass ranges between 100 and 109 GeV/ $c^2$  and additionally 158 to 175 GeV/ $c^2$  at the 95% confidence level [30]. The analysis presented in this thesis improves upon the techniques of, and adds roughly 32% more data to the 2010 combination analysis [31] included in this previous result.



(a) Top quark and  $W$  mass versus  $m_H$ .

(b) Electroweak data  $\chi^2$  fit versus  $m_H$ .

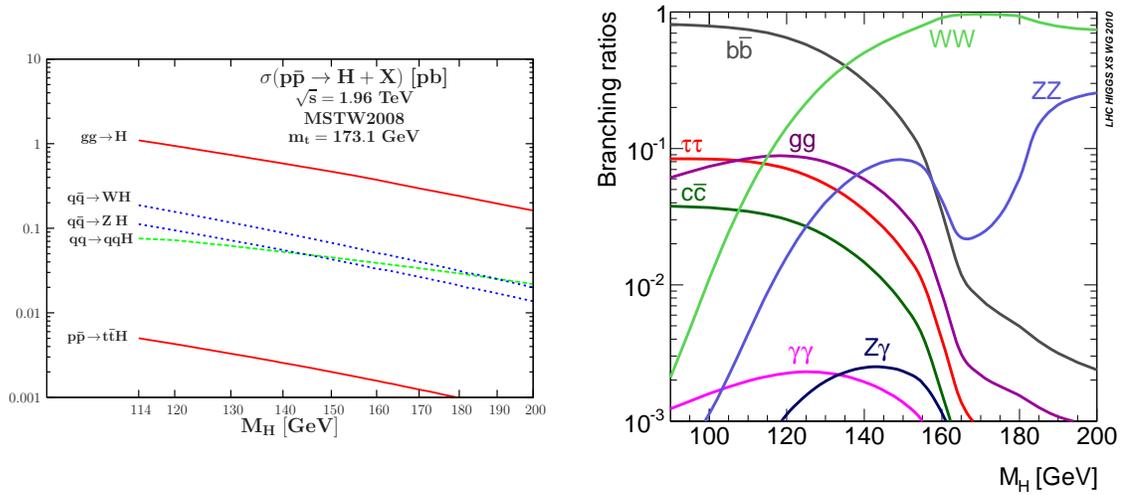
Figure 2.3: Relationships between electroweak data results and proposed Higgs boson masses. In Figure 2.3a, the solid blue ellipse results from direct measurements of  $m_t$  and  $m_W$ . The dashed red shape represents the top and  $W$  masses derived from indirect electroweak measurements. The sizes of the shapes represent one standard deviation. Both figures are from [28].

## 2.6 Higgs Phenomenology

### 2.6.1 Higgs Production and Decay

Due to the energy reach of the Tevatron and the predicted masses of the Higgs boson, current Tevatron searches largely focus in the mass range between 100-200  $\text{GeV}/c^2$ . The production mechanisms dominant in this range are gluon-gluon fusion ( $gg \rightarrow H$ ), associated (with a vector boson) production ( $q\bar{q} \rightarrow VH$ ,  $V = Z, W$ ), and vector boson fusion ( $q\bar{q} \rightarrow q'\bar{q}'H$ ). Their production cross sections versus  $m_H$  are illustrated in Figure 2.4a. The branching ratios are found using HDECAY version 3.53 [32] and are plotted versus mass in Figure 2.4b.

As one can see from Figures 2.4a and 2.4b, at masses above  $\sim 140$   $\text{GeV}/c^2$ ,



(a) Higgs Production Cross Sections

(b) Higgs Decay Branching Ratios

Figure 2.4: Production cross sections are given in pb versus Higgs boson mass,  $m_H$  (graphic from [33]). The Higgs boson decay branching ratios are listed in fractions and plotted versus Higgs boson mass (plot from [34].)

$H \rightarrow WW$  dominates. Analyses searching for this decay use the larger gluon-gluon fusion production mechanism and exploit the angular differences in decay products to distinguish it from  $p\bar{p} \rightarrow WW$ . For masses below  $\sim 135 \text{ GeV}/c^2$ ,  $H \rightarrow b\bar{b}$  is the primary decay channel. Searching for the process  $q\bar{q} \rightarrow H \rightarrow b\bar{b}$  is fruitless due to an overwhelming QCD background. Other methods are pursued, such as looking for  $H \rightarrow \tau\tau$ ,  $t\bar{t}H$ , or a bump (via quark loop) in the diphoton mass spectrum from  $H \rightarrow \gamma\gamma$ . The analysis in this dissertation searches for the associated production mechanism, namely  $q\bar{q} \rightarrow ZH$  as illustrated in the Feynman diagram in Figure 2.5.

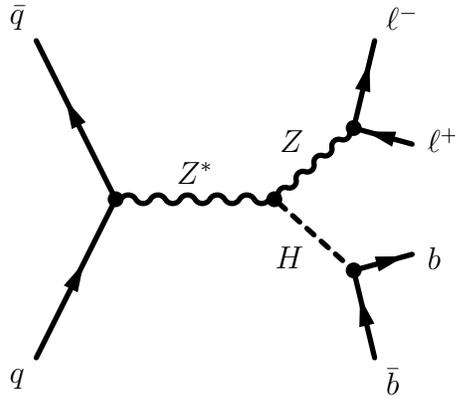


Figure 2.5: Tree level diagram for associated  $ZH$  production.

### 2.6.2 Backgrounds to Search

The signature we pursue has the Higgs boson decaying to two  $b$  jets ( $b\bar{b}$ ) and the  $Z$  decaying to electrons ( $e^+e^-$ ). This selection is heavily dominated by other processes such as  $t\bar{t}$ , diboson (not photon),  $Z \rightarrow e^+e^- + b\bar{b}$ , and misidentified objects ( $e$ 's or  $b$ 's). Many of these processes and their relative rates (not including the branching ratio, or  $BR$ , to electrons except where explicitly stated) are shown in Figure 2.6 and Table 2.3. These backgrounds will be further investigated in Section 4.3. The extreme imbalance in production rate combined with pernicious background signatures ( $Z \rightarrow e^+e^- + b\bar{b}$ ,  $ZZ \rightarrow e^+e^-b\bar{b}$ ) have led searches at the Tevatron to apply sophisticated analysis techniques in an effort to make competitive searches. These, along with the event selection are detailed in Chapter 4.

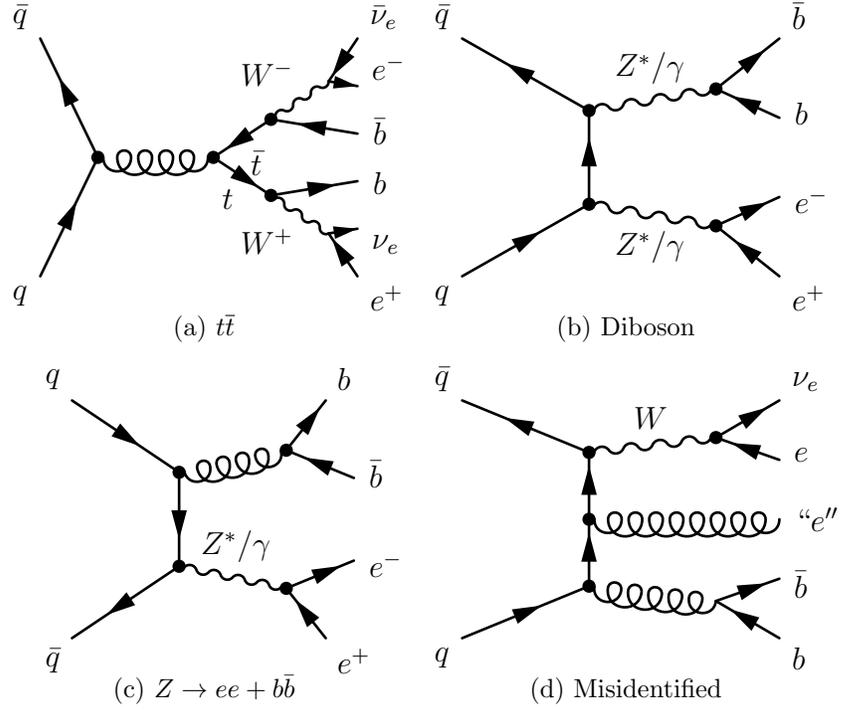


Figure 2.6: Feynman diagram of a subset of background processes.

Background	$\sigma(pb)$
$t\bar{t}$	7.0
$ZZ, WZ, WW$	3.6, 3.5, 11
$Z \rightarrow e^+e^- + b\bar{b}$	0.96
$Z \rightarrow e^+e^- + c\bar{c}$	2.1
$ZH_{120} \rightarrow llb\bar{b}$	$6.1 \times 10^{-3}$

Table 2.3: Production cross sections of processes involved in the search. The  $ZH$  cross section quoted is for a  $m_H$  of 120 GeV/ $c^2$  and includes the branching ratio of the  $Z$  boson decaying to charged leptons and  $H$  decaying to  $b\bar{b}$ .

# Chapter 3

## The Tevatron and the Collider Detector at Fermilab

### 3.1 The Tevatron

In order to study massive, unstable elementary particles, one studies high-energy collisions. While nature creates these in the form of cosmic rays, man-made collisions at a particle accelerator provide a consistent, high-statistics sample. This analysis was performed using collision data collected at CDF produced by the Tevatron.

The Tevatron is the accelerator at Fermi National Accelerator Laboratory (FNAL) in Batavia, Illinois. Prior to the Large Hadron Collider (LHC) turning on outside of Geneva, Switzerland, it was the world's highest energy collider. Now (Fall 2011), its claim to fame is that it is the world's highest-energy proton-anti-proton collider, and has provided roughly three times the data of the LHC. The Tevatron is aptly named after the energy regime of particles it accelerates. For collision data, each beam is accelerated to 980 GeV ( $\sqrt{s} = 1.96$  TeV), and for accelerator studies, the

proton beam's energy has been as high as 1.012 TeV<sup>1</sup>[35].

### FERMILAB'S ACCELERATOR CHAIN

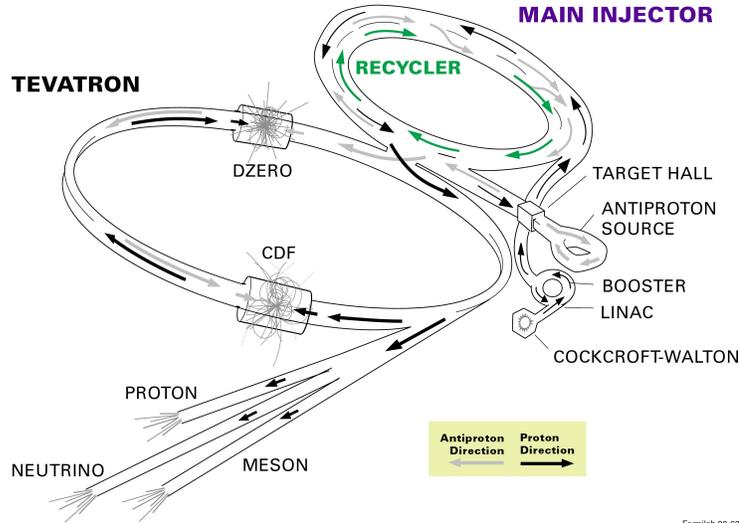


Figure 3.1: The Fermilab Tevatron's accelerator chain (Image from Fermilab Visual Media Services).

Since one machine cannot provide stable beam acceleration from the keV to TeV range, beam preparation for collision data requires a number of stages and specialized equipment (see Figure 3.1). The process begins with a cylinder of hydrogen (H) gas. The hydrogen gas is pulsed between two concentric molybdenum electrodes (cathode inside anode), a magnetron, to produce  $H^-$ . These  $H^-$  ions are extracted then accelerated to 750 keV by a Cockcroft-Walton accelerator and sent to the Linac. Fermilab's Linac is a LINear ACcelerator composed of two types of machines. The first is a radio frequency (RF) modulated series of drift tubes that bunches the beam of particles while accelerating them to 116.5 MeV. The second is a series of side-coupled cavities. "Side-coupled" refers to each of the RF cavities timing being tied to adjacent cells. Here there are 448 RF cavities each providing their own boost in

<sup>1</sup>All information presented here on the accelerator chain was gathered from Fermilab's Rookie Books [35].

energy. The  $H^-$  ions leave the Linac with 400 MeV of energy and are transported to the booster.

The booster is the first synchrotron the particles encounter in the accelerator chain. As the  $H^-$  ions from the Linac enter the booster, they are passed through a carbon foil that acts as an electron stripper making the  $H^-$  ions protons. These protons are then accelerated many times by RF chambers along the nearly 470 meter long circumference to an energy of 8 GeV and transported to the main injector.

The main injector is the second synchrotron in the accelerator chain. It is about seven times larger than the booster and roughly elliptical in shape. Originally, many components of the main injector sat in the same ring where the Tevatron is now. This accelerator was called the main ring and was used for earlier experiments. In order to provide the next increase in energy for the Tevatron's particles, the main ring was disassembled, and many components were used in the main injector located to the southwest of the Tevatron.

For Tevatron collision data, the main injector takes 8 GeV particles (protons from the booster, and anti-protons from both the anti-proton production source and the anti-proton recycler) and accelerates them to 150 GeV for entry into the Tevatron. For anti-proton ( $\bar{p}$ ) production, it accelerates 8 GeV protons to 120 GeV and sends them to the anti-proton source.

Anti-proton availability is one of the limiting reagents in producing collision data at the Tevatron. Fermilab has gone to great lengths to increase the availability of anti-protons, both with the addition of the  $\bar{p}$  recycler storing the  $\bar{p}$ 's as they are produced, and with the continued modification of their production. The anti-protons are produced by first striking the 120 GeV protons from the main injector against an Inconel (nickel-iron alloy) target. From the resulting spray of particles, a lithium lens optimized on 8 GeV  $\bar{p}$ 's focuses the particles and sends them to the debuncher.

This is an RF machine designed to focus and reduce the momentum spread of the particles. From here, they are sent to the accumulator where they are stored before being transferred to the recycler or main injector. This process produces about two  $\bar{p}$ 's for every  $10^5$  protons hitting the target. While this may seem inefficient, it is a factor of 30 improvement over the original Tevatron Run I rate and has led to the world's largest data collection in the TeV energy regime.

To prepare for collision data, the main injector first transfers 36 bunches of 150 GeV protons to the Tevatron. The process is then repeated in the main injector using anti-protons from the accumulator. At this point, the particles are in their final synchrotron of their acceleration, the Tevatron.

The Tevatron is a super-conducting synchrotron about 3.9 miles in circumference (1 km radius). In order to accelerate the proton and anti-proton bunches to 980 GeV, it uses eight RF cavities. To control the beams at this energy, very strong magnetic fields are needed. These are provided by the fields produced by sending 4600 A through niobium-titanium (NbTi), a superconductor, at 4.6 K. The Tevatron uses 774 super-conducting dipole magnets to bend the beams in circular paths and 216 quadrupole magnets to focus the beam. Once the Tevatron accelerates the beams to 980 GeV, and further focuses the beam and removes the halo, collisions in CDF can begin.

## **3.2 The Collider Detector at Fermilab: CDF**

The Collider Detector at Fermilab was built to measure the decay products from Tevatron collisions. CDF is comprised of sub-detectors that extract different information about the decay products passing through them. Very simply put, there are tracking detectors closest to the beam to record precise spatial information and

momentum of charged particles. Then there are a series of calorimeters used to measure energy, and the outermost detectors are the muon chambers for muon identification. When the information is combined from the sub-detectors, a reconstruction of the collision products, or event, can be made. A cross sectional elevation view of the CDF detector can be seen in Figure 3.2. The following sections will describe CDF in greater detail. For a complete reference, please see [36].

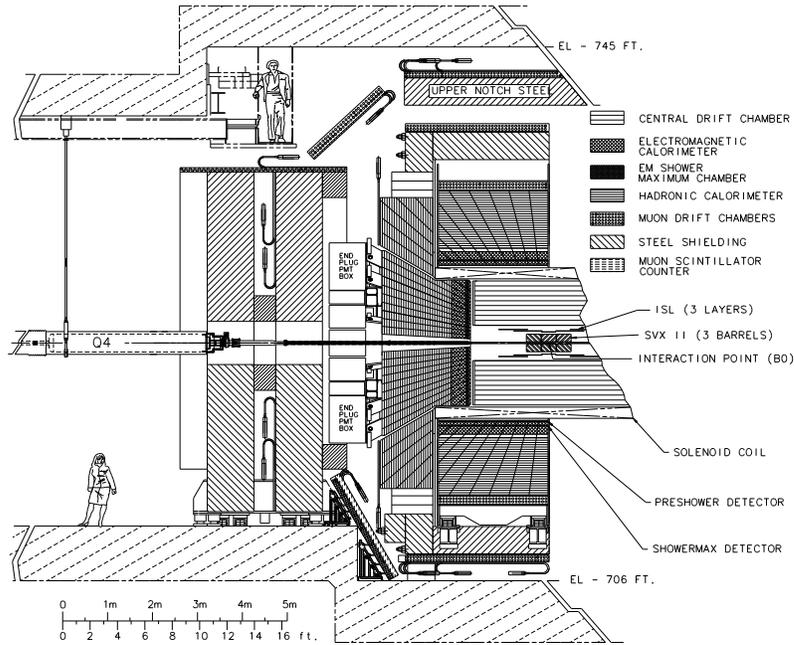


Figure 3.2: Half of a cross-sectional elevation view of CDF. Image from [36]

### 3.2.1 Tracking

#### 3.2.1.1 Solenoid

At CDF, a magnetic field is employed to force charged particles on a helical path through the tracking system allowing momentum and charge to be measured. The magnetic field is produced by a super-conducting solenoid. The solenoid is made by 1164 turns of aluminum-stabilized NbTi/Cu super-conductor around a radius of

1.483 m and an axial length of 4.796 m [37]. This, along with 4650 Amps, has provided a stable 1.41 Tesla field inside the tracking volume (2.8 m diameter and 3.5 m long) since Run I [36].

### 3.2.1.2 Silicon

CDF's silicon tracking system has three units: L00, SVX II, and the ISL (see Figures 3.3a and 3.3b). L00 [38], or Layer zero zero, is closest to the interaction point with half of the strips being placed just outside the beam pipe at a radius of 1.35 cm and the other half providing gap coverage at a radius of 1.62 cm. These are single-sided detectors providing only  $\phi$  information<sup>2</sup> and have  $z$  coverage roughly to  $\pm 47$  cm.

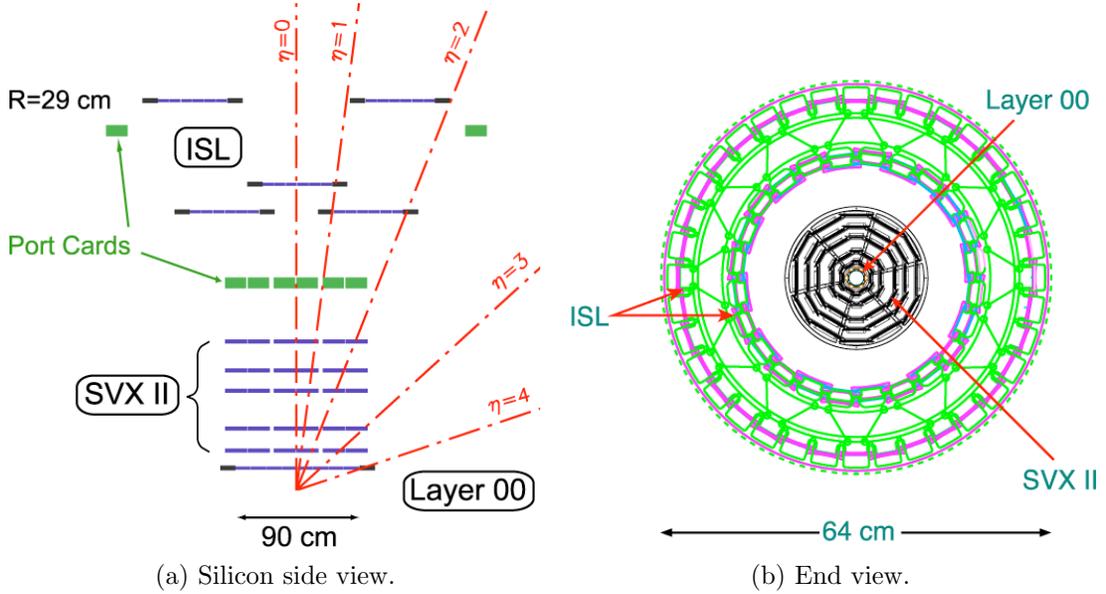


Figure 3.3: Images from [39]

The next silicon system is the Silicon Vertex Detector (SVX II) [36]. SVX II is an upgrade to the silicon system used in CDF Run I. SVX II has an inner radius of 2.44 cm, an outer radius of 10.6 cm, and is 87 cm long covering  $|\eta| < 2$ . It has three

<sup>2</sup>The coordinate system used at CDF has  $z$  along the beam line, the radial direction is orthogonal to  $z$ , and  $\eta$  describes the angle from the beam line related by  $\eta = -\ln[\tan(\theta/2)]$ . The  $\phi$  angle wraps around  $z$ .

barrels, each with 12  $\phi$  wedges. Each wedge has five double-sided layers of silicon providing stereo resolution for track fitting.

The outer-most silicon detector in CDF is the ISL, or Intermediate Silicon Layer [36]. The ISL consists of a single double-sided layer at a radius of 22 cm in the central region ( $|\eta| < 1$ ,  $|z| < 65$  cm), and two double-sided layers at 20 cm and 28 cm in the more forward region ( $1 < |\eta| < 1.9$ ). This sub-detector extends the silicon coverage in  $\eta$  and in the region inside of the outer tracker (COT). These extra layers allow for better 3-D track fitting and improve the momentum and impact parameter resolutions. The combined silicon system has an impact parameter resolution of about 40 microns at high  $p_T$ . This level of sensitivity is very important in low mass Higgs analyses where the ability to identify a displaced vertex from a  $B$  meson decay is central to reconstructing the Higgs signature.

### 3.2.1.3 Central Outer Tracker (COT)

The second component of CDF's tracking system is the COT [36]. The COT is a wire tracking chamber symmetric in  $\phi$  wrapping around the beam line and silicon detectors (see Figure 3.4a). Wire chambers work by having a surface at a potential within a gas that can be ionized by a traversing charged particle. As the particle passes through the gas, it ionizes gas molecules (see Figure 3.4b). The net free electrons travel to the potential surface and are read out by electronics. The COT's active volume extends 3.1 m in the  $z$ -direction and from 43.4 cm to 132.3 cm in the radial direction. The COT has eight layers in the radial direction, called super layers. The super layers are filled with cells tilted at 35 degrees with respect to the radial direction to compensate for the Lorentz angle of the electrons stripped by the ionizing particle. Each cell contains 12 sense wires and 17 potential wires that are 40  $\mu\text{m}$  diameter gold-plated tungsten. Odd-numbered super layers have wires at a

2 degree stereo angle for improved spacial measurement. The cells are separated by  $6.35 \mu\text{m}$  thick gold-on-Mylar field panels. The structure is filled with a 50:50 Ar/Et (plus a minute amount of isopropyl) gas mixture that leads to a maximum drift time of 177 ns. In 2004, a small amount of  $\text{O}_2$  ( $\sim 200$  ppm) began being continuously cycled through the COT to remedy effects of premature aging [40].

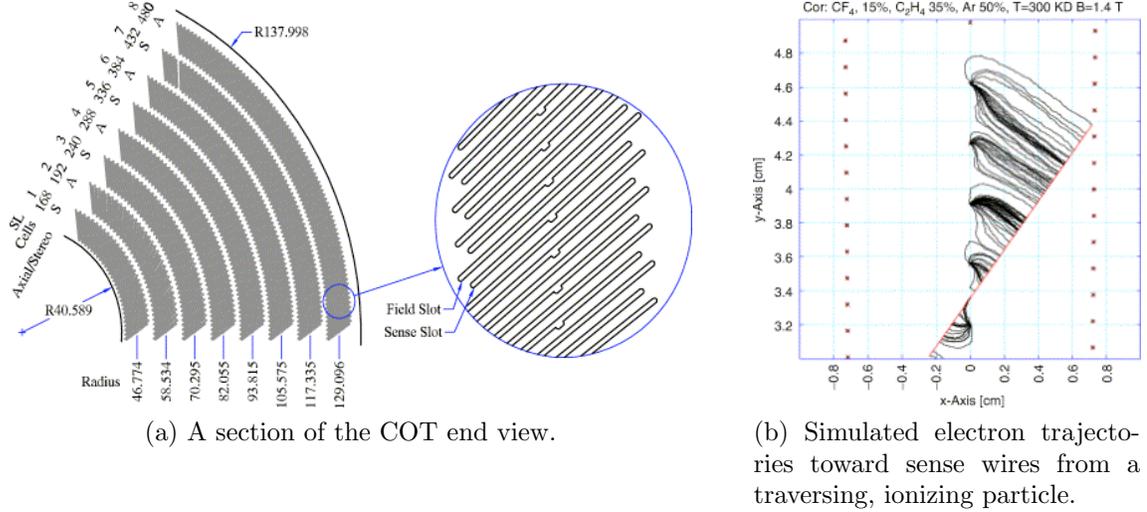


Figure 3.4: Image 3.4a from [41], and Image 3.4b from [40].

### 3.2.2 Calorimetry

Outside of CDF's tracking system and solenoid is the calorimetry. Since different types of particles interact differently in matter, two types of calorimetry are used to measure particles' energies: an electromagnetic (EM) calorimeter, and a hadronic (HAD) calorimeter. In an effort to provide good coverage this far away from the interaction point, these two calorimeters are further separated into a central region ( $|\eta| < 1.1$ ) and a plug or forward region.

### 3.2.2.1 Central Electromagnetic Calorimetry (CEM)

In the central region, just outside of the solenoid, is the central electromagnetic calorimeter (CEM) [42]. It is a lead and scintillator sampling calorimeter meaning that tiles of scintillator are placed between layers of lead such that as EM particles pass through, they interact with the lead producing particles that scintillate in the tiles. This light is then picked up by wavelength-shifting fibers connected to photomultiplier tubes and sent to readout electronics. This calorimeter (along with the central hadronic calorimeter) is segmented into projective towers in  $\eta$  and  $\phi$  of size  $0.11 (\eta) \times 15$  degrees in  $\phi$ . Each module has 10 towers in it and each segment in  $\phi$  has two modules leading to a pseudorapidity coverage of nearly  $\pm 1.1$ . In  $\phi$ , there are 24 modules and complete coverage.

Within the CEM at 5.9 radiation lengths ( $X_0$ , between the eighth lead and ninth scintillator layers), there is a two-dimensional readout strip chamber to measure shower position more precisely [42]. This length is known as the “shower maximum,” the distance into the material at which the showers are maximally developed. This chamber has one side of copper-backed 1/16 inch PC board (G10) with insulated connectors to gold-coated tungsten wires between extruded aluminum channels. The chamber is filled with a 95:5 gas mixture of Ar/CO<sub>2</sub>. The energy resolution,  $\sigma_E$ , of the CEM system is  $\sigma_E/E = 13.5\%/\sqrt{(E \cdot \sin\theta)}$  and the position resolution is  $\pm 2$  mm at 50 GeV.

### 3.2.2.2 Plug Electromagnetic Calorimetry (PEM)

In the more forward region, the plug electromagnetic calorimeter (PEM) is also a lead and scintillator sampling calorimeter (see Figure 3.5) [43]. It also has 24 15 degree wedges in  $\phi$  and an embedded 2D shower max detector. The plug calorimeter extends the pseudorapidity coverage to  $|\eta| \leq 3.64$ . A difference here from the central

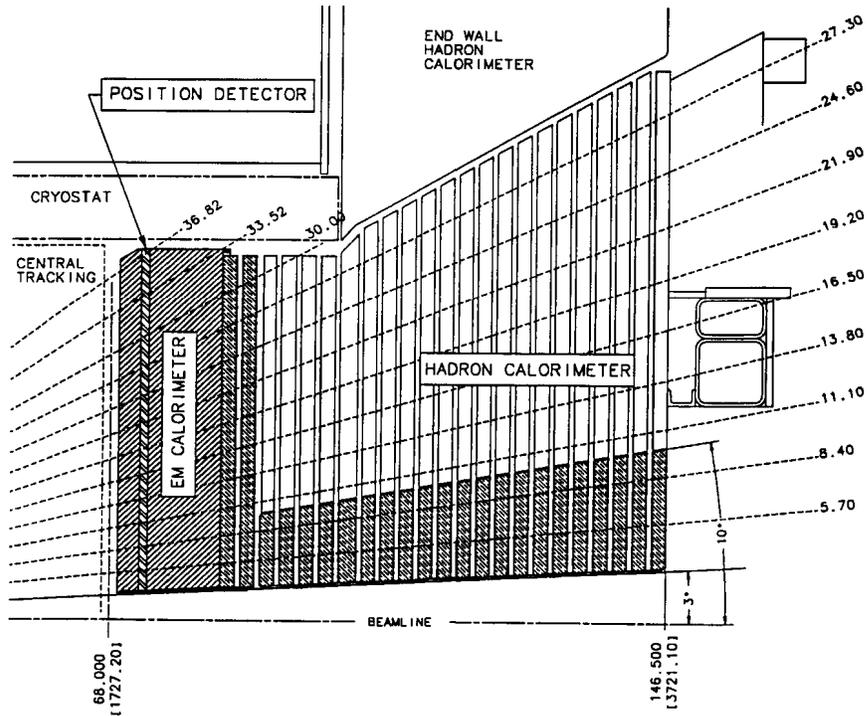


Figure 3.5: Cross-sectional view of a quarter of the plug calorimetry (image from [36])

calorimeter is the wavelength shifting fiber to collect the light is embedded in a groove on the tile rather than on the  $\phi$  boundary. Additionally, a preshower radiator (PPR) consisting of a 1.27 cm thick stainless steel sheet and thicker (2.5 times) single-layer of scintillating tiles is employed in front of the first lead plate.

### 3.2.2.3 Central Hadronic Calorimetry

Outside of the EM calorimeters lie the hadronic calorimeters. These are also sampling calorimeters using alternating layers of steel and scintillator. The central hadronic calorimeter (CHA) [44] is complete in the azimuthal angle and covers 45 to 135 degrees in  $\theta$ . It has 48 modules arranged in two cylinders each with 32 layers of alternating 2.5 cm steel and 1 cm scintillator. The endwall calorimeter covers  $\theta$

between 30 to 45 degrees and 135 to 150 degrees also with 48 modules. This section of calorimetry has 15 layers of alternating 5 cm thick steel and 1 cm thick scintillator. The central and endwall calorimeters are divided into towers like the EM calorimeter ( $\sim 0.1$  in  $\eta$  and 15 degrees in  $\phi$ ) and each  $\phi$  wedge has 24 towers.

### 3.2.2.4 Plug Hadronic Calorimeter

The plug region also has a hadron calorimeter just outside of the electromagnetic calorimeter [45]. This is also an iron and scintillator sampling calorimeter with 23 layers of alternating 2 inch iron and 6 mm scintillator. Stainless steel disks have been added to the 10 degree cone around the beam pipe to increase material thickness and bring the  $|\eta|$  coverage to 3 degrees (see Figure 3.5). Each 30 degree wedge in  $\phi$  in a scintillating layer has 32 tiles, and again, the wavelength shifting readout fibers are placed in a groove cut in the tile. This arrangement brings the energy resolution of this detector to  $81\% \sqrt{E} \oplus 7\%$  [46].

## 3.3 Muon System

The outermost detectors at CDF are the muon chambers. This is because, of commonly detected particles, the muons can travel the farthest in the detector. Indeed, most of the muon system is behind more than 6.2 pion interaction lengths. Being the farthest from the interaction point constrained some of the design of the system since it was limited by the physical space of the collision hall. The central muon detectors [36] (CMU) from Run I still exist in the detector. This subsystem is a cylindrical shell made of 2304 cells stacked four deep, and cover  $|\eta| < 1.0$  and most of the azimuth (see Figure 3.6). Each cell has a  $50 \mu\text{m}$  stainless steel wire and cathode strips on top and bottom for field shaping. Odd layers are staggered with respect

to even layers to improve resolution (see Figure 3.7b). However, since the maximum drift time for this detector is 800 ns while the bunch cross spacing for Run II is 396 ns, this system cannot be used alone. The central muon upgrade (CMP) [36] was built to handle the timing issue. The CMP covers nearly the same pseudorapidity regions as the CMU, but has some deficiencies (see Figure 3.6). This is because the CMP forms a box enclosing the central CDF barrel and the  $\eta$ -coverage toward the corners is reduced by geometry. This detector also has stacked, alternating staggered four layer wire cells (see Figure 3.7b). To resolve which collision a hit came from, it adds 216 scintillator tiles on the outer radius surface that are then read out by phototubes.

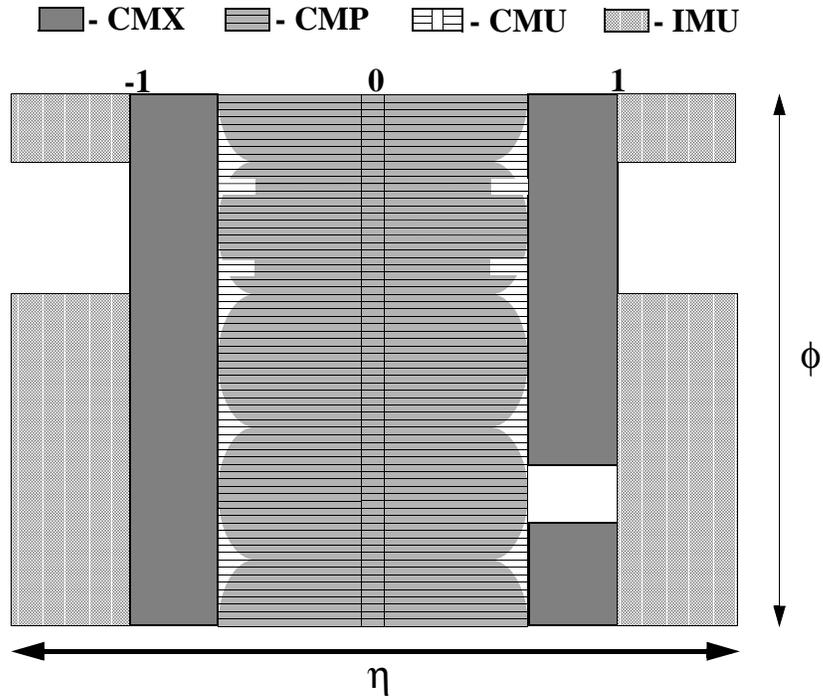
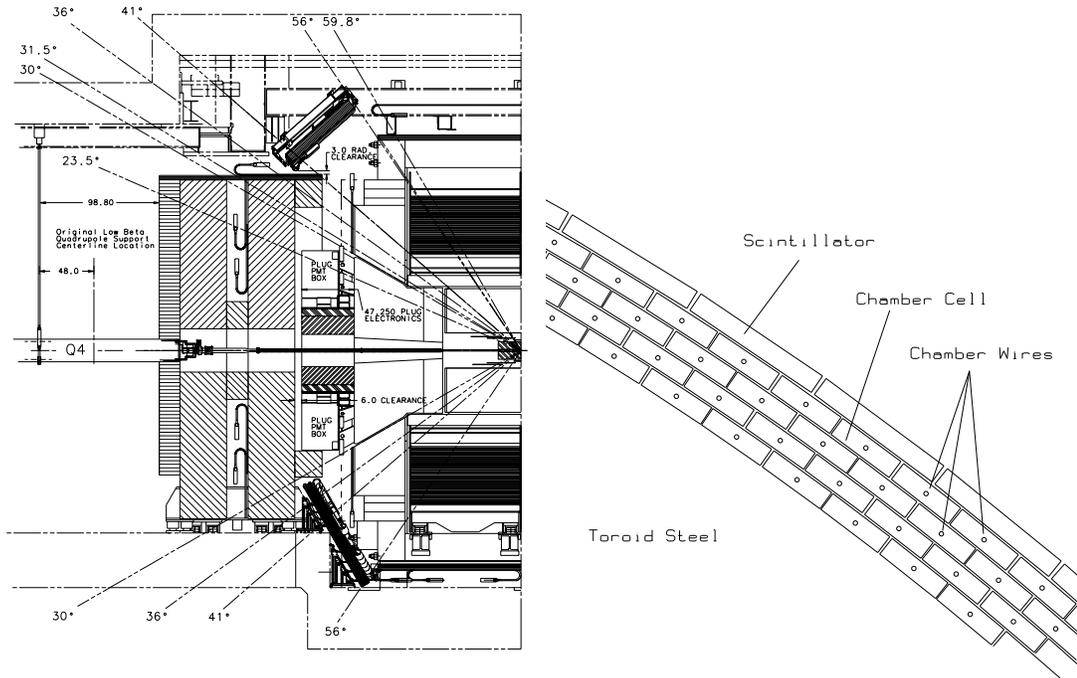


Figure 3.6: Muon system  $\eta - \phi$  coverage map. Image from [36].

To extend the  $\eta$ -coverage to  $\sim 1.0$ , the central muon extension (CMX) [36] was added for Run II. This system is also made from four layers of alternating 1/2 cell

offset drift tubes, however, instead of forming a box or barrel, it forms a partial cone (see Figure 3.7a). The scintillating tiles are trapezoids to accommodate this shape and are placed on the inside and outside of the system in four staggered layers to provide better resolution.



(a) Side view of muon system.

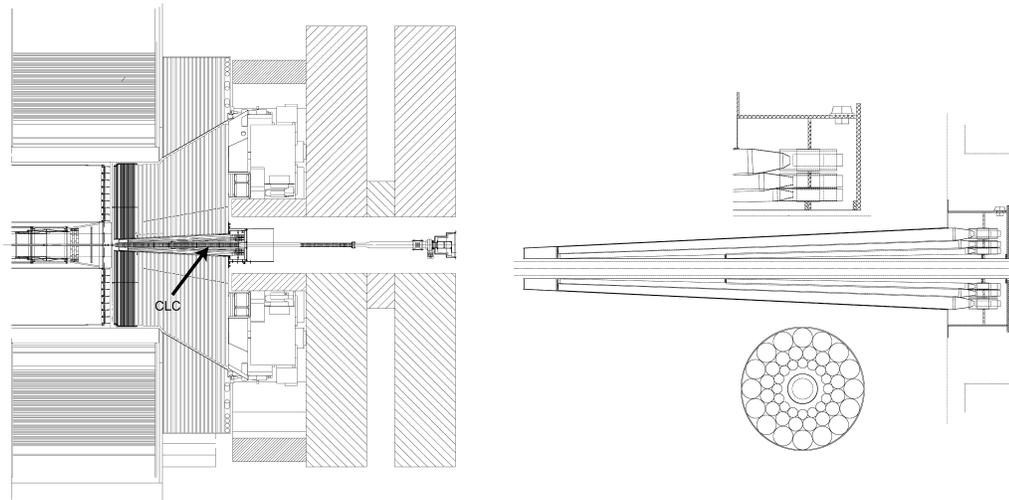
(b) Detail of muon cell and scintillator arrangement.

Figure 3.7: Figure 3.7a shows the placement of some of the muon components. Note the CMX system oriented on an angle with respect to the beamline. Figure 3.7b shows the arrangement of staggered cells and scintillator for the IMU, but the layout is similar for the other muon subsystems. Images from [36].

The last muon system is the intermediate muon detector (IMU) [36]. This system extends the pseudorapidity coverage to  $\sim 2$  using chambers similar to the CMP (see Figure 3.7b) arranged in a barrel geometry (the floor spacing prevents a perfect barrel shape). Scintillator for this system is placed on the outer radius, on the endwall, and on the inside of the unused toroid from Run 1 (see Figure 3.7a ).

### 3.4 Cherenkov Luminosity Counter

To properly set limits and measure cross sections of physical processes, the integrated luminosity, or total number of interactions, must be known. To measure this quantity, CDF has the Cherenkov Luminosity Counter (CLC) [47]. This system is made of a series of conical tubes within the three degree cone ( $3.7 \leq |\eta| \leq 4.7$ ) around the beam pipe in the plug calorimeter (see Figures 3.8a and 3.8b). The tubes are filled with isobutane at atmospheric pressure that acts as a radiator of Cherenkov light from particles produced in a  $p\bar{p}$  collision. A cone-shaped light collecting mirror is placed on the larger opening of each tube (see the top of Figure 3.8b). These help send the produced light to the attached photomultiplier tubes that measure the light output, and in effect, the luminosity.



(a) Location of the CLC in the forward region. (b) Arrangement of counters in the CLC.

Figure 3.8: Schematic of the Cherenkov Luminosity Counter (CLC). Images from [47].

## 3.5 CDF's Data Acquisition

Interesting events have a low probability of occurring, or cross section, compared to the complete inelastic collision cross section. To compensate for this, the Tevatron attempts to maximize the luminosity, or number of  $p\bar{p}$  collisions. With the bunch setup described in Section 3.1, the bunch crossing rate is 1.7 MHz and the  $p\bar{p}$  collision rate can be upwards of 10 MHz [48]. With each event taking about 170 kB to store, storing all events could amount to nearly 1.7 TB/s. Given mass storage constraints and a maximum data acquisition storage rate of 18 MB/s, CDF developed a trigger system with the goal of saving only interesting events.

The trigger system has three levels and is shown schematically in Figure 3.9. The Level-1 (L1) trigger takes in data from the COT, calorimeters, and muon system and, using custom hardware, makes a machine decision to pass the event to Level-2 or reject it. At L1, track data are available by use of the eXtremely Fast Tracker (XFT) and the Track Trigger Extrapolation System (XTRP) [48]. The XFT finds charged tracks from hits in the COT. It performs efficiently finding over 90% of tracks with  $p_T > 1.5$  GeV/ $c$  and provides accurate  $p_T$  information with resolution  $\delta p_T/p_T = 0.002 p_T$ . The XTRP system combines the track information with calorimeter and muon system data resulting in lepton identification information for the trigger system.

At Level-2, information from L1 is combined with data from the central shower maximum detector and the SVX. The SVX information allows vertex displacement to enter into the trigger decision which can help highlight events with  $b$ -quarks. Here, the information is combined into a limited event reconstruction on custom hardware, and a decision is made to either pass the event to Level-3 (L3) or reject it.

At L3, the event is fully reconstructed using a full detector readout to over 100

PCs. Here, the accept, or pass, rate is about 75 events per second. These events are sent to mass storage. The triggers used in this analysis and other event selection criteria are detailed in the following chapter.

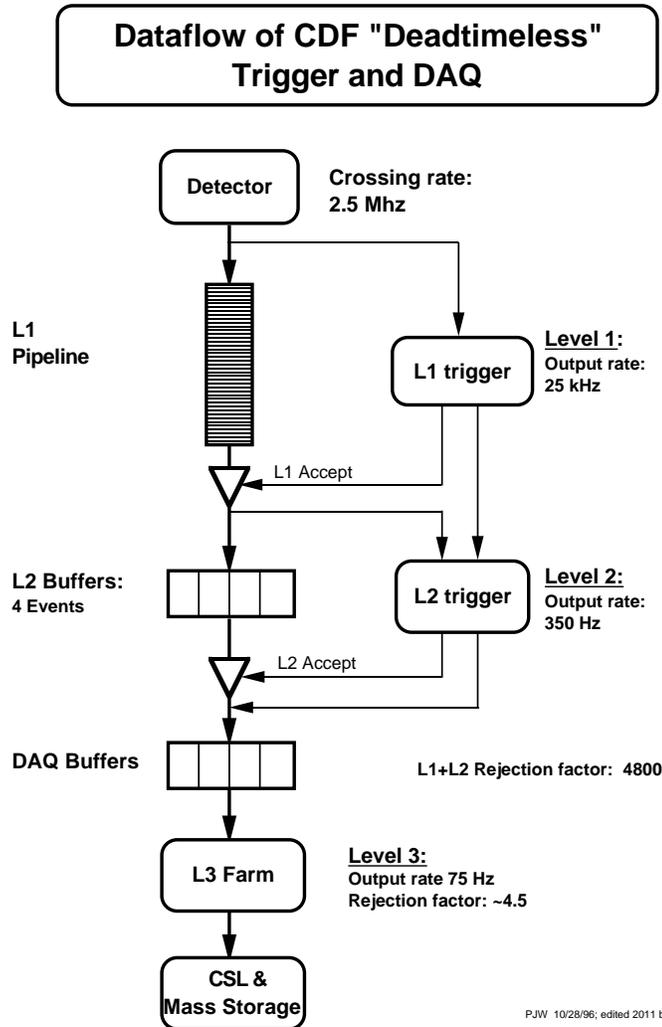


Figure 3.9: Image from [36], edited by the author. CSL refers to the Consumer-Server/Logger described in [46]. Peak rates are 15-25% higher [48]

# Chapter 4

## Event Selection and Reconstruction

### 4.1 Data

The data used in this analysis were collected at CDF between February 2002 and March 2011 and correspond to an integrated luminosity of  $7.5 \text{ fb}^{-1}$ .

#### 4.1.1 Triggers Used

Considering every event saved to mass storage was beyond the capability of resources available to this analysis. Instead, since the final state has electrons, we only consider data events that passed (or *fired*) an electron-motivated trigger. The three triggers we select are:

- **ELECTRON\_CENTRAL\_18**: Fires on a single electron candidate that has a track and energy deposited in the central calorimeter. The detector-measured (uncorrected) calorimeter energy must be at least  $\sim 89\%$  electromagnetic. This trigger has been in effect for all of Run II.

- **Z\_NOTRACK**: Fires on two largely electromagnetic calorimeter deposits of at least 18 GeV as measured at the detector level. This trigger was added during the first period of data taking.
- **Z\_NOTRACK\_MASS**: Also fires on two calorimeter-based objects. The detector must measure at least 18 (9) GeV for the first (second) object and the energy must be largely electromagnetic. The mass of the di-electron candidate object must be at least 40 GeV/ $c^2$ . This trigger has only been in effect for data recorded beyond  $\sim 2.45 \text{ fb}^{-1}$ .

The requirements for each of these at each trigger level are listed in Table 4.1. The modeling of trigger inefficiencies and behavior is described in Section 4.3.2.

Trigger Name	Level 1	Level 2	Level 3
<b>ELECTRON CENTRAL 18</b>	$E_T \geq 8 \text{ GeV}$ Had/Em $< 0.125$ Track $p_T \geq 8.34$	cluster $ \eta  < 1.317$ cluster $E_T \geq 18 \text{ GeV}$ cluster Had/Em $\leq 0.125$	$E_T \geq 18 \text{ GeV}$ Had/Em $\leq 0.125$ central calorimeter Track $p_T \geq 9 \text{ GeV}$ Lshr $< 0.4$ $\Delta Z < 8 \text{ cm}$
<b>Z NOTRACK</b>	$E_T \geq 18 \text{ GeV}$ Central Had/Em $\leq 0.125$ Plug Had/Em $\leq 0.0625$ two objects	cluster $ \eta  < 3.6$ cluster $E_T \geq 16 \text{ GeV}$ cluster Had/Em $\leq 0.125$ two clusters	two objects $E_T \geq 18 \text{ GeV}$
<b>Z NOTRACK MASS</b>	$E_T \geq 18 \text{ GeV}$ Central Had/Em $\leq 0.125$ Plug Had/Em $\leq 0.0625$ two objects	$E_{T1} \geq 16 \text{ GeV}$ $E_{T2} \geq 8 \text{ GeV}$ Had/Em $\leq 0.125$ Mass( $e_1, e_2$ ) $\geq 40 \text{ GeV}/c^2$	$E_{T1} \geq 18 \text{ GeV}$ $E_{T2} \geq 9 \text{ GeV}$ Had/Em $\leq 0.125$

Table 4.1: Requirements for the three electron triggers to pass each trigger level. An event passing level 3 is saved to mass storage and considered in this analysis. The “no track” label in a trigger name does not require a trackless object, but rather only takes into account calorimeter quantities in the trigger decision. The variable Lshr is the lateral shower profile of the electron candidate comparing the energy fractions in adjacent towers to that of test beam electrons.

## 4.1.2 Event Credibility Checks

### 4.1.2.1 Cosmic Muon Filter

CDF is not impenetrable to muons from cosmic rays. We apply a filter to remove data events that have been contaminated by products of cosmic rays. In this analysis, these products can mimic electron or jet signatures, or simply throw off the event energy calculation by bremsstrahlung in the calorimeter and producing a track. To remove these events, we apply a common CDF filter: the `isCosmic` flag [49]. Since the interactions of objects from cosmic rays are unrelated to the  $p\bar{p}$  collisions, the object's  $r$ - $\phi$  impact parameter and timing information are used to flag these. Additionally, back-to-back tracks or energy in towers inconsistent with the event help distinguish cosmic events and are used in the flagging algorithm.

### 4.1.2.2 Good Run Check

At times during data taking, the CDF detector may be performing sub-optimally. This may be because part of the detector is off and waiting for repairs, or a component could be behaving unreliably due to some hardware or software failure. In any case, record keeping of which components of the detector were good during data-taking is maintained in a database. This information is processed to make good run lists [50] specific to different types of analyses. For this analysis, we apply good run list `em_mu_si_v40`. In addition to a functional calorimeter system and central tracker, this good run list requires a good shower maximum detector, good SVX and ISL silicon systems, and a good muon system. As the name implies, the events in the list are good for electrons, muons, and silicon which is needed for  $b$  tagging. The version label corresponds to data collected through what in CDF is commonly referred to as period 34. In this analysis, we only consider events that are marked good by this

list.

## 4.2 Event Reconstruction

### 4.2.1 Electron Selection & $Z$ Reconstruction

#### 4.2.1.1 Electron Identification

Earlier searches at CDF for  $ZH \rightarrow \ell^+\ell^-b\bar{b}$  used a cut-based electron identification. Here, for the first time in this channel, a neural-network-based electron identification is implemented. Three networks were developed and applied according to the class, or type, of electron. There is a network for central ( $|\eta| < 1.1$ ) electrons, one for “phoenix” forward ( $|\eta| > 1.1$ ) electrons, and one for non-phoenix forward electrons. The classifications are exclusive and each candidate can have, at maximum, one neural-network score.

In an effort to train the networks to identify electrons that would be realistically saved in data, and to reduce the computing burden, trigger-inspired preselection cuts were applied to the events used in the training and application of the networks (see Table 4.2). Additionally, a requirement that the reconstructed  $Z$  vertex be well-contained within the detector ( $|z_0| < 60$  cm) was imposed.

Category	$\eta$	EM $E_T$ (GeV)	HAD/EM	Additional
Central	$ \eta  < 1.1$	$> 9$	$< 0.125$	
Forward Phoenix	$ \eta  > 1.1$	$> 9$	$< 0.0625$	
Forward Non-Phoenix	$1.2 <  \eta  < 2.8$	$> 9$	$< 0.125$	Momentum Defined

Table 4.2: The trigger-inspired preselection cuts for the different electron identification networks. The variable EM  $E_T$  is the transverse energy in the electromagnetic calorimeter. The variable HAD/EM is ratio of the energies in the hadronic calorimeter to the electromagnetic calorimeter.

Electron candidates with the phoenix designation are aptly named since they are

reconstructed by an offline algorithm, or more poetically, they are raised from the ashes of unconnected silicon hits and calorimeter deposits. Non-phoenix electrons use tracks in the COT (central tracker) to determine their  $p_T$  and path [51]. In the forward regions, this becomes more challenging due to lack of coverage. The phoenix algorithm avoids using the COT and calculates helical paths based on the shower maximum centroid in the calorimeter, hits in the SVX silicon system, and the event primary vertex. The helical paths consider both charge possibilities and, if both lead to a track matched to the calorimeter, the algorithm returns the track with the smaller  $\chi^2/\text{dof}$  (dof is an acronym for degrees of freedom and is equal to the number of silicon hits).

A variety of configurations were evaluated for the construction of the neural networks. Instead of a single-electron network, a  $Z$  object identifier was evaluated with comparable results, but ultimately dropped due to an inability to model misidentified electrons well. For the single-electron identification, networks only using shape and quality variables (track  $\chi^2$ ,  $E/P$ , Had/Em, *etc.*) were evaluated against those using shape, quality, and energy variables (track  $p_T$ , energy, *etc.*). Based on the change in the  $Z$  mass distribution from the cut-based selection for data,  $Z$  Monte Carlo, and  $ZH$  Monte Carlo, using a number of score cut values and by seeing how well the network isolated signal from background in the score distributions, the networks including energy-type variables were determined to have performed better. These variables were allowed in all three networks.

Alongside the variable study, different options for signal and background templates were also evaluated. For signal, generator level (hepg-matched) electrons in  $Z$ +light-flavored (l.f.) jets Monte Carlo and a data option were considered. The data option took one electron passing minimal cuts from  $Z$  events with two electrons ( $76 \leq M_{ee} \leq 106 \text{ GeV}/c^2$ ) where the second electron passed tight cuts. The method

is called tag-and-probe and the probe leg is taken as the unbiased electron.

For background templates, various data and Monte Carlo options were also considered. Templates made from electron candidates hepg-matched to non-electron objects in the  $Z+l.f.$  jets and  $W$ +jets Monte Carlo were tested. In data, templates made from electron candidates in  $Z$  objects outside of the  $Z$  window 76-106 GeV/ $c^2$  and templates made from electron candidates in jet-triggered data with exactly one candidate ( $Z$  veto) and  $\cancel{E}_T < 15$  GeV ( $W$  veto) were also considered. The templates that were most effective at identifying electrons cleanly, and therefore used, had signal templates from hepg-matched electrons in the  $Z+l.f.$  jets Monte Carlo and background templates populated with a non-electron enhanced data sample. For the central electron network, the background template used electron candidates from di-electron objects outside of the  $Z$  window. For both forward networks, background templates were populated with electron candidates from jet-triggered data with a  $W$  and a  $Z$  veto.

The variables included in each network are an optimal subset of a pool of variables selected by applying an iterative method to minimize the classification error. Given a pool of  $N$  variables, the method works by training and testing  $N$  single-variable networks. Then the variable corresponding to the network with the lowest testing error<sup>1</sup> is saved. Then,  $N - 1$  two-variables networks are trained and tested with the most-powerful variable always being included. The variables from the network with the lowest testing error are saved, and the next iteration ( $N - 2$  three-variable networks) are trained and tested. This continues until the next iteration can no longer produce a lower testing error. The testing error versus added variable for this procedure is plotted in Figure 4.1. The variables selected for each network are

---

<sup>1</sup>The testing error is equal to  $\sum_{i=1}^N (s_i - t_i)^2$  where the sum is over the testing set of events,  $s_i$  is the network score and  $t_i$  is the target value for a given event. For more information, please see Appendix A

listed in Table 4.3 and plotted for signal and background in Figures 4.2-4.4. The networks used in the variables-selection procedure were created using ROOTJETNET [52], while the final electron identification networks were created with TMVA [53].

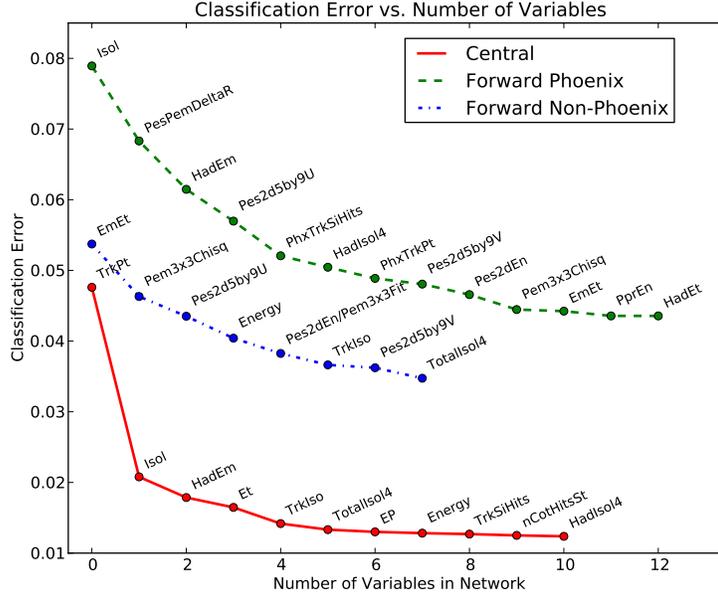


Figure 4.1: The testing error of each network as the best variable is added to the network architecture.

In addition to the neural-network selected electrons, a cut-based class of track-only electrons [54] is retained from previous searches [55]. The electrons have tracks that point to uninstrumented, or nonfiducial, parts of the calorimeter. This type of electron is informally called “crack track” and abbreviated “CrkTrk.” The crack track selection criteria are listed in Table 4.4.

#### 4.2.1.2 Reconstructing a $Z$ Boson

To reconstruct  $Z$  bosons, pairs of electrons are considered. For pairs of neural-network-selected electrons, one electron’s score must be above a high-score value and the other must be at least above a low-score value. The high- and low-score cuts

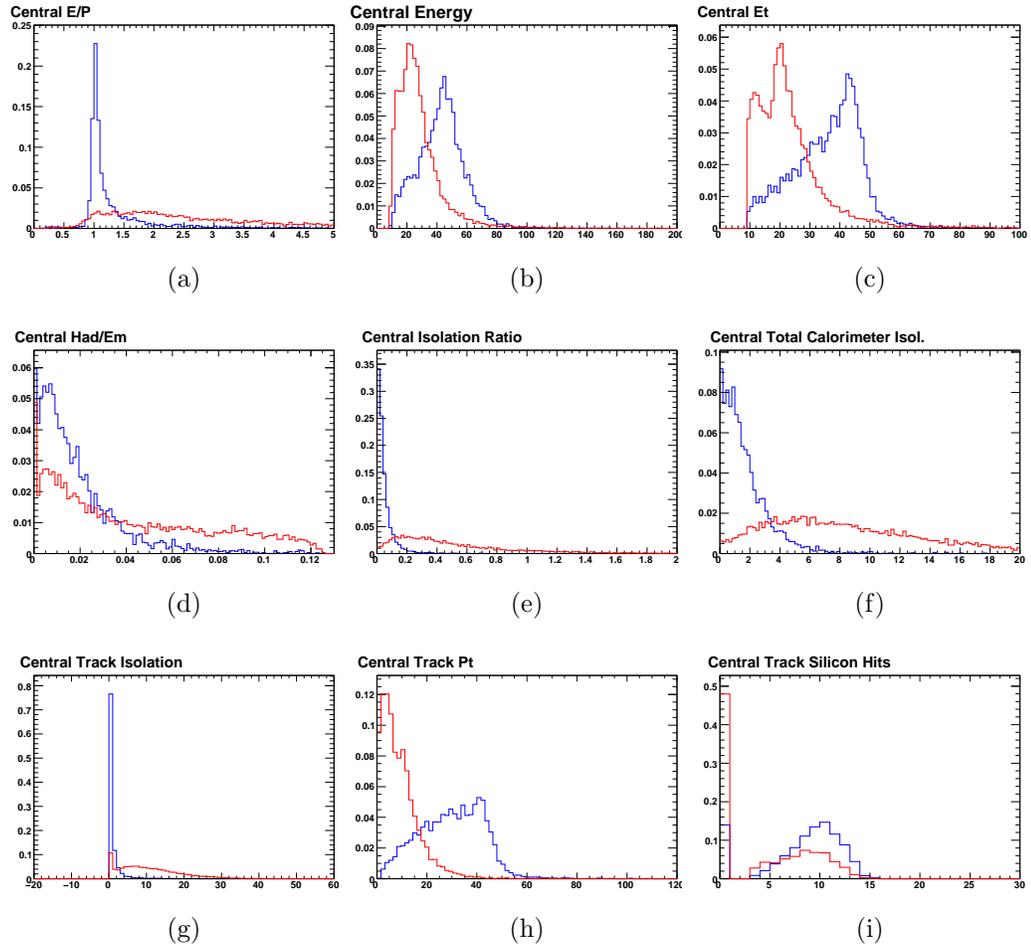


Figure 4.2: The variables used in the training of the central electron identification neural network. Signal is blue; background is red.

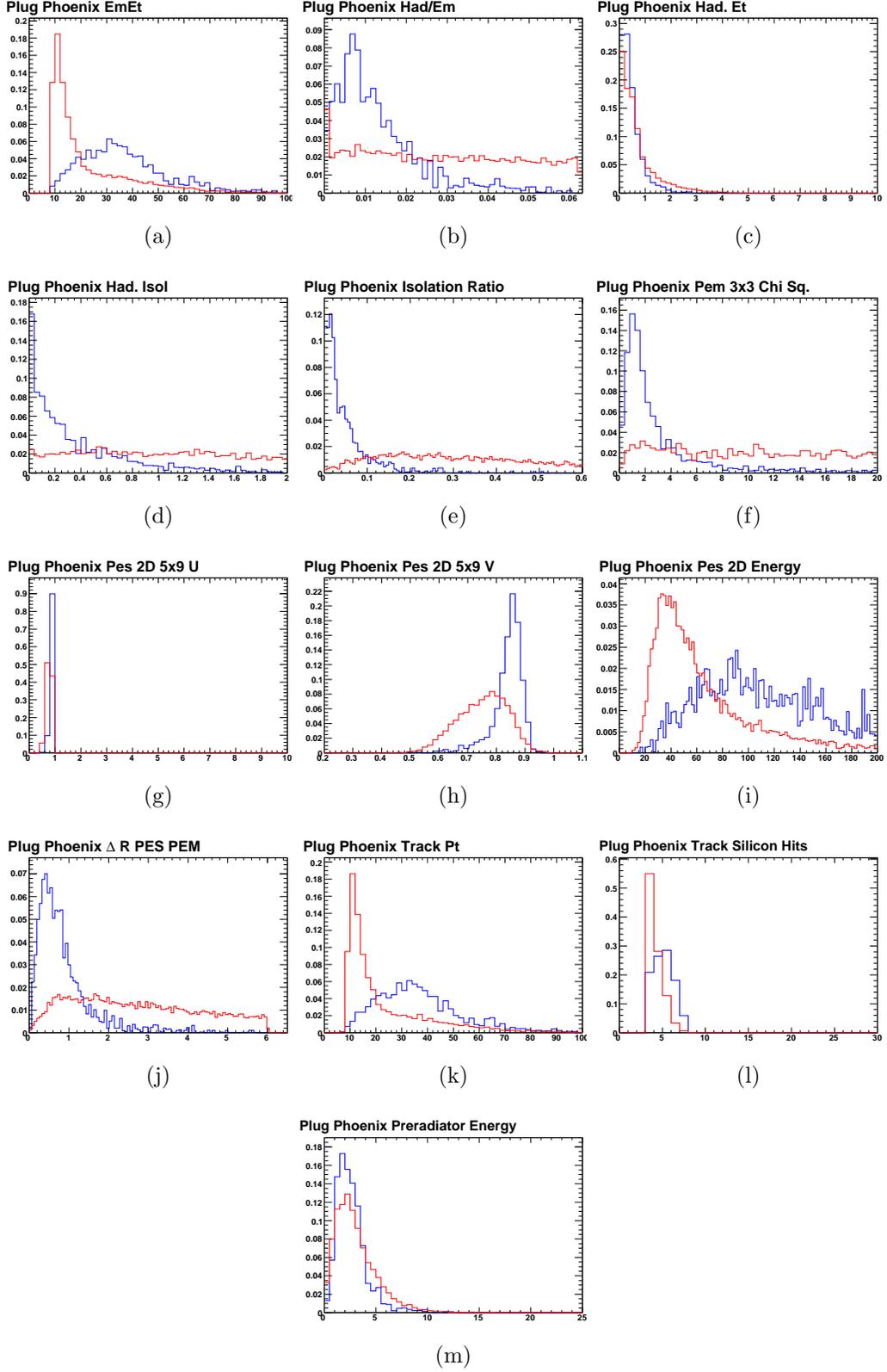


Figure 4.3: The variables used in the training of the plug phoenix electron identification neural network. Signal is blue; background is red.

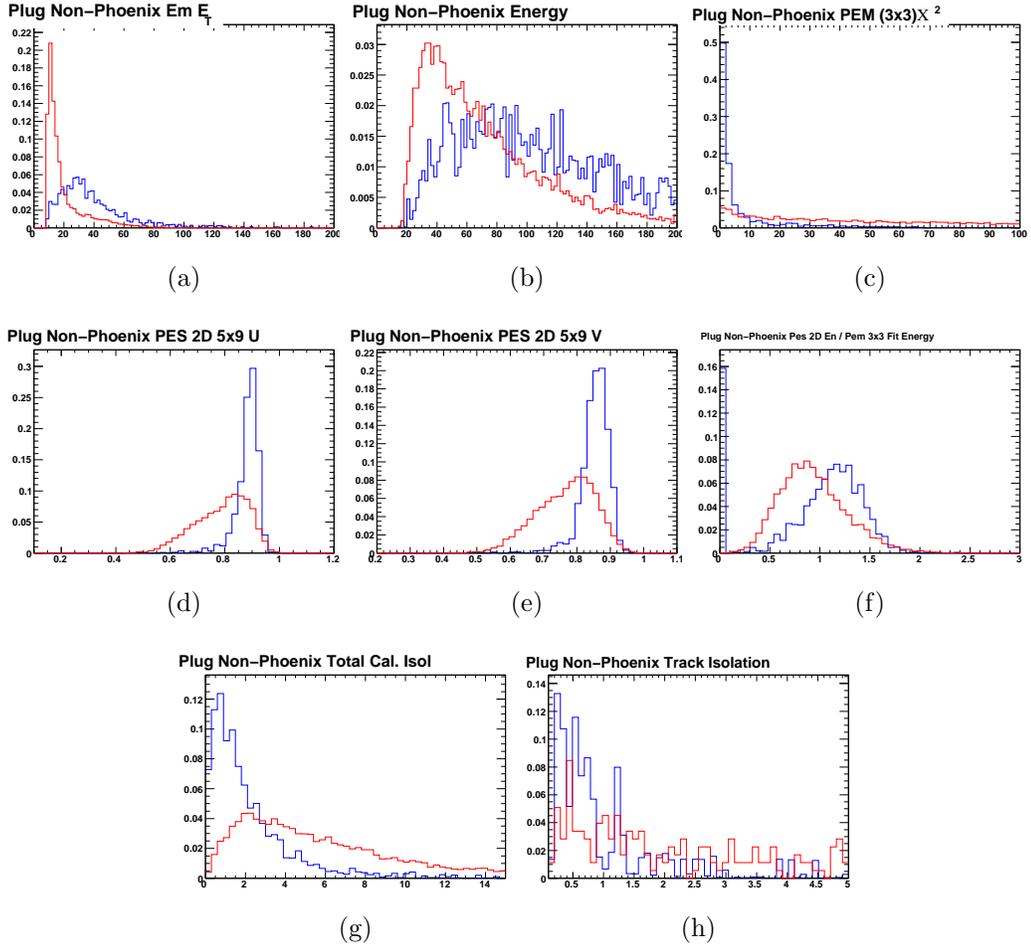


Figure 4.4: The variables used in the training of the plug non-phoenix electron identification neural network. The candidates in this category have an associated track of minimal quality (momentum is defined). The track isolation variable is truncated at 0.1 to see the discriminating power in the variable (values at zero would set the scale high). Signal is blue; background is red.

Central	Plug Phoenix	Plug Non-Phoenix
Track $p_T$	Isolation Ratio	EM $E_T$
Isolation Ratio	$\Delta R$ (shower profile centroid, cal. cluster centroid)	Pem 3×3 Fit $\chi^2$
$E_T$	HAD/EM	PES 2d 5×9 U
Track Isolation	PES 2d 5×9 U	Energy
Total Cal. Isolation ( $R=0.4$ )	Silicon Hits	Energy deposited in 2 × 2 cal. towers
E/P	Had. Isol. ( $R=0.4$ )	Track Isolation
Energy	Track $p_T$	PES 2d 5×9 V
Silicon Hits	PES 2d 5×9 V	Total Cal. Isolation ( $R=0.4$ )
	Energy deposited in 2 × 2 cal. towers	
	PES 3×3 Fit $\chi^2$	
	EM $E_T$	
	Plug Preradiator Energy	
	HAD $E_T$	

Table 4.3: The variables used in each electron identification neural network. Isolation ratio is the energy deposited in the electron cluster excluding the electron over the total energy in the cluster. The Pem 3×3 fit  $\chi^2$  determined by comparing the shower shape of the cluster in 3×3 towers to that of test beam data. PES 2d 5×9 U (V) is the ratio of activity matching the best 2×2 cluster in five strips in the plug shower profile detector over that of nine strips in the U (V) plane.

Crack Track Electron Requirements
No muon tracks
$ z_0  < 60$ cm
$( d_0  \leq 0.2$ cm & No Si. hits) or $( d_0  \leq 0.02$ cm & Si. hits)
COT: St. seg $\geq 3$ & Ax. seg $\geq 3$ and St. hits $\geq 15$ & Ax. hits $\geq 15$
Track not fiducial to calorimeter
$p_T \geq 20$ GeV
$\frac{Isol}{p_T} < 0.1$ or $\frac{EmIsol}{p_T} < 0.1$

Table 4.4: The selection requirements for electrons not fiducial to the calorimeters.

for each region are defined in Table 4.5. The cut values were selected by evaluating the resulting  $Z$  mass distribution in data and  $ZH$  signal. They were not numerically optimized. Attempts at this led to extreme cut-values with either a low purity or low efficiency. To reconstruct a  $Z$ , electrons from different classes can be paired. For instance, a central electron with a score of 0.5 could successfully be paired with a forward phoenix electron with a score of 0.9. Additionally, high-score central

electrons can be paired with cut-based crack track electrons to make a  $Z$ .

To enter the pretag selection category and be considered for a final analysis channel, the reconstructed  $Z$  object must have a mass between 76 and 106 GeV/ $c^2$ . All  $Z$ s made from two central electrons have an additional opposite charge requirement imposed.

Region	High Score	Low Score
Central	0.75	0.3
Forward Phoenix	0.5	0.0
Forward Non-Phoenix	0.6	0.3

Table 4.5: Electron identification neural net score criteria. The score range is -1 to 1.

## 4.2.2 Jet Selection & Higgs Candidate Reconstruction

Our signature has  $H \rightarrow b\bar{b}$ , so we begin the Higgs candidate reconstruction by first selecting events with at least two quark candidates. When a quark is produced in a collision, it hadronizes, fragments, and radiates gluons before passing through the detector producing a collimated cone of particles. These objects are called jets. We use jets found by the JETCLU [56] algorithm with a fixed cone size<sup>2</sup> of  $R = 0.4$ , and apply jet-energy corrections [57] to account for detector effects and multiple  $p\bar{p}$  interactions. Jets we consider must be outside of a cone of  $R = 0.4$  from a selected electron in order that it can be well-measured. Additionally, they must be within  $|\eta_{det}| < 2.0$  to preserve silicon tracking ability, and they must pass a staggered  $E_T$  cut of  $E_T(\text{jet}_1) \geq 25$  GeV and  $E_T(\text{jet}_2) \geq 15$  GeV. Events with a  $Z$  object and jets meeting these criteria make up our high-statistics ( $> 25$  times the number of events) model validation sample called “pretag” (see Figures 4.11-4.12 and B.1-B.8). Nearly

<sup>2</sup>The cone size is given in terms of  $\eta - \phi$  space,  $R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$ .

79% of  $ZH$  events with a reconstructed  $Z$  boson pass the pretag criteria.

#### 4.2.2.1 $b$ Tagging and Final Analysis Channels

We form our final analysis channels by requiring at least one of the jets to have likely originated from a  $b$  quark (see Table 4.6). The hadrons  $b$  quarks form live long enough to produce a displaced vertex of decay particles from the primary interaction point. Algorithms have been developed to identify these displaced vertices and their application is called “ $b$  tagging.” We use a combination of two algorithms to make three orthogonal final analysis channels. The SECVTX [58] (secondary vertex) algorithm has loose, tight, and ultra-tight operating points. We only use the loose and tight designations. The JETPROB [59] algorithm returns the probability that a jet did not originate from the primary vertex. We consider jets with JETPROB values less than 5%. When selecting the two candidate Higgs jets in each event, we check first if the jet is tight SECVTX tagged, then check for a loose SECVTX tag, and if both fail, we take the two highest  $E_T$  jets. For entry into the final analysis channels, we consider pairs of jets. The first channel considered requires that both jets have a tight SECVTX tag (double tight tag). If an event fails this, it is considered for the second channel which requires one jet to have a loose SECVTX tag and the other to have a JETPROB value  $< 5\%$  (double loose tag). If an event passes neither of these, it is considered for our last channel which requires only a single tight SECVTX tag (single tight tag). Of the events at pretag,  $\sim 60.7\%$  of events meet one of these channel’s criteria. Until their combined limits are calculated at the end of the analysis, these three categories are kept as distinct channels due to their different  $S/B$  values.

Analysis Level	Classification	Requirements
Model Validation	Pretag	$Z$ object, $ \eta_{det}  < 2.0$ , $\Delta R(\text{jet}, e) \geq 0.4$ , $E_T(\text{jet}_{1,2}) \geq 25, 15$ GeV
Pretag plus:		
Final Channels	Double Tight	two tight SECVTX tags
	Double Loose	one loose SECVTX tag another JETPROB $\leq 5\%$
	Single Tight	exactly one tight SECVTX tag

Table 4.6: Analysis event classification. The pretag sample serves as a model validation. The three tag categories are subsets of the pretag sample and are the final analysis channels.

#### 4.2.2.2 Neural Network Jet-Energy Corrections

The dijet mass ( $M_{jj}$ ) is the most powerful single variable in this analysis for identifying potential Higgs events. Yet the standard reconstruction at CDF leads to a dijet mass distribution with room for improvement. An earlier study found that improving the dijet mass resolution from 12% to 10% would lead to the same gain as increasing the amount of data by 20% [60].

In events where all decay products are measured, poorly reconstructed jets lead to an increase in  $\cancel{E}_T$ . Since our signature and most of our backgrounds meet this criterion, this feature was exploited in an earlier  $ZH$  analysis that developed a neural network to correct jet energies [61]. The network utilized correlations between  $\cancel{E}_T$ ,  $\vec{E}_T(\text{jet}_{1,2})$ , and the  $Z$  + dijet imbalance (see Table 4.7) to correct the jet energies (the directions were unchanged). The network was trained using a wide range of mass points to avoid biasing the corrections toward a particular mass window. The earlier analysis group’s efforts were fruitful; they saw an improvement of 50% in the dijet mass resolution [61]. This analysis applied a network developed for the  $2.7 \text{ fb}^{-1}$  analysis [62] and sees good improvement in the dijet mass distribution (see Figure 4.5).

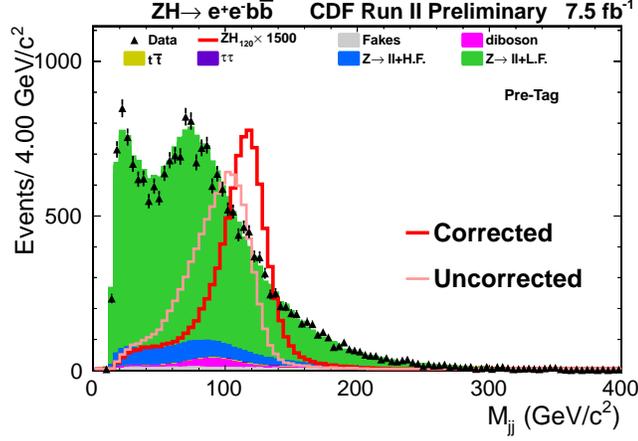


Figure 4.5: The dijet mass distribution,  $M_{jj}$ , for events passing the pretag requirements. The red and pink solid lines are the scaled signal distributions highlighting the effect of the dijet-neural-network corrections (red is corrected, pink is uncorrected). The network is applied to data and signal and background Monte Carlo alike. The uncorrected distribution is included only for signal here to simplify the presentation.

All kinematic variables are calculated with the corrected energies with the notable exception of  $\cancel{E}_T$ .  $\cancel{E}_T$  is a powerful variable in distinguishing our signal from backgrounds with neutrinos. As mentioned above, the network jet energy corrections are driven by  $\cancel{E}_T$  due solely to mismeasurement. In events with neutrinos, where the  $\cancel{E}_T$  is not due to mismeasurement, the corrections are not well-motivated, and applying them to their best-distinguishing variable dilutes its power. Therefore, to retain the discriminating power of  $\cancel{E}_T$  for these backgrounds, we do not recalculate  $\cancel{E}_T$ .

Variable	
$E_T(\text{jet}_1)$	$E_T(\text{jet}_2)$
$\Delta\phi(\text{jet}_1, \vec{\cancel{E}}_T)$	$\Delta\phi(\text{jet}_2, \vec{\cancel{E}}_T)$
$\frac{\vec{P}_T(Z+H) \cdot \vec{P}_T(\text{jet}_1)}{ P_T(\text{jet}_1) }$	$\frac{\vec{P}_T(Z+H) \cdot \vec{P}_T(\text{jet}_2)}{ P_T(\text{jet}_2) }$
$\frac{\vec{\cancel{E}}_{T,\text{jet}_1}}{ P_T(\text{jet}_1) }$	$\frac{\vec{\cancel{E}}_{T,\text{jet}_2}}{ P_T(\text{jet}_2) }$
$\cancel{E}_T$	Number of Jets

Table 4.7: The variables used in the neural network to correct jet energies.

## 4.3 Model

To test the Higgs boson and non-Higgs boson hypotheses, Monte Carlo and data-driven methods are used. Monte Carlo is the term given to simulated events that use parameters drawn randomly from distribution functions. Its fitting namesake is the world famous casino in Monaco where an uncle of one of the pioneers of Monte Carlo methods, Stanislaw Ulam, would play games of chance [63]. The Monte Carlo model is described in the following section. Data-driven methods are used to model events from misidentified electrons and  $b$  jets in Sections 4.3.4 and 4.3.5 respectively.

### 4.3.1 Monte Carlo

Monte Carlo is used to model all of the signal processes and all of the background processes not resulting from particle misidentification. Data-driven methods are used for misidentified processes since Monte Carlo does not accurately recreate these mistakes. With this in mind, the Monte Carlo processes considered all have two real electrons and two jets. For the final analysis channels, it is further required that the processes have two heavy flavored jets. These processes are:  $Z$ +jets, diboson ( $WW, WZ, ZZ$ ),  $t\bar{t}$ , and  $ZH$  for 11 mass hypotheses.

Events generated with Monte Carlo receive weights proportional to

$$w_i \sim \frac{\sigma \times \int \mathcal{L}_{data}}{N_{generated}} \quad (4.1)$$

where  $\sigma$  is the process cross section,  $\int \mathcal{L}$  is the integrated luminosity of the data, and  $N_{generated}$  is the total number of generated events. The events are further weighted to match the luminosity profile of the data. This is done by applying a weight that relates the number of vertices (a value directly related to the instantaneous luminosity) distribution in Monte Carlo at the pretag level to that of data at the pretag level.

The samples use the CTEQ5 parton distribution function (pdf) [64]. The  $Z$ +light flavor (l.f.;  $u, d, s$ ) Monte Carlo is generated using ALPGEN v2.10 prime [65] for the hard interaction and PYTHIA 6.325 [66] for showering. To account for differences between leading order (LO) and next leading order (NLO) cross sections in this sample, a 1.4 k-factor is applied. This sample is used only for modeling events passing the pretag criteria since there are no real  $b$  jets. It is replaced with a data-driven method at the tag level. The remaining Monte Carlo processes are modeled with PYTHIA 2.16, and it should be noted the  $t\bar{t}$  sample assumes a top quark mass of  $172.5 \text{ GeV}/c^2$ . All processes use GEANT3 [67, 68] for detector simulation. The cross sections applied for the processes are listed in Table 4.8 for background and Table 4.9 for signal.

### 4.3.2 Trigger Modeling

As mentioned earlier, we are not able to analyze every collision event produced at CDF. In an effort to maximize the number of interesting events analyzed, we apply triggers described in Section 4.1.1. By requiring data events to pass a trigger and

Process	Generator	$\sigma$
$Z+l.f.$	ALPGEN+PYTHIA	4.66 fb to 2111 pb
$Z+c\bar{c}$	ALPGEN+PYTHIA	148.4 to 1512 fb
$Z+b\bar{b}$	ALPGEN+PYTHIA	53.9 to 715.4 fb
$WW$	PYTHIA	11.34 pb
$WZ$	PYTHIA	3.47 pb
$ZZ$	PYTHIA	3.62 pb
$t\bar{t}$	PYTHIA	7.04 pb

Table 4.8: The production rates used in normalization of Monte Carlo model components. The l.f. denotes light-flavor jets ( $u,d,s$ ). The  $Z$ +jets cross sections ( $\sigma$ ) include a k-factor of 1.4 to account for the difference between NLO and LO calculations. There is a range of cross sections for the ALPGEN+PYTHIA processes because they are generated in different samples according to additional number of partons included in the scattering.

be saved, we are losing some events that would otherwise pass our selection. This could either be due to the event failing to meet all of the criteria, or due to a trigger inefficiency or detector effect (for instance, failing to resolve two electrons). To account for the loss of these data events in the MC model, earlier analyses have applied standard CDF efficiency functions in electron  $\eta$  and  $E_T$ . These do not translate well into the di-electron triggers we now additionally consider for added acceptance (see Table 4.10). In this analysis, we develop a single efficiency function of the combined (three triggers) trigger decision using a regression neural network.

This regression network approach has been used before [69], but is new to this analysis. To create the network, we first run over an orthogonal data stream (events saved due to any quality other than electrons). The independent data set used was collected due to its significant  $\cancel{E}_T$  characteristics. In this sample, 36,596 events ( $\sim 0.007\%$ ) had an identified  $Z$  object and  $\sim 83\%$  of these fired at least one of our triggers allowing us to adequately train a network to emulate the trigger decision. The network was trained using  $\Delta R(e_1, e_2)$ ,  $Z$  mass, electron energies, track  $p_{TS}$ ,  $\eta_{det}$ 's, and Had/Ems. The entire stream (all run periods) was considered to properly

$M_H$ (GeV/ $c^2$ )	$\sigma$ (fb)	$BR(H \rightarrow b\bar{b})$
100	169.8	0.8033
105	145.9	0.7857
110	125.7	0.7590
115	103.9	0.7195
120	90.2	0.6649
125	78.5	0.5948
130	68.5	0.5118
135	60.0	0.4215
140	52.7	0.3304
145	46.3	0.2445
150	40.8	0.1671

Table 4.9: Signal production rates ( $\sigma$ ) and branching ratios ( $BR$ ). An additional factor of 0.10095 is included to account for  $BR(Z$  to charged leptons).

Trigger	Fired (%)	Fired Exclusively (%)
Electron Central 18	$76.5 \pm 0.8$	$6.3 \pm 0.2$
Z No Track	$85.4 \pm 0.9$	$5.7 \pm 0.2$
Z No Track Mass	$67.4 \pm 0.7$	$4.4 \pm 0.1$

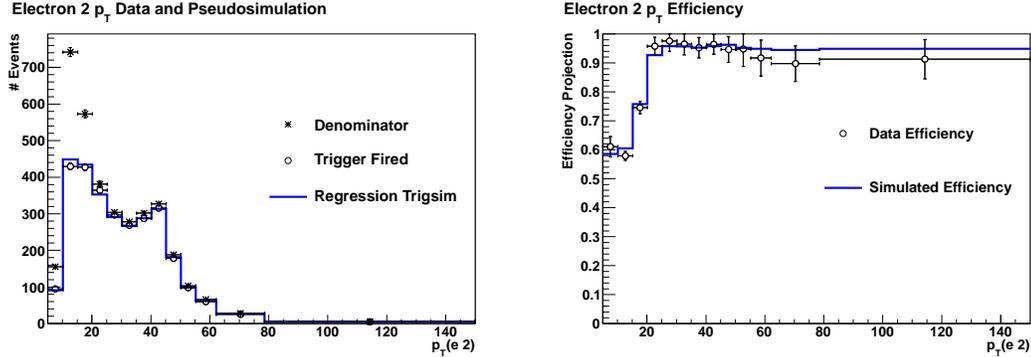
Table 4.10: The percent of pretag data events that fired a given electron trigger. There is about an 10% gain in pretag data by including the triggers without a track requirement.

account for the fraction of data where Z\_NOTRACK and Z\_NOTRACK\_MASS were not active <sup>3</sup>.

The output of the network corresponds to the probability that a given event would fire at least one of the triggers. To avoid spurious values in the tails of network distributions, an ensemble of neural networks is evaluated, and the median value used. This has been found to be more robust [70, 71]. A self-consistency check of the trigger model was satisfactory and is shown in Figure 4.6. For each MC event entering the selection, the ensemble output is applied as a weight reflecting the probability that we would save a given event in data. The effect of this can be seen

<sup>3</sup>The Z\_NOTTRACK trigger was added during the first run period ( $< 0.3 \text{ fb}^{-1}$ ), and the Z\_NOTRACK\_MASS trigger has been in effect since  $\sim 2.45 \text{ fb}^{-1}$  of data were collected.

in Figure 4.7.



(a) Trigger decision and model in  $p_T(e_2)$  of the training set. (b) Efficiency in  $p_T(e_2)$  of the training set.

Figure 4.6: Figure 4.7a shows the  $p_T$  of the second electron in the  $\cancel{E}_T$  (training) stream for all  $Z$  events (star), and  $Z$  events that fired one of our considered triggers (circles). The solid blue line is the result of the trigger regression model applied to all  $Z$  events in the stream (denominator). Figure 4.7b is the efficiency as a function of  $p_T(e_2)$  of the trigger selection in  $Z$  events in the  $\cancel{E}_T$  stream. It is found by taking the fired values from Figure 4.7a and dividing by the denominator. (Files for the plots were provide by H. Wolfe and edited by the author.)

Events with a  $Z$  object made with a CrkTrk electron do not use the multivariate trigger model. In these events, the crack track type electron would not fire or contribute to any of the electron triggers since the EM calorimeter deposits are low. To model the trigger effects in these events, a single-electron turn-on efficiency (in  $\eta_{det}$ ,  $E_T$ , and run number) for the ELECTRON\_CENTRAL\_18 trigger is applied for the neural-network-identified electron (the method is described in [72]).

### 4.3.3 Electron Corrections

#### 4.3.3.1 Electron Efficiency Scale Factors

While effort was made to have the electron identification behave well in both data and Monte Carlo (*i.e.* variables well-modeled in both), differences in efficiencies must still

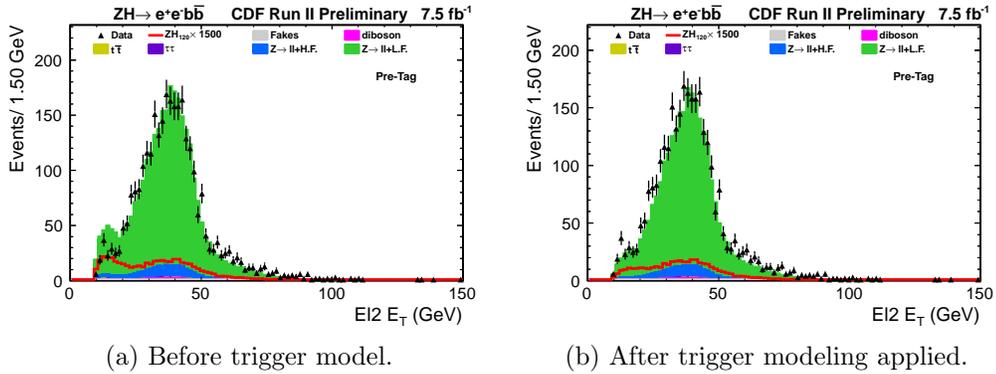


Figure 4.7: Effect of the trigger modeling on the second leading electron  $E_T$  in  $Z$ 's made from two forward electrons. This is at the pretag selection level.

be accounted. One can imagine that the selection might behave better in Monte Carlo where the events are sometimes more ideal. To account for this, a scale factor is found and applied to the Monte Carlo. The scale factor is a ratio of efficiencies:  $\epsilon_{data}/\epsilon_{MC}$ . The efficiency is found by taking events in which one electron passes the high score central cut (this is considered a “tagged” leg), and the other possible electron satisfies the preselection cuts<sup>4</sup> (this is a “probe” leg). The efficiency is found for each type of electron and is determined by the region of this latter candidate. The number of events forming a  $Z$  mass between  $86\text{-}96 \text{ GeV}/c^2$  determines the denominator of the calculation. The numerator is found by evaluating how many candidate electrons pass the selection score requirements. This is done in  $Z + jets$  Monte Carlo and data. When finding the efficiency in data, first, a falling exponential shape from QCD interactions is fit and removed. The removal of this shape is well-motivated since there are many events in the data with one good central electron plus jets that may satisfy the preselection cuts. The resulting scale factors are given in Table 4.11.

An alternate method of finding efficiencies was also tested. This technique fol-

<sup>4</sup>The preselection cuts were found to be highly efficient in both data and Monte Carlo.

Electron Class	High Score	Low Score
Central	1.00	1.02
Forward Phoenix	0.99	1.00
Forward Non-Phoenix	0.75	0.82

Table 4.11: The scale factors applied to electrons in Monte Carlo to account for differences in efficiencies between data and Monte Carlo. The systematic error is taken to be 0.43%, and the combined (with statistical errors) error is estimated to be 1%.

lowed a more traditional tag-and-probe method used at CDF [73]. The events considered required  $\cancel{E}_T < 15$  GeV in an effort to remove  $W$  events and exactly two electron candidates to avoid combinatorics in recreating a  $Z$  boson. In each event, one electron candidate is randomly chosen and tested as a tag leg passing standard CDF central electron cuts [74], and the other electron candidate is tested for probe requirements (see Table 4.12). If either fails, or if the invariant mass is outside of 76-106 GeV/ $c^2$ , the event is not considered. Otherwise, it enters the denominator. The numerator is again found by evaluating the neural network selection on the probe leg. The efficiencies can be found in Table 4.13. While not directly used in the analysis, the scale factors agree with the systematic uncertainty assessed and they give an estimate to the efficiency of the selection.

Central	Forward Phoenix	Forward Non-Phoenix
Not flagged a conversion Em $E_T > 9$ GeV		
Had/Em < 0.125 Fiducial Track $z_0 < 60$ cm $Q_1 + Q_2 = 0$	Had/Em < 0.0625	Had/Em < 0.125 $1.2 \leq  \eta  \leq 2.8$ $E/P$ defined Energy > 0

Table 4.12: The cuts used in the alternate method of finding efficiencies. These are currently not applied in the analysis, but are meant to serve as a scale for the identification efficiency.

	High Score	Low Score
Central $\epsilon_{data}$	$0.942 \pm 0.004$	$0.978 \pm 0.004$
Central $\epsilon_{MC}$	$0.940 \pm 0.002$	$0.978 \pm 0.002$
Scale Factor	$1.002 \pm 0.005$	$1.000 \pm 0.005$
Forward Phoenix $\epsilon_{data}$	$0.891 \pm 0.004$	$0.956 \pm 0.005$
Forward Phoenix $\epsilon_{MC}$	$0.917 \pm 0.003$	$0.973 \pm 0.003$
Scale Factor	$0.972 \pm 0.006$	$0.983 \pm 0.006$
Forward Non-Phoenix $\epsilon_{data}$	$0.540 \pm 0.005$	$0.658 \pm 0.005$
Forward Non-Phoenix $\epsilon_{MC}$	$0.812 \pm 0.004$	$0.890 \pm 0.005$
Scale Factor	$0.664 \pm 0.007$	$0.739 \pm 0.007$

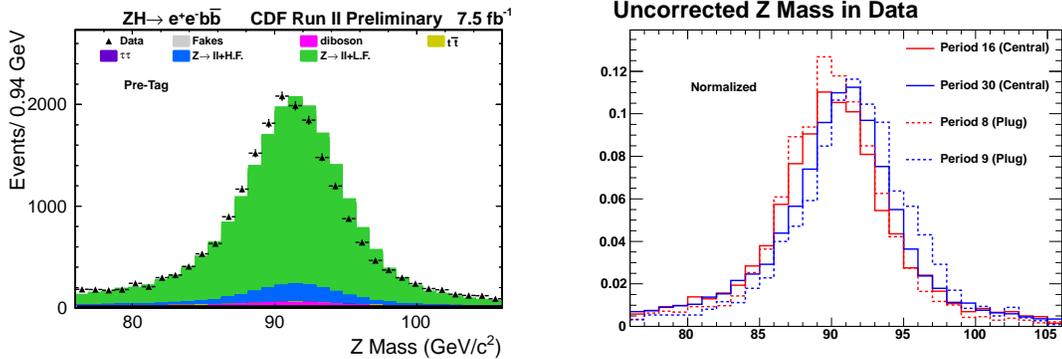
Table 4.13: The alternate method of finding efficiencies. These are currently not applied in the analysis, but are meant to serve as a scale for the identification efficiency.

#### 4.3.3.2 Data Energy Scale Factors

As can be seen in Figure 4.8a, there is poor agreement in the  $Z$  mass distribution. This was found to be at least partially due to calorimeter calibrations dependent in time [55] (see Figure 4.8b). To correct for this, an energy correction factor is calculated for each data period and applied to the data entries. The factor is calculated by fitting the  $Z$  mass distribution between 76 to 106 GeV/ $c^2$  in data for a given run period and Monte Carlo with a Gaussian function. In an effort to avoid bias by correcting to the model validation region, only orthogonal events with fewer than two jets enter the mass distribution. The ratio of the means (Monte Carlo over data) is then taken as the correction factor. This value is calculated separately for the central and forward regions of the detector by only considering  $Z$  events with two central electrons or two plug electrons in the calculation. The energy correction values applied to data are listed in Table 4.14, and their effect on the  $Z$  mass distribution is shown in Figure 4.11a.

Period	Central Scale Factor	Plug Scale Factor
0	$1.0036 \pm 0.0005$	$1.0014 \pm 0.0005$
1-4	$1.0065 \pm 0.0005$	$1.0092 \pm 0.0006$
5-7	$1.0083 \pm 0.0006$	$1.0064 \pm 0.0008$
8	$1.0078 \pm 0.0007$	$1.0158 \pm 0.0011$
9	$1.0088 \pm 0.0008$	$0.9979 \pm 0.0009$
10	$1.0091 \pm 0.0006$	$1.0079 \pm 0.0008$
11	$1.0108 \pm 0.0007$	$1.0086 \pm 0.0008$
12	$1.0130 \pm 0.0009$	$1.0050 \pm 0.0009$
13	$1.0131 \pm 0.0006$	$1.0010 \pm 0.0007$
14	$1.0081 \pm 0.0017$	$1.0083 \pm 0.0025$
15	$1.0120 \pm 0.0008$	$1.0122 \pm 0.0011$
16	$1.0136 \pm 0.0012$	$1.0139 \pm 0.0016$
17	$1.0116 \pm 0.0007$	$0.9986 \pm 0.0008$
18	$1.0070 \pm 0.0006$	$1.0006 \pm 0.0009$
19	$1.0039 \pm 0.0006$	$1.0054 \pm 0.0012$
20	$1.0051 \pm 0.0006$	$1.0105 \pm 0.0013$
21	$1.0065 \pm 0.0005$	$1.0115 \pm 0.0009$
22	$1.0039 \pm 0.0005$	$1.0062 \pm 0.0010$
23	$1.0045 \pm 0.0006$	$1.0089 \pm 0.0013$
24	$1.0039 \pm 0.0006$	$1.0088 \pm 0.0012$
25	$1.0043 \pm 0.0006$	$1.0019 \pm 0.0010$
26	$1.0051 \pm 0.0007$	$1.0045 \pm 0.0013$
27	$1.0040 \pm 0.0005$	$1.0030 \pm 0.0007$
28	$1.0040 \pm 0.0005$	$1.0052 \pm 0.0008$
29	$1.0038 \pm 0.0005$	$1.0070 \pm 0.0007$
30	$1.0035 \pm 0.0004$	$1.0051 \pm 0.0007$
31	$1.0078 \pm 0.0009$	$1.0059 \pm 0.0013$
32	$1.0055 \pm 0.0005$	$1.0053 \pm 0.0007$
33	$1.0044 \pm 0.0005$	$1.0063 \pm 0.0008$
34	$1.0042 \pm 0.0006$	$1.0071 \pm 0.0009$

Table 4.14: Electron energy correction factors by period. This is a data-only correction applied to each electron from a  $Z$  object decay.



(a) Pre-tag  $Z$  mass with no period-based energy corrections. (b) Pre-tag data for select periods with no period-based energy corrections.

Figure 4.8: The  $Z$  mass distributions before period-based energy corrections are applied to data. Figure 4.8a is the  $Z$  mass distribution at the pre-tag level before any period-based energy corrections are applied. It highlights the cumulative effect on the resolution of omitting the period-based energy corrections. Figure 4.8b shows the discrepancies the  $Z$  mass peak between a subset of data periods. These latter distributions are normalized to one and do not have a jet requirement.

#### 4.3.3.3 Plug-Energy Smearing

After applying the energy corrections based on calorimetry calibrations to the data, there is still noticeable disagreement in the  $Z$  mass distribution for  $Z$ 's with a forward electron. The simulation has better energy resolution for forward electrons than is seen in practice (see Figure 4.9a). To account for this, a smearing function is developed for application to the Monte Carlo. The function is made by first taking all of the anti-tagged  $Z$  events in data (an event with a  $Z$  object that does not enter the final analysis channels), and fitting their  $Z$  mass distribution with a Gaussian between 85 and 97 GeV/c<sup>2</sup>. The  $\sigma$  of this resultant fit is then taken as the smear factor (here,  $\sim 3.55$ ), and applied to the Lorentz vectors of forward electrons as

$$L_{smear} = L_0 \left( 1 + \frac{Rdm}{100} \right) \quad (4.2)$$

where  $Rdm$  is a random value drawn from a Gaussian centered at zero and with

the standard deviation equal to the smear factor. While improvements could be developed to better characterize the relationship between the data and Monte Carlo resolution in the forward region, this method was applied as a first step with good results (see Figure 4.9).

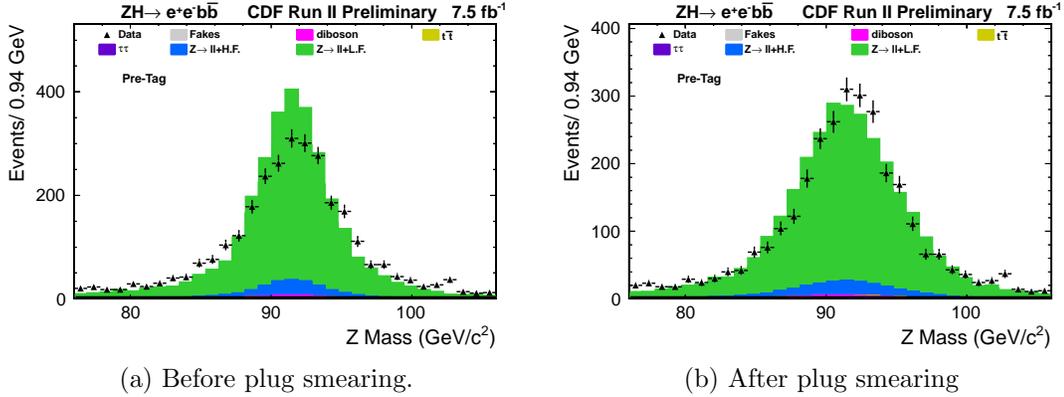


Figure 4.9: Above are  $Z$  mass distribution shape comparisons (Monte Carlo model is normalized to data) for  $Z$  objects made of two forward electrons at the pretag level. The data are not as well-measured as the Monte Carlo is modeled. This is the motivation for the smearing of the forward electron energies.

### 4.3.4 Misidentified Electrons

Non-electron objects, such as jets or photons, can pass the electron selection criteria. These objects are referred to colloquially as “fake electrons,” or simply as “fakes.” To model the events entering the selection due to these objects, data-driven methods are employed.

The events due to misidentified electrons entering the Central + CrkTrk category are modeled from same-sign events otherwise passing that selection.

In the other  $Z$  categories, the fake contribution derives from a single electron fake rate. As the name suggests, the fake rate is the rate at which jet objects pass the electron selection criteria. To calculate this rate, all jet-triggered data are processed.

Events with more than one possible electron objects are rejected to remove real electrons from  $Z$  bosons. Events with  $\cancel{E}_T > 15$  GeV are ignored to remove real electron contamination from  $W$  boson decays. The jet leading in  $p_T$  is also ignored in an effort to remove trigger turn-on effects, and trigger bias. The remaining jets in a given region of the detector enter the denominator. For instance, in the central fake rate calculation, only jets in the central region enter the denominator. The numerator is made of electrons in these events passing the selection criteria and matching a denominator jet from the same event within a  $\Delta R$  cone of 0.4. In addition to different electron categories, the fake rate is calculated for different  $E_T$  ranges. Four data streams differentiated by their jet  $E_T$  trigger threshold are processed for this calculation: jet20, jet50, jet70, and jet100. The average rate of the four is taken as the nominal value while the spread and combined error lead to a conservative rate uncertainty of 50%.

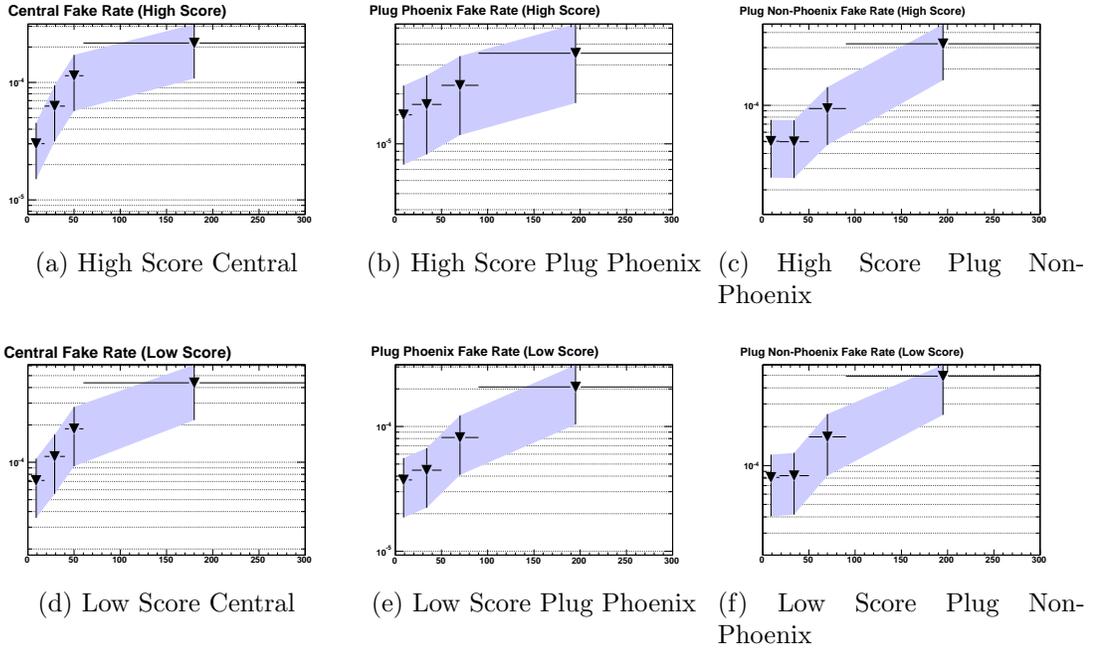


Figure 4.10: Measured fake rates for the neural net electron selection. The low-score fake rate is not exclusively low pass. It is low score and above.

To model the events entering our selection due to non-electron objects passing the neural-network electron identification, we apply the applicable fake rates to all single electron plus jet pairs in a data set. Each of these events then enters the background model with a weight equal to the probability that the jet would fake an electron. We do not apply the fake rate to jet plus jet pairs since the electron plus jet set should already include the proper ratio of events where the identified electron is actually due to a fake object.

### 4.3.5 Mistagged $b$ Jets

In addition to misidentified electrons, events enter the final analysis channels due to misidentified  $b$  jets. These jets are called mistags. Since both  $b$ -tagging algorithms attempt to find jets originating from displaced vertices, mistags result from detector resolution effects,  $K_s$  and  $\Lambda$  decays, and interactions with detector materials [75]. This first source results in mistags distributed symmetrically about the interaction point while the second two are biased toward the positive tag direction.

In an effort to determine the mistag frequency, first the rate of negative tags is measured, and then corrections are developed to account for tag direction asymmetry. The rate of mistagging is parameterized by six variables (see Table 4.15) and packaged in a look-up table or “mistag matrix” (binning and further description in [76], [77]). The rate in each cell or bin in the matrix is found by fitting MC templates of  $b$ ,  $c$ , and light-flavored jets to data in a variable sensitive to heavy-flavor. In SECVTX, this is the reconstructed mass of the track associated with the jet. In JETPROB, it is a multivariable fit involving impact parameter distance variables and track mass (for details, see [78] for JETPROB and [75, 79] for SECVTX).

To model the events entering the final analysis channels due to a misidentified  $b$  jet, the mistag rate is evaluated for all jets in events of the pretag data sample.

Mistag Parameterization Variables
$E_T$ of jet (uncorrected)
$\eta$ of jet
Number of track associated with jet, $N_{trk}$
$\sum E_T$ of jets
Position of vertex, $z_{vtx}$
Number of vertices in event, $N_{vtx}$

Table 4.15: Variables used in the parameterization of mistag probabilities.

Event kinematics are calculated using the two highest  $E_T$  jets and the total event weight is the probability that any combination of event jets would be mistagged and enter the channel.

### 4.3.6 $b$ -Tag Scale Factors

The  $b$ -tag algorithms applied in this analysis are more efficient at tagging  $b$  jets in Monte Carlo than data. To account for this in the modeling, a  $b$ -tag scale factor is introduced. This is again found by taking the ratio of efficiencies between data and Monte Carlo. The efficiency for SECVTX is found by evaluating dijet events where one jet is positively tagged and the other has an associated electron. The efficiency is then related by the ratio of true heavy-flavor electron jet tags over the number of events considered [80]. The efficiency in data of the JETPROB algorithm is found by adding templates of heavy-flavor and light-flavor jet  $p_{TS}$  to determine the composition of data [81]. In Monte Carlo, it is found by looking at the generator information. In both cases, the scale factors reported are constants (see Table 4.16), and any slight variable dependencies are included as a systematic effect (see Section 5.1).

In applying the  $b$ -tag scale factors, it is worth noting that the three tag categories are not independent. For instance, an event is only considered for the double loose

Tag Operating Point	Scale Factor
Tight SECVTX	$0.96 \pm 0.05$
Loose SECVTX	$0.98 \pm 0.07$
JETPROB < 5%	$0.78 \pm 0.04$

Table 4.16: The efficiency scale factors for single  $b$ -tag operating points.

Tag Category	Scale Factor
Double Tight	$0.96 \times 0.96$
Double Loose	$0.98 \times 0.78 \times 1.05$
Single Tight	$0.96 \times 1.07$

Table 4.17: The efficiency scale factors applied for the  $b$ -tag categories.

tag category if it fails the double tight tag requirements. To account for this, a modified scale factor is applied to the model following the prescription used in the Higgs plus  $W$  boson analysis [82, 83]. The final scale factors applied are listed in Table 4.17.

## 4.3.7 Model Validation

### 4.3.7.1 Pretag Model Validation

In this analysis, a fully independent control region for model validation is not used. Instead, the pretag sample serves as a high-statistics model validation (< 4% of events enter the final analysis channels). Table 4.18 shows the agreement between the data and model at the pretag level. Figures 4.11 illustrate the agreement in the reconstructed  $Z$  mass distribution and number of jets distribution, while Figures 4.12 show the  $\cancel{E}_T$  in linear and log scales. All of the variables used in the final discriminant are shown at the pretag level in Appendix B.

$ZH \rightarrow e^+e^-b\bar{b}$  Analysis CDF Run II Preliminary ( $7.5 \text{ fb}^{-1}$ )

Data	21122
$ZH_{120}$	$6.2 \pm 0.70$
$t\bar{t}$	$126 \pm 17$
Diboson	$397 \pm 34$
$Z/\gamma^* \rightarrow ee + \text{h.f.}$	$1786 \pm 561$
$Z/\gamma^* \rightarrow ee + \text{l.f.}$	$18783 \pm 4229$
Fakes	$354 \pm 177$
Model	$21446 \pm 4300$

Table 4.18: Event totals at the pretag level.

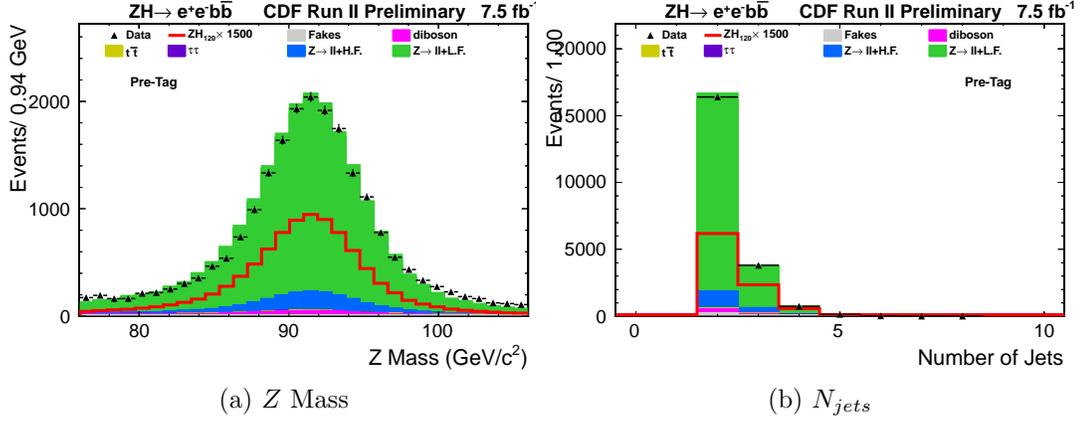


Figure 4.11: The  $Z$  mass and number of jet distributions for the data and model at the pretag selection level (before  $b$  tagging is imposed).

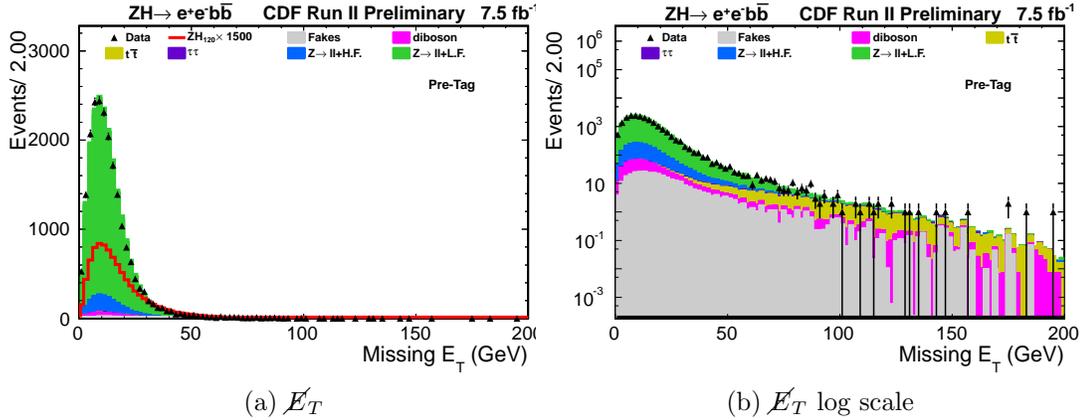


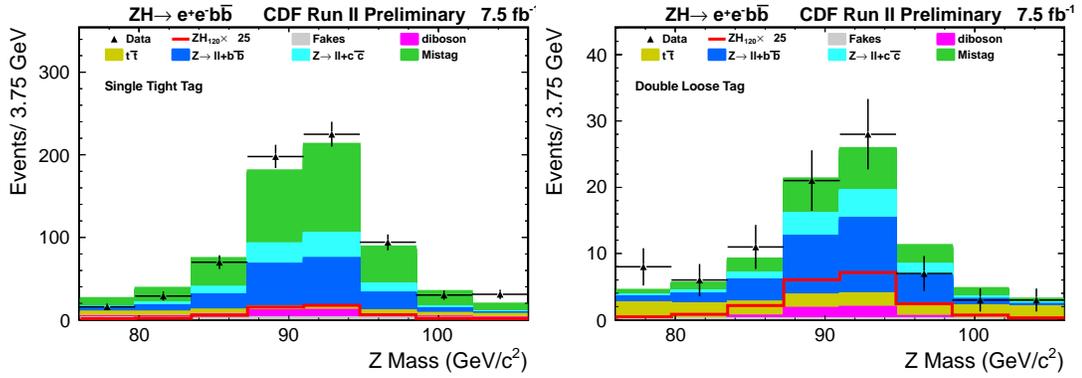
Figure 4.12: The  $\cancel{E}_T$  distribution on linear and log scales for the data and model at the pretag selection level (before  $b$  tagging is imposed).

### 4.3.7.2 Final Analysis Channels: Tag Level Model

After the model is validated,  $b$  tagging is imposed and the final analysis channels are populated. The event totals are listed in Table 4.19. The  $Z$  mass, dijet mass, and  $\cancel{E}_T$  are shown in Figures 4.13-4.15 while all of the inputs to the final discriminant are shown in Appendix C for each channel.

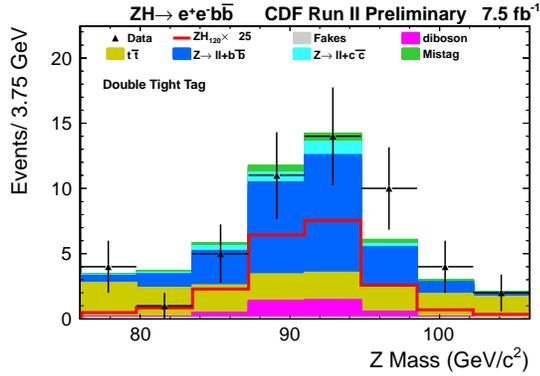
$ZH \rightarrow e^+e^-b\bar{b}$ Analysis CDF Run II Preliminary (7.5 fb <sup>-1</sup> )			
	Single Tight Tag	Loose Double Tag	Double Tight Tag
Data	693	87	51
$ZH_{120}$	$2.0 \pm 0.2$	$0.8 \pm 0.1$	$0.9 \pm 0.1$
$t\bar{t}$	$42 \pm 6$	$17 \pm 2$	$16 \pm 3$
Diboson	$27 \pm 3$	$5.7 \pm 0.7$	$4.3 \pm 0.6$
$Z/\gamma^* \rightarrow ee + \text{h.f.}$	$254 \pm 81$	$43 \pm 14$	$27 \pm 10$
Mistags	$333 \pm 47$	$20 \pm 5$	$2.2 \pm 0.6$
Fakes	$25 \pm 12$	$0.4 \pm 0.2$	$0.2 \pm 0.1$
Model	$681 \pm 120$	$86 \pm 20$	$50 \pm 13$

Table 4.19: Event totals at the tag levels.



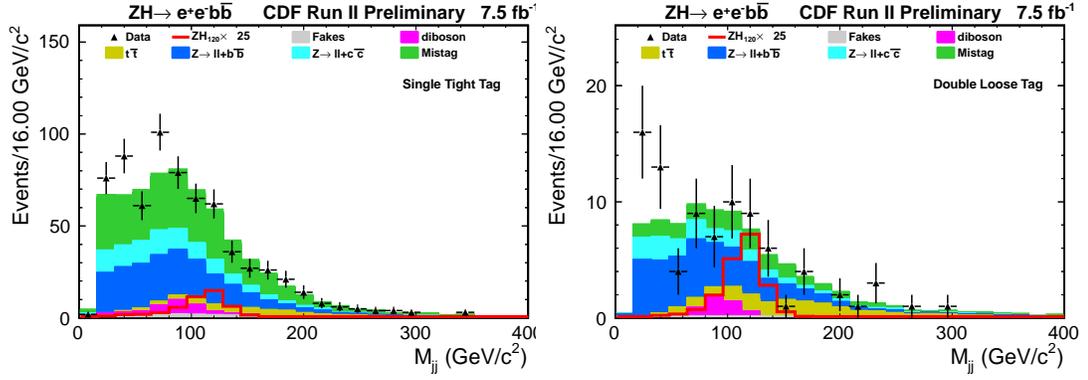
(a) Single Tight Tag

(b) Double Loose Tag



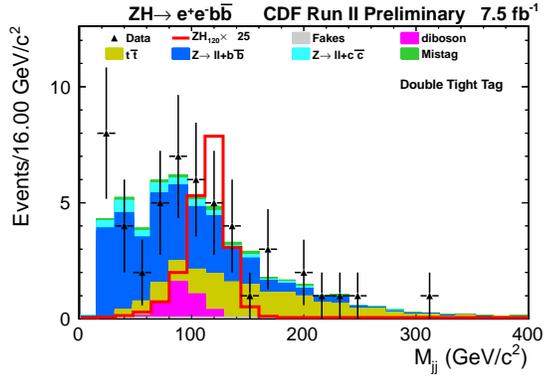
(c) Double Tight Tag

Figure 4.13: The  $Z$  mass distribution for the data and model for each final analysis channel.



(a) Single Tight Tag

(b) Double Loose Tag



(c) Double Tight Tag

Figure 4.14: The dijet mass ( $M_{jj}$ ) distribution for the data and model for each final analysis channel.

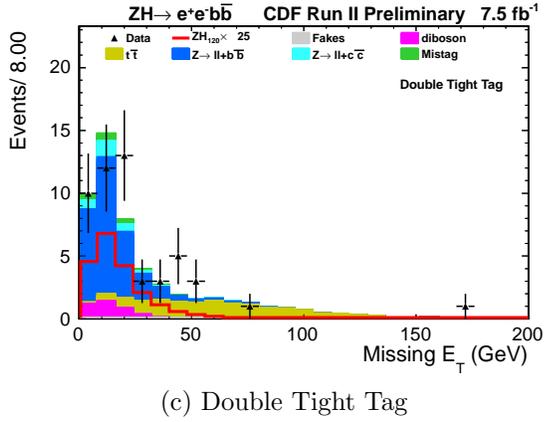
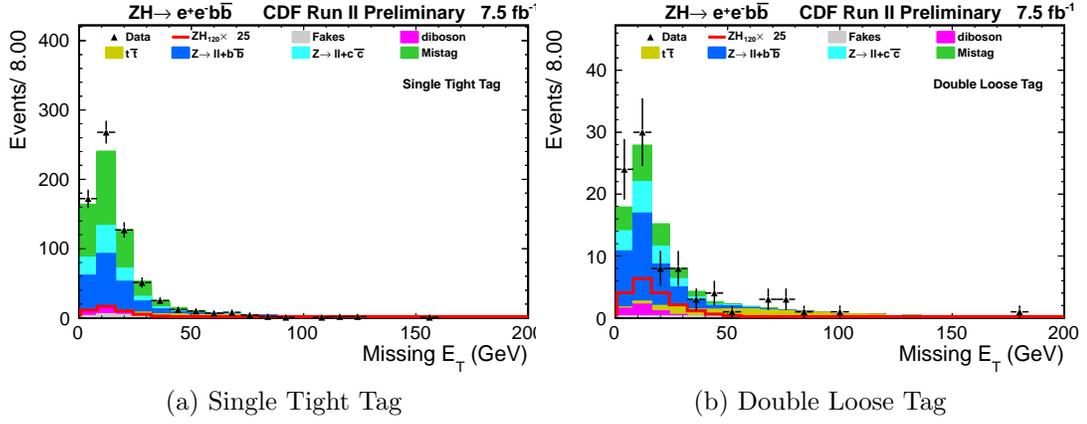


Figure 4.15: The  $\cancel{E}_T$  distribution for the data and model for each final analysis channel.

# Chapter 5

## Systematic Uncertainties

To set credible limits on Higgs production, we must properly take into account systematic effects. Underestimating their effect leads to overstated results; overestimating them dilutes the limit. Here, systematic effects fall into two classes: rate uncertainties and shape uncertainties. As the name implies, rate uncertainties affect the rate, or normalization of a process. Shape uncertainties can affect where an event is placed in a distribution (by migration). These can also affect the normalization by migrating events above and below variable cut values.

### 5.1 Rate Uncertainties

An uncertainty due to the luminosity measurement is assessed on all of the Monte Carlo (MC). There is a 3.8% uncertainty due to the calculation of the total inelastic cross section, and a 4.4% uncertainty due to the acceptance and efficiency of CDF's CLC (luminosity measurement) [84].

A 1% uncertainty is attributed to the trigger model. This value was determined by training the trigger network on only a subset of events, then evaluating it on the remaining events. This was tested many times and 1% accounts for the variation in

total rates.

Systematic allowances due to uncertainty in a process' theoretical cross section are also assigned. The diboson ( $ZZ$ ,  $WZ$ ,  $WW$ ) MC samples receive a 6% uncertainty [85]. The  $Z$ + heavy flavored (h.f.) jets MC receives a 40% uncertainty [86]. The  $t\bar{t}$  sample receives a 10% uncertainty meant to also account for the difference between  $t$ 's experimentally-measured mass and the MC's generated mass of  $172.5 \text{ GeV}/c^2$  [87]. The  $ZH$  signal is assessed a 5% uncertainty [33]. This  $ZH$  uncertainty is used in fitting the model in the signal plus background hypothesis and not used in the limit calculation.

The estimated contribution due to misidentified electrons is assessed a 50% rate uncertainty. This is to account for the difference in rates obtained by using different jet-triggered data streams.

An uncertainty on the  $b$ -tag scale factor is applied to the heavy-flavor MC with a value of 5.2% for single tight tag, 8.7% for loose double tag, and 10.4% for tight double tag categories. This uncertainty is due to varying ratios of charm quark fractions, bias in track multiplicity in the scale factor calculation, and uncertainty on the ratio of mistags to negative tags [88, 80, 77]. For JETPROB, the uncertainty is found by running on an independent sample and comparing predicted and observed rates.

The uncertainty in rate due to the electromagnetic (EM) energy scale is 3%. This is found by evaluating the effect on acceptance from the period-based  $Z$  peak corrections and plug-energy smearing. The uncertainty on the lepton ID is found to be 2% by varying the  $Z$  mass window considered from 86-96 to 76-106  $\text{GeV}/c^2$  in the scale factor calculation.

The effect of varying initial and final state radiation (ISR/FSR) is found by evaluating the change in acceptance in samples with more or less radiation. Following

CDF precedence, the effect is assessed in signal MC. Here, 120 GeV/ $c^2$  Higgs mass samples are used and the parameters in PYTHIA are varied to produce  $\pm 1\sigma$  shifts in ISR or FSR (the shifts are found by studying gluon radiation in DY events) [89]. The effects measured in these samples are used for all mass hypotheses.

## 5.2 Shape Uncertainties

The jet-energy scale (JES) systematic is assessed as a shape uncertainty. Two altered final discriminants are created using the  $\pm 1\sigma$  variations of the joint physics jet-energy corrections on the leading and sub-leading jets [90]. This can lead to a change in rate by migrating jets in events above or below the jet  $E_T$  cuts.

Unfortunately, simply using the varied JES outputs leads to poor behavior. The evidence of this is the result of a fit that the MCLIMIT machinery performs. It takes the sample templates and assigned uncertainties and fits for the most-likely values of templates and systematic effects. Using only the up and down shapes for the JES systematic, the fit constrains the uncertainty on the effect to nearly  $\sim 0.5\%$  of a standard deviation on some mass hypotheses. This seems to be due to large upward individual bin fluctuations in both the up and down shapes [91]. To handle this, a multistep adjustment is made. First, individual bin fluctuations are capped at 200% for the JES. Then, the up (down) templates are redefined to be the unshifted JES shape plus (minus) the difference between the JES up shape and the down shape over two (up (down) = nominal  $\pm (-) \frac{\text{up-down}}{2}$ ). The new templates are then renormalized to reflect the change in acceptance due to shifting the JES.

Finally, a shape uncertainty is assessed on the mistag contribution to account for the  $E_T$  turn-on and fit errors in the calculation [92].

Single Tight Tag  $ZH \rightarrow e^+e^-b\bar{b}$  Analysis CDF Run II Preliminary (7.5 fb<sup>-1</sup>)

Contribution	Fakes	Top	WW	WZ	ZZ	Z + bb	Z + c $\bar{c}$	Z+l.f.	ZH
Luminosity ( $\sigma_{\text{inel}}(p\bar{p})$ )	0	3.8	3.8	3.8	3.8	3.8	3.8	0	3.8
Luminosity Monitor	0	4.4	4.4	4.4	4.4	4.4	4.4	0	4.4
Trigger Emulation	0	1	1	1	1	1	1	0	1
Lepton ID	0	2	2	2	2	2	2	0	2
Lepton Energy Scale	0	3	3	3	3	3	3	0	3
ZH Cross Section	0	0	0	0	0	0	0	0	5
Fake Leptons	50	0	0	0	0	0	0	0	0
B-Tag Efficiency	0	5.2	5.2	5.2	5.2	5.2	5.2	0	5.2
$t\bar{t}$ Cross Section	0	10	0	0	0	0	0	0	0
Diboson Cross Section	0	0	6	6	6	0	0	0	0
$\sigma(p\bar{p} \rightarrow Z + HF)$	0	0	0	0	0	40	40	0	0
ISR/FSR	0	0	0	0	0	0	0	0	4.0
Mistag Rate (shape dep.)	0	0	0	0	0	0	0	+13.9 -13.8	0
Jet Energy Scale (shape dep.)	0	+1.9 -2.5	+19.6 -4.0	+5.2 -6.2	+5.3 -7.1	+12.1 -11.1	+4.1 -9.9	0	+3.0 -4.3

Table 5.1: Systematic uncertainties on the contributions for the Single Tight Tag channel. Systematic uncertainties are listed by name. Systematic uncertainties for  $ZH$  shown in this table are obtained for  $m_H = 115$  GeV/ $c^2$ . Uncertainties are relative, in percent and are symmetric unless otherwise indicated.

 Loose Double Tag  $ZH \rightarrow e^+e^-b\bar{b}$  Analysis CDF Run II Preliminary (7.5 fb<sup>-1</sup>)

Contribution	Fakes	Top	WW	WZ	ZZ	Z + bb	Z + c $\bar{c}$	Z+l.f.	ZH
Luminosity ( $\sigma_{\text{inel}}(p\bar{p})$ )	0	3.8	3.8	3.8	3.8	3.8	3.8	0	3.8
Luminosity Monitor	0	4.4	4.4	4.4	4.4	4.4	4.4	0	4.4
Trigger Emulation	0	1	1	1	1	1	1	0	1
Lepton ID	0	2	2	2	2	2	2	0	2
Lepton Energy Scale	0	3	3	3	3	3	3	0	3
ZH Cross Section	0	0	0	0	0	0	0	0	5
Fake Leptons	50	0	0	0	0	0	0	0	0
B-Tag Efficiency	0	8.7	8.7	8.7	8.7	8.7	8.7	0	8.7
$t\bar{t}$ Cross Section	0	10	0	0	0	0	0	0	0
Diboson Cross Section	0	0	6	6	6	0	0	0	0
$\sigma(p\bar{p} \rightarrow Z + HF)$	0	0	0	0	0	40	40	0	0
ISR/FSR	0	0	0	0	0	0	0	0	4.0
Mistag Rate (shape dep.)	0	0	0	0	0	0	0	+25.5 -21.4	0
Jet Energy Scale (shape dep.)	0	+1.3 -2.3	0	+7.5 -0.1	+4.1 -4.4	+8.2 -7.8	+3.3 -5.5	0	+2.1 -2.7

Table 5.2: Systematic uncertainties on the contributions for the Loose Double Tag channel. Systematic uncertainties are listed by name. Systematic uncertainties for  $ZH$  shown in this table are obtained for  $m_H = 115$  GeV/ $c^2$ . Uncertainties are relative, in percent and are symmetric unless otherwise indicated.

Tight Double Tag  $ZH \rightarrow e^+e^-b\bar{b}$  Analysis CDF Run II Preliminary (7.5 fb<sup>-1</sup>)

Contribution	Fakes	Top	$WZ$	$ZZ$	$Z + b\bar{b}$	$Z + c\bar{c}$	$Z+1.f.$	$ZH$
Luminosity ( $\sigma_{\text{inel}}(p\bar{p})$ )	0	3.8	3.8	3.8	3.8	3.8	0	3.8
Luminosity Monitor	0	4.4	4.4	4.4	4.4	4.4	0	4.4
Trigger Emulation	0	1	1	1	1	1	0	1
Lepton ID	0	2	2	2	2	2	0	2
Lepton Energy Scale	0	3	3	3	3	3	0	3
$ZH$ Cross Section	0	0	0	0	0	0	0	5
Fake Leptons	50	0	0	0	0	0	0	0
B-Tag Efficiency	0	10.4	10.4	10.4	10.4	10.4	0	10.4
$t\bar{t}$ Cross Section	0	10	0	0	0	0	0	0
Diboson Cross Section	0	0	6	6	0	0	0	0
$\sigma(p\bar{p} \rightarrow Z + HF)$	0	0	0	0	40	40	0	0
ISR/FSR	0	0	0	0	0	0	0	4.0
Mistag Rate (shape dep.)	0	0	0	0	0	0	+29.3 -25.4	0
Jet Energy Scale (shape dep.)	0	+1.4 -2.6	+7.8 -3.1	+3.4 -5.9	+6.8 -6.6	+1.0 -3.7	0	+1.6 -2.7

Table 5.3: Systematic uncertainties on the contributions for the Tight Double Tag channel. Systematic uncertainties are listed by name. Systematic uncertainties for  $ZH$  shown in this table are obtained for  $m_H = 115$  GeV/ $c^2$ . Uncertainties are relative, in percent and are symmetric unless otherwise indicated.

# Chapter 6

## Final Analysis Discriminant

As in earlier analyses of this channel at CDF, the final analysis discriminant is a neural-network output. The analysis with  $1 \text{ fb}^{-1}$  of data developed a two-dimensional neural-network (two network outputs specify an  $x$  and  $y$  position) to simultaneously distinguish  $ZH$  from the dominant  $Z + jets$  background and the kinematically different  $t\bar{t}$  background [61, 93]. This method was fruitful and used in a number of following analyses [94, 62, 95, 55, 96, 31]. One of the shortcomings of the method, however, is the two-dimensional output required many bins in order to have limit-setting power. Many bins led to some being sparsely populated and this can lead to poorly set up marginalization integrals or specious discoveries if data end up in any bin with signal and no background. The discriminant applied here attempts to keep the isolation of the  $Z + jets$  background from the  $t\bar{t}$  background while avoiding the trap of sparse bins.

### 6.1 Super Bins

The final discriminant for this analysis is a one-dimensional output with three distinct regions designed to exploit different kinematic features of the backgrounds. This idea

came from the complimentary  $ZH \rightarrow \mu\mu b\bar{b}$  analysis [97]. The first region is designed to be populated with  $t\bar{t}$ -like events. This is accomplished by developing and applying a new neural-network designed to separate  $Z + jets$  from  $t\bar{t}$ . This is trained with  $t\bar{t}$  and  $Z + jets$  MC at the tag level (no overtraining is evident). Events with scores greater than 0.5 enter Region I. Events failing the  $t\bar{t}$  network are placed in one of the two remaining regions. These are intended to take in events with real  $Z$ 's. Events indicative of having light flavored jets are collected in Region II while events with heavy-flavor jets go to Region III. This partitioning is accomplished by applying a custom flavor separator based on the Karlsruhe Neural Network (KIT) [98, 99]. KIT was developed only for use on tight  $SECV_{TX}$ -tagged jets, so for application here, a pseudo-KIT, or “jet-flavor separator” value is defined as

- the maximum of the KIT value of the jets.
- one (1) if neither jet has a KIT score, but one has  $JETPROB < 0.01$ .
- minus one (-1) if neither jet has a KIT score, and neither has  $JETPROB < 0.01$ .

If the pseudo-KIT value is greater than zero, the event enters Region III, otherwise, it enters Region II. The logic leading to the super bin, or Region, placement is outlined in Figure 6.1.

## 6.2 Discriminant

The placement of each event within a region corresponds to a neural-network score. This network is trained with backgrounds representative of all of the backgrounds at tag-level and  $ZH$  signal at the mass point being considered. One can imagine instead optimizing a different network for each of the three regions, however, in this initial attempt due to available resources, the same network was evaluated in each of

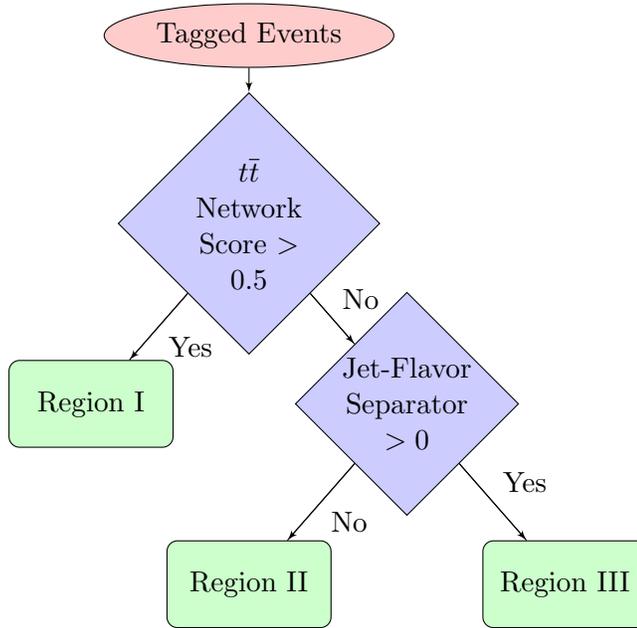


Figure 6.1: Flow chart of how the final discriminant region bins are set. Within each of the three regions, a given event is assigned to the bin corresponding to the value output by the final neural network discriminant.

the three regions. The network takes as inputs individual variables corresponding to event kinematics ( $Z p_T$ ,  $M_{jj}$ , *etc.*) and event shapes (sphericity, twist<sup>1</sup>, *etc.*) as well as two boosted decision tree (BDT) outputs (inputs are listed in Table 6.2). With the exception of the BDT values, the inputs were selected based on earlier networks where variables were selected iteratively and the lowest error combination was used (see Section 4.2.1.1). The BDTs were developed using the TMVA machinery [53] to process many distributions and exploit differences between  $ZH$  and background since including all of these in the neural network would reduce its power. One BDT used shape information and the other used energy information (see Table 6.1). The BDTs were trained using appropriately weighted MC backgrounds and a mixture of

<sup>1</sup> $\text{Twist}(x_1, x_2) = \tan^{-1}(\Delta\phi(x_1, x_2)/\Delta\eta(x_1, x_2))$  [100]. In other variables,  $\theta^*$  is the angle between the  $Z$  boson candidate and the proton beam direction in the zero momentum frame.

Higgs boson masses in the  $ZH$  signal (all mass hypotheses used the same BDTs). The BDT output distributions including the  $ZH$  120  $\text{GeV}/c^2$  mass point are shown in Figures 6.2a-6.2c and 6.3a-6.3c. These BDT outputs were the most powerful discriminating variables in the final network as ranked by TMVA's machinery ( $M_{jj}$ , the former leader, was ranked third), and their use improved the separation of signal and background by 5% [101].

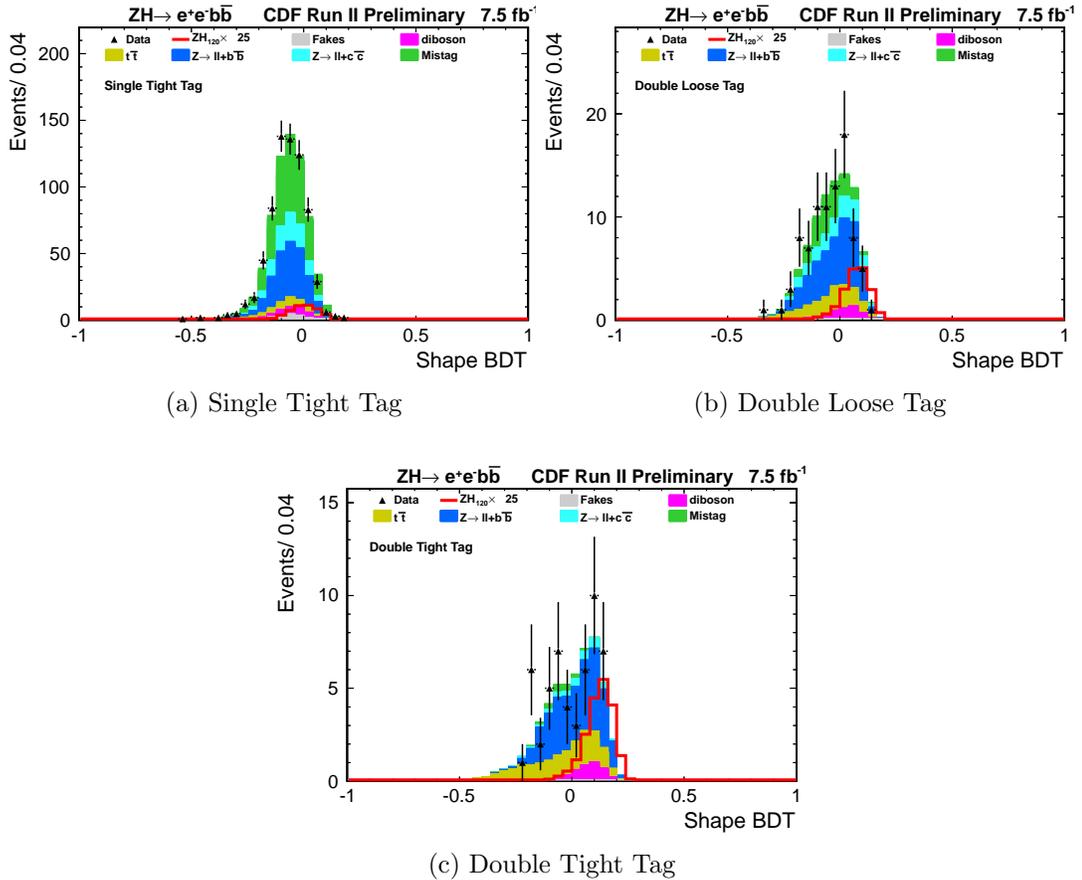


Figure 6.2: The shape BDT output distributions for each channel. This BDT takes in the variables listed in Table 6.1 and the returned value is used as an input in the final analysis discriminant.

Combining the region approach and the final neural network, the final discriminants on which limits are set are shown in Figures 6.4a-6.4c for each channel at the 120  $\text{GeV}/c^2$  Higgs boson mass point. All of the mass points between 100-150  $\text{GeV}/c^2$

Shape BDT	Energy BDT
$\Delta R(e_1, e_2)$	Dijet Mass
$\cancel{E}_T$ proj. onto vector $\Sigma(jets)$	$\cancel{E}_T$
$\Delta R(j_1, j_2)$	$\cancel{E}_T / \sqrt{(j_1 E_T + j_2 E_T)}$
$\Delta R(Z, H)$	$\cancel{E}_T / \sqrt{(\Sigma \text{ jet } E_T)}$
Aplanarity	$Z E_T + \text{Dijet } E_T$
Sphericity	Dijet $p_T$
$\Delta\eta(j_1, j_2)$	Mass( $e_1, j_1$ )
Twist( $e_1, e_2$ )	Mass( $e_2, j_2$ )
Twist( $j_1, j_2$ )	$Z p_T$
$\Delta\phi(j_1, j_2)$	Mass( $Z, jj$ )
$\Delta\theta(\cancel{E}_T, j_1)$ in $Z$ rest frame	Number of jets
$\Delta\theta(\cancel{E}_T, j_2)$ in $Z$ rest frame	$J_1 E_T$
$\Delta\theta(\cancel{E}_T, e_1)$ in $H$ rest frame	$J_2 E_T$
$\Delta\theta(\cancel{E}_T, e_2)$ in $H$ rest frame	$\cancel{E}_T + \text{el. } E_T\text{'s} + \text{jet } E_T\text{'s}$
$\cancel{E}_T$ projection onto jet 1	$\cancel{E}_T + \text{lepton } E_T\text{'s}$
$\cancel{E}_T$ projection onto jet 2	$\Delta E_T(j_1, j_2)$
$Z\eta$	$e_1 E_T$
$j_1\eta$	$e_2 E_T$
$j_2\eta$	
$\Delta R(j_1, Z)$	
$\Delta R(j_2, Z)$	
$\cos(\theta^*)$	
$\cos(\chi \xi = \pi/2)$	
$\cos(\theta_{jet_1})$ in $Z$ rest Frame	
$\cos(\theta_{jet_2})$ in $Z$ rest Frame	
$\cos(\theta_{e_1})$ in $H$ rest Frame	
$\cos(\theta_{e_2})$ in $H$ rest Frame	

Table 6.1: Distributions input to the BDT's.  $\theta$  is the angle between an object and the proton beam direction.  $\theta^*$  is the angle between the  $Z$  boson candidate and the proton beam direction in the zero momentum frame. The sum of the angles  $\chi$  and  $\xi$  is equal to the angle between the Higgs candidate and the lead  $p_T$  lepton in the  $Z$  boson rest frame [102] ( $\chi$  is the angle between the lead  $p_T$  electron and the  $Z$  spin vector;  $\xi$  is the angle between the Higgs boson and the  $Z$  spin vector and is chosen to be  $\pi/2$  by convention.).

---

Final NN Discriminant Inputs

---

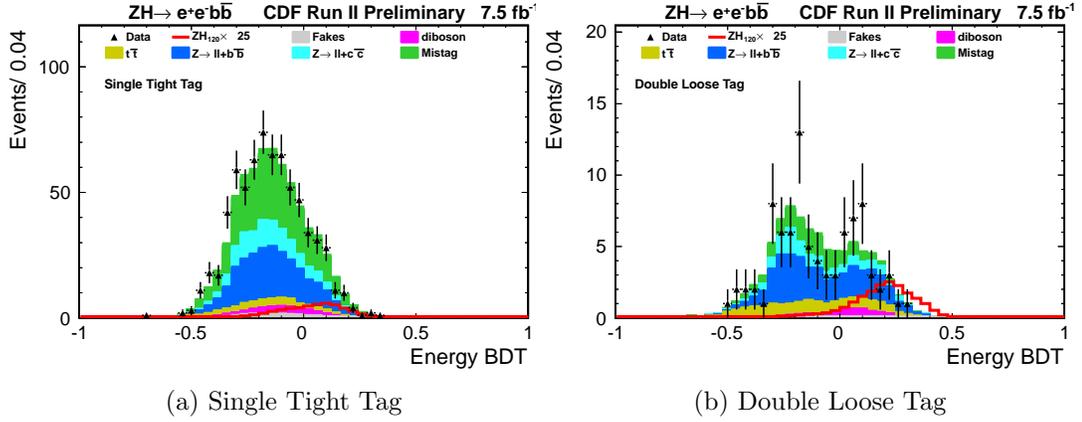
Cut Level  
Energy BDT  
Shape BDT  
 $\Delta R(e_1, e_2)$   
Twist  $e_1 e_2$   
Sphericity  
 $\Delta\phi(bb)$   
 $\text{Cos}(\theta^*)$   
 $\Delta R(\text{jet}_2, Z)$   
 $M_{jj}$   
 $\cancel{E}_T$   
 $E_T(Z) + E_T(H)$   
 $p_T(H)$   
 $p_T(Z)$

MET projection onto the jets

---

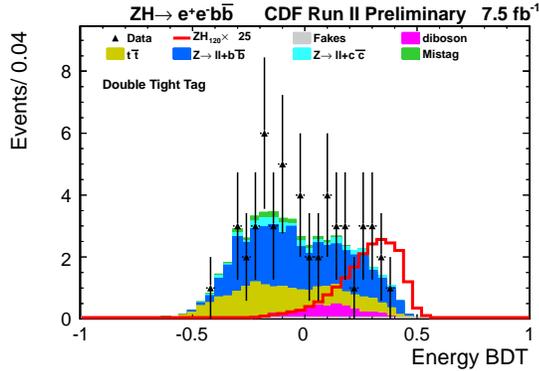
Table 6.2: Distributions input to the final discriminants. Cut Level is equal to -1 if pretag or single tight tag, 0 if loose double tag, and +1 if double tight tag.  $\theta^*$  is the angle between the  $Z$  boson candidate and the proton beam direction in the zero momentum frame.

in 5 GeV/ $c^2$  steps are in Appendix D. The input distributions at the pretag and tag levels are in Appendix C.



(a) Single Tight Tag

(b) Double Loose Tag



(c) Double Tight Tag

Figure 6.3: The energy BDT output distributions for each channel. This BDT takes in the variables listed in Table 6.1 and the returned value is used as an input in the final analysis discriminant.

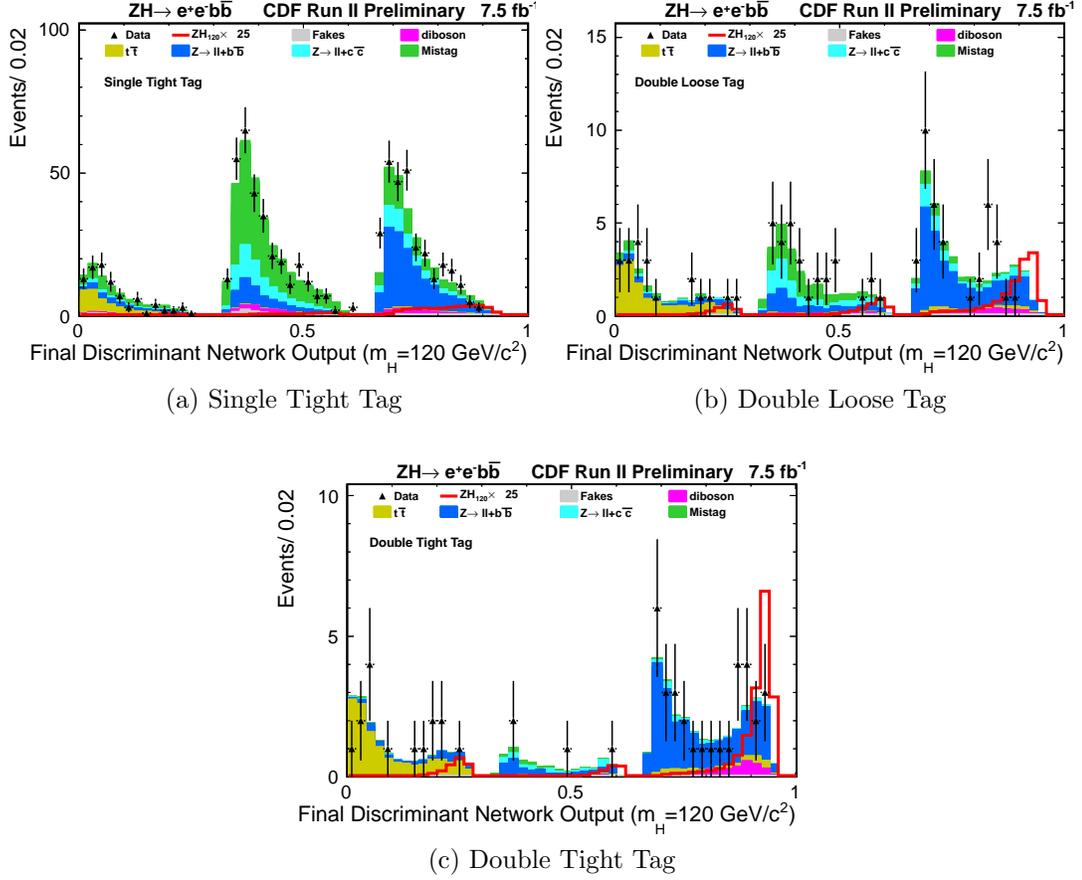


Figure 6.4: The final analysis discriminant at the  $120 \text{ GeV}/c^2$  mass point. The three regions correspond to  $t\bar{t}$ -like events on the left,  $Z$ +light-flavor dominated events in the middle, and event most like  $Z + b\bar{b}$  on the right third.

# Chapter 7

## Results

Since there is not an excess of likely Higgs events, we proceed to set upper limits on the production cross section times branching ratio. We use a Bayesian method [103] in the MCLIMIT machinery [104, 105] to set combined limits on the three tag category channels.

### 7.1 Limit Calculation

The Poisson probability of obtaining a given result with  $N_C$  channels and  $N_b$  bins is

$$\prod_{i=1}^{N_C} \prod_{j=1}^{N_b} \frac{e^{-(R \times s_{ij} + b_{ij})} (R \times s_{ij} + b_{ij})^{n_{ij}}}{n_{ij}!} \quad (7.1)$$

where  $n_{ij}$  is the number of observed events, and  $b_{ij}$  is the expected background in the  $ij$  bin. The quantity  $R \times s_{ij}$  is a scaling factor,  $R$ , times the expected signal in the  $ij$  bin. The value  $R$  sets the scale for how sensitive an analysis is to the signal. The product over the channels in this analysis goes over the three tag categories. The product over the bins goes over the bins in the final neural net output. To take into account the systematic uncertainties, we introduce a function  $\pi(\vec{\theta})$ , where  $\theta_k$

describes the  $k$ th nuisance parameter. The posterior density function is then

$$\mathcal{L}(R, \vec{s}, \vec{b} | \vec{n}, \vec{\theta}) \times \pi(\vec{\theta}) = \prod_{i=1}^{N_C} \prod_{j=1}^{N_b} \frac{\mu_{ij}^{n_{ij}} e^{-\mu_{ij}}}{n_{ij}!} \times \prod_{k=1}^{n_{np}} e^{-\theta_k^2/2} \quad (7.2)$$

Where  $\mu_{ij} = R \times s_{ij}(\vec{\theta}) + b_{ij}(\vec{\theta})$  with the expected signal and background contributions now including uncertainty dependence, and the last product is over nuisance parameters.

This function is integrated over the parameter space yielding a posterior density as a function of  $R$  only,  $P(R)$  (see Figure 7.1a). The integration is performed by sampling the parameter space according to the prior,  $\pi(\theta)$ . Here, truncated (to prevent negative values) Gaussian distributions are assumed and 2000 (1,000,000) samplings are made for the expected (observed) limits. This function is then integrated to find  $R_0$  such that

$$0.95 = \int_0^{R_0} dR P(R) \quad (7.3)$$

This  $R_0$  value is the result of this analysis. We place upper production cross section limits of  $R_0 \times \sigma_{ZH}$  at the 95% confidence level.

To quantify the expected sensitivity of the method, we use Monte Carlo techniques. The systematic uncertainties are varied according to their priors and pseudo-experiments (PEs) are drawn and Poisson fluctuated. The posterior density function is calculated for each PE and the resultant  $R_0$  values enter a distribution,  $s(R_0)$  (Figure 7.1b). The median value is taken as the expected value and one and two sigma bands are calculated based on 68% and 95% coverage respectively. For the results presented here 5000 PEs were performed.

This process is performed at 11  $m_H$  hypotheses in 5 GeV/ $c^2$  intervals. The resultant limits are listed in Table 7.1 and shown as a function of mass in Figure 7.2.

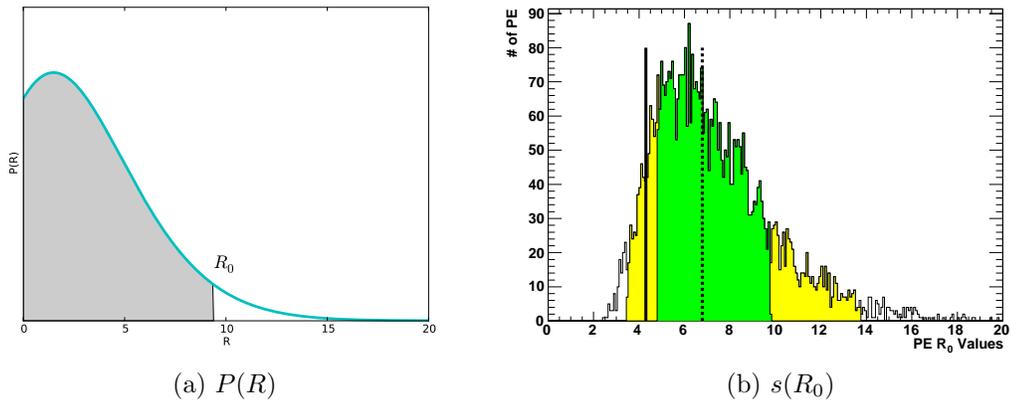


Figure 7.1: Figure 7.1a shows a sample  $P(R)$  distribution for a single PE. The  $R_0$  value is highlighted and would be a single entry into  $s(R_0)$ . Figure 7.1b is a sample of the distribution of  $R_0$  values,  $s(R_0)$ . The median value is the dashed vertical line, and the corresponding one and two sigma bands from PEs are highlighted in green and yellow. The observed value is the solid line.

The individual channel limits are shown in Appendix E.

$M_H$ (GeV/ $c^2$ )	Observed	$-2\sigma$	$-1\sigma$	Median	$+1\sigma$	$+2\sigma$
100	2.7	1.9	2.7	3.7	5.4	7.7
105	3.0	2.2	3.0	4.3	6.2	8.7
110	3.7	2.5	3.4	4.8	6.9	9.7
115	3.9	3.0	4.1	5.8	8.3	11.7
120	4.3	3.5	4.8	6.9	9.7	13.8
125	4.8	4.2	5.8	8.1	11.8	16.3
130	5.4	5.2	7.1	10.1	14.5	20.5
135	6.8	6.7	9.1	12.8	18.2	25.8
140	10.7	9.0	12.3	17.1	24.7	34.5
145	15.2	13.2	18.1	25.4	36.5	51.3
150	25.0	21.6	29.0	40.8	58.4	80.9

Table 7.1: The 95% CL upper limits on the  $ZH$  production rate expressed as a factor on  $\sigma_{ZH} \times BR(H \rightarrow b\bar{b})$ . The observed limits are obtained using CDF II data, while the median,  $-2, -1, +1$ , and  $+2$  are obtained from the distribution of upper limits obtained in 5000 background-only pseudo-experiments.

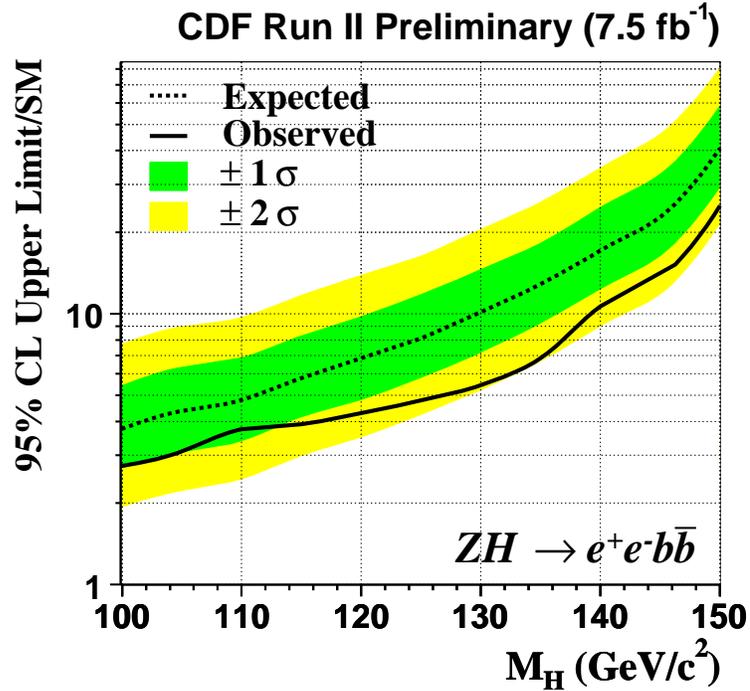
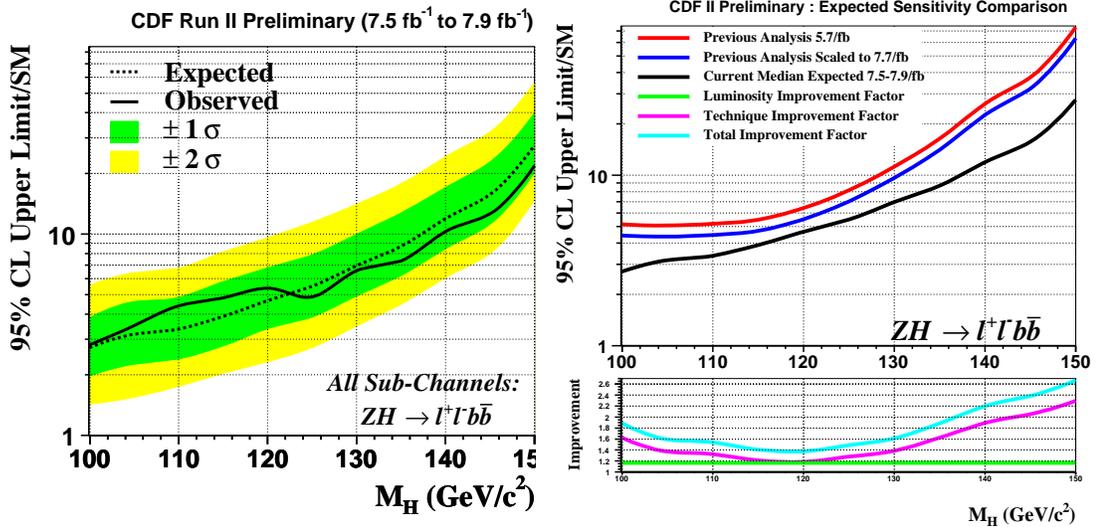


Figure 7.2: Upper limits on production cross section times branching ratio over SM expectation from the  $ZH \rightarrow e^+e^-b\bar{b}$  channel. The limit is calculated for 11 mass hypotheses and one (green) and two (yellow) standard deviation bands are drawn.

## 7.2 Combination with $ZH \rightarrow \mu^+ \mu^- b\bar{b}$

This analysis is combined with the  $ZH \rightarrow \mu^+ \mu^- b\bar{b}$  analysis [97] and the results are shown in Figure 7.3a. This combined result can be directly compared to the previous, and joint-flavor-only ( $e$  and  $\mu$ ), analysis with  $5.7 \text{ fb}^{-1}$  [31]. As a function of Higgs boson mass, the improvement in sensitivity is summarized in Figure 7.3b. At  $m_H=120 \text{ GeV}/c^2$ , the combined result improved in sensitivity from technique alone (magenta line) by  $\sim 20\%$ . Greater factors of improvement away from the 120  $\text{GeV}/c^2$  mass point can be attributed to training a final discriminant at each mass point instead of only at 120  $\text{GeV}/c^2$  as was previously done.



(a)  $ZH \rightarrow \ell^+ \ell^- b\bar{b}$  Limits

(b) Improvement in sensitivity.

Figure 7.3: Figure 7.3a shows the upper limits on production cross section times branching ratio over SM expectation from the  $ZH \rightarrow \ell^+ \ell^- b\bar{b}$  channel ( $\ell = e, \mu$ ). The limit is calculated for 11 mass hypotheses and one (green) and two (yellow) standard deviation bands are drawn. Figure 7.3b shows the improvement in analysis sensitivity across the mass range. The cyan curve in the lower portion represents the total improvement with respect to the  $5.7 \text{ fb}^{-1}$  analysis, the magenta line is due to technique alone, and the green line is the improvement due to the increase in data. Figure 7.3b was created by S.Z. Shalhout.

### 7.3 CDF and Tevatron Combinations

The  $ZH \rightarrow e^+e^-b\bar{b}$  channel is also included in a combined CDF standard model Higgs mass limit result [106] and a combined Tevatron result [107]. The contribution of each channel to the CDF result is shown in Figure 7.4a. The combined CDF result is shown in Figure 7.4b with the exclusion mass regions highlighted. The excluded regions are found by linearly interpolating between the 5 GeV/c<sup>2</sup> mass points and CDF alone excludes at the 95% confidence level  $156.5 < m_H < 173.7$  GeV/c<sup>2</sup>. CDF and D0's combined Tevatron result is shown in Figure 7.5 with exclusion regions of  $158 < m_H < 175$  and  $100 < m_H < 109$  GeV/c<sup>2</sup> found in the same manner.

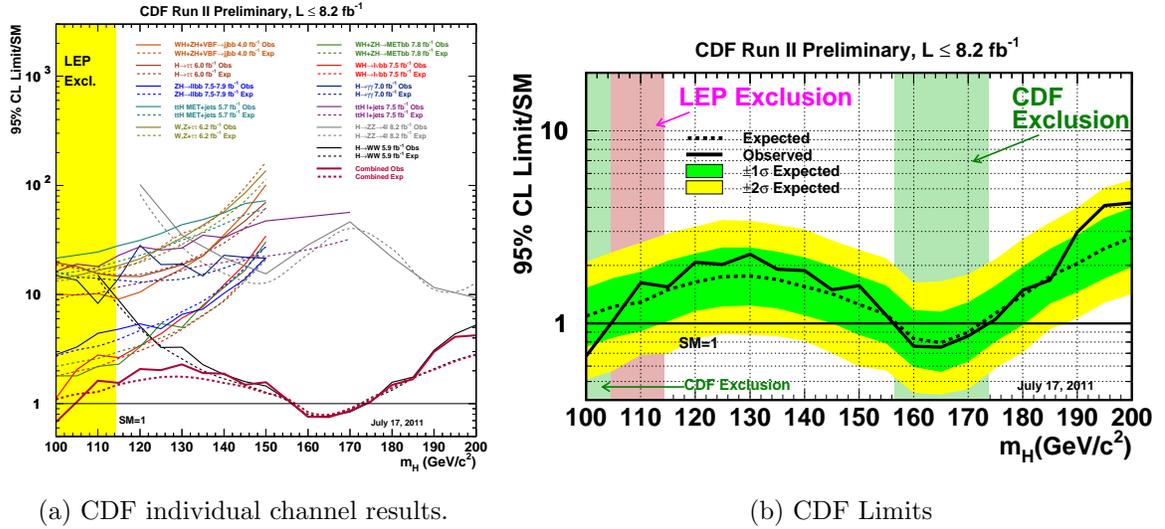


Figure 7.4: Plot 7.4a shows the CDF individual channel Higgs production limits times branching ratio over the SM prediction [106]. The solid lines are the observed limits; the dashed lines are the median expected limits. Figure 7.4b is the resulting limit of the combined CDF channels [106].

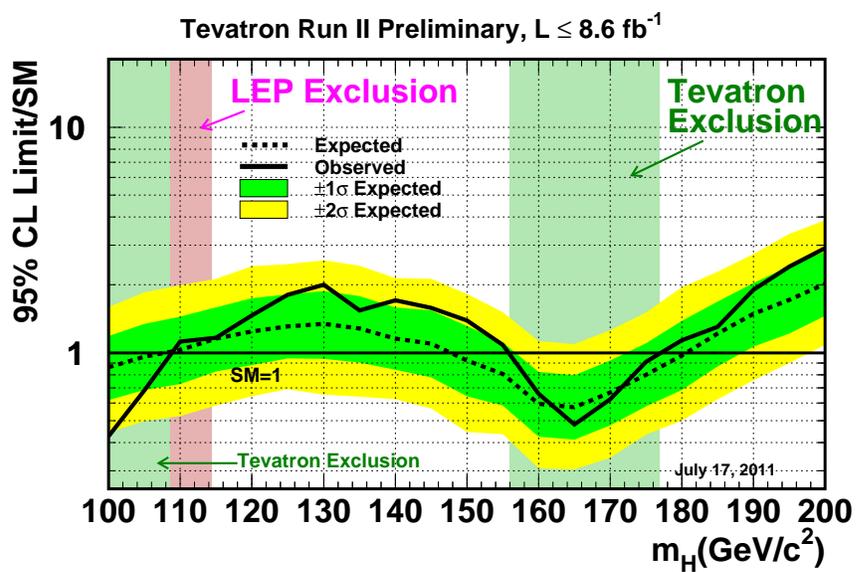


Figure 7.5: The combined Tevatron upper limits on Higgs boson production cross section times branching ratio over SM prediction. Image from [107]

# Chapter 8

## Conclusion

This analysis presented a search for the standard model Higgs boson in the associated production process  $p\bar{p} \rightarrow ZH \rightarrow e^+e^-b\bar{b}$  using  $7.5 \text{ fb}^{-1}$  of Tevatron data collected by the CDF II detector. Advanced methods were added to this analysis and applied alongside proven, sophisticated tools. New techniques that improved acceptance included consideration of additional triggered data streams, and a neural-network-based electron identification. The inclusion of the third data stream provided a  $\sim 5\%$  boost in data  $Z$  acceptance. The neural-network-based electron identification reduced the electron misidentification rate by a factor of five at the pretag level and provided modest gains in acceptance ( $\sim 5\%$  at the pretag level ignoring crack-track electrons). Improvements to the discriminant included using boosted decision tree (BDT) outputs, and dividing the output into three distinct regions. Additionally, discriminants were trained at each mass hypothesis instead of only the  $120 \text{ GeV}/c^2$  mass point.

It is difficult to determine the effect of individual improvements to the analysis because they are intertwined. For instance, BDTs were able to be added to the analysis since a large number of variables showed improvement in modeling. This

could be due to the electron identification, the trigger model, the Monte Carlo, or any combination of these. While it was not possible to isolate the improvement factor due to each technique alone, the complete analysis (including the  $\mu$  channel) saw a  $\sim 20\%$  improvement due to technique alone (ignoring the increase in data) at the 120  $\text{GeV}/c^2$  mass point. Due to training the discriminant at each mass hypothesis, limits at other masses improved considerably more ( $\sim 60.7\%$  due to technique at the 135  $\text{GeV}/c^2$ ).

Good agreement was observed between data and the null hypothesis, and upper production cross section ( $\sigma_{ZH}$ ) times branching ratio ( $H \rightarrow b\bar{b}$ ) were set for 11 mass hypotheses between 100-150  $\text{GeV}/c^2$  at the 95% confidence level. These results were included in the CDF and Tevatron combined Higgs boson searches. The Tevatron excludes the standard model Higgs boson between  $158 < m_H < 175$  and  $100 < m_H < 109$   $\text{GeV}/c^2$  at the 95% confidence level.

## 8.1 Notes on Future Work

Since the Tevatron shut down on September 30, 2011, there remains one last data set for analysis. With these data alone, there should be about a 10% improvement in limit sensitivity. This is insufficient to be sensitive at all considered mass hypotheses. There must be significant improvements in technique to make a strong statement on the Higgs boson between 100-200  $\text{GeV}/c^2$ . The following paragraphs suggest some possible improvements for this channel.

One way to improve the analysis sensitivity is by increasing acceptance. A natural extension of the multivariate regression trigger model is to include all triggered data or at least all triggered data in the electron stream. This was successfully carried out by the  $ZH \rightarrow \mu^+ \mu^- b\bar{b}$  analysis [97] and a 12% gain in data was seen. Additionally,

the electron selection has some capacity for model improvements. An over-aggressive requirement on track-only (crack track) electrons was imposed requiring that they not overlap with any other potential electron. In practice, it is sufficient that the crack track electron only avoid overlap with the other electron forming the  $Z$  boson. Changing this leads to a  $\sim 1\%$  gain in  $Z$  bosons. Also, the events with crack track electrons could have their  $\cancel{E}_T$  improved by correcting for the electron's reconstructed energy. Additionally, ultra loose forward electrons used in previous analyses [55] could be included. They are already in the framework, but there were concerns about their purity. Potentially the largest gains in acceptance can be made from improving the  $b$ -tag efficiency. Only about 60.7% of  $ZH$  events with a reconstructed  $Z$  make it into one of the final analysis ( $b$ -tagged) channels. At the moment, a new  $b$  tagger, HOBIT (Higgs Optimized  $b$ -Identification Tagger) is being implemented. This tagger was optimized for low mass Higgs boson searches and makes use of vertex information, extensive track quantities ([108]), as well as output from other algorithms. Preliminary results show a  $\sim 35\%$  gain in single  $b$ -tag efficiency for an operating point with a comparable mistag rate of tight SECVTX (efficiency scale factors are estimated to be comparable).

Further gains in sensitivity can be made by improving the final discriminant. Due to time constraints, matrix element outputs [109] were not included as inputs to the final discriminants. The power of the matrix element outputs, however, might be diluted since the BDT outputs could be extracting similar event information. An additional gain can be made by treating two and three jet events as separate channels. A preliminary test of this showed a 9% improvement in limits in this channel. The  $ZH \rightarrow \mu^+\mu^-b\bar{b}$  channel, however, saw more modest gains. Finally, the discriminant itself can be improved. The three discriminant regions could have individual networks that better exploit kinematic differences in backgrounds. Along

the same lines, it could also be fruitful to train the  $b$ -tag categories separately if they are sufficiently different in their input distributions. There is also the possibility for improvement in the training of the final discriminant. Evidence for this can be seen in the final discriminant output at the  $150 \text{ GeV}/c^2$  mass point (Figure D.11c) where there is a pernicious background increasing into the signal region. The cause of this has not been isolated, however, there is evidence that careful composition of the training sample might remedy this.

The Tevatron is getting more sensitive with each analysis to making a definitive statement on the existence of a standard model Higgs boson between  $100\text{-}200 \text{ GeV}/c^2$ . It will be exciting to see the results with the full data set. Even if small mass bands are not fully sensitive, the Large Hadron Collider (LHC) will likely be able to exclude or observe it in the coming years. At that point, the world will either know the mechanism of electroweak symmetry breaking or this particular theory will be unceremoniously swept into the dustbin of history.

# Appendix A

## Introduction to Neural Networks

The use of multivariate techniques, like neural networks, has become commonplace in particle physics. In neutrino experiments, the MiniBooNE collaboration has used boosted decision trees (BDTs) for particle identification [110]. Almost any modern Tevatron analysis is now touched by multivariate methods in some way (object identification, discriminants, trigger modeling). In the analysis in this dissertation, neural networks are used to identify electrons, as a jet-flavor separator, to distinguish  $t\bar{t}$ , to model trigger behavior, and as the final discriminant. The widespread use is a relatively new development. In this Appendix, a simple example is given illustrating some of the basic concepts. A well-done introduction can be found in [111].

Two types of neural networks often used in particle physics (and other fields) are classification and regression (or function approximation networks). The electron identification and final discriminant networks are examples of classification networks. The trigger modeling network is an example of a regression network. The following example is simple enough to illustrate basic ideas of both types.

Neural networks are a type of machine learning. Often, introductory literature will show a photo of a brain or synapse presumably in an effort to illustrate the par-

allels between how a human brain can assess multiple inputs to make a decision and how neural networks can take in variables and provide an output. In this example, a network is developed to solve an XOR problem (see Figure A.2a). The architecture is illustrated in Figure A.1. There are three inputs:  $x$  values,  $y$  values, and a constant term. This constant term allows for a degree of freedom in the resulting neural-network function. It is analogous to allowing for a non-zero  $y$ -intercept in a linear fit.

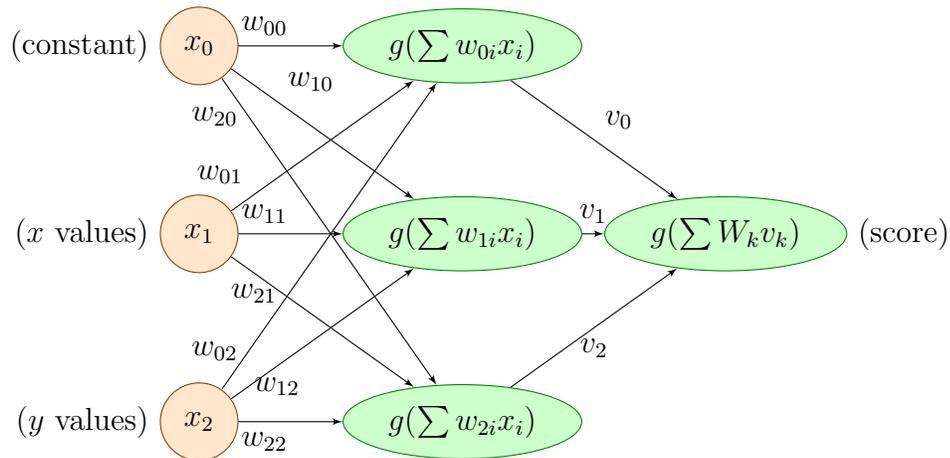


Figure A.1: The neural network architecture used in the XOR example.

In this example, there is one hidden layer with three nodes as signified by the middle column of ovals in Figure A.1. Each of the inputs is fed into each of the hidden nodes with an individual corresponding weight. Here, there are three inputs and three hidden nodes resulting in nine weights for this layer. At each node, the sum of the weights times inputs ( $\sum_{i=0}^2 x_{vi} w_{hi}$ ) is fed into a turn on function. Here a sigmoid function is used  $g(x) = 1/(1 + e^{-2x})$ . These values are multiplied by a weight,  $W_k$ , and then fed into the output node. The sum of the quantities is again fed into a sigmoid and the result is the score.

So now it is clear how inputs are fed into a network (a function) and a return

value (score) is given. Neural networks develop their power by adjusting the weights within this function to reduce the errors. This is done by training. First, a training sample is produced. This is a set of data with the desired result known. In the code snippet, the training set is the tenth of the `values` vector with the `test` vector set. With this training sample, scores,  $s_v$ , are found for each set of values,  $x_{vi}$  according to

$$s_v = g \left[ \sum_{h=0}^2 W_h g \left( \sum_{i=0}^2 x_{vi} w_{hi} \right) \right] \quad (\text{A.1})$$

A total error over the training set is then defined using the difference from the target value,  $t_v$ , as

$$E = \frac{1}{2} \sum_{v=1}^{N_{points}} (t_v - s_v)^2 = \frac{1}{2} \sum_{v=1}^N \left( t_v - g \left[ \sum_{h=0}^2 W_h g \left( \sum_{i=0}^2 x_{vi} w_{hi} \right) \right] \right)^2 \quad (\text{A.2})$$

The weights within the neural network can then be adjusted in an attempt to minimize this error. Here, the gradient descent method is used, however, there exist other methods. The change in weights,  $W_i$  is then

$$\Delta W_i = -\alpha \frac{\partial E}{\partial W_i} = -\alpha \frac{\partial E}{\partial s_v} \frac{\partial s_v}{\partial W_i} \quad (\text{A.3})$$

where  $\alpha$  is a dampening factor to prevent the weights from fluctuating wildly. The desired change in weight can then be found from

$$\frac{\partial E}{\partial W_h} = \sum_{v=1}^N (t_v - s_v) \frac{\partial s_v}{\partial W_i} = \sum_{v=1}^N (t_v - s_v) s_v (1 - s_v) g \left( \sum_{i=0}^2 x_{vi} w_{hi} \right) \quad (\text{A.4})$$

$$\begin{aligned}
\frac{\partial E}{\partial w_{hi}} &= \sum_v (t_v - s_v) \frac{\partial s_v}{\partial w_{hi}} \\
&= \sum_v (t_v - s_v) s_v (1 - s_v) W_{hk} g \left( \sum_k x_{vi} w_{hk} \right) \left( 1 - g \left( \sum_k x_{vi} w_{hk} \right) \right) x_{vi}
\end{aligned} \tag{A.5}$$

With the change in weights defined, the algorithm can be run in a loop (each attempt with the training set is called an epoch) with the error (hopefully) reducing with each iteration.

The code snippet below (Listing A.1) is commented in an attempt to clarify the toy algorithm. The number of epochs is set to 50,000 and an additional dampening factor, beta, is applied to reduce the change in weights as the epochs increase. Also, the subset of values used for training is changed every 4,000 epochs to avoid overtraining. This is inelegant but functional. After the training, the network is applied to all of the values. A text file is created with  $x_v, y_v, s_v$ , and python code is listed (Listing A.2) to plot the values. Figure A.2 shows the XOR target and the results. Points with a score above 0.5 received a green circle, and those below have a white circle. There are noticeable inefficiencies that would be cleaned up by a refined technique. Listing A.3 gives the commands to be entered on the command line to run the code.

```

#include <cmath>
#include <vector>
#include <stdio.h>
#include <stdlib.h>
#include <time.h>
#include <iostream>
#include <fstream>
#include <algorithm>

using namespace std;

float sigmoid( float x );

int main() {

    srand ( time(NULL) );

```

```

vector< vector<float> > values;
vector< int> target, test;
vector< float> scores;

vector< vector<float> > weights, dw, erwt;
vector< float > Weights, dW, erWt;

float alpha = 80.0;
int nHidden = 3;

ofstream results;
results.open("output.txt");

// Create values
for(int i=0; i<2000; i++) {
    float xV = rand()/float(RAND_MAX);
    float yV = rand()/float(RAND_MAX);
    vector<float> p;
    p.push_back(1);
    p.push_back(xV);
    p.push_back(yV);
    values.push_back(p);

    if(i<2000/10) test.push_back(1);
    else test.push_back(0);

    if( (xV>0.5 && yV>0.5) || (xV<0.5 && yV<0.5) ) target.push_back(1);
    else target.push_back(0);

    scores.push_back(0);
}

// Initialize weights
for(int i =0; i < 3; i++ ) {
    vector<float> p;
    for(int h=0; h<nHidden; h++)
        p.push_back(rand()/(20*float(RAND_MAX))-0.5/20.);

    weights.push_back(p);

    p.clear();
    for(int h=0; h<nHidden; h++) p.push_back(0);
    dw.push_back(p);
    erwt.push_back(p);
}

// Filling weights for hidden nodes
for(int i=0; i<nHidden; i++) {
    Weights.push_back(rand()/(20*float(RAND_MAX))-0.5/20.);
    dW.push_back(0);
    erWt.push_back(0);
}

float error = 10;
int nEpochs = 0, maxEpochs = 50000;

for( nEpochs = 0; nEpochs < maxEpochs; nEpochs++ ) {
    error = 0;

```

```

if( nEpochs%4000 == 0 ) random_shuffle(test.begin(), test.end() );

for(unsigned int w=0; w<weights.size(); w++ ) {
    for(unsigned int h=0; h<weights[0].size(); h++ ) {
        erwt[w][h]=0.0;
        weights[w][h]=weights[w][h]-dw[w][h];
    }
}

for (int w=0; w<nHidden; w++) {
    erWt[w]=0.0;
    Weights[w]=Weights[w]-dW[w];
}

for(unsigned int i=0; i<values.size(); i++) {

    float sum =0;
    if(test[i]!=1) continue;

    for(int h=0; h<nHidden; h++ ) {
        float temp_sum=0.0;
        for(unsigned int w=0; w<weights.size(); w++) {
            temp_sum += weights[w][h]*values[i][w];
        }
        sum += Weights[h]*sigmoid( temp_sum );
    }

    scores[i] = sigmoid( sum );
    error += pow(target[i]-scores[i],2);

    // Error on layer 0
    for(unsigned int w=0; w<weights.size(); w++ ) {
        for(unsigned int h=0; h<weights[0].size(); h++ ) {
            float temp_sum=0.0;
            for(unsigned int ww=0; ww<weights.size(); ww++) {
                temp_sum += weights[ww][h]*values[i][ww];
            }

            erwt[w][h] += -(target[i]-scores[i])*scores[i]*(1-scores[i])*
                (Weights[h])*sigmoid(temp_sum)*(1-sigmoid(temp_sum))*values[i][w];
        }
    }

    // Error on layer 1
    for(int h=0;h<nHidden;h++) {
        float temp_sum=0.0;
        for(unsigned int w=0; w<weights.size(); w++) {
            temp_sum += weights[w][h]*values[i][w];
        }
        erWt[h] += -(target[i]-scores[i])*scores[i]*(1-scores[i])*sigmoid(temp_sum);
    }
} // values

float beta = (maxEpochs-nEpochs+6000)*1.0/maxEpochs;

for(unsigned int w=0; w<weights.size(); w++ ) {
    dW[w] = alpha*beta*erWt[w]/values.size();
    for(unsigned int h=0; h<weights[0].size(); h++ ) {
        dw[w][h] = alpha *beta* erwt[w][h]/values.size();
    }
}

```

```

    if(nEpochs%100==0) cout << nEpochs << " | " << error << endl;
    error = error/values.size();
    nEpochs++;
}

char buffer[50];

for(unsigned int i=0; i<values.size(); i++) {
    float sum =0;

    for(int h=0; h<nHidden; h++ ) {
        float temp_sum=0.0;
        for(unsigned int w=0; w<weights.size(); w++) {
            temp_sum += weights[w][h]*values[i][w];
        }
        sum += Weights[h]*sigmoid( temp_sum );
    }

    sprintf(buffer,"%5.3f %5.3f %5.3f\n", values[i][1], values[i][2], sigmoid(sum) );
    results << buffer;
}

results.close();

return 0;
}

float sigmoid(float s) {
    return 1.0/(1+exp(-2*s));
}

```

Listing A.1: Neural test code in C++ (test.cxx).

```

#!/usr/bin/env python

import numpy as np
from numpy import ma
import pylab as pl
import sys

values = np.loadtxt('output.txt')
x=values[:,0]
y=values[:,1]
score = values[:,2]
xhi=x[score>=0.5]
yhi=y[score>=0.5]
xlo=x[score<0.5]
ylo=y[score<0.5]

fig = pl.figure()
ax = fig.add_subplot(111)
ax.scatter(xhi,yhi,s=20, marker='o',c='g')
ax.scatter(xlo,ylo,s=20, marker='o',c='w')
fig.savefig('nnOutput.pdf')
pl.show()
sys.exit()

```

---

Listing A.2: Output plotting code in python (plot.py).

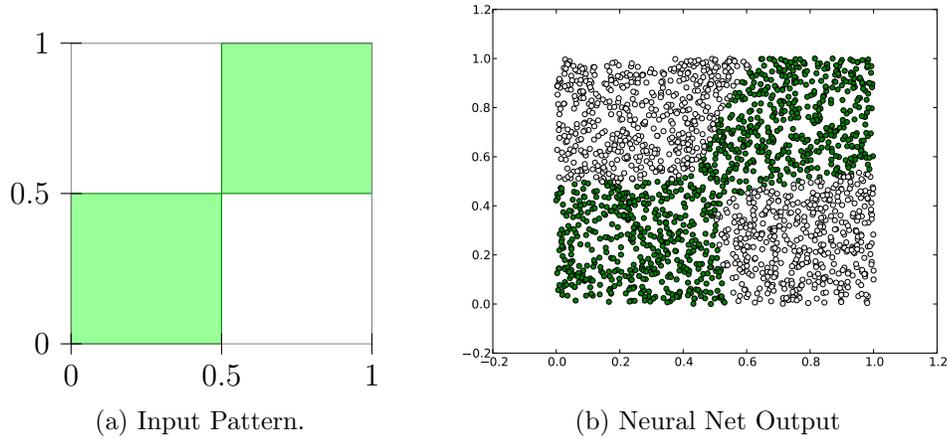


Figure A.2: A test neural network for an XOR pattern. The left is the test pattern. The right is the result of the developed neural network. Points with a neural network score  $\geq 0.5$  are colored green while points with a score  $< 0.5$  are white.

```
g++ test.cxx -o testNN.exe -Wall -ansi
./testNN.exe
./plot.py
```

Listing A.3: Commands to be entered to run the test code and produce the scatter plot.

# Appendix B

## Pretag Distributions

Included here are the distributions of final discriminant input variables and the neural network output discriminants at the pretag selection level.

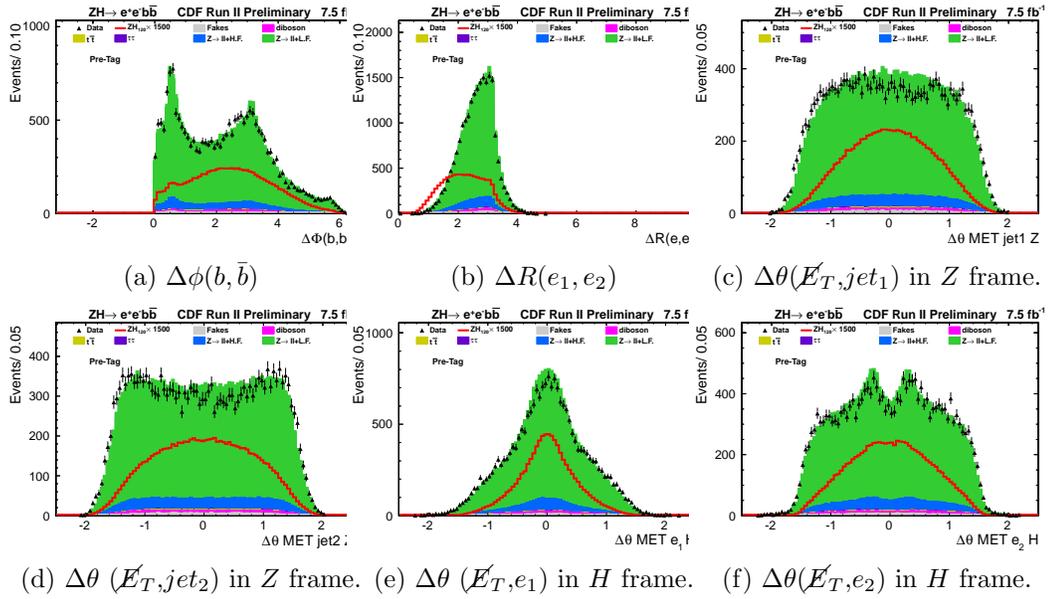
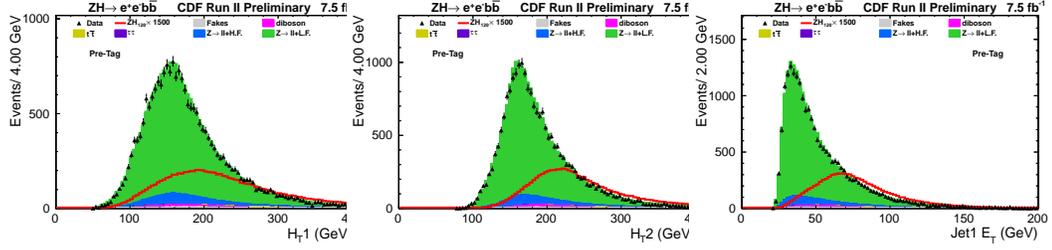


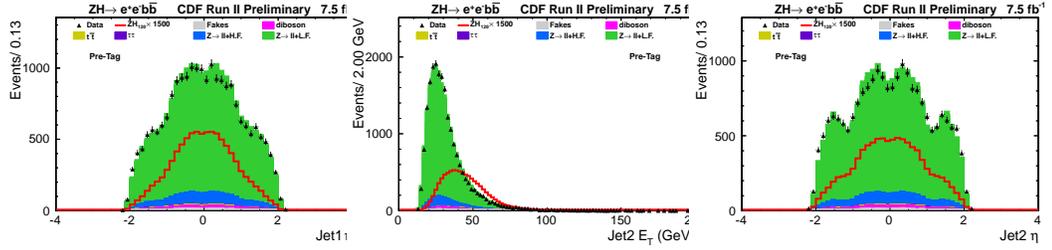
Figure B.1: Model plots at pretag.



(a)  $H_{T1}$ ,  $\cancel{E}_T$  + leading electron  $E_T$  + second leading electron  $E_T$  +  $E_T$  of the sum of all jets vector.

(b)  $H_{T2}$ ,  $\cancel{E}_T$  + leading electron  $E_T$  + second leading electron  $E_T$  +  $E_T$  of the sum of all jets vector.

(c) Leading jet  $E_T$

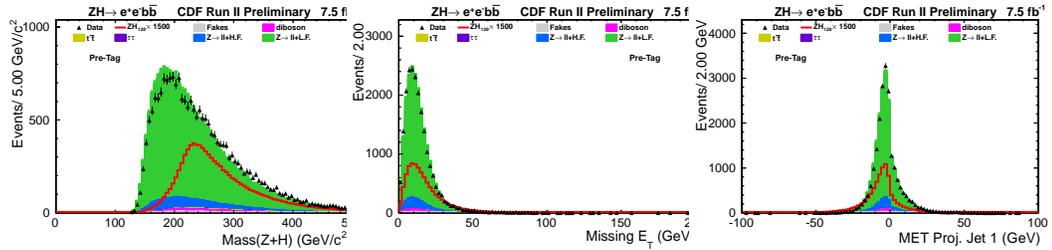


(d) Leading jet  $\eta$

(e) Second jet  $E_T$

(f) Second jet  $\eta$

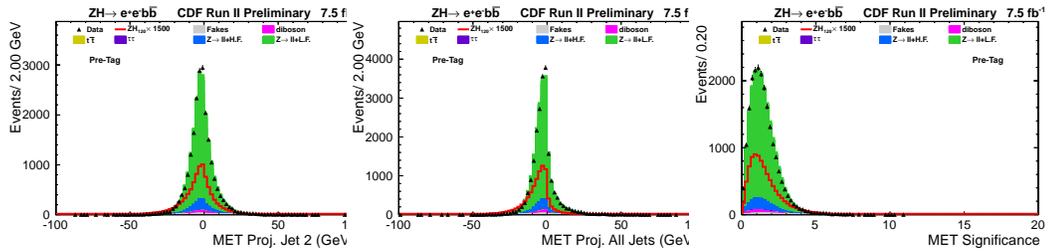
Figure B.2: Model plots at pretag.



(a) Mass ( $Z + H$ )

(b)  $\cancel{E}_T$

(c)  $\cancel{E}_T$  projected onto the leading jet.



(d)  $\cancel{E}_T$  projected onto the second and leading jet.

(e)  $\cancel{E}_T$  projected onto the sum of all jets.

(f)  $\cancel{E}_T$  significance,  $\cancel{E}_T / (j_1 E_T + j_2 E_T)$

Figure B.3: Model plots at pretag.

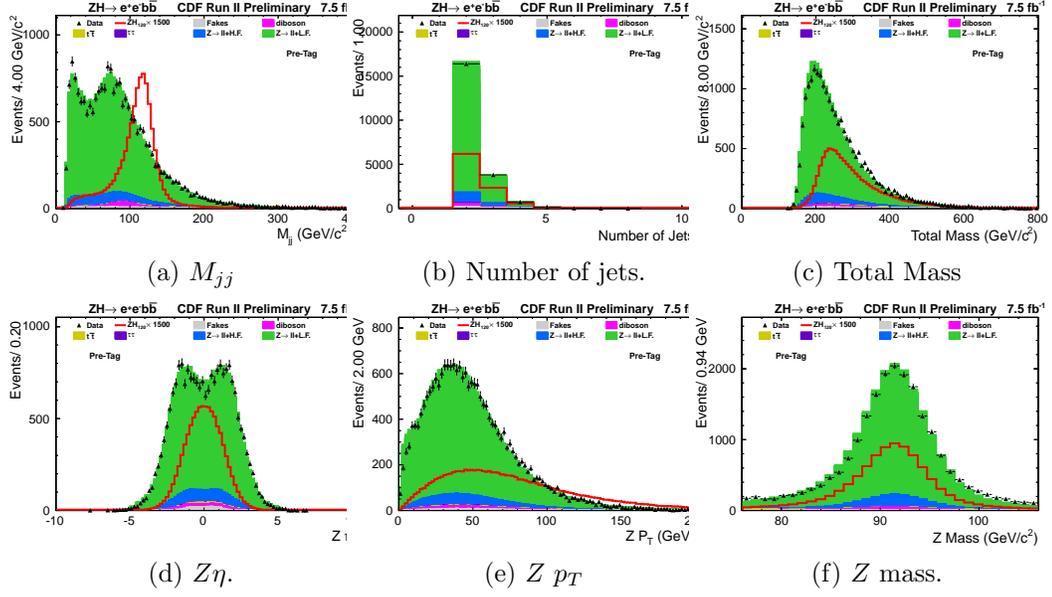


Figure B.4: Model plots at pretag.

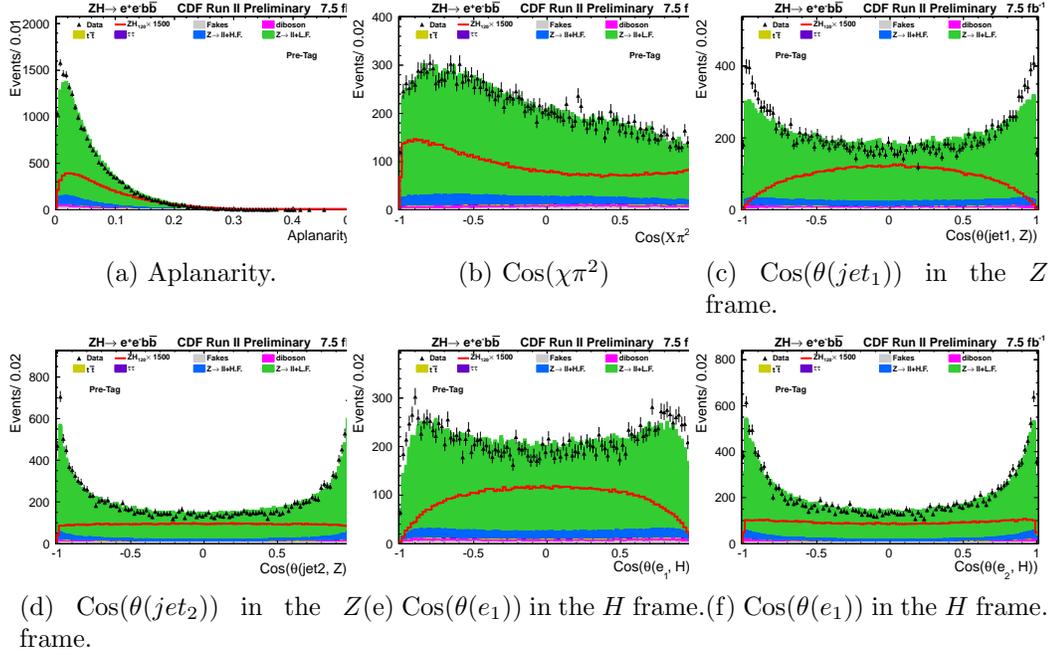


Figure B.5: Model plots at pretag.

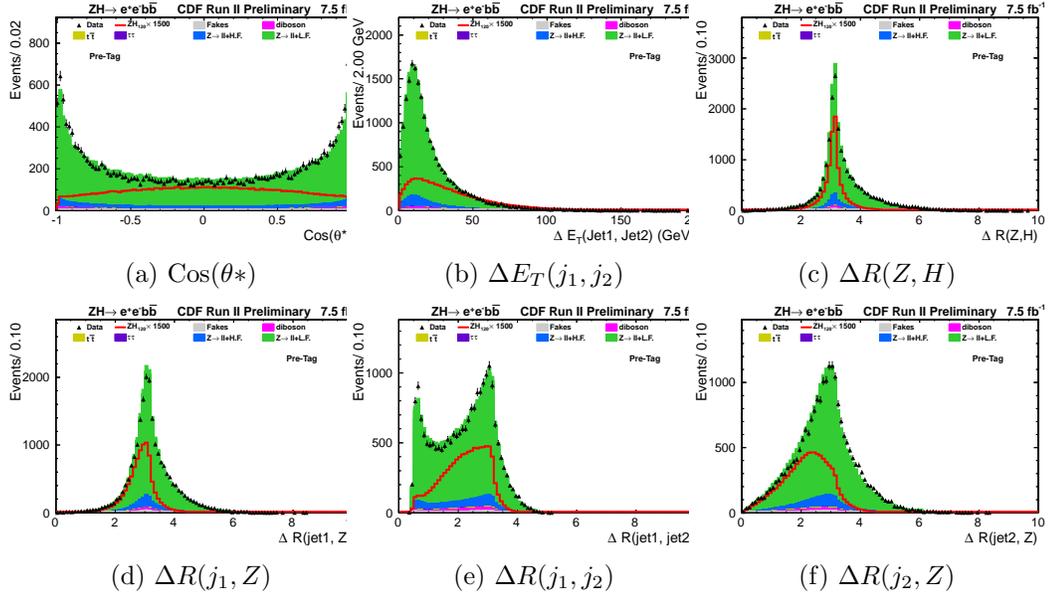


Figure B.6: Model plots at pretag.

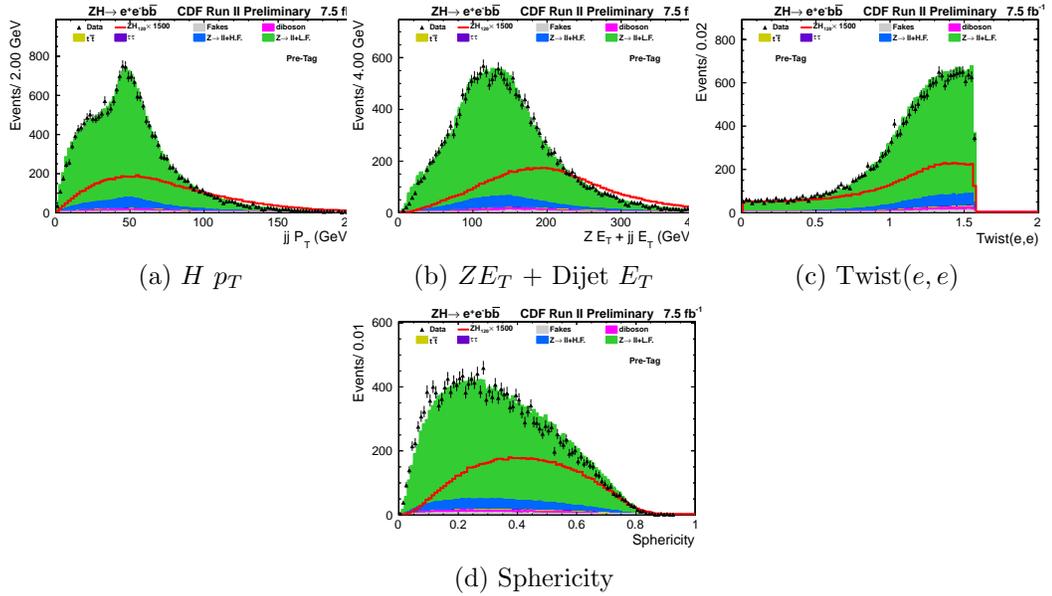


Figure B.7: Model plots at pretag.

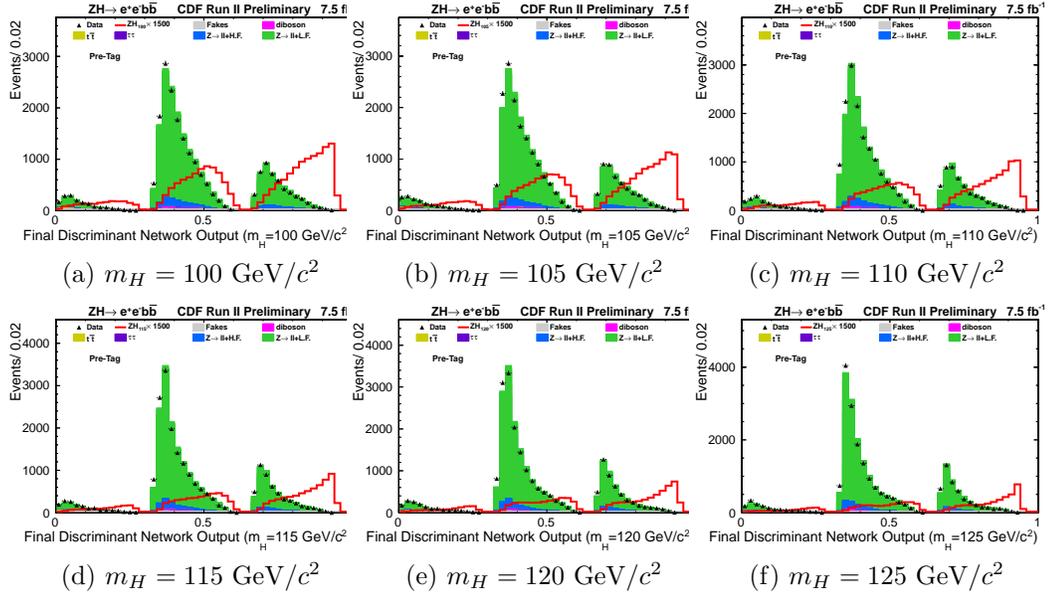


Figure B.8: Final discriminants for events passing the pretag selection.

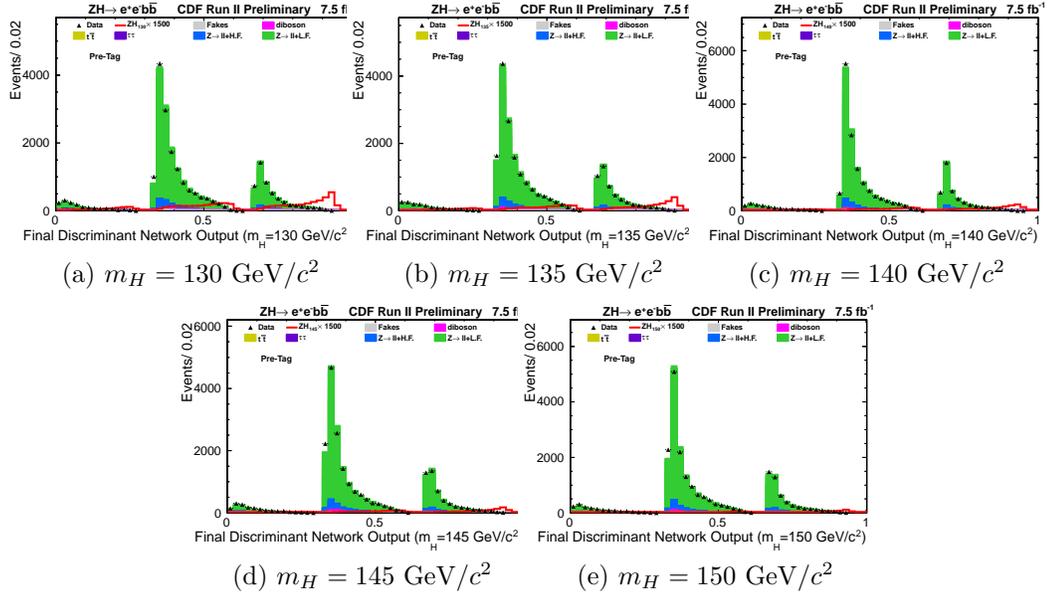


Figure B.9: Final discriminants for events passing the pretag selection.

# Appendix C

## Final Analysis Channel

### Distributions

Included here are the distributions of final discriminant input variables for the three analysis channels.

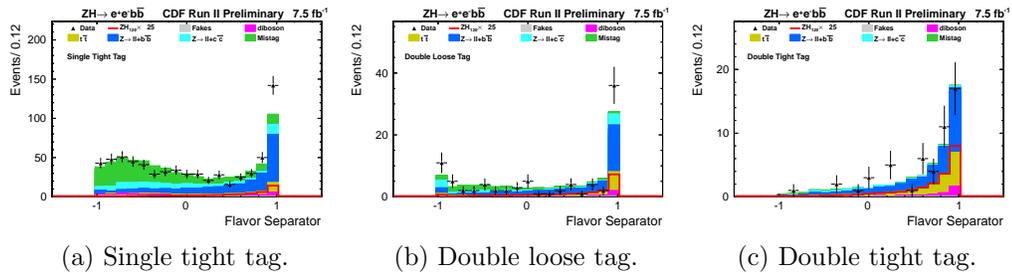
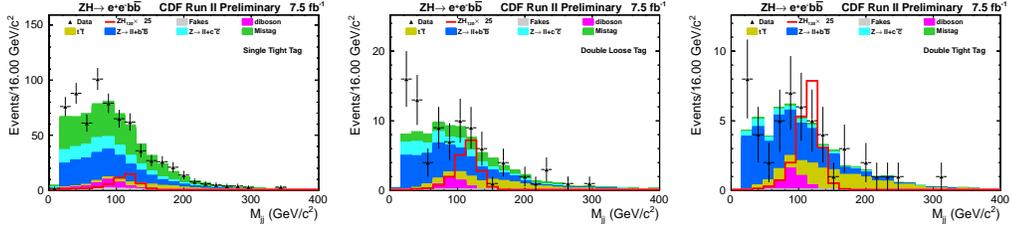


Figure C.1: PKIT (flavor separator).

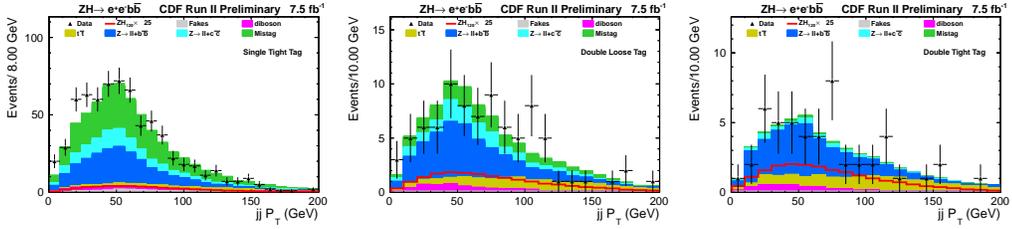


(a) Single tight tag.

(b) Double loose tag.

(c) Double tight tag.

Figure C.2:  $M_{jj}$

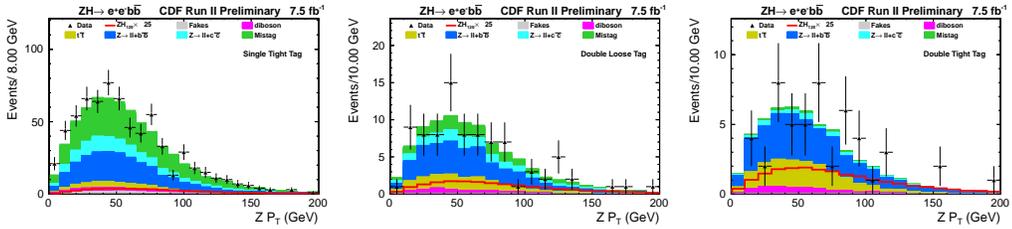


(a) Single tight tag.

(b) Double loose tag.

(c) Double tight tag.

Figure C.3:  $H p_T$



(a) Single tight tag.

(b) Double loose tag.

(c) Double tight tag.

Figure C.4:  $Z p_T$

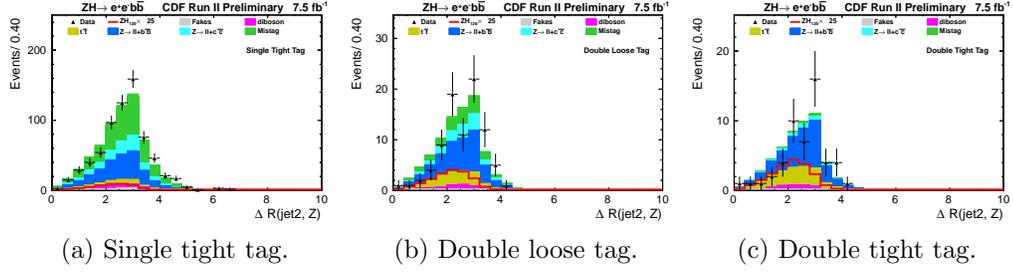


Figure C.5:  $\Delta R(j_2, Z)$

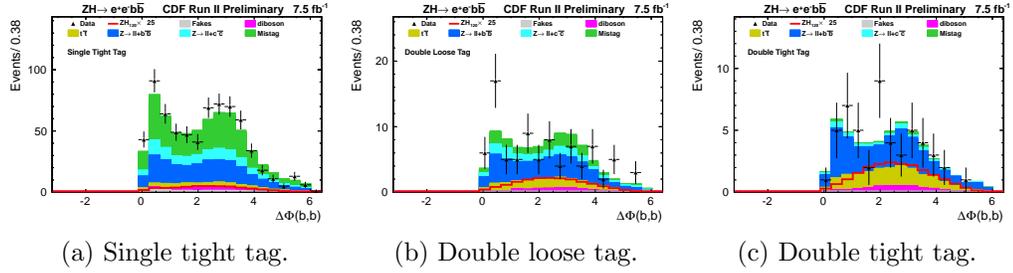


Figure C.6:  $\Delta\phi(b, \bar{b})$

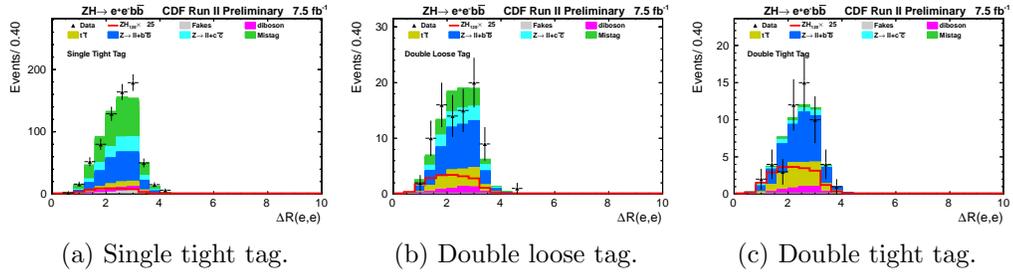


Figure C.7:  $\Delta R(e_1, e_2)$



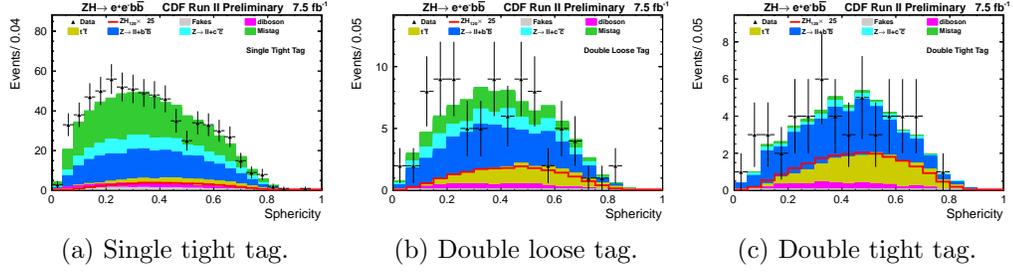


Figure C.11: Sphericity.

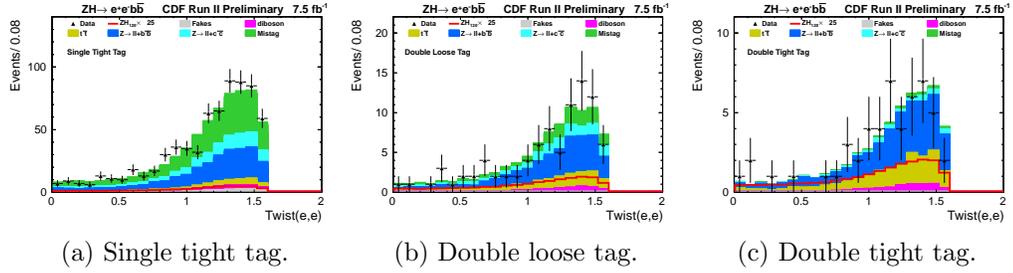


Figure C.12:  $\text{Twist}(e_1, e_2)$

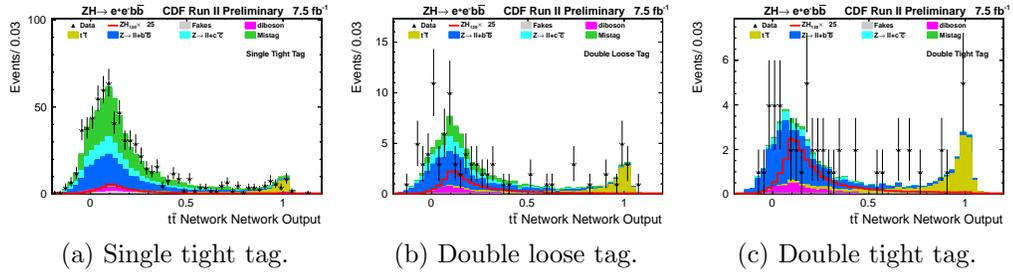


Figure C.13:  $t\bar{t}$  network output.

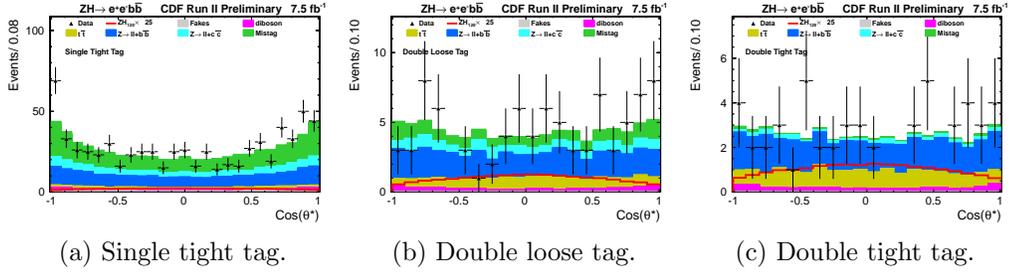


Figure C.14:  $\text{Cos}(\theta^*)$ .

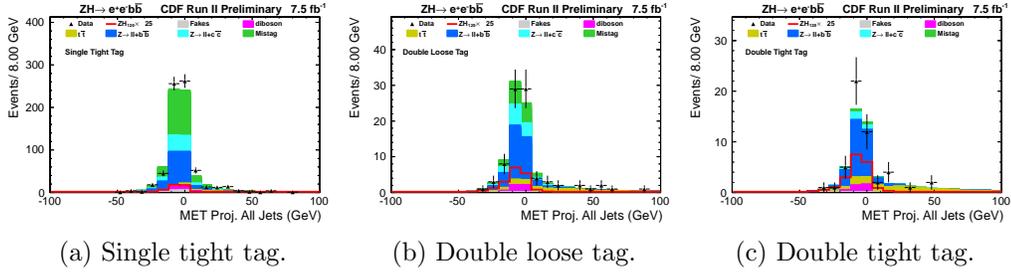


Figure C.15:  $\cancel{E}_T$  projected on vector sum of all jets.

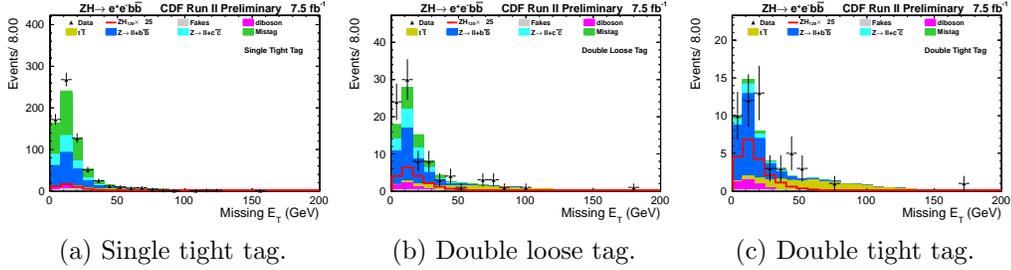


Figure C.16:  $\cancel{E}_T$

# Appendix D

## Final Analysis Discriminants

Figures D.1-D.11 show the final analysis discriminants for the three tag categories. The limits are calculated with these (and the systematic) distributions for each mass point. Each mass point has a separately trained network; there is one for each tag category.

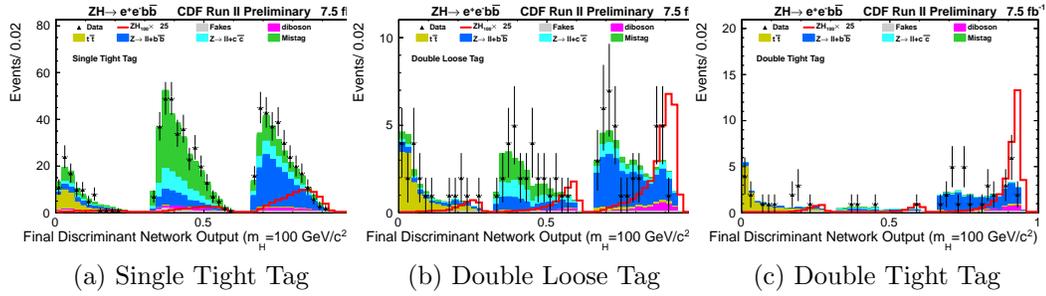


Figure D.1: Final discriminant for the  $m_H = 100 \text{ GeV}/c^2$  mass point.

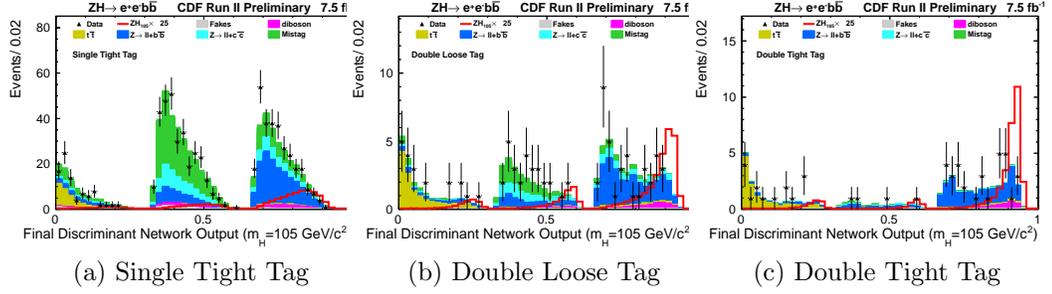


Figure D.2: Final discriminant for the  $m_H=105 \text{ GeV}/c^2$  mass point.

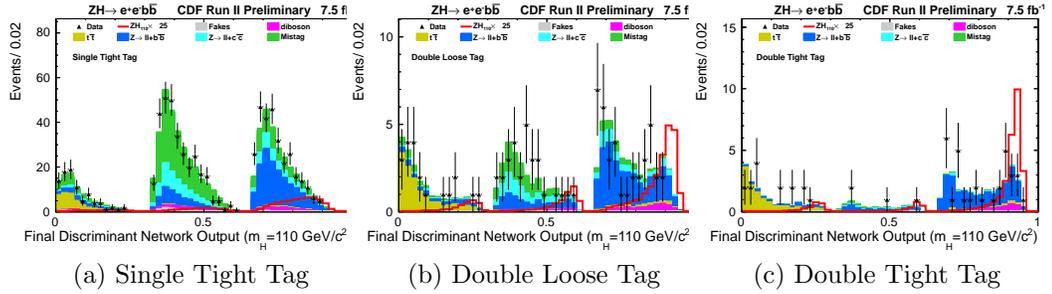


Figure D.3: Final discriminant for the  $m_H=110 \text{ GeV}/c^2$  mass point.

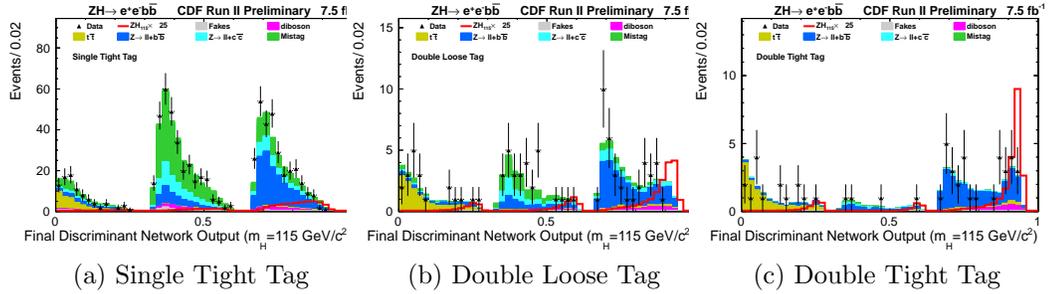


Figure D.4: Final discriminant for the  $m_H=115 \text{ GeV}/c^2$  mass point.

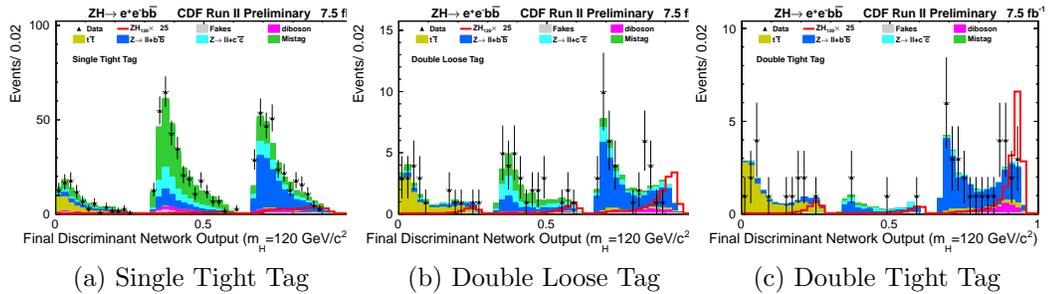


Figure D.5: Final discriminant for the  $m_H=120 \text{ GeV}/c^2$  mass point.

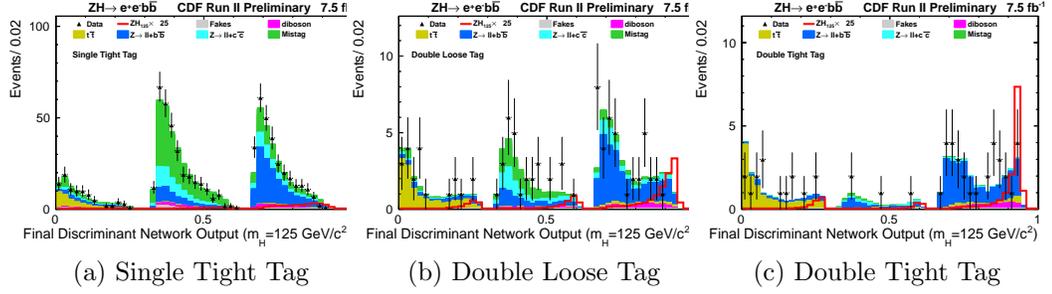


Figure D.6: Final discriminant for the  $m_H=125$   $\text{GeV}/c^2$  mass point.

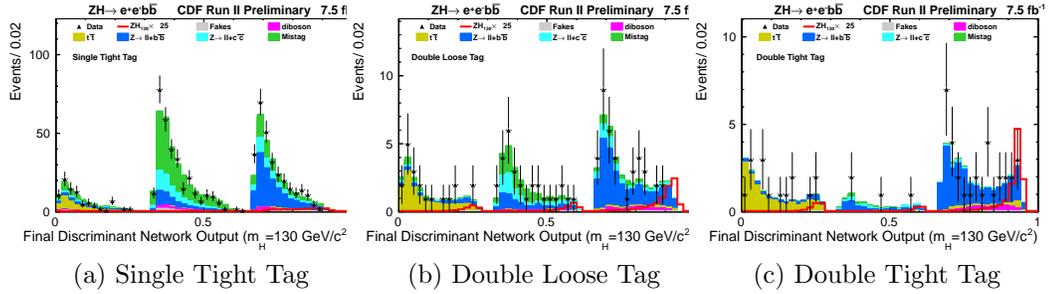


Figure D.7: Final discriminant for the  $m_H=130$   $\text{GeV}/c^2$  mass point.

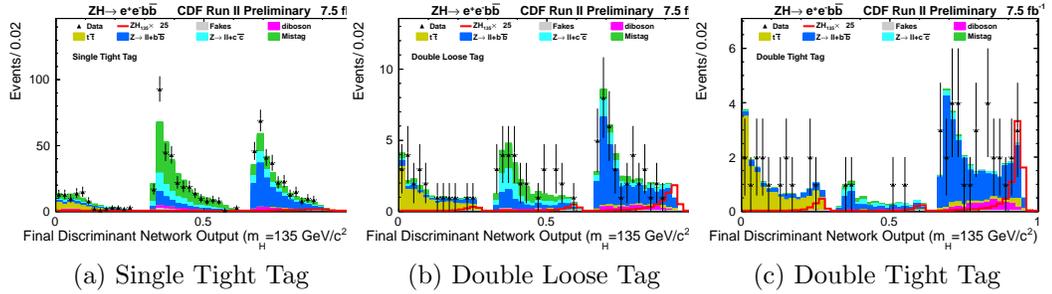


Figure D.8: Final discriminant for the  $m_H=135$   $\text{GeV}/c^2$  mass point.

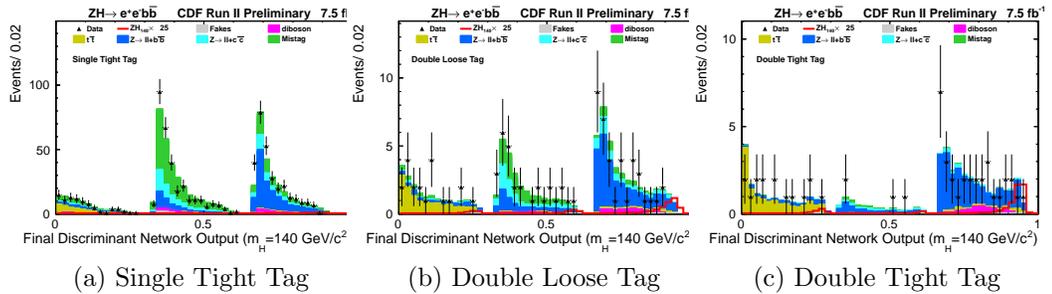


Figure D.9: Final discriminant for the  $m_H=140$   $\text{GeV}/c^2$  mass point.

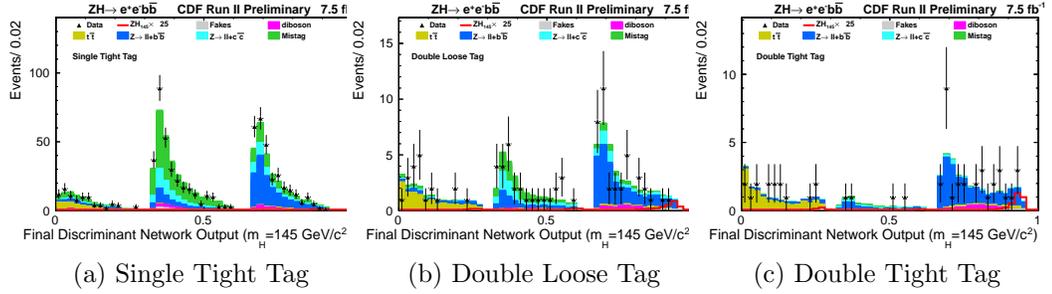


Figure D.10: Final discriminant for the  $m_H=145 \text{ GeV}/c^2$  mass point.

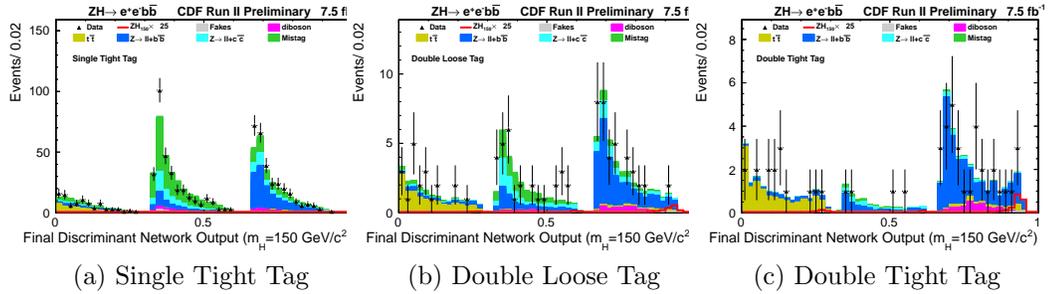


Figure D.11: Final discriminant for the  $m_H=150 \text{ GeV}/c^2$  mass point.

# Appendix E

## Individual Channel Limits

The final limits in this analysis (Section 7.1) are found by considering the three tag category channels and combining them. Here, the limits for the individual channels are shown in Tables E.1-E.3 and Figure E.1. There is a small observed excess in the double tight tag channel in the lower mass points and a slight deficit in the higher mass points of the single tight tag and double loose tag categories. In all cases, there is no significant deviation from the expected sensitivity.

$ZH$ Mass	Observed Limit	Expected Limit				
		$-2\sigma$	$-1\sigma$	Median	$+1\sigma$	$+2\sigma$
100	5.6	3.1	4.3	6.0	8.7	12.4
105	6.6	3.6	4.9	6.9	10.0	14.4
110	6.9	4.0	5.4	7.7	11.3	16.2
115	8.7	4.9	6.6	9.4	13.7	19.8
120	9.4	6.0	8.0	11.2	16.4	22.9
125	10.7	7.3	9.7	13.6	19.8	28.2
130	11.8	9.1	11.9	16.8	24.2	34.5
135	16.4	11.7	15.5	21.9	31.4	44.3
140	23.4	16.7	21.2	29.5	41.9	59.4
145	37.5	24.5	31.6	43.9	62.5	89.3
150	60.2	38.8	50.0	69.6	98.7	140.9

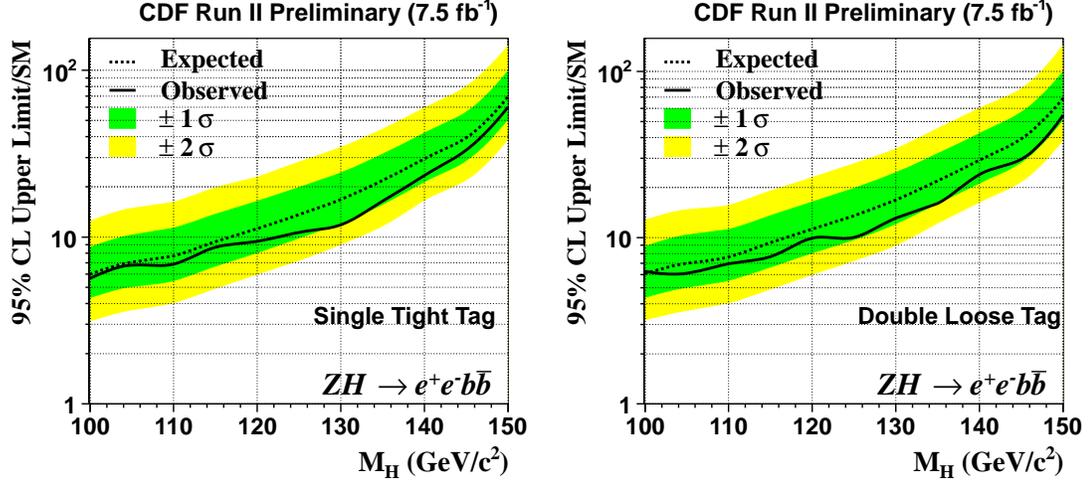
Table E.1: Individual channel limits for the single tight tag category.

$ZH$ Mass	Observed Limit	Expected Limit				
		$-2\sigma$	$-1\sigma$	Median	$+1\sigma$	$+2\sigma$
100	6.3	3.2	4.3	6.1	8.8	12.7
105	6.0	3.6	4.8	6.8	10.1	14.4
110	7.0	4.0	5.5	7.6	11.2	15.7
115	7.7	4.9	6.6	9.3	13.5	19.1
120	9.9	6.0	8.0	11.3	16.4	23.0
125	10.0	7.3	9.6	13.5	19.8	27.8
130	13.1	9.0	12.0	16.9	24.6	34.2
135	16.0	12.2	15.9	21.9	31.9	45.3
140	24.0	16.4	21.0	29.3	42.3	59.8
145	32.8	24.3	31.3	43.8	63.1	87.9
150	54.5	38.8	50.2	68.9	99.1	143.4

Table E.2: Individual channel limits for the double loose tag category.

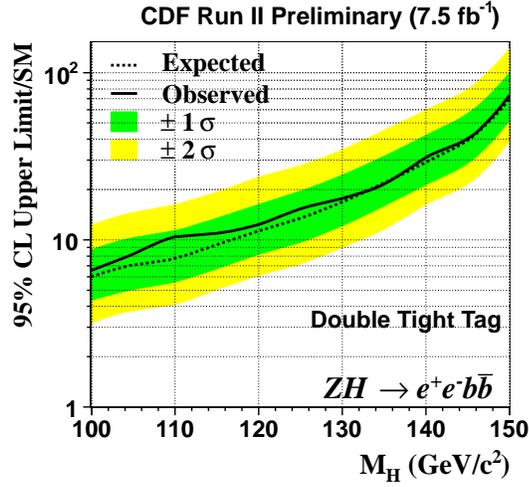
$ZH$ Mass	Observed	Expected Limit				
	Limit	$-2\sigma$	$-1\sigma$	Median	$+1\sigma$	$+2\sigma$
100	6.5	3.2	4.3	6.0	8.7	12.2
105	7.8	3.7	4.8	6.9	10.0	14.1
110	10.4	4.1	5.5	7.8	11.3	16.2
115	10.9	5.1	6.6	9.4	13.5	19.3
120	12.4	6.1	8.1	11.3	16.3	23.7
125	15.3	7.2	9.5	13.5	19.5	27.8
130	17.7	9.0	12.0	16.9	24.4	34.4
135	21.5	11.6	15.7	21.9	31.6	45.2
140	31.1	16.3	21.1	29.2	42.1	59.6
145	44.9	24.1	31.6	44.1	63.6	88.9
150	73.5	39.0	50.7	69.8	99.8	139.2

Table E.3: Individual channel limits for the double tight tag category.



(a) Single Tight Tag

(b) Double Loose Tag



(c) Double Tight Tag

Figure E.1: Upper limits on production cross section times branching ratio over SM expectation for the individual final analysis channels in the  $ZH \rightarrow e^+e^-b\bar{b}$  analysis. The limit is calculated for 11 mass hypotheses and one (green) and two (yellow) standard deviation bands are drawn.

# Bibliography

- [1] “All Nobel Prizes in Physics.” Nobelprize.org.  
[http://www.nobelprize.org/nobel\\_prizes/physics/laureates/](http://www.nobelprize.org/nobel_prizes/physics/laureates/).  
Accessed October 2011.
- [2] J. C. Street and E. C. Stevenson, “New evidence for the existence of a particle of mass intermediate between the proton and electron,” *Phys. Rev.* **52** (Nov, 1937) 1003–1004. <http://link.aps.org/doi/10.1103/PhysRev.52.1003>.
- [3] D. Griffiths, *Introduction to Elementary Particles*. Physics textbook. Wiley-VCH, 2008.
- [4] M. L. Perl *et al.*, “Evidence for anomalous lepton production in  $e^+ - e^-$  annihilation,” *Phys. Rev. Lett.* **35** (Dec, 1975) 1489–1492.  
<http://link.aps.org/doi/10.1103/PhysRevLett.35.1489>.
- [5] K. Nakamura *et al.* *JPG* **37** no. 075021, (2010) . <http://pdg.lbl.gov>.
- [6] A. Das and T. Ferbel, *Introduction To Nuclear and Particle Physics*. World Scientific, 2004.
- [7] L. Lederman, *The God Particle: If the Universe is the Answer, What is the Question*. Dell Pub, New York, N.Y, 1994.
- [8] D. J. Gross and F. Wilczek, “Ultraviolet behavior of non-abelian gauge theories,” *Phys. Rev. Lett.* **30** (Jun, 1973) 1343–1346.  
<http://link.aps.org/doi/10.1103/PhysRevLett.30.1343>.
- [9] H. D. Politzer, “Reliable perturbative results for strong interactions?,” *Phys. Rev. Lett.* **30** (Jun, 1973) 1346–1349.  
<http://link.aps.org/doi/10.1103/PhysRevLett.30.1346>.
- [10] R. Feynman, *QED: The Strange Theory of Light and Matter*. Alix G. Mautner memorial lectures. Princeton University Press, 1985.
- [11] Y. Ng, *Julian Schwinger: The Physicist, the Teacher, and the Man*. World Scientific, 1996.

- [12] M. Peskin and D. Schroeder, *An Introduction to Quantum Field Theory*. Addison-Wesley, Reading, Mass, 1995.
- [13] F. Englert and R. Brout, “Broken symmetry and the mass of gauge vector mesons,” *Phys. Rev. Lett.* **13** no. 9, (Aug, 1964) 321–323.
- [14] P. W. Higgs, “Broken symmetries and the masses of gauge bosons,” *Phys. Rev. Lett.* **13** no. 16, (Oct, 1964) 508–509.
- [15] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, “Global conservation laws and massless particles,” *Phys. Rev. Lett.* **13** no. 20, (Nov, 1964) 585–587.
- [16] W. Greiner, *Gauge Theory of Weak Interactions*. Springer, Berlin New York, 2000.
- [17] E. Commins and P. Bucksbaum, *Weak Interactions of Leptons and Quarks*. Cambridge University Press, 1983.
- [18] B. W. Lee, C. Quigg, and H. B. Thacker, “Weak interactions at very high energies: The role of the Higgs-boson mass,” *Phys. Rev. D* **16** no. 5, (Sep, 1977) 1519–1531.
- [19] J. F. Gunion, H. E. Haber, G. Kane, and S. Dawson, *The Higgs Hunter’s Guide*. Perseus Publishing, Cambridge, 2000.
- [20] C. Quigg, “Electroweak symmetry breaking and the Higgs sector,” *Acta Phys. Polon B* **30** no. 99/033-T, (1999) 2145–2192, [arXiv:hep-ph/9905369v1](#).
- [21] T. van Ritbergen and R. G. Stuart, “On the precise determination of the fermi coupling constant from the muon lifetime,” *Nucl. Phys. B* **564** (2000) 343, [arXiv:hep-ph/9904240v1](#).
- [22] M. Steinhauser and T. Seidensticker, “Second order corrections to the muon lifetime and the semileptonic B decay,” *Phys.Lett.B* **467** (1999) 271, [arXiv:hep-ph/9909436v1](#).
- [23] G. Altarelli and G. Isidori, “Lower limit on the Higgs mass in the standard model: An update,” *Phys. Lett. B* **337** no. 1-2, (1994) 141 – 144.
- [24] J. Casas, J. Espinosa, and M. Quirós *Phys. Lett. B* **342** (1995) 171. *idem.*, *Phys. Lett. B* **382**, 374 (1996).
- [25] M. Sher *Phys. Lett. B* **317** (1993) 159–163, [arXiv:hep-ph/9307342v1](#). Addendum-*ibid.* *B* **331** (1994) 448.
- [26] K. Holland, “Triviality and the Higgs mass lower bound,” *Nucl. Phys. B - Proceedings Supplements* **140** (2005) 155 – 161, [arXiv:hep-lat/0409112v1](#). LATTICE 2004 - Proceedings of the XXII<sup>nd</sup> International Symposium on Lattice Field Theory.

- [27] **NuTeV** Collaboration, G. P. Zeller *et al.*, “A precise determination of electroweak parameters in neutrino nucleon scattering,” *Phys. Rev. Lett.* **88** (2002) 091802, [arXiv:hep-ex/0110059](#). Erratum-ibid. 90 (2003) 239902.
- [28] **The LEP Electroweak Working Group** Collaboration, J. Alcaraz *et al.*, 2011. <http://lepewwg.web.cern.ch/LEPEWWG/>.
- [29] “Search for the Standard Model Higgs boson at LEP,” *Phys. Lett. B* **565** (2003) 61 – 75, [arXiv:hep-ex/0306033v1](#).
- [30] The TEVNPH Working Group, “Combined CDF and D0 upper limits on standard model Higgs-boson production with up to 6.7 fb<sup>-1</sup> of data,” FERMILAB-CONF 10-257-E, FERMILAB, 2010. [arXiv:1007.4587v1 \[hep-ex\]](#).
- [31] B. Kilminster *et al.*, “Search for ZHllbb in 5.7/fb of CDF 2 data,” CDF Note 10235, CDF, 2010.
- [32] A. Djouadi, J. Kalinowski, and M. Spira *Comput. Phys. Commun.* **108** (1998) 56, [arXiv:hep-ph/9704448v1](#).
- [33] J. Baglio and A. Djouadi, “Predictions for Higgs production at the Tevatron and the associated uncertainties,” *J. High Energy Phys.* **2010** (2010) 1–53, [arXiv:1003.4266v2 \[hep-ph\]](#). 10.1007/JHEP10(2010)064.
- [34] S. Dittmaier *et al.*, “Handbook of LHC Higgs Cross Sections: 1. Inclusive Observables,” 2011.
- [35] [http://www-bdnew.fnal.gov/operations/rookie\\_books/rbooks.html](http://www-bdnew.fnal.gov/operations/rookie_books/rbooks.html).
- [36] **CDF II** Collaboration, “The CDF II Detector Technical Design Report,” FERMILAB-Pub 96/390-E, Fermilab, Fermi National Accelerator Laboratory, P.O. Box 500, Batavia, Illinois 60510, November, 1996.
- [37] H. Minemura *et al.*, “Construction and testing of a 3 m diameter x 5 m superconducting solenoid for the Fermilab Collider Detector Facility (CDF),” *Nucl. Instrum. Methods* **A238** no. 1, (1985) 18 – 34.
- [38] P. Azzi-Bacchetta *et al.*, “Proposal for a very low mass, very small radius silicon layer in the CDF upgrade,” CDF Note 4924, CDF, 1998.
- [39] **CDF** Collaboration, A. Sill, “CDF run II silicon tracking projects,” *Nucl. Instrum. Methods* **A447** (2000) 1–8.
- [40] D. Ambrose *et al.*, “Update of COT prototype aging tests,” CDF Note 7540, CDF, 2004.
- [41] T. Affolder *et al.*, “CDF central outer tracker,” *Nucl. Instrum. Methods* **A526** no. 3, (2004) 249 – 299.

- [42] L. Balka *et al.*, “The CDF central electromagnetic calorimeter,” *Nucl. Instrum. Methods* **A267** no. 2-3, (1988) 272 – 279.
- [43] M. Albrow *et al.*, “The CDF plug upgrade electromagnetic calorimeter: test beam results,” *Nucl. Instrum. Methods* **A480** no. 2-3, (2002) 524 – 546.
- [44] L. Balka *et al.*, “The CDF central electromagnetic calorimeter,” *Nucl. Instrum. Methods* **A267** no. 2-3, (1988) 272 – 279.
- [45] S. Cihangir, F. Marchetto, P. McIntyre, T. Meyer, and R. Webb, “The CDF forward/backward hadron calorimeter,” *Nucl. Instrum. Methods* **A267** no. 2-3, (1988) 249 – 256.
- [46] **CDF IIb** Collaboration, “The CDF II Detector Technical Design Report,” tech. rep., Fermilab, Fermi National Accelerator Laboratory, P.O. Box 500, Batavia, Illinois 60510, September, 2002.
- [47] D. Acosta *et al.*, “The CDF Cherenkov luminosity monitor,” *Nucl. Instrum. Methods* **A461** no. 1-3, (2001) 540 – 544. 8th Pisa Meeting on Advanced Detectors.
- [48] R. Downing *et al.*, “Track extrapolation and distribution for the CDF-II trigger system,” *Nucl. Instrum. Methods* **A570** no. 1, (2007) 36 – 50.
- [49] A. Taffard, “Run II cosmic ray tagger,” CDF Note 6100, CDF, 2003.
- [50] W. Badgett, H. Frisch, R. St. Denis, T. Vaiciulis, “The good run list,” CDF Note 5613, CDF, 2001.
- [51] J. Goldstein, J. Incandela, C. Issever, C. Mills, T. Nelson, R. Snider, and D. Stuart, “Measurement of  $W \rightarrow e\nu$  charge asymmetry in Run 2 with calorimeter-seeded silicon tracks,” CDF Note 6282, CDF, 2003.
- [52] C. Ciobanu, R. Hughes, P. Koehn, C. Neu, and B. Winer, “A ROOT interface to JETNET,” CDF Note 5434, CDF, 2002.
- [53] <http://tmva.sourceforge.net/#credits>.
- [54] **CDF** Collaboration, A. Abulencia *et al.*, “Observation of  $WZ$  production,” *Phys. Rev. Lett.* **98** (Apr, 2007) 161801, [arXiv:hep-ex/0702027v1](https://arxiv.org/abs/hep-ex/0702027v1). FERMILAB-PUB-07-042-E.
- [55] S. Shalhout, *A Search for the Standard Model Higgs Boson in the Process  $ZH \rightarrow \ell^+\ell^-b\bar{b}$  in  $4.1 \text{ fb}^{-1}$  of CDF II Data*. PhD thesis, Wayne State University, 2010.
- [56] **CDF** Collaboration, F. Abe *et al.*, “Topology of three-jet events in  $p\bar{p}$  collisions at  $s = 1.8 \text{ TeV}$ ,” *Phys. Rev. D* **45** no. 5, (Mar, 1992) 1448–1458.

- [57] CDF Collaboration, A. Bhatti *et al.*, “Determination of the jet energy scale at the collider detector at Fermilab,” *Nucl. Instrum. Methods* **A566** (2006) 375–412, [arXiv:hep-ex/0510047](#).
- [58] W. Yao *et al.*, “A seed vertexing b-Tag algorithm for top,” CDF Note 2716, CDF, 1994.
- [59] D. Gerdes, “B identification using jet probability,” CDF Note 2023, CDF, 1993.
- [60] **CDF and D0 Working Group Members** Collaboration, L. Babukhadia *et al.*, “Results of the Tevatron Higgs sensitivity study,” FERMILAB-PUB-03-320-E.
- [61] J. Efron, *Search for the Higgs Boson in the  $ZH \rightarrow \ell^+ \ell^- b\bar{b}$  Channel at CDF Run II*. PhD thesis, The Ohio State University, 2007. CDF Note 8976.
- [62] R. Hughes *et al.*, “A search for  $ZH \rightarrow llbb$  in  $2.7 \text{ fb}^{-1}$  using a neural network discriminant,” CDF Note 9665, CDF, 2009.
- [63] N. Metropolis, “The beginning of the Monte Carlo method,” *Los Alamos Science* **15** (1987) 125–130.  
<http://library.lanl.gov/cgi-bin/getfile?15-12.pdf>. Special Issue.
- [64] CTEQ Collaboration, H. L. Lai *et al.*, “Global QCD analysis of parton structure of the nucleon: CTEQ5 parton distributions,” *Eur. Phys. J.* **C12** (2000) 375–392, [arXiv:hep-ph/9903282](#). MSUHEP-903100.
- [65] M. L. Mangano, F. Piccinini, A. D. Polosa, M. Moretti, and R. Pittau, “ALPGEN, a generator for hard multiparton processes in hadronic collisions,” *J. High Energy Phys.* **2003** (2003) 1–38, [arXiv:hep-ph/0206293v2](#). CERN-TH/2002-129.
- [66] T. Sjöstrand, P. Edén, C. Friberg, L. Lönnblad, G. Miu, S. Mrenna, and E. Norrbin, “High-energy-physics event generation with PYTHIA 6.1,” *Computer Physics Communications* **135** no. 2, (2001) 238 – 259, [arXiv:hep-ph/0010017v1](#). LU TP 00-30.
- [67] R. Brun, R. Hagelberg, M. Hansroul, and J. Lassalle, *GEANT Detector Description and Simulation Tool*. CERN Program Library Long Writeup W5013, 1994. CERN-DD-78-2-REV.
- [68] S. Agostinelli *et al.*, “GEANT4—simulation toolkit,” *Nucl. Instrum. Methods* **A506** no. 3, (2003) 250 – 303.
- [69] O. González, K. Potamianos, M. Vidal, and H. Wolfe, “Using neural-network parameterizations of the trigger efficiency for the met+jet sample,” CDF Note 10538, CDF, 2011.

- [70] L. Hansen and P. Salamon, “Neural network ensembles,” *IEEE Transactions on Pattern Analysis and Machine Intelligence* **12** (1990) 993–1001.
- [71] M. Assaad, R. Boné, and H. Cardot, “A new boosting algorithm for improved time-series forecasting with recurrent neural networks,” *Information Fusion* **9** no. 1, (2008) 41 – 55. Special Issue on Applications of Ensemble Methods.
- [72] V. B. Bo-Young Han, “Trigger efficiencies for the high Et central electron in the Gen6 data (run:217990 - 222426),” CDF Note 8629, CDF, 2006.
- [73] D. Hare, E. Halkiadakis, and T. Spreitzer, “Electron ID efficiency and scale factors for winter 2007 analyses,” CDF Note 8614, CDF, 2006.
- [74] CDF Collaboration, A. Abulencia *et al.*, “Measurements of inclusive  $W$  and  $Z$  cross sections in  $p\bar{p}$  collisions at  $\sqrt{s} = 1.96$  TeV,” *J. Phys. G: Nucl. Part. Phys.* **34** (2007) 2457, [arXiv:hep-ex/0508029v3](#). FERMILAB-PUB-05-360-E.
- [75] CDF Collaboration, D. Acosta *et al.*, “Measurement of the  $t\bar{t}$  production cross section in  $p\bar{p}$  collisions at  $\sqrt{s} = 1.96$  TeV using lepton + jets events with secondary vertex  $b$ -tagging,” *Phys. Rev. D* **71** (Mar, 2005) 052003, [arXiv:hep-ex/0410041v2](#). FERMILAB-PUB-04-275-E.
- [76] J. Freeman, S. Grinstein, T. Junk, and E. Palencia, “Secvtx scale factors and mistag matrices for 2.2 fb-1 of data up to p13,” CDF Note 9280, CDF, 2008.
- [77] B. Alvarez, J. Cuevas, G. Gomez, E. Palencia, T. Rodrigo, L. Scodellaro, and T. Vilar, “Jet probability tag rate matrix for Gen6 2.2invfb data,” CDF Note 9227, CDF, 2008.
- [78] CDF Collaboration, A. Abulencia *et al.*, “Measurement of the  $t\bar{t}$  production cross section in  $p\bar{p}$  collisions at  $\sqrt{s} = 1.96$  TeV using lepton + jets events with jet probability  $b$ -tagging,” *Phys. Rev. D* **74** (Oct, 2006) 072006.
- [79] S. Grinstein, J. Guimarães da Costa, and D. Sherman, “SecVtx mistag asymmetry for winter 2007,” CDF Note 8626, CDF, 2006.
- [80] J. Freeman, “Summer 2009 SecVtx scale factors calculated using the electron method through period 22,” CDF Note 9848, CDF, 2009.
- [81] F. Garberson and E. Palencia, “Jet probability  $b$ -tag efficiency measurement using muon transverse momentum for 1.2 fb<sup>-1</sup>,” CDF Note 8882, 2007.
- [82] A. Buzatu *et al.*, “Search for the Standard Model Higgs boson production in association with  $W$  boson using 3.2fb-1,” CDF Note 9756, CDF, 2009.
- [83] M. J. Frank, *A Search for the Standard Model Higgs Boson Produced in Association with a  $W$  Boson*. PhD thesis, Baylor University, 2011.

- [84] S. Klimenko, J. Konigsberg, and T. Liss, “Averaging of the inelastic cross sections measured by the CDF and the E811 experiments,” FN 0741, Fermilab, 2003.
- [85] J. M. Campbell and R. K. Ellis, “Update on vector boson pair production at hadron colliders,” *Phys. Rev. D* **60** (Nov, 1999) 113006, [arXiv:hep-ph/9905386v2](#).
- [86] J. Campbell and R. K. Ellis, “Radiative corrections to  $Zb\bar{b}$  production,” *Phys. Rev. D* **62** (Oct, 2000) 114012, [arXiv:hep-ph/0006304v1](#).
- [87] U. Langenfeld, S. Moch, and P. Uwer, “Measuring the running top-quark mass,” *Phys. Rev. D* **80** (Sep, 2009) 054009.
- [88] S. Grinstein, J. Guimarães da Costa, and D. Sherman, “Electron-method SecVtx scale factor for winter 2007,” CDF Note 8625, CDF, 2007.
- [89] CDF Collaboration, A. Abulencia *et al.*, “Top quark mass measurement using the template method in the lepton + jets channel at CDF II,” *Phys. Rev. D* **73** (Feb, 2006) 032003, [arXiv:hep-ex/0510048v1](#).
- [90] A. Bhatti *et al.*, “Determination of the jet energy scale at the Collider Detector at Fermilab,” *Nucl. Instrum. Methods* **A566** no. 2, (2006) 375 – 412, [arXiv:hep-ex/0510047v1](#).
- [91] Correspondence with T. Junk.
- [92] J. Freeman, “Summer 2009 SecVtx mistag asymmetry corrections through period 22,” CDF Note 9847, CDF, 2009.
- [93] CDF Collaboration, T. Aaltonen *et al.*, “Search for the Higgs boson produced in association with  $Z \rightarrow \ell^+\ell^-$  in  $p\bar{p}$  collisions at  $\sqrt{s} = 1.96$  TeV,” *Phys. Rev. Lett.* **101** (Dec, 2008) 251803, [arXiv:0807.4493v2 \[hep-ex\]](#). FERMILAB-PUB-08-250-E.
- [94] R. Hughes *et al.*, “A search for  $ZH \rightarrow llbb$  in  $2.4 \text{ fb}^{-1}$  using a neural network discriminant,” CDF Note 9475, CDF, 2008.
- [95] R. Hughes *et al.*, “Search for  $ZH \rightarrow \ell^+\ell^-b\bar{b}$  in  $4.1/\text{fb}$  of CDF 2 data,” CDF Note 9889, CDF, 2009.
- [96] CDF Collaboration, T. Aaltonen *et al.*, “Improved search for a Higgs boson produced in association with  $Z \rightarrow l^+l^-$  in  $p\bar{p}$  collisions at  $\sqrt{s} = 1.96$  TeV,” *Phys. Rev. Lett.* **105** (Dec, 2010) 251802, [arXiv:1009.3047v2 \[hep-ex\]](#). Fermilab PUB-10-365-E.
- [97] J. Pilot *et al.*, “Search for the Standard Model Higgs boson in the  $ZH \rightarrow \mu\mu bb$  Final state using  $7.9/\text{fb}$ ,” CDF Note 10572, CDF, 2011.

- [98] M. Feindt, T. Muller, S. Richter, and W. Wagner, “A neural network b tagger for single-top analyses,” CDF Note 7816, CDF, 2006.
- [99] T. Chwalek, M. Feindt, D. Hirschbuehl, T. Muller, M. Renz, S. Richter, and W. Wagner, “Update of the neural network b tagger for single-top analyses,” CDF Note 8903, CDF, 2007.
- [100] J. Gallicchio, J. Huth, M. Kagan, M. Schwartz, K. Black, and B. Tweedie, “Multivariate discrimination and the Higgs+W/Z search,” *Journal of High Energy Physics* **2011** (2011) 1–41, [arXiv:1010.3698v2 \[hep-ph\]](#).  
10.1007/JHEP04(2011)069.
- [101] S. Lockwitz *et al.*, “A Search for the Standard Model Higgs boson in the process ZH to eebb using 7.5/fb of CDF II data,” CDF Note 10593, CDF, 2011.
- [102] G. Mahlon and S. Parke, “Using spin correlations to distinguish Zh from ZA at the International Linear Collider,” *Phys. Rev. D* **74** (Oct, 2006) 073001, [arXiv:hep-ph/0606052v1](#). Fermilab-Pub-06/160-T.
- [103] J. Heinrich, “Bayesian limit software: Multi-channel with correlated backgrounds and efficiencies,” CDF Note 7587, CDF, 2005.
- [104] T. Junk, “Sensitivity, exclusion and discovery with small signals, large backgrounds, and large systematic uncertainties,” CDF Note 8128, CDF, 2007.
- [105] T. Junk, “Confidence level computation for combining searches with small statistics,” *Nucl. Instrum. Methods* **A434** (1999) 435–443, [arXiv:hep-ex/9902006](#).
- [106] T. Junk, E. James, B. Kilminster, C. Group, and W. Yao, “Combination of CDF’s searches for the Standard Model Higgs boson with up to 8.2 fb-1 of data,” CDF Note 10609, CDF, 2011.
- [107] The TEVNPH Working Group, “Combined CDF and D0 upper limits on standard model Higgs boson production with up to 8.6 fb-1 of data,” FERMILAB-CONF 11-354-E, FERMILAB, 2011. [arXiv:1107.5518v1 \[hep-ex\]](#).
- [108] J. Freeman, W. Ketchum, J. Lewis, S. Poprocki, A. Pronko, V. Rusu, and P. Wittich, “An artificial neural network based b jet identification algorithm at the CDF experiment,” *Nucl. Instrum. Methods* **A** no. 0, (2011) –, [arXiv:1108.4738v2 \[hep-ex\]](#). FERMILAB-TM-2515-E-PPD.
- [109] CDF Collaboration, T. Aaltonen *et al.*, “Search for the Higgs boson produced in association with  $Z \rightarrow \ell^+\ell^-$  using the matrix element method at CDF II,” *Phys. Rev. D* **80** (Oct, 2009) 071101, [arXiv:0908.3534v1 \[hep-ex\]](#).

- [110] **MiniBooNE** Collaboration, Aguilar-Arevalo *et al.*, “Search for electron neutrino appearance at the  $\Delta m^2 \sim 1 \text{ eV}^2$  scale,” *Phys. Rev. Lett.* **98** (Jun, 2007) 231801, [arXiv:0704.1500v3 \[hep-ex\]](#). A-UR-07-2246.
- [111] A. K. Jain, J. Mao, and K. Mohiuddin, “Artificial neural networks: A tutorial,” *Computer* **29** (1996) 31–44.