Search for the Higgs Boson in the $ZH \to \mu^+\mu^- b\bar{b}$ Channel at CDF Using Novel Multivariate Techniques

DISSERTATION

Presented in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy in the Graduate School of The Ohio State University

By

Justin Robert Pilot, B.S., M.S.

Graduate Program in Physics

The Ohio State University

2011

Dissertation Committee:

Professor Brian Winer, Advisor

Professor Richard Hughes

Professor Junko Shigemitsu

Professor Michael Poirier

ABSTRACT

We present a search for the Standard Model Higgs Boson using the process $ZH \to \mu^+\mu^- b\bar{b}$. We use a dataset corresponding to 9.2 fb^{-1} of integrated luminosity from proton-antiproton collisions with center-of-mass energy 1.96 TeV at the Fermilab Tevatron, collected with the CDF II detector. This analysis benefits from several new multivariate techniques that have not been used in previous analyses at CDF. We use a multivariate function to select muon candidates, increasing signal acceptance while simultaneously keeping fake rates small. We employ an inclusive trigger selection to further increase acceptance. To enhance signal discrimination, we utilize a multi-layer approach consisting of expert discriminants. This multi-layer discriminant method helps isolate the two main classes of background events, $t\bar{t}$ and Z+jets production. It also includes a flavor separator, to distinguish light flavor jets from jets consistent with the decay of a B-hadron. With this novel multi-layer approach, we proceed to set limits on the ZH production cross section times branching ratio. For a Higgs boson with mass 115 GeV/c^2, we observe (expect) a limit of 8.0 (4.9) times the Standard Model prediction.

To my family,
for all of their support.

Acknowledgments

I would not have reached this stage in my education without the advice and guidance of many people, and I cannot thank them enough for their contributions. First of all, my advisors Brian Winer and Richard Hughes – thank you for taking me as a graduate student and developing my scientific abilities, providing motivation and guidance along the way when I struggled. I have learned a great deal from you both, and I know that your training will allow me to be successful in the field in the future.

Thanks also go to Homer Wolfe, our group's CDF post-doc. Thank you for always being there to answer questions and look over plots and slides. I enjoyed the many long phone conversations we had discussing analysis details. You've provided me with an increased excitement toward the work I will do in the future.

To my fellow CDF student at Ohio State, Jon Wilson, thanks for answering the endless computing-related questions I have asked. I admire your expertise in this area. Also, thanks for being there to help puzzle out analysis details and problems. I've enjoyed working with you over the past several years. I also want to thank everyone else in our group at OSU that I've known over the years – thanks for all the good times we've shared both in and out of the lab.

I would also like to thank everyone I worked with while spending time at Fermilab, operating the CDF detector. I enjoyed a great working environment and met a wide variety of people. Without them and countless others, the experiment would not have been the success that it is today. Specific thanks go to the CDF Higgs group conveners, Ben Kilminster, Eric James, and Craig Group, for their great suggestions and guidance over the years.

Finally, I'd like to thank my parents and family for their never-ending love and support. Without it, I could not have accomplished nearly as much.

VITA

June 2006	REU Participant, University of Notre Dame, Notre Dame, Indiana
2006-2007	Teaching Assistant, Michigan State University, East Lansing, Michigan
May 2007	**B.S. Physics**, Michigan State University
2007-2009	Graduate Teaching Associate, The Ohio State University, Columbus, Ohio
2008-2011	Graduate Research Associate, The Ohio State University
August 2009	**M.S. Physics**, The Ohio State University

Publications

"Search for high-mass resonances decaying into ZZ in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV"
T. Aaltonen *et al.* [CDF Collaboration]
arXiv:1111.3432 [hep-ex]
FERMILAB-PUB-11-613-E(2011) (Submitted to Phys. Rev. D)

"Search for resonant production of $t\bar{t}$ pairs in 4.8 fb^{-1} of integrated luminosity of $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV"
T. Aaltonen *et al.* [CDF Collaboration]
Phys. Rev. D **84**, 072004 (2011)

"Evidence for the charmless annihilation decay mode $B_s^0 \to \pi^+\pi^-$"
T. Aaltonen *et al.* [CDF Collaboration]

arXiv:1111.0485 [hep-ex]
FERMILAB-PUB-11-585-E(2011)

"Search for a Higgs Boson in the Diphoton Final State in $p\bar{p}$ Collisions at $\sqrt{s} = 1.96$ TeV"
T. Aaltonen *et al.* [CDF Collaboration]
arXiv:1109.4427 [Unknown]
FERMILAB-PUB-11-482-E-PPD(2011)

"Search for heavy metastable particles decaying to jet pairs in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV"
T. Aaltonen *et al.* [CDF Collaboration]
arXiv:1109.3136 [hep-ex]
FERMILAB-PUB-11-451-E-PPD(2011) (Submitted to Phys.Rev.Lett.)

"Top-quark mass measurement using events with missing transverse energy and jets at CDF"
T. Aaltonen *et al.* [CDF Collaboration]
arXiv:1109.1490 [hep-ex]
FERMILAB-PUB-11-422-E-PPD(2011)

"Measurements of branching fraction ratios and CP-asymmetries in suppressed $B^- \to D(\to K^+\pi^-)K^-$ and $B^- \to D(\to K^+\pi^-)\pi^-$ decays"
T. Aaltonen *et al.* [CDF Collaboration]
arXiv:1108.5765 [hep-ex]
FERMILAB-PUB-11-404-E-PPD(2011) (Submitted to Phys.Rev.D)

"Search for resonant production of $t\bar{t}$ decaying to jets in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV"
T. Aaltonen *et al.* [CDF Collaboration]
arXiv:1108.4755 [hep-ex]
FERMILAB-PUB-11-397-E(2011)

"Search for $WZ + ZZ$ production with $\not\!\!E_T$ + jets with b enhancement at $\sqrt{s} = 1.96$ TeV"
T. Aaltonen *et al.* [CDF Collaboration]
arXiv:1108.2060 [hep-ex]
FERMILAB-PUB-11-375-E-PPD(2011) (Submitted to Phys.Rev.Lett.)

"Measurement of the top-quark mass in the lepton+jets channel using a matrix element technique with the CDF II detector"
T. Aaltonen *et al.* [CDF Collaboration]
arXiv:1108.1601 [hep-ex]
FERMILAB-PUB-11-374-E-PPD(2011) (Submitted to Phys.Rev.D)

"Measurements of the Angular Distributions in the Decays $B \to K^{(*)}\mu^+\mu^-$ at CDF"
T. Aaltonen *et al.* [CDF Collaboration]
arXiv:1108.0695 [hep-ex]
FERMILAB-PUB-11-364-PPD(2011)

"Search for new physics in high p_T like-sign dilepton events at CDF II"
T. Aaltonen *et al.* [CDF Collaboration]
arXiv:1108.0101 [hep-ex]
FERMILAB-PUB-11-367-E(2011) (Submitted to Phys.Rev.Lett.)

"A search for resonant production of $t\bar{t}$ pairs in 4.8 fb^{-1} of integrated luminosity of $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV"
T. Aaltonen *et al.* [The CDF Collaboration]
arXiv:1107.5063 [hep-ex]
FERMILAB-PUB-11-350-E(2011)

"Measurement of Polarization and Search for CP-Violation in $B_s^0 \to \phi\phi$ Decays"
T. Aaltonen *et al.* [CDF Collaboration]
arXiv:1107.4999 [hep-ex]
FERMILAB-PUB-11-345-E(2011)

"Observation of the Ξ_b^0 Baryon"
T. Aaltonen *et al.* [CDF Collaboration]
arXiv:1107.4015 [hep-ex]
FERMILAB-PUB-11-323-E-PPD(2011) (Submitted to Phys.Rev.Lett.)

"Search for a Heavy Top-Like Quark in $p\bar{p}$ Collisions at $\sqrt{s} = 1.96$ TeV"
T. Aaltonen *et al.* [CDF Collaboration]
arXiv:1107.3875 [hep-ex]
FERMILAB-PUB-11-333-E-PPD(2011) (Submitted to Phys.Rev.Lett.)

"Observation of the Baryonic Flavor-Changing Neutral Current Decay $\Lambda_b^0 \to \Lambda\mu^+\mu^-$"
T. Aaltonen *et al.* [CDF Collaboration]
arXiv:1107.3753 [hep-ex]
FERMILAB-PUB-11-328-E(2011)

"Search for new physics in $t\bar{t} + \not{E}_T \to b\bar{b}q\bar{q}q\bar{q} + \not{E}_T$ final state in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV"
T. Aaltonen *et al.* [CDF Collaboration]
arXiv:1107.3574 [hep-ex]
FERMILAB-PUB-11-326-E-PPD(2011) (Submitted to Phys.Rev.Lett.)

"Search for $B_s \to \mu^+\mu^-$ and $B_d \to \mu^+\mu^-$ Decays with CDF II"
T. Aaltonen *et al.* [CDF Collaboration]
arXiv:1107.2304 [hep-ex]
FERMILAB-PUB-11-315-E-PPD(2011)

"Study of Substructure of High Transverse Momentum Jets Produced in Proton-Antiproton Collisions at $\sqrt{s} = 1.96$ TeV"
T. Aaltonen *et al.* [CDF Collaboration]

arXiv:1106.5952 [hep-ex]
FERMILAB-PUB-11-297-E-PPD(2011)

"Measurement of the Cross Section for Prompt Isolated Diphoton Production in $p\bar{p}$ Collisions at $\sqrt{s} = 1.96$ TeV"
T. Aaltonen *et al.* [CDF Collaboration]
arXiv:1106.5131 [hep-ex]
FERMILAB-PUB-11-296-PPD(2011) (Submitted to Phys.Rev.D)

"Measurement of the Cross Section for Prompt Isolated Diphoton Production in $p\bar{p}$ Collisions at $\sqrt{s} = 1.96$ TeV"
T. Aaltonen *et al.* [CDF Collaboration]
arXiv:1106.5123 [hep-ex]
FERMILAB-PUB-11-295-PPD(2011) (Submitted to Phys.Rev.Lett.)

"Search for Higgs Bosons Produced in Association with b-Quarks"
T. Aaltonen *et al.* [CDF Collaboration]
arXiv:1106.4782 [hep-ex]
FERMILAB-PUB-11-294-E-PPD(2011)

"Evidence for $t\bar{t}\gamma$ Production and Measurement of $\sigma_{t\bar{t}\gamma}/\sigma_{t\bar{t}}$"
T. Aaltonen *et al.* [CDF Collaboration]
arXiv:1106.3970 [hep-ex]
FERMILAB-PUB-11-283-E-PPD(2011) (Submitted to Phys.Rev.D)

"Measurement of branching ratio and B_s^0 lifetime in the decay $B_s^0 \to J/\psi f_0(980)$ at CDF"
T. Aaltonen *et al.* [CDF Collaboration]
arXiv:1106.3682 [hep-ex]
FERMILAB-PUB-11-279-E(2011) (Submitted to Phys.Rev.D)

"Measurements of the properties of $\Lambda_c(2595)$, $\Lambda_c(2625)$, $\Sigma_c(2455)$, and $\Sigma_c(2520)$ baryons"
T. Aaltonen *et al.* [CDF Collaboration]
Phys. Rev. D **84**, 012003 (2011) [arXiv:1105.5995 [hep-ex]]

"First Search for Multijet Resonances in $\sqrt{s} = 1.96$ TeV $p\bar{p}$ Collisions"
T. Aaltonen *et al.* [CDF Collaboration]
Phys. Rev. Lett. **107**, 042001 (2011) [arXiv:1105.2815 [hep-ex]]

"Measurement of the $t\bar{t}$ production cross section in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV using events with large Missing E_T and jets"
T. Aaltonen *et al.* [CDF Collaboration]
arXiv:1105.1806 [hep-ex]
FERMILAB-PUB-11-210-E-PPD(2011) (Submitted to Phys.Rev.D)

"Top quark mass measurement using the template method at CDF"
T. Aaltonen *et al.* [CDF Collaboration]

Phys. Rev. D **83**, 111101 (2011) [arXiv:1105.0192 [hep-ex]]

"Search for a Very Light CP-Odd Higgs Boson in Top Quark Decays from $p\bar{p}$ Collisions at 1.96 TeV"
T. Aaltonen *et al.* [CDF Collaboration]
Phys. Rev. Lett. **107**, 031801 (2011) [arXiv:1104.5701 [hep-ex]]

"Search for the Rare Radiative Decay: $W \to \pi\gamma$ in $p\bar{p}$ Collisions at $\sqrt{s} = 1.96$ TeV"
T. Aaltonen *et al.* [CDF Collaboration]
arXiv:1104.1585 [hep-ex]
FERMILAB-PUB-11-171-E(2011) (Submitted to Phys.Rev.Lett.)

"Invariant Mass Distribution of Jet Pairs Produced in Association with a W boson in $p\bar{p}$ Collisions at $\sqrt{s} = 1.96$ TeV"
T. Aaltonen *et al.* [CDF Collaboration]
Phys. Rev. Lett. **106**, 171801 (2011) [arXiv:1104.0699 [hep-ex]]

"Measurements of Direct CP Violating Asymmetries in Charmless Decays of Strange Bottom Mesons and Bottom Baryons"
T. Aaltonen *et al.* [CDF Collaboration]
Phys. Rev. Lett. **106**, 181802 (2011) [arXiv:1103.5762 [hep-ex]]

"First Measurement of the Angular Coefficients of Drell-Yan e^+e^- pairs in the Z Mass Region from $p\bar{p}$ Collisions at $\sqrt{s} = 1.96$ TeV"
T. Aaltonen *et al.* [CDF Collaboration]
Phys. Rev. Lett. **106**, 241801 (2011) [arXiv:1103.5699 [hep-ex]]

"Measurement of Event Shapes in Proton-Antiproton Collisions at Center-of-Mass Energy 1.96 TeV"
T. Aaltonen *et al.* [CDF Collaboration]
Phys. Rev. D **83**, 112007 (2011) [arXiv:1103.5143 [hep-ex]]

"Measurement of the Top Pair Production Cross Section in the Lepton + Jets Channel Using a Jet Flavor Discriminant"
T. Aaltonen *et al.* [CDF Collaboration]
Phys. Rev. D **84**, 031101 (2011) [arXiv:1103.4821 [hep-ex]]

"Search for New Dielectron Resonances and Randall-Sundrum Gravitons at the Collider Detector at Fermilab"
T. Aaltonen *et al.* [CDF Collaboration]
Phys. Rev. Lett. **107**, 051801 (2011) [arXiv:1103.4650 [hep-ex]]

"Limits on Anomalous Trilinear Gauge Couplings in $Z\gamma$ Events from $p\bar{p}$ Collisions at $\sqrt{s} =$ 1.96 TeV"
T. Aaltonen *et al.* [CDF Collaboration]
Phys. Rev. Lett. **107**, 051802 (2011) [arXiv:1103.2990 [hep-ex]]

"Measurement of the mass difference between t and $\bar{t}$ quarks"
T. Aaltonen *et al.* [CDF Collaboration]
Phys. Rev. Lett. **106**, 152001 (2011) [arXiv:1103.2782 [hep-ex]]

"Search for Production of Heavy Particles Decaying to Top Quarks and Invisible Particles in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV"
T. Aaltonen *et al.* [CDF Collaboration]
Phys. Rev. Lett. **106**, 191801 (2011) [arXiv:1103.2482 [hep-ex]]

"Measurement of the B_s Lifetime in Fully and Partially Reconstructed $B_s \to D_s^-(\phi\pi^-)X$ Decays in $p\bar{p}$ Collisions at $\sqrt{s} = 1.96$ TeV"
T. Aaltonen *et al.* [CDF Collaboration]
arXiv:1103.1864 [hep-ex]
FERMILAB-PUB-11-051-E(2011) (Submitted to Phys.Rev.Lett.)

"Search for New Heavy Particles Decaying to $ZZ \to \ell\ell\ell\ell, \ell\ell jj$ in $p\bar{p}$ Collisions at $\sqrt{s} = 1.96$ TeV"
T. Aaltonen *et al.* [CDF Collaboration]
Phys. Rev. D **83**, 112008 (2011) [arXiv:1102.4566 [hep-ex]]

"Observation of $B_s^0 \to J/\psi K^{*0}(892)$ and $B_s^0 \to J/\psi K_S^0$ Decays"
T. Aaltonen *et al.* [CDF Collaboration]
Phys. Rev. D **83**, 052012 (2011) [arXiv:1102.1961 [hep-ex]]

"Search for the Higgs boson in the all-hadronic final state using the CDF II detector"
T. Aaltonen *et al.* [The CDF Collaboration]
arXiv:1102.0024 [hep-ex]
FERMILAB-PUB-11-021-E(2011) (Submitted to Phys.Rev.Lett.)

"Observation of the $Y(4140)$ structure in the $J/\psi\,\phi$ Mass Spectrum in $B^\pm \to J/\psi\,\phi K$ decays"
T. Aaltonen *et al.* [The CDF Collaboration]
arXiv:1101.6058 [hep-ex]
FERMILAB-PUB-11-019-E(2011)

"Search for heavy bottom-like quarks decaying to an electron or muon and jets in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV"
T. Aaltonen *et al.* [The CDF Collaboration]
Phys. Rev. Lett. **106**, 141803 (2011) [arXiv:1101.5728 [hep-ex]]

"Measurement of the Top Quark Mass in the Lepton+Jets Channel Using the Lepton Transverse Momentum"
T. Aaltonen *et al.* [The CDF Collaboration]
Phys. Lett. B **698**, 371 (2011) [arXiv:1101.4926 [hep-ex]]

"Search for High Mass Resonances Decaying to Muon Pairs in $\sqrt{s} = 1.96$ TeV $p\bar{p}$ Collisions"
T. Aaltonen *et al.* [The CDF Collaboration]
Phys. Rev. Lett. **106**, 121801 (2011) [arXiv:1101.4578 [hep-ex]]

"Production of Λ, $\bar{\Lambda}^0$, $\Xi^{\pm}$ and $\Omega^{\pm}$ Hyperons in $p\bar{p}$ Collisions at $\sqrt{s} = 1.96$ TeV"
T. Aaltonen *et al.* [CDF Collaboration]
arXiv:1101.2996 [hep-ex]
FERMILAB-PUB-11-155-E-PPD(2011)

"Measurement of the Forward-Backward Asymmetry in the $B \to K^{(*)}\mu^+\mu^-$ Decay and First Observation of the $B_s^0 \to \phi\mu^+\mu^-$ Decay"
T. Aaltonen *et al.* [CDF Collaboration]
Phys. Rev. Lett. **106**, 161801 (2011) [arXiv:1101.1028 [hep-ex]]

"Evidence for a Mass Dependent Forward-Backward Asymmetry in Top Quark Pair Production"
T. Aaltonen *et al.* [CDF Collaboration]
Phys. Rev. D **83**, 112003 (2011) [arXiv:1101.0034 [hep-ex]]

"Search for a New Heavy Gauge Boson W' with Electron $+\not{E}_T$ Event Signature in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV"
T. Aaltonen *et al.* [CDF Collaboration]
Phys. Rev. D **83**, 031102 (2011) [arXiv:1012.5145 [hep-ex]]

"Top Quark Mass Measurement in the Lepton + Jets Channel Using a Matrix Element Method and *in situ* Jet Energy Calibration"
T. Aaltonen *et al.* [CDF Collaboration]
Phys. Rev. Lett. **105**, 252001 (2010) [arXiv:1010.4582 [hep-ex]]

"Improved Search for a Higgs Boson Produced in Association with $Z \to \ell^+\ell^-$ in proton antiproton Collisions at $\sqrt{s} = 1.96$ TeV"
T. Aaltonen *et al.* [CDF Collaboration]
Phys. Rev. Lett. **105**, 251802 (2010) [arXiv:1009.3047 [hep-ex]]

"Search for the supersymmetric partner of the top quark in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV"
T. Aaltonen *et al.* [CDF Collaboration]
Phys. Rev. D **82**, 092001 (2010) [arXiv:1009.0266 [hep-ex]]

"Updated Search for the Flavor-Changing Neutral-Current Decay $D^0 \to \mu^+\mu^-$"
T. Aaltonen *et al.* [CDF Collaboration]
Phys. Rev. D **82**, 091105 (2010) [arXiv:1008.5077 [hep-ex]]

"Measurement of the $WW + WZ$ Production Cross Section Using a Matrix Element Technique in Lepton + Jets Events"

T. Aaltonen *et al.* [CDF Collaboration]
Phys. Rev. D **82**, 112001 (2010) [arXiv:1008.4404 [hep-ex]]

"Direct Top-Quark Width Measurement CDF"
T. Aaltonen *et al.* [CDF Collaboration]
Phys. Rev. Lett. **105**, 232003 (2010) [arXiv:1008.3891 [hep-ex]]

"Diffractive W and Z Production at the Fermilab Tevatron"
T. Aaltonen *et al.* [CDF Collaboration]
Phys. Rev. D **82**, 112004 (2010) [arXiv:1007.5048 [hep-ex]]

"Measurement of the $t\bar{t}$ Production Cross Section with an in situ Calibration of b-jet Identification Efficiency"
T. Aaltonen *et al.* [CDF Collaboration]
Phys. Rev. D **83**, 071102 (2011) [arXiv:1007.4423 [hep-ex]]

Fields of Study

Major Field: Physics

Studies in Experimental High Energy Particle Physics: Dr. Brian Winer

Table of Contents

List of Figures

List of Tables

Chapter 1

INTRODUCTION

Since the earliest moments in recorded history, mankind has been increasingly involved in the pursuit of understanding the fundamental properties and inner workings of the universe. This endeavor has led to a very thorough scientific understanding of the universe in which we live.

Particle physics is devoted to determining what fundamental entities make up the universe, as well as understanding their properties and interactions with one another. In ancient Greek times, it was believed that the fundamental entity of matter was the atom – a single unit of matter that could not be further broken into smaller pieces. This view persisted for thousands of years, and it was not until the discovery of the electron by J. J. Thomson in 1897 [44] that evidence for the atom being a composite object began to mount. Further experiments by Ernest Rutherford confirmed the composite nature of atoms by scattering alpha (^{4}He) particles off a sheet of gold foil only several atoms thick[40]. It was expected that the alpha particles would not be affected by the gold foil and simply pass right through. The actual result of the experiment was quite different, in that some alpha particles were scattered at large angles, directly back from their incident directions. This led Rutherford to postulate that atoms were composed of a central nucleus, containing the majority of the mass of the atom. Rutherford also stated that most of the rest of the atom was empty space.

It was not until the development of quantum mechanics in the early 20$^{\text{th}}$ century that a complete description of the atom and its behavior was realized. Quantum mechanics, with its probabilistic nature, was the only way to explain some observed phenomena, such as the photoelectric effect[22] and emission spectra of elements.

As advances in technology occurred, the components of atoms were probed further. The protons and neutrons in the nucleus were found to be composite objects themselves, being made up of individual quarks. Further studies continue to this day to probe the structure of quarks and determine if there is yet another layer waiting to be revealed.

Quantum mechanics evolved further into the study of quantum field theory, in which individual particles are treated as localized excitations in a physical field permeating the universe. These different fields for each particle type can interact with one another, as well as determine how particles travel through the universe. The set of fields and interactions that describe the universe in which we live is known as the Standard Model (SM). The SM can predict and explain an amazing array of physical aspects of the universe, from the production and decay rates of particles, to interactions between particles in collisions, as well as a wide variety of astrophysical phenomena, such as cosmic rays.

Even with the Standard Model being an immensely successful description of the universe, there are still many questions which remain to be answered. The SM does not include a description of gravity – the theory of microscopic interactions, quantum mechanics, and macroscopic interactions, Einstein's theory of general relativity, have yet to be reconciled. Perhaps in the future a successful theory of quantum gravity will be developed. Additionally, while the SM predicts massless neutrinos, it has been observed that neutrinos do, in fact, have a small mass[26]. Several theories exist to explain the small neutrino mass, but it is difficult to rectify these theories with the existing SM.

Even more fundamental questions can be asked, addressing the origin of particle masses in the Standard Model. We observe and can measure the masses of particles, but do not have a fundamental explanation as to why particles have mass, why different particles have different masses, and by what mechanism particles obtain their masses. One theory attempts to answer these questions by adding an additional field to the SM. This is known as the Higgs mechanism, postulated by Peter Higgs and others in 1964 [31, 24, 28]. The Higgs mechanism, when added to the SM, acts in such a way to give particles their mass in the theory. This Higgs mechanism also postulates a new particle, the Higgs boson, which is the physical quantum of the new Higgs field introduced in the theory. More detail will be given on the theory of the SM and

Higgs mechanism in Chapter 2.

This thesis presents a search for the Higgs boson, utilizing a specific process where the Higgs boson is produced in association with a Z boson decaying to two muons. We only consider Higgs bosons which decay to a pair of b-quarks. We use the CDF detector, located at the Fermilab Tevatron, both of which are described in Chapters 5 and 6. We then describe the analysis details in subsequent chapters. We present results in the form of limits on the SM predicted cross section for ZH production times the $H \to b\bar{b}$ branching ratio in Chapter 11. Because the Higgs production cross section is so small, this analysis is not yet sensitive enough to observe evidence for the Higgs boson. With the accumulation of additional integrated luminosity, we may be able to observe the presence of a signal.

The Tevatron has been at the 'energy frontier' for many years, but operations there ceased in September 2011. The energy frontier has since moved overseas to the LHC. With increased beam energies, the cross section for Higgs production increases, and less data is required to obtain similar sensitivities. The Higgs boson, if it exists, most likely will be observed at the LHC, where the CMS and ATLAS experiments take data with beam energies 3.5 times that of the Tevatron. Whether the Higgs boson is or is not found in the near future, any information from experiments relating to the origin of particle masses will be important to understand the fundamental nature of our universe.

Chapter 2

The Standard Model of Particle Physics

The Standard Model (SM) of particle physics is one of the most successful scientific theories ever proposed. In this Chapter, we describe the theoretical details of the Standard Model relating to this analysis.

2.1 Particle Content

There are two types of particles in the SM: fermions and bosons. Fermions, having half-integer units of spin, make up matter and antimatter. Bosons, having integer units of spin, are force-carrying particles and are responsible for the interactions between particles.

Matter Particles

There are two types of fermions in the SM: quarks and leptons. Both quarks and leptons have spin $\frac{1}{2}$, and come in three generations. The three generations are essentially copies of one another, and the masses of particles increase for later generations. Equation 2.1 shows the three generations of leptons. There are three charged leptons, the electron (e^-), the muon (μ^-), and the tau (τ^-). The electron is a component of atoms, while the muon and tau are not found in ordinary matter. Muons incident on the atmosphere of the Earth are fairly common, as they are a decay product of cosmic rays.

$$\begin{pmatrix} e \\ \nu_e \end{pmatrix} \begin{pmatrix} \mu \\ \nu_\mu \end{pmatrix} \begin{pmatrix} \tau \\ \nu_\tau \end{pmatrix} \tag{2.1}$$

Lepton	**Symbol**	**Charge** (e)	**Spin**	**Mass** (MeV/c^2)
Electron	e^-	-1	1/2	0.511
Muon	μ^-	-1	1/2	106
Tau	τ^-	-1	1/2	1 777
Electron Neutrino	ν_e	0	1/2	
Muon Neutrino	ν_μ	0	1/2	$< 2 \cdot 10^{-6}$
Tau Neutrino	ν_τ	0	1/2	

Table 2.1: Properties of leptons in the Standard Model.

The neutrinos, of which there is one for each corresponding charged lepton, carry no electric charge, and have an extremely low mass. Table 2.1 summarizes the properties of the leptons.

Equation 2.2 shows the three generations of quarks, with two 'flavors' per generation. The up (u) and down (d) quarks comprise the protons and neutrons in atoms.

$$\begin{pmatrix} u \\ d \end{pmatrix} \begin{pmatrix} c \\ s \end{pmatrix} \begin{pmatrix} t \\ b \end{pmatrix} \tag{2.2}$$

The other types of quarks, charm (c), strange (s), top (t), and bottom (b) do not exist in the atoms of everyday matter, but can be produced through high energy particle collisions. Quarks have charges which are fractions of the electron charge. Quarks in the upper position of their respective doublets have a charge of $+\frac{2}{3}e$, while those in the lower position have $-\frac{1}{3}e$ charge (e is the fundamental unit of electric charge, $1e = +1.6 \cdot 10^{-19}$ C). Quarks also carry another type of charge, known as color charge. Combinations of quarks, or bound states, which make up larger pieces of matter must have zero net color charge. The combinations of quarks in bound states also result in integer values of their electric charge. The quarks do not exist as free particles – they must exist in these bound states with other quarks, with one exception. Because the mass of the top quark is very large (173 GeV/c^2), it decays very rapidly, even before it has a chance to form a bound state with other quarks.

Quarks form bound states, known as hadrons, with other quarks in one of two ways. Mesons are bound states of one quark and one anti-quark ($q\bar{q}$). Common examples are the pion ($\pi^0 \equiv u\bar{d}$) and the kaon ($K^+ \equiv u\bar{s}$). The second class of hadrons are bound states of three quarks (qqq), called baryons. The most common baryons are the proton (uud) and neutron (udd)

Quark	**Symbol**	**Charge** (e)	**Spin**	**Mass** (MeV/c^2)
Up	u	$+2/3$	$1/2$	1.5-3.3
Down	d	$-1/3$	$1/2$	3.5-6.0
Charm	c	$+2/3$	$1/2$	105
Strange	s	$-1/3$	$1/2$	1270
Top	t	$+2/3$	$1/2$	$1.73 \cdot 10^5$
Bottom	b	$-1/3$	$1/2$	4200

Table 2.2: Properties of quarks in the Standard Model.

which make up the atoms in matter. A wide variety of mesons and baryons are known, each having a different quark composition. Table 2.2 summarizes the properties of the quarks in the Standard Model.

Force-Carrying Particles

The quarks and leptons make up the matter we observe in the universe, but there also exist interactions between these particles. In the Standard Model, these interactions are mediated by exchanging force-carrying particles. These force-carrying particles are bosons, having integer spin. There are different bosons to transmit each of the four fundamental forces. The electromagnetic force is mediated by the photon, a massless particle that is also the quantum component of the wave-particle duality of light. The weak nuclear force has two associated bosons, the $W^{\pm}$ and Z^0. The SM unifies the electromagnetic and weak interactions into a single force. Finally, the strong interaction, which keeps quarks in bound states within hadrons, is mediated by gluons, another type of massless boson. Table 2.3 summarizes the properties of the force-carrying particles in the Standard Model. The allowed interactions of the SM are shown in Figure 2.1, where allowed interactions between particles are shown by drawing a line between the various particle types. Bosons interact with fermions, but bosons can sometimes interact with the other bosons in the SM as well.

The Standard Model predicts an additional boson, the Higgs boson. The Higgs boson was postulated to explain the origin of particle masses in the SM, which will be detailed in Section 2.4.

Particle	Symbol	Charge (e)	Spin	Mass (GeV/c^2)
Photon	γ	0	1	0
W boson	$W^{\pm}$	± 1	1	80.4
Z boson	Z	0	1	91.2
Gluon	g	0	1	0

Table 2.3: Properties of the force-carrying particles in the Standard Model.

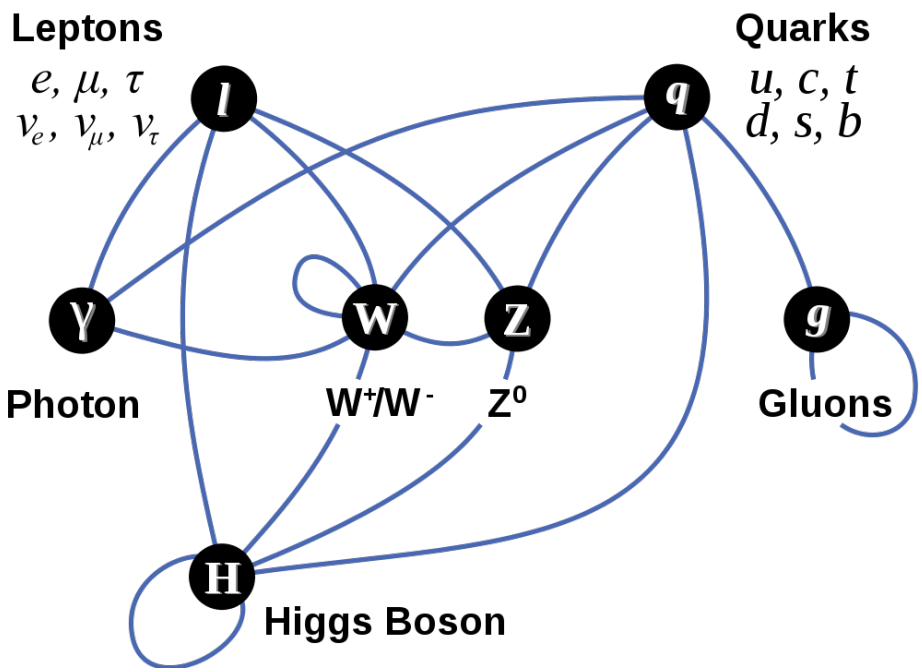


Figure 2.1: Allowed interactions between particles in the Standard Model. Particle types that can interact are shown connected with lines, while those that cannot interact are not connected. Lines that loop back to their starting position represent particles that can interact with themselves through self-couplings.

2.2 Quantum Field Theory and Gauge Symmetries

The theory of the Standard Model is grounded in the mathematics of quantum field theory. This theory uses the Lagrangian formalism to describe particle propagation and interaction. For example, a fermionic field $\psi(x)$ obeys the Dirac Lagrangian, shown in Equation 2.3.

$$\mathscr{L} = \overline{\psi}(i\gamma^\mu \partial_\mu - m)\psi \tag{2.3}$$

Here, γ^μ are the Dirac matrices, and $\overline{\psi} \equiv \psi^\dagger \gamma_0$ is the adjoint fermionic field. m represents the mass of the field. From the Lagrangian, the equations of motion can be derived using the Euler-Lagrange equation for fields, Equation 2.4.

$$\partial_\mu \left(\frac{\partial \mathscr{L}}{\partial(\partial_\mu \psi)} \right) - \frac{\partial \mathscr{L}}{\partial \psi} = 0 \tag{2.4}$$

Applying these equations to the Lagrangian in Equation 2.3 gives the equation of motion for a fermionic field, the Dirac equation:

$$i\gamma^\mu \partial_\mu \psi - m\psi = 0. \tag{2.5}$$

We expect the fields to obey certain symmetry laws which are found in nature. For example, the behavior of these fields must remain the same under translations and rotations. In the Lagrangian of Equation 2.3, there is an inherent global symmetry. If the field is modified by a constant phase, $\psi \rightarrow \psi \cdot e^{i\alpha}$, the Lagrangian remains the same, as the phases cancel one another. This is an example of a global symmetry, since α is a constant and does not depend on any position. Since physical particles are the localized excitations of these fields, and particles obey the same symmetry laws, we expect that these symmetries should be obeyed at each point in spacetime, meaning the parameter α should be a function of spacetime coordinates, $\alpha \rightarrow \alpha(x)$. If we try to impose this local symmetry, $\psi \rightarrow \psi \cdot e^{i\alpha(x)}$, the Lagrangian does not remain unchanged, as is seen in Equation 2.6.

$$\mathscr{L}[\psi \rightarrow \psi \cdot e^{i\alpha(x)}] = \mathscr{L} - \overline{\psi}(\gamma^\mu \partial_\mu \alpha(x))\psi \tag{2.6}$$

This additional term spoils the local symmetry of the Lagrangian. To keep the Lagrangian

invariant under this transformation, another field must be introduced, a gauge field A_μ. The covariant derivative must also become a gauge-covariant derivative, $\partial_\mu \rightarrow D_\mu \equiv \partial_\mu - iqA_\mu$. If the gauge field transforms under this local symmetry as $A_\mu \rightarrow A_\mu + \frac{1}{q}(\partial_\mu \alpha(x))$, then the new Lagrangian of Equation 2.7 remains invariant and possesses a gauge symmetry.

$$\mathscr{L} = \overline{\psi}(i\gamma^\mu \partial_\mu - m)\psi + q\overline{\psi}\gamma^\mu A_\mu \psi \tag{2.7}$$

The second term in Equation 2.7 represents an interaction between the fermion and the gauge field, with a coupling strength of q. Because this Lagrangian has just one degree of freedom, $\alpha(x)$, in the transformation, the gauge symmetry group for this Lagrangian is $U(1)$. It is using this method that the Lagrangian of the Standard Model theory was developed. In fact, by adding a gauge-invariange kinetic term for the gauge field A_μ, we arrive at the Lagrangian for quantum electrodynamics, shown in Equation 2.8, where $F_{\mu\nu} \equiv \partial_\mu A_\nu - \partial_\nu A_\mu$.

$$\mathscr{L}_{\text{QED}} = \overline{\psi}(i\gamma^\mu \partial_\mu - m)\psi + q\overline{\psi}\gamma^\mu A_\mu \psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} \tag{2.8}$$

The Standard Model describes the three fundamental forces important in particle physics – the strong, weak, and electromagnetic forces. In fact, each of these forces can be described mathematically by a different gauge symmetry group. The electromagnetic force possesses a $U(1)$ symmetry, the weak force a $SU(2)$ symmetry, and the strong force a $SU(3)$ symmetry. The combined symmetry group of the theory of the Standard Model is then $SU(3) \otimes SU(2) \otimes U(1)$. Each of the symmetry groups has a set of gauge fields as a result of imposing the local gauge symmetry into the Standard Model Lagrangian. For the electromagnetic sector of the SM, there is just one gauge field, B_μ. The weak sector has three gauge fields W_μ^i $(i = 1, 2, 3)$, as there are three degrees of freedom in the $SU(2)$ symmetry group. Similarly, there are eight degrees of freedom in the $SU(3)$ group, so there are eight gauge fields G_μ^j $(j = 1, \ldots, 8)$ mediating strong interactions in the Standard Model.

2.3 Constructing the Electroweak Lagrangian

The electroweak Lagrangian is constructed out of the fermionic fields corresponding to matter particles, and the bosonic fields corresponding to the force-carrying particles. The electroweak Lagrangian unifies the electromagnetic theory described by the $U(1)$ symmetry with the weak sector described by an $SU(2)$ symmetry. Therefore, the electroweak Lagrangian possesses the $SU(2) \otimes U(1)$ gauge symmetry.

The $SU(2)$ symmetry group requires three generators T^i $(i = 1, 2, 3)$, whereas the $U(1)$ group has just a single generator Y, the weak hypercharge. Instead of the electric charge used in the example above, the weak hypercharge is used to unify the electromagnetic and weak theories. The weak hypercharge is related to the usual electric charge by the Gell-Mann-Nishijima formula of Equation 2.9. The quantity I_3 is third component of the particle weak isospin vector $\vec{I}$.

$$Q = I_3 + \frac{Y}{2} \tag{2.9}$$

The generators T^i are 3×3 matrices, with $T^i \equiv \tau^i/2$ and τ^i representing the standard Pauli matrices.

The covariant derivative, shown in Equation 2.10 is formed in a similar way to that in the previous section. It has the correct form to ensure symmetry under local gauge transformations. The fields B_μ and W^i_μ are the gauge fields of $U(1)$ and $SU(2)$, respectively. Additionally, there are now two coupling strengths, g' for the $U(1)$ interaction and g for the $SU(2)$ interaction.

$$D_\mu = \partial_\mu - \frac{i}{2} g' Y B_\mu - i g W^i_\mu T^i \tag{2.10}$$

It is useful to discuss the Lagrangian in terms of left-handed and right-handed helicity states. The left-handed helicity state for leptons consists of an isospin doublet consisting of one neutrino species and its associated lepton. The left-handed quarks also form isospin doublets; one per generation. The right-handed helicity states only exist for the quarks and charged leptons, as there is no right-handed component for the neutrino. The field content for one generation of particles in this helicity basis is listed in Equations 2.11 and 2.12. For the remainder of this section, only the first generation of leptons will be used to demonstrate the construction of the

electroweak Lagrangian. Of course, one must sum over all fermion species to form the complete Lagrangian, with the different weak hypercharges Y and isospins I_3 for each fermion species.

$$\psi_L = \begin{pmatrix} \nu_L \\ \ell_L \end{pmatrix}, \begin{pmatrix} u_L \\ d_L \end{pmatrix}, \ldots \tag{2.11}$$

$$\psi_R = \ell_R, u_R, d_R, \ldots \tag{2.12}$$

As the right-handed helicity states are singlets under the $SU(2)$ group ($\vec{I} = 0$), the final term of 2.10 will not contribute. From this point, we can expand the Lagrangian as follows (using just one left-handed doublet and one right-handed singlet for an example):

$$\mathscr{L}_{EW} = \overline{\psi}_L(i\gamma^\mu D_\mu)\psi_L + \overline{\psi}_R(i\gamma^\mu D_\mu)\psi_R \tag{2.13}$$

$$= i\overline{\psi}_L\gamma^\mu\partial_\mu\psi_L + i\overline{\psi}_R\gamma^\mu\partial_\mu\psi_R \tag{2.14}$$

$$+ \frac{1}{2}g'Y\overline{\psi}_L\gamma^\mu B_\mu\psi_L + \frac{1}{2}g'Y\overline{\psi}_R\gamma^\mu B_\mu\psi_R \tag{2.15}$$

$$+ g\overline{\psi}_L\gamma^\mu[W_\mu^1 T^1 + W_\mu^2 T^2 + W_\mu^3 T^3]\psi_L \tag{2.16}$$

It is convenient to combine the first and second components of W^i into a set of charged fields, defined by Equation 2.17.

$$W_\mu^\pm = \frac{1}{\sqrt{2}}\left(W_\mu^1 \pm i \cdot W_\mu^2\right) \tag{2.17}$$

Using this substitution and expading the $SU(2)$ term (Equation 2.16) of the Lagrangian, the charged current interactions can be derived from the Lagrangian:

$$\mathscr{L}_{CC} = \frac{ig}{\sqrt{2}}[\overline{\nu}_L\gamma^\mu W_\mu^- \ell_L + \overline{\ell}_L\gamma^\mu W_\mu^+ \nu_L] \tag{2.18}$$

As Equation 2.18 describes the interactions via the $W^\pm$ bosons, there are three terms remaining in the Lagrangian – these terms describe the neutral current interactions in the Standard Model. These interactions are mediated both by the photon and Z boson. The electroweak theory must describe electromagnetic interactions as well as weak neutral current interactions. In order to give the correct electromagnetic couplings, the remaining gauge fields, B_μ and W_μ^3, mix according to Equation 2.19 to give the photon field A_μ and the Z boson field, Z_μ. The angle

θ_W is the weak mixing angle, and is a free parameter of the Standard Model.

$$\begin{pmatrix} A_\mu \\ Z_\mu \end{pmatrix} = \begin{pmatrix} \cos\theta_W & \sin\theta_W \\ -\sin\theta_W & \cos\theta_W \end{pmatrix} \begin{pmatrix} B_\mu \\ W^3_\mu \end{pmatrix} \tag{2.19}$$

After making substitutions in favor of the new fields A_μ and Z_μ, the piece of the Lagrangian describing neutral current interactions is obtained. Because electromagnetic interactions do not occur between neutrinos, the coefficient of $\overline{\nu}_L \gamma^\mu A_\mu \nu_L$ must vanish. This gives the condition $g' \cos\theta_W = g \sin\theta_W = q$, the electric charge from the QED Lagrangian of Equation 2.7. The electromagnetic portion of the neutral current interactions is the same for left-handed and right-handed fields, shown in Equation 2.20.

$$\mathscr{L}^{\mathrm{EM}}_{NC} = g \sin\theta_W (\overline{\ell}_L \gamma^\mu A_\mu \ell_L + \overline{\ell}_R \gamma^\mu A_\mu \ell_R) \tag{2.20}$$

The portion of the Lagrangian describing interactions with the Z boson is more complicated, as the interactions for left- and right-handed fields have differing strengths. This piece of the Lagrangian is shown in Equation 2.21.

$$\mathscr{L}^{Z}_{NC} = \frac{1}{2}(g' \sin\theta_W \mp g \cos\theta_W) \begin{pmatrix} \overline{\ell}_L \\ \overline{\nu}_L \end{pmatrix} \gamma^\mu Z_\mu \begin{pmatrix} \ell_L \\ \nu_L \end{pmatrix} - g' \sin\theta_W \overline{\ell}_R \gamma^\mu Z_\mu \ell_R \tag{2.21}$$

The three terms $\mathscr{L}_{CC}$, $\mathscr{L}^{\mathrm{EM}}_{NC}$, and $\mathscr{L}^{Z}_{NC}$ describe the electroweak interactions of the Standard Model. Just one more term is missing from the Lagrangian – the kinetic terms for the new gauge boson fields $W^\pm_\mu$, A_μ, and Z_μ. These terms have structures similar to the last term in Equation 2.8. All of these pieces form the complete picture of electroweak interactions in the Standard Model, but there is still one crucial ingredient missing in the electroweak sector – mass.

2.4 Higgs Mechanism and Spontaneous Symmetry Breaking

Apart from the kinetic and interaction terms in the electroweak Lagrangian, there are no mass terms present for the fermions or the gauge bosons. In fact, adding a mass term of the form $m\overline{\psi}\psi$ would spoil the invariance of the Lagrangian under the gauge symmetry, as it mixes the left- and right-handed components of the ψ field. Therefore, the masses of particles in

the Standard Model cannot be added simply by writing in mass terms. There must be another mechanism at work which accounts for the generation of particle masses in the Standard Model. This mechanism is known as the Higgs mechanism and was named for Peter Higgs, although Brout, Englert, Guralnik, Hagen, and Kibble developed the mechanism independently of Higgs in 1963-64[31, 24, 28].

The Higgs mechanism requires the introduction of a new complex scalar $SU(2)$ doublet to the Standard Model,

$$\Phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix}. \tag{2.22}$$

This complex doublet is described by the Lagrangian given in Equation 2.23. The field has both a kinetic term and a potential term $V(\Phi)$. There are two arbitrary parameters of the potential, ξ and λ, which will be crucial to the Higgs mechanism. The gauge covariant derivative D_μ is that of Equation 2.10, containing the gauge fields of the $SU(2)$ and $U(1)$ groups.

$$\mathscr{L}_{\text{Higgs}} = D_\mu \Phi^\dagger D^\mu \Phi - \underbrace{(\xi^2 \Phi^\dagger \Phi + \lambda (\Phi^\dagger \Phi)^2)}_{V(\Phi)} \tag{2.23}$$

This potential term has an interesting structure. For $\xi^2 > 0$, the vacuum expectation value of the field, $\langle 0|\Phi|0\rangle$ will be zero, as the minimum of the potential is for $|\Phi| = 0$. However, if $\xi^2 < 0$, the vacuum expectation value becomes non-zero, and the minimum of $V(\Phi)$ is now given by Equation 2.24. This case is shown in Figure 2.2.

$$|\Phi| = \sqrt{\frac{-\xi^2}{2\lambda}} = \frac{v}{\sqrt{2}} \tag{2.24}$$

This results in a non-zero expectation value for the field Φ. To preserve the conservation of electric charge, the electromagnetic $U(1)$ symmetry must remain in the Lagrangian. Therefore, the non-zero vacuum expectation value is applied to the neutral component of this complex scalar field, while the charged component attains a zero vacuum expectation value, as in Equation 2.25.

$$\langle 0|\Phi|0\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix}. \tag{2.25}$$

At this point, the Higgs Lagrangian of Equation 2.23 can be expanded, with Φ being evaluated at the vacuum expectation value $\langle \Phi \rangle_0$. The complex doublet that was originally introduced can

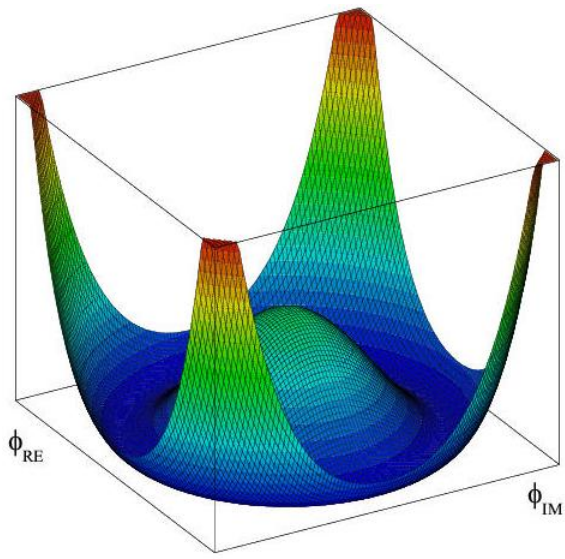


Figure 2.2: The shape of the Higgs potential $V(\Phi)$ for $\xi^2 < 0$. It has a minimum located away from the origin, meaning the Higgs field acquires a non-zero vacuum expectation value.

now be expanded about this vacuum expectation value, as in Equation 2.26. The new field ϕ_H is a physical scalar field that will interact with the other Standard Model fields.

$$\Phi' = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ v + \phi_H \end{pmatrix} \tag{2.26}$$

Evaluating the Higgs Lagrangian with Φ' gives terms describing interactions between the new Higgs field ϕ_H and gauge fields, as well as self-interactions of the Higgs field. There are also additional terms with the vacuum expectation value v which are mass terms for the gauge bosons, from the potential term $V(\Phi)$. Equation 2.27 shows the quadratic gauge field terms that result from the expansion of the first term of $\mathscr{L}_{\text{Higgs}}$.

$$\begin{aligned} D_\mu \Phi'^\dagger D^\mu \Phi' &= \left| [\partial_\mu - \frac{i}{2} g' Y B_\mu - i g W_\mu^i T^i] \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + \phi_H \end{pmatrix} \right|^2 \\ &= \frac{1}{2}\partial_\mu \phi_H \partial^\mu \phi_H - \frac{g'^2 v^2}{8} B_\mu B^\mu - \frac{g^2 v^2}{4} W_\mu^+ W^{-\mu} - \frac{g^2 v^2}{8} W_\mu^3 W^{3\mu} + \text{other terms} \end{aligned} \tag{2.27}$$

With the substitutions of A_μ and Z_μ for B_μ and W_μ^3, one can obtain the mass terms for the gauge bosons. The potential term $V(\Phi')$ also contains a mass term for the Higgs field. The mass terms

in the Higgs portion of the Standard Model Lagrangian are shown in Equation 2.28.

$$\mathscr{L}_{\text{Higgs}}^{\text{Mass}} = \frac{g^2 v^2}{4} W_\mu^+ W^{-\mu} + \frac{g^2 v^2}{8\cos^2\theta_W} Z_\mu Z^\mu - \xi^2 \phi_H^2 \tag{2.28}$$

From these terms, the vector boson masses $m_W = gv/2$ and $m_Z = gv/(2\cos\theta_W)$ are determined. The photon field A_μ does not have a quadratic mass term, therefore remains massless after the Higgs mechanism takes effect. This also preserves the $U(1)$ electromagnetic symmetry of the Lagrangian. The Higgs mass can also be determined from the final term in Equation 2.28 as $m_H = \sqrt{2}\xi$. Although the Higgs mass cannot be directly calculated in the Standard Model, due to the unknown parameters ξ and λ, the value of the vacuum expectation value $\langle\Phi\rangle_0$ has been indirectly measured to be 246 GeV, through precise experimental measurements and calculations of muon decay properties[45]. These precise calculations lead to a value for the Fermi coupling constant, $G_F = g^2\sqrt{2}/8m_W^2$, where g is the $SU(2)$ weak coupling constant. From this value, the Higgs vacuum expectation value can be obtained according to $v = (\sqrt{2}G_F)^{1/2}$.

This new physical particle, the scalar Higgs boson, which was introduced to the theory to provide mass terms for the gauge bosons, is the only remaining particle predicted by the Standard Model but yet undiscovered. The search for the Higgs boson has been at the forefront of the physics programs at collider experiments for many years, and this thesis presents a search for the Higgs boson utilizing a specific production mechanism and specific decay mode of the Higgs boson.

Chapter 3
HIGGS BOSON PHENOMENOLOGY

Properties of the Higgs boson, such as production mechanisms, decay modes, and couplings, have been studied in detail since the Higgs mechanism was proposed. This chapter details the phenomenology of the Higgs boson, highlighting the most promising channels for its discovery.

3.1 Higgs Boson Production

The Higgs boson can be produced in collisions through a variety of mechanisms, each with varying cross sections. Some of these mechanisms are more useful for searching for the Higgs boson at the Tevatron, depending on the signal and background acceptance rates. Figure 3.1 shows the Feynman diagrams for the five main production mechanisms at the Tevatron.

One of the simplest Higgs production mechanisms is that of quark fusion, also called direct Higgs production. Because the Higgs couples more strongly to more massive objects, the largest contribution to this diagram is a pair of b-quarks in the initial state. The cross section for the $b\bar{b} \rightarrow H$ process at the Tevatron ($\sqrt{s} = 1.96$ TeV) is approximately 10 fb for a Higgs mass $m_H = 120$ GeV/c^2. This cross section is one of the lowest for Higgs production at the Tevatron, since the b quarks in the initial state must come from the quark sea in the proton and antiproton.

The gluon fusion production mechanism has the highest cross section of any Higgs production mode at the Tevatron. The cross section for this process is approximately 1000 fb for $m_H = 120$ GeV/c^2. This process uses two gluons in the initial state, which split into quark pairs and form a quark loop. The Higgs is then produced through the fusion of the quarks in the

internal loop. The most likely candidate for the quark running in the internal loop is the top quark, due to its large mass and high Yukawa coupling to the Higgs boson.

The vector boson fusion process has a smaller cross section than that of the gluon fusion process, at only 80 fb (for $m_H = 120$ GeV/c^2). It ranks third among the five Higgs production mechanisms. This process consists of two quarks in the initial state that each radiate a vector boson (either a $W^{\pm}$ or Z, but each quark must radiate the same vector boson type). The two vector bosons then fuse to produce the Higgs boson. The two quarks from the initial state are still present, and are generally detected as jets produced with small angles to the incoming beams.

Similar to the vector boson fusion process, the quark-pair association process gives a final state with a Higgs boson and two additional quarks. The initial state is different, with two gluons each splitting into a quark-antiquark pair. Two of the quarks from the gluon splittings fuse to form the Higgs boson. This process has the smallest cross section of any Higgs production mechanism at the Tevatron, at only 5 fb. All quark flavors can contribute to this process, although the $gg \to t\bar{t}H$ process has the largest cross section.

Finally, the associated production mechanism is one of the most promising signatures for Higgs searches at the Tevatron. With the second highest cross section of 300 fb for $m_H = 120$ GeV/c^2, the Higgs boson is produced in association with a vector boson, $W^{\pm}$ or Z. The Higgs boson is actually radiated from the vector boson that is produced from the two-quark initial state. This radiation is the reason for the alternate name of this production mode, *Higgstrahlung*. The analysis described in this thesis uses this production mechanism, where the Higgs boson is produced in association with a Z boson.

Figure 3.2 shows a comparison of the cross sections of these different Higgs production mechanisms at the Tevatron. As the Higgs mass increases, the cross sections of the various processes decrease. The next section describes the various decay modes of the Higgs boson. It is important to choose a suitable combination of production process and decay mode to optimize the sensitivity of Higgs searches.

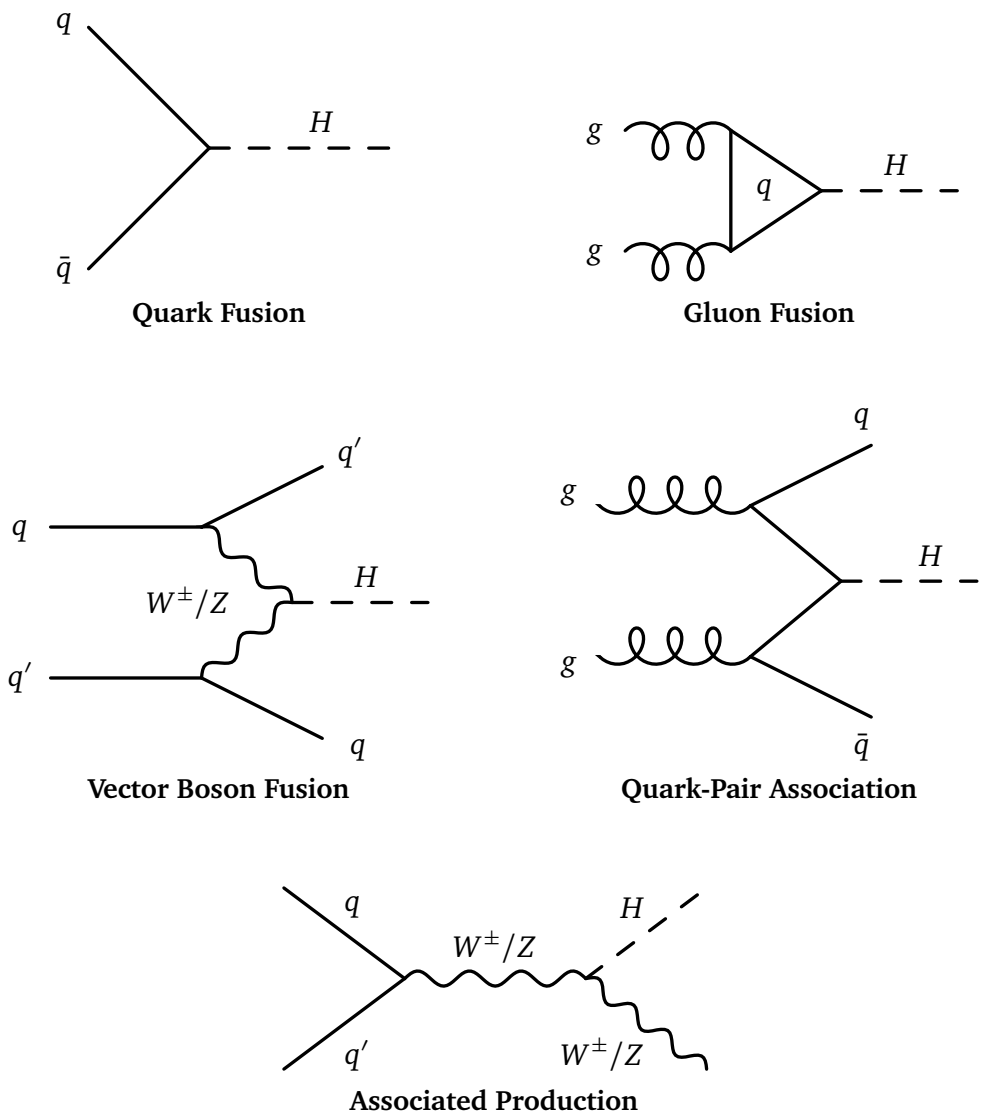


Figure 3.1: Feynman diagrams for the five main Higgs production mechanisms.

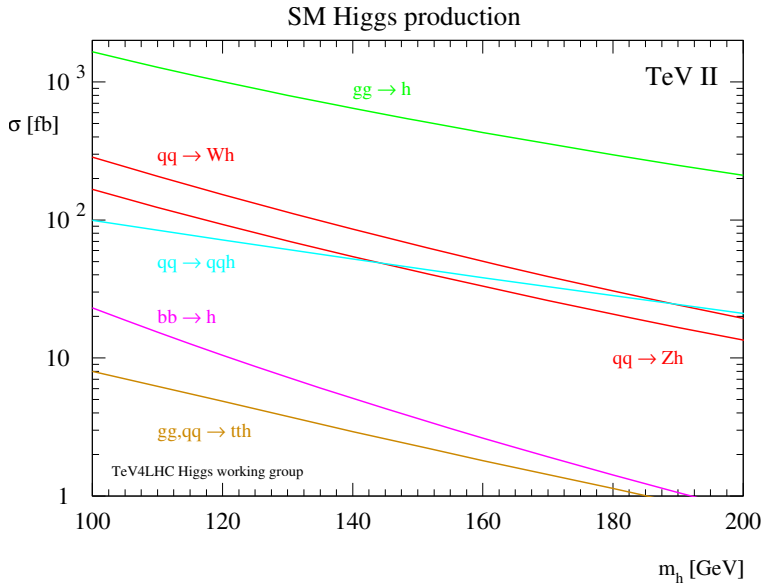


Figure 3.2: Cross sections of the various Higgs production mechanisms at the Tevatron, shown as a function of the Higgs mass.

3.2 Higgs Boson Decay

After the Higgs boson is produced via any of the above-mentioned mechanisms, it can decay in a variety of allowed modes, each with varying rates as a function of the mass of the Higgs boson and consisting of many different final state particles. The allowed decay modes, and their associated rates, called branching ratios, are shown in Figure 3.3 as a function of the Higgs boson mass.

At the Tevatron, searches for the Higgs boson are generally divided into two regimes: low-mass Higgs searches in the range $100 < m_H < 135$ GeV/c^2, where the $H \to b\bar{b}$ decay mode is predominantly utilized, and high-mass Higgs searches in the range $135 < m_H < 200$ GeV/c^2, using the $H \to W^+W^-$ decay mode.

Based on the available decay modes and production mechanisms of the Higgs boson, one can choose from a wide variety of final state combinations to search for the Higgs. Different production mechanisms are better suited to be chosen for different decay modes, and a certain decay mode might not provide the same level of sensitivity for each production mechanism. The

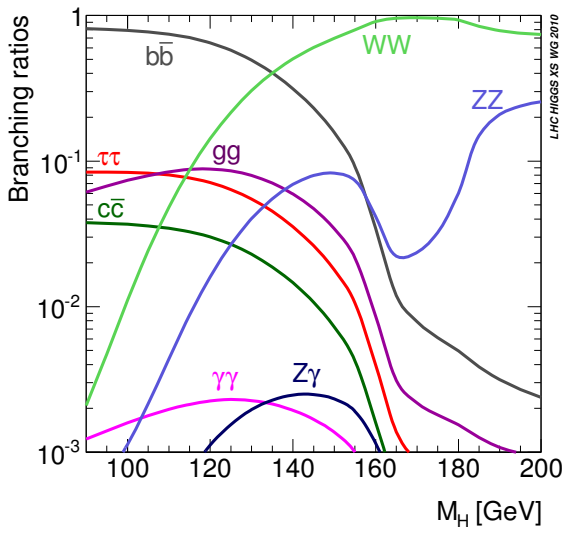


Figure 3.3: Branching ratios of the Higgs boson to each of its allowed decay products, as a function of the Higgs mass m_H. Higgs searches at the Tevatron generally use the $H \to b\bar{b}$ and $H \to W^+W^-$ decay modes, with additional sensitivity coming from the decay modes with smaller branching ratios.

main issue in determining the sensitivity of possible channels is exploring the relative signal and background production rates.

For example, while the gluon fusion production mechanism has the highest cross section, using this in combination with the $H \to b\bar{b}$ decay mode in the low-mass region is simply not feasible due to the enormous QCD dijet production. In this process, the Higgs is produced alone, and subsequently decays to two b-quarks, which appear as jets in the detector. The cross section for $b\bar{b}$ production at the Tevatron is many orders of magnitude higher than the $gg \to H$ cross section, eliminating all sensitivity for a Higgs search in $gg \to H \to b\bar{b}$. Figure 3.4 shows the relative cross sections of different background processes at the Tevatron.

To obtain the most sensitivity for the low-mass regime where the $H \to b\bar{b}$ decay mode dominates, it is common to choose a production mechanism with additional particles produced in association with the Higgs boson. For example, the most sensitive analyses in the low-mass regime are those which utilize the VH ($V = W^{\pm}, Z$) associated production mechanism. The decays of the additional vector boson produce distinct final states, and can be better distinguished

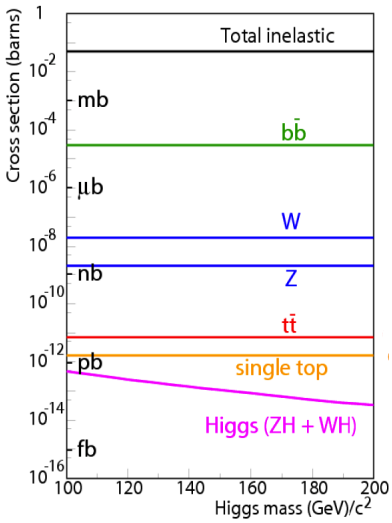


Figure 3.4: Comparison of Higgs cross section to background process cross sections at the Tevatron.

from the large amount of background events. Even still, the V+jets background contribution is a few orders of magnitude above the VH production rate, so it remains a challenge to design an analysis toward maximizing Higgs sensitivity. The three low-mass channels that are most sensitive are $ZH \to \nu\bar{\nu}b\bar{b}$, $ZH \to \ell^+\ell^- b\bar{b}$, and $WH \to \ell\nu b\bar{b}$. Additional channels can add sensitivity to this low mass regime due to their distinct final states, such as $H \to \gamma\gamma$ and $H \to \tau^+\tau^-$, even though their individual branching ratios are small (0.2% and 6.8%, respectively [21]).

In the high-mass regime, the dominant $H \to W^+W^-$ decay mode allows one the advantage of using the gluon fusion production mechanism, which has the highest cross section of any such mechanism at the Tevatron. The decays of the W bosons give the distinct final states, therefore the requirement of additional objects produced in association with the Higgs boson is not necessary for high-mass search channels. Different decays of the W bosons divide this region into several search channels, mainly divided into differing numbers of identified leptons. Additionally, the $H \to ZZ$ decay mode is utilized in this regime, with possible final states consisting of four leptons, two leptons and two jets, and four jets. Due to the dip in the $H \to ZZ$ branching ratio for $150 < m_H < 190$ GeV/c^2, these channels are not as sensitive as the channels utilizing the $H \to W^+W^-$ decay mode.

This thesis presents a search for the Higgs boson, using the ZH associated production mech-

anism and the $H \to b\bar{b}$ decay channel. The final state consists of two muons from the decay of the Z boson, and two b-quarks, which hadronize to form jets in the detector. The analysis is most sensitive in the low-mass regime, but we use Higgs mass hypotheses up to $m_H = 150$ GeV/c^2 in the analysis.

Chapter 4
Constraints on Higgs Properties

Since the Higgs boson has been postulated, it has been the goal for many experiments to discover the particle, or provide evidence limiting the range of possible masses of the particle. Both direct searches for the Higgs boson, as well as indirect measurements of electroweak parameters like the W boson mass and top quark mass provide information regarding the probable mass of the Higgs boson, if it does indeed exist. Additionally, some theoretical calculations can provide information as to the probable range of Higgs boson masses.

4.1 Theoretical Constraints

There exist several theoretical predictions of possible mass ranges for the Higgs boson. These are usually derived by finding some point where the theory of the Standard Model breaks down. We currently know that the Standard Model is not a complete description of nature, as it does not include a theory of gravity. It also has several unsolved problems, such as the hierarchy or fine-tuning problem, the lack of an adequade description of neutrino masses, and the lack of a dark matter candidate. This means that some new physics must enter into the theoretical picture, at some energy scale Λ. The scale of this new physics affects the theoretical bounds on allowed Higgs boson masses.

One theoretical constraint on the Higgs boson mass is that of triviality. The Higgs boson coupling parameter λ from Equation 2.23 must remain finite to maintain the interactions between the Higgs boson and other particles. If the λ parameter goes to infinity, which it does at some high energy scale, the Higgs boson will not interact, and the required spontaneous

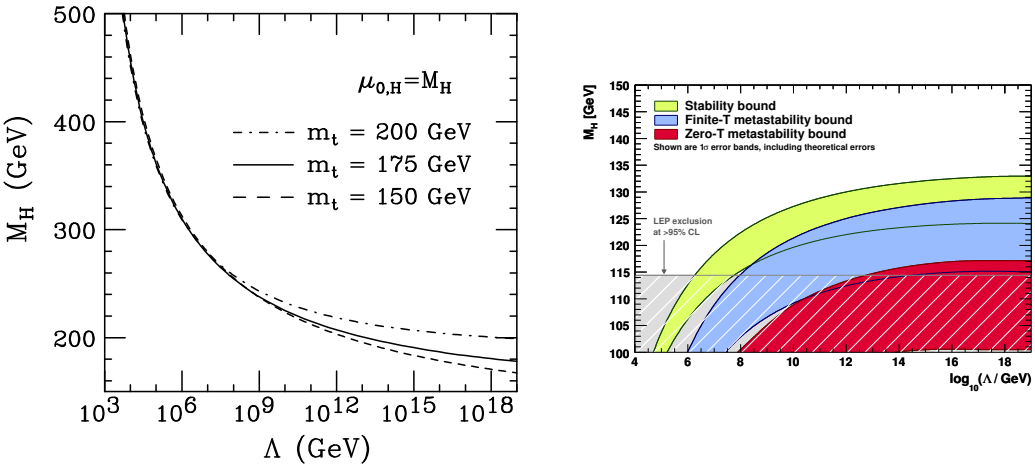


Figure 4.1: Bounds on the Higgs mass determined from theory, from the triviality constraint (left)[29] and the vacuum stability constraint (right)[23]. Allowed regions are those under the curve (left), and above the colored regions (right).

symmetry breaking does not occur. Therefore, the parameter λ must remain finite below the scale at which new physics enters the theory, Λ. Inserting this constraint into the field theory equations yields the bound on the Higgs boson mass shown in Equation 4.1[27]. Figure 4.1 shows the upper limit on the Higgs boson mass from the triviality constraint, as a function of the new physics scale Λ.

$$m_H^2 < \frac{8\pi^2 v^2}{3\ln(\Lambda^2/v^2)} \tag{4.1}$$

Besides the triviality constraint, there is another constraint which is applicable based on theoretical concerns. If the λ parameter in the Higgs potential were to become negative, the shape of the Higgs potential would change in such a way that spontaneous symmetry breaking would not occur. The Higgs field would acquire a vacuum expectation value exactly equal to zero, instead of the non-zero value v which is required for the theory to remain stable. This condition is known as the vacuum instability bound. Again, the lower limit on the Higgs masses for which this condition occurs is a function of the scale at which new physics is allowed to enter the theory. Figure 4.1 shows the lower limit of allowed Higgs masses based on this vacuum stability constraint, as a function of the new physics scale Λ[23]. There are several different scenarios determined from different theoretical predictions.

These theoretical constraints on the Higgs boson mass seem to favor the low-mass region, as many theorists predict that new physics will enter into the theory at energy scales of approximately 1 TeV. This probable low-mass region is something that both direct and indirect experimental constraints support.

4.2 Direct Experimental Constraints

One of the earliest experimental constraints on the Higgs boson came from the four experiments at LEP, an electron-positron collider that previously occupied the current LHC tunnel on the border of France and Switzerland. The four collaborations, ALEPH, DELPHI, L3, and OPAL performed searches for the Higgs boson using the $e^+e^- \rightarrow ZH$ associated production process. These experiments were operational until releasing their final results in 2003. After collecting a total of approximately 2.5 fb^{-1} of data, at a center-of-mass energy of $\sqrt{s} \geq 189$ GeV, the combination of analyses from each of the four collaborations resulted in a lower bound on the Higgs boson mass of 114.4 GeV/c^2, at the 95% confidence level[15]. Figure 4.2 shows the CL_s confidence level limit as a function of Higgs boson mass. The expected lower limit is $m_H > 115.3$ GeV/c^2, higher than the observed limit due to an excess in the data.

Since the final results of the LEP experiments in 2003, the Tevatron experiments CDF and D0 have been the main source of new results from Higgs boson searches. Both CDF and D0 have many independent analysis channels utilizing the various production mechanisms and decay modes of the Higgs described in Chapter 3. These individual analyses are combined in a statistically-rigorous method to produce an all-encompassing result of Higgs searches at the Tevatron. More details of this combination method, as well as of the individual analysis channels, are described in Chapter 11.

The most recent results from the Tevatron, released in Summer 2011, have expanded the range of excluded Higgs masses. The combined Tevatron results [43], which used up to 8.6 fb^{-1} of collected integrated luminosity, have excluded a Standard Model Higgs boson with $m_H < 108$ GeV/c^2 (this is a less stringent constraint than that from the LEP experiments), as well as Higgs bosons with $156 < m_H < 177$ GeV/c^2. Both of these reported exclusions are

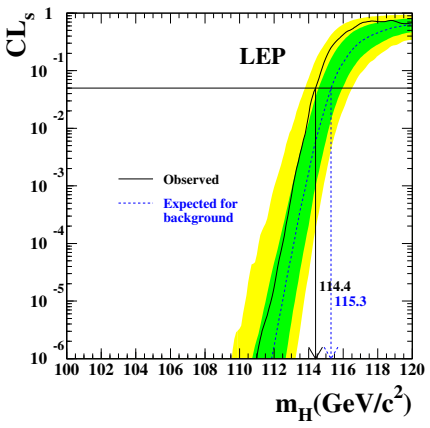


Figure 4.2: Results from the combination of Higgs searches at the four LEP experiments, in the form of confidence levels as a function of Higgs mass. The expected lower limit at the 95% confidence level on the mass of the Higgs boson is $m_H > 115.3$ GeV/c^2, while the observed limit is $m_H > 114.4$ GeV/c^2.

at the 95% confidence level. This exclusion range of high Higgs boson masses represents the first direct constraint since the LEP result. Figure 4.3 shows the most recent result of combined Higgs searches at the Tevatron.

Additionally, the LHC has begun to produce data at an ever-increasing rate. The first Higgs results from the LHC were reported in Summer 2011, with the ATLAS and CMS experiments each reporting separate combinations of analysis channels within the individual experiments. With just approximately 2 fb^{-1} of collected integrated luminosity, CMS reports three excluded mass ranges for the Standard Model Higgs boson, $145 < m_H < 216$ GeV/c^2, $226 < m_H < 288$ GeV/c^2, and $310 < m_H < 400$ GeV/c^2, at the 95% confidence level[1]. The ATLAS collaboration excludes several ranges of Higgs boson masses independently of CMS, $146 < m_H < 230$ GeV/c^2, $256 < m_H < 282$ GeV/c^2, and $296 < m_H < 459$ GeV/c^2, again at the 95% confidence level[2]. Figure 4.4 shows the most recent results for Higgs searches at CMS and ATLAS.

The LHC experiments mainly used the $H \rightarrow W^+W^-$ decay mode to obtain these excluded mass ranges. Due to the larger energy of the LHC beams and the increase in the Higgs boson production cross section, the LHC experiments have already eclipsed the excluded mass range obtained by the Tevatron experiments. However, the Tevatron is more sensitive to the channels

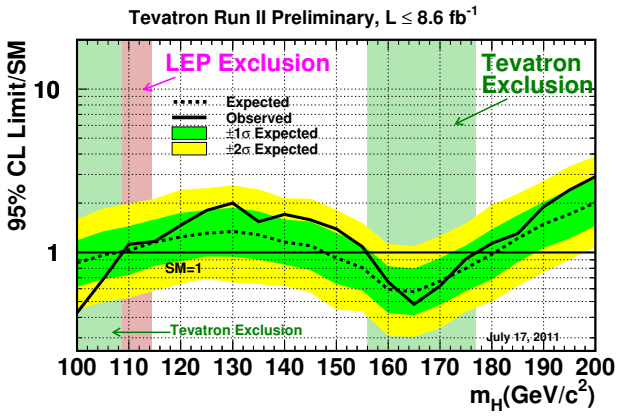


Figure 4.3: Combinations of Higgs searches at the Tevatron, shown in the form of expected and observed limits on the Higgs production cross section, as a ratio to the expected Standard Model value.

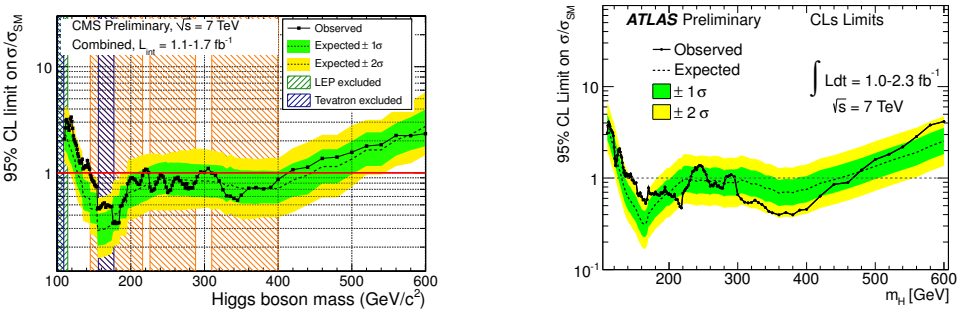


Figure 4.4: Combinations of Higgs searches at CMS (left) and ATLAS (right), shown in the form of expected and observed limits on the Higgs production cross section, as a ratio to the expected Standard Model value.

utilizing $H \rightarrow b\bar{b}$ decay modes, and will remain so for the next several years. Even though the LHC will continue to accumulate more high-quality data, the Tevatron will remain competitive due to the unique production mechanisms that are not feasible for study at the LHC.

4.3 Indirect Experimental Constraints

Since the Higgs mechanism is responsible for the breaking of electroweak symmetry, many measurements of electroweak parameters can provide constraints on the possible masses of the Higgs boson.

Two of the most important electroweak parameters are the masses of the W boson and top quark. Since the top quark has a predicted coupling to the Higgs boson of order unity, it has a strong influence on the Higgs sector of the Standard Model. Additionally, calculations of electroweak scattering processes receive corrections from Higgs loops which can be added to the Feynman diagrams. Therefore, measuring parameters of certain electroweak processes gives another handle into the Higgs sector of the Standard Model.

Figure 4.5 shows the dependence of the value of m_W predicted by the Standard Model as a function of Higgs mass. As the Higgs mass increases, the predicted W boson mass decreases [12]. Also shown in Figure 4.5 is the world average experimental measurement of m_W and its 1σ uncertainty bands. The experimental measurement favors a low mass Higgs boson.

The most recent measurement of the top quark mass, $m_t = 173.2 \pm 0.9$ GeV/c^2[34], also is used to constrain the Higgs boson mass. While the W boson mass decreases with increasing Higgs masses, the top quark mass has the opposite behavior, increasing with Higgs boson mass, as shown in Figure 4.5 (right).

The masses of the W boson and the top quark are only two of the electroweak parameters used in this way to constrain the Higgs boson mass. In the most recent result from the Gfitter Group[13], a total of 25 electroweak parameters are used as inputs to fit for the most likely Higgs mass value. These inputs include m_Z, Γ_Z, $\sin^2\theta_W$, m_b, m_c, m_s, as well as others including forward-backward asymmetry parameters and coupling constant measurements[13]. A fit using these parameters can return the most probable values for the Higgs mass. This fit can

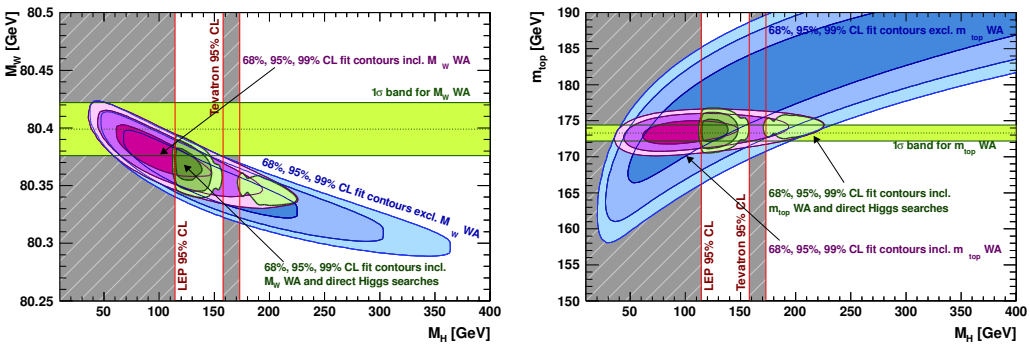


Figure 4.5: The Standard Model predictions of m_W (left) and m_t (right), shown by the blue bands for various Higgs mass hypotheses. The blue bands show the 1, 2, and 3σ predictions of the Standard Model, while the pink bands show the predictions including experimental measurements. The light green horizontal band shows the experimental value of m_W (m_t) and its associated uncertanty. Gray regions represent Higgs mass regions that are currently excluded, while the dark green bands represent the 1, 2, and 3σ predictions of the Higgs boson mass using all experimental measurements. Figures from Ref. [13].

be done in two possible ways – once using only the electroweak parameters as inputs, ignoring regions of Higgs boson masses already excluded by experimental searches, or using both the electroweak input parameters and results from existing Higgs boson searches. The results favor a low-mass Higgs boson, $m_H \in [72, 127]$ GeV/c^2 (at the 1σ level) when excluding the direct experimental Higgs searches, and $m_H \in [115, 132]$ GeV/c^2 (at the 1σ level) when including direct searches. These results are shown graphically in Figure 4.6.

From all of the constraints presented in this chapter, it is likely that the Higgs boson is light, $m_H \leq 130$ GeV/c^2. The analysis presented in this thesis is optimized to search for low-mass Higgs bosons. We develop a search strategy using the decay $H \to b\bar{b}$, which is dominant across the entire low-mass Higgs search range at the Tevatron.

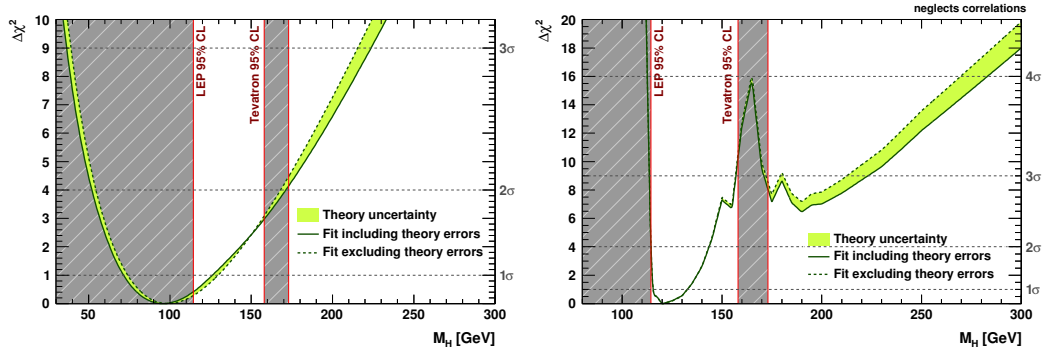


Figure 4.6: Fit results for the Higgs boson mass, using only electroweak input parameters (left), and with the inclusion of direct experimental Higgs boson searches (right). The minimum of the green band represents the fit value, with 1, 2, and 3σ uncertainties shown on the right axis. The gray regions represent already excluded regions of Higgs boson masses from experimental searches. Figures from Ref. [13].

Chapter 5
THE FERMILAB ACCELERATOR COMPLEX

The Fermi National Accelerator Laboratory, known as Fermilab, is a Department of Energy National Laboratory located approximately 40 miles West of Chicago, Illinois, in the town of Batavia. The laboratory is home to a wide array of detectors, from fixed target experiments to neutrino detectors, to the all-purpose high energy physics detectors, CDF and D0. In order to create the beams of particles used to produce the collisions, a complex chain of accelerators is used. This chain of accelerators is capable of bringing protons and antiprotons to an energy of 980 GeV, storing protons, producing and storing antiprotons, and sending beam to the various experiments scattered around the lab. Figure 5.1 shows a schematic drawing of the various accelerators used at Fermilab. This chapter details the operation of these accelerators and how particle collisions are produced.

5.1 Pre-Accelerator

The entire chain of acceleration starts at the pre-accelerator. Here, a bottle of hydrogen gas is used as a source of protons. The hydrogen gas is ionized to produce hydrogen ions (H^-) which can be accelerated more easily. These H^- ions are exposed to a potential of -750 kV inside an electrically charged dome, and therefore are accelerated out of the dome and obtain an energy of 750 keV. The pre-accelerator produces beam at a rate of 15 Hz.

The voltage is obtained through a Cockroft-Walton generator, which is a type of voltage multiplying technique. Figure 5.2 shows the pre-accelerator at Fermilab. After beam obtains an energy of 750 keV, it is transferred to the linear accelerator for the next stage in the process.

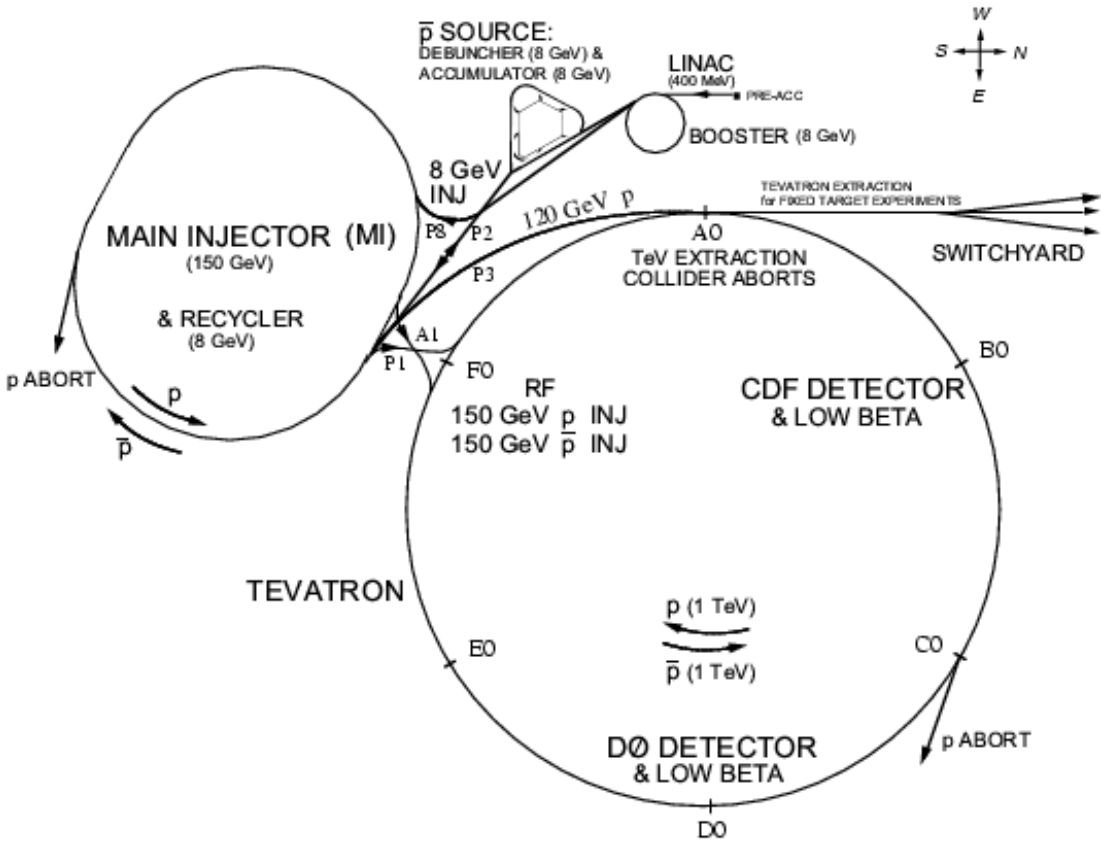


Figure 5.1: Schematic drawing of the accelerator chain at Fermilab.

Figure 5.2: The pre-accelerator at Fermilab, which takes H^- ions and accelerates them to an energy of 750 keV.

5.2 Linear Accelerator

The linear accelerator, or *linac*, is the second stage of the accelerator chain at Fermilab. As its name implies, this section is a long, straight batch of accelerators that take the beam energy from 750 keV, as it leaves the pre-accelerator, to 400 MeV before heading to the next stage of acceleration.

The linear accelerator employs radio frequency (RF) technology to accelerate beam. This technology uses oscillating electric fields to impart kinetic energy to the beam. However, just allowing the particles to experience the electric force would not provide any acceleration, as the energy would increase when the field is in one direction and decrease as the electric field oscillates back to the opposite direction. To avoid this unwanted situation, the particles are shielded from the electric field by drift tubes at certain positions along the beamline. This shielding allows the particles to experience the electric field only when it is useful for acceleration, and therefore the beam has a net gain of kinetic energy when it reaches the end of the linac. Figure 5.3 shows an example of the drift tubes in the linear accelerator. Note that the lengths of the

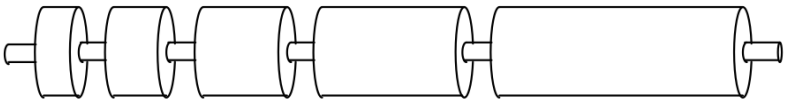


Figure 5.3: An example of drift tubes in a linear accelerator. Acceleration occurs in the gaps between the tubes. When the particles are in the tubes, they are shielded from the electric field (it is at that time in the opposite direction required for acceleration) and feel no accelerating force.

tubes change – more shielding is necessary as the particle traverses the linac because the particle is being accelerated and travels a longer distance when the field is in the opposite direction of that needed for acceleration.

The acceleration using drift tubes only occurs in the first portion of the linear accelerator. The last portion of the linac uses a different technology, side-coupled RF cavities, instead of the drift tubes used in the beginning portion. The side-coupled RF cavities are individual cells rather than one long cell containing many different drift tubes. This technology allows for larger accelerating gradients – the H^- ions are accelerated from 750 keV to 116.5 MeV in the $\sim$75 m long drift tube portion of the linac, and from 116.5 MeV to 400 MeV in the shorter, 64 m long side-coupled cavity portion of the linear accelerator[18].

5.3 Booster

After reaching an energy of 400 MeV in the linear accelerator, the H^- ions are transfered to the booster, the first circular accelerator encountered by the beam on its journey to the Tevatron. In the process of injecting beam into the booster, the H^- ions are sent through a stripping foil made of carbon, where the two electrons are removed. After penetrating the stripping foil, the beam now consists of protons, and can join already circulating proton beam in the booster. Any remaining H^- ions in the beam will be removed, as they have the opposite charge of the already circulating beam in the booster and will be deflected out of orbit by the magnetic force.

The booster is a circular accelerator, 468 m in circumference. It consists of focusing and defocusing magnets, which also steer the beam around the circular ring. The booster consists

of 19 RF accelerating cavities, which take the beam from an energy of 400 MeV at injection to 8 GeV, where it can be sent to the next stage of acceleration, the main injector. The booster is capable of accepting and accelerating beam at a rate of 15 Hz.

5.4 Main Injector

The 8 GeV beam from the booster is transfered to the main injector, a larger circular accelerator with a circumference of approximately 3.2 km. It again uses RF cavities to accelerate the beam, and dipole magnets to steer the beam around the main injector ring. The main injector serves as the 'crossroads' of the Fermilab accelerator complex, as it can send beam to a variety of locations at varying energies. It can accelerate beam to 150 GeV for injection into the Tevatron, can send beam at 120 GeV to the antiproton complex for the production of antiprotons, and can send beam to the fixed target experiments, or to the neutrino experiments at an energy of 8 GeV.

The main injector is used to accelerate both protons and antiprotons, and transfer beam into the Tevatron at the required injection energy of 150 GeV. The main injector tunnel also contains a second ring of permanent magnets to store antiprotons until they are ready to be extracted, accelerated, and injected into the Tevatron for high-energy collisions. The permanent magnets have a field to provide the correct deflection of antiprotons with an energy of 8 GeV around the main injector tunnel, as this is the energy with which they are selected from the antiproton complex. This second set of magnets is known as the recycler, as it was originally designed to re-capture beam from the Tevatron and reuse it at a later time.

After accelerating proton and antiproton beam to the injection energy of 150 GeV in the main injector, it is ready to be transfered to the Tevatron, the final accelerator in the chain toward producing proton-antiproton collisions.

5.5 Antiproton Source

The proton beam is, in a sense, simple to obtain, as it can be produced from an ordinary bottle of hydrogen gas. Such a source of antiprotons does not exist, however, and the antiprotons

must be produced on-site by utilizing the accelerated proton beam.

To produce antiprotons, a proton beam is extracted from the main injector after being accelerated to an energy of 120 GeV. This proton beam is then sent to the antiproton complex, where it is incident on a target composed of nickel. When the beam is incident on this target, a large spray of secondary particles is produced. In this spray, some antiprotons are produced, with a varying energy spectrum. The spray of particles, including antiprotons, is focused into a tighter beam by a lithium lens and then sent through a dipole magnet designed to divert negatively-charged particles with an energy of $\sim$8 GeV to another beamline. The energy spectrum of produced antiprotons has a peak near 8 GeV, making this a nice choice to maintain a high efficiency of antiproton capture.

Even with the system designed to capture antiprotons with energies near the peak of the production spectrum, the efficiency of antiproton production remains extremely low, with 10^5 incident protons required to capture just 1 or 2 antiprotons. The time needed to capture enough antiprotons for injection into the Tevatron remains one limiting factor in the accumulation of proton-antiproton collisions.

After producing enough antiprotons, the beam is transfered to the recycler, where the permanent magnets are designed to accept the 8 GeV particles from the antiproton source. The antiproton beam can then be transfered to the main injector, accelerated to 150 GeV, and injected into the Tevatron in preparation for proton-antiproton collisions.

5.6 Tevatron

The largest accelerator, both in terms of energy and circumference, at Fermilab is the Tevatron. The Tevatron is a circular accelerator 6.28 km in circumference, about twice that of the main injector. The Tevatron accepts both proton and antiproton beams at an energy of 150 GeV, and accelerates them using RF cavities to an energy of 980 GeV, after which the proton and antiproton beams are set to collide in the center of the two general-purpose detectors CDF and D0.

The Tevatron consists of 774 superconducting dipole magnets, used to steer the beam

around the circular beampipe, and 216 superconducting quadrupole magnets, used to maintain a focused beam. The entire lattice of magnets is cryogenically cooled with liquid helium to maintain superconductivity. There are a set of specialized quadrupole magnets placed immediately before and after the beam is about to collide in the center of the two detectors, known as low-β quadrupole magnets (β, the amplitude function, is a measure of how tightly focused the beam is at any position in the accelerator). These magnets can focus the beam extremely tightly, in order to produce a high rate of proton-antiproton collisions.

Table 5.1 shows some of the parameters of the beams for typical collider operations. To prepare for high-energy physics operations, 36 'bunches' of protons are accelerated through the complex to 150 GeV and injected into the Tevatron. The bunches are arranged into 3 'trains' of 12 bunches each, with a large space between each train. There is just 396 ns between bunches, but 2.6 μs between the trains. This extra time between trains is the abort gap, and allows time for certain magnets to direct the beam to a beam dump in the case of a malfunction, or to intentionally end collider operations. The antiproton beam has a similar structure, 3 trains of 12 bunches each, and is injected into the Tevatron, traveling the opposite direction as the proton beam. While the proton and antiproton beams share the same beampipe, they actually travel on helical paths, spiraling around one another.

After the beam is injected and accelerated to 980 GeV, the low-β specialized quadrupole magnets are ramped up to focus the beams in the center of the detectors. This step is known as 'squeezing' the beams. After the squeezing process, the beams are collimated to improve the beam profile and energy distribution. This involves moving a piece of metal into the edge of the beam, to remove particles that do not have the correct energy or position.

Once the beam is squeezed and collimated, the machine is declared ready for high-energy physics operation. At this point, the experiments turn on the various detector elements and begin recording collision data.

5.6.1 Luminosity

The luminosity of colliding beams is a measurement quantity which relates to the rate of collisions between beams. It is affected by many quantities, including the frequency of beam

Properties of Tevatron Collision Operations	
Beam Energy (p and $\bar{p}$)	980 GeV
p Bunches	36
$\bar{p}$ Bunches	36
Protons per Bunch	$2.6 \cdot 10^{11}$
Antiprotons per Bunch	$9 \cdot 10^{10}$
Collision Rate	396 ns
β^* Amplitude Function at IP	0.28 m
Antiproton Production Rate	$2.5 \cdot 10^{11}$ hr^{-1}

Table 5.1: Parameters of the Tevatron during a typical collider run.

crossing, bunch intensity, and beam size. It is defined in Equation 5.1,

$$L = f \frac{N_1 N_2}{4\pi\sigma_x\sigma_y} \tag{5.1}$$

assuming the two beams collide with zero crossing angle. The quantity f represents the frequency of collisions, N represents the number of particles in each of the two colliding bunches, and σ represents the transverse profile sizes of the beam in the x and y directions (beam moves in the z direction). Instantaneous luminosity is this measure at a given instant of time during a collider run. This quantity is measured in units of cm^{-2}s^{-1}. The instantaneous luminosity is integrated over time to give a measure of the number of collisions collected in a data set. This is referred to as the integrated luminosity, and relates the size of an experimental dataset.

To increase the luminosity, accelerators are designed to focus the beam more strongly (decreasing σ_x and σ_y), increase the bunch intensity (increase N_1 and N_2), or increase the collision rate (increase f). The Tevatron was originally designed to produce a luminosity of $1 \cdot 10^{30}$ cm^{-2}s^{-1}, however it has achieved luminosities over $400 \cdot 10^{30}$ cm^{-2}s^{-1}, an impressive achievement for the physicists involved in its operation. The Tevatron has delivered approximately 12 fb^{-1} of integrated luminosity to each of the CDF and D0 detectors. Using the associated Higgs production cross section of 300 fb as an example, this means that 3600 Higgs bosons are expected to have been produced by this mechanism in each detector. Due to data-taking efficiencies, reconstruction, and analysis selection, the actual number of expected Higgs events in an analysis is much smaller.

The Tevatron ended its run on September 30, 2011, as the LHC has been accumulating data at an ever-increasing rate. Even so, the unique proton-antiproton collision phenomenology will allow some Tevatron results to remain competitive with those from the LHC for several years to come.

Chapter 6
THE CDF DETECTOR

The CDF detector is one of two general-purpose detectors located on the Fermilab Tevatron designed for the analysis of proton-antiproton collisions at high energies. It has been in use since the first collisions at the Tevatron in 1985. The detector consists of several components optimized to identify different particles. All of these components must be working together and synchronized to ensure events are reconstructed correctly for further analysis. This chapter will detail the operation of the detector components, as well as the data acquisition process used to collect events offline for physics analysis.

6.1 Overview

The CDF (Collider Detector at Fermilab) detector was designed to identify many different types of particles, at a wide range of angles relative to the colliding beams. Therefore, it consists of many layers of detector material surrounding the interaction point where collisions take place. Figure 6.1 shows a cross-sectional view of the detector, with various components labeled. These components will be described in this chapter. Different detector components are used to make measurements of various properties of particles, such as energy, momentum, and lifetime. Using all these properties, one can reconstruct events in terms of the various physics objects present. These physics objects include electrons, muons, jets, photons, and other particle types.

Most collision events do not have any interesting objects, and do not need to be saved for further analysis. Additionally, there is some associated dead time with reading out the detector signals in an event, so it is not desirable to do this for every collision. To select interesting

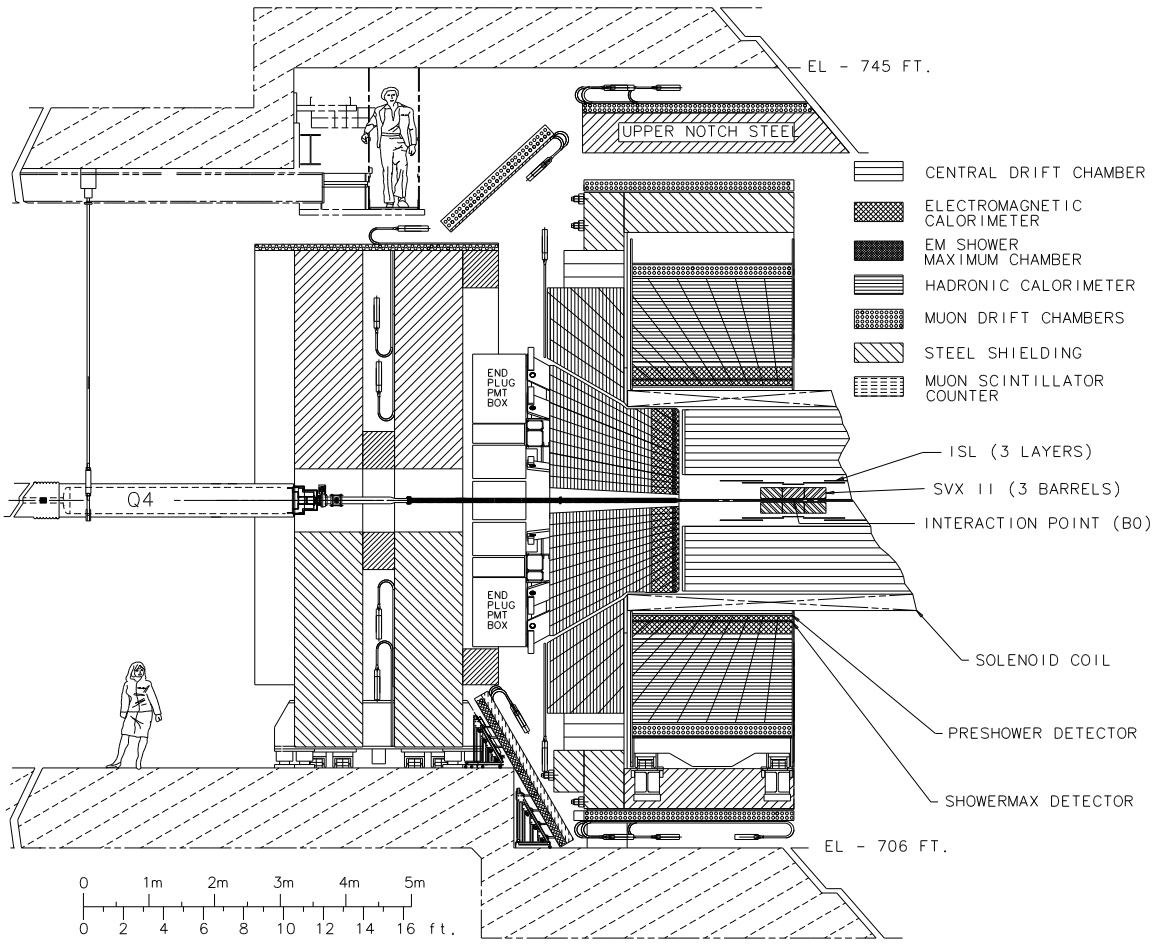


Figure 6.1: A cross-sectional view of the entire CDF detector.

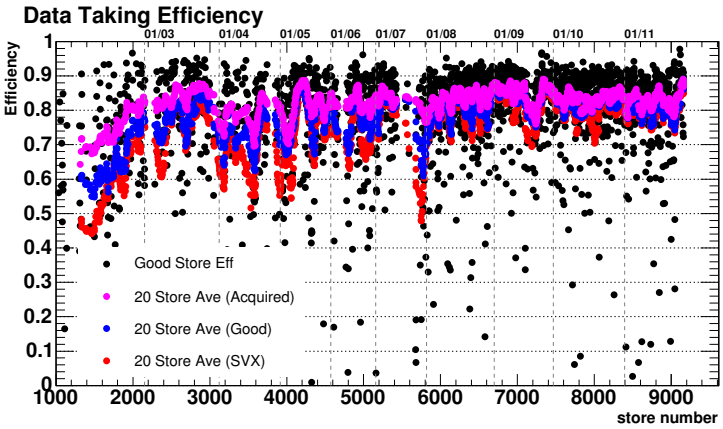


Figure 6.2: CDF data-taking efficiency through time.

events in real time, a sophisticated system of electronics exists to make the decision of whether to save an event to disk or not. This is known as the trigger system, and it makes the total size of offline data sets feasible for long-term storage.

As all of the detector components are used in concert with one another, it is important that they all be working properly. If just one component becomes inoperable, the entire event may not be reconstructed properly, or the trigger system may not work most efficiently. These effects become important, as one would like to obtain as much integrated luminosity as possible with the full detector being operational. To date, the CDF detector has had a high rate of data-taking efficiency – acquiring approximately 83% of the delivered integrated luminosity by the Tevatron. Figure 6.2 shows the data-taking efficiency over time.

Geometry

The CDF detector uses a cylindrical coordinate system, with the $\hat{z}$ direction defined by the incoming proton beam direction and the origin $(0,0,0)$ located in the center of the detector. The polar angle θ is defined relative to the positive $\hat{z}$ axis, with $\theta = 90^\circ$ defining the $\hat{x}$-$\hat{y}$ plane. Instead of the angle θ, one usually refers to the pseudorapidity, defined in Equation 6.1, with $\eta = 0$ being the central region of the detector, perpendicular to the incoming beams and $\eta = \pm\infty$ being parallel to the beamline. The azimuthal angle, ϕ, is defined with $\phi = 0$ at the

$+\hat{x}$ axis. The positive $\hat{x}$ axis points away from the center of the Tevatron ring.

$$\eta \equiv -\ln \tan \frac{\theta}{2} \tag{6.1}$$

One can define a distance between two points in the CDF detector with the pseudorapidity and azimuthal coordinates. This distance is defined as $\Delta R \equiv \sqrt{\Delta\phi^2 + \Delta\eta^2}$.

6.2 Tracking

To measure the momentum of charged particles produced in collisions, tracking detectors are used. The CDF detector has two separate tracking systems – a silicon tracking system which is useful for the identification of secondary vertices consistent with the decay of a B-hadron. The central outer tracker (COT) measures charged particle tracks and is used in the identification of charged leptons, with the addition of calorimeter information.

The tracking volume of the CDF detector is surrounded by a superconducting solenoidal coil consisting of niobium-titanium wire. The coil is 1.5 m in radius, and 4.8 m long. The current in the solenoid is 4650 A, which corresponds to a magnetic field of 1.4 T[18]. This strong magnetic field causes curvature of the charged particles. The curvature is measured to determine particle momentum in the tracking system.

Silicon Systems

The silicon tracking system is the first detector component a particle will encounter after being produced in a collision event. Silicon tracking systems are valuable for their excellent resolution – these systems are most useful for identifying displaced vertices due to long-lived particles that decay outside the beampipe. This is especially important for Higgs searches, as one would like to identify jets consistent with the decays of b-hadrons by requiring a displaced vertex to be present for a given jet. This technique is discussed further in Chapter 7.

Silicon systems consist of segments (strips or pixels) of silicon, a semiconducting material. When a charged particle passes through the silicon, some of the material is ionized to create electron-hole pairs. Because a high electric potential is applied to the silicon segments, these

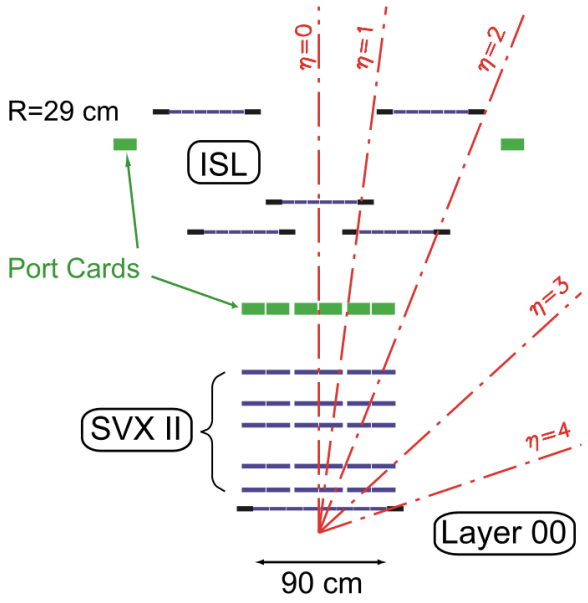


Figure 6.3: Silicon systems of the CDF detector.

electron-hole pairs move in opposite directions to produce a current. This current is converted to an electronic signal, and a 'hit' is measured on the specific silicon segment through which the particle passed. By stitching together the hits from different silicon segments, a particle track can be reconstructed.

There are three separate silicon systems in the CDF detector: Layer 00 (L00), the Silicon Vertex Detector (SVX), and the Intermediate Silicon Layers (ISL). The layout of these three systems is shown in Figure 6.3. The silicon tracking system has full coverage out to $|\eta| < 2.0$.

The L00 detector consists of just a single layer of silicon strips mounted directly on the beampipe, at a radial distance of just 2.5 cm from the interaction point. The expected hit resolution for this sub-detector is the best of any silicon layer, at 7.2 μm [9]. Moving outward from L00, one will encounter the SVX detector, consisting of five separate layers placed at radial distances of between approximately 4 and 17 cm from the interaction point. The SVX hit resolution is slightly higher than that of L00 at 10.8 μm[9]. Using the information from additional axial sensors, one can obtain the z-coordinate of a hit position with a resolution of 780 μm using the L00 and SVX detectors together. The ISL detector consists of another set of

silicon layers, at radial distances of 19 to 32 cm from the interaction point. The hit resolution for ISL is the worst of all the silicon sub-detectors, at 32.3 μm for the radial hit position and 2.2 mm for the position along the z axis[9]. In combination, the L00, SVX, and ISL are used to determine the impact parameters of charged particle tracks. The impact parameters are useful for identifying displaced tracks consistent with the decays of b-hadrons, an important aspect for Higgs searches. The combined resolution on impact parameter measurement for the three silicon sub-detectors is 25 μm for tracks perpendicular to the silicon layers. If some layers are missed by a track, the resolution will degrade, but remain less than ∼100 μm[9].

Central Tracking System

While the silicon detectors are used for precision measurements of track impact parameters for the purpose of identifying b-jets, the next tracking system in the CDF detector is the main source of measuring track curvatures and determining charged particle momenta. This specific detector is the Central Outer Tracker, or the COT[7]. Figure 6.4 shows where the COT exists in relation to the other sub-detectors comprising CDF. The COT is a large part of the volume of the CDF detector, extending 3.1 m along the beam axis, and radially between 43.4 and 132.3 cm from the interaction point.

The COT is a cylindrical multi-wire drift chamber consisting of 8 superlayers, with each superlayer divided along the ϕ direction into individual cells. These individual cells contain a field sheet on each side, which is responsible for creating an electric field in the cells. Also present are potential wires, responsible for producing the correct shape of the electric field inside the cell. The entire volume of the COT is filled with an equal mixture of argon and ethane gas. When charged particles traverse the COT, this gas is ionized. The resulting ionization travels along the electric field lines until it encounters the sense wires in a COT cell. An avalanche effect occurs due to the increasing field strentgth near a sense wire, and is recorded as a 'hit' by that individual sense wire. This electric signal, in combination with timing information from beam crossings, can be used to reconstruct the many charged particle paths that result from a proton-antiproton collision.

In total, there are 78 120 wires in the COT, with 30 240 being sense wires that can be used

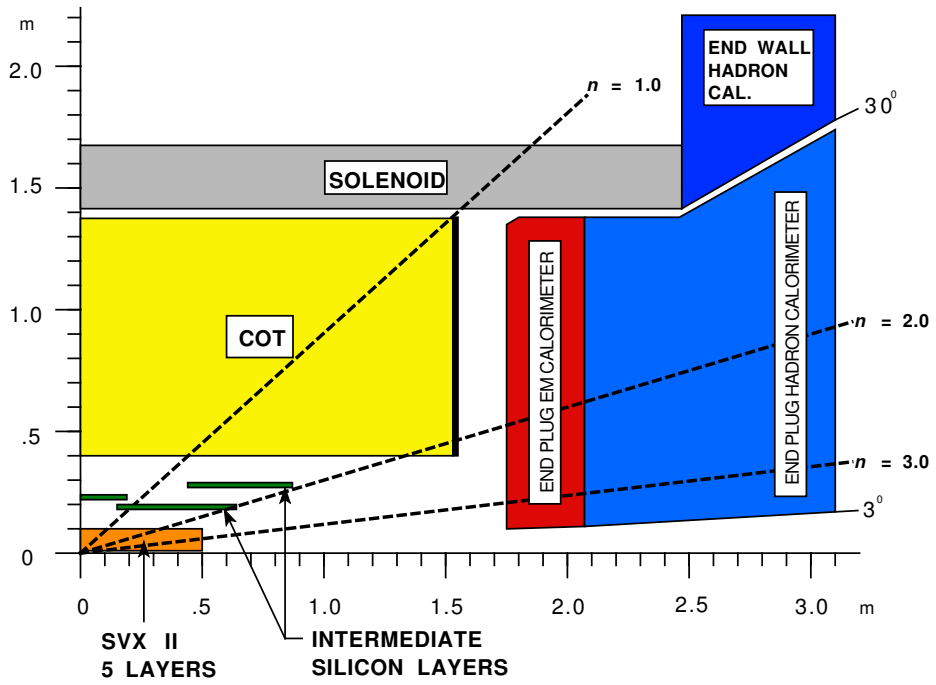


Figure 6.4: A section of the CDF detector, showing some of the many sub-detectors and their relative locations.

to reconstruct hits. The COT can be used to determine the position of a track in both the $r-\phi$ and $r-z$ planes, due to alternating superlayers being offset by $\pm 2^{\circ}$ relative to the neighboring layers. These alternating superlayers allow the track to be reconstructed in three dimensions. Since the entire volume of the COT is immersed in the magnetic field provided by the solenoid, charged particles will have a helical path as they traverse the COT. To reconstruct the track due to a charged particle, the COT is used to measure the parameters of the helix corresponding to the path of the charged particle. From the five helix parameters, one can determine the transverse momentum of the charged particle. The transverse momentum (p_T) resolution of the COT, determined from the resolution on the track curvature measurement, is approximately $0.3\% \cdot p_T^2$ $(\mathrm{GeV}/c)^{-1}$[18]. The resolution is a function of the transverse momentum, as particles with large p_T leave straighter tracks which are harder to measure accurately.

Althought the COT is mainly used to measure the momenta of charged particles, it was also designed to work in concert with the silicon sub-systems. Track matching between silicon and

COT hits occurs to better reconstruct charged particle tracks. Since particles only traverse all superlayers of the COT for $|\eta| < 1$, the silicon systems can be useful to increase reconstruction efficiency in the further forward regions of the detectors where particles only travel through a smaller portion of the COT.

6.3 Calorimetry

Moving outward from the center of the detector, after particles encounter the silicon system and COT, they begin to interact with the calorimeters, where particle energies are measured. While the tracking systems make their measurements with the particle intact, the energy determination is a destructive measurement. This means the particle is contained in the calorimeter and does not travel through to subsequent detectors. Additionally, while the tracking systems can only measure charged particles, the calorimeters can measure neutral particles as well as charged particles.

The CDF calorimeter is divided into several subdetectors. There are two general types of calorimeters, electromagnetic and hadronic. The electromagnetic calorimeters are used to measure particles which interact mainly by the electromagnetic force, such as photons and electrons. Hadronic calorimeters measure particles interacting via the strong force, such as charged pions and kaons. Besides the separate electromagnetic and hadronic calorimeters, the CDF calorimeters are divided into central and forward calorimeters. The central calorimeters (both electromagnetic and hadronic) cover a pseudorapidity range $|\eta| < 1.1$, while the forward ('plug') calorimeters cover $1.1 < |\eta| < 3.4$. The layout of the calorimetry system is shown in Figure 6.5.

Electromagnetic

The central electromagnetic calorimeter (CEM) [14] consists of several towers, each covering a range of 0.1 units in η (out to $\eta = \pm 1.1$) and 15° in the ϕ direction. This calorimeter covers the entire azimuthal range. Each of these towers consists of alternating layers of lead absorber and plastic scintillator. When particles are incident on the first layer of the calorimeter, a

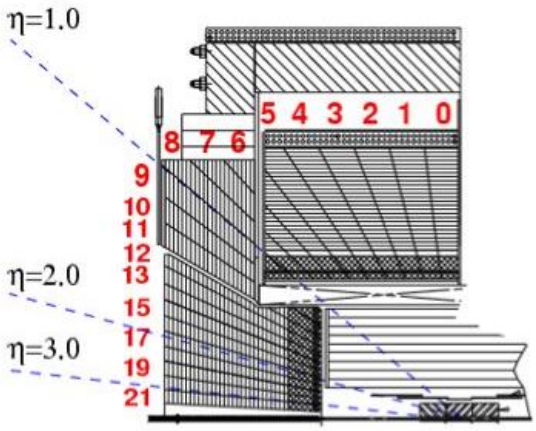


Figure 6.5: Schematic of the CDF calorimeter systems, with the individual towers numbered.

shower consiting mainly of electrons, positrons and photons develops. When the shower passes through the scintillator layers, light is produced and directed toward a photomultiplier tube via light guides. The phototube produces an electric signal that is calibrated to give a measurement of the total energy deposited by the shower. Since the shower is due to the originating particle, this gives a measurement of the particle energy.

Similar to the COT, the resolution of the energy measurement from the CEM detector is a function of the energy itself. The CEM resolution is $\frac{\Delta E}{E} = 13.5\% \cdot E_T^{-1/2}$ (GeV)$^{1/2}$, meaning for large energies the measurement resolution is better.

There is another electromagnetic calorimeter which extends the active detector range to $|\eta| < 3.6$ – the plug electromagnetic calorimeter (PEM)[10]. The PEM detector is built similarly to the CEM detector, although it has coarser segmentation in the η coordinate. It again consists of alternating layers of lead absorber and plastic scintillator. The resolution from PEM measurements is comparable to that of the CEM, at $\frac{\Delta E}{E} = 16\% \cdot E_T^{-1/2}$ (GeV)$^{1/2}$.

Hadronic

Although the hadronic calorimeter operates under the same basic principle as the electromagnetic calorimeter, is it constructed differently due to the properties of the hadronic showers produced by incident particles. The shower develops via the strong interaction, so it consists of

pions, kaons, protons, neutrons, and other hadrons instead of electrons, positrons, and photons as in an electromagnetic shower.

The central hadronic calorimeter at CDF (CHA)[16] consists of alternating layers of steel absorber and plastic scintillator. It is segmented in such a way to have the towers from the electromagnetic and hadronic calorimeters coincide. The scintillator works the same way as in the CEM detector, guiding the light produced in the scintillator toward a photomultiplier tube. The CHA towers are much longer, at 1.3 m of material compared with only 0.2 m of material for the CEM towers. This is due to the longer absorption length required for hadronic showers. The energy resolution for the hadronic calorimeter is poorer than the electromagnetic, at $\frac{\Delta E}{E} = 50\% \cdot E_T^{-1/2}$ (GeV)$^{1/2}$.

The hadronic calorimeter also consists of a forward detector, similar to the PEM detector. A third sub-detector, known as the 'end-wall' hadronic calorimeter sits in between the central and forward hadronic calorimeters and increases the coverage of the calorimeter. These sub-detectors are shown in Figure 6.5.

Missing Energy

The calorimeters can also be used to infer the presence of undetected particles. In a collision event, we expect the total energy transverse to the beam to be zero. We can use this constraint along with the calorimeter measurements to determine the amount of "missing energy" – a measure of the imbalance of energy in the transverse plane. The missing energy, $\not\!\!E_T = |\vec{\not\!\!E}_T|$, is defined in Equation 6.2, where i represents the number of the calorimeter tower with $|\eta| < 3.6$ and $\hat{n}^i$ is a unit vector perpendicular to the beam axis and pointing at the i^{th} calorimeter tower.

$$\vec{\not\!\!E}_T = -\sum_i E_T^i \hat{n}^i \tag{6.2}$$

The missing energy is a quantity that is especially useful in selecting events containing neutrinos. Neutrinos will not interact, escaping undetected and leaving an imbalance of measured energy in the calorimeter. This quantity needs to be reconstructed carefully, and requires corrections for certain types of particles.

6.4 Muon Systems

Not all particles will be completely absorbed by the calorimeters. Some particles may travel through the calorimeters and only deposit a small fraction of their energy. These particles are known as minimum-ionizing particles, and the main type of particle that is detected this way at CDF is the muon.

The rate of energy loss as a particle travels through a medium is roughly proportional to E/m^2, where m represents the particle mass. For electrons, this is a large factor, since the electrons usually have energies of order 10 GeV while the mass is only 0.5 MeV. For muons, however, the mass is a factor of approximately 200 larger, so the rate of energy loss drops dramatically for muons compared to electrons with the same energy.

Other particles may travel all the way through the calorimter some small fraction of the time, but muons generally always traverse the calorimeter completely. Special detectors are located on the outermost layer of the detector to identify these muons. To shield the detectors from the naturally-occurring cosmic ray muon background, these detectors are surrounded by pieces of steel.

There are four main sub-detectors making up the muon system at CDF: the CMU and CMP detectors cover the range $|\eta| < 0.6$, the CMX extends from $|\eta| = 0.6$ to $|\eta| = 1.0$, and the IMU extends even further from $|\eta| = 1.0$ to $|\eta| = 1.5$. Figure 6.6 shows the layout of the muon sub-detectors. The muon detectors operate on the same basic principle of the COT. They consist of drift chambers containing several individual cells. The chambers are filled with an equal mixture of argon and ethane gas, and each individual cell contains a single wire that records the ionization due to a muon traversing the chamber. A combination of hits in neighboring chambers, called a 'stub', is required to identify the presence of a charged particle, most likely a muon. Additionally, the CMX, CMP, and IMU sub-detectors have scintillation counters and photomultiplier tubes which record the presence of a charged particle. These counters are used for timing information, to ensure that the reconstructed charged particle is associated with the correct proton-antiproton interaction.

Information from the hits in the muon chambers can be used to match a charged parti-

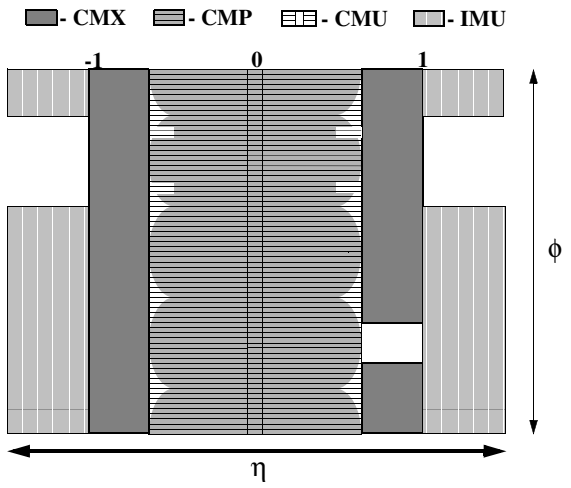


Figure 6.6: Positions of the individual muon sub-detectors in $\eta - \phi$ space.

cle track in the silicon tracking system or COT. This is useful for trigger activation, and for offline reconstruction. This thesis uses reconstructed muon candidates that have tracks from the COT, but do not necessarily have matching 'stubs' in the muon chambers. Loosening the stub requirement allows for increased acceptance of muon candidates, but also increases the rate of background events. We have developed several techniques to reduce this additional background, described in Chapter 7.

6.5 Cherenkov Luminosity Counter

Physics analyses depend on an accurate measurement of the luminosity collected by the detector, for precise measurements of process cross sections and estimations of signal abundances for new physics processes. The measurement of instantaneous luminosity is done with the Cherenkov Luminosity Counter (CLC)[5].

There are two CLC modules, with one placed at each side of the detector. These modules cover a peseudorapidity range $3.75 < |\eta| < 4.75$, and each consists of 48 long, narrow tubes which are used as Cherenkov detectors. Figure 6.7 shows a picture of one CLC detector module. These narrow tubes are filled with isobutane gas. As the particles produced in collision events are highly relativistic, they exceed the speed of light in the gas, and radiate Cherenkov light.

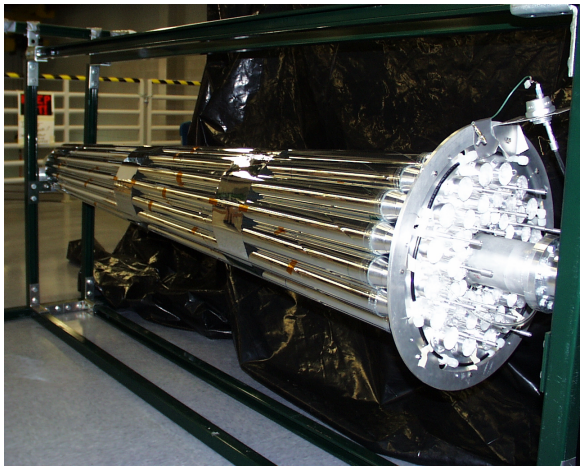


Figure 6.7: A photograph of one of the two CLC modules used to measure $p\bar{p}$ collision luminosity in the CDF detector.

This light is detected via photomultiplier tubes at the end of each individual segment of the CLC detector.

The CLC uses the total inelastic cross section $\sigma_{p\bar{p}}$ at the Tevatron to determine luminosity. This cross section has been measured to be approximately 60 mb [11, 3]. The luminosity L is determined by Equation 6.3, where N represents the number of interactions per bunch crossing, and f is the frequency of bunch crossings at the Tevatron.

$$L = \frac{Nf}{\sigma_{p\bar{p}}} \tag{6.3}$$

To measure luminosity, the CLC uses two operating methods. For low-luminosity environments, the CLC measures bunch crossings which have no proton-antiproton interactions. As the number of interactions per bunch crossing follows a Poisson distribution, the probability for no interactions is given by e^{-N}. From this, N can be determined, and thereby the luminosity can be determined. In higher luminosity environments, there are few bunch crossings which yield no interactions, so the CLC measures the number of hits in the CLC detectors. This quantity is proportional to the number of interactions in a bunch crossing, so it can also be used to determine N, and therefore the instantaneous luminosity L.

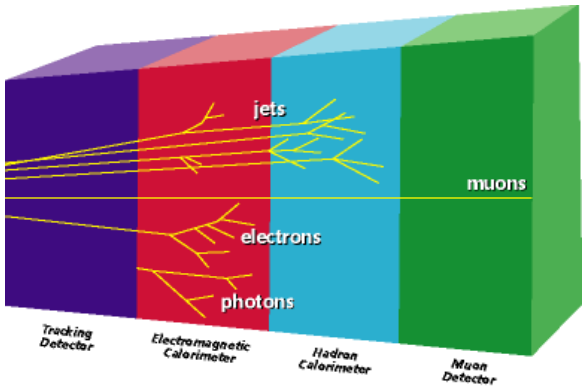


Figure 6.8: The signatures of different particle types in the various CDF sub-detectors.

In conclusion, all of the sub-detectors at CDF must work together to produce high-quality physics data. Figure 6.8 shows how the different components are used in the identification of particles. The presence or absence of activity in an event is used to distinguish particles. For example, both a photon and an electron leave a deposit in the electromagnetic calorimeter, but only the electron will leave a track in the COT. Similarly, a muon and an electron will both leave tracks in the COT, but only the muon will travel outside the calorimeters and through the muon chambers.

If just one of these components is not functional, the data acquired may not be usable for offline analysis. It is a testament to all involved in the design, construction, maintenance, and supervision of the sub-detectors that CDF is able to maintain such a high efficiency of data acquisition with all detectors functioning properly.

We have discussed how individual components of the CDF detector operate. Reading out the signals from all of these channels and reconstructing events for offline analysis is another major component of operating the detector. This aspect of operation is discussed in the following section.

6.6 Data Acquisition

The Tevatron delivers collisions to the CDF experiment every 396 nanoseconds. Due to the sheer number of collisions occuring, as well as the fact that most collisions do not produce

interesting physics objects, it is not desirable nor feasible to record every collision event as part of the final dataset. There exists a complex system of electronics to make real-time decisions on whether to save an event for offline analysis or to ignore the event completely. This system is known as the data acquisition system (DAQ), and involves several layers of trigger systems. The trigger systems identify certain features of events characteristic of physics processes of interest. A flowchart of this data acquisition system is shown in Figure 6.9. The DAQ runs on a clock cycle of 132 ns, while the actual collision rate is three times that at 396 ns. There was a proposal to reduce the collision rate to 132 ns, which is the reason for the difference in the clock cycle and collision rate.

When a collision occurs in the detector, the raw data from the detector is sent to the first layer of the trigger system, 'Level 1'. The Level 1 trigger is a pipeline which acts to increase the amount of time the system can use to process individual events. The pipeline consists of 42 clock cycles, or about 5.5 μs. Events come in every 396 ns, so a decision must be made every 396 ns.

The Level 1 trigger receives information from three detector elements – the COT tracking system, the central calorimeter, and the CMU and CMX muon systems. If certain physics objects are present, such as high-p_T charged particles (tracks), muon stubs, or large energy deposits in the calorimeter, the decision is made to keep the event and the relevant data is sent to the next system. If none of the requirements on certain reconstructed objects are met, the event is rejected and not used for further analysis. The Level 1 trigger selection reduces the data flow tremendously, from the 2.5 MHz beam crossing rate down to the range of 10-30 kHz. The rate of accepted events from the Level 1 trigger depends on the instantaneous luminosity of the beams. In fact, certain triggers can turn off and on depending on the instantaneous luminosity to keep the accept rate from rising too high, or the triggers can accept only some of the events that pass its requirements. This is known as prescaling triggers, and is an important effect used to maintain a high data-taking efficiency simultaneously for vastly different types of physics processes.

The Level 2 trigger recieves all of the data from the Level 1 system, plus additional data from the silicon tracking system and the calorimeters. The Level 2 system consists of 4 separate

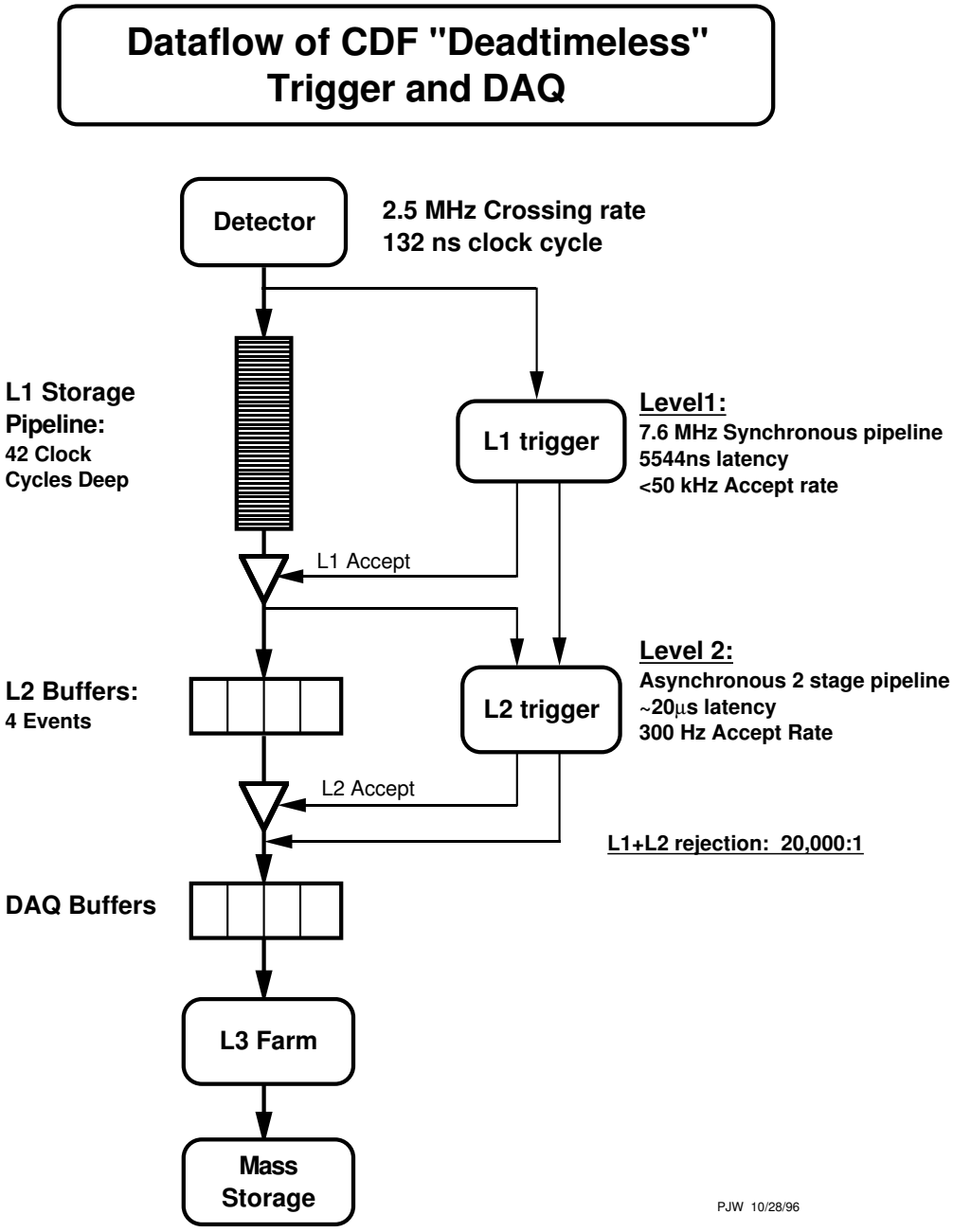


Figure 6.9: A flowchart of the CDF trigger system.

processors that can receive data from Level 1. If these four processors are simultaneously analyzing an event, the incoming events from Level 1 are lost, and deadtime is incurred. To keep the deadtime to a minimum, the Level 2 system must make a trigger decision approximately every 20 μs.

The Level 2 system is the first stage where clusters are formed out of individual calorimeter towers. This clustering is used to reconstruct jets in the detector. This clustering algorithm involves finding a 'seed' tower with an energy deposit above some trigger threshold, usually several GeV. Then neighboring towers are analyzed and sometimes added to the seed tower to form a cluster. Some Level 2 triggers require these clusters to flag the event for further analysis. The silicon tracking information is used at Level 2, and is especially important for triggers that require the presence of a displaced vertex (for identifying b-quark jets), or tracks with large impact parameters.

The Level 2 system reduces the Level 1 accept rate of approximately 20 kHz down to approximately 300 Hz. Events that pass the Level 2 trigger requirements are sent to the next system for complete event reconstruction. In combination, the Level 1 and Level 2 trigger systems reduce the data flow rate by an approximate factor of 10^4.

The final step in the CDF data acquisition system is the Level 3 computing farm. Events passing the Level 1 and Level 2 trigger selections are transferred to a computing farm, where the complete set of information read out from the detector is used to completely reconstruct the event of interest. The Level 3 computing farm consists of approximately 350 individual processing nodes that are used to analyze individual events. Events that pass the Level 3 trigger selection are permanently stored for further offline analysis. Since only events that pass Level 3 are available for physics analyses, it is important to design a trigger that will select the specific physics processes that are interesting to analyze.

The rate of events passing Level 3 selection is approximately 150 Hz for an average Tevatron store, meaning the total reduction in data flow from the collision rate to the rate of writing events to offline storage is $2 \cdot 10^4$. Even with this large reduction in data flow, the CDF detector still produces a large amount of data, requiring a large amount of permanent storage. The average event size written out to storage after a Level 3 decision is approximately 250 kilobytes,

which leads to approximately 100 terabytes recorded each year for offline analyses. This assumes an average of 100 hours of collider operations per week for the Tevatron, an operating efficiency of about 60%.

Triggers are one of the earliest choices made when designing an analysis. The trigger selection defines the subset of data used for offline analysis. Usually one chooses a small number of triggers, such as ones requiring a muon or electron. Because the dependence of trigger efficiencies on kinematic quantities like lepton p_T and η is usually difficult to model, a select set of triggers is used in the analysis. This analysis, however, uses the entire set of available triggers having specific characteristics for the offline analysis. This increases the size of the available dataset significantly, and the complication of modeling trigger efficiencies is avoided due to a novel multivariate technique described in Chapter 8.

Now that the CDF detector is understood, the specifics of this analysis can be described. The offline analysis is the focus of the remainder of this dissertation.

Chapter 7
Object Identification

To search for the Higgs boson in the process $ZH \to \mu^+\mu^- b\bar{b}$, we first use objects reconstructed with the CDF detector to identify events consistent with the expected final state. This involves using reconstructed muon candidates, as well as jets with displaced vertices, known as b-tagged jets. This chapter details the object reconstruction and selection, which forms the basis for this analysis.

7.1 Muon Selection

The first step in the event selection for this analysis is to require two muons that are consistent with the decay of a Z boson. Muon candidates are reconstructed using information from the tracking systems, calorimeters, and of course the muon detectors. At CDF, there are three general categories of reconstructed muon candidates. CMUP muon candidates are found in the pseudorapidity range $0 < |\eta| < 0.6$ and require a COT track matched with hits in both the CMU and CMP muon detectors (a 'stub'). An upper limit is imposed on the calorimeter energy deposit associated with this muon candidate, to ensure that it is a minimum-ionizing particle. CMX muons are similar to CMUP in that they require a track matched with a minimum-ionizing calorimeter deposit and stub in the CMX detector. CMX muons have a different pseudorapidity range $0.6 < |\eta| < 1.0$, however. Finally, CMIO muon candidates also require the presence of a track and minimum-ionizing calorimiter deposit, but CMIO muon candidates do not require a stub in any muon detector. As the criteria for identifying CMIO muon candidates is much less stringent, many CMIO candidates are actually mis-identified particles that mimic the signature

Cut Quantity	CMUP	CMX	CMIO
Primary Vertex $\lvert z_0 \rvert$ (cm)	< 60		
p_T (GeV)	> 20	$\in [10, 20]$	> 20
Hadronic Calorimeter Energy (GeV)	< 6		$0 < E < 12$
Electromagnetic Calorimeter Energy (GeV)	< 2		$0 < E < 4$
Isolation Fraction	< 0.1		< 0.4
Track Axial COT Superlayers	≥ 3	any	any
Track Stereo COT Superlayers	≥ 2	any	any
Impact Parameter $\lvert d_0 \rvert$, with Silicon Hits	< 0.02		
Impact Parameter $\lvert d_0 \rvert$, no Silicon Hits	< 0.2		
COT Hits	≥ 1		

Table 7.1: Example selection criteria for CMX, CMUP, and CMIO type muons. There are two sets of cuts for the CMIO muon type, with looser cuts if the CMIO p_T is above 20 GeV.

of muons in the detector.

It is typical in analyses requiring the presence of muons to use a set of cuts to identify quality CMUP, CMX, and CMIO muon candidates. A typical list of cuts is shown in Table 7.1. A set of cuts such as those in Table 7.1 removes a large fraction of ZH signal acceptance, decreasing the sensitivity of the analysis. For this analysis, a multivariate approach is used to select high-quality muon candidates instead of a set of cuts on muon quantities. In addition, this multivariate approach removes the distinction between CMX, CMUP, and CMIO muons, grouping all muons into a single category.

7.1.1 Multivariate Muon Identification

To increase acceptance for $Z \to \mu^+\mu^-$ events, this analysis uses a multivariate function for the selection of muon candidates. The multivariate function is an artificial neural network (see Appendix A for details) trained on several of the muon kinematic quantities listed in Table 7.1. The quantities used to train the neural network are described below.

- **Transverse Momentum** p_T – The transverse momentum of the muon candidate track (measured by the COT).

- **Electromagnetic Energy** E_{EM} – The energy deposit in the electromagnetic calorimeter

matched to the muon candidate track.

- **Hadronic Energy** E_{HAD} – The energy deposit in the hadronic calorimeter matched to the muon candidate track.

- **Isolation Fraction** – The total energy deposited in the electromagnetic and hadronic calorimeters in a cone of $\Delta R \equiv \sqrt{\Delta\eta^2 + \Delta\phi^2}$ around the muon candidate track, divided by the muon candidate track p_T.

- **Number of Axial Hits** – The total number of hits forming the muon candidate track from all COT axial superlayers.

- **Number of Stero Hits** – The total number of hits forming the muon candidate track from all COT stereo superlayers.

- **Number of Silicon Hits** – The total number of hits in the silicon tracking system corresponding to the muon candidate track.

- **Impact Parameter** $|d_0|$ – The impact parameter of the muon track, relative to the primary vertex of the event.

- **Track** χ^2 – The χ^2 per degree of freedom from the fitting of the track to the COT hits.

- **Muon** (η, ϕ) **position** – The pseudorapidity and azimuth coordinates of the muon position in the detector.

- **Distance to Nearest Jet** – The minimum distance ΔR between the muon candidate and the nearest jet in the event.

The muon selection function (henceforth muon NN) is trained on individual muon candidates. The signal sample is composed of $Z \rightarrow \mu\mu$ events from a ZH Monte Carlo sample where the two muons in an event have opposite charge and form an invariant mass $m_{\mu\mu}$ in the range $[76, 106]$ GeV/c^2. The background training sample is composed of events selected from data containing two muons with the same charge ('same-sign' pairs). The data sample gives us a large number of these events, which are due to QCD production, the main source of

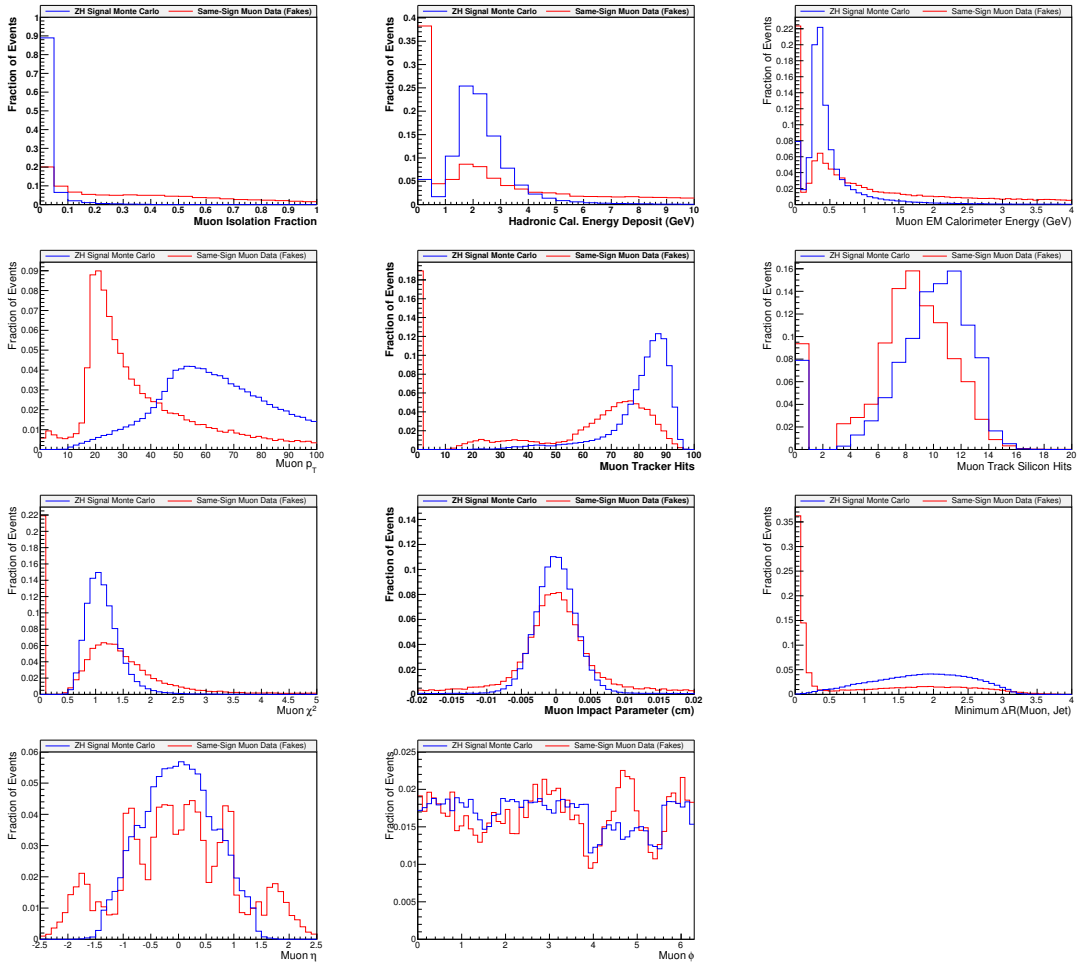


Figure 7.1: Comparison of input variable distributions for the signal training sample (blue) and the background training sample (red). Histograms are drawn normalized to equal areas.

mis-identified muons in the detector. This selection produces events containing two muons, but each of these muons is a separate record used for training the muon NN. Figure 7.1 shows the separation between the signal and background training samples for the muon NN. The muon NN is trained using the TMVA software package[32], a set of tools used to produce multivarate algorithms and functions. We choose to designate a high-quality muon as one having a muon NN output score greater than or equal to 0.0. With this selection criteria, we have developed a high-efficiency method to select muon candidates. This function retains 96% of the signal sample while simultaneously rejecting 94% of the background sample, for a cut value of greater

than or equal to 0.0.

The muon NN forms the first part of the analysis selection. Once two high-quality muons are selected with the muon NN, they are required to form an invariant mass $m_{\mu\mu}$ in the 'Z-window': $[76, 106]$ GeV/c^2. The invariant mass m_{12} for two particles is defined in Equation 7.1.

$$m_{12} = \sqrt{(E_1 + E_2)^2 - (\vec{p}_1 + \vec{p}_2)^2} \tag{7.1}$$

The only other requirement on muon candidates is that the two muon candidates forming the Z boson originate from the same primary vertex. This is imposed by requiring the difference in the muon track primary vertex positions $|z_0^{(\mu_1)} - z_0^{(\mu_2)}|$ is less than 4 cm.

Figure 7.2 shows the output of this multivariate muon selection function. Before requiring the muon NN selection, the black curve is obtained, consisting mainly of fake muons from QCD production. The red curve shows events which do not pass the muon NN selection. This red curve basically accounts for all events without a real Z boson, overlapping with the black curve everywhere except in the vicinity of the Z boson resonance. Events that pass the muon NN selection are shown in the blue curve. This represents a high-quality Z selection, signified by the presence of a sharp, clean Z resonance peak. Some fake events do pass the muon NN selection, shown by the dotted black line, but this contribution is an order of magnitude below the real Z boson contribution.

After the selection of $Z \to \mu\mu$ candidate events, the next step of the selection involves searching for jets consistent with the decay of a Higgs boson.

7.2 Jet Selection

Because the Higgs boson is expected to be a heavy object, $m_H > 100$ GeV/c^2, it will be essentially produced at rest in collision events. This means when it decays, the decay products will be high-p_T objects. This analysis searches for Higgs bosons decaying to a pair of b and $\bar{b}$ quarks. When quarks are produced in collision events, they rapidly hadronize to produce sprays of particles in the detector. These sprays of particles are called jets, and are reconstructed by searching for clusters of energy deposits in the calorimeters around a high energy deposit in a

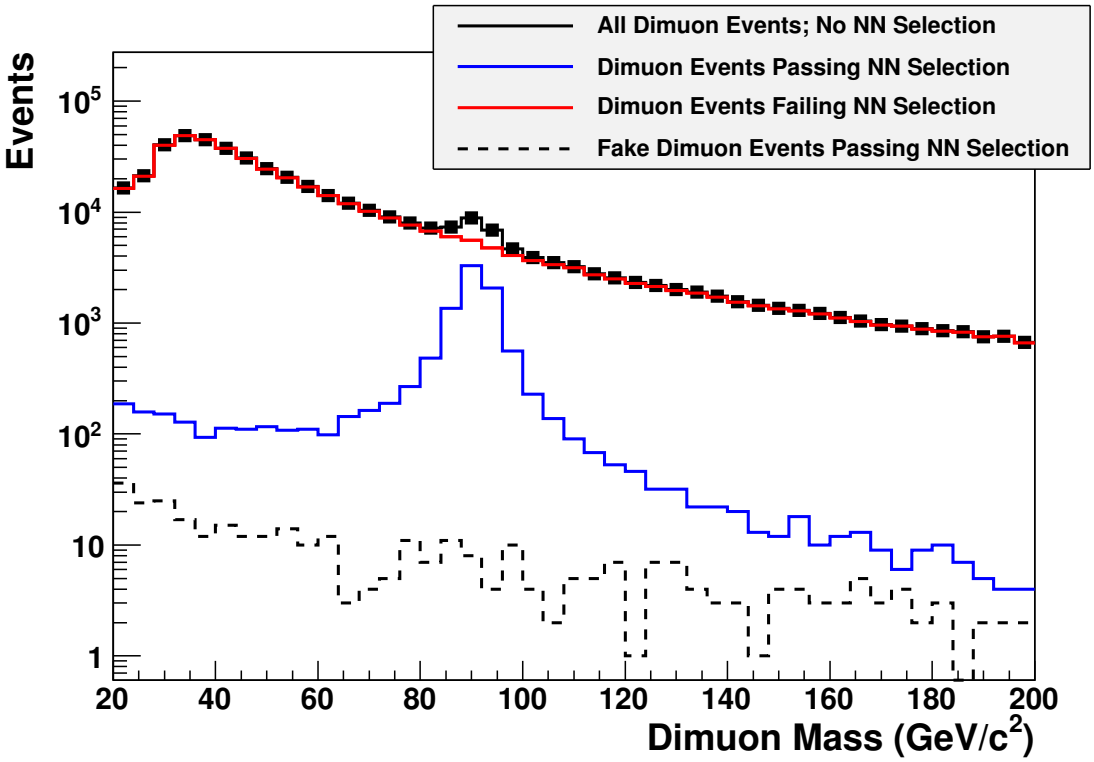


Figure 7.2: The muon NN selection function performance. Events passing the selection are shown in blue and form a clean Z peak. The red curve shows events failing the selection, accounting for basically all of the non-Z background that would have otherwise entered the event selection.

Jet Selection Criteria	
N_{jets}	≥ 2
Jet E_T	≥ 15 GeV
Lead Jet E_T	≥ 25 GeV
Jet $\lvert\eta\rvert$	≤ 2.0

Table 7.2: Jet selection criteria for this analysis.

'seed' tower. The reconstruction algorithm used in this analysis clusters towers together within a distance of $\Delta R < 0.4$ from the seed tower[19].

After requiring the presence of a Z boson as described above, we require the presence of two or more jets having pseudorapidities $|\eta| < 2.0$. The jets are also required to have large energy deposits, with the highest energy jet in the event required to have a transverse energy deposit $E_T \equiv E \sin\theta$ of at least 25 GeV. The other jets in the event must have transverse energy deposits of at least 15 GeV. The jet selection requirements are summarized in Table 7.2.

The jet energies are not simply the energies measured by the individual calorimeter towers summed together. There are various corrections that are applied to the calorimeter measurements to determine the energies of reconstructed objects[17]. The corrections include calibration factors, such as those measured from $Z \rightarrow e^+e^-$ decays to ensure a reconstructed Z boson mass at the correct value. There are pseudorapidity-dependent corrections applied to ensure a uniform response of the calorimeter as a function of pseudorapidity. Additionally, if multiple interactions occur during the same beam crossing, there will be extra energy deposited in the calorimeter from the additional interactions. Finally, due to some regions of the calorimeter lacking instrumentation, the jets are corrected for energy loss due to these regions. All of these corrections (known as 'level 5 corrections' at CDF) are applied to the jets in this analysis prior to making any selection using the E_T value of the jets. There are additional corrections possible, such as those removing energy due to the underlying event and correcting for energy outside of the pre-determined cone size of the jet (this analysis uses a cone size of $\Delta R = 0.4$ around the seed tower). These additional corrections, however, are not applied to jets in this analysis.

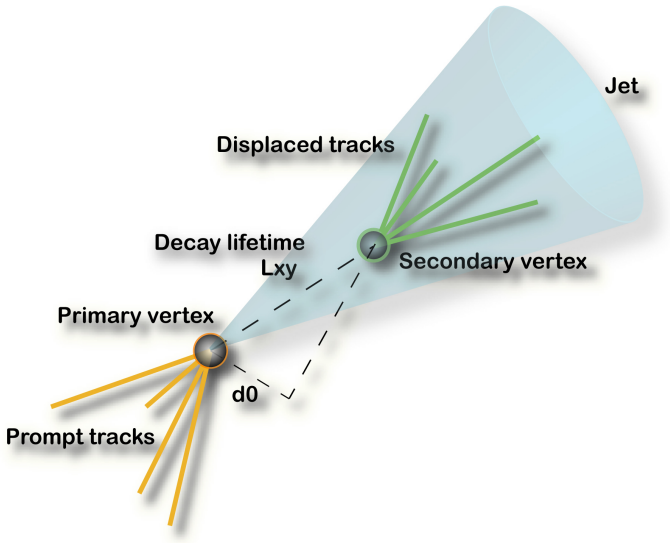


Figure 7.3: An example of a jet produced by the decay of a long-lived particle, such as a B meson formed from the hadronization of a b-quark.

7.3 b-tagging Algorithms

Selecting events with two high-quality muons consistent with the decay of a Z boson and two or more high-E_T jets ensures consistency with the final state expected from the process $ZH \rightarrow \mu^+\mu^- b\bar{b}$. However, there is more that can be done to increase the purity of signal relative to the amount of background events. There exist algorithms to identify displaced vertices such as those produced from the decays of B mesons produced in the hadronization of b-quarks. Due to the relatively long lifetime of B mesons, they travel some distance – usually several millimeters – from the primary vertex before decaying. This decay length can be resolved by the silicon detector as a displaced vertex. An example of this is shown in Figure 7.3. If a displaced vertex is found in a jet, the jet is said to be 'b-tagged'. There are complications with b-tagging, as not all jets containing B hadrons will be tagged, and not all tagged jets are actually from B meson decays. These complications will be discussed in a later chapter.

Secondary Vertex Tagging

The first of two b-tagging algorithms used in this analysis is known as the secondary vertex (SV) algorithm[6]. The algorithm is applied to one jet at a time. After finding the position of the primary vertex using all the tracks in the event, the algorithm searches for tracks inside the cone of a jet that have large impact parameters. There are quality cuts made on the tracks used in the algorithm, including a $p_T > 0.5$ GeV/c cut as well as an impact parameter significance cut of $|d_0|/\sigma_{d_0} > 2.5$. The impact parameter significance value is a metric for identifying tracks that do not point back to the primary vertex. Using these tracks, the SV algorithm attempts to reconstruct a secondary vertex in the cone of a jet. If the algorithm is successful, the jet is said to be b-tagged. The algorithm requires at least two tracks satisfying the track quality cuts to successfully tag a jet.

There are three operating points for the SV algorithm: ultra-tight, tight, and loose. These three operating points have different vertex reconstruction quality criteria which lead to different b-tag purities in each of the categories. The SV loose algorithm will tag many more jets than the ultra-tight SV algorithm, however the rate of b-tagging for jets that are actually light jets (jets originating from u, d, s, c quarks, or gluons) is drastically higher. This analysis uses two of the three operating points, the tight and loose SV algorithms.

Figure 7.4 shows the efficiency for identifying b-jets with the SV algorithm using both the tight and loose operating points. This efficiency depends on kinematics of the jets, such as the E_T and $|\eta|$. This dependence comes about mainly due to the necessary use of silicon tracking, which covers only a certain portion of the geometry of the detector. The tagging efficiency peaks at approximately 50% for central, high-E_T jets with the loose SV algorithm.

As previously mentioned, light jets can be spuriously tagged by the SV algorithm. Figure 8.6 shows the rate of this 'mistagging' as a function of both jet E_T and η. The mistag contribution is crucial to the modeling of the data, and will be discussed further in Chapter 8.

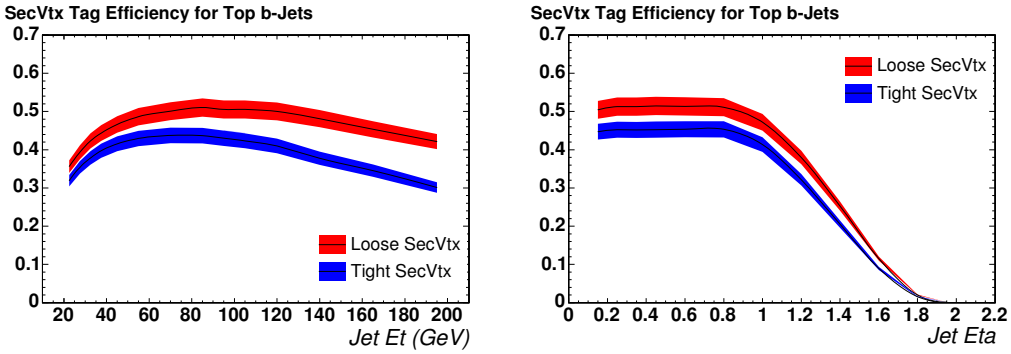


Figure 7.4: Tagging efficiency of the secondary vertex (SV) algorithm, as a function of jet E_T and jet η, for the tight and loose operating points.

Jet Probability Tagging

There is a second algorithm used in this analysis to identify jets consistent with the decays of b-hadrons. Instead of reconstructing displaced vertices as is done in the SV algorithm, the jet probability (JP) algorithm[4] uses the parameters of individual tracks that make up a jet. The JP algorithm determines a probability for a set of tracks associated with a given jet to have originated from the primary vertex in the event. As the output of the JP algorithm is a probability, the metric ranges from 0 to 1, with JP scores near 0 signifying a jet produced by long-lived particles such as b-hadrons. To construct this probability, the JP algorithm uses the impact parameter significance, d_0/σ_{d_0}. This quantity can be positive or negative, but is biased toward the positive for heavy flavor jets.

Since the JP algorithm outputs a continuous variable, any operating point can be chosen. Due to complications with the calculation of mistag rates, there are two operating points widely used at CDF. Jets are tagged with the JP algorithm if they have output probabilities less than 0.05 or less than 0.01 – the 5% and 1% operating points, respectively. Figure 7.5 shows the distribution of heavy flavor jet (those from b and c quarks) and light flavor jet output probabilities from the JP algorithm.

The tagging efficiencies for the 5% and 1% operating points of the JP algorithm are similar to those of the tight and loose operating points for the SV algorithm. Again, the tag rate is

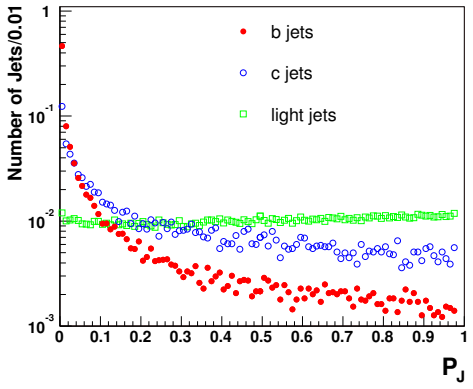


Figure 7.5: Distributions of jet probability output scores (P_J) for b-quark jets, c-quark jets, and light flavor jets. Heavy flavor jets have scores near zero, meaning they have low probability for originating at the primary vertex of the event.

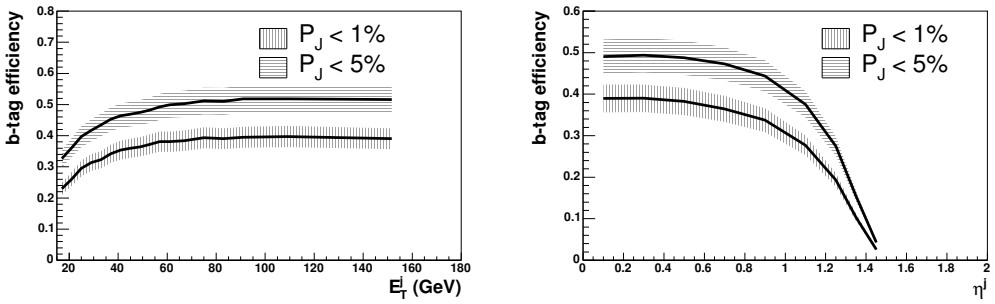


Figure 7.6: Tagging efficiency of the jet probability algorithm, as a function of jet E_T and jet η, for the 5% and 1% operating points.

near 50% for high-E_T central ($|\eta| < 1.0$) jets. Figure 7.6 shows the tagging efficiency for the JP algorithm as a function of jet E_T and η.

This analysis uses several combinations of the SV and JP algorithm operating points to define categories of events for this analysis. In order of increasing signal purity, the combinations used in this analysis are:

1. A single tight secondary vertex tag (SVT)

2. One loose secondary vertex tag and one jet probability (5%) tag (SVL + JP)

3. Two tight secondary vertex tags (SVT + SVT)

Splitting up events into these different tagging categories improves the overall sensitivity of the analysis, as sets of events with differing amounts of signal and background contributions are isolated into different regions of the analysis. If every event satisfying the preliminary event selection were put into a single category, the overall sensitivity of the analysis would be greatly reduced.

The reconstruction of muon and jet objects, with the addition of the application of b-tagging algorithms, forms the basis for this analysis. Once the selection criteria has been finalized, we search for events consistent with the expected final state in data, as well as in several Monte Carlo samples which comprise the signal and background models of this analysis. The following chapter details the background processes which contribute to this analysis, and the background model used to validate the selection applied to the data.

Chapter 8
Background Model and Data Selection

When searching for events consistent with the final state from the process $ZH \to \mu^+\mu^- b\bar{b}$, other events originating from non-signal processes are inevitably selected into the analysis. These processes contribute to the background expectation in the analysis. A robust understanding of the background model is necessary to make any statement about the expected sensitivity of the analysis to a new process, such as ZH. This chapter details the methods used to form the background model, which comes both from simulated Monte Carlo backgrounds, as well as some data-derived background contributions.

8.1 Monte Carlo Simulations

Many of the background processes for this analysis are modeled using Monte Carlo simulated events. To create these samples, an event generator is used to simulate the hard-scatter process which occurs in a collision event. The event generators used in this analysis are PYTHIA[41] and ALPGEN[36]. These generators use Monte Carlo statistical methods to simulate the collision events that are themselves probabilistic in nature. For example, after a given particle is produced in a collision event, it may radiate a gluon, it may radiate a photon, or it may do nothing at that moment. These subsequent events are simulated using random decisions in the software, based on many factors such as the particle energy and position in the detector. Ideally, these simulated events would exactly mimic the actual collision events produced in the detector. However, this is not always the case and a detailed verification study must be made to confirm that the background model, composed mainly of Monte Carlo events, is an accurate

simulation of the data.

The Monte Carlo simulations produced by PYTHIA and ALPGEN are not complete enough to use in physics analyses. These events must be simultaneously evaluated for interactions with detector material as they are produced with the event generators. The detector material is modeled with GEANT[8], a tool for simulating detector material which can be interfaced with the event generators. As the particles produced by the event generators travel through the simulated material in GEANT, the response of the detector is simulated. This part of the simulation includes the effects of magnetic and electric fields, as well as the shower formation for particles incident on the calorimeters. Figure 8.1 shows a comparison of electron conversion interactions with the detector material, for data and simulated Monte Carlo events.

Once the simulations are complete, the information from the simulated detector response is reconstructed using the same method as applied to collision data. This ensures that the same variables can be used for data and Monte Carlo events in offline physics analysis. Monte Carlo simulations give high-statistics samples of single processes. These samples are crucial to understanding the kinematics of background events, but especially for understanding the kinematics of predicted processes that have not yet been observed, such as production of Higgs bosons. We rely on Monte Carlo simulations to determine the expected amount of signal events produced, as well as their kinematics, to design a sensitive analysis.

8.2 Background Processes

There are several processes which can satisfy the event selection for this analysis, producing two oppositely charged muons and two or more high-E_T jets.

Z/γ^*+Jets Production

The main source of background before b-tagging is imposed is Z+jets production. This background includes the production of a Z boson that decays to a pair of muons, along with any number of jets. The jets could be produced from quarks or gluons directly from the hard scatter process, or they could be produced from gluon radiation from the initial or final states. The glu-

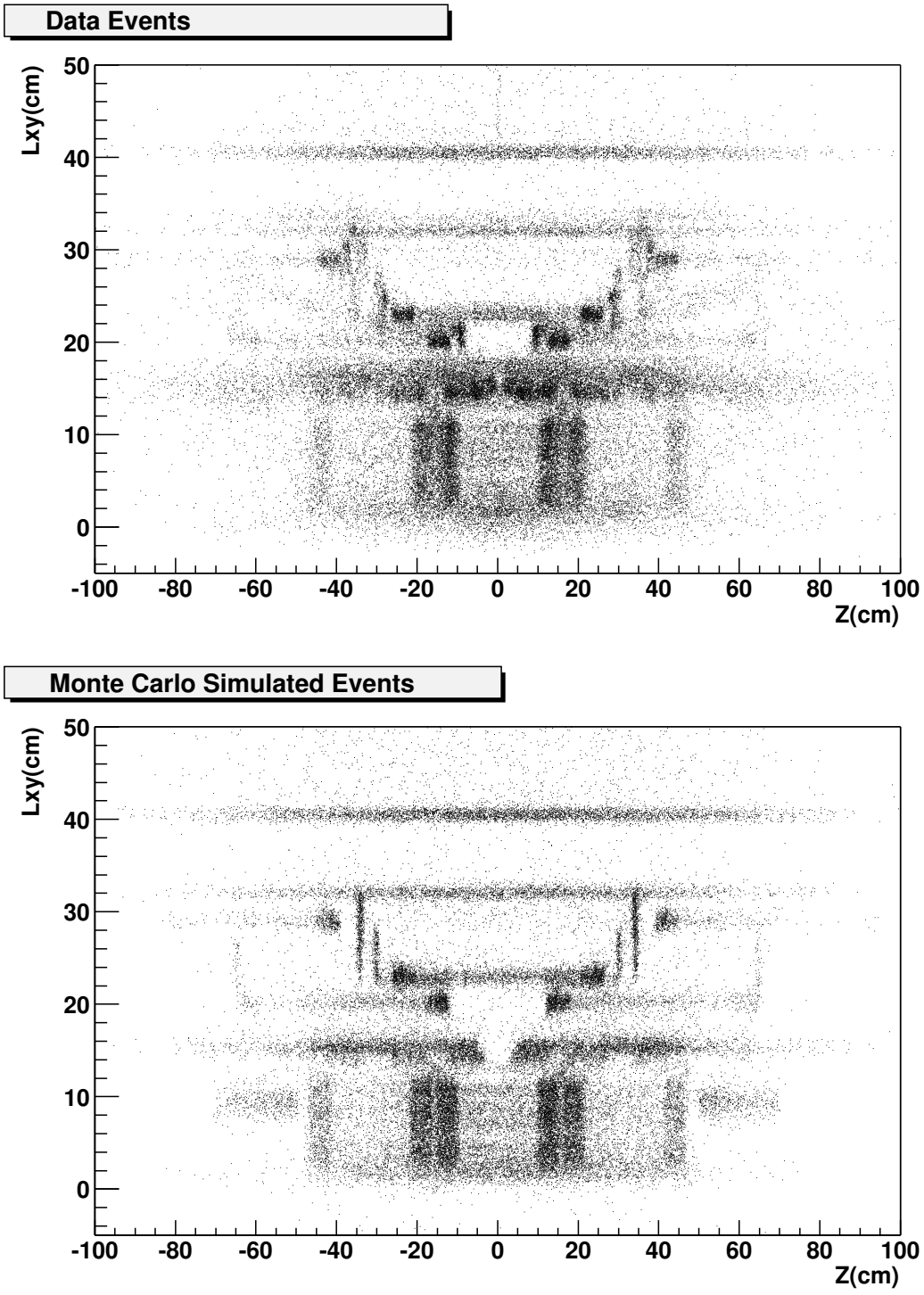


Figure 8.1: Comparison of interactions of conversion electrons and detector material for data (top) and Monte Carlo simulated (bottom) events. The simulation reproduces the data accurately, showing the correct distribution of detector material.

ons could futher split to form a pair of quark jets as well. This category of background includes events with up to as many as 10 jets satisfying the event selection criteria. Example Feynman diagrams of Z+jets production processes are shown in Figure 8.2.

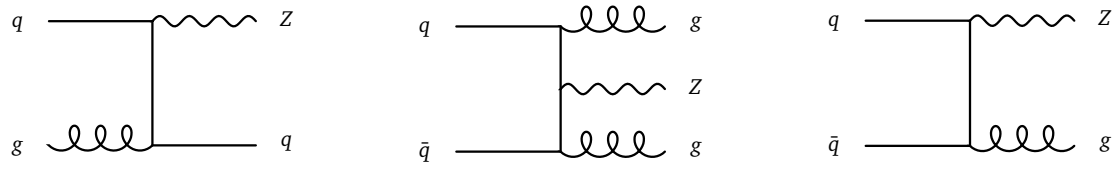


Figure 8.2: Example Feynman diagrams for production of the Z+jets background process. There are, of course, many different diagrams contributing to this category of background, including additional gluon radiations and splittings.

Although this background process contributes mostly to the analysis before the imposition of any b-tagging requirements, processes which produce b or c quarks do contribute even after the b-tagging requirement. The $Z + b\bar{b}$ process is one of the irreducible backgrounds of this analysis, since it contains two b-quark jets, the same final state that is expected from a $H \to b\bar{b}$ decay.

This background process also includes a contribution from the Drell-Yan continuum $\gamma^* \to \mu^+\mu^-$ background, if the muon pair is produced with additional jets. Therefore, we include Monte Carlo samples containing pairs of muons with invariant masses in the range $[20, 600]$ GeV/c^2 before applying the Z-window selection, to account for differences in the Monte Carlo production and event reconstruction which lead to differences in the calculated invariant mass.

Diboson Production

Another contribution to the background in this analysis is from the production of a pair of vector bosons. There are three possible diboson processes: ZZ, WZ, and WW production. Figure 8.3 shows Feynman diagrams for the diboson processes.

For ZZ production, one of the Z bosons decays to a pair of muons, while the other Z boson decays to a pair of quarks which produce jets in the detector. The ZZ process is the largest

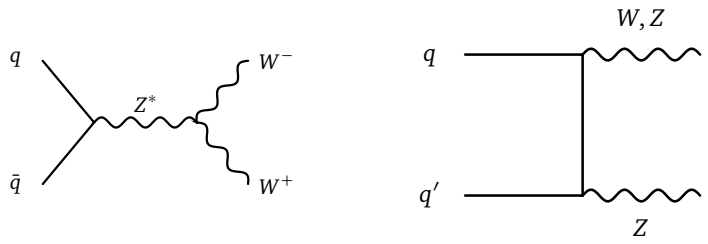


Figure 8.3: Feynman diagrams representing the diboson processes contributing to the background composition of this analysis.

contribution to the diboson background in this analysis. Additional jets may also be produced via initial or final state gluon radiation.

For WZ production, the muon pair must be produced by the decay of the Z boson. This means the W boson must decay hadronically to produce jets in the detector. Since the branching ratio for $W \rightarrow$ hadrons is approximately 68%, this is also a sizable contribution.

The WW process is a very small contribution to the analysis. It contributes when both W bosons decay via $W \rightarrow \mu\nu$, which has a branching ratio of only 10.5%. The required jets in the event must come from additional gluon splittings. Due to all of these factors, WW is a very small contribution before b-tagging, and essentially vanishes after the requirement of b-tagged jets.

Top Quark Production

Another background process that contributes mainly after the requirement of b-tagged jets is that of top quark pair production, or $t\bar{t}$ production. Since the dominant decay of the top quarks is $t \rightarrow Wb$, this process is a source of two b-quark jets. The two muons required in the event can either be produced from the W boson decays, or in semi-leptonic decays of other jets produced from non-leptonic W boson decays. One interesting feature of this background process is that there are no real Z bosons produced, so the invariant mass $m_{\mu\mu}$ will not have a peak at the Z boson mass, as in the other background processes. Figure 8.4 shows Feynman diagrams for two $t\bar{t}$ production processes.

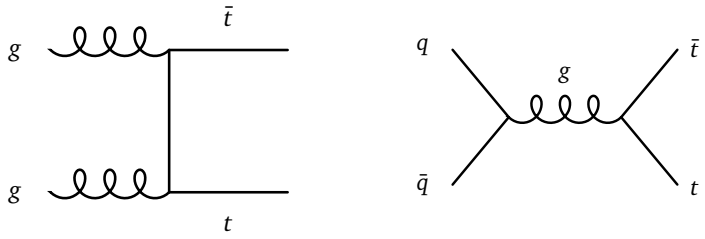


Figure 8.4: Feynman diagrams representing the $t\bar{t}$ processes contributing to the background composition of this analysis.

The processes mentioned above, including Z/γ^*+jets, diboson, and $t\bar{t}$ production are all modeled with Monte Carlo simulated data samples. There are some contributions to the background for this analysis which do not use Monte Carlo simulations, and must be modeled using data-based techniques.

Misidentified Z bosons

Occasionally, the muons that are selected in an event and used to form the Z boson candidate are not actually from the decay of a real Z boson. There are other processes that produce muons, which can be selected by the muon NN and form an invariant mass in the Z-window of $m_{\mu\mu} \in [76, 106]$ GeV/c^2. These processes include $\gamma^* \to \mu^+\mu^-$ decays and decays-in-flight of K and π mesons. Additionally, semi-leptonic decays of heavy flavor quarks (b and c) can produce muons which pass the event selection.

Because these other categories of events produce an equal number of positively and negatively charged muons, there will be an approximately equal number of $\mu^\pm\mu^\pm$ and $\mu^\pm\mu^\mp$ pairs. To model the contribution of this misidentified Z boson, or 'fake Z' background, we invert the opposite charge requirement on the muon pair and select events containing two muons with the same charge ('same-sign' muons).

Ignoring the Z boson mass peak, we observe that the amount of opposite-sign and same-sign muon pairs selected is equal, within uncertainties. The muon NN rejects most of this background process, although the remaining contribution is non-negligible. Figure 8.5 shows opposite-sign and same-sign dimuon events that fail the muon NN selection. Note that the

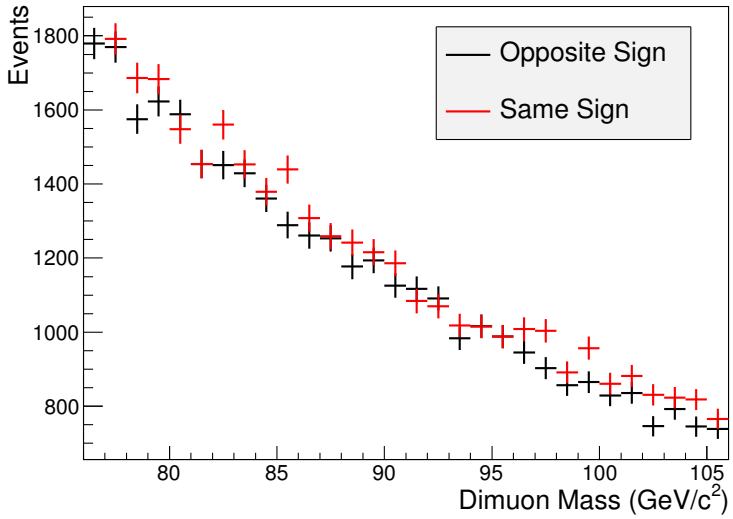


Figure 8.5: Same-sign (red) and opposite-sign (black) dimuon events which fail the muon NN selection. We use the same-sign selection to model the opposite-sign contribution of fake QCD muon production to the event selection, assuming equal proportions of same-sign and opposite-sign events. The difference between these two contributions is taken as a systematic uncertainty on the fake Z background normalization.

amounts of these two contributions are roughly equal, providing justification for using this method to model the fake Z boson background. We take the difference between these two contributions as a systematic on the normalization of the fake Z background (see also Chapter 10). Figure 7.2 also shows that the rate of this fake Z boson category is about 2 orders of magnitude smaller than the real Z boson contribution.

Mistagged Jets

This analysis uses the b-tagging algorithms described in Chapter 7 to select jets consistent with the hadronization of b-quarks. Because those algorithms are not 100% efficient, nor 100% accurate, the contribution of non-b-quark jets that are tagged under the algorithms must be added to the background model. To accurately measure the rate of mis-tagged jets entering the final event selection, the data is used to model this background process.

To measure the rate of mistagged light jets, a mistag matrix is used. This mistag matrix,

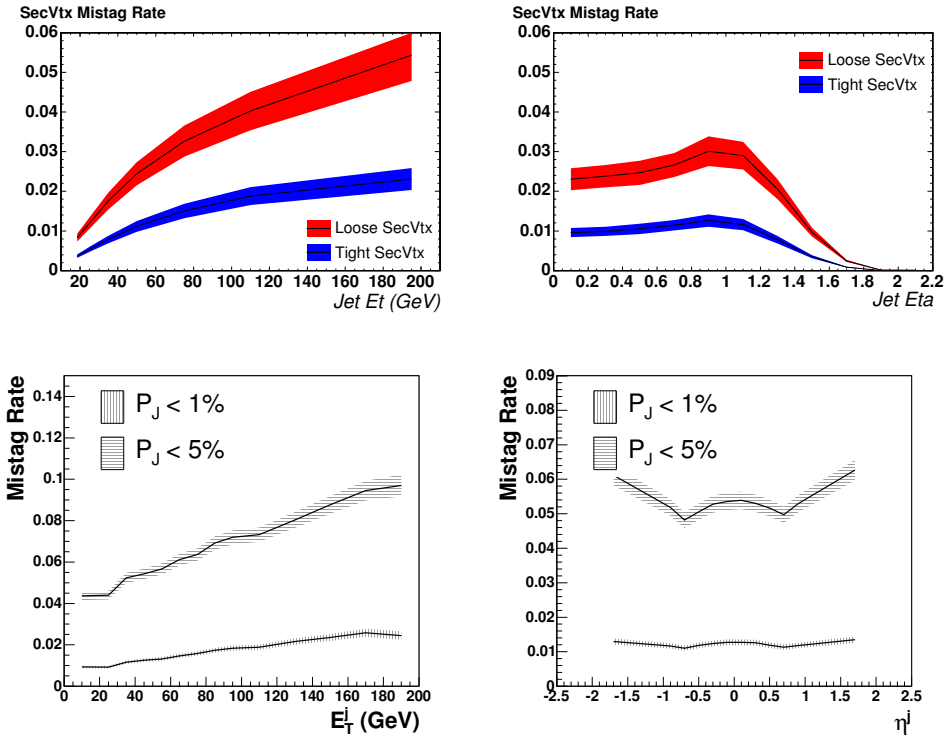


Figure 8.6: Mistag rates of the secondary vertex (SV) and jet probability (JP) algorithms, as a function of jet E_T and jet η, for the tight, loose (SV), 1%, and 5% (JP) operating points.

when applied to a given jet, returns a probability for that jet to receive a b-tag by the algorithm. The mistag matrix is parameterized using a variety of jet quantities.

The mistag matrix uses bins of jet E_T, jet η, the number of tracks used by the algorithm of interest, the number of reconstructed vertices in the event, as well as the z coordinate of the primary vertex in the event. A separate mistag matrix is used for the SV tight and SV loose operating points, as well as for the JP algorithm (at the 5% operating point). Figure 8.6 shows the rate of mistagged jets for the SV and JP algorithms.

We apply the mistag matrix to the data sample prior to the b-tag requirement to obtain an estimate of the mistag contribution to the three b-tag categories used in the analysis. More details on the combination of mistag samples for the different b-tag types is given later in this chapter.

8.3 Data Selection

We use data collected with the CDF II detector from February 2002 through September 2011. We require the data to be taken when the silicon system was fully functional, as the application of the b-tagging algorithms described in Chapter 7 requires silicon tracking information. With this requirement, the data corresponds to an integrated luminosity of 9.2 fb^{-1}, measured by the Cherenkov Luminosity Counter (described in Chapter 6).

The data at CDF is divided into a set of 'streams', or sets of triggers with similar physics requirements. Each stream is labeled by a single letter. For example, the 'B' data stream consists of triggers requiring the presence of high-p_T leptons. The 'E' stream consists of triggers requiring the presence of missing energy at the calorimeter level. This analysis uses the B and E streams. The B stream is chosen for obvious reasons, but the E stream is chosen due to the minimum-ionizing property of muons as they pass through the calorimeter. This property causes an imbalance of energy deposits in the calorimeter, and will show up as missing energy when the requirements for the trigger are evaluated. Further in the offline anaylsis, the missing energy is corrected for the presence of muons, but this data stream serves as a significant source of acceptance of dimuon events.

It is typical for CDF analyses to choose a specific trigger to select data events. For example, one might require the MUON_CMUP18 trigger. However, this trigger requires at lease one muon with transverse momentum $p_T > 18$ GeV/c to save the event for offline analysis. Similarly, in the E stream, one may choose to work only with the MET40 trigger, which requires $\not\!\!E_T > 40$ GeV at the calorimeter level to save the event. Making specific trigger requirements like this can reduce acceptance for an analysis, so we choose to use every available trigger comprising the B and E streams. A single data event could be obtained from both the B and E streams, but we choose to keep the streams orthogonal, with the B stream taking priority over the E stream.

Selecting such a large number of triggers for the offline analysis causes complications in modeling the trigger efficiency for Monte Carlo simulated events. This complication requires the introduction of a new multivariate technique which determines probabilities for any trigger in the B or E stream to fire for a given event. More detail of this method is given in a following

section.

8.4 Forming the Background Model

This analysis uses many Monte Carlo simulated data samples, corresponding to each of the background processes mentioned above. One process has millions of simulated events comprising the dataset, and the samples need to be weighted to the actual expected number of events corresponding to the integrated luminosity collected in the experimental dataset. This involves applying weights to Monte Carlo events. Some weights are the same for each event, such as is the case when normalizing to the process cross sections. Other weights depend on the kinematics of a given Monte Carlo record. This section details the normalization of Monte Carlo samples and application of kinematic-dependent weights.

Table 8.1 lists the Monte Carlo samples used in this analysis, along with their process descriptions and individual process cross sections. The first step in the normalization procedure is to weight the Monte Carlo events to the correct amount expected for a given luminosity. If a Monte Carlo dataset has N_{gen} generated records, of which N_{selected} pass the event selection, and a process cross section of σ, then the expected number of data events due to this process in a given amount of integrated luminosity L is given by Equation 8.1.

$$N_{\text{exp}} = \frac{\sigma \cdot L}{N_{\text{gen}}} N_{\text{selected}} \tag{8.1}$$

In practice, the normalization is done on an event-by-event basis by applying weights to each Monte Carlo event that passes the final analysis selection. Because of differing reconstruction efficiencies in the data and Monte Carlo samples, objects in the Monte Carlo need to have scale factors applied to account for these discrepancies. Additionally, the Monte Carlo samples do not have trigger simulations, and need trigger efficiencies applied as weights to each individual event. Equation 8.2 shows the full event weight, w_{ev}, applied to each Monte Carlo event. Data-derived background contributions do not receive any weight factors, as they have the same reconstruction efficiencies as the data used in the selection. The following subsections

Process	Generator	Sample	σ (pb)	K-factor
$Z(\rightarrow \mu\mu) + q\bar{q}$	P+A	ztktp5	158.0	1.4
	P+A	ztktp6	21.6	1.4
	P+A	ztkt7p	3.47	1.4
	P+A	ztkt8p	0.548	1.4
	P+A	ztkt9p	0.099	1.4
$Z/\gamma^*(\rightarrow \mu\mu) + q\bar{q}$	P+A	xtkt5p	160.0	1.4
	P+A	xtkt6p	8.40	1.4
	P+A	xtkt7p	1.60	1.4
	P+A	xtkt8p	0.233	1.4
	P+A	xtkt9p	0.040	1.4
$Z(\rightarrow \tau\tau) + q\bar{q}$	P+A	ztktt3	158.0	1.4
	P+A	ztktt4	21.5	1.4
	P+A	ztktt5	4.14	1.4
$Z(\rightarrow \mu\mu) + c\bar{c}$	P+A	ztktc5	1.08	1.4
	P+A	ztktc6	0.331	1.4
	P+A	ztktc7	0.107	1.4
$Z(\rightarrow \mu\mu) + b\bar{b}$	P+A	ztktb5	0.511	1.4
	P+A	ztktb6	0.134	1.4
	P+A	ztktb7	0.0385	1.4
$t\bar{t}$	P	ttop25	7.04	1.0
WW	P	ihht1a	11.34	1.0
WZ	P	jhht1a	3.47	1.0
ZZ	P	khht1a	3.62	1.0

Table 8.1: A list of the Monte Carlo samples used in the background model for this analysis. Also shown are the process cross sections and K-factors applied to the samples, as well as the event generators used to produce the samples (P ≡ PYTHIA, A ≡ ALPGEN).

detail the origin of each of these terms.

$$w_{\text{ev}} = \frac{1}{N_{\text{gen}}} \cdot \sigma \cdot L \cdot \left[\epsilon^{(\text{B})}_{\text{trig}} + \epsilon^{(\text{E})}_{\text{trig}} (1 - \epsilon^{(\text{B})}_{\text{trig}}) \right] \cdot F^{(1)}_{\mu} F^{(2)}_{\mu} \cdot F_{\text{tag}} \cdot F_{z_0} \cdot K_{\text{ALPGEN}} \tag{8.2}$$

Trigger Efficiencies

This analysis uses a trigger selection which simultaneously selects events from all available triggers in a data stream. To model the efficiency for a given Monte Carlo event to be accepted by any one of the triggers in the streams of interest is a complicated approach, as the triggers have dependences on various kinematic quantities that are correlated with one another. If this analysis used just a single trigger, modeling the behavior of this trigger for Monte Carlo events would be simple. Many trigger efficiencies can be parameterized using a single kinematic quantity, such as the muon p_T. This parameterization gives a function which can be used to determine the trigger efficiencies for any event.

Because of the multitude of triggers involved, and the complicated correlations between kinematic variables, we use a multivariate approach to measure the trigger efficiencies using collision data events. This method uses the two data streams from which events are selected for the analysis, the `B` and `E` streams. Since these two sets of data events are orthogonal, we can use one stream to measure the trigger efficiency for the second stream. For an event that satisfies a trigger requirement in the `B` stream, we can ask how often that event also was recorded by one of the triggers in the `E` stream – and determine the dependence of its presence in that stream on different kinematic quantities. This process is similar to trigger efficiency measurements for single triggers, but we use two sets of triggers instead of two individual triggers.

The kinematic dependence of the trigger efficiencies is determined by training an ensemble of artificial neural networks, using the kinematic variables of interest as inputs. The neural networks use regression to determine stream efficiencies for the entire kinematic phase space. The regression target is '1' if the event did appear in the alternate stream, and '0' if the event did not appear in the alternate stream. The output of the function then ranges from 0 to 1 and can be interpreted as a probability for an event with those kinematics to appear in the stream of interest. These trigger efficiency functions are measured using data, and the response using

Kinematic Variables Used For Regression

B **Stream**	E **Stream**
Presence in E Stream	Presence in B Stream
Muon 1,2 p_T	Number of Jets
Muon 1,2 η	Jet 1,2,3 E_T
Muon 1,2 ϕ	Jet 1,2,3 η
Muon 1,2 Fiduciality	$\Delta R(j_1, j_2)$
Muon 1,2 CMX Stub	Vertex $\not{E}_T$
Muon 1,2 CMUP Stub	

Table 8.2: Kinematic quantities used in the regression-based determination of trigger efficiencies for modeling the B and E stream data selection. The kinematics of jet 3 are only used in the E stream function when there is a third jet present in the event.

Monte Carlo simulated events as inputs is used to emulate the trigger selection for this analysis.

There are two functions derived using this method – one to determine an efficiency for events to appear in the B stream, $\epsilon_{\text{trig}}^{(\text{B})}$ and one for the E stream efficiency $\epsilon_{\text{trig}}^{(\text{E})}$. The kinematic variables used for the regression training are listed in Table 8.2. The function for the B stream uses mainly muon kinematic properties, while the function for the E stream contains mainly jet-dependent quantities. This is due to the kinematic properties of the triggers making up each of the streams.

We validate this technique by comparing the regression-derived function for the stream efficiency as a function of the kinematic variables used in training to the actual measured efficiency from the data events themselves. Figure 8.7 shows the dependence of the trigger efficiency on the transverse momentum of the muon – this is one of the variables used in the regression procedure.

This regression technique is used to obtain the efficiencies $\epsilon_{\text{trig}}^{(\text{B})}$ and $\epsilon_{\text{trig}}^{(\text{E})}$ appearing in Equation 8.2. Because the B stream takes priority over the E stream, we veto events selected in the E stream that are also selected in the B stream. To model this selection process in the Monte Carlo samples, we apply the efficiency term $\left[\epsilon_{\text{trig}}^{(\text{B})} + \epsilon_{\text{trig}}^{(\text{E})}(1 - \epsilon_{\text{trig}}^{(\text{B})})\right]$, to mimic Monte Carlo events selected by the B stream or the E stream, but *not* the B stream. A single Monte Carlo record contributes to each category.

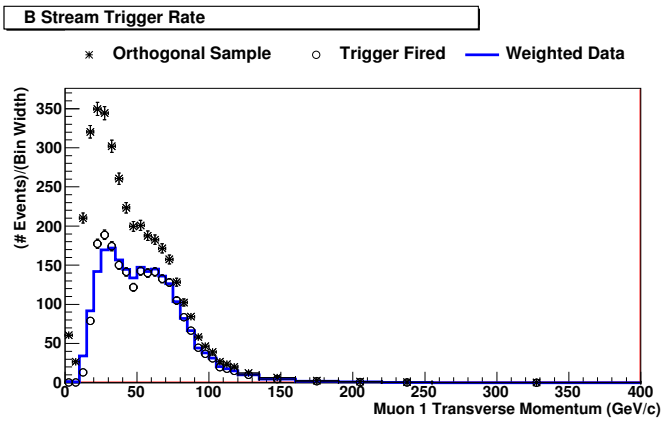


Figure 8.7: Validation of the regression-derived trigger efficiency, using the transverse momentum of the muon as an example.

Muon NN Scale Factor

Due to differences in the data and Monte Carlo reconstruction, the efficiencies for a muon to pass the muon NN selection differ in each case. To form an accurate background model using the muon NN, these differences must be taken into account with the use of scale factors. Scale factors are introduced to weight the Monte Carlo to equalize these reconstruction efficiencies. Ideally, the scale factors will be close to 1.0, as the Monte Carlo is expected to be a very close simulation of the detector response.

We measure a scale factor for the muon NN selection using the 'tag-and-probe' method. In this method, we first search for a single muon using a set of tightened cuts, to increase the quality of this so-called 'tag' muon. The tag muon for this muon must be a CMX or CMUP type muon, satisfying the cuts for those types listed in Table 7.1. After finding a tag muon, we search for any second muon forming an invariant mass $m_{\mu\mu}$ in the range $[86, 96]$ GeV/c^2. The mass window is restricted compared to the window used in the nominal event selection, to increase the likelihood that the muon pair is from a Z boson decay. If this second muon passes these criteria, it is the 'probe' muon. Because the probe muon has no cuts applied, we require the missing energy $\not\!\!E_T$ to be less than 10 GeV to reduce background contamination from QCD fake muon production in this measurement.

After a probe muon candidate is found, we evaluate the muon NN for this muon to determine if it would satisfy the selection. From this, we can determine an efficiency $\epsilon = N_{\text{pass}}/N_{\text{found}}$, the ratio of the number of probe muons passing the muon NN selection to the total number of probe muons found. Measuring this efficiency for a data sample and a Monte Carlo sample leads to the determination of the scale factor, as in Equation 8.3.

$$F_\mu = \frac{\epsilon_{\text{DATA}}}{\epsilon_{\text{MC}}} \tag{8.3}$$

We measure the muon NN scale factor as a function of the 11 kinematic variables used in its application. Figure 8.8 shows the scale factor measurements as a function of these different variables. The scale factor does not depend on many of these variables, which is expected. However, due to some dependence of the scale factor on the p_T of the muons, we choose to measure the scale factor separately for muons with $p_T > 40$ GeV/c and those with $p_T < 40$ GeV/c. To obtain the final value for F_μ we fit a horizontal line to the scale factors measured for each individual kinematic variable and take the average over all variables. Using this procedure, we obtain the results shown in Equation 8.4.

$$F_\mu = \begin{cases} 0.969 \pm 0.009 & \text{if } p_T < 40 \text{ GeV}/c \\ 0.992 \pm 0.007 & \text{if } p_T > 40 \text{ GeV}/c \end{cases} \tag{8.4}$$

These scale factors are applied as part of the event weight in Equation 8.2, one factor for each muon depending on its transverse momentum. This scale factor takes the place of the normal identification and reconstruction scale factors for sets of muon cuts (such as those in Table 7.1). The uncertainties on the measurement of the muon scale factors will be taken as a systematic uncertainty in the calculation of the final results. More information on the systematic uncertainties applied to this analysis can be found in Chapter 10.

b-tag Scale Factors

Similar to the scale factor for the muon NN, there are also differences in the b-tagging efficiencies for data and Monte Carlo simulated events. Therefore, another scale factor needs to be

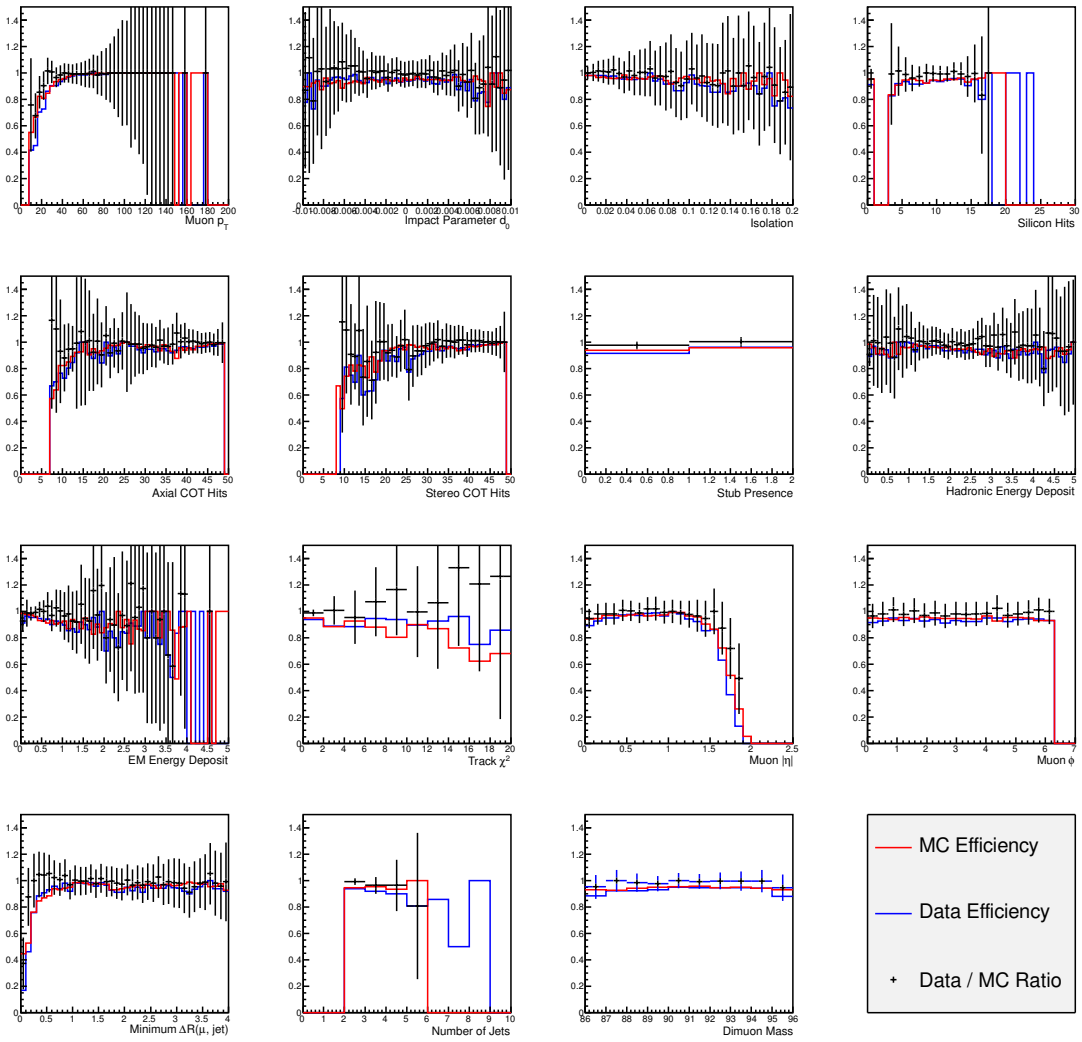


Figure 8.8: Scale factor measurements for the muon NN, as a function of the kinematic variables used. To obtain the final scale factor value, the measurements are averaged over all kinematic variables after fitting a horizontal line to each. There are two regions defined, for muon p_T $<$ / $>$ 40 GeV/c, each with a different scale factor value.

***b*-tag Scale Factors**	
Secondary Vertex, Tight Tag (F_T)	0.96 ± 0.05
Secondary Vertex, Loose Tag (F_L)	0.98 ± 0.07
Jet Probability, $< 5\%$ Tag (F_J)	0.78 ± 0.04

Table 8.3: The values for the b-tag scale factors used in this analysis, for each of the tagging algorithms and operating points.

applied to events which satisfy the b-tagging requirements. There exists a different scale factor for each operating point of each b-tagging algorithm. The scale factors used in this analysis are listed in Table 8.3. The scale factors are measured by the CDF high-p_T b-tag group, with procedures outlined in Ref. [25] and [38].

In the Monte Carlo, a scale factor is applied for each of the first two jets that are tagged in the event (we choose to only evaluate the two leading jets for b-tags), depending on the type of tag on the jet. For example, an event with two SVT tagged jets would be weighted by a b-tag scale factor of $F_{\text{tag}} = (0.96)^2$. However, there is an additional complication caused by the hierarchy of tag catgories. Again, the three tag categories used are as follows:

1. A single tight secondary vertex tag (SVT)

2. One loose secondary vertex tag and one jet probability (5%) tag (SVL + JP)

3. Two tight secondary vertex tags (SVT + SVT)

Events that fail the first tag category requirements still have a chance to satisfy the second or third category requirements. This effect needs to be accounted for in the Monte Carlo samples. After applying the appropriate b-tag scale factors to one category, the 'un-tagged' portion of the event (the difference between 1 and the scale factor combination applied) still can satisfy the requirements of another tagging category. This effect is known as Monte Carlo b-tag migration, and leads to complicated b-tag scale factors for each of the three tagging categories used in the analysis. Equations 8.5, 8.6, and 8.7 describe the computation of the b-tag scale factors for

each of the three tag categories listed above.

$$F_{\text{SVT+SVT}} = F_T \cdot F_T \tag{8.5}$$

$$F_{\text{SVL+JP}} = F_L \cdot F_J + 2 \cdot \frac{N_{\text{SVT+SVT}}}{N_{\text{SVL+JP}}} \cdot F_T \cdot (1 - \sqrt{F_L \cdot F_J}) \tag{8.6}$$

$$\begin{aligned} F_{\text{SVT}} &= F_T + 2 \cdot \frac{N_{\text{SVL+JP}}}{N_{\text{SVT}}} \cdot F_T \cdot (1 - \sqrt{F_L \cdot F_J}) \\ &\quad + 2 \cdot \frac{N_{\text{SVT+SVT}}}{N_{\text{SVT}}} \cdot F_T \cdot (1 - F_T) \cdot (1 - \sqrt{F_L \cdot F_J}) \end{aligned} \tag{8.7}$$

The factors F_T, F_L, and F_J are defined in Table 8.3. These b-tag scale factors are used in place of F_{tag} in Equation 8.2 when determining weights for Monte Carlo events in the various tag categories.

Other Factors

The other remaining factors that are applied to Monte Carlo events in the background model are a scale factor for the primary vertex position cut, and a K-factor for Monte Carlo samples produced with ALPGEN.

In our event selection, we choose to cut events having the primary vertex position $|z_0| > 60$ cm. When the Monte Carlo samples used to model the background are created, this same selection is made to generate simulated events with $|z_0| < 60$ cm. With this cut having different efficiencies for data and Monte Carlo events, there is a scale factor applied to Monte Carlo events to account for this difference. The scale factor used in this analysis is $F_{z_0} = 0.972 \pm 0.001$. This balances the efficiencies for the cut in Monte Carlo and data.

When simulating processes with many jets, the Monte Carlo generators only use the hard-scatter processes at leading-order, meaning the generated events are from tree-level Feynman diagrams only. However, the processes that we are trying to simulate can occur at all orders. Therefore, the cross-section of the simulated processes will not match the cross-section of the data events. This discrepancy is accounted for by a K-factor. The K-factor is a correction to the Monte Carlo simulated cross-sections, and is the ratio of next-to-leading-order (NLO) to leading-order (LO) cross section calculations. Just going to NLO from LO is sufficient to get an accurate background model. The K-factor used in this analysis is $K_{\text{ALPGEN}} = 1.4 \pm 0.4$, and is

applied only to processes simulated with ALPGEN, as described in Table 8.1.

8.5 Background Model Validation

After running the event selection on the entire dataset and all background Monte Carlo samples in Table 8.1, as well as creating the data-derived background contribution samples, we verify the background model by making several plots of kinematic quantities to check agreement in different variables. This chapter shows these validation plots before moving on to advanced techniques of the analysis, such as event discriminants.

Table 8.4 shows the expected numbers of events for each background process, as well as the observed number of data events, prior to the application of b-tagging. The background estimation is in good agreement with the observed number of data events.

The following figures show the background model in several kinematic variables.

Process	**Expected Events**		
$Z(\rightarrow \mu\mu, \tau\tau) + q\bar{q}$	14279.0	$\pm$	2855.8
$Z(\rightarrow \mu\mu) + c\bar{c}$	1001.3	$\pm$	400.5
$Z(\rightarrow \mu\mu) + b\bar{b}$	558.1	$\pm$	223.2
$t\bar{t}$	127.0	$\pm$	12.7
ZZ	172.0	$\pm$	10.3
WZ	153.9	$\pm$	9.2
WW	6.8	$\pm$	0.4
Fake $Z \rightarrow \mu\mu$	1133	$\pm$	56.7
Total Background	17431.8	$\pm$	2893.0
Observed Data	17881		

Table 8.4: Event yields for this analysis, prior to the application of b-tagging. This pre-tag region serves to validate the analysis framework.

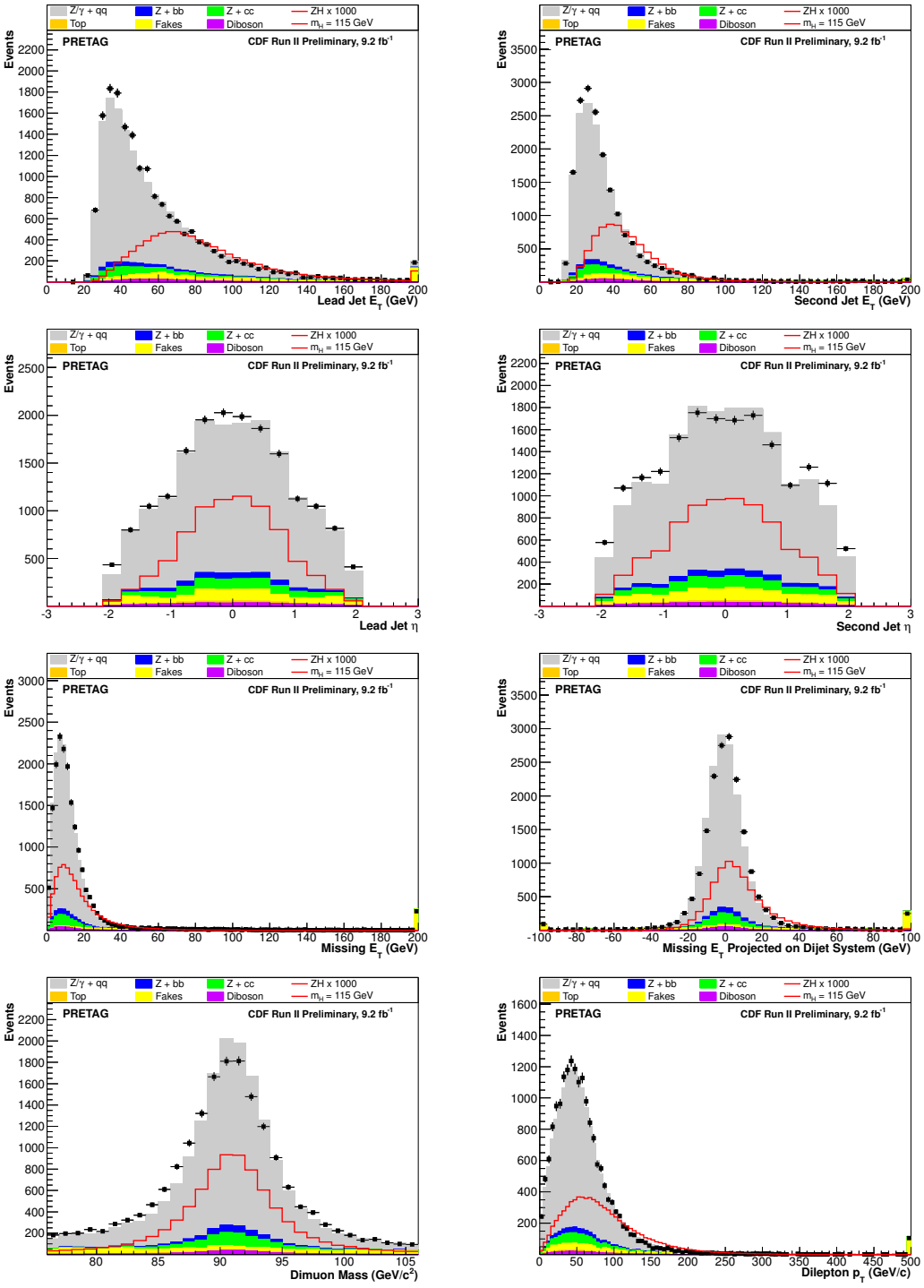


Figure 8.9: Kinematic distributions to compare the background model used in this analysis with the observed data events. The last bin on the right side of the distributions includes overflow contributions.

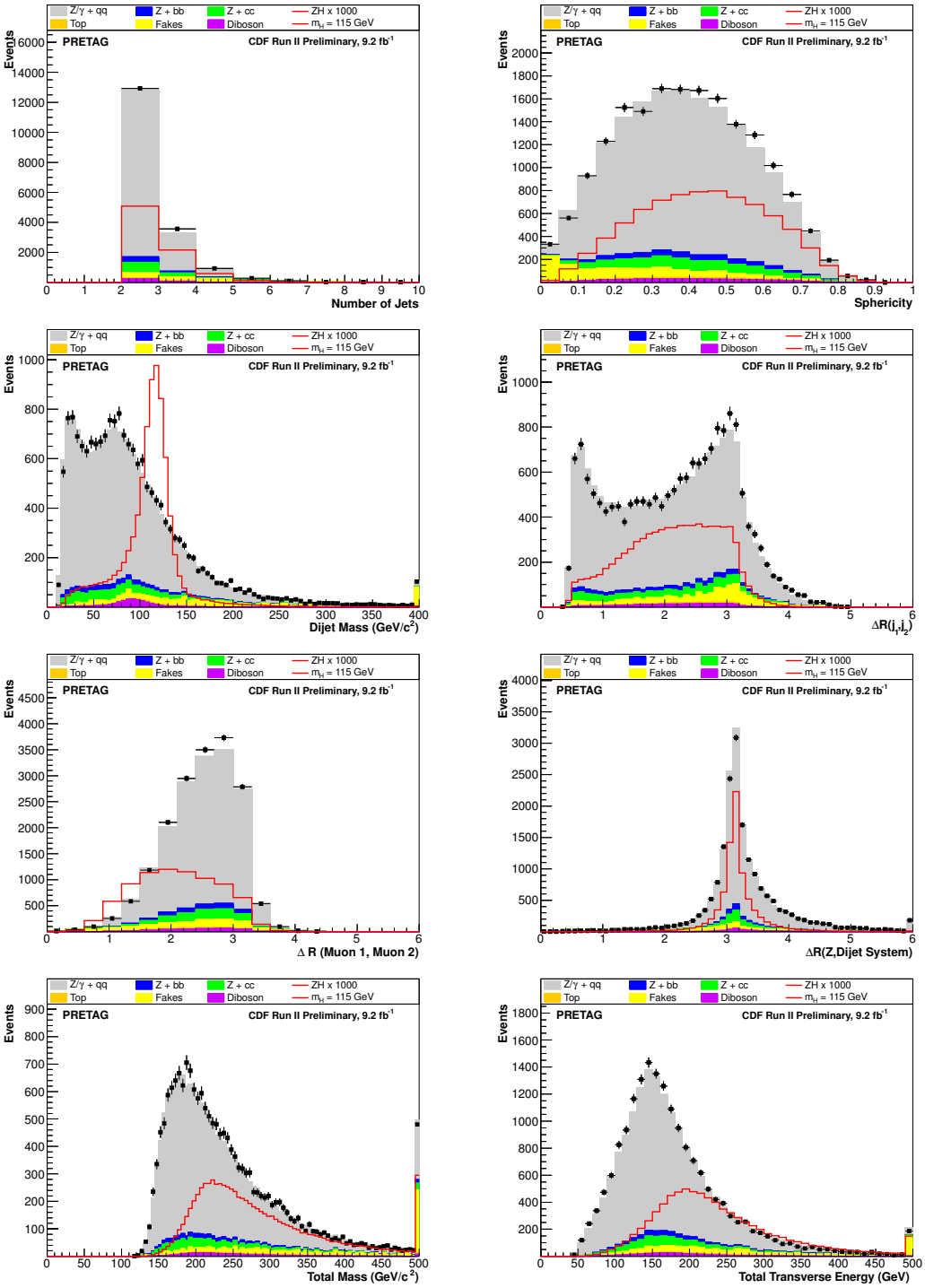


Figure 8.10: Kinematic distributions to compare the background model used in this analysis with the observed data events. The last bin on the right side of the distributions includes overflow contributions.

Process **Expected Events**	SVT	SVL+JP	SVT+SVT
$Z \to \mu\mu$+ Mistags	260.0 ± 35.1	22.1 ± 6.2	1.3 ± 0.4
$Z \to \mu\mu + c\bar{c}$	71.7 ± 28.7	11.8 ± 4.7	2.0 ± 0.8
$Z \to \mu\mu + b\bar{b}$	148.3 ± 59.3	33.4 ± 13.4	22.0 ± 8.8
$t\bar{t}$	40.2 ± 4.0	18.6 ± 1.9	15.3 ± 1.5
ZZ	15.1 ± 0.9	4.9 ± 0.3	3.4 ± 0.2
WZ	5.4 ± 0.3	0.5 ± 0.01	–
WW	0.3 ± 0.01	–	–
Misidentified $Z \to \mu\mu$	33 ± 1.7	1 ± 0.1	1 ± 0.1
Total Background	574.0 ± 74.8	92.2 ± 15.6	45.1 ± 9.0
Observed Data	631	107	48

Table 8.5: Event yields for the three tagging categories used in this analysis.

With the background sufficiently modeled before the application of b-tagging, we proceed to form distributions of kinematic quantities for the three b-tag categories used in the analysis. Again, the expected number of events agrees with the observed data events, within uncertainties. Table 8.5 shows the event yields for each of the three tagging categories, and Figures 8.11 through 8.16 show distributions of kinematic variables for each of the three b-tagging categories used in this analysis.

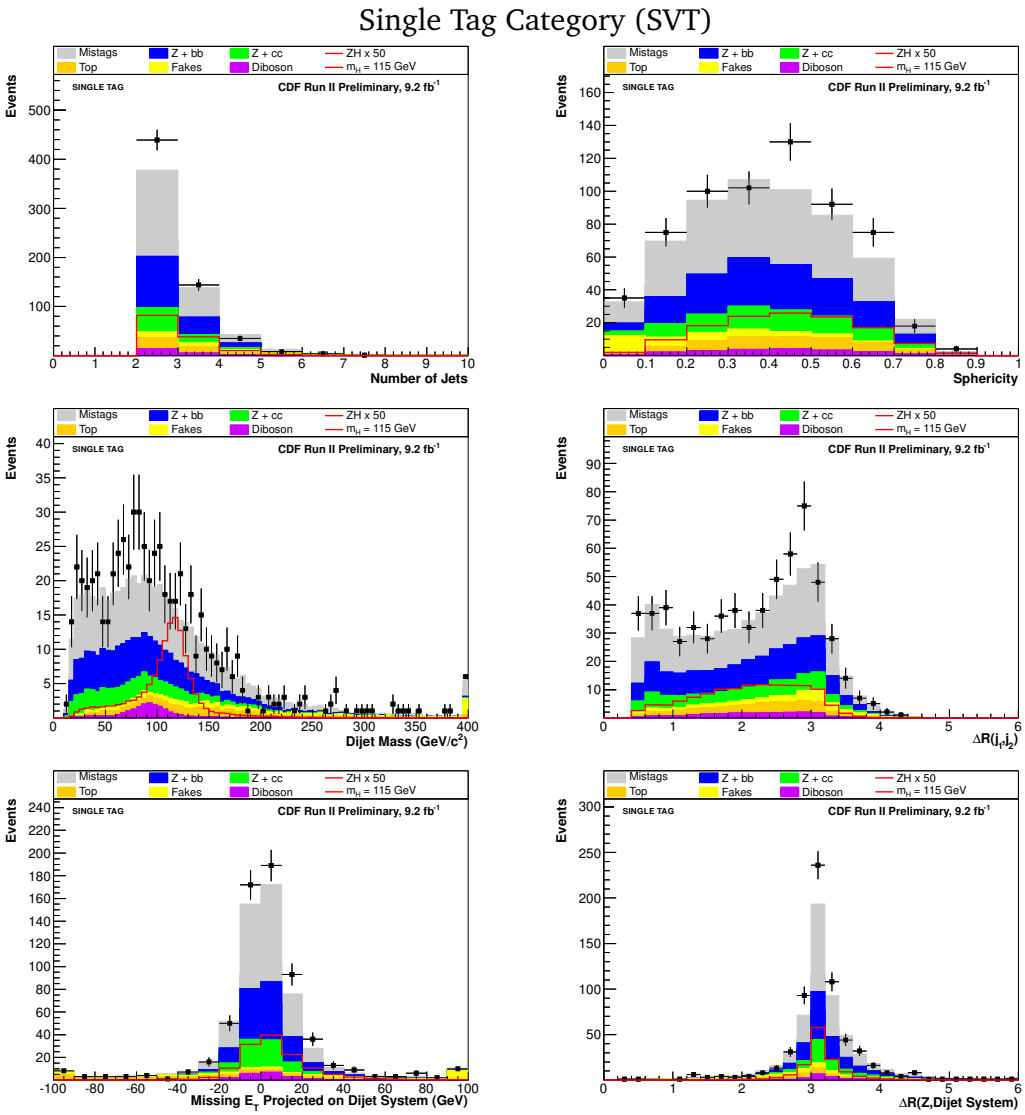


Figure 8.11: Distributions of some kinematic variables for the single b-tag category. The last bin on the right side of the distributions includes overflow contributions.

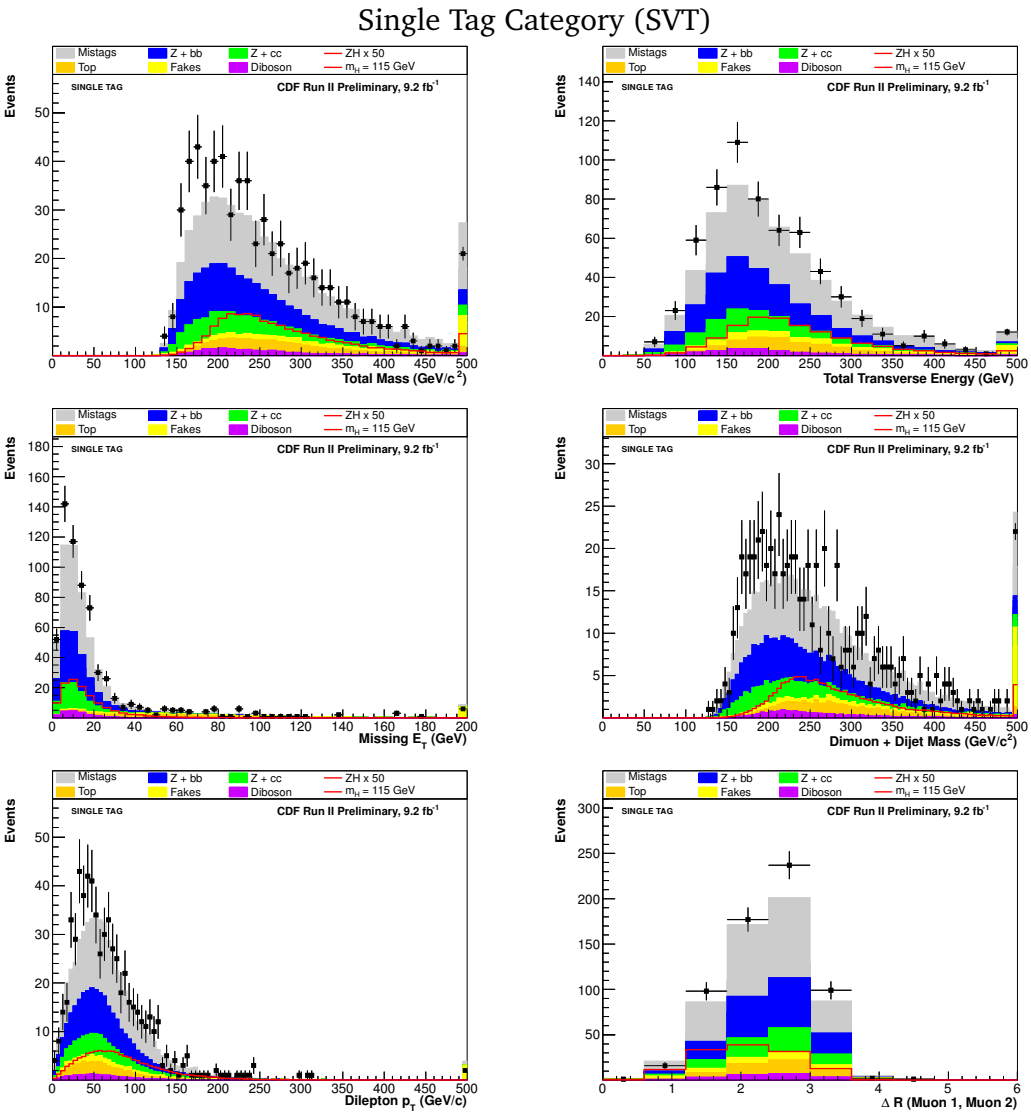


Figure 8.12: Distributions of some kinematic variables for the single b-tag category. The last bin on the right side of the distributions includes overflow contributions.

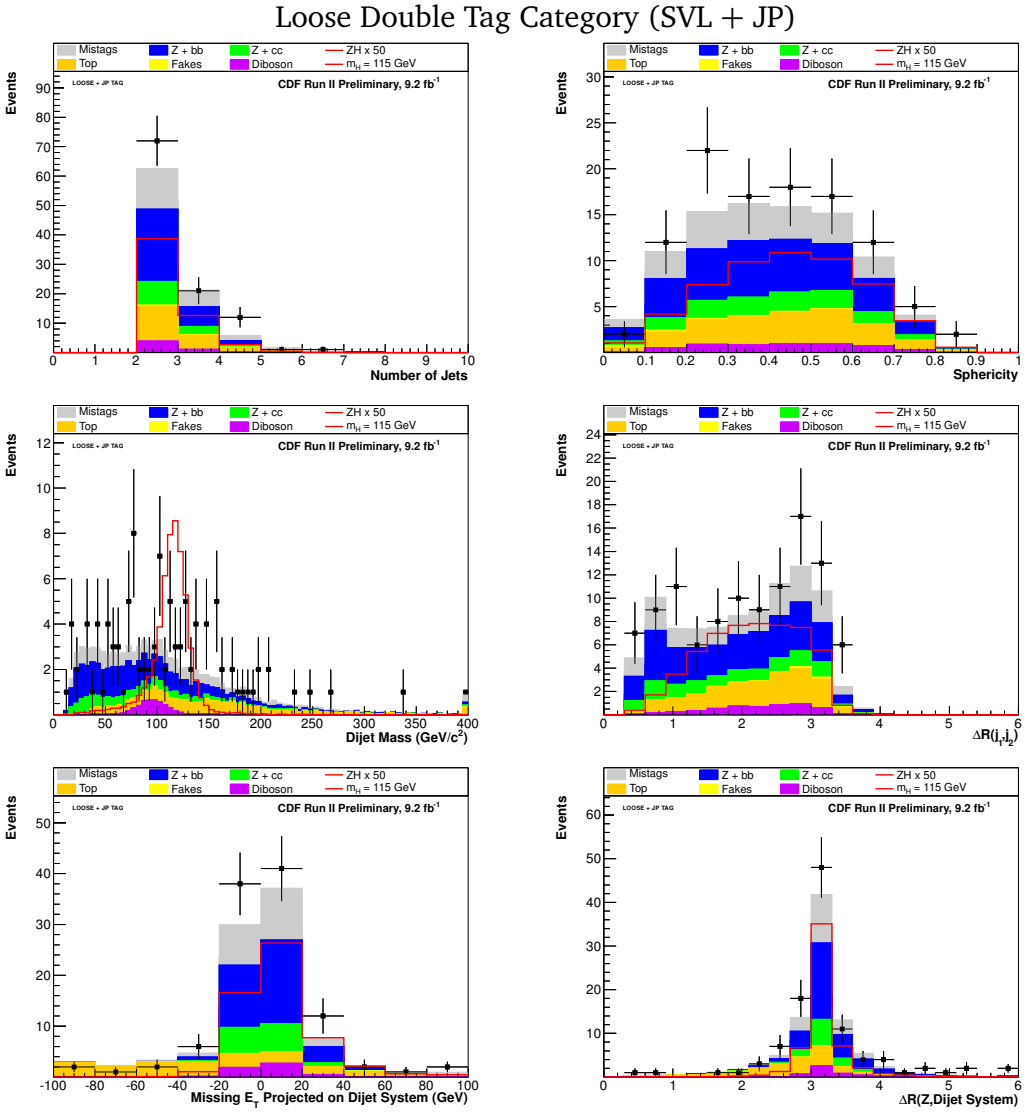


Figure 8.13: Distributions of some kinematic variables for the loose double *b*-tag category, consisting of one secondary vertex tag and one jet probability (5%) tag. The last bin on the right side of the distributions includes overflow contributions.

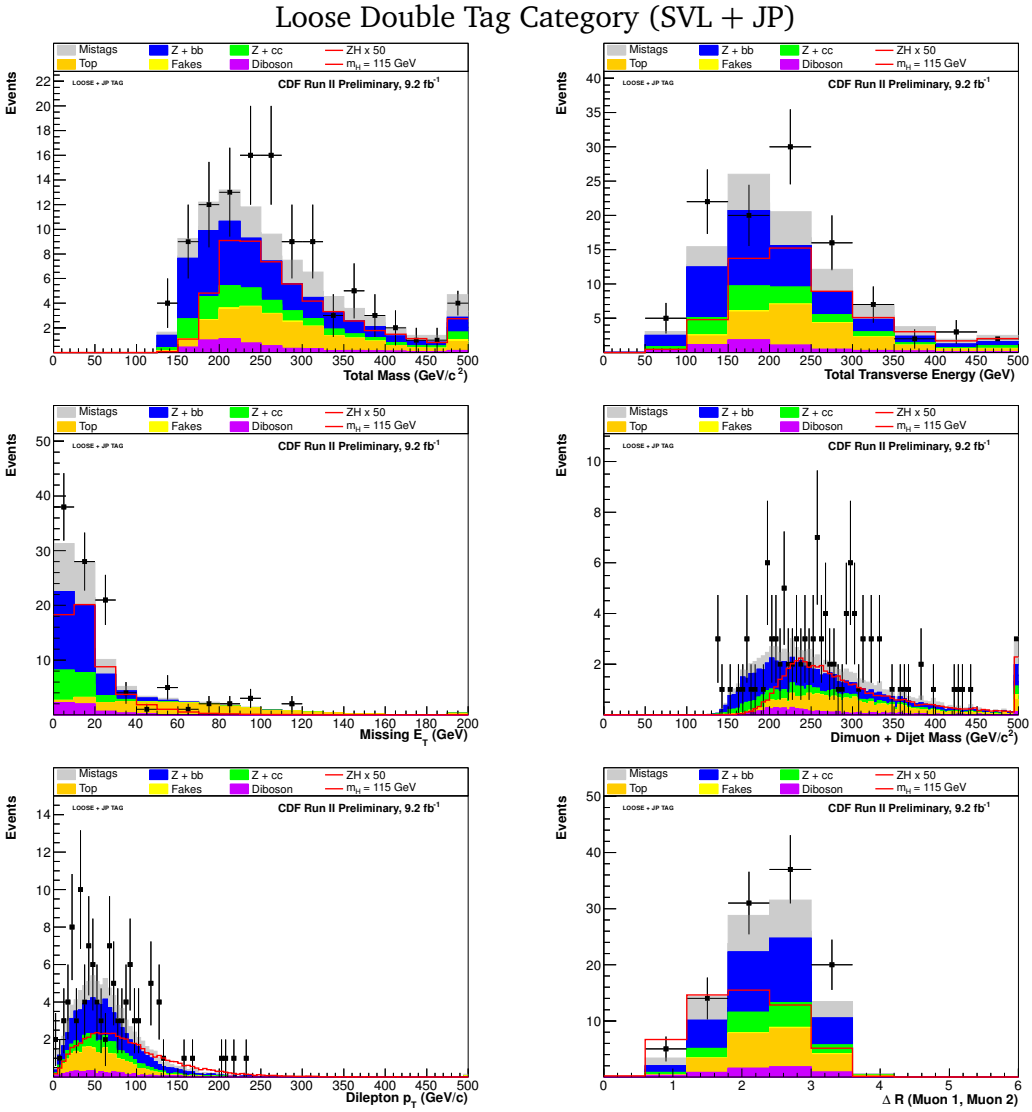


Figure 8.14: Distributions of some kinematic variables for the loose double *b*-tag category, consisting of one secondary vertex tag and one jet probability (5%) tag. The last bin on the right side of the distributions includes overflow contributions.

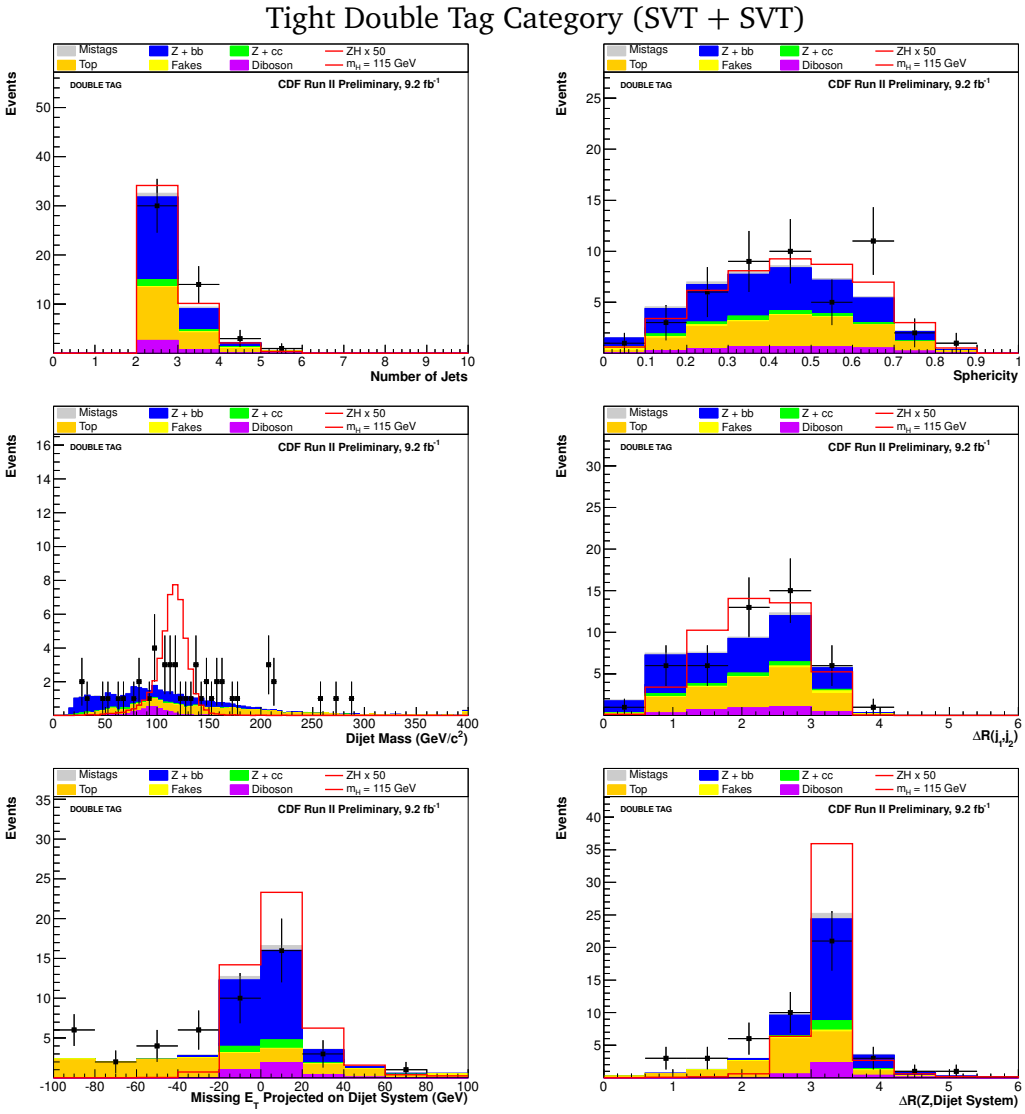


Figure 8.15: Distributions of some kinematic variables for the tight double *b*-tag category, consisting of two secondary vertex tags. The last bin on the right side of the distributions includes overflow contributions.

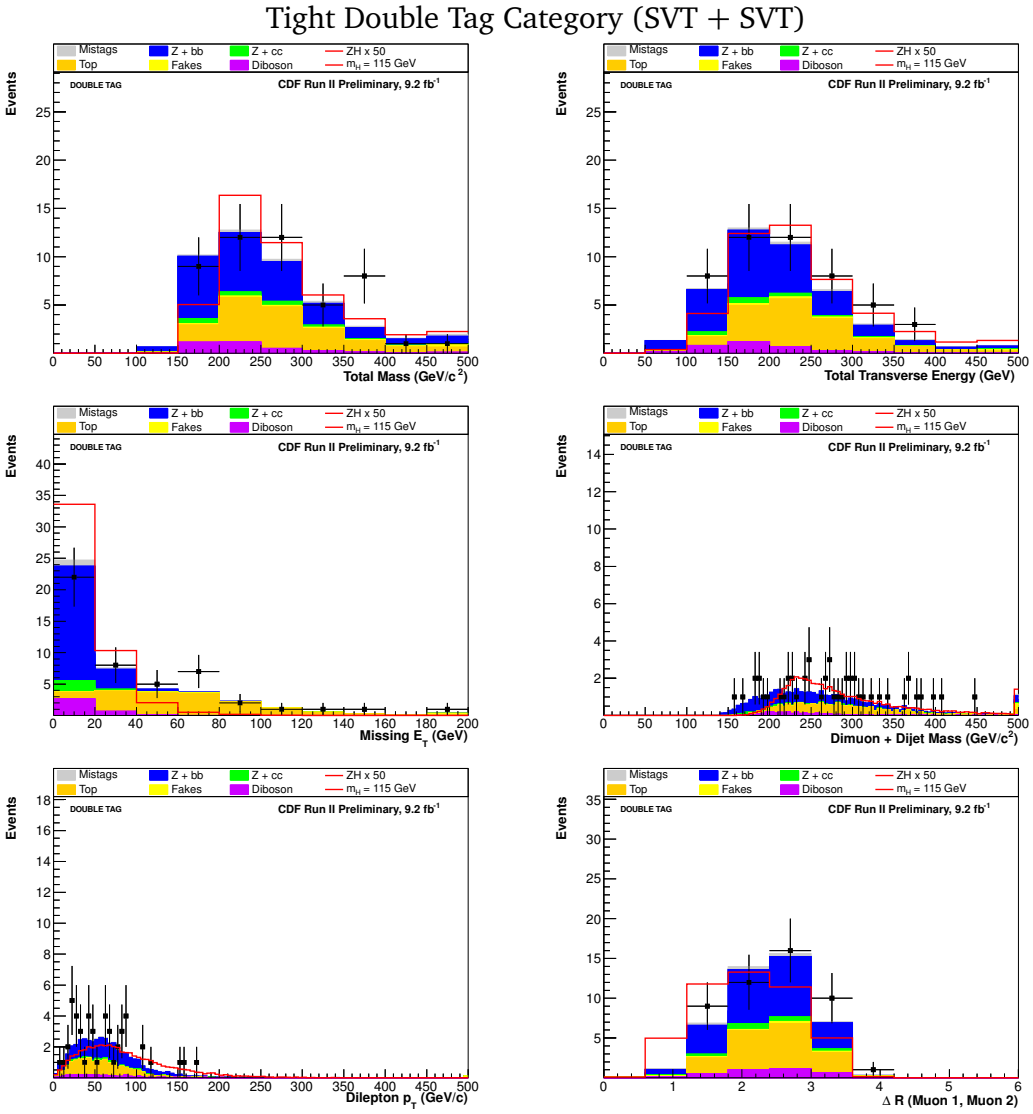


Figure 8.16: Distributions of some kinematic variables for the tight double *b*-tag category, consisting of two secondary vertex tags. The last bin on the right side of the distributions includes overflow contributions.

Chapter 9

Multivariate Techniques for Discriminating Signal

After validating the background model used in this analysis, we aim to improve the sensitivity to the $ZH \to \mu^+\mu^- b\bar{b}$ process. We use several new multivariate techniques to discriminate signal candidate events from the large, overwhelming background contribution to this analysis. This chapter describes these techniques in detail.

9.1 Signal Model

As for the background processes, we use a set of Monte Carlo samples to model the ZH signal process for which this analysis searches for. There are 11 different Monte Carlo samples used, each corresponding to a different value of the Higgs boson mass, m_H, in 5 GeV/c^2 steps, for $100 \leq m_H \leq 150$ GeV/c^2. These samples are generated with PYTHIA, and do not require the application of any K-factor.

The signal Monte Carlo samples used have forced $Z \to \ell^+\ell^-$ decays, meaning the Z boson is only allowed to decay to charged leptons in the Monte Carlo events. We must take this into account in the number of generated Monte Carlo events by multiplying by the branching ratio for $Z \to \ell^+\ell^-$, which is approximately 10%[37]. Additionally, these samples have a forced decay of the Higgs boson, $H \to b\bar{b}$. We must multiply by the branching ratio for $H \to b\bar{b}$ for each of these samples as well. Table 9.1 shows the ZH production cross sections used[42], and $H \to b\bar{b}$ branching ratios used for each Higgs mass[21].

We apply the same event selection to the signal Monte Carlo samples as described in the

m_H (GeV/c^2)	SAM Name	σ_{ZH} (fb)	$\mathscr{B}(H \to b\bar{b})$
100	ahgt7a	168.5	0.8121
105	ahgt7b	144.6	0.7957
110	ahgt7c	124.6	0.7702
115	ahgt7d	107.8	0.7322
120	ahgt7e	93.5	0.6789
125	ahgt7f	81.4	0.6097
130	ahgt7g	71.1	0.5271
135	ahgt7h	62.2	0.4362
140	ahgt7i	54.7	0.3436
145	ahgt7j	48.1	0.2556
150	ahgt7k	42.4	0.1757

Table 9.1: The Monte Carlo samples used for the ZH signal processes, along with the cross sections and $H \to b\bar{b}$ branching fractions used for normalization. Values calculated by the programs in Ref.[42] and [21].

preceding chapters, with the application of event weights as in Equation 8.2. This leads to an estimate of the number of signal events expected to be present in the data. Table 9.2 shows the expected number of $ZH \to \mu^+\mu^- b\bar{b}$ events expected in both the pretag region and each of the three b-tag categories comprising the signal region.

9.2 Signal Event Discriminants

Just performing a counting experiment to observe the presence of a ZH signal is not feasible, as there are only a few signal expected and greater than 10 000 background events which contribute before tagging. Even with the requirement of b-tagging, the background contribution overwhelms the expected signal rate. To increase the sensitivity of the analysis, we can use correlations between kinematic quantities of the signal events, which can distinguish them from background events. This section details these techniques.

One of the best examples of this feature is with the invariant mass of the two leading jets in the event, m_{jj}. For Higgs signal events, the two leading jets are from the b-quarks produced in the Higgs decay. Therefore, the invariant mass of these two jets should be close to the Higgs boson mass. However, for background events, there is no expectation that the invariant mass should be any specific value. Therefore, events which have m_{jj} close to the Higgs mass have a

Expected $ZH \to \mu^+\mu^- b\bar{b}$ Events in 9.2 fb^{-1}				
m_H (GeV/c^2)	Pretag	SVT	SVL+JP	SVT+SVT
100	12.4	4.16	1.66	1.37
105	10.9	3.61	1.45	1.23
110	9.40	3.12	1.27	1.08
115	7.98	2.67	1.09	0.93
120	6.53	2.18	0.90	0.76
125	5.29	1.76	0.73	0.63
130	4.08	1.36	0.57	0.48
135	3.02	1.01	0.42	0.36
140	2.13	0.71	0.30	0.25
145	1.42	0.47	0.20	0.17
150	0.87	0.29	0.13	0.10

Table 9.2: Expected $ZH \to \mu\mu b\bar{b}$ signal events for this analysis, for the three b-tag categories as well as the pre-tag region.

higher likelihood to be signal events. To use features like this from many kinematic variables, we use an artificial neural network with several kinematic variables as inputs.

The event discriminants are trained using equal numbers of all the background processes that contribute to the analysis, weighted by their trigger efficiencies so as not to include events which do not contribute to the final selection.

The event discriminants use the following set of kinematic variables as inputs:

- $\not\!\!E_T$ – The missing energy in the event, corrected for jets and muons.
- N_{jets} – The number of tight jets in the event.
- $\vec{\not\!\!E}_T \cdot (\vec{j}_1 + \vec{j}_2)$ – The projection of the $\vec{\not\!\!E}_T$ vector on to the vector sum of the two leading jets in the event.
- $p_T(Z)$ – The transverse momentum of the reconstructed Z boson.
- $p_T(H)$ – The transverse momentum of the dijet system.
- $\sum E_T$ – The scalar sum of the E_T of the Z and all tight jets in the event.
- m_{jj} – The dijet mass (or reconstructed Higgs boson mass).

- $\Delta R(j_1, j_2)$ – The separation between the lead and secondary jet.
- $\Delta R(Z, H)$ – The separation between the reconstructed Z and H bosons.
- $\Delta R(\mu_1, \mu_2)$ – The separation between the two muons in the event.
- Sphericity – An angular measure which takes into account the distribution of all reconstructed objects and their positions in the detector.
- m_{Tot} – The total mass of all reconstructed objects (2 leptons and all jets).
- m_{Zjj} – The total reconstructed mass of the Z boson, lead, and secondary jets in the event.
- $p_T(\mu_i)$ – The transverse momentum of each of the two muons in the event.

Figures 8.9 and 8.10 show the kinematic distributions of these variables, from which one can see the separation between background processes and the signal process.

The discriminants are trained to output a score ranging from 0.0 to 1.0, with a low score signifying background-like events, and high scores signifying signal-like events. By using these discriminant functions, we can create areas of higher signal purity, thereby increasing the overall sensitivity of the analysis. Figure 9.2 shows an example discriminant output for signal and background Monte Carlo events.

A separate discriminant function is trained for each Higgs mass, to optimize the behavior across the entire range of Higgs mass hypotheses in this analysis. Figure 9.1 shows why it is important to optimize these discriminants for each individual Higgs mass hypothesis used. Events with vastly different m_H show up as background-like events in discriminants optimized for other Higgs mass values. If the discriminants were trained using kinematic information from all the Higgs mass samples combined, some Higgs events may be pushed toward the background events rather than being isolated as signal-like.

9.2.1 Ensemble Method

The discriminants used for this analysis use the ensemble method, which was first used for the $t\bar{t}H$ analysis at CDF[46]. In this method, the discriminant consists of several (in this case,

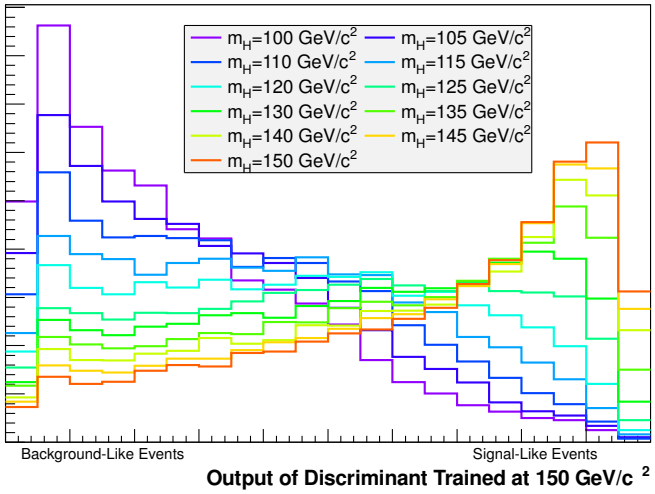


Figure 9.1: Output of various Higgs mass Monte Carlo samples for the final event discriminant trained specifically for $m_H = 150$ GeV/c^2. Some Higgs mass points look very signal-like, while others appear background-like. This shows the importance of training individual discriminants for each Higgs mass hypothesis.

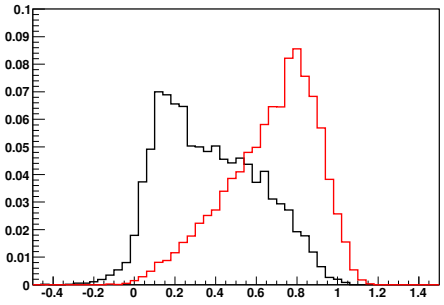


Figure 9.2: An example output histogram for a given neural network. Signal events are shown in red, background events are shown in black. This discriminant increases signal purity by shifting kinematically distinct events away from the majority of the background contribution.

30) individual neural networks. For each constituent network, a signal output histogram and background output histogram, normalized to unit area, is created. To obtain the output score for a single event for one of the constituent networks, the network is evaluated for that event to determine in which bin of the output histograms that event falls. The output score for that network, ζ, is then the ratio of the signal events, S_i, in that particular bin to the total number of events in that bin, $S_i + B_i$: $\zeta \equiv S_i/(S_i + B_i)$. The score ζ is constrained to the range $[0, 1]$. To obtain the final output score, $\overline{\zeta}$, for the ensemble of individual networks, this process is repeated for all N constituent networks, and the ζ scores are averaged, as in Equation 9.1.

$$\overline{\zeta} = \frac{1}{N}\sum_{i=1}^{N}\zeta_i = \frac{1}{N}\sum_{i=1}^{N}\frac{S_i}{S_i + B_i} \tag{9.1}$$

With this output score for each data and simulated event in the final signal region of the analysis, one could proceed to produce results using these distributions. However, for this analysis we have introduced a new technique to further increase the sensitivity beyond that obtained by using solely these final output discriminants.

9.3 Multi-Layer Discriminant Method

The multi-layer disciminant method uses 'expert' discriminant functions to characterize events further than determining if they are signal-like or background-like from the signal-optimized discriminants described in the previous section. These 'expert' discriminants are trained to identify single types of processes out of the many that contribute to a given background model. Based on the outputs of these functions, events can be characterized into different categories of background. By applying these functions carefully, one can both improve the overall sensitivity of the analysis by isolating regions of high signal and low background contribution, as well as reduce systematic uncertainties by combining the majority of a single type of background into a single category.

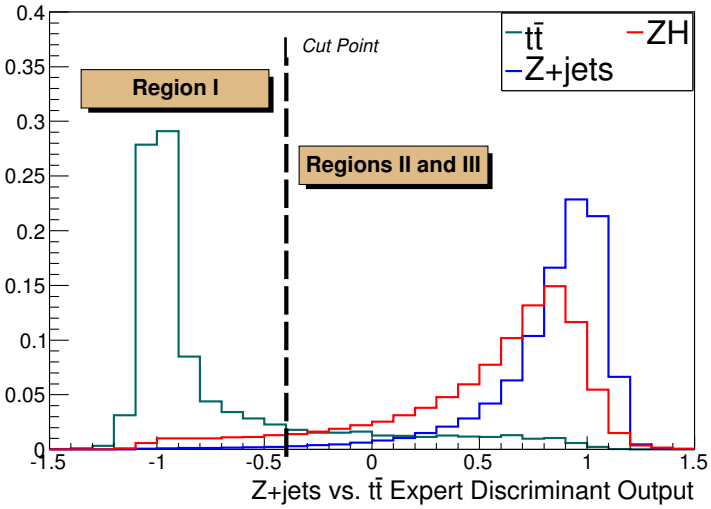


Figure 9.3: The expert discriminant used to separate $t\bar{t}$ and Z+jets events, the first step in the multi-layer discriminant method. We use a cut of $\alpha < -0.4$ to define Region I, a $t\bar{t}$-enriched region.

9.3.1 Top-Pair/Z+Jets Expert Discriminant

The first such expert discriminant function used in this analysis attempts to distinguish $t\bar{t}$ events from the main source of background, Z+jets production. As top-pair events are very kinematically distinct, having no reconstructed Z boson and larger $\not{E}_T$, this class of background events is easiest to distinguish. To create this function, we train a neural network to separate the two processes using the same set of kinematic variables listed above. Top-pair events are targeted to low output scores, while Z+jets events are targeted to high output scores.

The output of this expert discriminant, for $t\bar{t}$ and Z+jets events, as well as for signal events, is shown in Figure 9.3. Due to the differing kinematics, the two classes of backgrounds are well-separated by the expert discriminant. We place a cut on the output of this function, $\alpha < -0.4$ to define the first 'region' of the final output distributions. This region, Region I, will be enriched in $t\bar{t}$ events. As most of the other backgrounds have a reconstructed Z boson and additional jets, they will have output scores α generally above this cut. Having a region enriched in $t\bar{t}$ events can also help constrain systematic uncertainties, as discussed in Chapter 10.

Events that have high output scores, $\alpha > 0.4$, are kinematically similar to Z+jets events, and can fall into Region II or Region III. These events need to be further evaluated to classify them accordingly.

9.3.2 Jet-Flavor Separator

Events that are distinct from the $t\bar{t}$ region defined by the output of the expert discriminant can be further classified. The remaining events consist of Z+jets background and ZH signal events. After the requirement of b-tagging to form the final signal regions of the analysis, the main categories of background are $Z + c\bar{c}$ and $Z + b\bar{b}$ events. As we expect the ZH signal to produce two b-quark jets, it is advantageous to try to classify events based on the flavor content of the quark that originated the jet in the detector.

These functions, known as 'flavor separators', exist and have proven successful at discriminating jets originating from light quarks (u, d, s, c) and gluons from b-quark jets. Sometimes these functions are even used to b-tag jets. We use a flavor separator developed by the Karlsruhe Institute of Technology[39] (KNN) to distinguish between c-quark and b-quark jets.

The KNN function is a neural network and uses several kinematic quantities of the jet to produce a continuous output score, β_i, which defines the 'b-like' or 'c-like' flavor content of a jet. Figure 9.4 shows the output of this flavor separator function for $Z + bb$ and $Z + cc$ events. This flavor separator uses kinematic variables of b-tagged jets, such as the displacement distance of the secondary vertex, L_{xy}, the number of tracks at the secondary vertex, and the transverse momentum and invariant mass of those tracks. A complete listing of the input kinematic variables is found in [39].

We use this flavor separator function to divide the remaining Z+jets events into a region containing events with either light flavor or charm jets (Region II), and a region containing events with bottom jets (Region III). As the flavor separator returns a score for each jet in the event, we evaluate the function using the first two jets in the event to obtain two output scores, β_1 and β_2. To define Regions II and III, we use $\beta \equiv \max(\beta_1, \beta_2)$ as a metric. Events with $\beta > 0.4$ are considered to have a high purity of b-like jets and are classified in Region III, the heavy flavor region. Events failing this requirement are in Region II, the light flavor region.

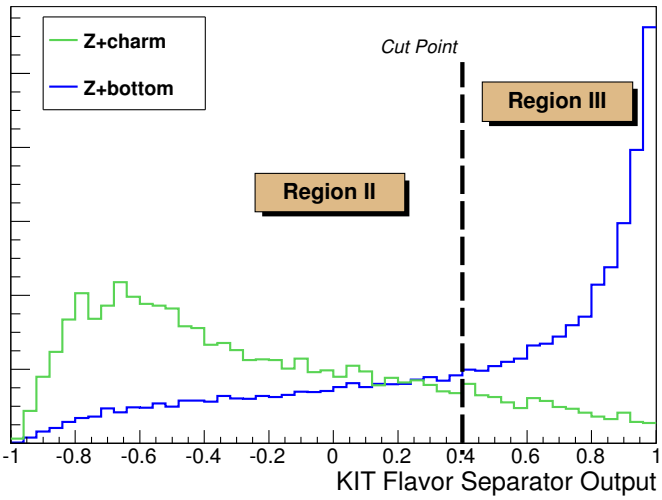


Figure 9.4: The flavor separator used in this analysis to distinguish b-quark jets from c- and light-quark jets.We use a cut of $\beta > 0.4$ to define a region enriched in b-jets.

Changing the cut values of α or β, apart from moving them to the extremes of the distributions, does not have a large impact on the overall sensitivity of the analysis.

Putting these two functions together to define the three regions, we obtain the final discriminating distributions. These distributions range from 0 to 3, with Region I from 0 to 1, Region II from 1 to 2, and Region III from 2 to 3. Table 9.3 summarizes the requirements on each of the multivariate functions used to divide the regions, which are mutually exclusive to one another. To be clear, the expert discriminant and flavor separator are only used to categorize events into their individual regions – each event is still evaluated with the final event discriminant described in Section 9.2. The output score from the final output discriminant, ζ, is modified to obtain the value plotted on the final distributions, ζ', according to Equation 9.2.

$$\zeta' = \begin{cases} \zeta & \text{if in Region I} \\ \zeta + 1 & \text{if in Region II} \\ \zeta + 2 & \text{if in Region III} \end{cases} \tag{9.2}$$

Criteria for Event Classification

Region	x-axis Range	$t\bar{t}$–Z+jets Discriminant	Flavor Separator
Region I ($t\bar{t}$-enriched)	$[0,1]$	$\alpha < -0.4$	NONE
Region II (Light Flavor)	$[1,2]$	$\alpha > -0.4$	$\beta < 0.4$
Region III (Heavy Flavor)	$[2,3]$	$\alpha > -0.4$	$\beta > 0.4$

Table 9.3: Criteria for defining the three regions used in the final discriminating distributions of the analysis. The same event discriminant is used in each region, with the output scores being changed according to Equation 9.2 to define the three regions based on the cuts on α and β shown here.

Figure 9.5 shows the output distributions of the main classes of background, and signal, after they have been classified according to the multi-layer method described above. Region I consists mostly of $t\bar{t}$ events. Region II contains the majority of events with a Z boson and additional light flavor jets, including a large fraction of the $Z + c\bar{c}$ background. The right-most region, Region III, contains a large fraction of the irreducible $Z + b\bar{b}$ background, as well as a large fraction of the ZH signal events.

We use this process to produce the final distributions that are used to test for the presence of a Higgs signal. While there are different final event discriminants trained specifically for each Higgs mass hypothesis, the same expert discriminant and flavor separator are used for classification in each case. There are no Higgs-mass-dependent quantities involved in the classification process, so it is not necessary to optimize these functions for various Higgs masses.

We now proceed to form the final distributions with the full background model, including the application of b-tagging. These distributions are used to test for the presence of a signal. Figure 9.6 shows one of the final event discriminants, for $m_H = 120$ GeV/c^2. The discriminants for each Higgs mass can be found in Appendix B.

We do not observe any significant excess in these distributions that would signify the presence of a signal. We will proceed to set upper limits on the ZH production cross section times $H \to b\bar{b}$ branching ratio. Before carrying out this procedure, we must adequately account for systematic uncertainties that affect the final result. This is the topic of the following Chapter.

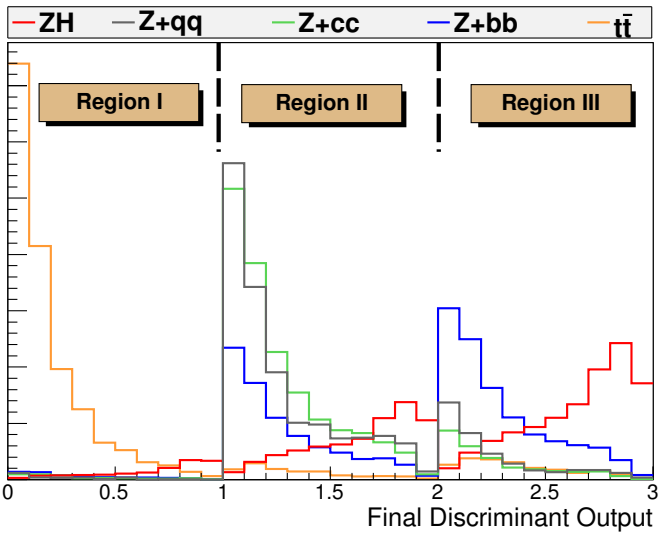


Figure 9.5: Behavior of the classification procedure for the main classes of background, and the signal events. Curves in this figure are drawn normalized to unit area, to compare the relative proportions of each background type and signal.

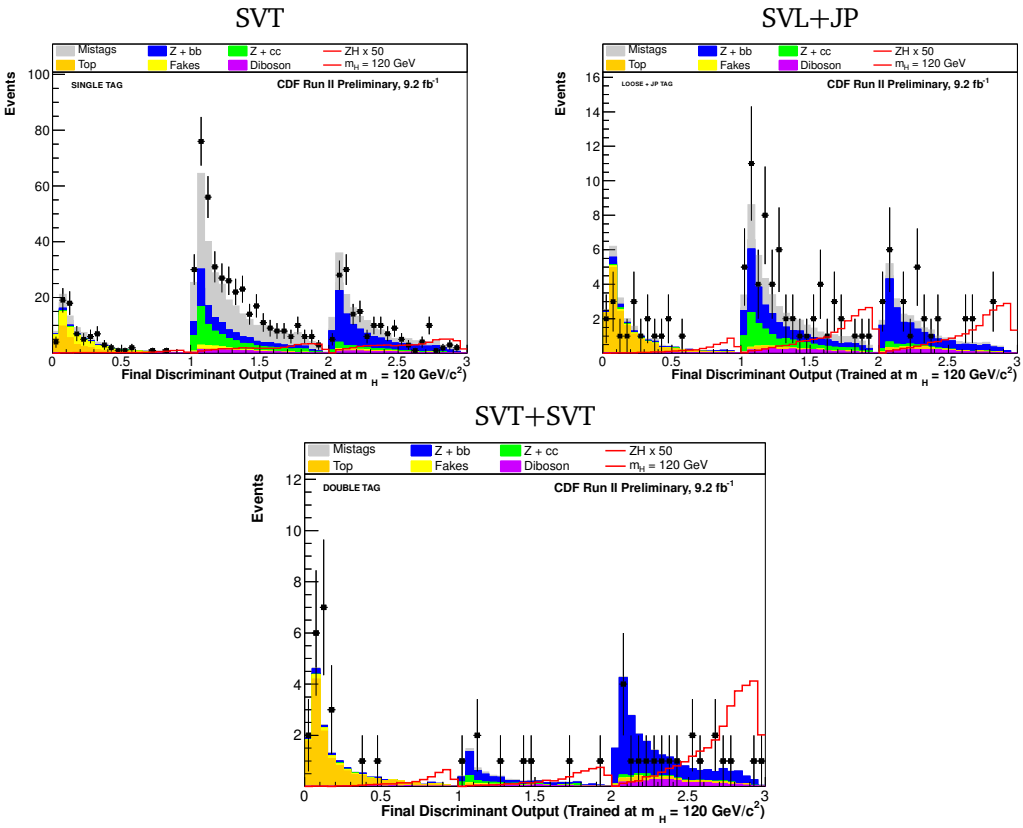


Figure 9.6: Example discriminant output, for the three b-tagging categories, for $m_H = 120$ GeV/c^2. Note that the ZH signal is scaled by a factor of 50 on the plots.

Chapter 10
SYSTEMATIC UNCERTAINTIES

There are many sources of uncertainty which affect the final results obtained for this analysis. Detector effects, theoretical inputs, uncertainties on scale factors, and discriminant effects all can impact the final analysis. This chapter details the application of these systematic uncertainties.

There are two general categories of systematic uncertainties used in this analysis: rate uncertainties, which alter the overall normalizations of event yields and do not affect the kinematics of events, and shape uncertanties, which can change event kinematics leading to different final event output discriminant scores. Some uncertainties have both a rate and shape component.

10.1 Rate Uncertainties

There are several sources of systematic uncertainty which impact the event weights, but do not change the kinematics of individual events. This class of systematics is known as rate systematics.

Luminosity

The luminosity is measured by the Cherenkov Luminosity Counter, as described in Chapter 6. This measurement has an inherent uncertainty, and affects the normalization of Monte Carlo samples in the weighting procedure applied to Monte Carlo events. There are actually two sources of uncertainty in this category.

The first is due to the measurement of the total $p\bar{p}$ inelastic cross section. This quantity has been measured by two Tevatron experiments – CDF[3] and the E811 experiment[11], with an associated error. Since the cross section is an input to determine the instantaneous, and therefore integrated, luminosity, the error on $\sigma_{p\bar{p}}$ propagates through to the luminosity measurement. This contribution to the luminosity uncertainty is 3.8%, and only applies to Monte Carlo samples that are weighted by the integrated luminosity corresponding to the data selection.

The second source of uncertainty on the luminosity measurement comes from the CLC itself. The CLC operates by measuring $p\bar{p}$ interactions. There is an efficiency for this process to be adequately detected by the CLC, leading to an uncertainty in the measured number of interactions per bunch crossing. This is another input to determine the luminosity, therefore the detector acceptance efficiency uncertainty is propagated through to the integrated luminosity calculation. The detector uncertainty is 4.4%, and this uncertainty is applied to all Monte Carlo samples, as is the case with the inelastic cross section uncertainty.

In summary, the two sources of uncertainty on the integrated luminosity calculation can be combined in quadrature to produce a total uncertainty of 5.8% on the measurement of the integrated luminosity used in the event weights of Monte Carlo samples.

Trigger Simulation

As described in Chapter 8, we apply trigger efficiencies to Monte Carlo events, based on a neural network-based regression technique. The determination of these efficiencies is an uncertain process, and a systematic error needs to be applied accordingly to all Monte Carlo events. We measure this uncertainty by using individual networks in the ensemble used to evaluate individual events. This discrepancy between constituent networks is used as an uncertainty on the trigger efficiency. The systematic uncertainty for the trigger simulation is 5%.

Process Cross Sections

In addition to the integrated luminosity, the cross sections of various processes are used in the normalization of Monte Carlo events. These process cross sections have associated uncertain-

ties, from both theoretical and experimental origins.

For the $t\bar{t}$ Monte Carlo sample, we apply a 10% systematic uncertainty. This uncertainty includes contributions from the theoretical prediction of the cross section. Additionally, because the Monte Carlo sample we use has a value of the top quark mass $m_t = 172.5$ GeV/c^2, this uncertainty also includes a contribution based on the discrepancy between the simulated and observed top quark masses.

For the diboson processes, WW, WZ, and ZZ, we apply a 6% systematic uncertainty. For Monte Carlo events with a Z boson and heavy flavor jets (c- and b-quarks), we apply a 40% systematic uncertainty. This uncertainty originates from the theoretical calculation of the cross section, as well as from the ALPGEN K-factor that is applied to these samples. This systematic uncertainty is the dominant uncertainty for this analysis.

ZH signal Monte Carlo samples receive a 5% systematic uncertainty to cover the error on the theoretical prediction of the cross sections. The contribution of additional intitial state or final state jets, due to gluon splitting or gluon radiation, can also affect the acceptance and normalization of signal events. This effect is accounted for with an additional rate systematic. The uncertainty varies based on the tag category – it is 1% for SVT, 2% for SVL+JP, and 5% for SVT+SVT. This uncertainty is comparable to that used for the signal cross section.

Lepton Identification

Errors on scale factors applied to Monte Carlo events propagate through to uncertainties on the total background estimation. Scale factors, such as that for the muon NN efficiency, affect the overall background normalization. Uncertainties on these scale factors must be accounted for. For the muon NN, we apply a 1% systematic uncertainty to account for the error on the scale factor measurement.

Because the muon NN uses calorimeter energy measurements as an input quantity, we apply a 1.5% systematic error to account for the uncertainty in the energy reconstruction due to calorimeter resolution.

Finally, there is a systematic uncertainty on the estimate of the mis-identified Z boson contribution to the background. This uncertainty is measured using data events that fail the muon

NN selection. While we assume that the contribution of same-charge and opposite-charge muon pairs is the same, we take the difference of these two contributions in the events that fail the muon NN as an uncertainty on the normalization of this background. This results in a 5% systematic uncertainty on the normalization of the mis-identified Z boson contribution.

b-tag Scale Factors

As described in Chapter 8, we apply scale factors to b-tagged jets to account for differing tag efficiencies in the Monte Carlo and the data. These measured scale factors have associated uncertainties which need to be propagated through to the background estimation in the final signal regions of the analysis. Table 8.3 lists the scale factors, along with their uncertainties, for each of the tagging algorithms used. We combine these fractional uncertainties to obtain errors for each tag category. For the single tight category (SVT), the error is 5.2%. For the double tag (SVT+SVT) category, we have twice that error, 10.4%, due to the correlated errors on the two jets with SVT tags. For the loose double tag category (SVL+JP), the two tag types are uncorrelated, so we add the two scale factor uncertainties in quadrature to obtain a systematic error of 8.7% on the SVL+JP category.

10.2 Shape Systematics

Apart from uncertainties affecting only the normalizations of various background contributions, there exists another class of systematic uncertainties which affect both the normalization and kinematic distributions of events. These uncertainties are known as shape systematics, and consist of two contributions for this analysis: jet energy scale and mistag uncertainties.

Jet Energy Scale

The jet energy scale systematic uncertainty is one of the largest shape uncertainties in use at CDF. At CDF, corrections are applied[17] to jets to account for various detector and physics effects, such as non-linearity of the calorimeter response, the presence of energy outside the cone of the reconstructed jet, and additional energy from the non-hard-scatter portion of the

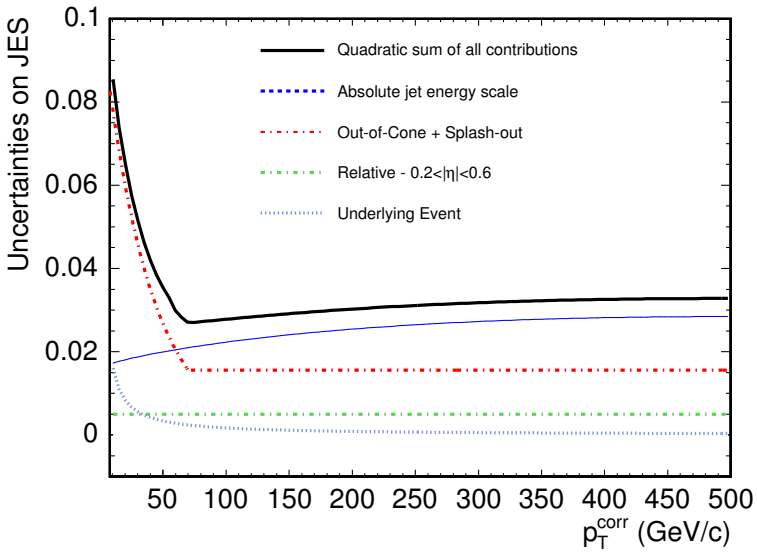


Figure 10.1: Jet Energy Scale systematic uncertainties, as a function of jet p_T (including absolute corrections). The various mechanisms leading to systematic errors are shown individually as well. Figure from [17].

event. These corrections have an uncertainty which shift the jet energies up and down, and are known as jet energy scale (JES) systematics. Figure 10.1 shows the total JES uncertainty as a function of jet p_T, and the contribution from each individual source of uncertainty.

Shifting the jet energies up and down within their JES uncertainties affects the Lorentz vectors of reconstructed physics objects in events. For example, quantities like m_{jj} and the dijet system transverse momentum will be affected by the change in jet energies. Because several of these affected quantities are inputs to the signal discriminants, the output scores will change with changes in jet energies within the JES systematics. This is the origin of the term 'shape systematic' – the shape of the discriminating distributions change with the JES systematic application.

We increase the jet energies by $\pm 1\sigma$ and produce templates for each of the background Monte Carlo samples. These templates are used in the limit setting process described in Chapter 11. Jets that did not satisfy the event selection using the nominal corrections – if they have low E_T, for example – can satisfy the event selection under a JES $+1\sigma$ shift. The inverse also is

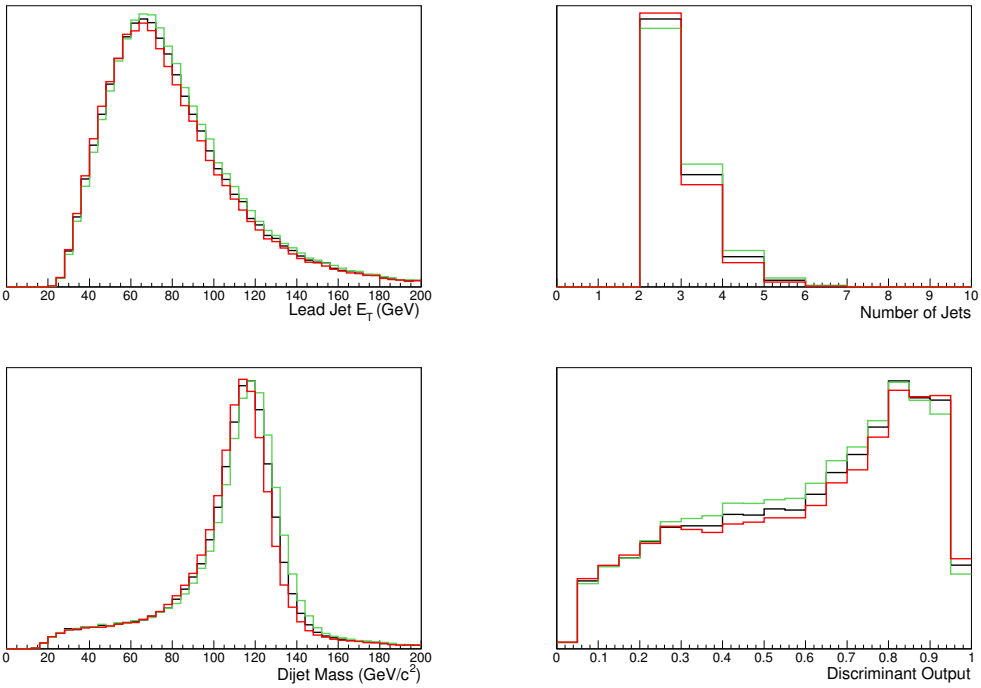


Figure 10.2: Typical effect of jet energy systematic shifts on kinematic distributions, as well as signal discriminant output, shown for a ZH Monte Carlo sample. The black distribution shows un-shifted jets, while the green and red curves show the distributions after shifting jets up or down by their 1σ energy uncertainties, respectively. The curves have arbitrary units, but are not normalized relative to one another.

possible – events could propagate out of the event selection if their E_T is shifted below the event selection threshold. Therefore, the JES systematic not only affects the shape of the discriminating distribution, but also affects the overall normalization of Monte Carlo backgrounds. The rate portion of the JES uncertainty is approximately a 10% systematic effect, and is asymmetric due to the kinematic variances described above.

Figure 10.2 shows the effect of these JES shifts on the kinematics a ZH signal Monte Carlo sample ($m_H = 115$ GeV/c^2). The number of jets, leading jet E_T, and dijet mass m_{jj} are shown, along with the final discriminant output for this Higgs mass. As the jet energies are shifted up and down, both shapes and overall rates change due to differing jet kinematics and events propagating in and out of the event selection.

Mistagged Jets

The final source of shape uncertainty is that due to the uncertainty of the measurement of the mistagged background contribution. The mistag matrix (described in Chapter 8) produces probabilities of events to be tagged, along with uncertainties on those probabilities. Using these uncertainties, we form a $+1\sigma$ mistag background shape and a -1σ background shape. These templates are used in the limit-setting procedure, similar to the process used for the JES systematic results.

Table 10.1 summarizes all of the systematic uncertainties applied to the various background contributions of the analysis. As these systematic uncertainties can change with the b-tag category, there is a separate table for each of the three b-tagging categories used in the analysis.

With all applicable systematics in place, we can move on to producing the final results for this analysis.

CDF $ZH \to \mu\mu b\bar{b}$ Analysis								CDF Run II Preliminary, 9.2 fb^{-1}	
Single Tag (ST) Category									
Contribution	Fakes	$t\bar{t}$	WW	WZ	ZZ	$Z+b\bar{b}$	$Z+c\bar{c}$	Mistags	ZH
Luminosity ($\sigma_{\text{inel}}(p\bar{p})$)		3.8	3.8	3.8	3.8	3.8	3.8		3.8
Luminosity Monitor		4.4	4.4	4.4	4.4	4.4	4.4		4.4
Lepton ID		1	1	1	1	1	1		1
Lepton Energy Scale		1.5	1.5	1.5	1.5	1.5	1.5		1.5
Fake Leptons	5								
Mistag Rate								$^{+13.7}_{-13.7}$	
Jet Energy Scale (shape dep.)		$^{+1.4}_{-1.6}$	$^{+15.5}_{-0.1}$	$^{+3.3}_{-4.3}$	$^{+3.6}_{-5.2}$	$^{+9.0}_{-8.6}$	$^{+8.3}_{-8.2}$		$^{+2.5}_{-2.8}$
b-tag Rate		5.2	5.2	5.2	5.2	5.2	5.2		5.2
$t\bar{t}$ Cross Section		10							
Diboson Cross Section			6	6	6				
Z+HF Cross Section						40	40		
ZH Cross Section									5
ISR/FSR									1
NN Trigger Model		5	5	5	5	5	5		5

CDF $ZH \to \mu\mu b\bar{b}$ Analysis								CDF Run II Preliminary, 9.2 fb^{-1}	
Loose + JetProb (LJP) Category									
Contribution	Fakes	$t\bar{t}$	WW	WZ	ZZ	$Z+b\bar{b}$	$Z+c\bar{c}$	Mistags	ZH
Luminosity ($\sigma_{\text{inel}}(p\bar{p})$)		3.8	3.8	3.8	3.8	3.8	3.8		3.8
Luminosity Monitor		4.4	4.4	4.4	4.4	4.4	4.4		4.4
Lepton ID		1	1	1	1	1	1		1
Lepton Energy Scale		1.5	1. 5	1.5	1.5	1.5	1.5		1.5
Fake Leptons	5								
Mistag Rate								$^{+25.5}_{-22.7}$	
Jet Energy Scale (shape dep.)		$^{+0.5}_{-0.8}$		$^{+3.8}_{-5.8}$	$^{+2.7}_{-3.3}$	$^{+6.6}_{-7.6}$	$^{+4.4}_{-6.7}$		$^{+1.2}_{-1.4}$
b-tag Rate		8.7	8.7	8.7	8.7	8.7	8.7		8.7
$t\bar{t}$ Cross Section		10							
Diboson Cross Section			6	6	6				
Z+HF Cross Section						40	40		
ZH Cross Section									5
ISR/FSR									2
NN Trigger Model		5	5	5	5	5	5		5

CDF $ZH \to \mu\mu b\bar{b}$ Analysis								CDF Run II Preliminary, 9.2 fb^{-1}	
Double Tag (DT) Category									
Contribution	Fakes	$t\bar{t}$	WZ	ZZ	WW	$Z+b\bar{b}$	$Z+c\bar{c}$	Mistags	ZH
Luminosity ($\sigma_{\text{inel}}(p\bar{p})$)		3.8	3.8	3.8	3.8	3.8	3.8		3.8
Luminosity Monitor		4.4	4.4	4.4	4.4	4.4	4.4		4.4
Lepton ID		1	1	1	1	1	1		1
Lepton Energy Scale		1.5	1.5	1.5	1.5	1.5	1.5		1.5
Fake Leptons	5								
Mistag Rate								$^{+28.6}_{-25.0}$	
Jet Energy Scale (shape dep.)		$^{+0.2}_{-0.6}$		$^{+20.5}_{-0.1}$		$^{+4.1}_{-4.5}$	$^{+2.5}_{-2.9}$		$^{+0.6}_{-1.1}$
b-tag Rate		10.2	10.2	10.2	10.2	10.2	10.2		10.2
$t\bar{t}$ Cross Section		10							
Diboson Cross Section			6	6	6				
Z+HF Cross Section						40	40		
ZH Cross Section									5
ISR/FSR									5
NN Trigger Model		5	5	5	5	5	5		5

Table 10.1: Systematic uncertainties on the analysis. Systematic uncertainties for $ZH \to \mu\mu b\bar{b}$ shown in this table are obtained for $m_H = 120$ GeV/c^2. Uncertainties are relative, in percent, and are symmetric unless otherwise indicated.

Chapter 11
Analysis Results

After producing the final discriminating distributions for all background contributions, signal hypotheses, and data events, we do not observe a significant excess which would signify the presence of a signal. We proceed to set upper limits on the ZH production cross section times $H \to b\bar{b}$ branching ratio using each of the three b-tag categories previously described. This chapter details the process used to produce these limits, along with the results of this analysis in combination with other Higgs search channels at both CDF and D0.

11.1 Limit Setting Procedure

The aim of the limit setting procedure is to evaluate the consistency of the data with the background-only hypothesis and the combined background and signal hypothesis. We aim to set an upper limit at the 95% confidence level, meaning we find the lowest cross section for the signal normalization that is inconsistent with the data distribution at the 2σ level. All cross sections above this specific one are excluded, giving an upper limit at the 95% confidence level on the Higgs production cross section for this process.

To evaluate this upper limit, we use the standard CDF software package, *MCLIMIT*[33]. This software uses the final event discriminant distributions, after the application of the multi-layer discriminant method, for each background contribution, the signal expectation, and the observed data. Using these distributions, as well as systematic errors as inputs to the program, a binned-likelihood Bayesian method, closely following the method described in [30], is used to produce the upper limit. This method is described below.

In a simple counting experiment to test for the presence of an expected signal $s' = s \cdot \epsilon$ (ϵ represents the selection efficiency due to event selection and reconstruction), consisting of n data counts and b background counts, the probability to obtain the observed data is that of Equation 11.1, with the expectation following a Poisson distribution.

$$P(n \mid s, b) = \frac{e^{-(s\epsilon+b)}(s\epsilon + b)^n}{n!} \tag{11.1}$$

In our analysis, we have one of these counting experiments for each bin of each histogram used for the final discriminating distributions. To obtain the probability to obtain the observed distribution in each of the N bins used in the analysis, one must take the product of the probabilities over all N bins, as in Equation 11.2.

$$P(\text{data} \mid \text{signal, background}) = \prod_{i=1}^{N} \frac{e^{-(s\epsilon_i+b_i)}(s\epsilon_i + b_i)^{n_i}}{n_i!} \tag{11.2}$$

In physics analysis, the quantities ϵ and b, the signal selection efficiency and background estimation have uncertainties, both statistical and systematic, as described in Chapter 10. These uncertainties, known as nuisance parameters in the limit setting procedure, must be correlated between the different bins in the final discriminants. These nuisance parameters are assigned priors – for rate uncertainties these priors are Gaussian functions with σ equal to the given systematic uncertainty for that process. The prior distribution for the signal is chosen to be flat, $\pi(s) = 1$.

The next step is to sample from this systematic parameter space using the prior distributions for each of the individual uncertainty contributions and background templates. This involves a large number of samplings of different nuisance parameters, of order 10^3 or 10^4. After these M different sets of nuisance parameter choices are produced, one can determine the posterior distribution for the signal, s:

$$p(s) = \frac{1}{M \cdot \mathcal{N}} \sum_{j=1}^{M} \left[\prod_{i=1}^{N} \frac{e^{-(s\epsilon_{j,i}+b_{j,i})}(s\epsilon_{j,i} + b_{j,i})^{n_i}}{n_i!} \right] \tag{11.3}$$

The quantity $\mathcal{N}$ is a normalization constant. The sum over the M samples of the nuisance parameter space ensures that the systematic uncertainties, and correlations between them, are

implemented into the limit calculation.

To extract an upper limit on the value of s, we must integrate over the posterior distribution of Equation 11.3 until we reach just a 5% probability for a given signal s to be consistent with the observed data in all correlated bins. This value of s represents the upper limit, at the 95% confidence level. These integrals have the form shown in Equation 11.4, and for a generic confidence level, one can obtain the upper limit value s_{upper} by solving the equation $I(s_{\text{upper}}) = (1-\beta)\cdot I(0)$. The variable β represents the desired confidence level. For this analysis, we use $\beta = 0.95$ to obtain limits at the 95% confidence level.

$$I(s') = \int_{s'}^{\infty} \frac{1}{M} \sum_{j=1}^{M} \left[\prod_{i=1}^{N} \frac{e^{-(s\epsilon_{j,i}+b_{j,i})}(s\epsilon_{j,i}+b_{j,i})^{n_i}}{n_i!} \right] \mathrm{d}s \tag{11.4}$$

The process described above is used to set the observed upper limit on the Higgs production cross section times $H \to b\bar{b}$ branching ratio. To do this for the data distribution, this process is carried out just a single time. However, to avoid bias when designing an analysis, it is useful to have the capability to estimate the sensitivity of an analysis using only the background and signal estimates. When adding new techniques to analyses, it is useful to have a statement of how much a given technique has contributed to the overall sensitivity of the analysis. This process involves computing a set of 'expected' limits using a large number of pseudoexperiments.

For each pseudoexperiment, a pseudo-dataset is computed using a random sampling of the systematic parameter space for the background-only hypothesis (no signal events or data events are used to construct the pseudo-dataset). This pseudo-dataset is then treated the same way as the actual observed data, using the process described above, to find an upper limit on the signal s. This entire pseudoexperiment is repeated a large number of times (of order 10^3 or more) to obtain a distribution of upper limits. We take the median of this distribution as the 'expected' limit, an estimate of the expected sensitivity of the analysis using the observed data. We also have uncertainty bands on the expected limit, with 1σ and 2σ uncertainties represented by the values of the upper limits that contain 68% and 95% of the pseudoexperiments, respectively.

Since the pseudoexperiments are formed from a model hypothesis consisting of no signal, the presence of a signal will cause the observed limit, calculated from the observed data, to

deviate from the expected limits, calculated assuming no signal. It is typical to quote expected and observed limits on the Higgs production cross section as a ratio of the Standard Model predicted value. Only when the expected sensitivity (expected limit) reaches a factor of one times the Standard Model is it possible to make a definitive statement about the existence or exclusion of the Higgs boson. This statement will be determined by the behavior of the observed limit – it will deviate upward from the expected limit in the presence of a Higgs signal. If the observed limit has a value below the Standard Model predicted value, the Higgs boson would be excluded at that point.

11.2 Expected and Observed Limits

We use the process described above to produce observed and expected limits for the combination of the three b-tagging categories that make up the signal region for this analysis. We set expected upper limits on the ZH production cross section times $H \to b\bar{b}$ branching ratio, at the 95% confidence level, ranging from 3.8 times the Standard Model predicted value for $m_H = 100$ GeV/c^2 to 39.2 times the Standard Model value for $m_H = 150$ GeV/c^2. The observed limits deviate slightly from the expected limits, but are within 2σ over the entire mass range. The observed limits range from 4.7 times the Standard Model predicted value for $m_H = 100$ GeV/c^2 to 52.8 times the Standard Model value for $m_H = 150$ GeV/c^2. These expected and observed limits are shown in Table 11.1, and summarized graphically in Figure 11.1. Because the same set of data is analyzed for each Higgs mass hypothesis, the observed limits will be correlated across the various Higgs mass hypotheses tested for the presence of a signal.

11.3 Combinations of CDF Higgs Searches

Just as we combine the three b-tagging categories to produce a single result for this analysis, we can combine this analysis with all other Higgs search analyses at CDF to produce a combined Higgs search result, including various Higgs production mechanisms and decay modes.

It is important to be especially careful in this process to ensure that the distinct analyses do not overlap with one another. If the event selections are not orthogonal, data events could enter

m_H (GeV/c^2)	Observed Limit	Expected Limits				
		-2σ	-1σ	Median	+1σ	+2σ
100	**4.8**	1.9	2.6	**3.7**	5.4	7.5
105	**6.0**	2.1	2.9	**4.0**	5.9	8.4
110	**6.8**	2.3	3.2	**4.4**	6.3	9.6
115	**8.0**	2.5	3.5	**4.9**	7.1	9.6
120	**9.9**	3.0	4.1	**5.9**	8.4	12.3
125	**9.9**	3.7	4.9	**7.1**	10.0	14.9
130	**14.0**	4.6	6.4	**8.9**	12.8	19.5
135	**16.5**	6.2	8.4	**11.8**	17.9	25.4
140	**22.5**	8.5	11.6	**16.5**	24.2	33.6
145	**34.2**	12.2	17.1	**24.1**	34.7	50.5
150	**51.6**	18.8	26.4	**38.4**	54.6	80.7

Table 11.1: Observed and expected limits, for all tagging categories combined. The expected limits include the $\pm1\sigma$ and $\pm2\sigma$ values as well.

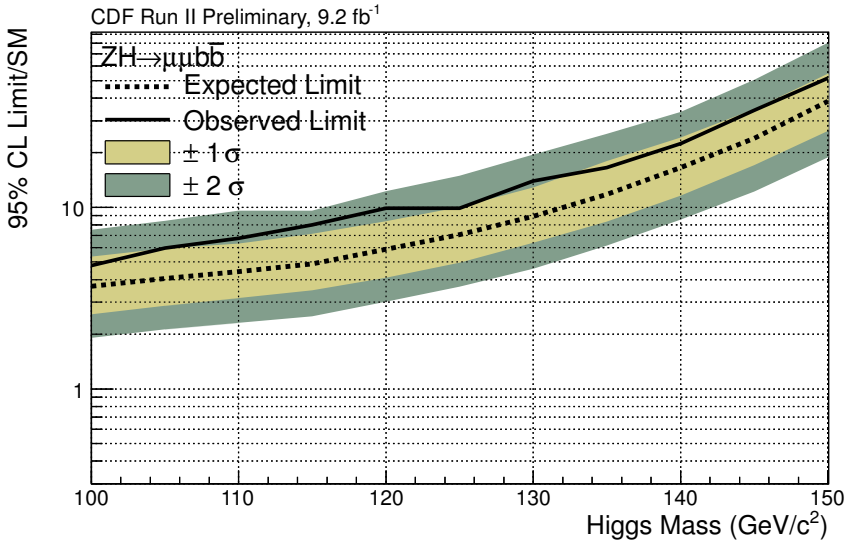


Figure 11.1: Expected limits for this analysis, shown in the dotted line. The solid line shows the observed limits. The yellow green bands represent the 1σ and 2σ distributions of expected limits obtained from pseudoexperiments.

into several analyses, and the expected Higgs signal would be overstated. Much work goes into comparing the different analyses entering the combination to ensure this orthogonality, to make a valid statement on the existence of the Higgs boson.

The upper limit on the Higgs boson production cross section (as a ratio to the Standard Model) is computed in the same way described above. There are simply more channels consisting of signal expectation histograms, background contribution histograms, and of course data distributions. Systematic uncertainties can be correlated across the different analyses, which helps constrain the overall uncertainty of the combined result. For example, an uncertainty such as that on the b-tag scale factor applies to *all* analyses using b-tagging, so background distributions containing b-tags have this systematic uncertainty correlated. If one channel is able to constrain the value of the uncertain parameter, this can then be applied to all other analyses with the same systematic uncertainty.

Figure 11.2 shows the combination of Higgs searches at CDF for the Summer 2011 conference season[20]. The main analyses which drive the sensitivity are the associated production channels WH and ZH, including the vector boson decay modes $W \to \ell\nu$, $Z \to \ell\ell$, and $Z \to \nu\nu$, with the predominant low-mass Higgs decay mode $H \to b\bar{b}$. There are other channels which add to the overall sensitivity, including $t\bar{t}H$, $H \to \gamma\gamma$, $H \to \tau\tau$, $H \to ZZ$, and $H \to WW$. Each of these individual analyses are shown as a distinct line on the plot, with the combination shown by the thick red line near the bottom of the plot. This combination does include the analysis presented in this thesis, but with a smaller amount of integrated luminosity compared to the result presented here. That result was combined with the complementary $ZH \to e^+e^-b\bar{b}$ analysis[35] that is designed in a similar fashion to this analysis.

For the combination of CDF analyses, we expect a limit on the Higgs production cross section of 1.49 times the Standard Model predicted value, and observe a limit of 1.55 times the Standard Model value, for $m_H = 115$ GeV/c^2. Furthermore, the CDF combination excludes the Higgs mass ranges $100 < m_H < 104.5$ GeV/c^2 and $156.5 < m_H < 173.7$ GeV/c^2, at the 95% confidence level.

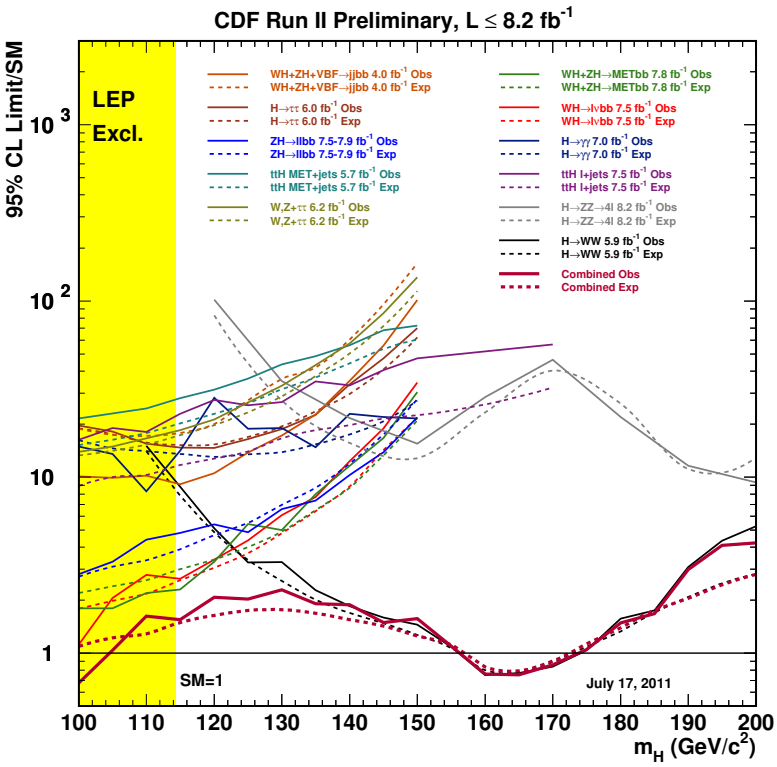


Figure 11.2: Expected (dashed) and observed (solid) limits for the combination of CDF Higgs searches. Individual analyses are shown with thin lines, while the combination of these analyses is shown with the dark red lines near the bottom of the plot.

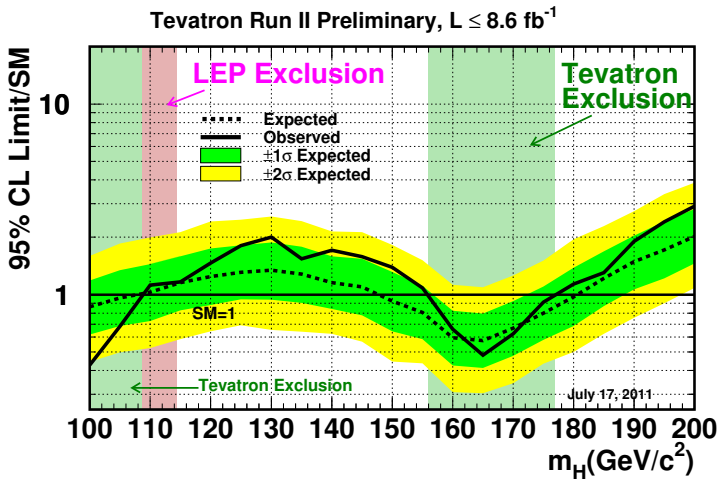


Figure 11.3: Expected (dashed) and observed (solid) limits for the combination of Tevatron Higgs searches. The 1σ and 2σ uncertainty bands are shown in green and yellow, respectively.

11.4 Combination of Tevatron Higgs Searches

To make a definitive statement about the Higgs boson, we need to push the senstivity of the analysis to a level where we expect the Higgs cross section upper limit to be equal to that of the Standard Model value. As can be seen in Figure 11.2, with only the CDF experiment we can reach a level of about 1.4 times the Standard Model. Given the amount of collisions that will be produced by the Tevatron, the only possible way to come close to Standard Model sensitivity is to combine analysis results from the two Tevatron CDF and D0.

Figure 11.3 shows the combined limit on the Higgs boson cross section, incorporating all of the CDF and D0 Higgs search analysis. Again, the analyses are required to be orthogonal and systematic uncertainties may be correlated between channels. The Tevatron experiments exclude the Higgs mass ranges $100 < m_H < 108$ GeV/c^2 and $158 < m_H < 177$ GeV/c^2. These results include the analysis of up to 8.6 fb^{-1} of integrated luminosity. With the full dataset of the Tevatron expected to be approximately 10 fb^{-1}, the combination of CDF and D0 analyses

with this full dataset could yield Standard Model sensitivity across the entire low-mass range, $100 < m_H < 150$ GeV/c^2. Much work will still be incorporated to the constituent analyses after the Tevatron ceases operations. New techniques to squeeze out any possible gains in sensitivity will be incorporated in order to produce the best result possible with the full Tevatron dataset.

Chapter 12
Conclusions and Outlook

We have produced an analysis searching for the Higgs boson in the $ZH \to \mu^+\mu^- b\bar{b}$ process. This analysis has implemented several new multivariate techniques that have improved the overall sensitivity of the analysis compared to previous iterations. The sensitivity improvement is larger than that expected from the increase in integrated luminosity alone, due to the new techniques.

We have implemented a multivariate selection for muon candidates, increasing the event acceptance by over 30%. This technique is in the process of being incorporated into other analyses using high-p_T muons, where similar acceptance gains are expected. We have also incorporated a multi-layer discriminant method, used to separate the main classes of background, $t\bar{t}$ and Z+jets events, using an expert discriminant in combination with a jet-flavor separator trained to separate b-like jets from c-like jets. This technique alone has improved the sensitivity of the analysis by 8% compared to using just a single final event discriminant. Other techniques new to this analysis include multivariate regression-derived trigger efficiency weights and ensemble-based final event discriminants.

This analysis sets an expected (observed) limit on the ZH production cross section times $H \to b\bar{b}$ branching ratio of 4.9 (8.0) times the Standard Model value, for a Higgs mass $m_H = 115$ GeV/c^2. This analysis result has been included in the combined CDF and Tevatron Higgs search results, shown at Summer 2011 conferences.

The result shown in Figure 11.3 is an extremely incredible result. It incorporates the hard work of hundreds of physicists, from those designing and carrying out the analyses, to those

responsible for operating the detectors and collecting the data, to the accelerator physicists responsible for producing collisions in the Tevatron. All of these people have contributed to producing this high-quality physics result.

As the Tevatron has shutdown as of September 2011, the global center of particle physics has shifted to the LHC. If the Tevatron is unable to make a definitive statement regarding the Higgs boson, the LHC certainly will. If the Higgs boson is discovered, this would be a tremendous event, confirming the Higgs mechanism describing electroweak symmetry breaking as described in this thesis. If the Higgs boson is excluded by these upcoming analyses, the result would be just as interesting. There must be some mechanism for electroweak symmetry breacking, as we observe vector boson masses. Discovering the source of this electroweak symmetry breaking will still be the one of the goals of the physics program at the LHC.

The next few years will be an exciting time in the field of high-energy physics. With the discovery or exclusion of the Higgs boson, there will be much more work to understand physics at the weak scale. Apart from this, there are many new physics models which predict additional particles, such as supersymmetry and dark matter. These new particles may be discovered at the LHC in the near future.

All of these efforts will continue to shape our understanding of nature at the most fundamental level. The pursuit of deeper and deeper knowledge will continue to motivate physicists as this understanding unfolds over time.

BIBLIOGRAPHY

[1] CMS Collaboration, *Search for standard model Higgs boson in pp collisions at $\sqrt{s} = 7$ TeV and integrated luminosity up to 1.7 fb^{-1}*, (2011), no. CMS PAS HIG-11-022.

[2] ATLAS Collaboration, *Update of the Combination of Higgs Boson Searches in pp Collisions at $\sqrt{s} = 7$ TeV with the ATLAS Experiment at the LHC*, (2011), no. ATLAS-CONF-2011-135.

[3] F. Abe et al., *Measurement of the antiproton-proton total cross section at $\sqrt{s} = 546$ and 1800 GeV*, Phys. Rev. D **50** (1994), 5550–5561.

[4] 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** (2006), no. 7, 072006.

[5] D. Acosta et al., *The performance of the CDF luminosity monitor*, Nucl. Instr. Meth. A **494** (2002), no. 1-3, 57 – 62.

[6] D. Acosta et al., CDF Collaboration, *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** (2005), no. 5, 052003.

[7] T. Affolder et al., *CDF Central Outer Tracker*, Nucl. Instr. Meth. A **526** (2004), no. 3, 249 – 299.

[8] S. Agostinelli et al., *Geant4 – a simulation toolkit*, Nucl. Instr. Meth. A **506** (2003), no. 3, 250 – 303.

[9] T. Akimoto et al., *The CDF Run IIb Silicon Detector: Design, preproduction, and performance*, Nucl. Instr. Meth. A **556** (2006), no. 2, 459 – 481.

[10] M. Albrow et al., *The CDF plug upgrade electromagnetic calorimeter: test beam results*, Nucl. Instr. Meth. A **480** (2002), no. 2-3, 524 – 546.

[11] C. Avila, W.F. Baker, R. DeSalvo, D.P. Eartly, C. Guss, H. Jostlein, M.R. Mondardini, J. Orear, S.M. Pruss, R. Rubinstein, S. Shukla, and F. Turkot, *A measurement of the proton-antiproton total cross section at $\sqrt{s} = 1.8$ TeV*, Phys. Lett. B **445** (1999), no. 3-4, 419 – 422.

[12] M. Awramik, M. Czakon, A. Freitas, and G. Weiglein, *Precise prediction for the W-boson mass in the standard model*, Phys. Rev. D **69** (2004), 053006.

[13] M. Baak et al., *Updated Status of the Global Electroweak Fit and Constraints on New Physics*, arXiv:1107.0975 (hep-ph) (2011).

[14] L. Balka et al., *The CDF central electromagnetic calorimeter*, Nucl. Instr. Meth. A **267** (1988), no. 2-3, 272 – 279.

[15] R. Barate et al., *Search for the Standard Model Higgs boson at LEP*, Phys. Lett. B **565** (2003), 61 – 75.

[16] S. Bertolucci et al., *The CDF central and endwall hadron calorimeter*, Nucl. Instr. Meth. A **267** (1988), no. 2-3, 301 – 314.

[17] A. Bhatti et al., *Determination of the jet energy scale at the collider detector at Fermilab*, Nucl. Instr.Meth. A **566** (2006), 375–412.

[18] R. Blair et al., CDF Collaboration, *The CDF II Detector Technical Design Report*, (1996).

[19] Gerald C. Blazey et al., *Run II jet physics*, arXiv:0005012 (hep-ex) (2000).

[20] CDF Collaboration, CDF Public Note 10609, *Combination of CDF's Searches for the Standard Model Higgs boson with up to 8.2 fb^{-1} of Data*.

[21] A. Djouadi, J. Kalinowski, and M. Spira, *HDECAY: A Program for Higgs boson decays in the standard model and its supersymmetric extension*, Comput.Phys.Commun. **108** (1998), 56–74.

[22] A. Einstein, *Über einen die erzeugung und verwandlung des lichtes betreffenden heuristischen gesichtspunkt*, Annalen der Physik **322** (1905), no. 6, 132–148.

[23] J. Ellis, J. R. Espinosa, G. F. Giudice, A. Hoecker, and A. Riotto, *The Probable Fate of the Standard Model*, Phys. Lett. B **679** (2009), 369–375.

[24] F. Englert and R. Brout, *Broken symmetry and the mass of gauge vector mesons*, Phys. Rev. Lett. **13** (1964), no. 9, 321–323.

[25] John Freeman, CDF Internal Note 9848, *Summer 2009 SecVtx Scale Factors Calculated Using the Electron Method through Period 22*, (2009).

[26] Y. Fukuda et al., *Measurements of the Solar Neutrino Flux from Super-Kamiokande's First 300 Days*, Phys. Rev. Lett. **81** (1998), 1158–1162.

[27] John F. Gunion et al., *The higgs hunter's guide*, Addison-Wesley, 1990.

[28] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, *Global conservation laws and massless particles*, Phys. Rev. Lett. **13** (1964), no. 20, 585–587.

[29] Thomas Hambye and Kurt Riesselmann, *Matching conditions and higgs boson mass upper bounds reexamined*, Phys. Rev. D **55** (1997), no. 11, 7255–7262.

[30] Joel Heinrich, *Bayesian Limit Software: multi-channel with correlated backgrounds and efficiencies*, CDF Public Note 7587, April 2005.

[31] Peter W. Higgs, *Broken symmetries and the masses of gauge bosons*, Phys. Rev. Lett. **13** (1964), no. 16, 508–509.

[32] Andreas Hocker et al., *TMVA: Toolkit for multivariate data analysis*, PoS **ACAT** (2007), 040.

[33] Thomas Junk, *Confidence Level Computation for Combining Searches with Small Statistics*, Nucl. Instr. Meth. A **434** (1999), 435–443.

[34] Mark Lancaster, *Combination of CDF and D0 results on the mass of the top quark using up to 5.8 fb^{-1} of data*, (2011).

[35] S. Lockwitz et al., *A Search for the Standard Model Higgs Boson in the Process $ZH \to e^+e^-b\bar{b}$ Using 7.5 fb^{-1} of CDF II Data*, CDF Public Note 10593, July 2011.

[36] Michelangelo L. Mangano, Mauro Moretti, Fulvio Piccinini, Roberto Pittau, and Antonio D. Polosa, *ALPGEN, a generator for hard multiparton processes in hadronic collisions*, JHEP **07** (2003), 001.

[37] K. Nakamura et al., The Particle Data Group, *The Review of Particle Physics*, J. Phys. G **37** (2010), 075021.

[38] Enrique Palencia and Ford Garberson, CDF Internal Note 8882, *Jet Probability B-Tag Efficiency Measurement Using Muon Transverse Momemtum for 1.2 fb^{-1} analyses*, (2007).

[39] Svenja Richter, Ph.D. Thesis (Advisors: Feindt, M., Wagner, W.), *Search for electroweak single top-quark production with the CDF II experiment*, (2007).

[40] E. Rutherford, *LXXIX. The scattering of α and β particles by matter and the structure of the atom*, Philosophical Magazine Series 6 **21** (1911), no. 125, 669–688.

[41] Torbjorn Sjostrand, Stephen Mrenna, and Peter Z. Skands, *PYTHIA 6.4 Physics and Manual*, JHEP **05** (2006), 026.

[42] M. Spira, http://people.web.psi.ch/spira/proglist.html, *V2HV Program*.

[43] The CDF and D0 Collaborations, The Tevatron New Phenomena and Higgs Working Group, *Combined CDF and D0 Upper Limits on Standard Model Higgs Boson Production with up to 8.6 fb^{-1} of Data*, arXiv:1107.5518 (hep-ex) (2011).

[44] J.J. Thomson, *Cathode rays*, Phil.Mag. **44** (1897), 293–316.

[45] Timo van Ritbergen and Robin G. Stuart, *Complete 2-loop quantum electrodynamic contributions to the muon lifetime in the fermi model*, Phys. Rev. Lett. **82** (1999), 488–491.

[46] J.S. Wilson et al., CDF Public Note 10574, *Search for the Higgs Boson Produced in Association with Top Quarks Using 7.5 fb^{-1}*, (2011).

Appendix A
Artificial Neural Networks

The analysis described in this thesis relies heavily on the use of multivariate techniques. For the multivariate muon selection, the expert discriminants, the jet-flavor separator, and the final event discriminants, artificial neural networks are used. This chapter details the artificial neural network technique and how these functions are obtained. A more detailed description of neural networks and other machine learning multivariate techniques can be found in Ref. [32].

Neural Network Structure

The basic structure of a neural network consists of many layers of nodes. These individual nodes mimic the behavior of neurons, or brain cells, from which the name of this technique is derived. An input layer consists of one node for each input quantity, such as the various kinematic quantities used for event discrimination. The output layer usually contains one node, which produces the output score of the neural network. Sometimes two or more output nodes are used to simultaneously train for two target outputs. Between the input and output layers, there are various hidden layers. These hidden layers provide the discriminating power of the algorithm. Figure A.1 shows the structure of a typical neural network with input, hidden, and output layers.

The individual nodes have associated weights which are used to produce the final output score of the neural network for a given set of input variables. For a simple neural network with N input nodes and one output node, a weighted sum of the input variables x_i and the weights w_i is used. A constant bias term, w_0, is also added to the network as another input

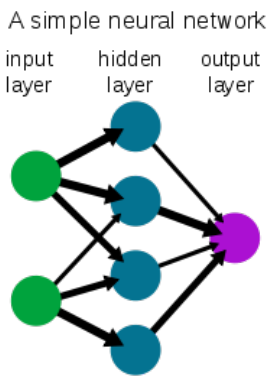


Figure A.1: A simple neural network consisting of an input layer with 2 nodes, a single hidden layer with 4 nodes, and a single output node.

node. Furthermore, each node of the network has an activation function, similar to that of the threshold for a biological neuron to activate. Common choices for activation functions are sigmoids, hyperbolic tangent, and linear functions. Equation A.1 shows the equation describing the output score, χ, of this network consisting of N input nodes and one output node.

$$\chi = \tanh\left[w_0 + \sum_{i=1}^{N} x_i \cdot w_i\right] \tag{A.1}$$

This notion can be extended for the addition of a hidden layer. Each input quantity is weighted by a unique factor and passed to each individual hidden nodes. The output of a single hidden node is computed as in Equation A.1, and this output is passed on to the next layer, with a second unique weight factor. This process repeats for all the hidden nodes and layers until the final output score is obtained. For one hidden layer with N_H hidden nodes, the equation describing the output score can be written as in Equation A.2, for a general activation function f.

$$\chi = f\left(\sum_{j=1}^{N_H} w'_j \cdot f\left(\sum_{i=1}^{N} w_{ij} \cdot x_i\right)\right) \tag{A.2}$$

Training Process

To produce a neural network that performs well, the weights for the output of each input and hidden node need to be determined. This is done with a large sample of training events for which the expected output score is known. For example, in the final event discriminant neural networks, the Higgs signal events have an expected output score of 1.0, while the background events have expected output scores of 0.0. To determine the weights in the training process, these events are used to minimize a least-squares function, as in Equation A.3. Using the set of input variables $\vec{x}_i$ and expected output score $y(\vec{x}_i)$ (usually 1 or 0) for each event, the set of weights $\vec{w}$ can be determined by minimizing the function of Equation A.3.

$$\frac{1}{2}\sum_{i=1}^{N_{\text{ev}}}[y(\vec{x}_i) - \chi(\vec{x}_i, \vec{w})]^2 \tag{A.3}$$

After the optimal weights have been determined, the function can be evaluated on an orthogonal set of events, the 'testing' sample, to validate the performance of the new neural network. Evaluating the error on the testing sample as in Equation A.4 can provide a metric of ranking the performance of different networks having various numbers of hidden nodes and layers. As a general rule, it is best to use the simplest network architecture possible to populate all the hidden nodes with an adequate number of events in the training process.

$$\sigma = \sqrt{\frac{1}{N_{\text{ev}}}\sum_{i=1}^{N_{\text{ev}}}[y(\vec{x}_i) - \chi(\vec{x}_i, \vec{w})]^2} \tag{A.4}$$

Once the training process is complete, the newly produced neural network can then be evaluated using new input events for which the output is unknown to classify them accordingly. This forms the basis of using neural networks as functions to classify signal and background.

There are other machine learning techniques, apart from neural networks, that are used in other physics analyses. These include boosted decision trees, random forests, support vector machines, and others. The details of these techniques will not be described here. However, the use of neural networks and these other techniques in physics analyses has immensely improved

the sensitivity of analyses to processes with small cross sections. By exploiting the correlations between kinematic quantities in signal and background, neural networks can produce output scores that distinguish signals from background more aptly than using a single kinematic variable for classification.

Appendix B

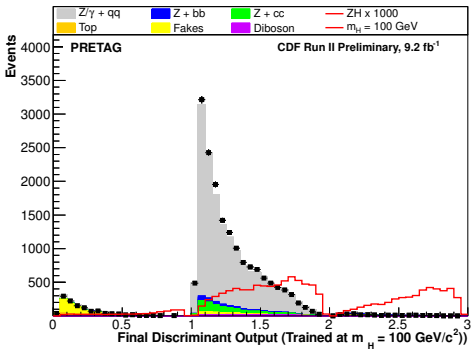
Discriminant Output Plots for all Higgs Mass Hypotheses

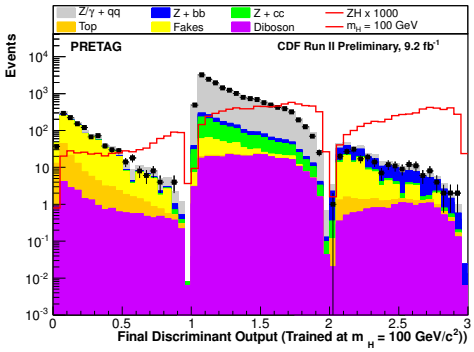
The following pages show the final output discriminants for each Higgs mass hypothesis tested in this analysis, from 100 GeV/c^2 to 150 GeV/c^2 in 5 GeV/c^2 increments. There are five plots shown for each m_H, two prior to the application of b-tagging (linear and logarithmic scales), as well as one for each of the three b-tagging cateogries. The b-tagging categories used are, in order of increasing signal purity:

1. A single tight secondary vertex tag (SVT)
2. One loose secondary vertex tag and one jet probability (5%) tag (SVL + JP)
3. Two tight secondary vertex tags (SVT + SVT)

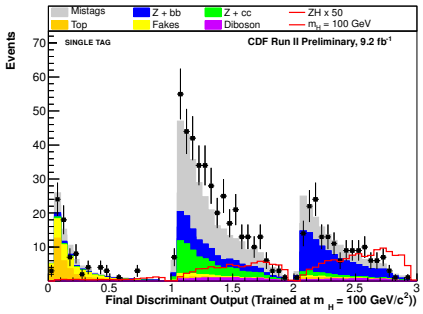
$m_H = 100\ \text{GeV}/c^2$ **Discriminant**

PRETAG

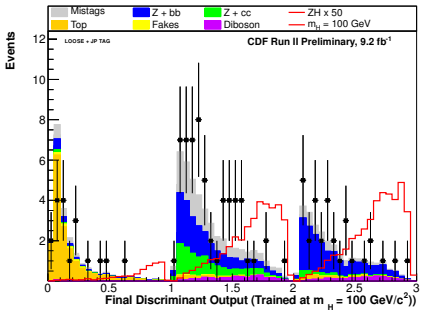




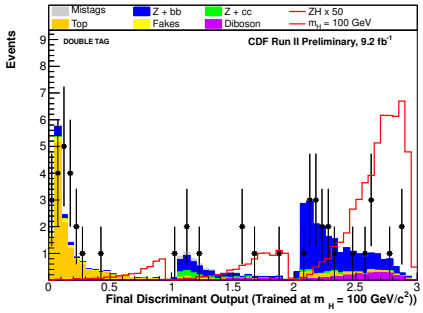
SVT



SVL+JP

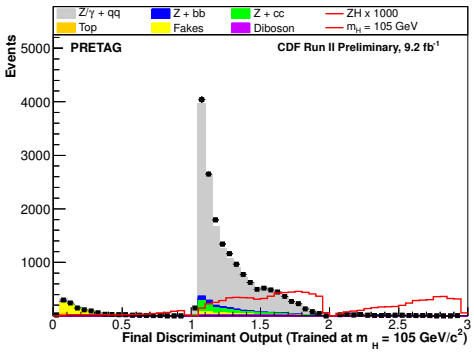


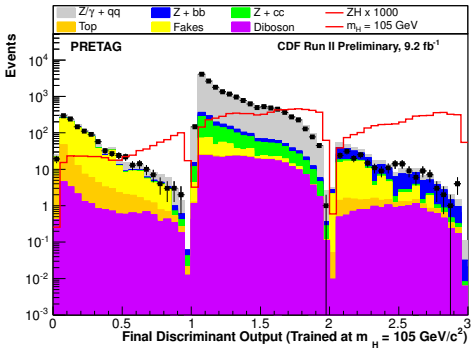
SVT+SVT



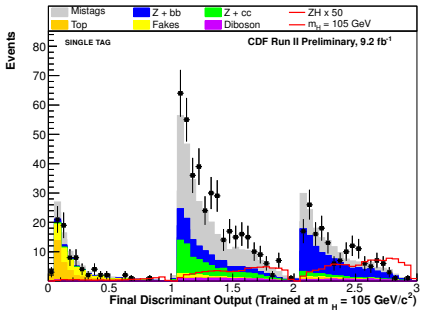
$m_H = 105\ \text{GeV}/c^2$ Discriminant

PRETAG

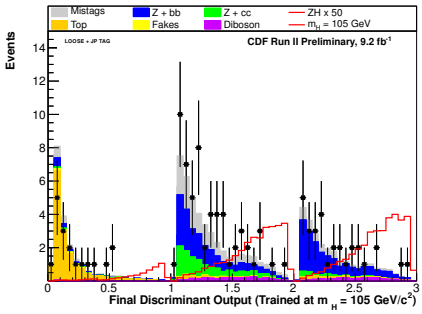




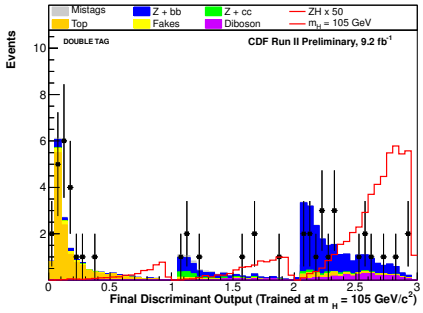
SVT



SVL+JP

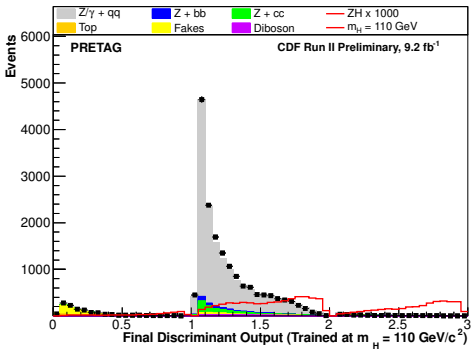


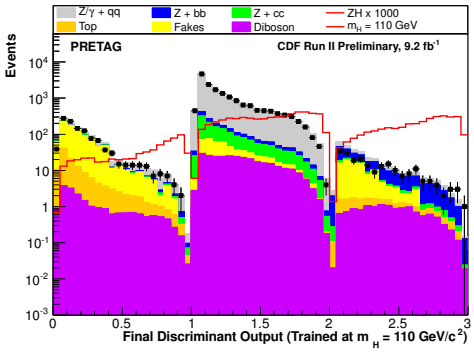
SVT+SVT



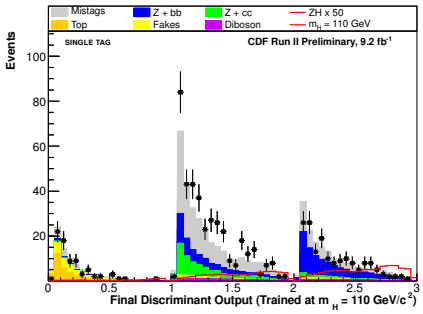
$m_H = 110\ \text{GeV}/c^2$ Discriminant

PRETAG

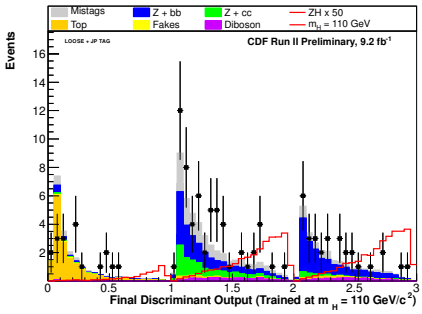




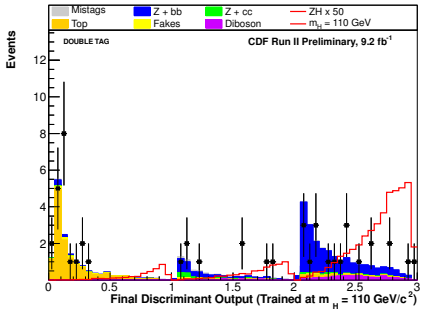
SVT



SVL+JP

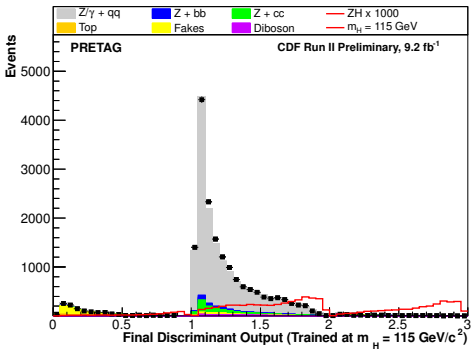


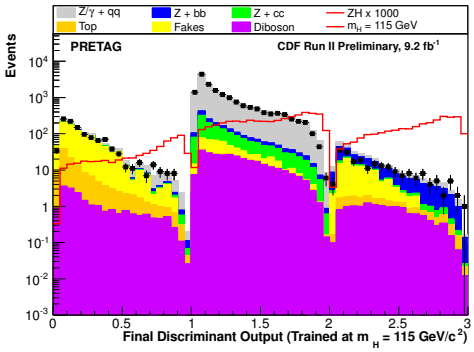
SVT+SVT



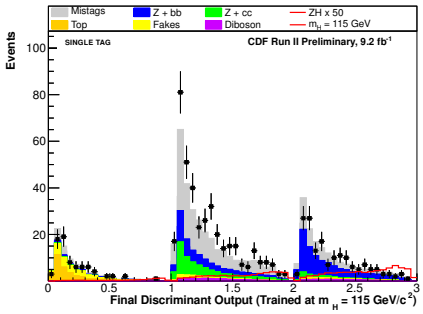
$m_H = 115\ \text{GeV}/c^2$ Discriminant

PRETAG

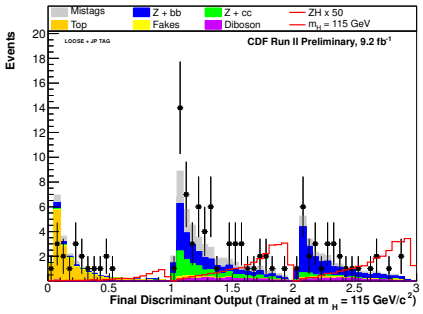




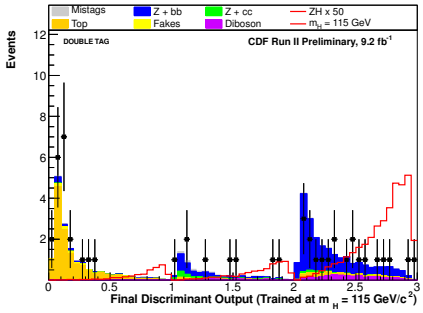
SVT



SVL+JP

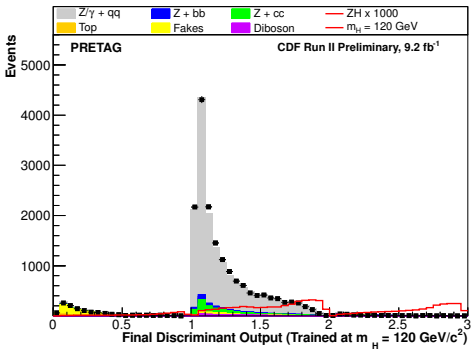


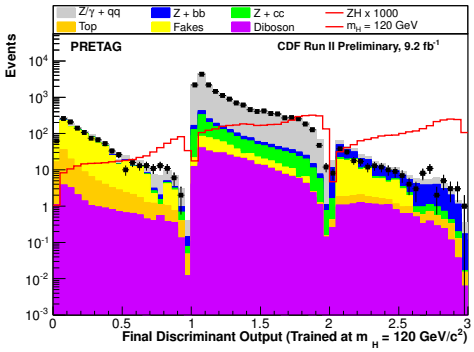
SVT+SVT



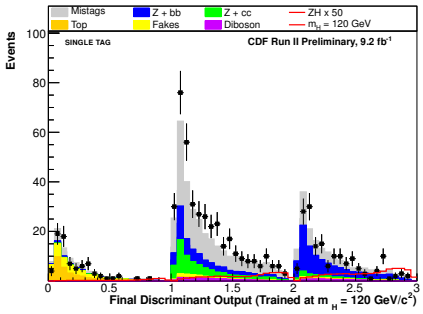
$m_H = 120\ \text{GeV}/c^2$ Discriminant

PRETAG

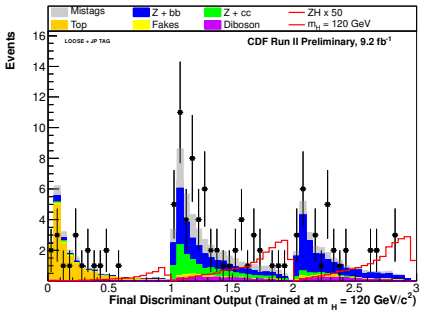




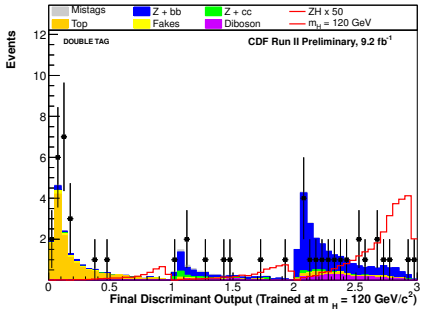
SVT



SVL+JP

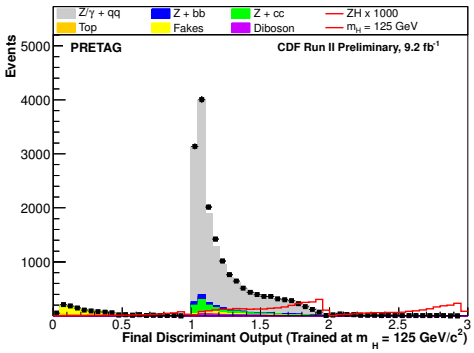


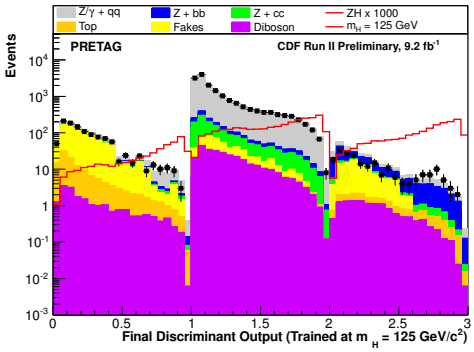
SVT+SVT



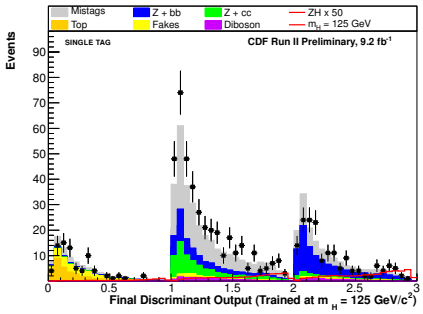
$m_H = 125\ \text{GeV}/c^2$ **Discriminant**

PRETAG

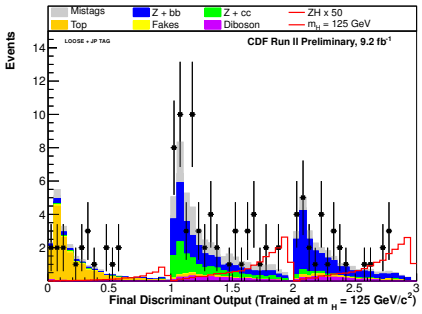




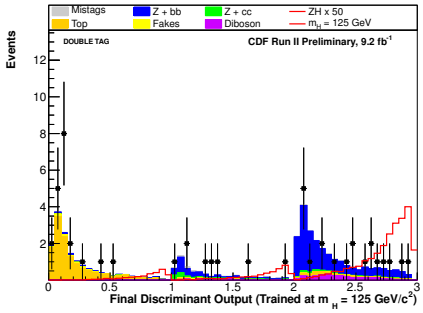
SVT



SVL+JP

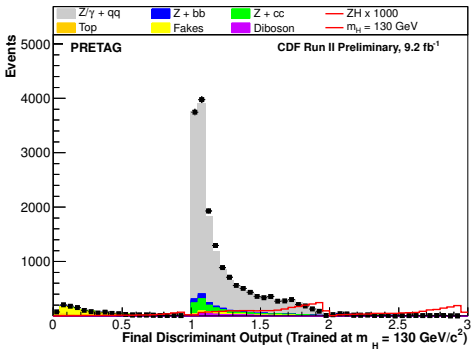


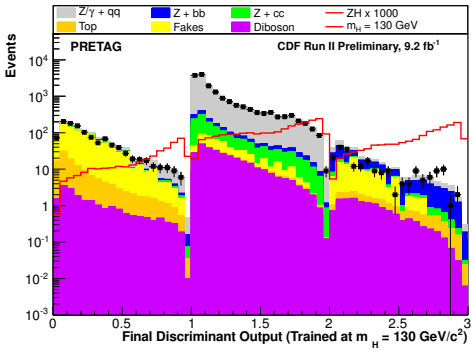
SVT+SVT



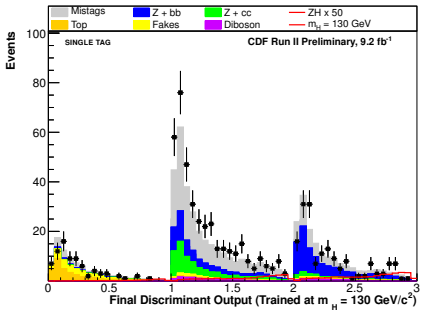
$m_H = 130\ \text{GeV}/c^2$ Discriminant

PRETAG

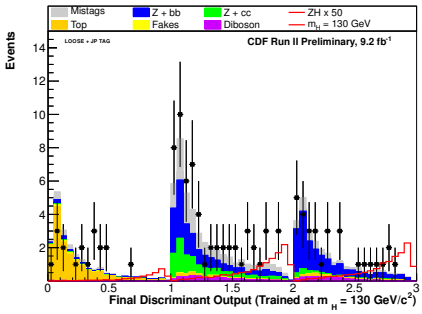




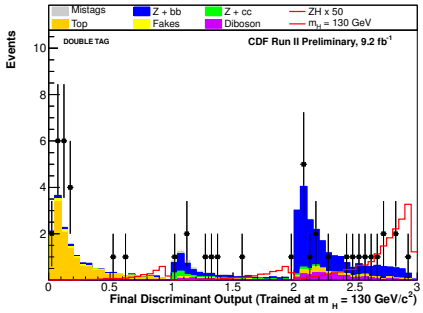
SVT



SVL+JP

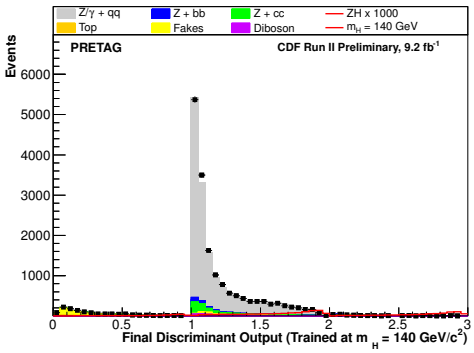


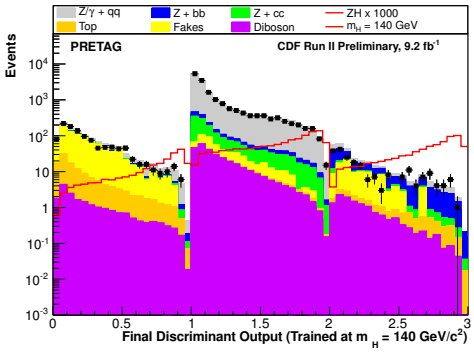
SVT+SVT



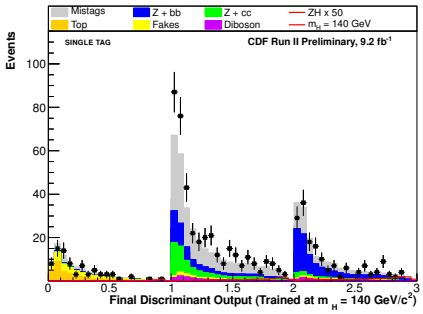
$m_H = 140\ \text{GeV}/c^2$ **Discriminant**

PRETAG

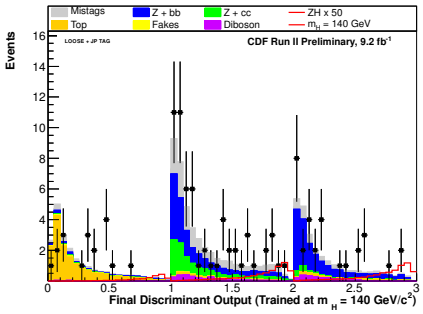




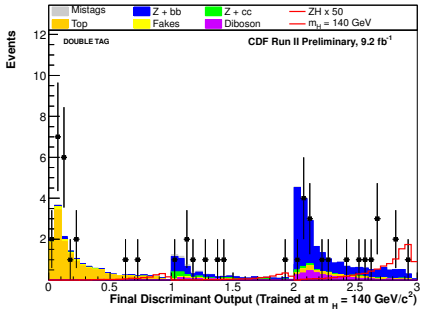
SVT



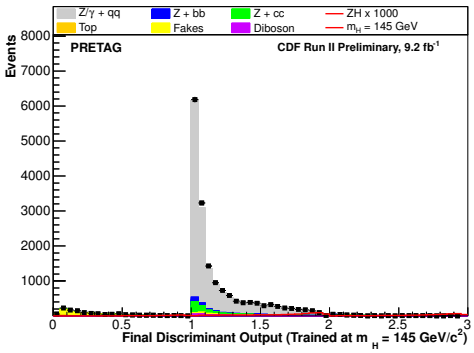
SVL+JP



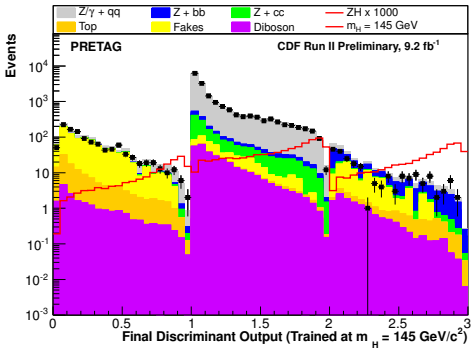
SVT+SVT



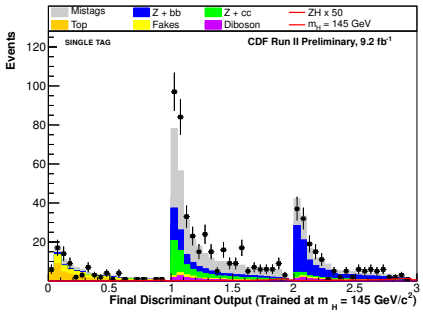
$m_H = 145\ \text{GeV}/c^2$ **Discriminant**



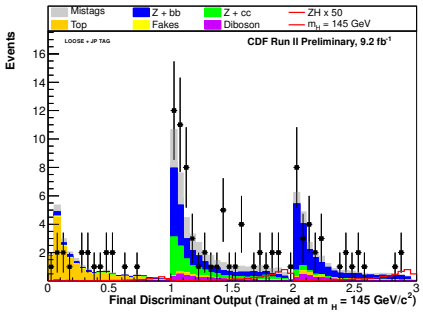
PRETAG



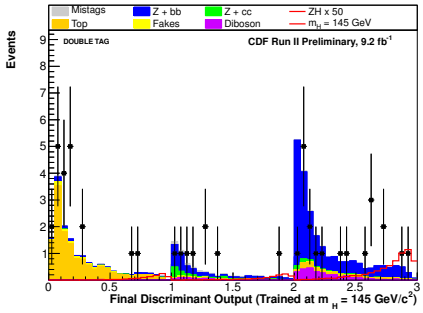
SVT



SVL+JP

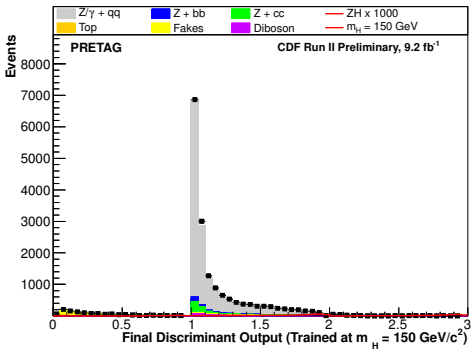


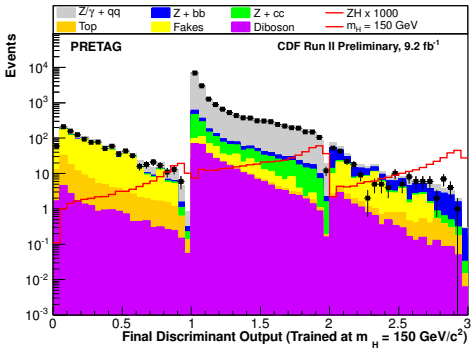
SVT+SVT



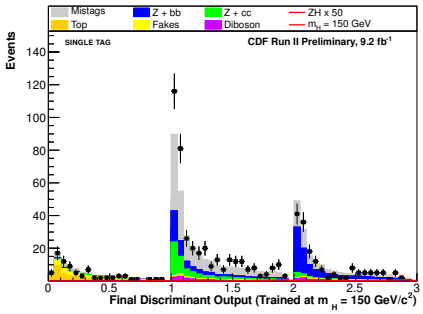
$m_H = 150\ \text{GeV}/c^2$ Discriminant

PRETAG

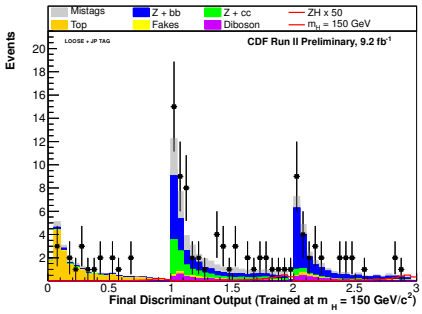




SVT



SVL+JP



SVT+SVT