SEARCH FOR $WZ + ZZ$ PRODUCTION WITH MISSING TRANSVERSE ENERGY AND b JETS AT CDF

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Stephen Thomas Poprocki

January 2013

SEARCH FOR $WZ + ZZ$ PRODUCTION WITH MISSING TRANSVERSE ENERGY

AND b JETS AT CDF

Stephen Thomas Poprocki, Ph.D.

Cornell University 2013

Observation of diboson processes at hadron colliders is an important milestone on the road to discovery or exclusion of the standard model Higgs boson. Since the decay processes happen to be closely related, methods, tools, and insights obtained through the more common diboson decays can be incorporated into low-mass standard model Higgs searches. The combined $WW + WZ + ZZ$ diboson cross section has been measured at the Tevatron in hadronic decay modes. In this thesis we take this one step closer to the Higgs by measuring just the $WZ + ZZ$ cross section, exploiting a novel artificial neural network based b-jet tagger to separate the WW background. The number of signal events is extracted from data events with large $\not{E}_{\mathrm{T}}$ using a simultaneous fit in events with and without two jets consistent with B hadron decays. Using $5.2\,\mathrm{fb}^{-1}$ of data from the CDF II detector, we measure a cross section of $\sigma(p\bar{p} \to WZ, ZZ) = 5.8^{+3.6}_{-3.0}$ pb, in agreement with the standard model.

Biographical Sketch

Stephen Poprocki was born on October 21, 1985 in Oberlin, Ohio and raised in Amherst, Ohio. He was a diligent and inquisitive child who questioned how things work and why people behave the way they do, reluctant to follow in their footsteps without careful consideration. Frustrated with wasting his time at school doing busywork and having his learning slowed by other students, he dropped out of high school halfway through the 10th grade. He then educated himself, focusing on science, music, Japanese, and programming, before going off to college one year early. He attended the College of Wooster in Ohio and double majored in physics and mathematics. During that time he spent one semester abroad at Kanda University of International Studies in Chiba, Japan—an experience that had a profound and lasting impact on him.

In Fall 2007, Stephen joined the physics Ph.D. program at Cornell University with the intent of furthering his understanding of how nature works. After two incredibly intense years of coursework and teaching, he began working with Peter Wittich on the CDF analysis that became the topic of this thesis. Stephen traveled often to Fermilab to work closely with colleagues, give presentations, and perform a three month shift as the data acquisition Ace for CDF. After completion of the analysis, Stephen began working on testing the VICTR track-trigger chip in 2011.

During his time at Cornell, Stephen grew in many unexpected ways. One was the discovery of his love for physical activities, especially outdoors. Another was his deepened appreciation of nature, beyond the scope of physics, obtained through the natural beauty of Ithaca and his countless bike rides through the Finger Lakes countryside.

To nature

ACKNOWLEDGMENTS

This work would not have been possible without the help of many to whom I owe my thanks. First and foremost is my advisor Peter Wittich for allowing me to work with him and for coming up with a research project which matched my skills and interests. Peter was especially flexible, patient, helpful, friendly, and respectful. I am lucky to have worked with him. I also owe thanks to my colleagues at Fermilab who contributed to this analysis: Vadim Rusu, Sasha Pranko, John Freeman, and Wes Ketchum. I extend my thanks to everyone who has worked on the Tevatron and CDF. For the VICTR project, I thank Ron Lipton, Zhenyu Ye, Jim Hoff, Jim Alexander, Anders Ryd, Julia Thom, and Nic Eggert.

I also thank Anders Ryd for his willingness to help when I took his particle physics class, and Yuval Grossman for helping me see the bigger picture of particle theory without focusing on the equations. I also thank both of them for being on my committee.

I thank my classmates for their help in surviving the first years of grad school: Danny, Ben, Walter, Hitesh, Turan, Kendra, Yao, Ben H., Mike, and anyone I forgot to mention. Also thanks to Johannes for helping immensely during E&M office hours, Don and Yao for being good officemates, Flip for his help with my theory questions, Colwyn for introducing me to cycling, Shankar for teaching me how to cook, my table tennis buddies Augustine, Saty, and Al for giving me something to do while at Fermilab, and a handful of friends I'm happy to have spent some of my time with: Gang, Xin, Darren, Souvik, Avi, Y.J., Yi-Ting, Shayna, and Jim Rohal to name just a few.

Last but not least, I thank my family (Mom, Dad, Mat, Tim) and my wonderful girlfriend Hyeri, for their love and support throughout everything.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

1.1 HISTORY

Physics developed out of natural philosophy in ancient Greece starting around 600 BC. The pre-Socratic Greek philosophers Thales and Democritus are credited as the fathers of science. Thales rejected the common explanations of natural phenomena based on mythology or the supernatural, and instead insisted on finding natural explanations. Democritus (with his teacher Leucippus) is credited for the first description of matter being made up of indivisible discrete units called atoms. From that time until around when Greece was annexed by Rome in 146 BC, many great thinkers changed the way we viewed the world including Pythagoras (proposed Earth is spherical, advances in mathematics), Aristotle (wrote compilations of scientific knowledge), Archimedes (hydrostatics), and Ptolemy (optics, planetary motion).

Advances were also made in China—most notably in magnetism with the invention of the compass (c. 200 BC – 100 AD)—and in India where Aryabhatiya proposed in 499 AD that the Earth revolves around its axis. Scientific knowledge from ancient Greece was translated into Arabic and incorporated into Islamic philosophy, with the most notable contributions coming from Alhazen (Book of Optics, c. 1000 AD), and Al-Biruni

(hydrodynamics). Their work was later studied and introduced into Western Europe by Roger Bacon in the 1200's.

The European Renaissance improved on the Greek way by emphasizing the importance of experimentally testing scientific theories. From c. 1400 AD – c. 1900 AD, in chronological order, the field received great advancements by Copernicus (Earth revolves around Sun), Galileo & Tycho (observational astronomy), Kepler (laws of planetary motion, gravity), Newton (mechanics), Young (light), Faraday/Henry/Maxwell (electricity & magnetism), Stoney (theory of electron), Röntgen (X-rays), Curie (radioactivity), and J. J. Thompson (discovery of electron) among others. Furthermore, there were important developments in optics, thermodynamics, fluid dynamics, and other subfields.

The 20th century brought even more changes in physics. Newtonian mechanics was succeeded by relativity and quantum mechanics, and the early atomic model by modern particle physics. Further, many new subfields were developed such as condensed matter physics, astrophysics & cosmology, computational physics, etc. To keep the story short, let us now just focus on the developments related to particle physics. From 1900 to around 1930, physicists like Planck, Einstein, Bohr, and Compton uncovered the quantum nature of light and the particle/wave duality of photons. Next Schröedinger, Heisenberg, and Dirac quickly tied together the theory of quantum mechanics. Along with Einstein's development of special relativity, it seemed to some that theoretical physics was complete, aside from a few details[1]. If only nature were so simple as to only have protons, electrons, and photons...

But no, the neutron was discovered by James Chadwick in 1932, as predicted by Rutherford in 1920 in order to solve the problem of an atom's mass not matching the prediction given by the proton/electron model. Also in 1932 the positron was discovered

by Carl Anderson, as foreshadowed by negative energy solutions to the Dirac equation. This was the first discovery of anti-matter. There was still another puzzle left unanswered: what holds the nucleus together? Physicists suspected another force, which they called the strong force, and Yukawa wrote down a theory which predicts a strong force mediating particle he called the pion. In 1936, Carl Anderson found a new particle while studying cosmic rays, and people believed it was the pion. But its properties did not quite match up. It was not until 1947 that it became clear that there were actually two new particles seen in cosmic rays: the muon and the pion. I.I. Rabi said of the muon, "who ordered that?" since it was unexpected and seemingly unnecessary.

In the 1950's a handful of new particles ($K, \Lambda, \Delta, \Sigma, \Xi$ and their charged or neutral variants) were discovered by observing decays from cosmic rays using cloud chamber detectors, and also thanks to the first modern particle accelerator—the $1.3\,\mathrm{GeV}$ Brookhaven Cosmotron. In the 60's this trend continued with the discovery of the ρ, ω, η, and ϕ mesons. But it was not a matter of just having a lot of particles, it was also a matter of classifying them in some way according to how they do or do not decay. This "zoo" of particles seemingly defied organization until Murray Gell-Mann proposed his Eightfold Way[2], a sort of periodic table of elementary particles.

In the midst of discovering these numerous middleweight and heavyweight particles

---

[1]Curiously enough, the same thing happened previously, just before the quantum revolution:

"In this field, almost everything is already discovered, and all that remains is to fill a few unimportant holes." — Philipp von Jolly to his student Max Planck, 1878.

"Our future discoveries must be looked for in the sixth decimal place." — A. A. Michelson, 1894.

"Physics, as we know it, will be over in six months." — Max Born, 1928.

[2]This name is no doubt inspired by the Eightfold Path in Buddhism which leads to the cessation of suffering, self-awakening, and insight into true reality through eight "right" elements: right view, intention, speech, action, livelihood, effort, mindfulness, and concentration.

(with respect to the electron (light) and proton (heavy) masses, and called mesons and baryons, respectively), there was one curious particle which deserves separate mention: the neutrino. In the cosmic ray bubble chamber experiments of the 50's, when looking at beta decay[3], pion decay[4], and muon decay[5], there seemed to be invisible particles participating. There were two big clues that was the case: kinks in the tracks, and resulting particle energy distributions inconsistent with two-body decays. First theorized by Pauli in 1930, these invisible particles turned out to interact with matter extraordinarily weakly and hence were not directly detected until 1956 by Cowan and Reines.

Building on his Eightfold Way, Gell-Mann proposed in 1964 that all hadrons are made up of more fundamental particles called quarks. Initially only three types (up, down, strange) were needed to explain all the known hadrons, but in 1974 a new, much heavier, meson was discovered. This particle was discovered (and named) by two groups independently (S. Ting at Brookhaven, B. Richter at SLAC) and carries the name J/ψ. To explain this particle, a fourth quark, the charm quark, was added to the quark model. This gave rise to predictions of undiscovered hadrons containing a charm, and indeed these were later discovered ($\Lambda_c, \Sigma_c, D, \ldots$). In 1976 a third lepton, the τ, was discovered by M. Perl at SLAC, and like the muon, was rather unexpected. However, one year later, another even heavier meson, the Υ, was found by Leon Lederman at Fermilab and understood to contain a fifth quark known as the bottom quark. Shortly after, other hadrons containing the bottom quark were discovered ($\Lambda_b, \Sigma_b, B, \ldots$). This gave three generations of leptons (e, μ, τ) but just 2.5 pairs of quarks. The missing sixth quark, the top quark, unexpectedly turned out to be some 40 times heavier than the bottom quark and was not discovered until 1995, by the CDF and DØ collaborations at Fermilab.

---

[3] $n \to p + e + \bar{\nu}$

[4] $\pi \to \mu + \nu$

[5] $\mu \to e + \nu + \bar{\nu}$

Finally, let us discuss the force mediating particles again. Yukawa had posited the pion as the mediator of the strong force, but now with the understanding of quarks, the strong force was fundamentally understood to be the exchange of gluons between quarks. The gluon was indirectly observed in 1979. For the electromagnetic force, the mediating particle was understood to be the photon. What about the weak force? In the 1960's, Glashow, Weinberg, and Salem unified the theories of the electromagnetic and weak forces into the electroweak force. In this theory, the weak force is mediated by W and Z bosons. With a theoretical prediction of the masses of these particles, a group headed by C. Rubbia at CERN discovered both the W and Z bosons in 1983. In this scheme, the gravitational force would be mediated by a hypothetical particle called the graviton, but its detection is practically impossible due to its extremely weak interaction with matter.

We have now discussed all of the known fundamental particles. The timeline of their discoveries is shown in Fig. 1.1.

The theory describing the fundamental particles and their interactions, known as the Standard Model of particle physics, and described in §2, started development in the 1970's and has proven to be very successful. However, in addition to the issues discussed in §2.3, there is one important piece missing. In the Standard Model, the way in which particles, and hence matter, obtain mass is described by the Higgs Mechanism and will be discussed in §2.7.1. This theory predicts another particle, known as the Higgs boson, but as of 2012 it has not yet been detected. A discovery of the Higgs would be the crowning achievement of the theory, while a non-discovery would produce many questions about the new, unknown physics. Indeed, the search for the Higgs is currently the principal effort of particle physicists and will likely remain so until our questions are answered. The Higgs will be discussed more in §2.7.
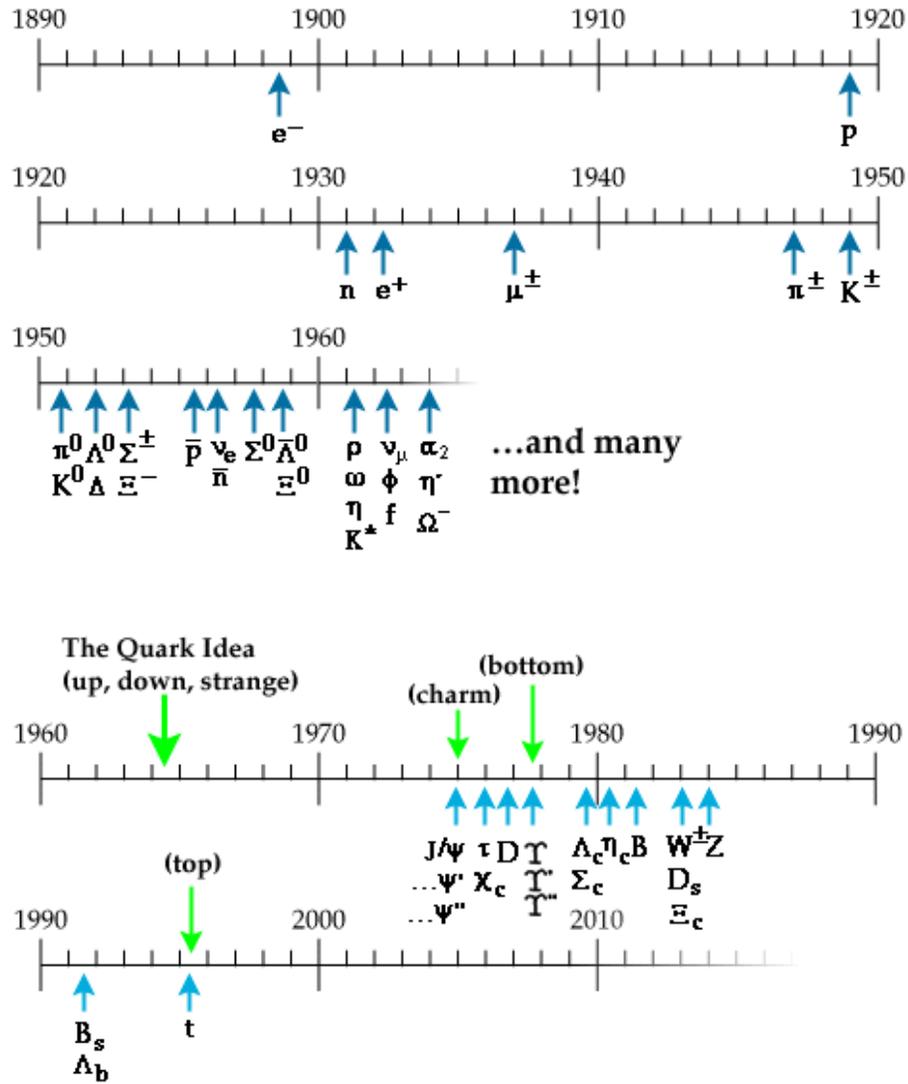
Figure 1.1: Timeline of particle discoveries. Figure from [1].

1.2 MOTIVATION

As we have seen, much of the progress in particle physics experiments has been, or has come from, the discovery of new particles. However, so far that trend has not continued beyond 1995 when the top quark was discovered. Indeed, the only new particle left for us to find, if the Standard Model is correct, is the Higgs boson. However, despite great efforts, we have not found it, and only now at the time of writing this do we feel we are closing in on it. Other people have instead focused on finding evidence that the Standard Model is not correct or complete, and have conducted many tests looking for new physics. At this time, no compelling evidence has been found.

This thesis takes a different approach in that although we are interested in the Higgs, but we do not search for it directly. Furthermore, we are not specifically searching for any new physics. Instead, we attempt to measure very rare decays which are related to the Higgs. These events are when two gauge bosons are produced, so called diboson events, and we are focusing on WZ and ZZ in this thesis. We do this because the same decays occur when one of the Z's is replaced with an H, since both particles decay to two b quarks. This is shown in Fig. 1.2.

The main reason we search for WZ/ZZ instead of WH/ZH is that it occurs more often. The cross section times branching ratio, which how we quantify how often a decay occurs, for $WZ + ZZ$ with $Z \to b\bar{b}$ at CDF is about six times higher than for $WH + ZH$ with a Higgs of mass $m_H = 125\,\mathrm{GeV}/c^2$ decaying to $b\bar{b}$ [2]. However, it still turns out we did not have sensitivity to this channel alone. Although we do perform that measurement, we also measure the $WZ + ZZ$ decay where the Z decays to any hadrons. This cross section times branching ratio is $5.1\,\mathrm{pb}$ and we end up with some sensitivity to this measurement.
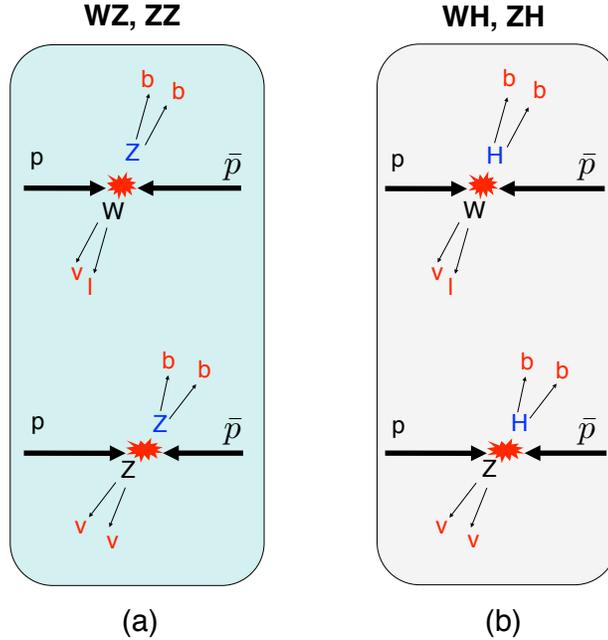
Figure 1.2: (a) The WZ and ZZ diboson decays we are measuring. (b) The related WH and ZH decays with the Higgs. These decays are similar since both the Z and a low mass Higgs decay to two b quarks.

There is further motivation though. A large and important background is this measurement is WW. In order to veto these events, we can require the jets to be consistent with b quark jets instead of other hadronic jets. Since W does not decay to two b quarks like the Z and H, we can cut out this background. It turns out that this identifying of jets as coming from b quarks is very important to many analyses. To explore potential improvements in these b-tagging algorithms, we have created our own novel artificial neural network based b-tagger. This tagger is now being incorporated into Higgs searches at CDF and will be discussed in §4.

Lastly, measurements of diboson production cross sections provide tests of the self-interactions of the gauge bosons. Deviations from the standard model (SM) prediction for the production rates could indicate new physics [3, 4], specifically in hadronic final states [5].

CHAPTER 2

THE STANDARD MODEL

We have introduced the Standard Model as the theory which describes the fundamental particles and their interactions. While a complete treatment of the Standard Model is clearly beyond the scope of this thesis, in this chapter we will first give a verbal and mathematical summary of the Standard Model, discuss the remaining issues, present boson and diboson phenomenology, discuss the Higgs mechanism, derive the boson and fermion mass terms and couplings to the Higgs, and finish with Higgs boson phenomenology and the current search results.

2.1   DESCRIPTION

Elementary particles are small, and furthermore, they are often traveling very fast. Classical mechanics breaks down in both of these régimes. For this reason, the Standard Model is formulated as a quantum field theory (see Fig. 2.1).

In general terms, the Standard Model can be described as this quantum field theory framework together with the constituent particles and interactions it describes. These fundamental particles and their properties are summarized in Fig. 2.2. As we are already familiar with, we have the three generations of quark pairs, the 3 generations of leptons and their corresponding neutrinos, the force mediating gauge bosons, and the Higgs boson. The quarks come in three "colors", though this is just a fancy term for a certain
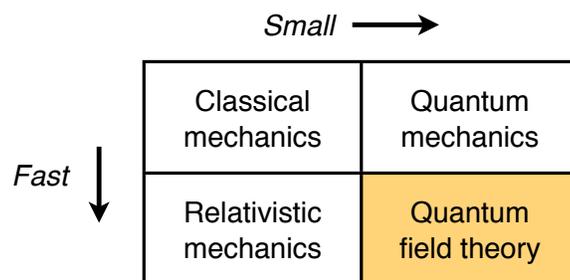
Figure 2.1: The four effective régimes of mechanics. Classical mechanics breaks down at the small scale and requires quantum mechanics. Similarly it breaks down as speeds approach the speed of light and requires relativity. When things are both small *and* fast, we need the combination of quantum mechanics and relativity: quantum field theory.

quantum number and has nothing to do with visual colors. The gluon carries two colors making eight different color states. In nature, we cannot directly see particles with color due to a principle called color confinement, so colored particles must be in combinations which cancel out the color for us to see them.

The way the particles interact can be depicted as Feynman diagrams. Quantum field theory gives us the machinery to turn these diagrams into mathematical calculations.

2.2 MATHEMATICAL SUMMARY

The Standard Model is summarized by:

1. **Gauge group**: $SU(3)_C \times SU(2)_L \times U(1)_Y$.

2. **Matter representations**: These are the left-handed quark doublet $Q = (u_L, d_L)$, right-handed up U, right-handed down D, left-handed lepton doublet $L = (\nu_{e,L}, e_L)$,
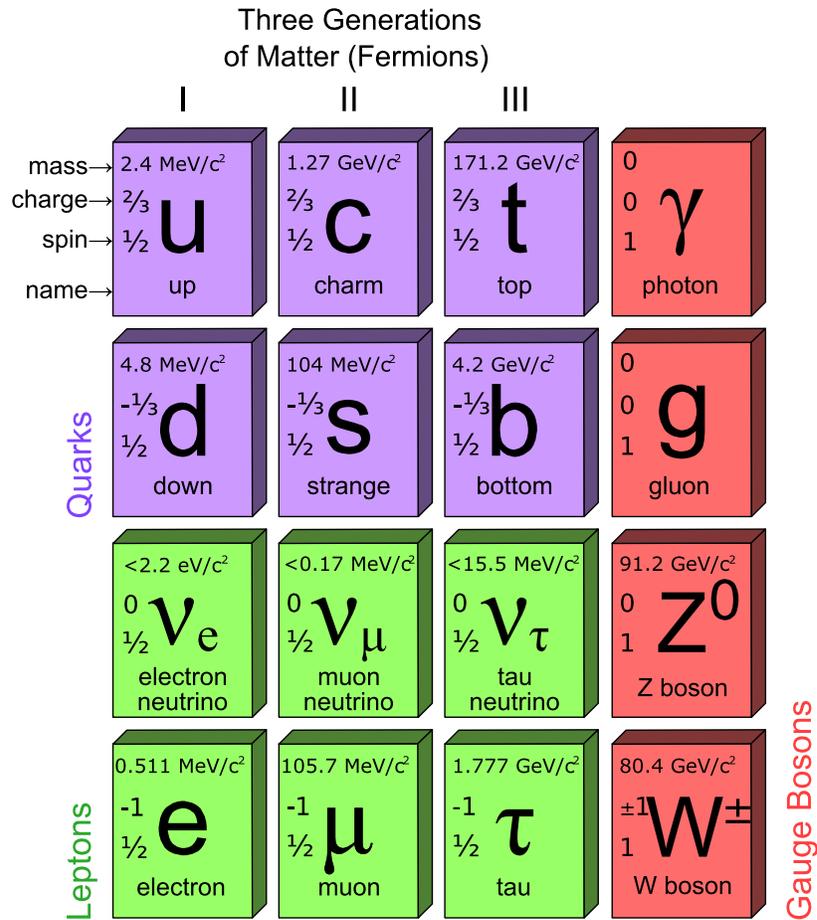
Figure 2.2: The fundamental particles in the Standard Model. The Higgs boson could be shown off to the side for it does not mediate a force and we have not confirmed its existence or properties. Figure from [6].

and right-handed charged lepton E. A nice way to remember this is "QUDLE", and write

$$Q_{Li}(3,2)_{+1/6}, \quad U_{Ri}(3,1)_{+2/3}, \quad D_{Ri}(3,1)_{-1/3}, \quad L_{Li}(1,2)_{-1/2}, \quad E_{Ri}(1,1)_{-1},$$

where the notation is in the form $(c, L)_Y$ and $i = 1, 2, 3$ is the generation index. There is also a single scalar representation

$$\phi(1,2)_{+1/2}.$$

3. **Spontaneous symmetry breaking**: The scalar ϕ has a vacuum expectation value

$$\langle \phi \rangle = \begin{pmatrix} 0 \\ v/\sqrt{2} \end{pmatrix} \tag{2.1}$$

and breaks the symmetry $SU(2)_L \times U(1)_Y \rightarrow U(1)_{\text{EM}}$.

We write the most general renormalizable Lagrangian consistent with these ingredients, and we can divide it into 3 parts:

$$\mathcal{L}_{\text{SM}} = \mathcal{L}_{\text{kinetic}} + \mathcal{L}_{\text{Higgs}} + \mathcal{L}_{\text{Yukawa}}.$$

For the kinetic terms we use the covariant derivative

$$D_\mu = \partial_\mu + ig_s G_\mu^a L^a + ig W_\mu^b T^b + ig' B_\mu Y,$$

where G_μ^a are the eight gluon fields, W_μ^b the three weak interaction bosons, B_μ the single hypercharge boson, L^a the $SU(3)_C$ generators ($\frac{1}{2}\lambda^a$ for triplets, 0 for singlets), T^b the $SU(2)_L$ generators ($\frac{1}{2}\tau^b$ for doublets, 0 for singlets), and Y the $U(1)_Y$ charge. The τ^b are the three 2×2 Pauli matrices and the λ^b are the eight 3×3 Gell-Mann matrices

2.3 ISSUES

Although the picture of the Standard Model as described sounds rather compact and elegant, it is actually more of a collection of theories, most notably quantum electrodynamics (QED), the Glashow-Weinberg-Salam electroweak theory, quantum chromodynamics (QCD), and the Higgs mechanism. While it has been largely successful, there still exist a number of deficiencies and so theorists are continuing to develop extension or alternate theories. These deficiencies can be summarized as follows:

Hierarchy problem There is a large quantum correction to the Higgs mass mostly from the Feynman diagram of the Higgs with a top quark loop. However, this is at odds with the theoretical predictions and experimental results for the mass of the Higgs. One possible solution comes from Supersymmetry, an extension of the Standard Model, which posits an additional particle, called a superpartner, for each Standard Model particle. The superpartner of the top quark, called the top squark, would contribute an analogous diagram to the top quark loop diagram to cancel its large contribution. However, such a cancelation of large numbers leaving a specific small number requires incredible fine-tuning of the large numbers. Such fine-tunings in nature are seen as implausible and possibly signs of new physics.

Neutrino mass & neutrino oscillations Neutrinos are massless in the Standard Model. However, neutrino oscillation experiments, where neutrinos of one lepton flavor are measured to change into another lepton flavor, have shown neutrinos have a small, non-zero mass. Neutrino masses and oscillations can be added into the Standard Model analogously to the quarks, but the exact nature of the neutrinos is still being investigated.

Dark matter & dark energy Cosmological experiments have shown that the total amount of matter and energy in the universe is much greater than that described by the Standard Model. In fact, the Standard Model contribution is only 4% of the total. Experiments are currently trying to discover the dark matter, and some extensions of the Standard Model offer candidate particles for the dark matter.

Matter/anti-matter asymmetry The Standard Model predicts almost an even amount of matter and anti-matter. However, the universe is made up of mostly matter. The Standard Model offers a mechanism to obtain this asymmetry through CP violation, but the observed amount is not enough. Some theorists propose alternate models such as baryogenesis and leptogenesis to explain the asymmetry.

Strong CP problem The Standard Model allows for CP violating QCD terms in the Lagrangian. However, the couplings for these terms have been measured to be very small. The reason this would be the case is unknown.

Unification As described before, the gauge group of the Standard Model is $SU(3)_C \times SU(2)_L \times U(1)_Y$, corresponding to the strong and electroweak forces. It turns out that the coupling constants for these three gauge groups vary with energy, due to renormalization, and around 10^{19} GeV, they become approximately equal. Grand Unified Theories such as $SU(5)$ and $SO(10)$ attempt to unify these gauge groups above this energy into just one. A further unification goal is to combine general relativity with quantum field theory, yielding a Theory of Everything. String theory is one such candidate.

2.4 JETS

Although not primitive objects themselves, jets are typically treated as such in physics analyses, so it will be helpful to understand what they are and how they are formed. Due to color confinement in QCD, when partons are produced in an event, they must quickly form colorless combinations of hadrons in a process called hadronization. The fundamental physics of hadronization are not well understood, so various models are employed, such as the popular Lund string model. In this model, colored particles are connected by a "string" which connects particles into color neutral combinations and, acting like a spring, has a potential $V(r) = \kappa r$ where $\kappa \approx 1\,\text{GeV/fm}$. When the partons move far enough apart, a $q\bar{q}$ pair can be popped out of the vacuum to break the string into two fragments to lower the total energy. This process is illustrated in Fig. 2.3. When repeated iteratively, a shower of particles called a jet is produced.
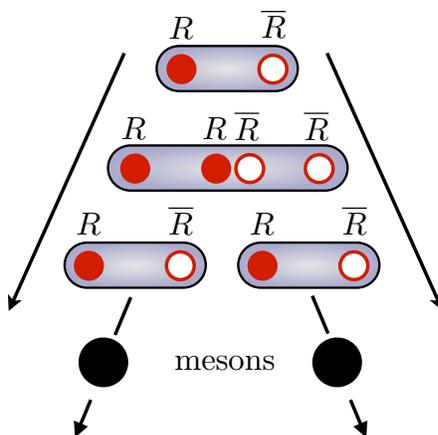
Figure 2.3: Illustration of the Lund string fragmentation process. Here R stands for red and $\overline{R}$ for anti-red. The fragmentation process is repeated iteratively, producing a shower of hadrons. Baryon production is not well understood but can be modeled by popping a diquark anti-diquark pair instead of the quark anti-quark [7].

2.5 W & Z BOSONS

The W and Z bosons mediate the electroweak force along with the photon. The main difference is that the Higgs mechanism gives mass to the W and Z bosons while the photon is massless. However, the W, and the Z especially (being neutral), can be thought of as sort-of massive photons. While the photon couples to a lepton/anti-lepton pair, the Z boson may also couple to quark/anti-quark and neutrino/anti-neutrino pairs. The Feynman diagram representing this is shown in Fig. 2.4, and the probabilities for the Z decaying to each type are listed in Tab. 2.1.
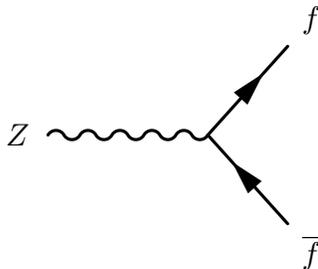
Figure 2.4: Feynman diagram of the coupling of the Z boson to fermions. Here $f = \{\ell, q, \nu\}$ (leptons, quarks, neutrinos). The Z, being electrically neutral, couples to particle/anti-particle pairs to conserve charge.

The W also couples to leptons, quarks, and neutrinos, but whereas the Z coupled to particle/anti-particle pair to preserve neutral electric charge, the $W^{\pm}$ couples to different particle types to conserve its $+1$ or -1 charge. The Feynman diagrams for the W couplings are shown in Fig. 2.5 and the branching fractions are listed in Tab. 2.2. The W coupling to quarks requires some explanation. Typically the W will decay to a quark and anti-quark in the same generation; *e.g.*, $W^+ \to u\bar{d}$ and $W^+ \to c\bar{s}$. Note that in the W's rest-frame, decaying to $t\bar{b}$ is not allowed due to energy conservation ($m_W = 80\,\mathrm{GeV}/c^2$). However, it is also possible for the W to decay to quarks in *different* generations. This is known as quark mixing in the weak interaction and is embodied in the Standard Model

Z decay mode	Branching fraction (%)	
$\ell^+\ell^-$	10	
$\quad e^+e^-$		3.4
$\quad \mu^+\mu^-$		3.4
$\quad \tau^+\tau^-$		3.4
invisible ($\nu\bar{\nu}$)	20	
hadrons	70	
$\quad (u\bar{u} + c\bar{c})/2$		12
$\quad (d\bar{d} + s\bar{s} + b\bar{b})/3$		15

Table 2.1: Decay modes of the Z boson in its rest-frame. "Invisible" is assumed to be $\nu\bar{\nu}$. For decay to hadrons, the breakdown for up-type and down-type averages are listed. Note that $Z \to t\bar{t}$ is not allowed due to energy conservation since the mass of the Z is $m_Z = 91\,\mathrm{GeV}/c^2$ while for the top quark, $m_t = 173\,\mathrm{GeV}/c^2$.

via the celebrated Cabibbo–Kobayashi–Maskawa (CKM) matrix. The mechanism for CP violation is also contained in the CKM matrix, but for our purposes it suffices to know that the coupling of the W to quarks is proportional to $|V_{qq'}|^2$ where

$$
V = \begin{bmatrix} |V_{ud}| & |V_{us}| & |V_{ub}| \\ |V_{cd}| & |V_{cs}| & |V_{cb}| \\ |V_{td}| & |V_{ts}| & |V_{tb}| \end{bmatrix} \approx \begin{bmatrix} 0.97 & 0.23 & 0.003 \\ 0.23 & 0.97 & 0.04 \\ 0.009 & 0.04 & 1.00 \end{bmatrix}.
$$

Now then, we have three semi-leptonic decays and two dominant hadronic decays for the W. Counting each hadronic decay three times (for the three colors), we have a total of nine decays, and it is worth pointing out in Tab. 2.2 that indeed each measured semi-leptonic mode branching fraction is approximately $1/9 \approx 11\%$ and the measured hadronic branching fraction for $c\bar{s}$ is indeed $3 \times 1/9 \approx 33\%$, where $u\bar{d}$ (hard to measure) should also be about 33%.

Finally, there is one more coupling of the W and Z bosons that we will need. This
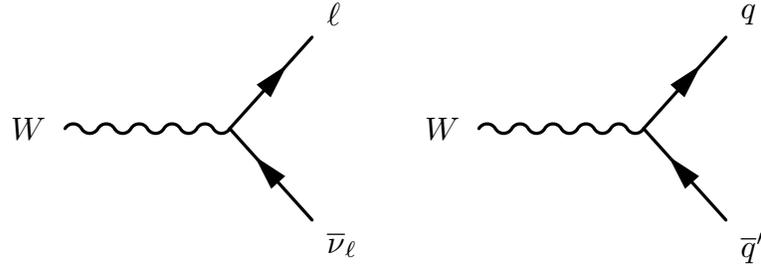
Figure 2.5: Feynman diagram of the coupling of the W boson to fermions. Left: the "semi-leptonic" decay $W \to \ell \bar{\nu}$. Right: the hadronic decay $W \to q\bar{q}'$.

W^+ decay mode	Branching fraction (%)
$\ell^+ \nu_\ell$	32
$e^+ \nu_e$	10.8
$\mu^+ \nu_\mu$	10.6
$\tau^+ \nu_\tau$	11.3
hadrons	68
$c\bar{s}$	31^{+13}_{-11}

Table 2.2: Decay modes of the W boson in its rest-frame.

is the coupling of a Z and a W^+ and W^-. The Feynman diagram for this is shown in Fig. 2.6.
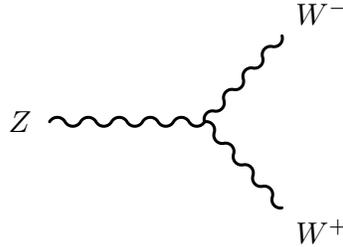
Figure 2.6: Feynman diagram for the coupling of a Z and two W's. Note that this decay would not happen in the rest-frame of the Z due to energy conservation, but can happen in a larger diagram where at least one of the bosons is "virtual" (existing for only a short time).

2.6   DIBOSONS

We now have all the pieces needed to understand how diboson events are produced. By putting together the Feynman rules, we obtain Fig. 2.7. As an aside, note that the Tevatron is a $p\bar{p}$ collider, and since a proton is made up of uud and an anti-proton of $\bar{u}\bar{u}\bar{d}$, the necessary $q\bar{q}$ in these diagrams already exist in the proton and anti-proton collisions. This is in contrast to the LHC which is a pp collider. In that case where does the anti-quark come from? Deep inelastic scattering experiments which probe the structure of nucleons showed us that in addition to the uud valance quarks in the proton, there are also gluons holding the quarks together, along with virtual $q\bar{q}$ pairs generated through these gluons. The probability of observing a given type of parton in a nucleon as a function of its momentum is known as the parton distribution function (PDF).

Next, let us calculate the various branching ratios for the diboson decays using Tables 2.1 and 2.2.
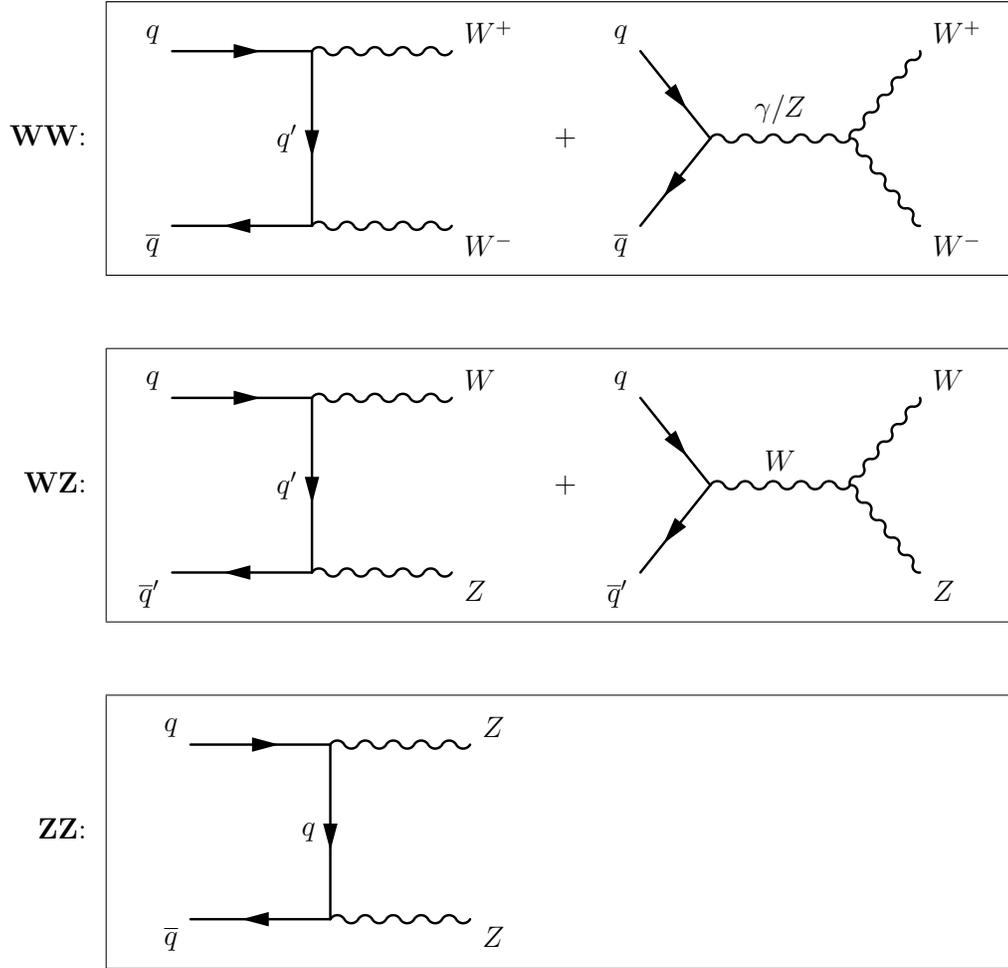
Figure 2.7: Tree-level Feynman diagrams for WW, WZ, and ZZ diboson production. Note there is only the one diagram for ZZ since there do not exist ZZZ or $ZZ\gamma$ couplings in the Standard Model. Indeed, an excess of ZZ events in data could be evidence for such anomalous triple gauge couplings.

For example, $\mathrm{Br}(W^-W^+ \to e^-\bar{\nu}e^+\nu) = 11\% \times 11\%$, and it is the same for muons. However, since taus are rather difficult to detect due to their short lifetime, most analyses—including this one—do not include taus. Thus for the present purposes of comparing diboson branching ratios in the channels we do and do not measure in this analysis, we will ignore taus from the lepton category. Hence when we write $\ell\ell$ we mean ee, $\mu\mu$, $e\mu$, and μe. Thus $\mathrm{Br}(WW \to \ell\bar{\nu}\bar{\ell}\nu) = 4(11\% \times 11\%) = 4.8\%$. As another example, $\mathrm{Br}(WW \to \ell\bar{\nu}jj) = 2(2 \times 11\% \times 68\%) = 30\%$ where one factor of two comes from not caring about from which W the products come from, and the other is for counting electrons and muons. Carrying out this process in this way for all combinations, we obtain Tab. 2.3. We see that the highest branching ratios for all the dibosons are the fully hadronic modes. However, this signal would be swamped by the QCD multi-jet background. Generally, adding a lepton makes the decay more rare, but a more clean (less background) channel.

2.7   HIGGS BOSON

Finally, we will discuss the Higgs boson due to its importance and relevance to this analysis. We will first discuss the Higgs mechanism and electroweak symmetry breaking, followed by Higgs production, decay, search channels, and results.

2.7.1   HIGGS MECHANISM AND ELECTROWEAK SYMMETRY BREAKING

We start with a general description of electroweak symmetry breaking motivated by the discussion in [9]. We will then clarify this picture mathematically. It is important to

	WW	WZ	ZZ
Cross section (pb)	12.4	3.7	1.4
$\mathrm{Br}(VV \to \nu\nu jj)$	—	14%	28%
$\mathrm{Br}(VV \to \ell\nu jj)$	30%	15%	—
$\mathrm{Br}(VV \to \nu\nu bb)$	—	—	6%
$\mathrm{Br}(VV \to \ell\nu bb)$	—	3.3%	—
$\mathrm{Br}(VV \to \ell\nu\nu\nu)$	—	4.4%	—
$\mathrm{Br}(VV \to \ell\ell\nu\nu)$	4.8%	—	2.8%
$\mathrm{Br}(VV \to 3\ell\nu)$	—	1.5%	—
$\mathrm{Br}(VV \to 3\ell j)$	—	—	—
$\mathrm{Br}(VV \to 4\ell)$	—	—	0.5%
$\mathrm{Br}(VV \to 4j)$	46%	48%	49%

Table 2.3: Diboson branching ratios for various decay modes. Note that V is shorthand for W or Z, and ℓ stands for e or μ. A "—" indicates a forbidden decay. The total cross sections [8] are also listed for comparison purposes. Our analysis is sensitive to the first two channels ($\nu\nu jj$ and $\ell\nu jj$) in the no-tag region and the second two channels ($\nu\nu bb$ and $\ell\nu bb$) in the two-tag region. Note that WW does not enter in the $\nu\nu bb$ and $\ell\nu bb$ decay modes.

point out that, since the Higgs has not been discovered, this theory has not yet been experimentally verified.

The Higgs mechanism is the process by which the $SU(2) \times U(1)$ symmetry corresponding to the electroweak force is broken into just the electromagnetic $U(1)$ symmetry, while simultaneously giving rise to the mass of the W and Z weak vector bosons. The Higgs field is special in that it has a non-zero vacuum expectation value (VEV), which allows other particles to interact with the Higgs field to gain mass. The Higgs field potential can be visualized as in Fig. 2.8. Since the origin is an unstable equilibrium, a state there will fall into the lower energy state, breaking the rotational symmetry. Since anywhere along the bottom of the well corresponds to the same energy, there is an added

degree of freedom and this corresponds to a massless Nambu-Goldstone boson. There are actually two more dimensions like this, since the Higgs field ϕ carries four degrees of freedom, for a total of three Goldstone bosons. These Goldstone bosons are massless due to the potential being flat in the excitation direction. In contrast, the radial degree of freedom has curvature and corresponds to the usual massive Higgs boson. In the unbroken electroweak theory, the W, Z, and photon do not exist in their usual form, and only assume it when the electroweak symmetry is broken. The W_1, W_2, W_3 and B bosons of the electroweak theory "eat" the Goldstone bosons (H^+, H^-, H^0) to obtain charge and mass to become the usual W^+, W^-, Z^0 bosons and the photon.
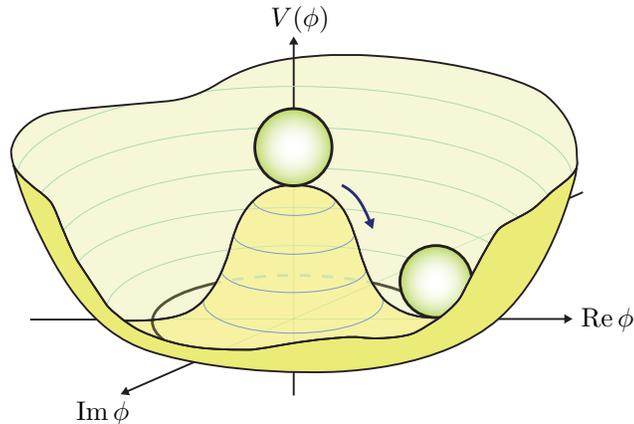
Figure 2.8: A Mexican hat potential for the scalar Higgs field demonstrating electroweak symmetry breaking. Figure courtesy of Flip Tanedo.

SPONTANEOUS BREAKING OF A LOCAL $U(1)$ GAUGE SYMMETRY

To see how this works mathematically, let us start with an example of symmetry breaking generating mass in a simplified theory, and then proceed by analogy to the full Higgs mechanism in the Standard Model. In particle physics, it is often most useful to work directly with the Lagrangian of a system. The Lagrangian summarizes the dynamics of the system, and the integral of the Lagrangian (over 4d space-time) gives the action,

which by the principle of least action yields the Euler-Lagrange equations of motion for the system. For example, the Lagrangian

$$\mathcal{L} \equiv T - V = \frac{1}{2}(\partial_\mu \phi)(\partial^\mu \phi) - \frac{1}{2}m^2 \phi^2 \tag{2.2}$$

for a real scalar field $\phi(x)$, when put into the Euler-Lagrange equation

$$\partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \right) - \frac{\partial \mathcal{L}}{\partial \phi} = 0,$$

yields the Klein-Gordon equation

$$(\partial_\mu \partial^\mu + m^2)\phi = 0.$$

Following the discussion in [10], we consider a Lagrangian describing a complex scalar field

$$\mathcal{L} = (\partial_\mu \phi)^*(\partial^\mu \phi) - \mu^2 \phi^* \phi - \lambda(\phi^* \phi)^2, \quad \phi \equiv \frac{1}{\sqrt{2}}(\phi_1 + i\phi_2), \tag{2.3}$$

which is equivalent to the sum of the Lagrangians (Eq. 2.2) for the scalar fields ϕ_1 and ϕ_2. However, we utilize the complex field Lagrangian since it is invariant under a global $U(1)$ transformation $\phi \rightarrow e^{i\alpha}\phi$. Symmetries such as this are important since Noether's Theorem tells us that for each symmetry there is a corresponding conserved current. However, the full power of symmetries comes when we require *local* gauge ("phase") invariance—a fundamental ingredient in the Standard Model. Let us proceed in this fashion to make the Lagrangian invariant under the local $U(1)$ transformation

$$\phi \rightarrow e^{i\alpha(x)}\phi.$$

We see that as it stands, Eq. 2.3 is not invariant under this transformation since

$$\partial_\mu \phi \rightarrow e^{i\alpha(x)}\partial_\mu \phi + ie^{i\alpha(x)}\phi \, \partial_\mu \alpha. \tag{2.4}$$

If we replace the derivative ∂_μ in the Lagrangian by the covariant derivative

$$D_\mu \equiv \partial_\mu - ieA_\mu,$$

24

where the gauge field A_μ transforms as

$$A_\mu \rightarrow A_\mu + \frac{1}{e}\partial_\mu\alpha,$$

then we see that the second term in Eq. 2.4 is canceled so that

$$D_\mu\phi \rightarrow e^{i\alpha(x)}D_\mu\phi,$$

and hence the modified Lagrangian

$$\mathcal{L} = (D_\mu\phi)^*(D^\mu\phi) - \mu^2\phi^*\phi - \lambda(\phi^*\phi)^2 - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} \tag{2.5}$$

is invariant under local $U(1)$ transformation. Note that we have included the kinetic term $F_{\mu\nu}F^{\mu\nu}$ for the gauge field where

$$F_{\mu\nu} \equiv \partial_\mu A_\nu - \partial_\nu A_\mu$$

is the electromagnetic tensor.

If $\mu^2 > 0$, then Eq. 2.5 is just the QED Lagrangian for a charged scalar particle of mass μ, with a self-interaction term and a massless photon A_μ. However, we wish to consider the case of $\mu^2 < 0$ and $\lambda > 0$ so that the potential is shaped as in Fig. 2.8:

$$V = \frac{1}{2}\mu^2(\phi_1^2 + \phi_2^2) + \frac{1}{4}\lambda(\phi_1^2 + \phi_2^2)^2 + \text{const.}$$

The minima of the potential occur at

$$0 \equiv \frac{\partial V}{\partial\phi_i} \implies \phi_1^2 + \phi_2^2 = v^2 \quad \text{where} \quad v \equiv \sqrt{\frac{-\mu^2}{\lambda}}.$$

Next we perturbatively expand the field around one of the minima. By doing this, we spontaneously break the $U(1)$ symmetry. We are free to arbitrarily choose which minimum, since nature must also, so we pick $\phi_1 = v, \phi_2 = 0$, and substitute

$$\phi(x) = \frac{1}{\sqrt{2}}\left[v + \eta(x) + i\xi(x)\right]$$

25

into Eq. 2.5 to get

$$\mathcal{L}' = \frac{1}{2}(\partial_\mu \eta)^2 + \frac{1}{2}(\partial_\mu \xi)^2 - v^2\lambda\eta^2 + \frac{1}{2}e^2v^2 A_\mu A^\mu - evA_\mu\partial^\mu\xi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}$$

$$+ \text{const.} + \text{interaction terms.} \quad (2.6)$$

We identify these terms as three particles with masses

$$m_\eta = \sqrt{2\lambda v^2} = \sqrt{-2\mu^2}, \quad m_\xi = 0, \quad m_A = ev.$$

This generation of a massive vector boson without explicitly writing the mass term in the gauge symmetric Lagrangian is precisely the Higgs mechanism. The difference between explicitly breaking the symmetry by adding the mass term and spontaneously breaking it through the perturbative expansion is important. The former produces unrenormalizable divergences[1] making the theory meaningless, while the latter does not. This is useful since we need massive $W^\pm$ and Z bosons in our theory.

SPONTANEOUS ELECTROWEAK SYMMETRY BREAKING IN THE STANDARD MODEL

Now we generalize the method of §2.7.1 to the breaking of the $SU(2) \times U(1)$ symmetry group, following the discussions in [10] and [11]. Obtaining the covariant derivative following the prescription in §2.2, the kinetic term for the Higgs field in the Lagrangian is

$$|D_\mu\phi|^2 = \left|\left(\partial_\mu + \frac{i}{2}gW_\mu^b\tau^b + \frac{i}{2}g'B_\mu\right)\phi\right|^2, \quad (2.7)$$

where $|\ |^2 \equiv (\ )^\dagger(\ )$. The Higgs field is written to contain four scalar fields in an isospin doublet:

$$\phi \equiv \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} \equiv \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix}.$$

---

[1]These divergences arise in loops when integrating over momenta $\int \mathrm{d}^4q\,(\cdots)$. As $q \to \infty$, the propagator for the massless photon goes as $1/q^2$ and makes the integral well-behaved while the massive propagator goes as $q_\mu q_\nu/q^2$ and the integral diverges.

We use the same potential $V(\phi) = \mu^2 |\phi|^2 + \lambda |\phi|^4$ as before, so the minimum is at $|\phi|^2 = v^2$ and we are free to break the symmetry and choose to perturbatively expand around the ground state of

$$\phi_0 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix}.$$

Then using Eq. 2.7,

$$\begin{aligned}
\mathcal{L}_{\text{higgs}} &= |D_\mu \phi|^2 - V(\phi) \\
&= \left| \left(\partial_\mu + \frac{i}{2} g W_\mu^b \tau^b + \frac{i}{2} g' B_\mu \right) \phi \right|^2 - V(\phi) \\
&= \frac{1}{8} \left| \begin{pmatrix} -2i\partial_\mu + g W_\mu^3 + g' B_\mu & g(W_\mu^1 - iW_\mu^2) \\ g(W_\mu^1 + iW_\mu^2) & -2i\partial_\mu - g W_\mu^3 + g' B_\mu \end{pmatrix} \begin{pmatrix} 0 \\ v + h \end{pmatrix} \right|^2 \\
&\quad - \frac{1}{2} \mu^2 (v + h)^2 - \frac{1}{4} \lambda (v + h)^4 \\
&= \frac{1}{8} v^2 g^2 |W_\mu^1 - iW_\mu^2|^2 \left(1 + \frac{h}{v} \right)^2 + \frac{1}{8} v^2 (g' B_\mu - g W_\mu^3)^2 \left(1 + \frac{h}{v} \right)^2 + \frac{1}{2} (\partial_\mu h)^2 \\
&\quad - \lambda v^2 h^2 - \lambda v h^3 - \frac{1}{4} \lambda h^4 + \text{const.}
\end{aligned}$$

Next we define[2] the mass eigenstate fields

$$W_\mu^\pm = \frac{1}{\sqrt{2}} (W_\mu^1 \mp iW_\mu^2)$$

$$Z_\mu = \frac{1}{\sqrt{g^2 + g'^2}} (g W_\mu^3 - g' B_\mu)$$

$$A_\mu = \frac{1}{\sqrt{g^2 + g'^2}} (g' W_\mu^3 + g B_\mu)$$

so that we have

$$\begin{aligned}
\mathcal{L}_{\text{higgs}} &= \left(m_W^2 W_\mu^- W^{+\mu} + \frac{m_Z^2}{2} Z_\mu Z^\mu \right) \left(1 + \frac{h}{v} \right)^2 + \frac{1}{2} (\partial_\mu h)^2 \\
&\quad - \frac{m_H^2}{2} h^2 - \frac{g_{HHH}}{3!} h^3 - \frac{g_{HHHH}}{4!} h^4 + \text{const.}, \tag{2.8}
\end{aligned}$$

---

[2]These can be obtained by diagonalizing the mass matrix.

27

where the masses and couplings are read off[3] as

$$m_W = \frac{1}{2}gv, \quad m_Z = \frac{1}{2}v\sqrt{g^2 + g'^2}, \quad m_H = \sqrt{2\lambda}v, \quad m_A = 0,$$

$$g_{HHH} \equiv 6\lambda v = \frac{3m_H^2}{v}, \quad g_{HHHH} \equiv 6\lambda = \frac{3m_H^2}{v^2}.$$

From Eq. 2.8 we can also read off the couplings

$$g_{HVV} = \frac{2m_V^2}{v}, \quad g_{HHVV} = \frac{2m_V^2}{v^2}.$$

The Feynman diagrams corresponding to these interactions are shown in Fig. 2.9.

FERMION MASSES AND COUPLINGS

Lastly, we will derive the fermion masses and their couplings to the Higgs. The first thing is to note that the Standard Model is a chiral theory and the weak force operates on left- and right-handed particles differently. The left-handed fields transform as $SU(2) \times U(1)$ and the right-handed fields as $U(1)$:

$$\chi_L \to e^{i\alpha_\mu(x)T^\mu} e^{i\beta(x)Y} \chi_L,$$

$$\psi_R \to e^{i\beta(x)Y} \psi_R.$$

Under these transformations, fermion mass terms are not invariant. This is because, for example, (using projection operators) $\bar{e}e = \bar{e}_R e_L + \bar{e}_L e_R$ which is not invariant. Fortunately the Higgs mechanism generates masses for the fermions through the symmetry

---

[3]The factors of 1/2, 1/3! and 1/4! come from the combinatorics of swapping identical particles.

28

breaking in the Yukawa terms in the Lagrangian. Consider the Yukawa term for the first generation leptons where the Higgs field gets a VEV as before:

$$\mathcal{L}_{\text{Yukawa},e} = -\lambda_e \bar{L}_L \phi E_R + \text{h.c.}$$

$$= -\frac{\lambda_e}{\sqrt{2}} \begin{pmatrix} \bar{\nu}_{e,L} & \bar{e}_L \end{pmatrix} \begin{pmatrix} 0 \\ v + h \end{pmatrix} e_R + \text{h.c.}$$

$$= -\frac{\lambda_e v}{\sqrt{2}} (\bar{e}_L e_R + \bar{e}_R e_L) \left(1 + \frac{h}{v}\right)$$

$$= -m_e \bar{e}e - g_{Hee} \bar{e}eh,$$

where "h.c." stands for the Hermitian conjugate (†) of the entire line and, since we have the same terms for the other lepton generations and analogously for the quarks, we write

$$m_f = \frac{\lambda_f}{\sqrt{2}} v, \quad g_{Hff} = \frac{m_f}{v}.$$

This completes Fig. 2.9. Note that here neutrinos remain massless and do not couple to the Higgs since there is no right-handed neutrino in the Standard Model.

2.7.2 PRODUCTION

From Fig 2.9, based on the mass dependence of the coupling constants, along with the allowed kinematical phase-space and factors of $1/v$, we can determine that the dominant processes for Higgs production will involve the Higgs coupling to V bosons and/or the third generation of quarks and leptons. These processes are shown in Fig. 2.10.

The cross sections for Higgs production from $p\bar{p}$ (Tevatron) and pp (LHC) as a function of the Higgs mass are shown in Fig. 2.11. The dominant production mechanisms are (in roughly descending order) $gg \to H$, $q\bar{q} \to WH/ZH$, $q\bar{q} \to q\bar{q}H$, $b\bar{b} \to H$, and $gg/q\bar{q} \to t\bar{t}H$.

Coupling	Diagram(s)
$g_{Hf\bar{f}} = \dfrac{m_f}{v}$	
$g_{HVV} = \dfrac{2m_V^2}{v}$	
$g_{HHVV} = \dfrac{2m_V^2}{v^2}$	
$g_{HHH} = \dfrac{3m_H^2}{v}$	
$g_{HHHH} = \dfrac{3m_H^2}{v^2}$	

Figure 2.9: Feynman diagrams and the corresponding coupling constants for the Standard Model Higgs boson, where $V = W^\pm$ or Z, and $v \approx 246\,\mathrm{GeV}$ is the vacuum expectation value of the Higgs field. Note that the first two diagrams correspond to the $Zf\bar{f}$ and ZW^+W^- couplings if the H were replaced with a Z, but this cannot be done for the third HZZ diagram since there is no Standard Model ZZZ coupling. Furthermore, there is the $HHVV$, and triple and quadruple self-couplings.

2.7.3 BRANCHING RATIOS

To determine which channels are important for Higgs searches, we first need to study the branching ratios for the decay of the Higgs. These are shown in Fig. 2.13 as a function of the Higgs mass.

We see that for $m_H < 135 \, \text{GeV}/c^2$, $H \to b\bar{b}$ dominates. Note that although the Higgs to fermions coupling is proportional to m_f and thus would be larger for t than b, the decay $H \to t\bar{t}$ is kinematically inaccessible for $m_H < 2m_t \approx 340 \, \text{GeV}/c^2$. Similar reasoning explains the behavior of $H \to WW/ZZ$: for $m_H < 2m_W \approx 161 \, \text{GeV}/c^2$, at least one of the W bosons must be virtual and thus suppressed by the phase space factor, and similarly for $H \to ZZ$ for $m_H < 2m_Z \approx 180 \, \text{GeV}/c^2$. These Feynman diagrams are shown in Fig. 2.14. Thus, in the region $2m_W < m_H < 2m_Z$, $H \to WW$ dominates since it becomes a 2-body decay while in $H \to ZZ$ at least one of the Z's is virtual. This effect
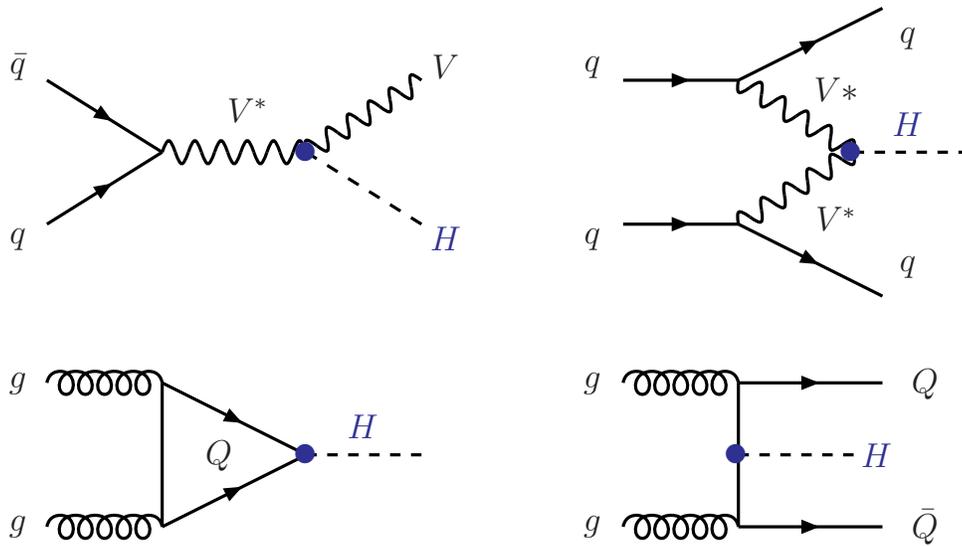
Figure 2.10: Dominant processes for Higgs production at hadron colliders: associated production with a W or Z (top-left), vector boson fusion (top-right), gluon-gluon fusion (bottom-left), and associated production with heavy quarks (bottom-right). Here Q standard for heavy quarks (*i.e.*, t or b). Figure from [12].
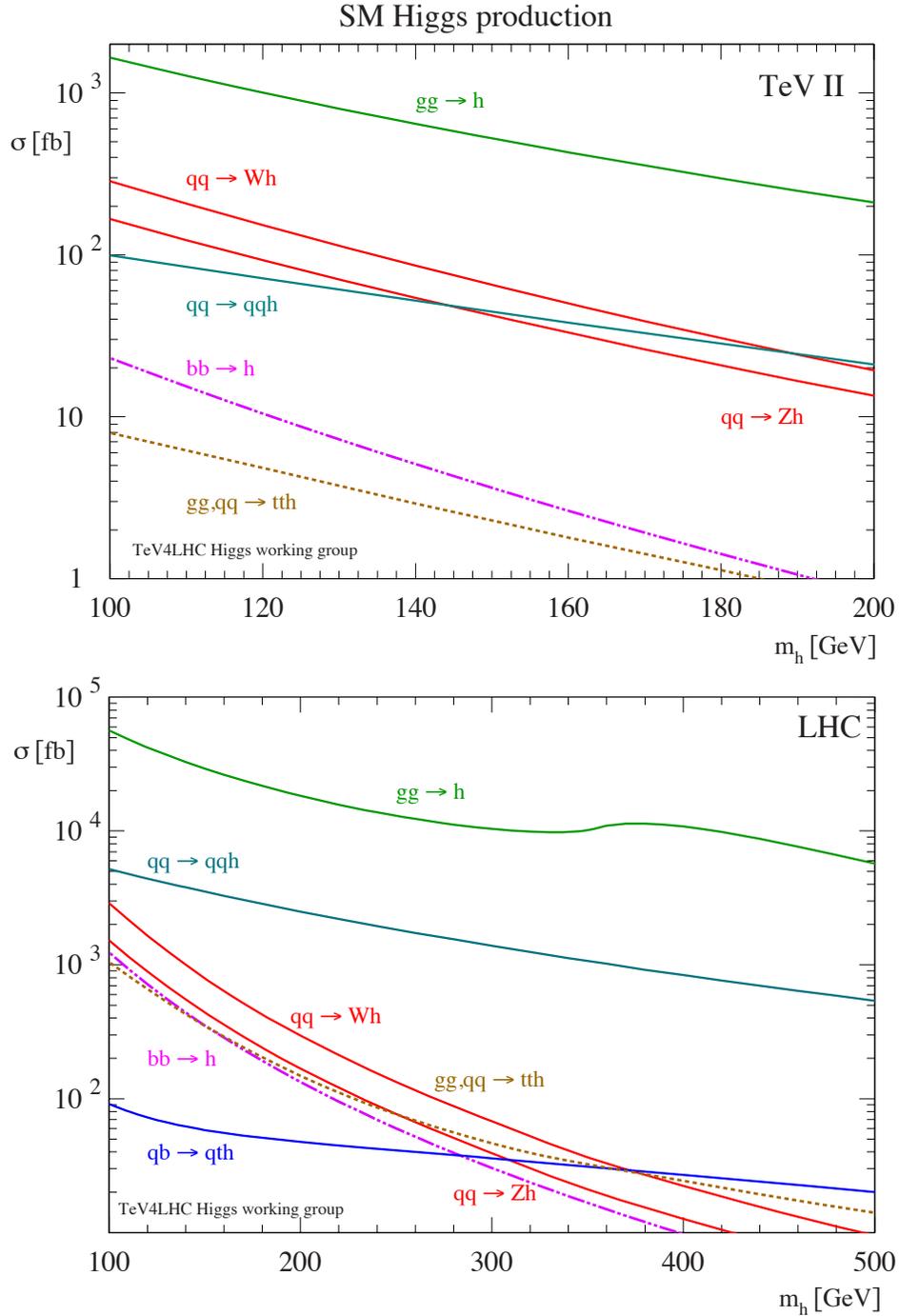
Figure 2.11: Cross sections for Higgs production as a function of the Higgs mass for the dominant processes at the Tevatron ($p\bar{p}$, $\sqrt{s} = 1.96\,\text{TeV}$) and the LHC ($pp$, $\sqrt{s} = 14\,\text{TeV}$). Although $gg \to H$ is a one loop diagram and thus suppressed, due to the large gluon production from QCD it becomes the largest contribution to Higgs production. Note the extra line for the LHC: $qb \to qtH$ as shown in Fig. 2.12. The maximum of this cross section at the Tevatron is 0.1 fb and thus off the plot. Figures modified from [13].
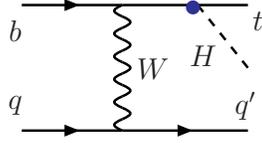
Figure 2.12: Single-top associated production $qb \to qtH$ in the t-channel. Figure from [12].
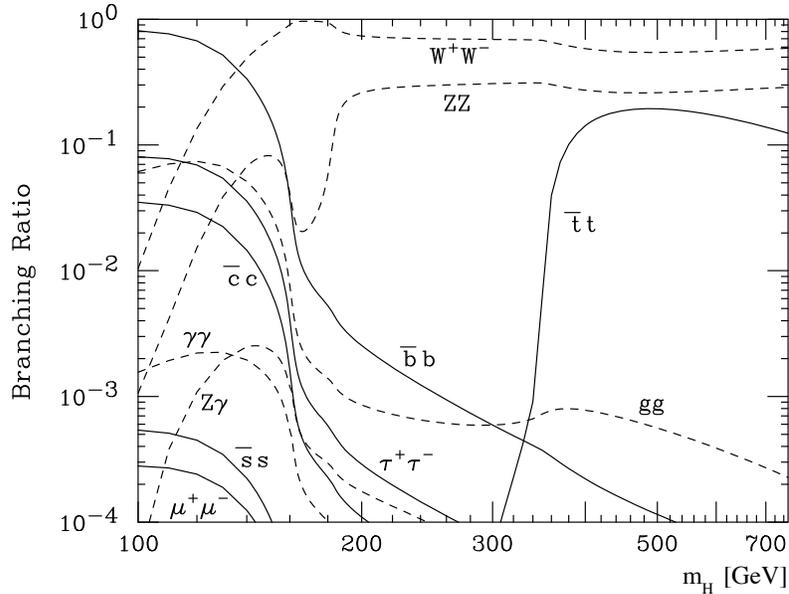
Figure 2.13: Branching ratios of the SM Higgs boson. $H \to b\bar{b}$ is dominant for $m_H < 135\,\text{GeV}/c^2$, and $H \to WW$ for $m_H > 135\,\text{GeV}/c^2$. Figure from [14].

is shown in Fig. 2.15. For Higgs masses above $2m_Z$, the WW and ZZ branching ratios level out, with $\text{Br}(H \to WW) \approx 2/3$ and $\text{Br}(H \to ZZ) \approx 1/3$.

Less dominant decays (in roughly descending order) are gg, $\tau^+\tau^-$, $c\bar{c}$, $\gamma\gamma$, $Z\gamma$, $s\bar{s}$, and $\mu^+\mu^-$. The gg, $Z\gamma$, and $\gamma\gamma$ are through loops (since g and γ are massless and hence do not couple directly to the Higgs) and shown in Fig. 2.16.
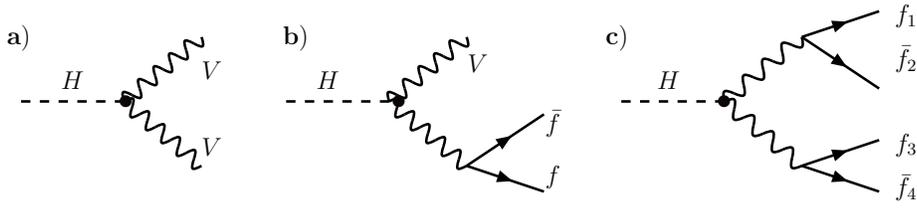
Figure 2.14: $H \to VV$ for (a) 2-body, (b) 3-body, (c) 4-body decays. When $m_H < 2m_V$, at least one of the V's is virtual. Figure from [12].
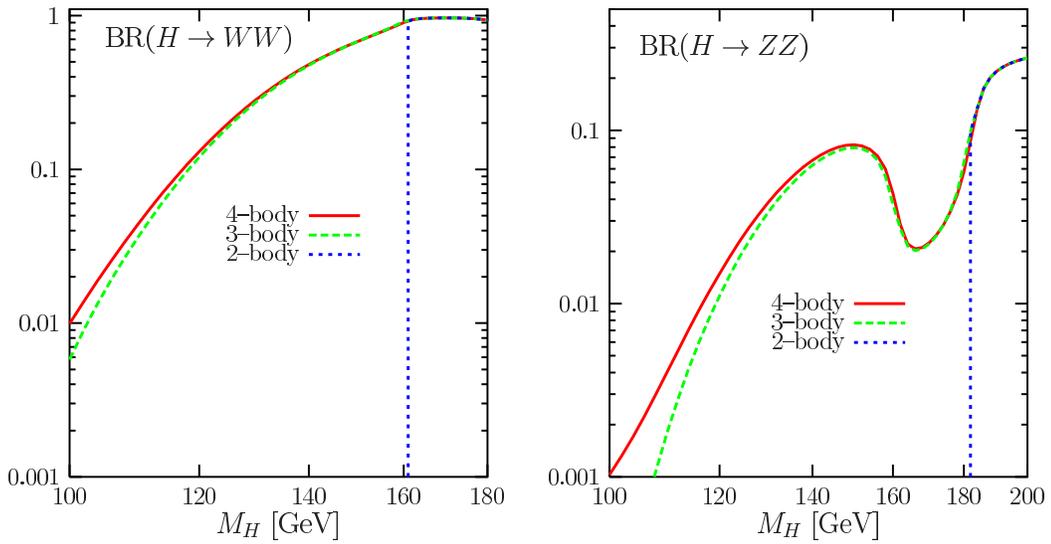
Figure 2.15: Branching ratios of the SM Higgs boson to WW (left) and ZZ (right) for 2-,3-, and 4-body decays. The dip in $\mathrm{Br}(H \to ZZ)$ around $160\,\mathrm{GeV}/c^2$ is explained by $\mathrm{Br}(H \to WW)$ receiving contribution from the 2-body decay. Figure from [12].
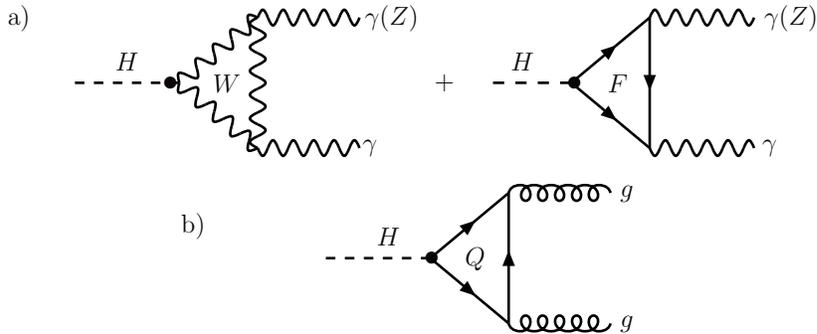
Figure 2.16: Higgs decay to (a) $\gamma\gamma/Z\gamma$ and (b) gg. Figure from [12].

2.7.4 DISCOVERY CHANNELS

For a particular decay channel to be useful, there are many practical issues that must be considered. This is where the detector's capabilities (*e.g.*, tracking resolution) and parameters (*e.g.*, luminosity and center-of-mass energy) come into play. Furthermore, there are issues related to actually performing the analysis such as backgrounds, uncertainties, statistics, etc. Here we summarize the main channels and their specific issues and relevance at the Tevatron and the LHC. Much of this section is based on [12].

Gluon-gluon fusion The largest Higgs production cross section at both the Tevatron and LHC is from $gg \to H$. However, the dominant hadronic Higgs decays do not yield a unique signature due to the large irreducible multi-jet (QCD) background. One thus has to look for more unique signatures in the less dominant decays.

- $\underline{H \to \gamma\gamma}$ For $m_H \lesssim 150\,\mathrm{GeV}/c^2$ (before the branching ratio falls off quickly) the branching ratio is small, however this channel provides a clear signature in the two-photon invariant mass distribution assuming the detector has a good photon system and good mass resolution. A large collected luminosity is needed since the decay is rare.

- $\underline{H \to ZZ^{(*)} \to 4\ell}$ For $m_H > 2m_Z$, this channel is dubbed "golden" since it is very clean. The main background is from ZZ but this is well-modeled. A large collected luminosity is also important, especially for heavy Higgs ($m_H \gtrsim 600\,\mathrm{GeV}/c^2$) where the width becomes large.

- $\underline{H \to WW^{(*)} \to \ell\ell\nu\nu}$ This is a very promising channel in the $2m_W < m_H < 2m_Z$ range where the branching ratio is close to one. These analyses exploit the fact that since the Higgs is spin zero, the W's will have opposite helicities,

and due to the correlation between the direction the lepton is emitted and the spin, these two leptons tend to have a smaller $\Delta\phi$ than those from the WW background.

Associated production with a W or Z Associated W or Z decays with leptons is useful in hadronic Higgs decays for distinguishing them from the QCD background. Note that given $\mathrm{Br}(W \to \ell\nu) \approx 20\%$ while $\mathrm{Br}(Z \to \ell\ell) \approx 6\%$, and the WH production cross section is about 1.5 times larger than ZH, the number of interesting events from WH will be more than ZH. This channel is the most promising for the Tevatron for a light Higgs decaying to $b\bar{b}$.

- $\underline{H \to b\bar{b}}$ This is the dominant decay mode for $m_H \lesssim 135\,\mathrm{GeV}/c^2$. The final states are $\ell\nu b\bar{b}$, $\ell\ell b\bar{b}$, and $\nu\bar{\nu}b\bar{b}$. The main backgrounds are $V+$jets ($V = W, Z$), VV, $t\bar{t}$, and single top. Of course, good b-tagging ability is crucial. Note that this channel is not as useful at the LHC due to a much larger QCD background and the cross section for the backgrounds growing faster with center-of-mass energy than the Higgs cross section.

- $\underline{H \to WW^{(*)}}$ This channel is not viable at the Tevatron due to too few events, but is useful at the LHC. The final states studied are trilepton, like-sign dilepton, and $\ell\nu\ell\nu jj$. There is an irreducible background from triple vector boson production.

- $\underline{H \to ZZ^{(*)}}$ This channel is not viable at the Tevatron since the HV production cross section is small for $m_H > 2m_Z$, and the branching ratio is small for $m_H < 2m_Z$. It could be useful at the LHC.

- $\underline{H \to \gamma\gamma}$ Again, not viable at the Tevatron since it is too rare, but is useful at the LHC for a low mass Higgs.

Vector boson fusion This channel is not viable at the Tevatron due to a small production cross section and a small signal to background ratio. However, it is useful at the LHC, especially when combining all of the decay channels.

- $H \to \tau^+\tau^-$ This channel is useful for low mass Higgs and for studying the couplings of the Higgs to leptons.

Associated production with heavy quarks The cross section for production in association with $t\bar{t}$ is only large enough for light Higgs searches. Note that at least one of the W's from the $t \to bW$ must decay semi-leptonically in order to suppress the QCD background.

- $H \to b\bar{b}$ This is perhaps the only channel at the LHC in which the Higgs can decay to $b\bar{b}$ and be extracted from the large QCD background. However, a rather large luminosity is required.

2.7.5 Constraints on the Higgs mass

The SM Higgs boson mass can be inferred through a global fit to the all of the precision electroweak data, shown in Fig. 2.18, resulting in [15, 16]

$$m_H = 94^{+29}_{-24}\,\mathrm{GeV}/c^2, \quad m_H > 152\,\mathrm{GeV}/c^2 \text{ excluded at 95\% CL} \qquad \text{(LEP+TEV)}.$$

The upper limit increases to $171\,\mathrm{GeV}/c^2$ if the $m_H \leq 114\,\mathrm{GeV}/c^2$ lower limit is included. Note that these numbers do not make use of the direct exclusions. The Gfitter Group also perform this fit including the LEP, Tevatron, and LHC exclusions. They find [17, 18]

$$m_H = 125^{+8}_{-10}\,\mathrm{GeV}/c^2 \qquad \text{(Gfitter -- LEP+TEV+LHC)}.$$
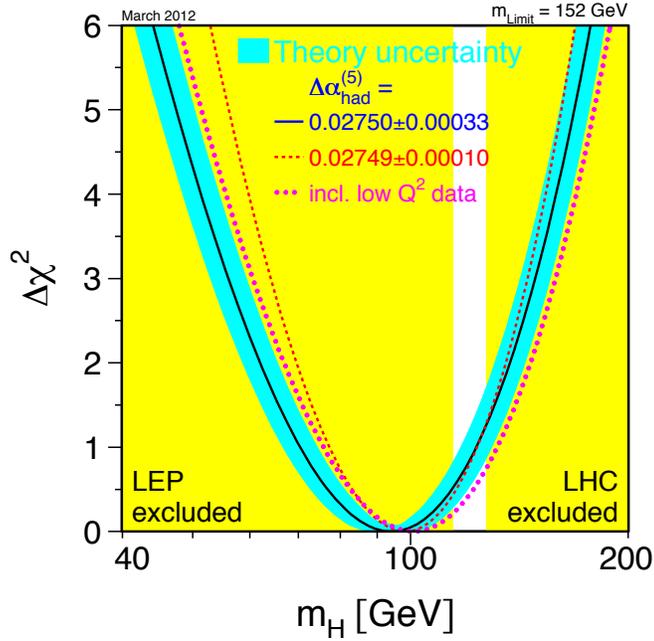
Figure 2.17: $\Delta\chi^2 = \chi^2 - \chi^2_{\min}$ versus m_H for the global fit to the precision electroweak data. The yellow regions are excluded at 95% CL from the direct LEP-II search and the Tevatron. Figure from [15].

The direct search exclusions at LEP and the Tevatron use the CL_s method [19, 20]. Here is a brief description of the method. First define a test statistic Q, which is typically taken to be the ratio of likelihoods

$$Q = \frac{P(\text{data} \,|\, \text{signal+background hypothesis})}{P(\text{data} \,|\, \text{background hypothesis})}$$

It is typical to work with

$$X \equiv -2\ln Q,$$

and define

$$CL_{s+b} = P_{s+b}(X \geq X_{\text{obs}}), \quad CL_b = P_b(X \geq X_{\text{obs}}), \quad CL_s \equiv \frac{CL_{s+b}}{CL_b},$$

where P_x is the probability distribution for the hypothesis x. These quantities are exemplified in Fig. 2.18. The signal+background hypothesis is excluded at 95% CL or more if

$CL_{s+b} < 0.05$. For discovery, $1 - CL_b < 1.3 \times 10^{-3}$ for 3σ or $1 - CL_b < 2.9 \times 10^{-7}$ for 5σ.
An important point is that in the LEP and Tevatron exclusions, CL_s is used instead of
CL_{s+b}. This is because CL_s can not be used to exclude a hypothesis for which there is
no experimental sensitivity (*e.g.*, the $-2 \ln Q$ pdfs would be greatly overlapping and then
CL_b would be small, making CL_s larger). Note that CL_s is conservative due to this.

LEP1 operated at $\sqrt{s} \approx m_Z$, so the main process was $e^+ e^- \rightarrow Z \rightarrow H Z^* \rightarrow H f \bar{f}$.
LEP2 operated at $\sqrt{s} = 209 \, \text{GeV}$, where $e^+ e^- \rightarrow HZ$, and looked in $(H \rightarrow b\bar{b}, Z \rightarrow f\bar{f})$, $(H \rightarrow b\bar{b}, Z \rightarrow \tau^+ \tau^-)$, and $(H \rightarrow \tau^+ \tau^-, Z \rightarrow b\bar{b})$. The results of all 4 LEP
experiments (ALEPH, DELPHI, L3, and OPAL) were combined and, using the CL_s
method, found [21]

$$m_H < 114.4 \, \text{GeV}/c^2 \text{ excluded at 95\% CL} \qquad \text{(LEP)},$$
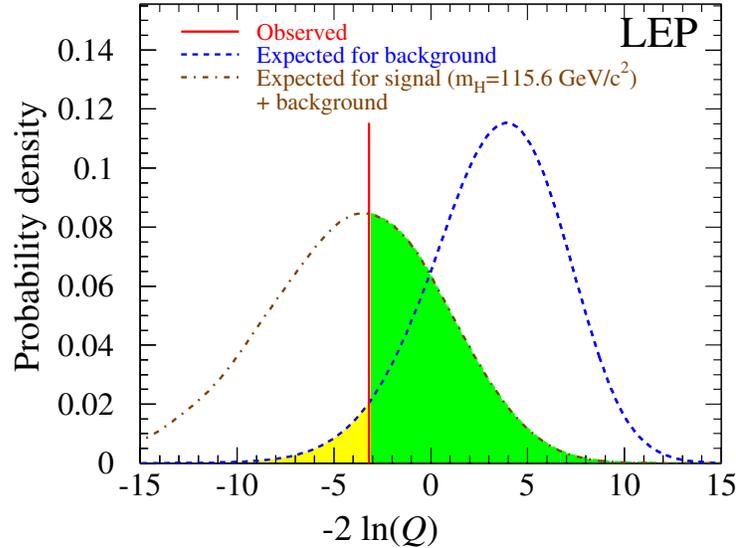
Figure 2.18: Example of the $X = -2 \ln Q$ pdfs for the background-only hypothesis (dashed blue) and signal + background hypothesis (dashed/dotted brown), and the observed value of X. The green area to the right of the observation is CL_{s+b} and the yellow area to the left of the observation is $1 - CL_b$. Figure from [20].

as shown in Fig. 2.19. Also, ALEPH saw a 3σ observation for $m_H \approx 117\,\mathrm{GeV}/c^2$, but when combined with the other 3 experiments becomes less than 2σ.
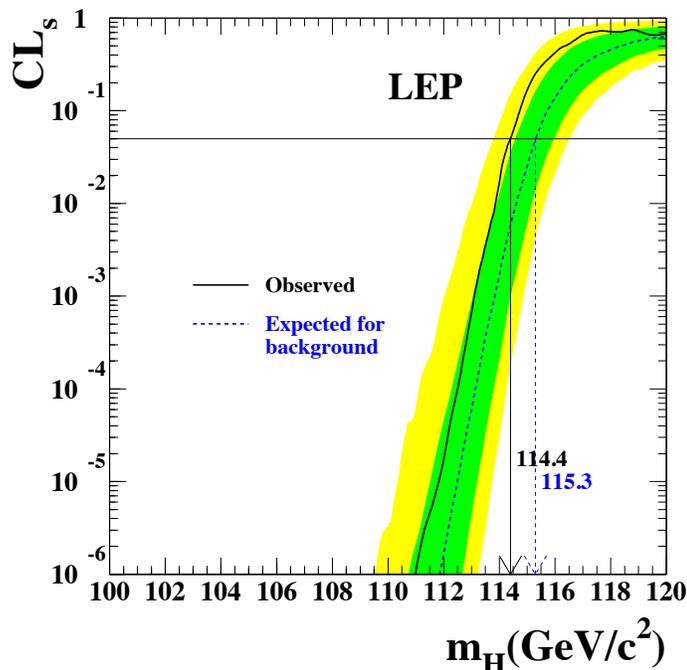
Figure 2.19: The CL_s vs. m_H for the signal+background hypothesis at LEP. The dashed line is the median background expectation, and the solid line is the measurement. Since a $CL_s < 0.05$ gives exclusion at 95% CL or greater, this sets the 95% CL lower bound on the SM Higgs mass at $114.4\,\mathrm{GeV}/c^2$. Figure from [21].

The Tevatron has also produced an exclusion based on a direct search in all of the sensitive channels, with CDF & DØ combined. They performed the analysis with the CL_s method and also with a Bayesian method. These results are shown in Figs. 2.20 and 2.21, respectively. They find [22]

$$100 < m_H < 106\,\mathrm{GeV}/c^2 \;\; \text{and} \;\; 147 < m_H < 179\,\mathrm{GeV}/c^2 \text{ excluded at 95\% CL (Tevatron).}$$

Finally, the CMS and ATLAS experiments at the LHC have also set limits at [23, 24]

$$127 < m_H < 600\,\mathrm{GeV}/c^2 \text{ excluded at 95\% CL} \quad \text{(LHC).}$$

This is also displayed on Fig. 2.21.
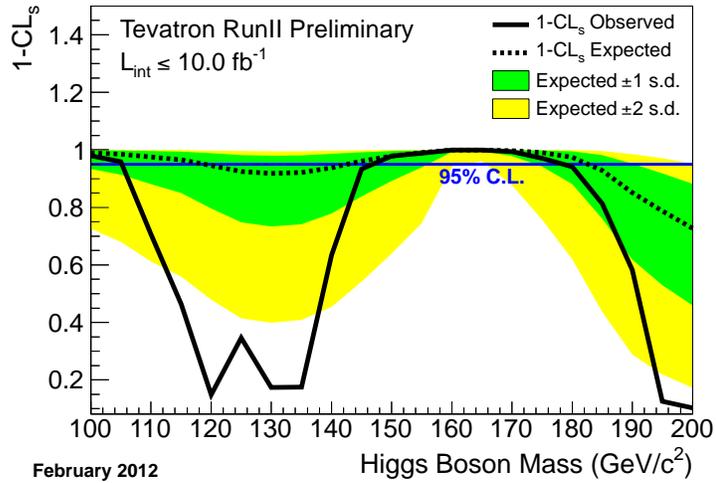
Figure 2.20: The $1 - CL_s$ vs. m_H for the combined CDF & DØ searches in all channels at the Tevatron. The SM Higgs mass is excluded at 95% CL or greater anywhere $1 - CL_s > 0.95$. Note that $H \to WW$ is the dominant decay for the excluded region of $147 < m_H < 179 \, \text{GeV}/c^2$. Figure from [22].
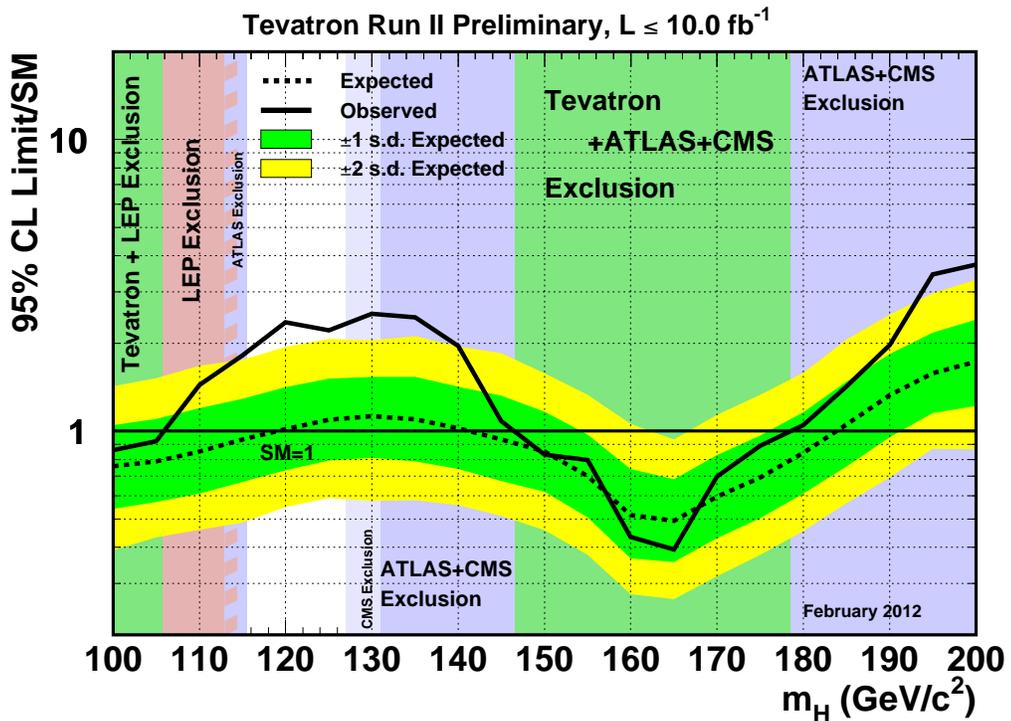
Figure 2.21: The 95% CL upper limit on the Higgs cross section divided by the Standard Model prediction. Anywhere this is less than one is excluded at 95% CL. This is the Bayesian version of Fig. 2.20. Results from the LHC have also been included. Figure from [22].

2.7.6 UPDATE

As of August 2012, the CMS and ATLAS experiments report 5σ or greater evidence of a new boson consistent with the Standard Model Higgs [25, 26]. The searches made use of the $\gamma\gamma$, ZZ, WW, $\tau^+\tau^-$, and $b\bar{b}$ decay modes and combined the results of searches using roughly $5\,\mathrm{fb}^{-1}$ of data collected at each of $\sqrt{s} = 7\,\mathrm{TeV}$ and $8\,\mathrm{TeV}$. The results of both experiments are in good agreement and they measure the mass to be about 125–126 GeV/c^2. The cross section measurements are shown in Fig. 2.22. More statistics are needed to see if all channels are consistent with the Standard Model prediction, specifically to address the question of whether or not the $H \to \gamma\gamma$ is high and the $H \to \tau\tau$ is low. If so, this would indicate new physics. Furthermore, to pin this down as the Standard Model Higgs boson, the spin should be measured. For now, since the new particle decays to two photons, we know it must have integer spin (boson) not equal to one.
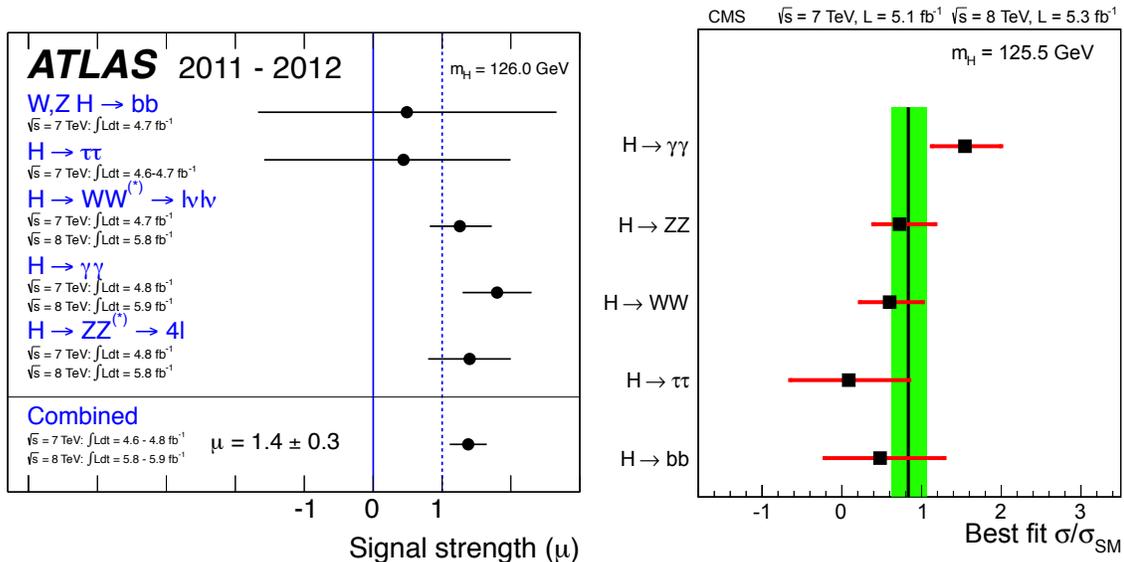
Figure 2.22: Cross section measurements for the new particle in the various Higgs decay channels, expressed as the signal strength $\mu \equiv \sigma/\sigma_{\mathrm{SM}}$ for the ATLAS experiment (left) and the CMS experiment (right). Figures from [26] and [25].

CHAPTER 3

Experimental apparatus

This thesis uses data from the CDF experiment at the Tevatron. The Tevatron is the proton anti-proton accelerator at the Fermi National Accelerator Laboratory (Fermilab), and CDF is a general purpose particle detector. Since each is an incredibly complex machine, our aim in this chapter is only to give brief overviews, with an emphasis on the capabilities of CDF most relevant to this analysis.

3.1   Fermilab & the Tevatron

Fermilab is a U.S. national laboratory in Batavia, Illinois, funded by the Department of Energy and specializing in particle physics research. It was founded in 1967 with the help of Robert Wilson who left Cornell to become the first director of the lab[1], and named in honor of the Italian physicist Enrico Fermi. The lab houses the Tevatron accelerator and its associated DØ and CDF experiments, as well as a number of neutrino experiments such as MiniBooNE, SciBooNE, and MINOS.

---

[1]Wilson successfully convinced Congress to fund the lab, justifying the $250 million dollar request with

"It only has to do with the respect with which we regard one another, the dignity of men, our love of culture. [...] It has to do with: Are we good painters, good sculptors, great poets? I mean all the things that we really venerate and honor in our country and are patriotic about.

In that sense, this new knowledge has all to do with honor and country but it has nothing to do directly with defending our country except to help make it worth defending." — Robert Wilson to US Congressional Joint Committee on Atomic Energy, 1969. [27]

The Tevatron is the main particle accelerator at Fermilab and collides protons with anti-protons at a center-of-mass energy of $\sqrt{s} = 1.96\,\text{TeV}$. Completed in 1983 with a circumference of four miles and costing \$120 million dollars, the Tevatron was the highest energy particle accelerator in the world from 1983 ($512\,\text{GeV}$) to 2010 ($1\,\text{TeV}$) when the Large Hadron Collider (LHC) at CERN surpassed it. Due to this, together with a lack of funding, the Tevatron was shut down in September, 2011.

Figure 3.1 shows an aerial view of Fermilab along with a schematic of the Tevatron. The Tevatron is made up of five sub-accelerators, each one designed to efficiently boost the beam energy in a certain energy region: the Cockcroft-Walton, Linac, Booster, Main Injector, and the Tevatron. They are summarized in Tab. 3.1

Name	Year turned on	Initial energy	Final energy	Maximum velocity (% of c)	Acceleration time (sec.)
Cockcroft-Walton	1971	0	$750\,\text{keV}$	4	1.6×10^{-7}
Linac	1971	$750\,\text{keV}$	$401\,\text{MeV}$	71	8×10^{-7}
Booster	1971	$401\,\text{MeV}$	$8\,\text{GeV}$	99.45	0.033
Main Injector	1999	$8\,\text{GeV}$	$150\,\text{GeV}$	99.998	1
Tevatron	1983	$150\,\text{GeV}$	$1\,\text{TeV}$	99.99996	20

Table 3.1: Accelerator stages of the Tevatron. Reproduced from [28].

The acceleration process [29–31] begins with a bottle of hydrogen gas which is converted into H^- ions by using a magnetron to generate a plasma near the surface of a molybdenum-lined container. These H^- ions can be accelerated via an electric field, unlike the original neutral helium gas, and are accelerated to $750\,\text{keV}$ via the Cockcroft-Walton[2] electrostatic accelerator.

---

[2]Cockcroft-Walton refers to a type of voltage multiplying circuit commonly used in high voltage applications such as televisions and photocopiers.

The Tevatron accelerator

Booster ring   CDF experiment

Antiproton source

DZero experiment

Main injector   Tevatron ring

Beam tunnel of Tevatron ring

Tevatron
Chicago
ILLINOIS
USA
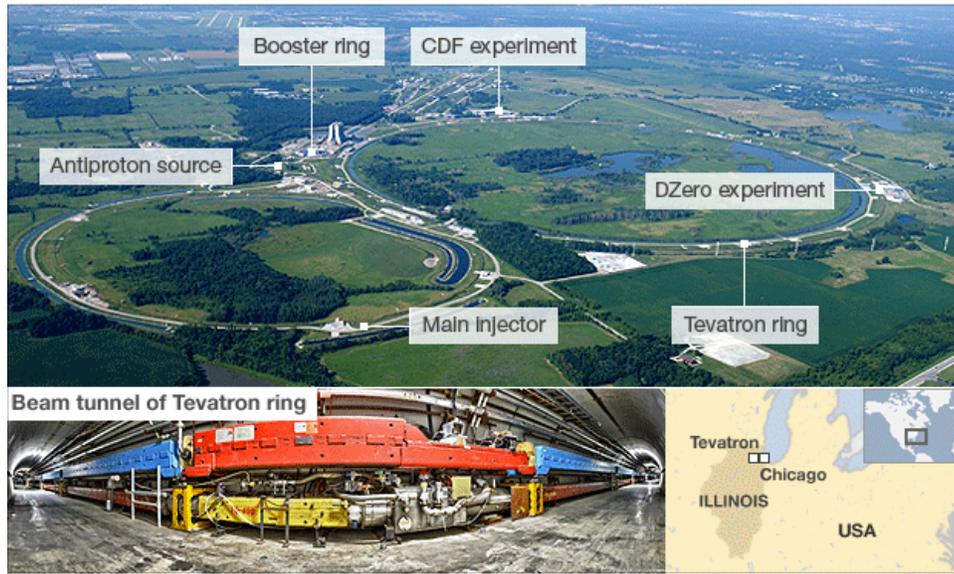
Source: Fermilab

FERMILAB'S ACCELERATOR CHAIN

MAIN INJECTOR

RECYCLER

TEVATRON

DZERO

CDF

PROTON

NEUTRINO   MESON

TARGET HALL

ANTIPROTON
SOURCE

BOOSTER
LINAC

COCKCROFT-WALTON

Antiproton   Proton
Direction   Direction
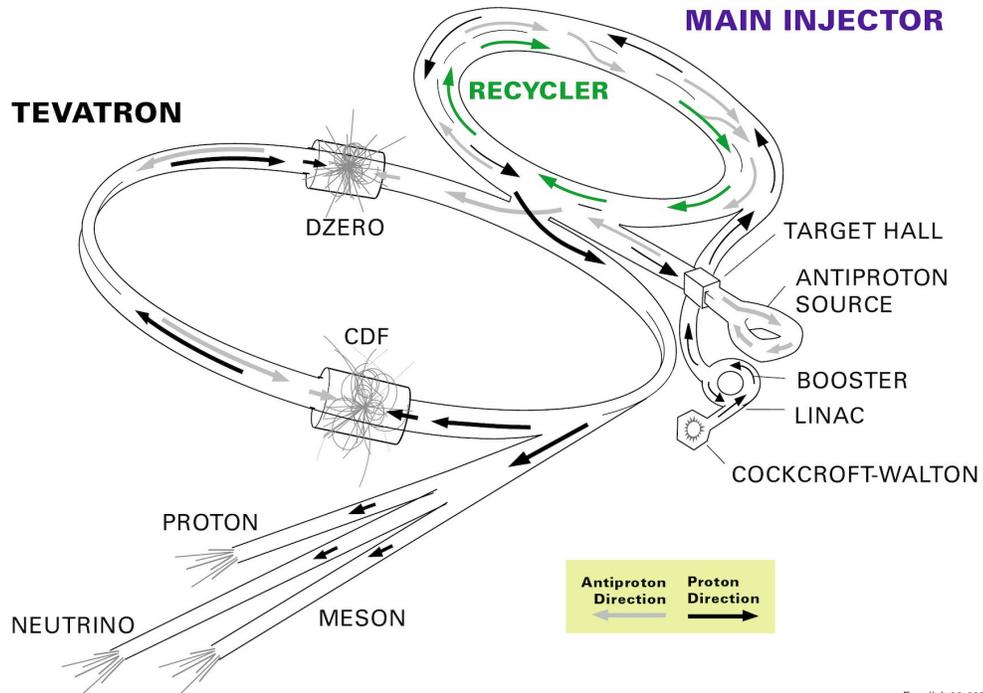
Fermilab 00-635

Figure 3.1: Top: aerial view of Fermilab and the Tevatron, photograph inside the tunnel of the Tevatron ring, and location map. Bottom: schematic showing the main parts of the Tevatron accelerator.

The helium ions then enter the Linac which is a 150 meter long linear accelerator made up of a series of electrically charged cylinders of increasing length whose charges oscillate to ensure the H^- ions are always being electrostatically pushed from behind and pulled forwards. The Linac accelerates the H^- ions to $401\,\mathrm{MeV}$ and passes them through a carbon foil which strips the electrons leaving just the protons.

The protons then enter the Booster, a synchrotron 468 meters in circumference which works by applying a magnetic field in the vertical direction, since the Lorentz force is $\vec{F} = q(\vec{E} + \vec{v} \times \vec{B})$, to bend the beam in a circle. Each revolution the beam of protons passes through an accelerating cavity, and after having circled around about 20,000 times in just 0.03 seconds, the protons leave the Booster with an energy of $8\,\mathrm{GeV}$.

Next the protons enter the Main Injector, a synchrotron $3.3\,\mathrm{km}$ in circumference, which can accelerate protons either to $120\,\mathrm{GeV}$ to be made into anti-protons or to $150\,\mathrm{GeV}$ for injection into the Tevatron. The Main Injector can also accelerate the $120\,\mathrm{GeV}$ anti-protons to $150\,\mathrm{GeV}$ for injection into the Tevatron. Furthermore, the Main Injector tunnel also houses the anti-proton Recycler ring which stores anti-protons to be injected into the Tevatron, but does not actually recycle them.

Anti-protons are made by colliding the $120\,\mathrm{GeV}$ proton beam with a nickel target. Anti-protons from these collisions are collected, focused into a beam with a Lithium Lens, and stored in the Accumulator ring to be injected back into the Main Injector.

The Tevatron is a synchrotron $6.3\,\mathrm{km}$ in circumference and was the first large accelerator to use superconducting magnets (about $1\,000$ of them). Protons and anti-protons enter the Tevatron, grouped into 3 trains of protons and 3 trains of anti-protons, separated by $2.6\,\mu\mathrm{s}$ and traveling in opposite directions, with each train made up of 12

bunches separated by 396 ns. These bunches are accelerated to 980 GeV each and focused to collide at the centers of the DØ and CDF detectors at a center-of-mass energy of $\sqrt{s} = 1.96\,\text{TeV}$.

In addition to collision energy, another accelerator parameter of interest is the luminosity, which is a measure of the amount of collisions. Specifically, the rate of collisions is given by

$$R = \frac{\mathrm{d}N}{\mathrm{d}t} = \mathcal{L}_{\text{inst}}\sigma, \tag{3.1}$$

where $\mathcal{L}_{\text{inst}}$ is the *instantaneous* luminosity, and σ is the cross section for whichever process we are considering. In terms of the beam parameters,

$$\mathcal{L}_{\text{inst}} = \frac{N_p N_{\bar{p}}}{A} f,$$

where f is the total bunch collision frequency, N_p and $N_{\bar{p}}$ are the number of protons and anti-protons in each bunch, and A is the effective cross sectional area of the beams [32]. For Gaussian beams, $A = 4\pi\sigma_x\sigma_y$ where $\sigma_{x,y}$ are the x, y-widths of the beam [33].

An alternate formulation for luminosity is often used in accelerator physics, expressed in terms of the beam emittance ϵ and the value of the betatron function at the interaction point β^*. The emittance is a measure of the size of the bunches in position and momentum phase space, and the betatron function describes the beam optics. Then the formula is [34]

$$\mathcal{L}_{\text{inst}} = F\frac{N_p N_{\bar{p}}\gamma}{4\pi\epsilon_n\beta^*} f,$$

where $F < 1$ is a geometric factor quantifying the amount of loss in overlap between the colliding bunches, and $\epsilon_n = \gamma\beta\epsilon$ (where $\beta \equiv v/c$ and γ is the Lorentz factor) is the momentum-independent normalized emittance.

The highest achieved instantaneous luminosity [35] delivered to CDF was $4.4 \times 10^{32}\,\text{cm}^{-2}\,\text{s}^{-1}$. The total, or *integrated* luminosity is a measure of the amount of data

in a collider experiment, and is the instantaneous luminosity integrated over time, given in units of inverse barns (often inverse femtobarns, fb^{-1}), where one barn is defined as $10^{-24}\,\mathrm{cm}^2$. The total number of particles produced in the experiment is then just the cross section times the integrated luminosity.

3.2 CDF

The Collider Detector at Fermilab (CDF) [36, 37] is one of two general purpose particle detectors in the Tevatron accelerator, the other being DØ. It is an international collaboration involving around 600 physicists from about 30 U.S. institutions and 30 international universities and laboratories. The detector is about 12 meters in all directions (about the size of a three-story house), and weighs around 6000 tons [38]. The main aims of the experiment are to search for physics beyond the Standard Model, study and measure properties of the top and bottom quarks and the W and Z bosons, study high energy jets and photons, and search for the Higgs boson.

CDF first began taking data in 1985 in the initial form of the experiment, called Run 0. Two major upgrades later took place, with Run I operating from 1992–1996 at $\sqrt{s} = 1.8\,\mathrm{TeV}$ collecting $0.13\,\mathrm{fb}^{-1}$ of data [39], and Run II from 2001–2011 at $\sqrt{s} = 1.96\,\mathrm{TeV}$ collecting $10\,\mathrm{fb}^{-1}$ of data [35].

The detector is cylindrically shaped with the beam pipe running through the axis and the collision point at the center, utilizing multiple layers of detector subsystems. Fig. 3.2 shows the detector and the arrangement of the various detector subsystems. In order of increasing distance from the beam pipe, these are: the particle tracking system for momentum measurement and b-jet identification, the solenoid magnet to bend the
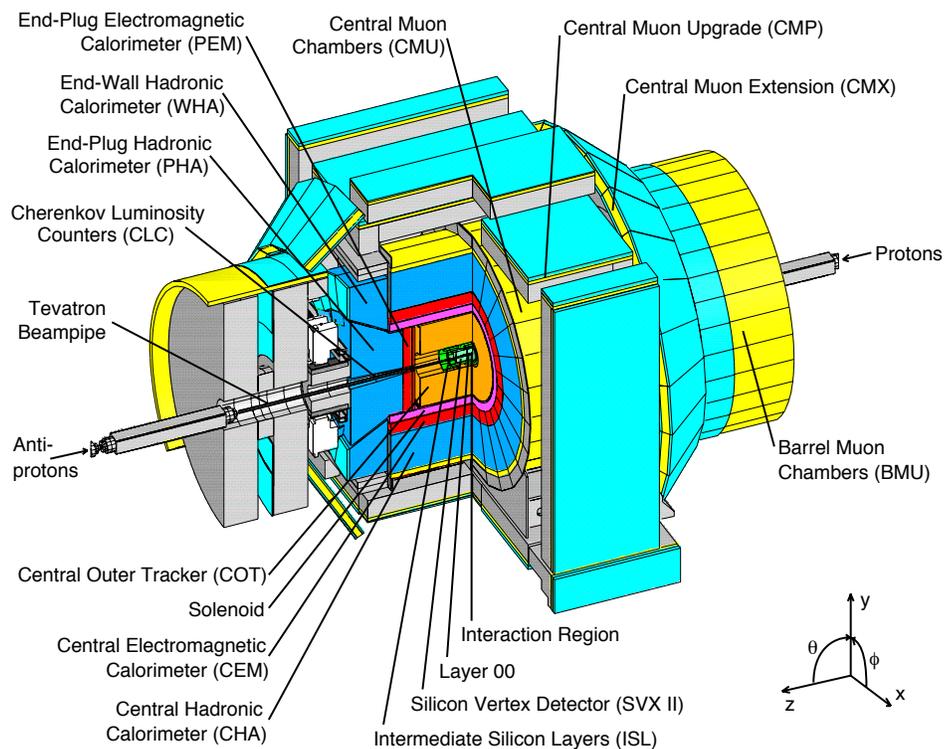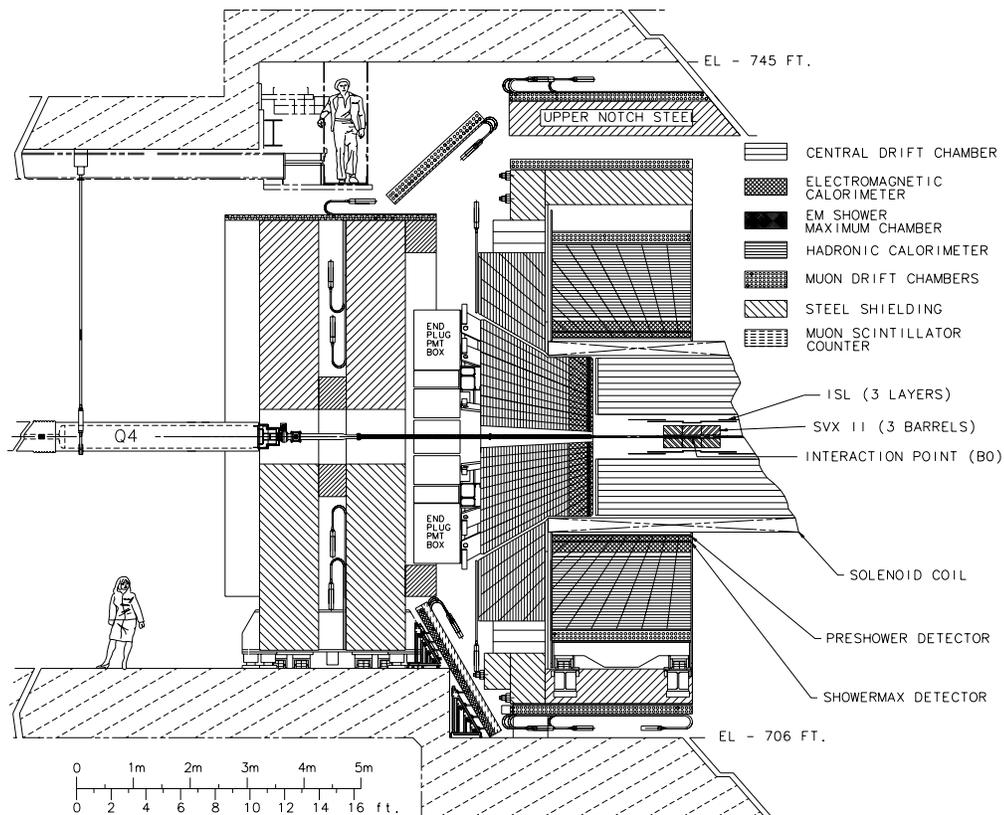
Figure 3.2: Top: elevation view of half the CDF II detector [37]. Bottom: isometric view [40].

tracks, the calorimeters for energy measurement, and the muon detectors. Just outside of the detector is the trigger and data acquisition system. Each of these will be discussed in more detail in §3.2.2.

3.2.1 Coordinate system

The CDF coordinate system is defined so that the $+z$-axis is taken to be in the direction the proton beam, $+x$ towards the outside of the ring, and $+y$ upwards. Relative to the proton beam, r is the radial distance, θ is the polar angle, and ϕ is the azimuthal angle. The pseudorapidity is defined as $\eta \equiv -\ln(\tan(\theta/2))$. The transverse momentum (p_T) and transverse energy (E_T) are the components of the momentum and energy vectors projected onto the x-y transverse plane.

3.2.2 Detector systems

Silicon Tracking System The silicon tracking system [41] is made up of three sub-detectors: the innermost silicon layer (L00), the silicon vertex detector (SVX II), and the intermediate silicon layer detector (ISL); they are shown in Fig. 3.3. Each are made up of silicon microstrip sensors which work by collecting the charge from electron-hole pairs created when a charged particle passes through silicon.

L00 consists of one layer at $r = 1.35\,\mathrm{cm}$, covering the area $|\eta| < 4$. SVX II consists of three barrels in z and 12 wedges in ϕ of five double-sided "stereo" layers[3] in r from $2.5\,\mathrm{cm}$ to $10.6\,\mathrm{cm}$, with a coverage of $|\eta| < 2$. SVX II is the main silicon detector with 0.4 million channels. The ISL detector consists of one layer at $r = 22\,\mathrm{cm}$ for

---

[3]The SVX II stereo layers are angled at $1.2°$ and $90°$ from the beam axis to provide tracking resolution in the z direction.
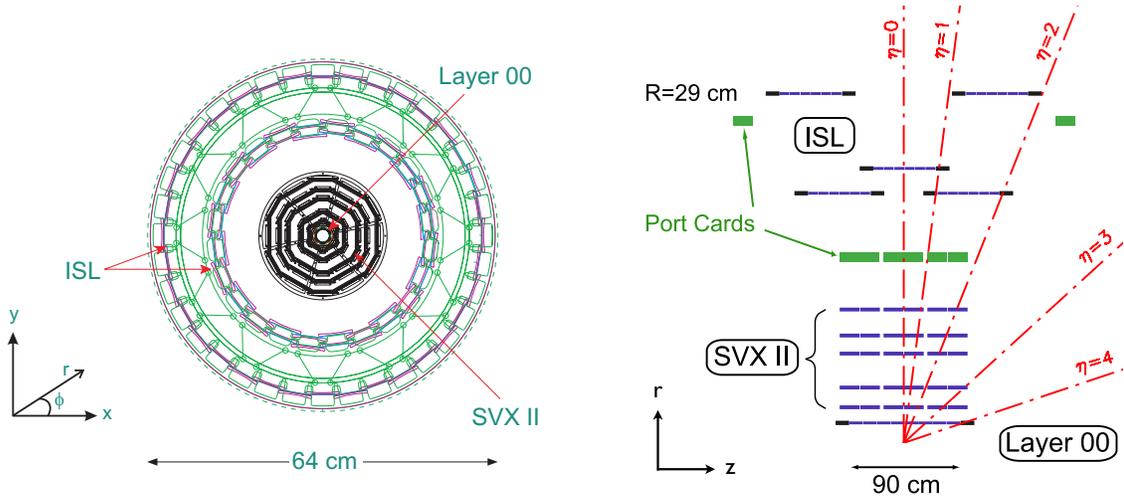
Figure 3.3: The silicon tracking system. Left: $r\phi$ view. Right: rz view. Figures from [42].

$|\eta| < 1$, and two layers, one at $r = 20\,\text{cm}$ and the other at $r = 28\,\text{cm}$, providing coverage between $1 < \eta < 2$. In total, the silicon tracking system consists of eight layers with a total of 0.7 million channels, has a hit position resolution of $10\,\mu\text{m}$, an impact parameter[4] resolution of $40\,\mu\text{m}$ (including a $30\,\mu\text{m}$ contribution from the beam width), and provides coverage in the region $|\eta| < 2$.

The silicon tracking system is especially important to our analysis for the tagging of b-jets. Since B hadrons have a lifetime of about $1.5\,\text{ps}$, they travel a measurable distance before decaying, and these distances are resolvable with the CDF tracking system. For example, since $c\tau \approx 450\,\mu\text{m}$, a B hadron with $p_{\text{T}} = 10\,\text{GeV}/c$ will travel on average $1\,\text{mm}$ from the interaction point in the transverse plane before decaying.

Central Outer Tracker (COT) The purpose of the Central Outer Tracker (COT) [43] is to provide tracking ability at larger distances ($r = 40\,\text{cm}$ to $r = 137\,\text{cm}$ and full coverage in $|\eta| \leq 1.0$). The COT is a $3.1\,\text{m}$ long cylindrical open-cell drift

---

[4]A track's impact parameter is the distance of closest approach to the beam line in the transverse plane.

chamber filled with gold wires and an equal mixture of argon and ethane gas. Alternating sense and potential wires set up an electric field, and the electron-hole pairs—created by the ionization of the gas by the charged particles—drift to these wires. The COT has over 30 000 sense wires strung between the two endplates with half strung parallel to the beam direction and the other half being stereo layers making $\pm 2°$ angles with the beam axis. The wires are spaced about 4 mm apart, allowing for a hit position resolution of 140 μm, and a p_T resolution of $\sigma_{p_T}/p_T = 0.15\% \times p_T/(1\,\mathrm{GeV}/c)$ for the COT only, and $\sigma_{p_T}/p_T = 0.07\% \times p_T/(1\,\mathrm{GeV}/c)$ for the COT plus silicon tracking system.

Tracks are reconstructed beginning in the COT with a cluster of three nearby hits consistent with a particle track from the interaction point. Additional hits are then progressively added to the chain by linking nearby hits in the direction of the track. Segments from the axial and stereo layers are linked and a five parameter helix is fit to the COT hits. A clustering algorithm identifies the silicon hits, and these are linked to the COT tracks by an outside-in algorithm which extrapolates each COT helix into the silicon and creates a search tree using the two best candidate hits in each layer. The final track is the one with the best fit (minimum χ^2).

In this analysis, the COT is important in providing improved accuracy in extrapolating tracks back to their primary or secondary vertices, as well as for measuring various track quantities which are used in our b-jet discriminator (see §4).

Solenoid Magnet The solenoid magnet [44] surrounds both the silicon tracking system and the COT, and provides a constant magnetic field parallel to the beam axis. This causes the tracks of charged particles to become curved, and the amount of curvature gives a measurement of the particle's momentum; particles with higher momentum curve less. The solenoid is 4.8 m long, located at a radius of 1.5 m, and

is superconducting to produce a 1.4 T magnetic field which is uniform to within 0.1% in the region $|r| < 1.5\,\text{m}$ and $|z| < 1.5\,\text{m}$.

Electromagnetic and Hadronic Calorimeters Calorimeters are positioned just outside of the solenoid to measure the total energy of both charged and neutral particles (specifically electrons and photons), and jets from hadronization. These modules are scintillator calorimeters segmented around the detector in a tower-based projective geometry. The inner electromagnetic calorimeters are composed of alternating lead sheets and plastic scintillator layers, while the outer hadronic calorimeters are composed of alternating steel and scintillator layers. The particles interact with the lead or steel and shower into many lower-energy particles whose energy is converted into light by the scintillators. That light is fed into photomultiplier tubes (PMTs) to measure the amount of light and hence the energy. The central and plug electromagnetic calorimeters have a thickness of $19X_0$ and $21X_0$ radiation lengths[5], or $1\lambda_I$ interaction length[6], and the central and plug hadronic calorimeters have a thickness of $4.5\lambda_I$ and $7\lambda_I$, respectively.

The central calorimeter system covers the region $|\eta| < 1$ and consists of the central electromagnetic (CEM) calorimeter, the central hadronic (CHA) calorimeter, and the end-wall hadronic (WHA) calorimeter. The forward end plug calorimeter system covers $1.1 < |\eta| < 3.6$ and consists of the plug electromagnetic (PEM) calorimeter and plug hadronic (PHA) calorimeter [44]. These can be seen in Fig. 3.2.

The energy resolution of the CEM is $\sigma_{E_\text{T}}/E_\text{T} = 13.5\%/\sqrt{E_\text{T}/(1\,\text{GeV})} \oplus 2\%$ [45], where $\oplus$ is addition in quadrature, and the energy resolution of the CHA is $\sigma_{E_\text{T}}/E_\text{T} = 50\%/\sqrt{E_\text{T}/(1\,\text{GeV})} \oplus 3\%$ [46, 47].

---

[5]The radiation length X_0 is a characteristic length for EM showers in a material through which all but $1/e$ of the energy is lost.

[6]The interaction length is analogous to the radiation length, but for hadronic showers.

Jets are identified by using the JETCLU algorithm [48] which looks for clusters of energy in the electromagnetic and hadronic calorimeters within a cone size of $\Delta R \equiv \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2} = 0.4$. To convert the measured jet energy into an estimate of the original parton energy, jet energy scale (JES) corrections [47] are made to account for the nonlinear response of the calorimeters, uninstrumented or damaged regions of the detector, energy not contained within the jet cone, multiple primary interactions, and the underlying event. The jet energy resolution can be approximated by $\sigma_{E_\mathrm{T}}/E_\mathrm{T} = 10\% + 1.0\,\mathrm{GeV}/E_\mathrm{T}$ [49, 50].

In this analysis we use the calorimeters to measure the transverse energy of jets. We require jets to have $E_\mathrm{T} > 20\,\mathrm{GeV}$ which corresponds to an uncertainty of $\sigma_{E_\mathrm{T}}/E_\mathrm{T} < 15\%$.

Muon Detectors The muon detectors make up the furthest layer of subdetectors from the center of the detector since muons interact weakly with matter and have a long lifetime ($2.2\,\mu$s). Due to this, muons will survive out to these detectors whereas other particles will have already decayed or been absorbed by the inner detectors. The muon detectors are made up of drift chambers and include the central muon detector (CMU) and central muon upgrade (CMP) for $|\eta| < 0.6$, the central muon extension (CMX) for $0.6 < |\eta| < 1.0$, and the barrel muon detector (BMU) for $1.0 < |\eta| < 1.5$. The CMU is sensitive to muons with $p_\mathrm{T} > 1.4\,\mathrm{GeV}/c$ while the CMP is separated by $60\,\mathrm{cm}$ of steel shielding and is sensitive to muons with $p_\mathrm{T} > 2.0\,\mathrm{GeV}/c$. These can be seen in Fig. 3.2.

In this analysis, we use the muon system to help suppress the large $t\bar{t}$ background, and also for obtaining data samples for evaluating our b-jet discriminator (see §4.5).

Trigger & Data Acquisition With a bunch collision rate of 1.7 million collisions per

second and roughly one million readout channels, CDF requires a rather sophisticated system for data readout and acquisition, triggering, and offline processing.

Analog signals and timing information from detector components are converted to digital signals via ADCs (analog to digital converters) and TDCs (time to digital converters), and read out with VME crates.

The CDF data storage system can write data at a rate of $20\,\mathrm{MB/s}$, corresponding to a maximum event rate of $100\,\mathrm{Hz}$. In order to reduce the rate to fall within this limit, 99.99% of events must be rejected [51]. Most of these are uninteresting QCD multi-jet or soft scattering events. The trigger employs progressing filtering in three tiers which operate on longer time scales with an increasing amount of information, in order to discard the uninteresting events, and doing so fast enough to not miss the interesting events. The trigger schematic is shown in Fig. 3.4 and consists mainly of:

Level 1 Dedicated hardware processors in a synchronous pipeline using coarse-grained data from a subset of detector components.

Level 2 Mixture of dedicated hardware processors and modified commercial processors in an asynchronous pipeline, exploiting additional data—notably from the silicon vertex tracker (SVT) which allows for triggering on displaced tracks.

Level 3 CPU farm of roughly 300 dual processor computers running the full reconstruction code in Linux.
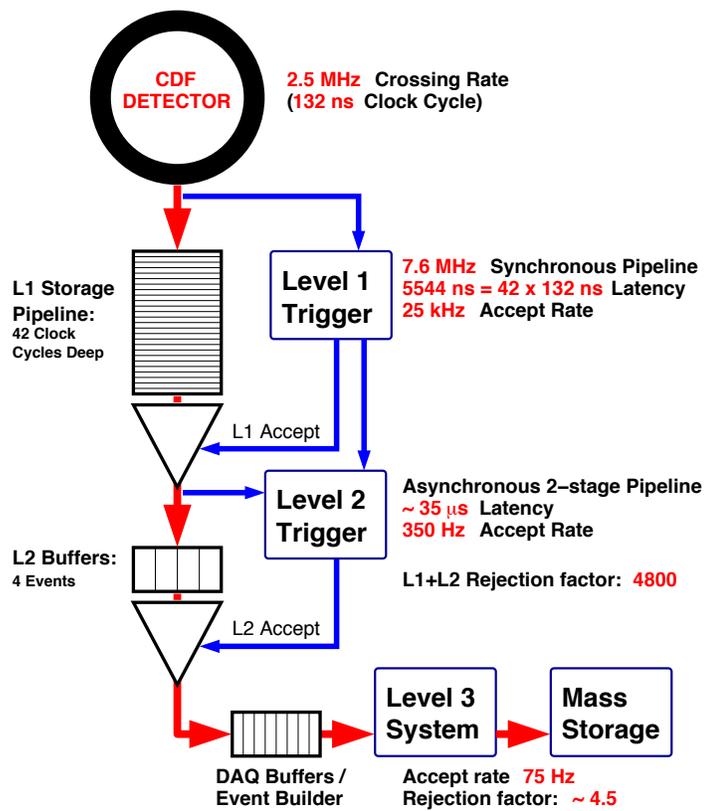
Figure 3.4: Dataflow of CDF trigger and DAQ. Typical accept rates and rejection factors are shown. Figure from [51].

3.3 PARTICLE IDENTIFICATION

Now that we have an understanding of the CDF detector, it remains to be seen exactly how information from the various detector subsystems is used to identify physics objects—specifically photons, electrons, muons, jets, and neutrinos (missing energy). The key is to exploit the differences in how the particles interact with the various detectors. This can be seen in Fig. 3.5. Jets from b quarks will be discussed in §4.
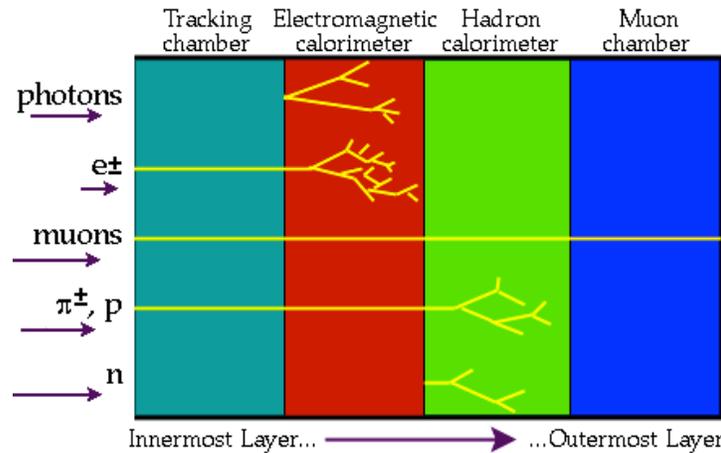
Figure 3.5: Particle signatures in a typical detector at a hadron collider. Figure from [1].

Photons Photons are neutral particles and hence do not leave tracks in the tracking system. However, they do interact with the lead in the electromagnetic (EM) calorimeter, decaying to an electron and positron (pair production). Those electrons and positrons in turn interact with the lead in the EM calorimeter and emit photons (bremsstrahlung). These processes are repeated causing an electromagnetic shower of lower and lower energy particles until all energy is absorbed by the calorimeter.

Electrons Being charged particles, electrons are seen passing through the tracking system via ionization. They then interact with the lead in the electromagnetic

calorimeter and produce an electromagnetic shower. To identify an electron, the energy deposit in the CEM or PEM must be matched to a track in the COT.

Muons Since muons are about 200 times more massive than electrons and are a minimum ionizing particle (MIP), they pass through the calorimeters and leave a hit in the muon detectors. That muon "stub" must then be matched to a track in the COT to identify a muon.

Jets / Hadrons Although hadrons are heavier than muons, they are stopped[7] in the hadronic calorimeter since they interact via the strong force. This interaction produces a hadronic shower wherein lighter hadrons are created and the particles are stopped or absorbed, leaving a cluster of energy deposits in the calorimeter.

Charged hadrons such as $\pi^\pm, K^\pm$, and p will leave ionization tracks in the tracking system and EM calorimeter while neutral hadrons such as π^0, K^0, and n will not.

Neutrinos (missing energy) Since neutrinos interact via the weak force and are stable, they do not interact with the matter in the detector and just pass through. Since they are not detected, the energy and momentum they carry off is missing. Since a hard scatting event involves only two of the partons, the rest of the constituents may travel down the beam line partially undetected. However, the transverse momentum of the hard scattering partons is to a good approximation zero initially, and can be measured in the final state, so only the transverse component of the momentum conservation requirement is used.

Typically the *energy* imbalance from the calorimeters is used rather than the momentum imbalance since it is more straightforward to measure. The amount of

---

[7]Some hadrons do pass through the calorimeter and leave a hit in the CMU. This is called hadronic punchthrough and can be greatly suppressed by requiring a muon also has a hit in the further shielded CMP.

missing transverse energy ($\not{E}_T$ or MET) in an event is given by the magnitude of the negative vectorial sum of all energy deposits over some threshold. Specifically, $\not{E}_T \equiv -\left| \sum_i E_T^i \vec{n}_i \right|$, where $\vec{n}_i$ is the unit vector in the azimuthal plane that points from the beam line to the ith calorimeter tower. This number is corrected for muons which deposit only a small fraction of their energy in the calorimeters by subtracting that energy and adding it back in from the muon's momentum measurement. Also, jet energy corrections are applied before calculating the $\not{E}_T$.

An event with large $\not{E}_T$ is indicative of the presence of one or more high transverse momentum neutrinos. However, there are other ways to get $\not{E}_T$ such as through energy mis-measurement, so additional techniques for neutrino selection are often employed such as the use of $\not{E}_T$-significance [52, 53] and $\Delta\phi(\vec{\not{E}_T}, \text{jet})$ in this analysis. The $\not{E}_T$-significance is a measure of the ratio of the value of $\not{E}_T$ to its uncertainty, and tends to be small for $\not{E}_T$ due to mis-measurement rather than due to undetected, long-lived neutral particles such as neutrinos. The $\Delta\phi(\vec{\not{E}_T}, \text{jet})$, discussed more in §5.2.1, is the smallest azimuthal angle between the $\not{E}_T$ vector and the jet axis for all jets above some low E_T threshold, and will be small for $\not{E}_T$ caused by a mis-measured jet.

IDENTIFICATION OF b QUARK JETS

In this chapter we discuss the novel artificial neural network based b jet tagger we developed for use in this analysis. This work has been published in [54].

4.1 B JET CHARACTERISTICS

The identification of jets resulting from the hadronization of b quarks and the subsequent decay of the B hadrons is an integral part of many physics analyses at high-energy particle collider experiments. This identification of b jets, known as b tagging, is used in searches for a low mass Higgs decaying to $b\bar{b}$, or in top quark analyses since $t \rightarrow W^+ b$. Searches for new physics also typically involve b tagging, especially in theories where the third generation has a special role.

The discrimination of b jets from non-b jets mainly exploits the long measurable lifetime (1.5 ps) of B hadrons and their relatively large mass ($5\,\mathrm{GeV}/c^2$). The long lifetime means that tracks from the B decay typically have a large impact parameter and meet at a displaced secondary vertex where the B decayed. This is shown in Fig. 4.1. The relatively large mass of the b quark means that they are very rarely created during string fragmentation and hence B hadrons are more likely to be from the higher momentum hard process than a typical light hadron [7]. Furthermore the number of particles produced during the fragmentation depends on the mass of the parton and is fewer for b jets [55],

and hence a B hadron will carry a larger fraction of the momentum of the parton than a light hadron. This implies that decay products within a b jet are more collimated within the jet cone and form a larger invariant mass. Additionally, the larger mass of the B hadron means it will typically decay into more particles than light hadrons.

Other characteristic features of b jets can be exploited such as the presence of a soft lepton (e or μ) in the jet in about 20% of B decays, either via semileptonic decay from the B directly or from a D or Λ_c resulting from the B.
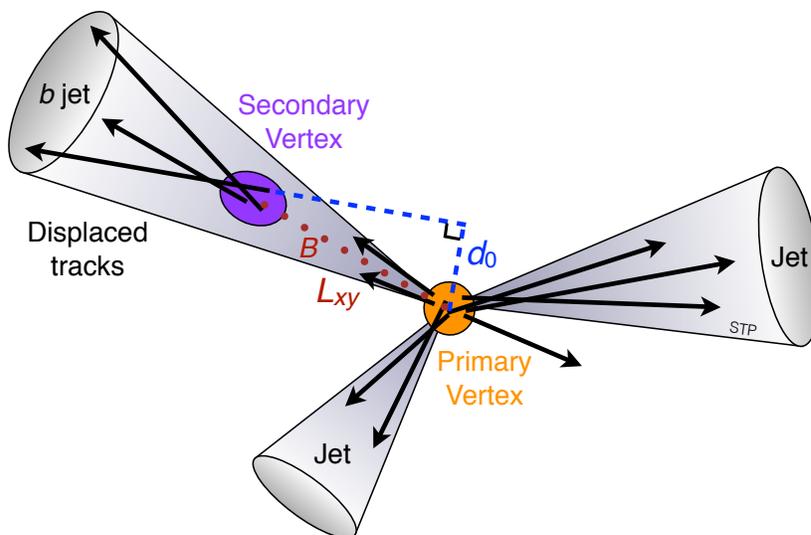
Figure 4.1: Illustration of displaced tracks and a secondary vertex from the decay of a B hadron. The decay length L_{xy} and track impact parameter d_0 are shown.

4.2 B TAGGERS AT CDF

Various b tagging algorithms have been developed at CDF since their successful application in discovering the top quark in 1995 [56]. In this section we discuss the main b-tagging algorithms used at CDF. Similar techniques have been developed at the DØ experiment [57] and at the CMS and ATLAS experiments at the LHC [58, 59].

SecVtx [50] is a secondary vertex tagger. It is the most commonly used b tagger at CDF. Using only significantly displaced tracks that pass certain quality requirements within each jet's cone, an iterative method is used to fit a secondary vertex within the jet. Candidate b-jets are selected based on the significance of the two-dimensional decay length in the r-ϕ plane, where the significance is defined as the value divided by its measured uncertainty: L_{xy}/σ_{Lxy}. The algorithm can be executed with different sets of track requirements and threshold values. In practice, three operating points are used, referred to as "loose", "tight", and "ultra tight".

The jet probability [60] tagger on the other hand does not look for a secondary vertex, but instead uses the distribution of the impact parameter significance of tracks in a jet (d_0/σ_{d_0}). By comparing these values to the expected distribution of values from light jets, it is possible to determine the fraction of light jets whose tracks would be more significantly displaced from the primary vertex than those of the jet under study. While light-flavor jets should yield a fraction uniformly distributed from 0 to 1, due to the long B lifetime, b jets often produce significantly displaced tracks and hence tend toward a fraction of 0. Although this algorithm produces a continuous variable for discriminating b jets, in practice only three operating points are supported (jet probability $< 0.5\%$, 1%, and 5%).

Soft-lepton taggers [61] take a different approach to b tagging. Rather than focusing on tracks within a jet, they identify semi-leptonic decays by looking for a lepton matched to a jet. The branching ratio of approximately 10% per lepton makes this method useful, although if used alone this class of tagger is not competitive with the previously mentioned taggers. However, because a soft-lepton tagger does not rely on the presence of displaced tracks or vertices, it has a chance to identify b jets that the other methods

can not. In practice, only the soft muon tagger is used at CDF since high-purity electron or tau identification within jets is very difficult.

Neural networks (NNs) can use as many flavor discriminating observables as is computationally feasible; hence the efficiency (fraction of tagged b jets) of NN taggers is often equal to or greater than that of conventional taggers for a given mistag rate (fraction of tagged non-b jets). One such NN-based algorithm at CDF, called the "KIT flavor separator" [62], analyzes SecVtx-tagged jets and identifies secondary vertices that are likely from long-lived B hadrons, separating them from jets with secondary vertices that originate from charm hadrons or that are falsely reconstructed. This flavor separator has been used in many CDF analyses, notably in the CDF observation of single top quark production [40]. Another NN-based algorithm, the "Roma tagger" [63, 64], has been used at CDF in light Higgs searches. While the SecVtx tagger attempts to find exactly one displaced vertex in a jet, the Roma tagger uses a vertexing algorithm that can find multiple vertices, as may be the case when multiple hadrons decay within the same jet cone (for example, in a $B \rightarrow D$ decay). Three types of NNs are used: one to distinguish heavy from light vertices, another to distinguish heavy-candidate from light-candidate unvertexed tracks, and a third that takes as inputs the first two NN outputs along with other flavor discriminating information, including SecVtx and jet probability tag statuses, number of identified muons, and vertex displacement and mass information. The performance of the Roma tagger is roughly equivalent to SecVtx at its operating points but allows for an "ultra loose" operating point yielding greater efficiency, useful in certain analyses.

4.2.1 The $b$ness tagger

The b tagger we developed for use in this analysis, known as the $b$ness tagger [54], is also a NN-based tagger, but is unique in its emphasis on exploiting information from individual tracks and in its ability to evaluate jets where a secondary vertex could not be reconstructed. This allows for operating at a greater efficiency than previous taggers—a feature useful in searches for rare processes such as ours or in Higgs searches. Our analysis [65] is the first application of this tagger, though it has also been used in a search for $WZ + ZZ \rightarrow \ell^+\ell^- + jj$ [66], and in at least one more diboson analysis currently underway at CDF.

The $b$ness tagger is a two-level NN as illustrated in Fig. 4.2, whereby all tracks in a jet passing loose track quality requirements are evaluated by a NN trained to discriminate between tracks from b jets and tracks from non-b jets, and these track NN outputs are fed into a jet NN which also uses the secondary vertex decay length significance and other jet-level information. The output of this jet NN is the final figure of merit for how likely the jet is from a b (the jet's "$b$ness"). This output is a continuous variable which allows for the b tagger operating point to be tuned to the desired efficiency and mistag rate in order to optimize individual analyses. The efficiency and mistag rate are measured in $t\bar{t}$ and $Z + 1$ jet samples, respectively. This novel method results in small systematic uncertainties on the measured efficiency and mistag rates.
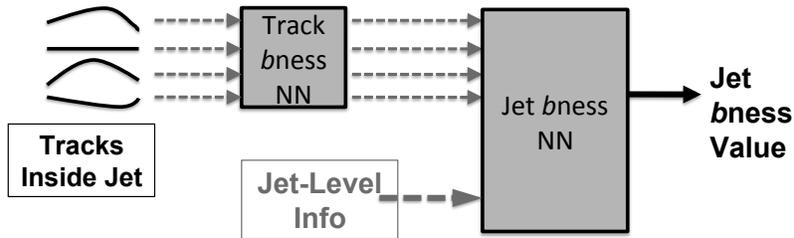
Figure 4.2: Schematic of the data flow in the two-level NN $b$ness tagger.

The Higgs Optimized b Identification Tagger (HOBIT) [67] is the latest b tagger at CDF. HOBIT is a NN-based tagger which takes as inputs the output of the track $b$ness, Roma, soft muon, jet probability, and SecVtx taggers, and is optimized for tagging b jets from $H \to b\bar{b}$ decays. It is estimated that replacing the previous tagging algorithms in light Higgs searches at CDF would result in improvements of 10–20% in Higgs boson sensitivity [67] and some analyses [68, 69] have rapidly made the switch.

4.3 ARTIFICIAL NEURAL NETWORKS

An artificial neural network (ANN or NN) [70] is a computer algorithm which roughly models biological neural systems. NNs are useful in discovering and exploiting correlations in multivariate systems for use in modeling complex relationships between inputs and outputs. NNs are used in a wide variety of applications of function approximation, pattern and object recognition or classification, and data processing. They are used in scientific, commercial, medical, and financial institutions. A few interesting examples of the broad applications of NNs are: speech and handwriting recognition, face identification, medical diagnosis, stock trading, automatic control of heating, ventilating, and cooling (HVAC) systems, and even washing machines, rice cookers, and audio synthesizers [71–74]. In particle physics NNs are typically used to discriminate between signal and background.

The basic structure of a multi-layer feed-forward NN is shown in Fig. 4.3. The boxes are analogous to neurons and the arrows to synapses. Input values x_i are stored in the top layer and the node values y_j^1 in the first layer are calculated as a function of the input values and the connected weights w_{ij}^1 between input layer node i to node j in the first hidden layer. Those results are propagated down to subsequent layers until the result is

obtained. The calculations typically use a weighted sum synapse function and a sigmoid neuron activation function, so that, for the example in Fig. 4.3, we have

$$y_{\text{NN}} = y_1^2 = \text{Sig}\left(\sum_{j=1}^{N_{\text{H}}} y_j^1 w_{j1}^2\right) = \text{Sig}\left[\sum_{j=1}^{N_{\text{H}}} \text{Sig}\left(\sum_{i=1}^{N_{\text{var}}} x_i w_{ij}^1\right) w_{j1}^2\right], \quad \text{where} \quad \text{Sig}(x) \equiv \frac{1}{1 + e^{-\beta x}},$$

and where N_{H} is the number of nodes in the hidden layer, N_{var} is the number of input variables, and β is the slope parameter.
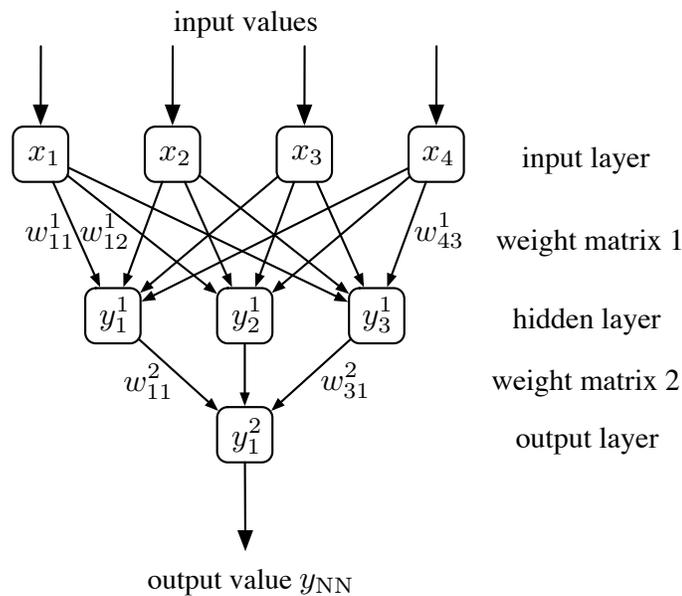
Figure 4.3: Illustration of an artificial neural network.

NNs are trained to yield the desired output by adjusting the weights according to an algorithm known as back-propagation (backward propagation of errors). Training begins with an initial random seed for the weights and uses the method of steepest descent to iteratively adjust the weights. Specifically, the error function for a single training event is defined as

$$E(\vec{x}, \vec{w}) = \frac{1}{2}\left[y_{\text{NN}}(\vec{x}, \vec{w}) - \hat{y}\right]^2,$$

where $\hat{y}$ is the desired output (typically 1 for signal, 0 for background). Then weights are updated by

$$\vec{w}' = \vec{w} - \eta \vec{\nabla}_{\vec{w}} E(\vec{x}, \vec{w}),$$

where η is a positive parameter called the learning rate. More advanced training algorithms have been developed but their explanation is beyond the scope of this thesis. Training continues for a fixed number of iterations or until the weights converge. Care must be taken not to "overtrain" the NN, wherein features specific to the training sample are learned, seemingly increasing the classification performance in the training sample while decreasing the general performance in an independent test sample. This can be avoided by halting training when the error in the test sample reaches a minimum.

After training, we would like some sense of which variables are used more strongly by the NN. Due to the complex network structure of a NN, it is often not clear how each variable independently affects the final output. One measure is the "importance" from TMVA, defined as

$$I_i \equiv \bar{x}_i^{\,2} \sum_{j=1}^{N_{\text{H}}} \left(w_{ij}^1\right)^2, \quad i = 1, \ldots, N_{\text{var}},$$

where $\bar{x}_i$ is the sample mean of the ith input variable. However, we find this to be an unsatisfactory metric since variables are internally normalized between -1 and 1, and so roughly symmetric variables will have a mean near zero, and hence a small importance value regardless of the variable's true importance. Instead we use a simplification[1] of Garson's algorithm [75]. For our NNs we found this simplified version to be in good

---

[1]Eq. 4.1 is a simplification of Garson's algorithm wherein we (1) assume that the hidden-to-output layer weights are equal, since as we use two hidden layers, Garson's algorithm would need to be applied iteratively which increases complexity, and (2) we do not normalize the weights for each hidden layer since that step is not particularly well motivated and introduces further abstraction.

agreement with the full Garson algorithm. Thus for clarity we will use the simpler metric, here called the relative weight and defined by

$$W_i \equiv \frac{\sum_{j=1}^{N_{\mathrm{H}}} |w_{ij}^1|}{\sum_{i=1}^{N_{\mathrm{var}}} \sum_{j=1}^{N_{\mathrm{H}}} |w_{ij}^1|}. \tag{4.1}$$

A separate measure of the importance potential of a variable is the amount of separation between the signal and background distributions, defined by

$$\langle S^2 \rangle \equiv \frac{1}{2} \int \frac{\left[S(x) - B(x) \right]^2}{S(x) + B(x)} \, \mathrm{d}x, \tag{4.2}$$

where S and B are the signal and background distributions, normalized to unit area. Note however that unlike the relative weight, this quantity is independent of the actual NN.

4.4   Specification and training of the $b$ness neural networks

An overview of the $b$ness tagger has already been described in §4.2.1. Here we specify the details.

The NNs are feed-forward multi-layer perceptrons with a single output and two hidden layers of 14 and 15 nodes with a hyperbolic tangent neuron activation function, and implemented using the MLP algorithm with the Toolkit for Multivariate Data Analysis (TMVA) software [70]. Performance was similar for different combinations of number of hidden layers and nodes.

The NNs are trained using a simulated data sample of ZZ events from the Monte Carlo (MC) simulation software PYTHIA [7] for event generation and parton showering.

This MC data is run through a GEANT [76] simulation of the CDF II detector to simulate detector effects and to obtain the geometric and kinematic acceptances. In this training sample, signal is defined as jets matched within $\Delta R < 0.4$ of b quarks from $Z \to b\bar{b}$ decays, while background is defined as jets not matched to b quarks. The jet NN was trained with 9 000 signal and 27 000 background training events, while the track NN was trained with 30 000 signal and 150 000 background training events. Both NNs were trained using 200 training cycles—a number chosen such that the NNs were adequately trained but not overtrained.

Tracks evaluated by the tagger must pass basic track quality requirements. Tracks with hits only in the COT are rejected as their resolution is inadequate for b tagging. Only tracks within a cone of $\Delta R = 0.4$ about the jet axis are evaluated. Furthermore, tracks at CDF are required to have $p_T > 0.4 \, \text{GeV}/c$. Lastly, tracks consistent with K_S and Λ^0 decays are removed since they can mimic B decays due to their long lifetimes of $90 \, \text{ps}$ and $260 \, \text{ps}$, respectively. This is done by exploiting the large branching ratios for $K_S \to \pi^+ \pi^-$ and $\Lambda^0 \to p \pi^-$ by removing track pairs within jets with opposite charge, forming a two-track vertex, and with an invariant mass within $10 \, \text{GeV}/c^2$ of $0.497 \, \text{GeV}/c^2$ (K_S) or $1.116 \, \text{GeV}/c^2$ (Λ^0).

4.4.1 THE TRACK NEURAL NETWORK

The discriminating characteristics of b jets has already been discussed in §4.1. Based on that discussion we use as inputs to the track NN the following variables:

- d_0 — signed impact parameter[2]

---

[2]The sign of the impact parameter of a track is defined to be positive if the angle between the parent jet axis and the line joining the primary vertex to the point of closest approach of the track to the vertex is less than 90°, and negative otherwise.

- d_0 significance — impact parameter significance: d_0/σ_{d_0} where σ_{d_0} is the uncertainty on the measurement of d_0

- z_0 — z-distance from the primary vertex

- z_0 significance — z_0/σ_{z_0}

- p_{T} — track p_{T}

- $p_\perp$ — track p_{T} with respect to the jet axis

- η_{axis} — track rapidity with respect to the jet axis

- Parent jet E_{T}.

Distributions of these variables can be seen in Fig. 4.4. Note that the training events have been weighted to have a flat parent jet E_{T} spectrum. This is because we want the tagger to perform well regardless of the jet E_{T}, and hence we do not want the tagger to bias itself towards discriminating b jets in a certain jet E_{T} region. By weighting the jet E_{T} distributions to be flat, the NN learns that jet E_{T} is not important. By explicitly inputting jet E_{T} into the NN, it should furthermore not exploit any correlations to jet E_{T} in the other input variables.

The input variables with separation and relative weight rankings are shown in Tab. 4.1. By both measures, the displacement variables (d_0, z_0, and their significances) are more important than the kinematic variables (p_{T}, $p_\perp$, η_{axis}). The correlation matrix between the input variables is shown in Fig. 4.5. Correlations are generally low, with the highest being between z_0 and z_0 significance (62%), and p_{T} and $p_\perp$ (50%).
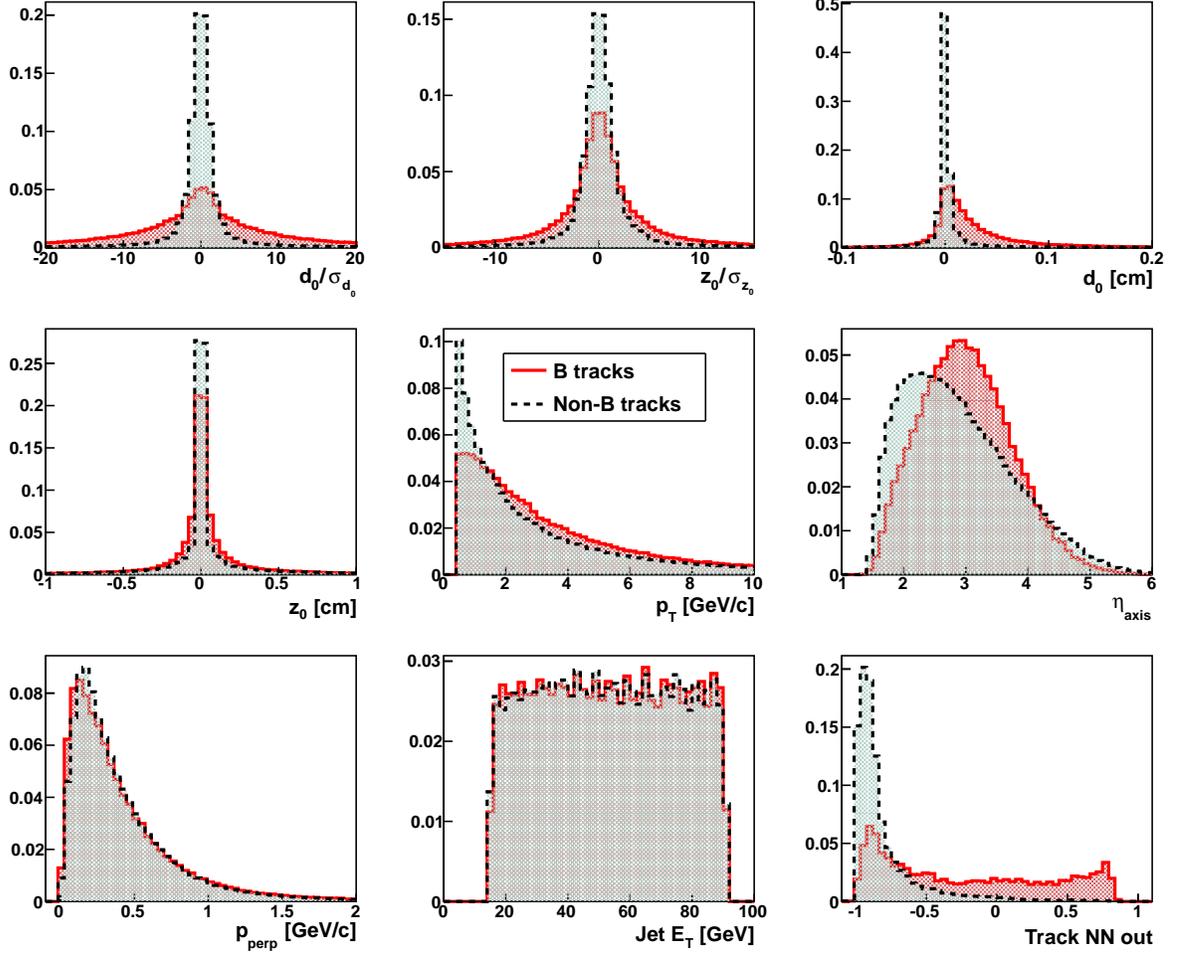
Figure 4.4: Distributions of the inputs to the track *b*ness NN in the *ZZ* MC simulated data used for training, along with the track NN output. The solid red curves are tracks matched within $\Delta R < 0.141$ to particles coming from B decays and the dashed black curves are tracks not matched to particles from B decays. Distributions are scaled to equal number of events and normalized to unit area. Events are weighted to a flat parent jet E_T distribution so as to not kinematically bias the tagger.

Variable	Separation		Relative weight	
d_0 significance	**1**	0.24	**1**	0.30
z_0 significance	**3**	0.07	**7**	0.03
d_0	**2**	0.21	**2**	0.26
z_0	**5**	0.02	**3**	0.15
p_{T}	**6**	0.02	**4**	0.09
η_{axis}	**4**	0.04	**6**	0.06
$p_\perp$	**8**	0.0005	**5**	0.08
Jet E_{T}	**7**	0.006	**8**	0.03

Table 4.1: Track NN input variable separation and relative weight measure. The separation (Eq. 4.2) measures the amount of overlap between signal and background distributions and is zero for exact overlap and one for no overlap. The relative weight (Eq. 4.1) is a measure of how much influence that variable has in the NN. Note that when variables are highly correlated, the NN may not rank them all as important as would be expected, as in the case of z_0 significance having a separation rank of #3 but a relative weight of #7. This is because z_0 significance is highly correlated with z_0 (relative weight #3), as seen in Fig. 4.5.
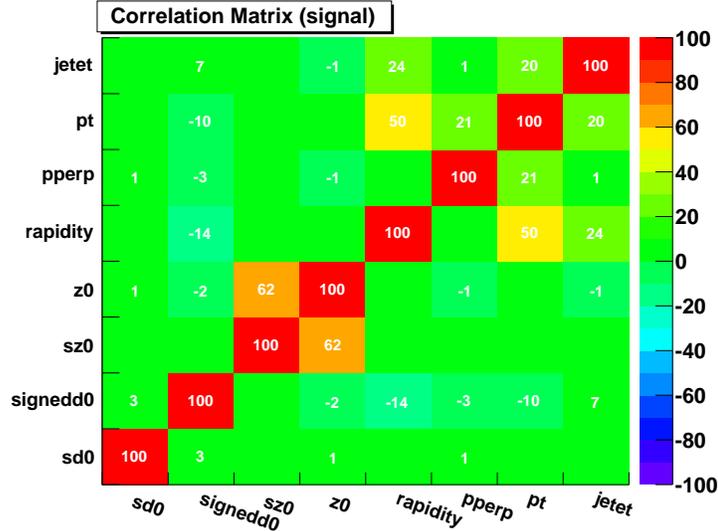
Figure 4.5: Matrix of linear correlation coefficients between the track NN input variables in the signal training sample.

4.4.2 THE JET NEURAL NETWORK

Based on the discussion of b jet characteristics in §4.1, we use as inputs to the jet NN the following variables:

- Top five track $b$nesses in the jet

- $L_{xy}/\sigma_{L_{xy}}$ — secondary vertex decay length significance (if found)

- Invariant mass of tracks in secondary vertex (if found)

- Number of tracks with $b$ness > 0

- Minimum muon likelihood — the likelihood that a muon candidate within the jet is a true muon. The value is calculated using the soft muon tagger [61], and a smaller value is more consistent with a true muon. If multiple muon candidates exist within the jet, the minimum muon likelihood value is used.

- K_S candidate — a K_S within a jet is more likely to have come from a cascade $B \to D \to K_S$ decay than not from a B decay. The K_S candidate is obtained using the same selection to remove K_S tracks.

- Parent jet E_T.

Note that the secondary vertex fitting is only performed on tracks with $b$ness greater than -0.5. The fitting is performed initially on all of such tracks within the jet. Next, an iterative procedure is performed wherein tracks which contribute a value of 50 or more to the total fit χ^2 are removed until no more such tracks exist, or until there are fewer than two tracks remaining in the fit. Note that if no secondary vertex is found, then the input values for the L_{xy} significance and the vertex mass lie in an underflow bin.

As with the track NN, the parent jet E_T is weighted to be flat and included as an input so as to not kinematically bias the tagger.

Distributions of the jet NN input variables are shown in Fig. 4.6. The input variables with separation and relative weight rankings are shown in Tab. 4.2. The most important variables are the track $b$ness inputs and the number of tracks, while the least important variables are the muon likelihood and K_S candidate. The correlation matrix between the input variables is shown in Fig. 4.7. The highest correlation was between L_{xy} significance and the vertex mass (94%), followed by high correlations between track $b$ness variables and the number of tracks. Distributions of the NN output for simulations of electroweak W and Z events similar to those in our main analysis are shown in Fig. 4.8.
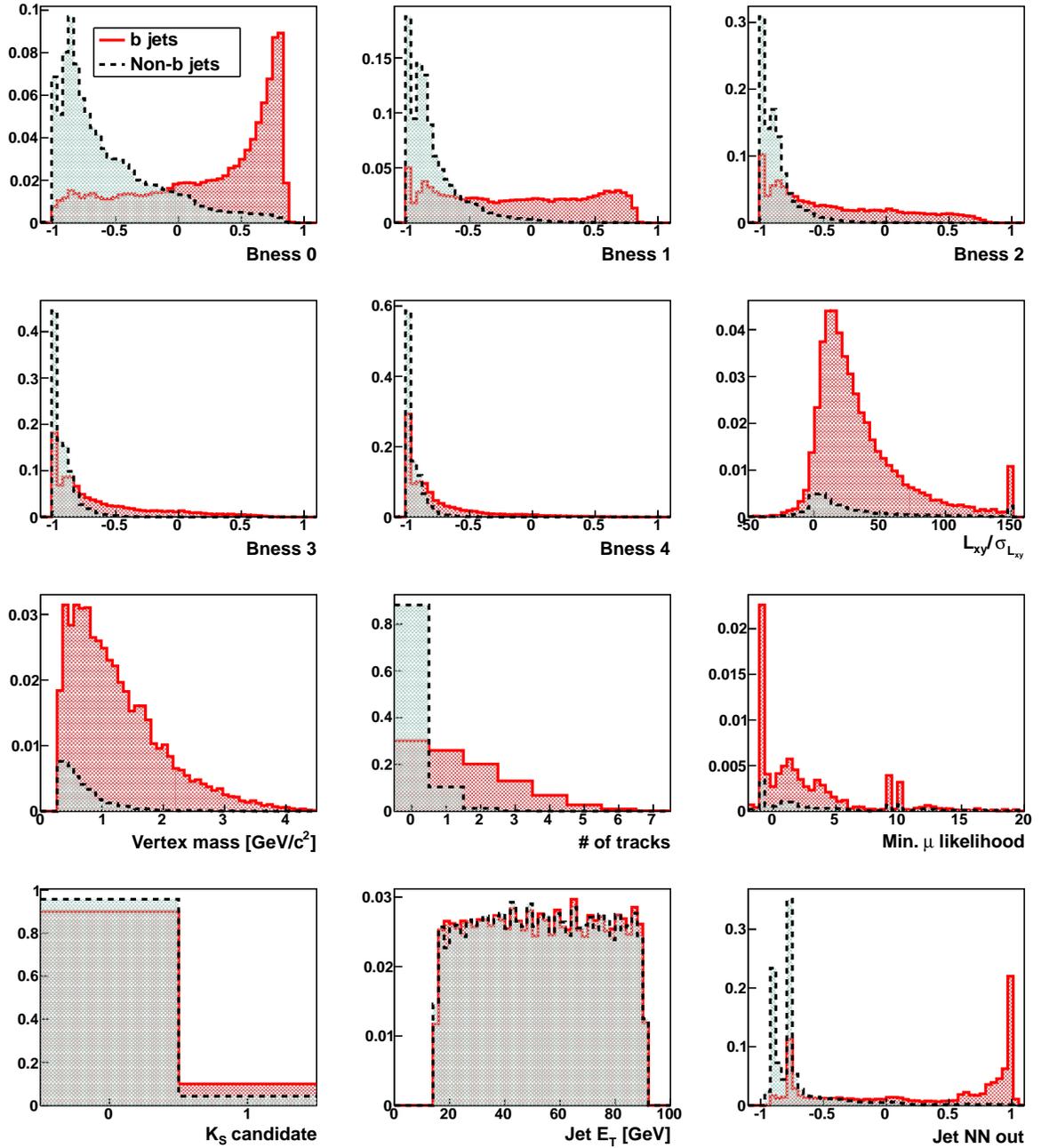
Figure 4.6: Distributions of the inputs to the jet *b*ness NN in the *ZZ* MC simulated data used for training, along with the jet NN output. The solid red curves are jets matched to a *b* and the dashed black curves are jets not matched to a *b*. Bness *i* refers to the *b*ness of the ith track, ordered in *b*ness (0 being the highest). Distributions are scaled to equal number of events and normalized to unit area. In the case of the decay length significance and vertex mass, if no secondary vertex is fit, the values fall outside the plotted region; similarly for the minimum muon likelihood.

Variable	Separation		Relative weight	
Bness 0	**1**	0.42	**1**	0.132
Bness 1	**2**	0.41	**9**	0.074
Bness 2	**4**	0.34	**2**	0.126
Bness 3	**5**	0.26	**5**	0.095
Bness 4	**8**	0.20	**4**	0.099
$L_{xy}/\sigma_{L_{xy}}$	**6**	0.26	**7**	0.089
Vertex mass	**7**	0.26	**6**	0.090
# of tracks	**3**	0.38	**3**	0.102
Min. μ likelihood	**10**	0.003	**10**	0.060
K_S candidate	**11**	0.0007	**11**	0.058
Jet E_{T}	**9**	0.006	**8**	0.074

Table 4.2: Jet NN input variable separation and relative weight measures. The separation (Eq. 4.2) measures the amount of overlap between signal and background distributions and is zero for exact overlap and one for no overlap. The relative weight (Eq. 4.1) is a measure of how much influence that variable has in the NN. Note that when variables are highly correlated, the NN may not rank them all as important as would be expected, as in the case of *b*ness 1 having a separation rank of #2 but a relative weight of #9. This is because *b*ness 1 is highly correlated with *b*ness 2 (relative weight #2) and the number of tracks (relative weight #3), as seen in Fig. 4.7.
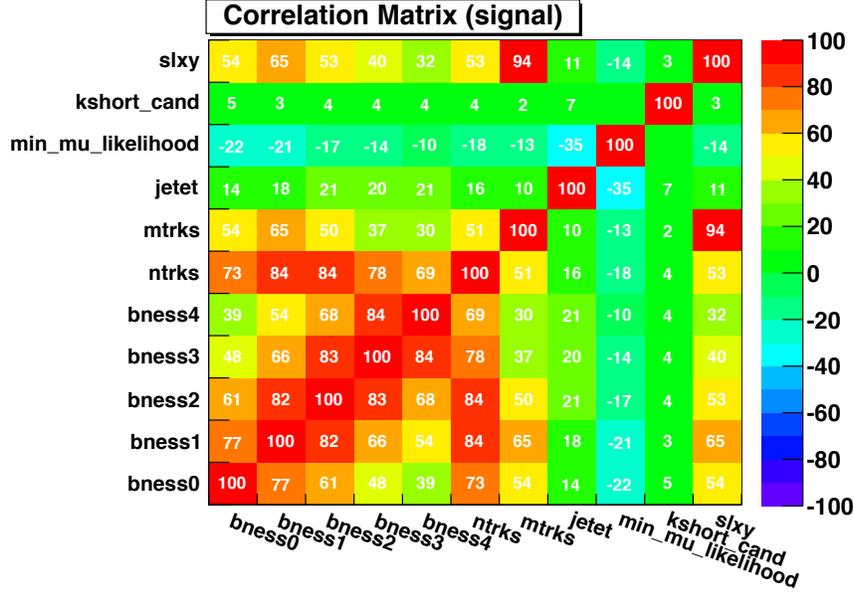
Figure 4.7: Matrix of linear correlation coefficients between the jet NN input variables in the signal training sample.
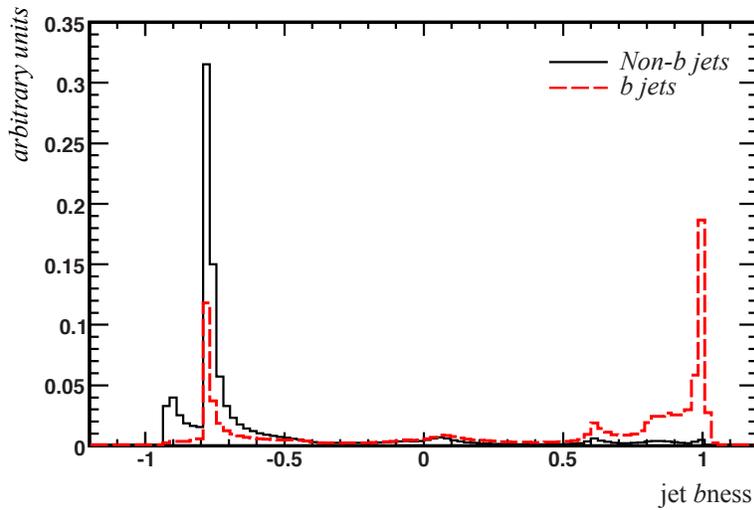
Figure 4.8: Output of the jet $b$ness NN for b jets (red dashed line) and non-b jets (black solid line). The high $b$ness region is dominated by b jets as desired. However, we note that there are also b jets peaked near -0.8. These are mostly from jets with no secondary vertex, zero tracks with positive $b$ness, and no K_S candidate. This is inevitable; some b jets do look just like non-b jets. Sharp features in the distribution are the result of integer valued inputs.

4.5 bNESS TAGGER VALIDATION METHOD

Since b taggers are trained on Monte Carlo simulations, we need to quantify how well they actually perform in data. There are two main quantities which characterize a b tagger. One is the fraction of b jets which are correctly tagged. This is known as the efficiency, or tag rate. The second quantity is the fraction of non-b jets which are incorrectly tagged. This is called the mistag rate, or false tag rate.

Since the $b$ness tagger has a tunable threshold, we must obtain these quantities as a function of the $b$ness cut. The efficiency $e(b)$ and mistag rate $m(b)$ are defined as

$$e(b) = \frac{N(b \text{ jets}) \text{ with } b\text{ness} > b}{N(b \text{ jets})} = \frac{N_B(b)}{N_B},$$

and

$$m(b) = \frac{N(\text{non-}b \text{ jets}) \text{ with } b\text{ness} > b}{N(\text{non-}b \text{ jets})} = \frac{N(b) - N_B(b)}{N - N_B}, \tag{4.3}$$

where N_B is the total number of b jets, $N_B(b)$ is the number of b jets above the threshold $b$ness, N is the total number of jets, and $N(b)$ is the total number of jets above the $b$ness threshold.

Furthermore, we evaluate the efficiency and mistag rate in the Monte Carlo simulation ($e_{\mathrm{MC}}(b)$ and $m_{\mathrm{MC}}(b)$, respectively), and determine the necessary scale factors, $s_e(b) = e(b)/e_{\mathrm{MC}}(b)$ and $s_m(b) = m(b)/m_{\mathrm{MC}}(b)$, to correct the results from simulation to agree with data. Finally, we obtain the uncertainties on the efficiency and mistag rate to use as a systematic uncertainty on the tagger.

These numbers are obtained by comparing data and simulation in two control regions:

$Z + 1$ jet Dominated by non-b jets in order to obtain the mistag rate. We choose events containing two oppositely charged electrons or muons consistent with the decay of a Z boson, plus one jet.

$\mathbf{t\bar{t}}$ Dominated by b jets in order to obtain the efficiency. We choose events containing the decay of a pair of top quarks, where we require exactly one lepton, at least four jets, and large $\not{E}_\mathrm{T}$ to select for the neutrino. The two jets with the highest $b$ness values in this sample should very likely be b jets.

The cuts applied for these two selection regions are described in Table 4.3. We use the $\not{E}_\mathrm{T}$-significance, as defined in [52, 53], to reduce background contributions from QCD multi-jet production where a jet is mis-identified as an electron or muon.

$Z + 1$ jet selection
$N_\mathrm{leptons} = 2$, both electrons or both muons
Leptons have opposite charge
Δz_0 between leptons $< 5\,\mathrm{cm}$
Lepton $p_\mathrm{T} > 20\,\mathrm{GeV}/c$
$75\,\mathrm{GeV}/c^2 < M_{ll} < 105\,\mathrm{GeV}/c^2$
$\not{E}_\mathrm{T} < 25\,\mathrm{GeV}$
Reconstructed $p_\mathrm{T}(Z) > 10\,\mathrm{GeV}/c$
$N_\mathrm{jets}(E_\mathrm{T} > 10\,\mathrm{GeV}) = 1$
Jet $E_\mathrm{T} > 20\,\mathrm{GeV}$, $|\eta| < 2.0$

$t\bar{t}$ selection
$N_\mathrm{leptons} = 1$
Lepton $p_\mathrm{T} > 20\,\mathrm{GeV}/c$
$\not{E}_\mathrm{T} > 20\,\mathrm{GeV}$
$\not{E}_\mathrm{T}$-significance $> 1(3)$ for $\mu(e)$ events
Reconstructed $M_\mathrm{T}(W) > 28\,\mathrm{GeV}/c^2$
Highest two $b$ness jets' $E_\mathrm{T} > 20\,\mathrm{GeV}$
$N_\mathrm{jets}(E_\mathrm{T} > 15\,\mathrm{GeV}) \geq 4$
Total sum $E_\mathrm{T} > 300\,\mathrm{GeV}$

Table 4.3: Summary of event selection requirements for the $Z + 1$ jet and $t\bar{t}$ samples. The total sum E_T is defined as the sum of the lepton p_T, $\not{E}_\mathrm{T}$, and E_T of all jets with $E_\mathrm{T} > 15\,\mathrm{GeV}$.

Monte Carlo simulations used for signal and background estimates are performed with the ALPGEN [77] event generator for W and Z plus jets samples and with PYTHIA for $t\bar{t}$ and processes with small contributions, with both samples interfaced to PYTHIA for parton showering. For the comparison to data, all sample cross sections are normalized to the results of NLO calculations performed with the MCFM v5.4 program [8] and using the CTEQ6M parton distribution functions [78].

We use data corresponding to an integrated luminosity of $4.8\,\mathrm{fb}^{-1}$. These events are selected using the high-p_T electron and muon triggers. Table 4.4 contains a summary of the total number of events.

	Electrons	Muons
$Z + 1$ jet selection		
Data Events	9512	5575
MC Events	9640 ± 880	5540 ± 490
$t\bar{t}$ selection		
Data Events	507	835
MC Events	542 ± 56	862 ± 85

Table 4.4: Number of events in data and MC in the $Z + 1$ jet selection region, after proper scale factors have been applied. The uncertainties on the MC reflect only the two dominant systematic uncertainties: the uncertainty on the jet energy scale and the uncertainty on the luminosity. Overall, the agreement in number of events is good.

4.5.1 MISTAG RATE DETERMINATION

The jet $b$ness distribution for the jets in the $Z + 1$ jet sample is shown in Fig. 4.9. The $t\bar{t}$ and diboson backgrounds are also included in the plot, but their contribution is negligible. We see that while the sample is mostly dominated by non-b jets as desired, there still remains a sizable contribution of b jets in the high $b$ness region. This can be seen more clearly in Fig. 4.10 which shows the contribution of jets matched to a b within a jet cone of $\Delta R < 0.4$, and jets not matched to a b. The b-jet purity ($N_{\mathrm{b\text{-}jets}}/N_{\mathrm{jets}}$) is also shown in Fig. 4.10 as a function of the lower threshold on the $b$ness cut. We see that the purity remains below 25% for $b$ness < 0.5, then sharply rises to more than 65% at the highest $b$ness cut values. Due to this, we expect the uncertainties in the mistag rate to be substantially higher there, due to both the small sample size and the high
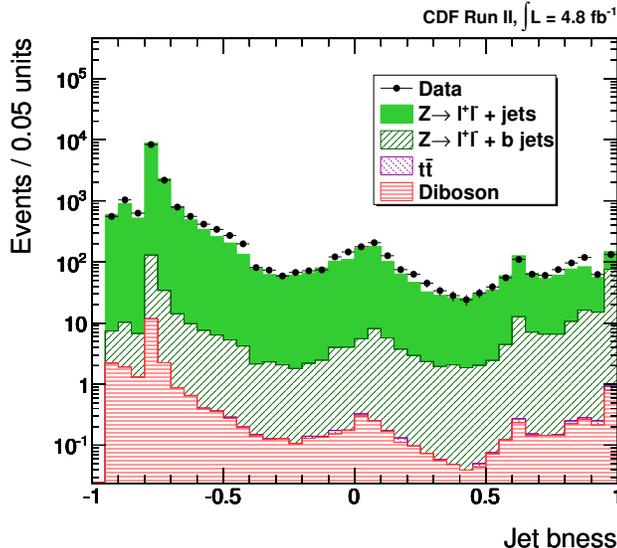
Figure 4.9: A comparison of the jet $b$ness in data and MC in the $Z + 1$ jet selection region. The MC is able to reproduce the main features of the $b$ness distribution in data. We use this distribution to determine the mistag rate for placing a cut on jet $b$ness in data, and use the differences between data and MC to determine corrections to the mistag rate in MC.

contamination rate of b jets, combined with the larger uncertainty on the number of b jets in that smaller sample.

Due to the significant b jet contribution in the sample, when calculating the mistag rate we must subtract from the data the expected number of b jets as obtained from the MC simulation. Specifically, we obtain the fraction f_B of b jets in the MC, and the b tag efficiency in MC, $e_{\mathrm{MC}}(b)$, so that we can write

$$N_B = f_B N, \quad \text{and} \quad N_B(b) = s_e(b) e_{\mathrm{MC}}(b) f_B N. \tag{4.4}$$

Note that we have multiplied $e_{\mathrm{MC}}(b)$ by the scale factor $s_e(b)$ to account for differences between the data and MC. Then if we define $m_{\mathrm{raw}}(b)$ to be the mistag rate in data before
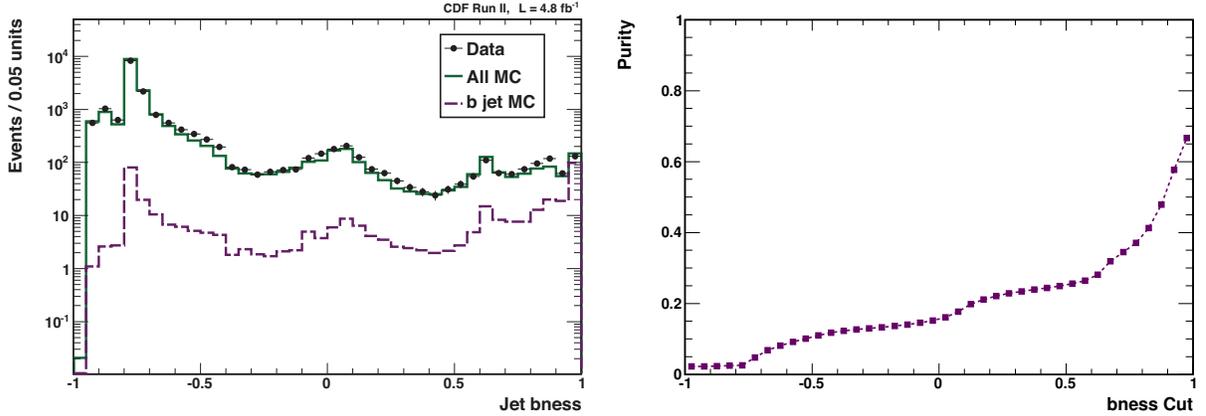
Figure 4.10: *Left*: A comparison of the jet *b*ness in data (black points) and MC (green solid line) in the $Z + 1$ jet selection region, with the portion of the MC jets matched to *b* quarks (purple dashed line) shown independently. *Right*: The *b*-jet purity for a given *b*ness cut, as determined from matched jets in the MC. As we wish to use the $Z + 1$ jet sample as a model for mistags, it is necessary to subtract the significant *b*-jet contribution at high *b*ness values.

subtracting the b jet contamination, Eq. 4.3 becomes

$$m(b) = \frac{m_{\mathrm{raw}}(b)N - s_e(b)e_{\mathrm{MC}}(b)f_B N}{N - f_B N} = \frac{m_{\mathrm{raw}}(b) - e(b)f_B}{1 - f_B}.$$

To evaluate this quantity which depends on the efficiency $e(b)$, we analogously write

$$e(b) = \frac{e_{\mathrm{raw}}(b) - m(b)f_L}{1 - f_L}, \tag{4.5}$$

where $e_{\mathrm{raw}}(b)$ is the efficiency uncorrected for the presence of non-b jets, and f_L is the fraction of non-b jets in the sample. This system of two equations in two unknowns is then solved in order to obtain both the efficiency and the mistag rate.

Furthermore, the *uncertainty* on the measured mistag rate in data can be evaluated given the uncertainty on the b tag efficiency and the uncertainty on f_B. The former is obtained by also writing the equation for the efficiency uncertainty which in turn depends on the uncertainty of the mistag rate, and solving the system iteratively (details in [54]), while the latter is taken to be 20% given the uncertainty of the ratio of $Z + b$ jet to inclusive Z production at CDF measured in [79].
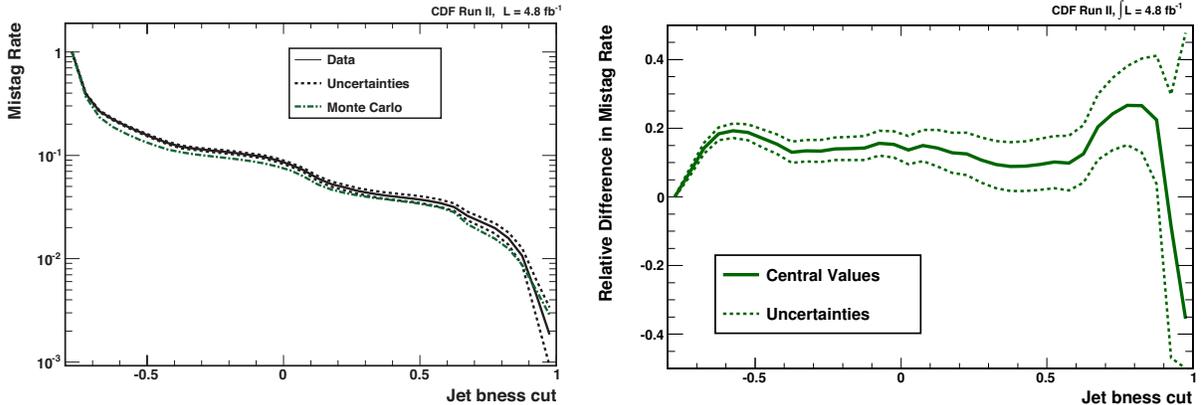
Figure 4.11: *Left*: The mistag rate in data (solid black line, dashed lines represent uncertainty) and Monte Carlo simulation (dot-dashed green line) as a function the cut on the jet $b$ness. We see our simulation typically under-predicts the mistag rate measured in data, requiring us to consider a correction to apply to the MC. *Right*: The calculated MC scale factor on the mistag rate (solid line) and its uncertainty (dashed lines) relative to the mistag rate in the MC. The value of the scale factors and their uncertainties at the relevant $b$ness cuts in this analysis are summarized in Tab. 4.5. We see very large uncertainties on the mistag rate scale factor around the high jet $b$ness cut of 0.85, due to the small number of events and significant heavy-flavor removal that must be done in this region.

The mistag rate and the relative difference between the mistag rate in data and MC $(s_m(b) - 1)$, along with their uncertainties, is shown in Fig. 4.11.

4.5.2   Tagging efficiency determination

The tagging efficiency determination is performed in the $t\bar{t}$ data selection, described in §4.5, with the selection cuts listed in Tab. 4.3 in order to obtain a sample with a large fraction of b jets. Since $t\bar{t}$ events with one lepton will typically have four or more jets, we order the jets by decreasing $b$ness value, following our diboson analysis where we also order jets by $b$ness since we are most interested in $Z/H \to b\bar{b}$ decays. The jet $b$ness distributions in MC and data in the $t\bar{t}$ sample for the two jets with highest $b$ness values are shown in Fig. 4.12. Agreement between data and MC is good, and we see a large
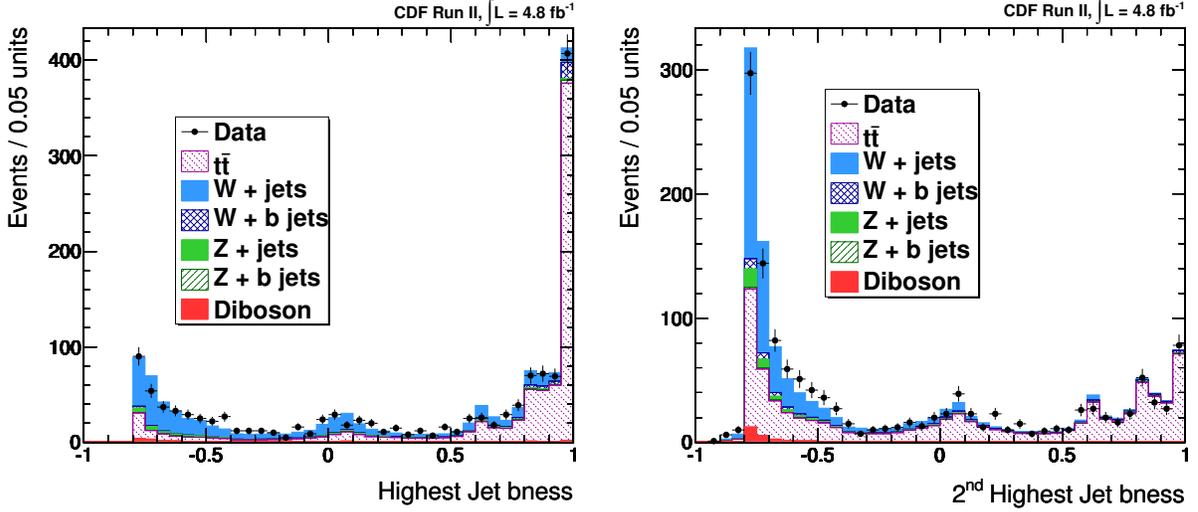
Figure 4.12: Jet *b*ness of the first (left) and second (right) jet, as ordered by *b*ness, in the $t\bar{t}$ lepton + jets selection region. The simulation reproduces most of the features of the data, and we see much of the *b*-enriched samples clustered towards high *b*ness.

contribution of $t\bar{t}$ peaked in the high *b*ness region, especially for the highest *b*ness jet, which indicates the tagger is indeed identifying *b* jets. As before, we separate the MC into jets matched to a *b* and jets not matched to a *b*, and show the *b* jet purity in Fig. 4.13. We see that for the highest *b*ness jet, the *b* purity ranges from 50% at the lowest *b*ness cut, to 90% at the highest *b*ness cut. The *b* purity of the second highest *b*ness jet similarly ranges from 35% to 90%.

Although the *b* purity is high, we still subtract the non-*b* jet contribution when calculating the efficiency (Eq. 4.5). This is done analogously to the subtraction of the *b* contamination in the $Z + 1$ jet sample for obtaining the mistag rate. The uncertainty on the efficiency is likewise obtained analogously to the mistag rate uncertainty.

The efficiencies for the highest and 2nd highest *b*ness jets are shown along with their uncertainties in Fig. 4.14. The relative difference between the efficiencies in data and MC
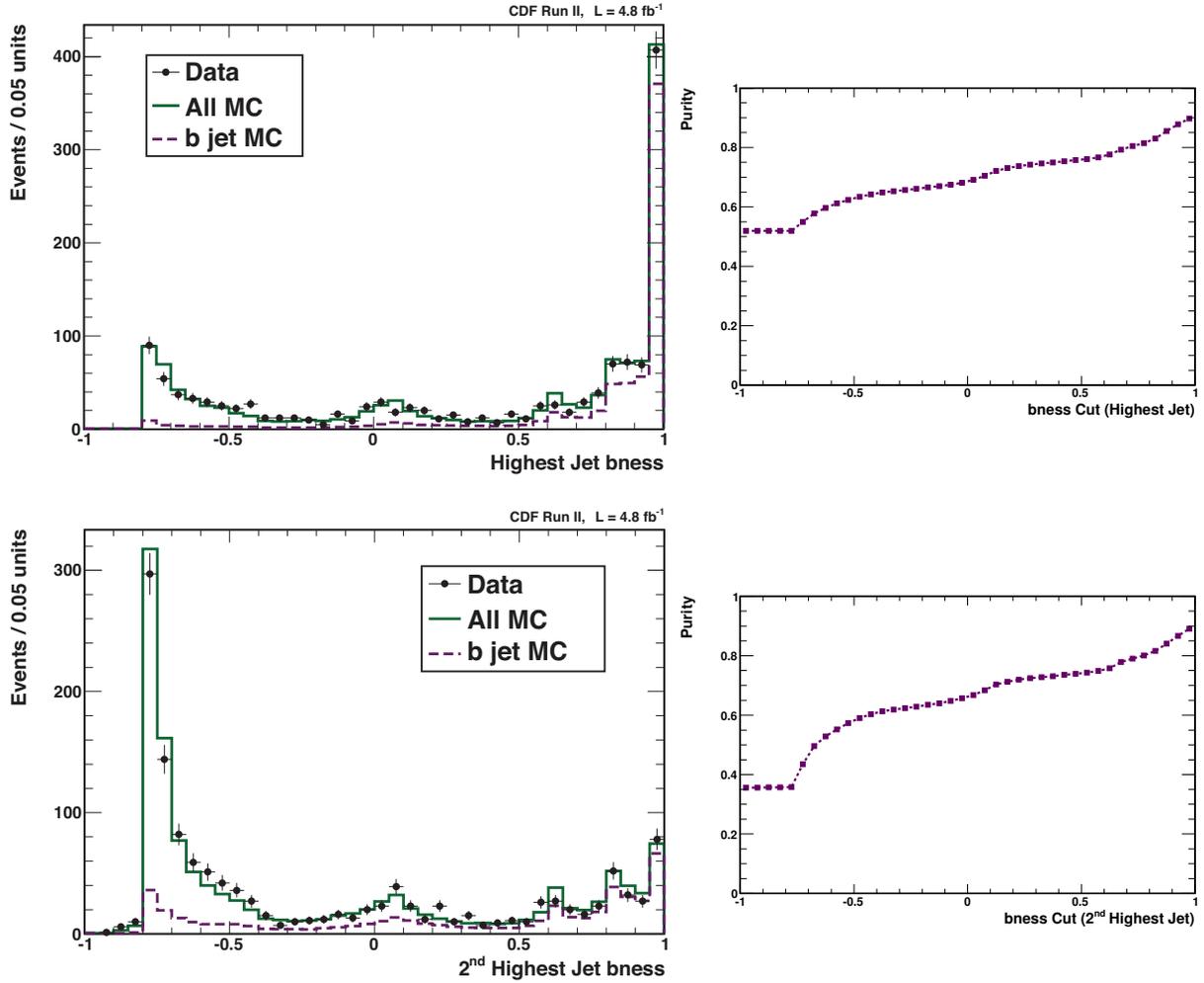
Figure 4.13: *Top Left*: A comparison of the highest jet *b*ness in data (black points) and MC (green solid line) in the $t\bar{t}$ lepton + jets sample, with the portion of the MC jets matched to *b* quarks (purple dashed line) shown independently. *Top Right*: The *b*-jet purity for a given *b*ness cut on the highest jet *b*ness, as determined from matched jets in the MC. *Bottom Left*: A comparison of the second highest jet *b*ness in data (black points) and MC (green solid line) in the $t\bar{t}$ lepton + jets selection region, with the portion of the MC jets matched to *b* quarks (purple dashed line) shown independently. *Bottom Right*: The *b*-jet purity for a given *b*ness cut on the second highest jet *b*ness, as determined from matched jets in the MC. In these plots, we see a high purity in our chosen sample, which is approximately 55% $t\bar{t}$ events.
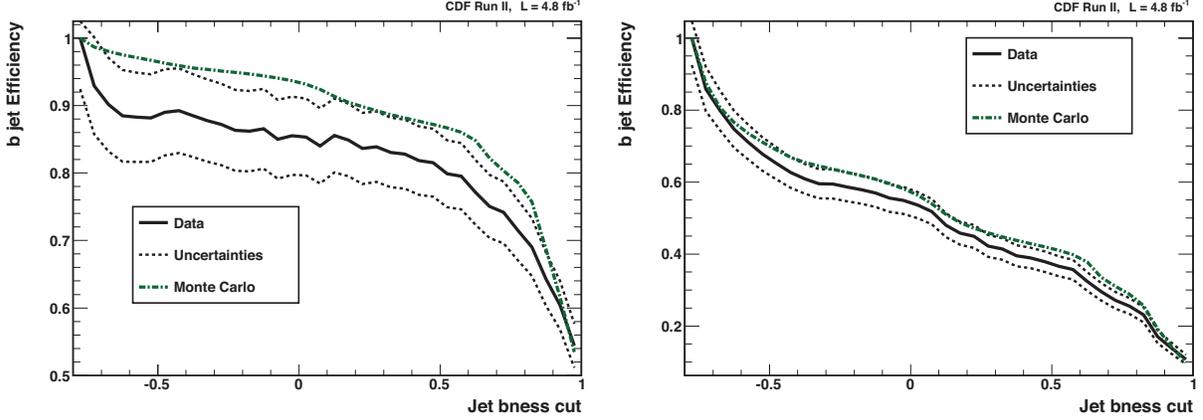
Figure 4.14: The efficiency of a *b*ness cut in data (solid black line, dashed lines represent uncertainty) and Monte Carlo (dot-dashed green line) as a function of the cut on jet *b*ness for the highest (left) and 2nd highest (right) *b*ness jets in an event. We see our simulation typically over-predicts the efficiency measured in data, and thus needs to be corrected.

$(s_e(b) - 1)$ is shown in Fig. 4.15. We see that the relative scale factor and its uncertainty are on the order of 10% or less, comparable to the SecVtx *b* tagger.

As described in §5.4.3, the optimization of the *b* tagging cuts in the diboson analysis resulted in cuts of highest jet *b*ness > 0.85 and 2nd highest jet *b*ness > 0.0. Thus we are interested in the values of the efficiency and mistag rate, scale factor, and uncertainties at those values, shown in Tab. 4.5. These numbers will be used in the diboson analysis.

Quantity	*b*ness Cut	Data	MC	% Difference	% Error
Mistag Rate	0.0	0.0819	0.0720	14%	4.1%
	0.85	0.00997	0.00869	15%	21%
Tag Efficiency	0.0	0.622	0.684	−9.0%	8.7%
	0.85	0.652	0.687	−5.2%	6.2%

Table 4.5: Mistag rates and efficiencies for two jet *b*ness cuts, determined from comparisons of data and MC in the $Z + 1$ jet and $t\bar{t}$ control regions. For the *b*ness cut at 0.85, we consider the highest *b*ness jet, and for the *b*ness cut at 0.0, we consider the 2nd highest *b*ness jet in our $t\bar{t}$ sample.
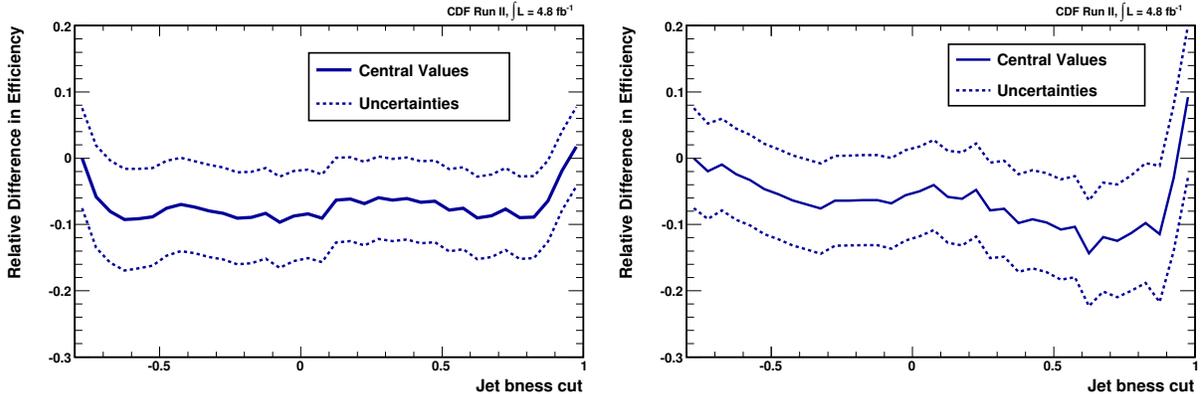
Figure 4.15: The difference in efficiency between data and Monte Carlo (center solid line) and its uncertainty (dashed lines) relative to the efficiency in Monte Carlo as a function of the cut on jet $b$ness for the highest (left) and 2nd highest (right) $b$ness jets in an event. The values of the scale factors and their uncertainties at the relevant $b$ness cuts in this analysis are summarized in Tab. 4.5.

4.6 PERFORMANCE

To assess the performance of the $b$ness tagger, we plot the receiver operating characteristics (ROC) curve in Fig. 4.16 for both the highest $b$ness jet and the 2nd highest $b$ness jet. These curves show the non-b jet rejection versus the b-jet efficiency for a range of $b$ness cuts.

In order to compare the performance of the $b$ness tagger to SecVtx, we compare the efficiency and mistag rates for both taggers using the two highest $b$ness jets in the $t\bar{t}$ MC simulation. The "tight" SecVtx tagger operating point on this sample of jets has an efficiency of 0.59 and a mistag rate of 0.052, while the "loose" operating point has an efficiency of 0.68 and a mistag rate of 0.088. For the highest jet $b$ness > 0.85 cut, we have a efficiency near the loose-tag efficiency (0.69), but a lower mistag rate (0.009) than even the tight SecVtx tag; for the 2nd highest jet $b$ness > 0.0 cut, we have a similarly high efficiency (0.68) and a mistag rate comparable to the loose SecVtx tag (0.082). Although
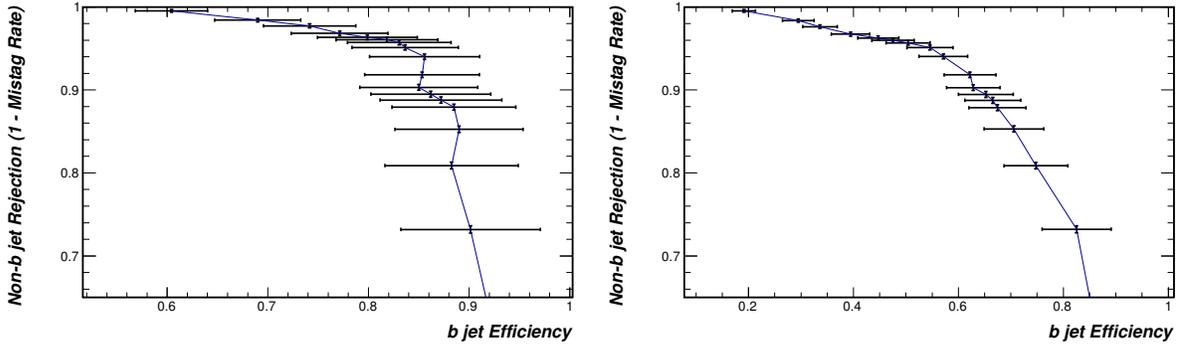
Figure 4.16: Plots of the non-b-jet rejection versus the b-jet efficiency for a range of a cuts on jet $b$ness for the highest (left) and 2nd highest (right) $b$ness jets in an event.
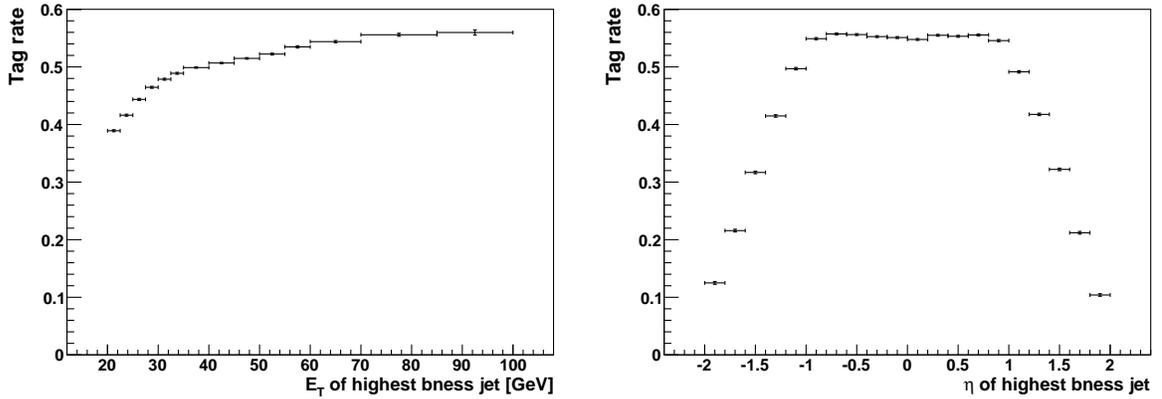
Figure 4.17: Tag performance for the jet with the highest $b$ness value as a function of transverse energy (left) and η (right) for a tagging requirement of $b$ness > 0.85, derived from simulated data.

a direct comparison is difficult, this shows that the performance of our b tagger is roughly comparable to or better than SecVtx.

Finally, we characterize the tagging performance as a function of the transverse energy and pseudorapidity of the jets in simulated data of dijet $b\bar{b}$ events. Fig. 4.17 shows the b tag efficiency for the highest $b$ness jet, where a jet is considered tagged if it has $b$ness > 0.85. The tagging efficiency ranges from 38% at low transverse energy to more than 50% at higher E_{T}. The efficiency is flat in the central region ($|\eta| < 1.0$) and drops off outside the central tracking system.

CHAPTER 5

ANALYSIS PROCEDURE

In this chapter we discuss the measurement of the cross section of WZ and ZZ events in final states with large $\not{E}_{\mathrm{T}}$ and two or more jets, using b tagging to suppress WW contributions. The $\not{E}_{\mathrm{T}}$ comes from either the $W \to \ell\nu$ or $Z \to \nu\bar{\nu}$ decays, while the jets come from the $Z \to q\bar{q}$ decay (or from the W in the case of $WZ \to q\bar{q}'\nu\bar{\nu}$). The final number of signal events is extracted in a simultaneous fit to the dijet invariant mass in two non-overlapping channels: events with at least two b-jet candidates (two-tag channel), and events with fewer than two b-jet candidates (no-tag channel). This increases the acceptance to WZ and ZZ events and was found to significantly improve our sensitivity compared to using only a single channel. We are sensitive to $WZ \to \ell\nu b\bar{b}$ and $ZZ \to \nu\bar{\nu}b\bar{b}$ in the two-tag channel, and $WZ \to \ell\nu q\bar{q}, q\bar{q}'\nu\bar{\nu}$ and $ZZ \to \nu\bar{\nu}q\bar{q}$ in the no-tag channel. This work has been published in [65].

5.1 DATASET AND EVENT SELECTION

The data for this analysis were collected with the CDF Run II detector at the Fermilab Tevatron—a $p\bar{p}$ collider operating at a center-of-mass energy of $\sqrt{s} = 1.96\,\mathrm{TeV}$, described in §3. The dataset analyzed corresponds to $5.2\,\mathrm{fb}^{-1}$ (see §5.3) and only includes events where the COT, calorimeter, and silicon systems were marked as operating properly.

Events passing any of the triggers which require large $\not{E}_T$ are included in order to maximize signal acceptance (see §5.3).

The event selection begins by requiring $\not{E}_T > 50\,\mathrm{GeV}$ and two or more central jets with $E_T > 20\,\mathrm{GeV}$ and $|\eta| < 2$. The $\not{E}_T$ cut selects for neutrinos and rejects generic dijet $2 \to 2$ events. Jets are reconstructed with the JETCLU algorithm [48] with a cone size of $\Delta R = 0.4$, and jet energy scale corrections applied as described in §3.2.2. Jets with an EM fraction (the ratio of the energy deposited in the electromagnetic calorimeter to the energy deposited in the hadronic calorimeter) greater than 90% are rejected as they are likely to have come from electrons mis-identified as jets.

The $b$ness tagger (§4) is run on all jets passing these requirements, and jets are sorted by $b$ness value in descending order since we are most interested in the $Z \to b\bar{b}$ decays. The invariant mass of the highest two $b$ness jets is constructed and required to be within the signal region of 40–$160\,\mathrm{GeV}/c^2$. Events passing the b-tagging requirements of highest jet $b$ness > 0.85 and 2nd highest jet $b$ness > 0 make up the two-tag channel, while events not passing those cuts make up the no-tag channel.

To suppress contributions from the QCD multi-jet background which may produce $\not{E}_T$ due to mis-measurement, we require $\Delta\phi(\vec{\not{E}}_T, \mathrm{jet}) > 0.4$ and $\not{E}_T$-significance > 4 (see §3.3 and §5.2.1).

We apply cuts to suppress contamination from non-collision backgrounds such as from beam halo and cosmic rays. At least one reconstructed vertex formed by charged particle tracks is required. Beam halo events are removed by requiring the event electromagnetic fraction, $E_{\mathrm{EM}}/(E_{\mathrm{EM}} + E_{\mathrm{HAD}})$, to be between 0.3 and 0.85. Cosmic ray events are removed

by requiring the arrival time of jets in the EM and hadronic calorimeters to be consistent with $p\bar{p}$ collisions: within $4.5\,\mathrm{ns}$ for the EM calorimeter and $15\,\mathrm{ns}$ for the hadronic calorimeter.

Lastly, since we select for b jets and allow for two or more jets, $t\bar{t}$ and single t production become significant backgrounds. To reduce these backgrounds, we apply cuts on the number of leptons and jets with $E_\mathrm{T} > 10\,\mathrm{GeV}$. These cuts require the total number of jets and leptons (including crack tracks[1]) to be less than four, the total number of leptons less than two, and individually the number of electrons and number of muons to be less than two. This cocktail of cuts was chosen by hand and increases the rejection efficiency compared to using a smaller number of cuts. The lepton identification used for these cuts is very loose and is described in [80].

The complete list of cuts is shown in Tab. 5.1.

5.2   Background estimation

The remaining backgrounds in this analysis after all selection cuts fall into four main categories:

Electroweak (EWK) : $W(\to \ell\nu)$+jets and $Z(\to \nu\nu, \ell\ell)$+jets processes, where the jets come from initial or final state radiation (ISR/FSR). Their contribution is estimated using Monte Carlo simulations and cross-checked using a γ+jets data set, described in §5.2.2.

Multi-jet (QCD) : Events with generic QCD jet production which result in $\not{E}_\mathrm{T}$ due

---

[1]A "crack track" is a track which passes through an uninstrumented region of the calorimeter, usually at $\eta = 0$ where the two central calorimeter halves meet, or in the gap between the central and plug calorimeters.

Variable	Cut value
$\not{E}_T$	$> 50\,\mathrm{GeV}$
N_{jet}	≥ 2
Jet E_T	$> 20\,\mathrm{GeV}$
Jet $|\eta|$	< 2
Jet EM fraction	< 0.9
m_{jj} (highest $b$ness jets)	$40\text{–}160\,\mathrm{GeV}/c^2$
Highest jet $b$ness (two-tag channel)	> 0.85
2nd highest jet $b$ness (two-tag channel)	> 0
$\not{E}_T$-significance	> 4
$\Delta\phi(\vec{\not{E}}_T, \mathrm{jet})$	> 0.4
N_{vert}	≥ 1
$E_{\mathrm{EM}}/E_{\mathrm{total}}$	$0.3\text{–}0.85$
$|T_{\mathrm{EM}}|$	$< 4.5\,\mathrm{ns}$
$|T_{\mathrm{HAD}}|$	$< 15\,\mathrm{ns}$
$N_{\mathrm{electrons}}$	< 2
N_{muons}	< 2
$N_{\mathrm{ele}} + N_{\mathrm{mu}} + N_{\mathrm{crack\ track}}$	< 2
$N_{\mathrm{jets},\mathrm{E_T}>10} + N_{\mathrm{ele}} + N_{\mathrm{mu}} + N_{\mathrm{crk}}$	< 4

Table 5.1: Complete list of cuts applied in the event selection.

to mis-measurements of the jet momenta. This background is obtained using a data-driven method described in §5.2.1.

t̄t and single t : Top quark pair and single top production. We estimate this background using a Monte Carlo simulation.

WW → ℓνjj : This diboson process is not being measured in this analysis and is treated as background. Since it peaks in the same mass range as WZ and ZZ, it is indistinguishable from signal in the no-tag channel, while b tagging heavily suppresses it in the two-tag channel. This background is evaluated using a Monte Carlo simulation.

Monte Carlo simulations used for signal and background estimates are performed with a combination of PYTHIA [7], ALPGEN [77] and MADGRAPH [81] event generators interfaced with PYTHIA for parton showering. The geometric and kinematic acceptances are obtained using a GEANT-based simulation of the CDF II detector [76]. For the comparison to data, all sample cross sections are normalized to the results of NLO calculations performed with the MCFM v5.4 program [8] and using the CTEQ6M parton distribution functions (PDFs) [78]. The complete list of MC data samples together with their cross sections are listed in Tab. B.1 and their expected number of events in Tab. 5.2.

5.2.1 MULTI-JET BACKGROUND

The $\not{E}_\mathrm{T}$ from QCD multi-jet production is very rarely due to neutrinos. Instead, it often comes from jets which are mis-measured, or when a hadron or photon is lost in an uninstrumented region of the detector. We expect the dominant effect to be jet mis-measurement. Although the large mis-measurement required to pass our high $\not{E}_\mathrm{T}$ cut only occurs in a very small fraction of events, this can be a significant background in a

Sample Description	No-tag channel	Two-tag channel
$Z \rightarrow$ee	14.1	0.3
$Z \rightarrow\mu\mu$	1114	8.8
$Z \rightarrow\tau\tau$	2540	11.9
$Z \rightarrow\nu\nu$	25097	204.0
$W \rightarrow$eν	33117	125.6
$W \rightarrow\mu\nu$	24227	144.0
$W \rightarrow\tau\nu$	61493	219.3
$t\bar{t}$	494.4	153.8
single top	372.3	72.2
WW	2677.5	6.7
WZ	814.5	23.7
WZ ($b\bar{b}$)	58.0	20.5
ZZ	332.3	21.2
ZZ ($b\bar{b}$)	50.1	19.6
$WZ + ZZ$	1146.8	44.9
$WZ + ZZ$ ($b\bar{b}$)	108.1	40.1
Non-QCD background	151146	946.7
QCD estimate	78631	76.3

Table 5.2: Expected contributions of different processes, for $5.2\,\mathrm{fb}^{-1}$. These estimates do not include our systematic corrections (such as for b tagging).

$\not{E}_T$+jets based analysis due the high cross section of multi-jet production. Since these processes can not be simulated accurately, we derive both the normalization and the dijet mass shape of the multi-jet background from data. The final measure of the amount of multi-jet background will be determined from the fit to data.

We cut on $\not{E}_T$-significance and $\Delta\phi(\vec{\not{E}}_T, \text{jet})$ in order to suppress this background. These distributions are shown in Fig. 5.1, which also demonstrates our final ability to model the multi-jet background.
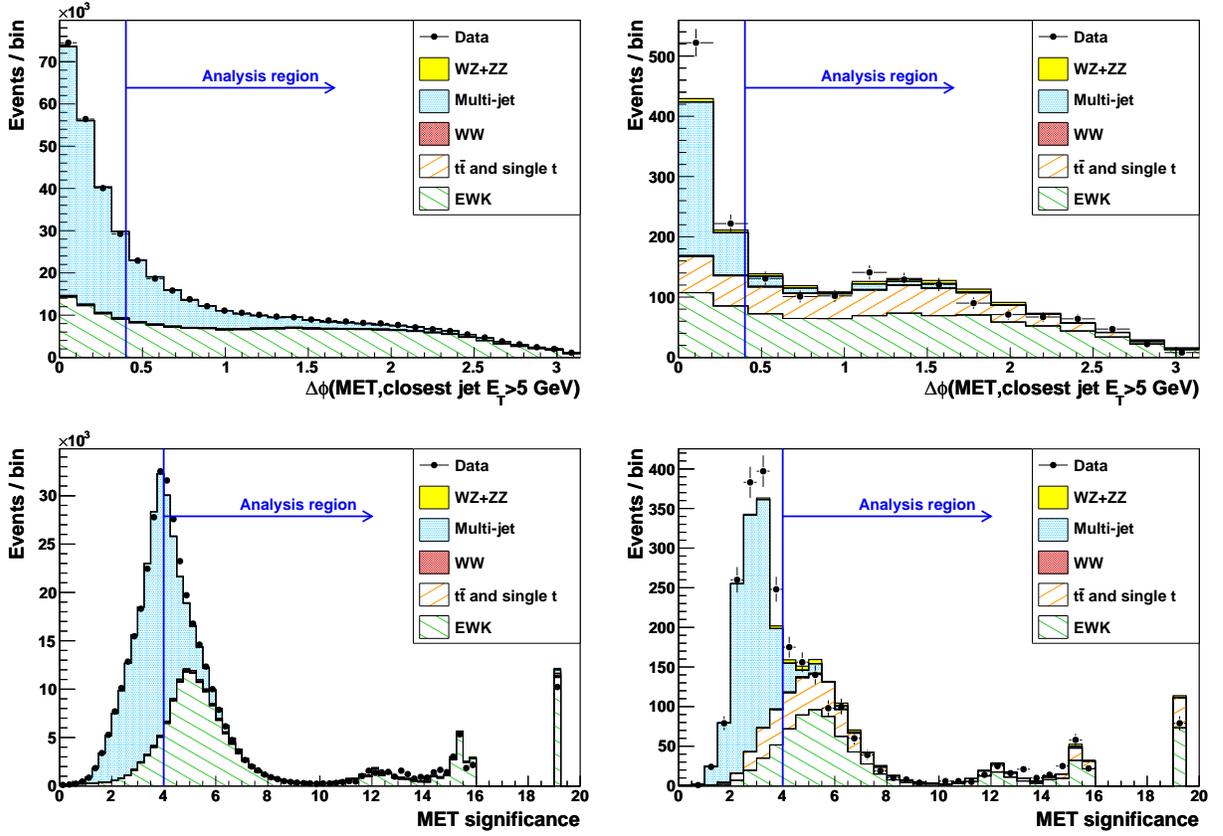
Figure 5.1: *Left*: no-tag region. *Right*: two-tag region. *Top row*: Minimum azimuthal angular separation $\Delta\phi(\vec{\not{E}}_T, \text{jet})$ between all jets with $E_T > 5\,\text{GeV}$ and the $\not{E}_T$, for events that pass all of the analysis cuts except for the $\Delta\phi(\vec{\not{E}}_T, \text{jet})$ cut. The analysis cut is at $\Delta\phi(\vec{\not{E}}_T, \text{jet}) > 0.4$. *Bottom row*: $\not{E}_T$-significance distribution for events that pass all of the analysis cuts except for the $\not{E}_T$-significance cut. The analysis cut is at $\not{E}_T$-significance > 4. The highest bin is the overflow bin.

To estimate the remaining multi-jet background contribution, we construct a new variable, $\not{P}_{\mathrm{T}}$, to complement the traditional calorimeter-based $\not{E}_{\mathrm{T}}$. The $\not{P}_{\mathrm{T}}$ is defined as the negative vectorial sum of tracks with $p_{\mathrm{T}} > 0.3\,\mathrm{GeV}/c$. Tracks used in the calculation of $\not{P}_{\mathrm{T}}$ have to pass minimal quality requirements and come within $|z_0| < 4\sigma$ of the primary vertex in the beam line axis. In the case of true $\not{E}_{\mathrm{T}}$ from neutrinos, we expect the azimuthal angle between the $\not{E}_{\mathrm{T}}$ and $\not{P}_{\mathrm{T}}$ to be small. The difference between these two angles is referred to as $\Delta\phi_{\mathrm{MET}}$. The shape of the $\Delta\phi_{\mathrm{MET}}$ distribution for the QCD background is shown in Fig. 5.2. It was obtained from QCD events which pass our selection cuts and simulated with PYTHIA. We see that the shape is slowly falling but still non-zero at high $\Delta\phi_{\mathrm{MET}}$ as expected.

Electroweak backgrounds and diboson signal will dominate at low $\Delta\phi_{\mathrm{MET}}$ due to correctly measured $\not{E}_{\mathrm{T}}$ from neutrinos. To validate the shape of $\Delta\phi_{\mathrm{MET}}$ distribution in events with true $\not{E}_{\mathrm{T}}$, we compare $Z \rightarrow \mu^+\mu^-$ in data and MC. We select $Z + jj$ events with two high p_{T} muons forming the Z mass peak. We also apply all of our event selection cuts. Note that we do not apply the muon correction to $\not{E}_{\mathrm{T}}$ in the $Z \rightarrow \mu^+\mu^-$ sample so that these events mimic those passing our final signal selection. The $\Delta\phi_{\mathrm{MET}}$ distribution in the data and MC $Z \rightarrow \mu^+\mu^-$ events is shown in Fig. 5.3. We see good agreement. By comparing the distributions in Figs. 5.2 and 5.3, we see that, unlike in the case of QCD, there are virtually no $Z \rightarrow \mu^+\mu^-$ events with $\Delta\phi_{\mathrm{MET}} > 1$. This difference is used to obtain the m_{jj} template for QCD events.

To determine the dijet mass shape of the multi-jet background, we subtract from data the dijet mass distribution of all other background predictions obtained with Monte Carlo simulations, in the multi-jet enhanced region of $\Delta\phi_{\mathrm{MET}} > 1$. The normalization of the template obtained this way is then corrected to account for those events with $\Delta\phi_{\mathrm{MET}} \leq 1$. This correction introduces a 7% uncertainty on the normalization of the
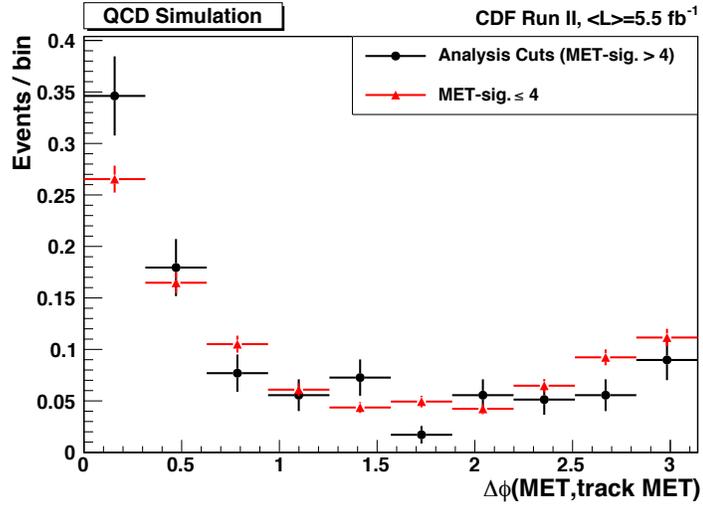
Figure 5.2: Example of the $\Delta\phi_{\mathrm{MET}}$ distribution in Monte Carlo simulated QCD events that pass our selection cuts ($E\!\!\!/_{\mathrm{T}}$-significance > 4, black), and events which fail the $E\!\!\!/_{\mathrm{T}}$-significance cut ($E\!\!\!/_{\mathrm{T}}$-significance < 4, red).
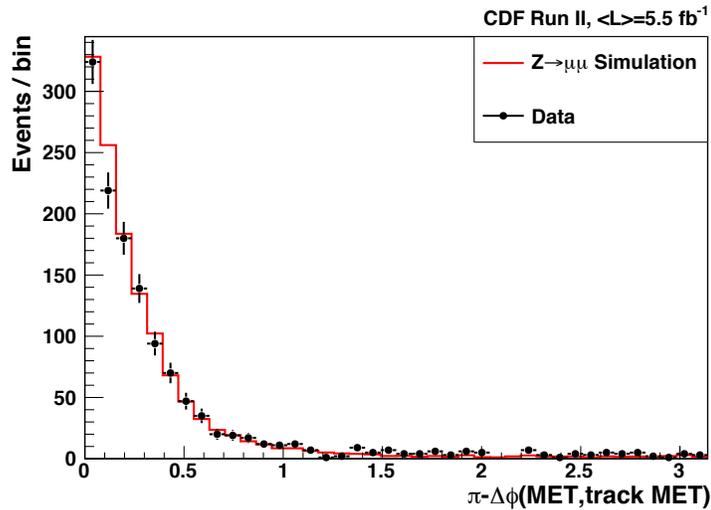

Figure 5.3: The $\Delta\phi_{\mathrm{MET}}$ distribution in MC (red) and data (points) for $Z \to \mu^+\mu^-$ events.

multi-jet background, where the uncertainty was assessed by obtaining the correction factor both in data and in a multi-jet Monte Carlo sample. The uncertainty on the shape of the distribution is estimated by comparing the difference in dijet mass shapes for $\Delta\phi_{\mathrm{MET}} > 1$ and $\Delta\phi_{\mathrm{MET}} < 1$ in a control sample defined by $3 < \not{E}_{\mathrm{T}}$-significance < 4. These dijet mass shapes and their ratio are shown along with the resulting multi-jet m_{jj} shape and its uncertainties in Fig. 5.4.

The procedure described above was applied to both the no-tag and two-tag regions. However, the number of QCD events in the two-tag region is too small to yield a template with adequate statistics (shown in Fig. 5.5). Thus we use the same shape as in the no-tag region, scaled to the predicted number of events in the two-tag channel. We explored two other methods of obtaining a better estimate for the two-tag channel QCD shape. One is to apply looser $b$ness cuts, to get some intermediate shape with usable statistics. Another way is to look at the shapes in W+jets in the no-tag and two-tag channels. These two methods are shown in Fig. 5.5, and we see that their shapes change in opposite ways. Since it is not clear which we trust more, we assume we do not know how the shape in the two-tag channel may change due to $b$ness cuts, and thus we fit for its shape independently of the no-tag channel QCD shape.

The entire procedure to model the QCD background was tested by comparing data to total background predictions for two distributions that are very sensitive to the QCD contribution: $\not{E}_{\mathrm{T}}$-significance and $\Delta\phi(\vec{\not{E}}_{\mathrm{T}}, \mathrm{jet})$. The results of these comparisons for both tag regions is shown in Fig. 5.1. We see good agreement.
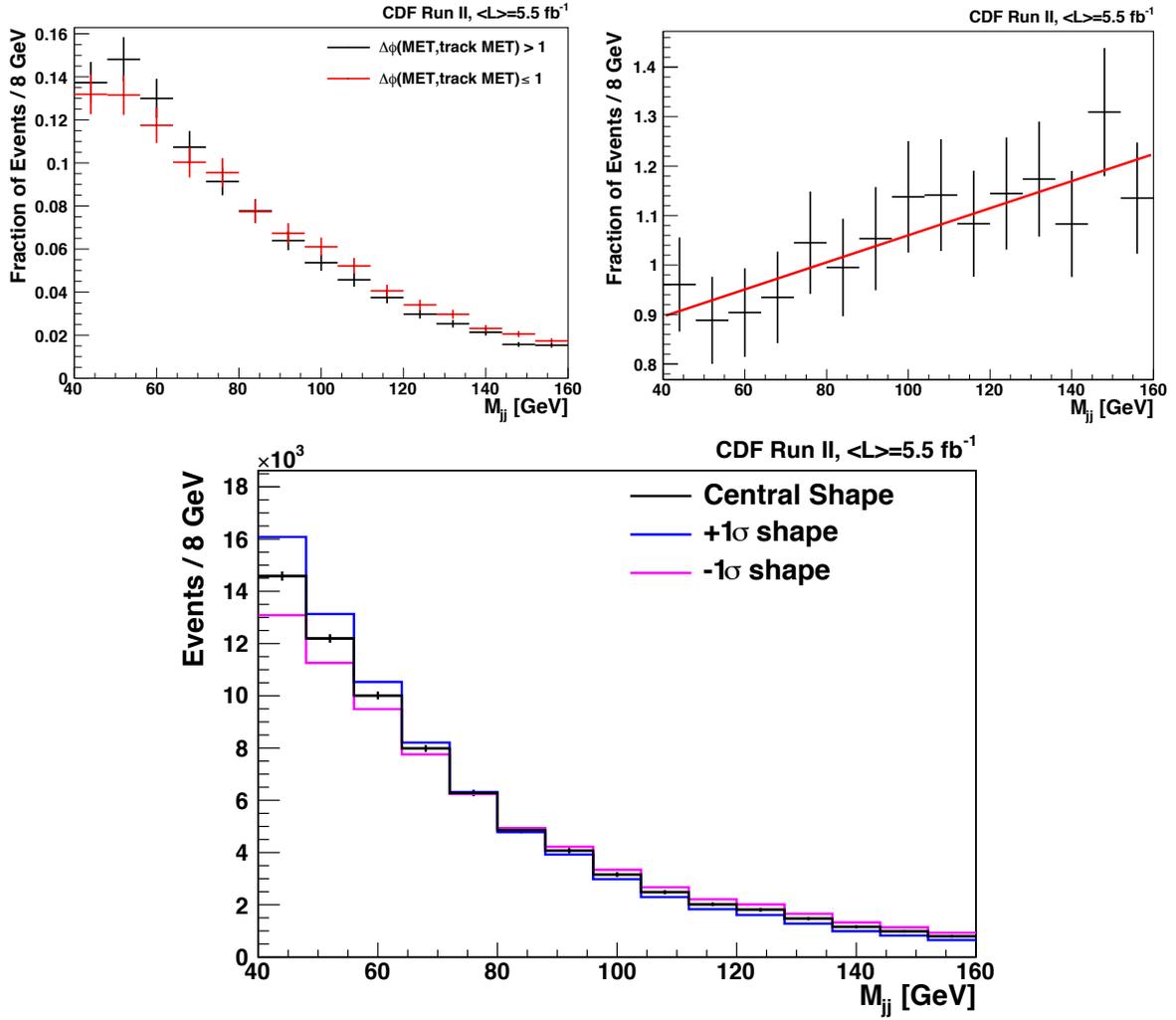
Figure 5.4: *Top-left*: QCD m_{jj} shapes for events with $\Delta\phi_{\mathrm{MET}} < 1$ and $\Delta\phi_{\mathrm{MET}} > 1$ in the $3 < \slashed{E}_{\mathrm{T}}$-significance < 4 region. *Top-right*: Straight line fit to the ratio of those QCD shapes from the $3 < \slashed{E}_{\mathrm{T}}$-significance < 4 region used to obtain the QCD shape uncertainties by multiplying and dividing the central QCD shape by the fit. *Bottom*: The multi-jet background dijet mass template and its corresponding shape uncertainties.
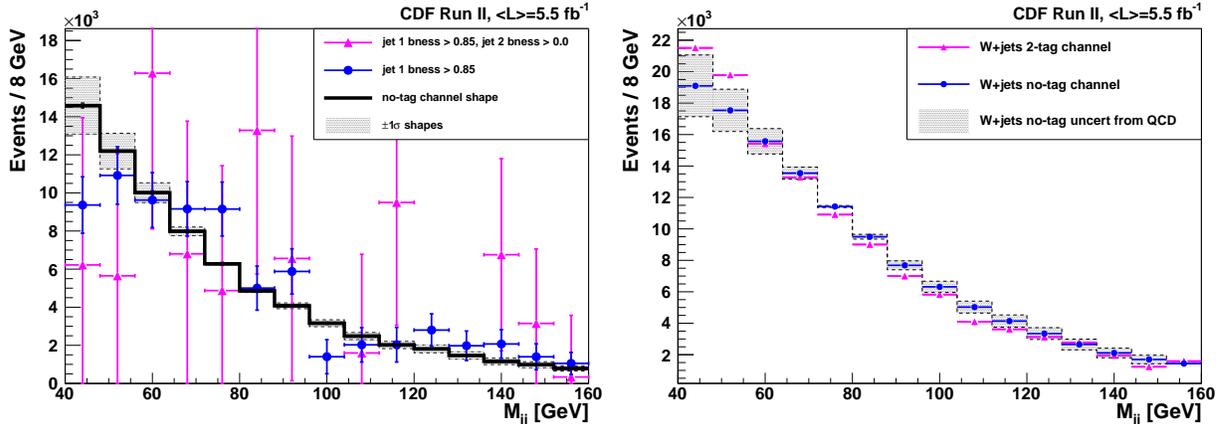
Figure 5.5: *Left*: QCD m_{jj} shapes obtained for the no-tag channel, two-tag channel cuts, and intermediate *b*ness cuts, along with the $\pm 1\sigma$ shapes. The two-tag and intermediate shapes are normalized to the no-tag shape. *Right*: m_{jj} for W+jets in the no-tag and two-tag channels, with the two-tag histogram normalized to the no-tag histogram.

5.2.2 ELECTROWEAK SHAPE SYSTEMATIC

Since the modeling of the EWK V+jets $(V = W, Z)$ is complicated and could potentially have a large uncertainty, we cross-check the MC prediction with a γ+jets data sample to obtain the systematic uncertainty. Note that this uncertainty is only for the shape of the distribution—the rate is allowed to float freely in the fit. This follows the method of the diboson $\not{E}_\mathrm{T}$+jets analysis of [52]. The idea is that jet kinematics in the γ+jets sample should be similar to those in the V+jets sample since the interactions are similar. Although the photon and W/Z are bosons with similar couplings, there are some differences such as the W and Z being massive, and the W being charged. These differences are accounted for by a weighting procedure described below.

Along with differences in the physics, there are also differences in the detector response to γ+jets and V+jets. Specifically, the V+jets enter our selection by decaying to one or more neutrinos, while the photon will not. In order to account for this difference, we use

100

the vectorial sum of the photon E_T plus $\not{E}_\mathrm{T}$ as a stand-in for the usual $\not{E}_\mathrm{T}$ quantity. A few other differences exist in the selection cuts applied to γ+jets versus $\not{E}_\mathrm{T}$+jets data, as shown in Tab. 5.3. These cuts are designed to allow for a data sample dominated by γ+jets events and having adequate statistics.

$\not{E}_\mathrm{T}$+jets	γ+jets
$\not{E}_\mathrm{T} > 50\,\mathrm{GeV}$	$\left\|\vec{\not{E}_\mathrm{T}} + \vec{E}_\mathrm{T\,photon}\right\| > 50\,\mathrm{GeV}$
$\Delta\phi(\vec{\not{E}_\mathrm{T}}, \mathrm{jet}) > 0.4$	$\Delta\phi(\vec{\not{E}_\mathrm{T}} + \vec{E}_\mathrm{T\,photon}, \mathrm{jet}) > 0.4$
$0.3 < \frac{\mathrm{EM}}{E_\mathrm{total}} < 0.85$	$0.3 < \frac{\mathrm{EM}}{E_\mathrm{total}}$
$\not{E}_\mathrm{T}$-significance > 4	$-$
jet $b$ness cuts	$-$
$-$	γ passes standard CDF cuts
$-$	$\Delta R(\mathrm{photon}, \mathrm{jet}) > 0.7$

Table 5.3: List of differences between cuts applied to the $\not{E}_\mathrm{T}$+jets vs. γ+jets sample. A "–" denotes a lack of cut.

To account for these differences between the γ+jets data and the V+jets MC, we normalize all distributions to equal number of events and weight the data bin-by-bin in the dijet mass distributions according to

$$(\gamma + \mathrm{jets})_{\mathrm{weighted}} = (\gamma + \mathrm{jets\ data}) \times \frac{V + \mathrm{jets\ MC}}{\gamma + \mathrm{jets\ MC}}.$$

The weighting ratio describes the difference in physics between γ+jets and V+jets, so that by multiplying the γ+jets data by this, we effectively obtain an estimate of V+jets from data. Since the Monte Carlo simulated events enter only in the ratio, any production difference is taken into account while effects such as detector resolution, PDF uncertainties and modeling of initial- and final-state radiation cancel. Also, note that since the γ+jets data sample will be contaminated with $\gamma + V \to$ jets events peaking in the signal region, their expected contribution obtained from MC simulation is subtracted from the γ+jets distribution.
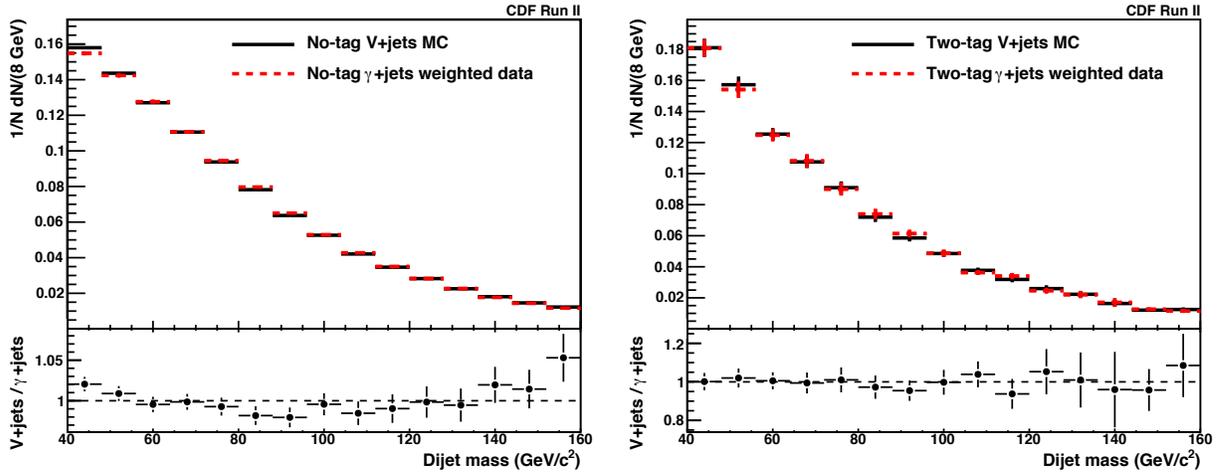
Figure 5.6: Comparison of the γ+jets template with the electroweak MC template in the no-tag (left) and two-tag (right) regions.

After we apply this weighting correction to the γ+jets data, there remains a difference between the corrected γ+jets data and our V+jets simulation, shown in Fig. 5.6, and we take this difference as a systematic uncertainty on the shape of the V+jets background prediction.

5.2.3 bNESS TAGGER CORRECTION AND SYSTEMATIC UNCERTAINTY

As we saw in §4.5, there are differences in the b-tagging efficiency and mistag rates between data and MC. Specifically, these numbers were shown in Tab. 4.5. When applying $b$ness cuts in an analysis to data, a correction should be made to the MC in order to match the efficiency and mistag rates in data. The scale factors in Tab. 4.5 were calculated per jet, and are not the most direct way of applying the correction. Instead we adjust the $b$ness cuts in MC to match the efficiency and mistag rates in data, depending on if the sample is dominated by heavy flavor (tag rate correction) or light flavor (mistag rate correction). The corrected heavy flavor backgrounds are $t\bar{t}$ and single t, while the corrected light flavor backgrounds are just WW. The EWK V+jets sample correction is

taken care of by allowing it to float unconstrained in the fit. Additionally, we determine equivalent cuts on $b$ness in the Monte Carlo that match the $\pm 1\sigma$ uncertainty values on the efficiency and mistag rates. These values are summarized in Tab. 5.4. When applying systematic changes to our Monte Carlo samples, we move the cuts on the jet $b$ness in a completely correlated fashion. The result of this correction produces a shift and uncertainty in the rates of the background processes, and these are included in the main table of systematics, Tab. 5.6.

Note that this method of shifting the $b$ness cuts would also automatically account for any differences in the dijet mass shape due to the $b$ness cuts. However, we found this shape difference to be negligible in our case.

Sample	$b$ness Cut in Data	Equivalent MC Cut		
		-1σ	Central Value	$+1\sigma$
Non-b Jets	0.0	-0.114	-0.0795	-0.052
	0.85	0.805	0.8325	0.861
b Jets	0.0	0.0275	0.1225	0.2675
	0.85	0.8465	0.876	0.903

Table 5.4: Alternative $b$ness cuts applied to Monte Carlo samples, chosen to match the measured mistag rates and tagging efficiencies in data. The uncertainties are determined using the calculated uncertainties on the mistag rates and tagging efficiencies, shown in Tab. 4.5.

5.3 TRIGGER AND LUMINOSITY

To maximize signal acceptance we include events passing any of a collection of high $\not{E}_\mathrm{T}$ triggers. Most of these triggers require $\not{E}_\mathrm{T} > 45\,\mathrm{GeV}$, though some allow lower $\not{E}_\mathrm{T}$ while including some additional requirement to keep the rate low enough.

In order to measure the efficiency of this combination of triggers, and to ensure that it does not sculpt the dijet mass distribution, we look at a data sample triggered on one or more high p_T muons. We do this because muons register as $\not{E}_T$ in the trigger since the muon correction has not yet been applied. The combined $\not{E}_T$ trigger efficiency can be obtained by dividing the dijet mass and $\not{E}_T$ distributions for events passing the $\not{E}_T$ triggers by the distributions for all events, applying selection cuts similar to our analysis. We then parameterize the trigger turn-on curves via a fit to

$$\epsilon = \frac{c}{1 + e^{\frac{a-x}{b}}}.$$

These plots along with the fit parameters are shown in Fig. 5.7.
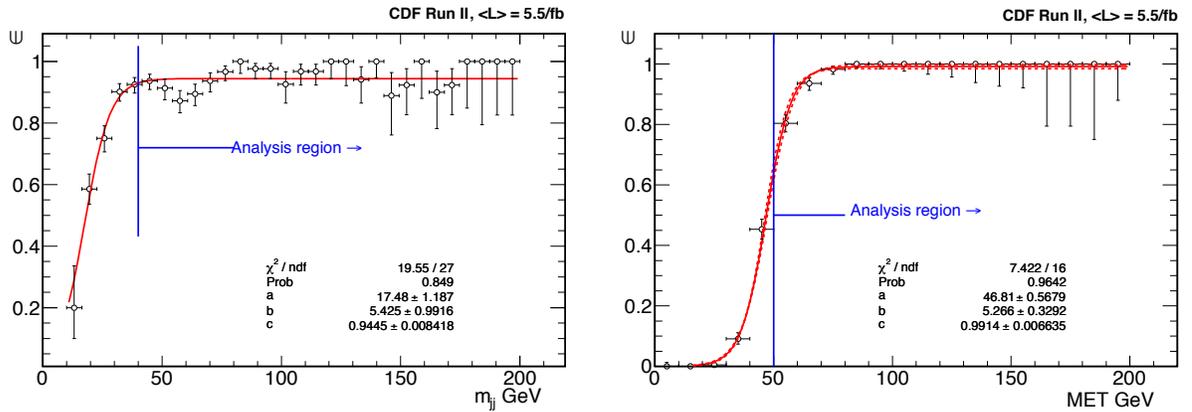

Figure 5.7: Trigger efficiency as a function of m_{jj} (left) and $\not{E}_T$ (right). The dashed lines in the $\not{E}_T$ plot show the statistical uncertainty.

We see that the trigger does not sculpt the dijet mass distribution. However, since the $\not{E}_T$ distribution is not flat at the low end of the analysis cut, we correct the MC simulations to match data by weighting them according to the $\not{E}_T$ turn-on curve.

The product of the luminosity times trigger efficiency is obtained by comparing the $\not{E}_T$-triggered data to a standard candle: $Z \rightarrow \mu\mu$ events with two good muons with high p_T. The advantage of this method is that the $Z \rightarrow \mu\mu$ cross section is accurately known,

easy to measure due to a simple selection with negligible backgrounds, and because the high p_T muon trigger is well understood. The selection requires the muons to have $p_T > 20\,\text{GeV}/c$ and form an invariant muon dimass consistent with the Z mass. This is on top of our base analysis cuts.

To include the effects of acceptance (the fraction of events not lost due to falling outside of the instrumented region of space), efficiency (the fraction of observed events passing the analysis cuts), and trigger efficiency, Eq. 3.1 is modified to give the number of observed events:

$$N = \sigma \cdot (A \times \epsilon) \cdot \epsilon_t \cdot \mathcal{L}, \tag{5.1}$$

where σ is shorthand for the $\sigma(p\bar{p} \to Z) \times \text{Br}(Z \to \mu\mu)$, ϵ_t is the trigger efficiency, $A \times \epsilon$ is the product of the acceptance and efficiency, and $\mathcal{L}$ is the integrated luminosity.

To do the scaling of the $\not{E}_T$-triggered data to the high p_T CMUP muon triggered data, we note that $\sigma \cdot (A \times \epsilon)$ is the same for both samples so setting them equal,

$$\frac{N_\mu}{\epsilon_\mu \mathcal{L}_\mu} = \frac{N_{\not{E}_T}}{\epsilon_{\not{E}_T} \mathcal{L}_{\not{E}_T}} \quad \Longrightarrow \quad \epsilon_{\not{E}_T} \mathcal{L}_{\not{E}_T} = \frac{N_{\not{E}_T}}{N_\mu/\epsilon_\mu} \mathcal{L}_\mu. \tag{5.2}$$

Since statistics are limited for the case of two CMUP muons, we also include events where one of the muons is a CMX muon. Then Eq. 5.2 becomes

$$(\epsilon_{\text{trig}} \times \mathcal{L})_{\not{E}_T} = \frac{N_{\not{E}_T,\text{CMUP+CMUP}} + N_{\not{E}_T,\text{CMUP+CMX}}}{\underbrace{\frac{N_{\mu,\text{CMUP+CMUP}}}{\epsilon_{\text{double CMUP}}}} + \underbrace{\frac{N_{\mu,\text{CMUP+CMX}}}{\epsilon_{\text{single CMUP}}}}} \times \mathcal{L}_\mu. \tag{5.3}$$

The numerical values are shown in Tab. 5.5, and give us

$$(\epsilon_{\text{trig}} \times \mathcal{L})_{\not{E}_T} = 5209\,\text{pb}^{-1} \pm 2.2\% \text{ (stat)} \pm 6.0\% \text{ (lumi)}$$

$$= 5209\,\text{pb}^{-1} \pm 6.4\% \text{ (combined)}, \tag{5.4}$$

where the 6% luminosity uncertainty is standard at CDF [82].

	CMUP+CMUP	CMUP+CMX
$N_{\not{E}_{\mathrm{T}}}$	215	188
N_μ	180	157
$\mathcal{L}_\mu$	$4673\,\mathrm{pb}^{-1}$	
$\epsilon_{\mu,\mathrm{trig}}$	0.985	0.878

Table 5.5: Observed numbers of events for calculating the luminosity times trigger efficiency in the $\not{E}_{\mathrm{T}}$-triggered dataset.

5.4 SIGNAL EXTRACTION

We extract the number of signal events with a binned maximum likelihood fit to data using the method described in [40, 83]. Histograms for backgrounds and signal are supplied, as well as various systematics as rate and/or shape uncertainties. We perform the fit in two different ways, depending on the measurement: for the main measurement of WZ/ZZ to jets with $\not{E}_{\mathrm{T}}$, we perform a simultaneous fit in the two-tag channel and the no-tag channel, while for measuring $WZ/ZZ \to b\bar{b} + \not{E}_{\mathrm{T}}$ we use only the two-tag channel.

5.4.1 TEMPLATES

The templates supplied to the fitter and the uncertainties on their normalizations are listed below:

EWK (W/Z+jets) : Normalizations are allowed to float in the fit, unconstrained, with no correlation between the two tagging channels.

Multi-jet (QCD) : We use our data-driven estimate, Gaussian constrained with an uncertainty of 7% in the no-tag channel. Because there are very few events in the two-tag channel template, we assign a normalization uncertainty equal to the

statistical uncertainty ($\sqrt{N}/N$, 11%) of the template. The uncertainties in the two channels are treated as uncorrelated.

$\mathbf{t\bar{t}}$ **and single t** : The uncertainties on the theoretical cross sections of these processes are 6% [84] and 11% [85, 86], respectively. We combine these two processes to a single template and treat these uncertainties as uncorrelated, which translates to an uncertainty of 5.8% on the normalization of the no-tag channel template, and 5.4% on the normalization of the two-tag channel template, due to the relative contributions of each process.

WW : We use the NLO cross section and apply a Gaussian constraint to the number of WW events centered on this value with a width equal to the theoretical uncertainty of 6% [8].

WZ/ZZ signal : As this is our signal, its normalization is allowed to float unconstrained in the fit. We assume that each signal process contributes proportionally to its predicted SM cross section: 3.7 pb for WZ and 1.4 pb for ZZ [8] corrected for our selection's acceptance and efficiencies.

Note that in the case of the $WZ/ZZ \to b\bar{b} + \not{E}_\mathrm{T}$ measurement, the signal template is broken in up into events with two b's matched to jets (signal) and the rest (background). These templates are shown in Fig. 5.8 which also demonstrates the effectiveness of b tagging in suppressing the WW background in the two-tag channel.

5.4.2 SYSTEMATIC UNCERTAINTIES

Besides the nuisance parameters in the fit associated with the rate uncertainties as described above, there are parameters for EWK shape, QCD shape, JES, and b-tagging
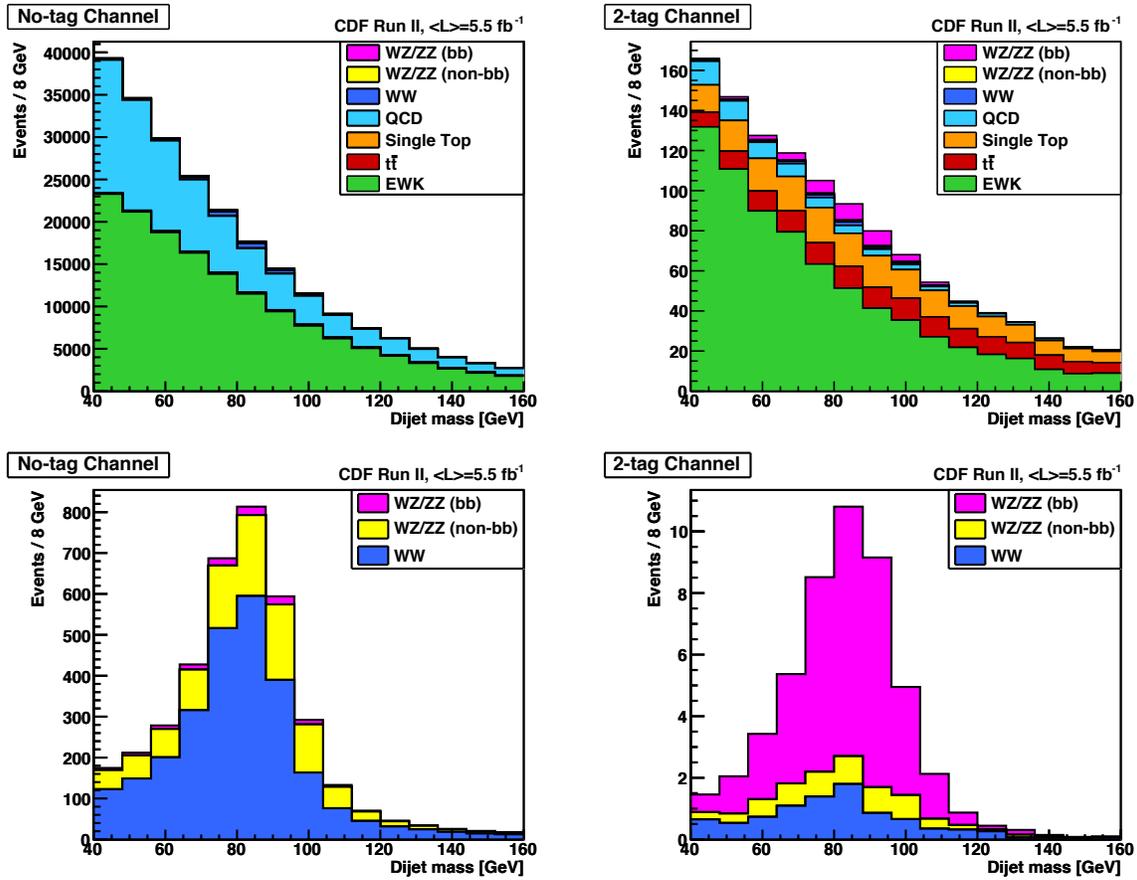
Figure 5.8: *Top row*: templates for the no-tag channel (left) and two-tag channel (right). *Bottom row*: only the diboson samples, stacked.

uncertainties. The shape uncertainties are handled by supplying two additional templates, corresponding to upward or downward fluctuations of the nuisance parameter by one sigma. Template morphing is handled by the MCLIMIT package [40, 83], and we allow extrapolation beyond the supplied shapes.

EWK shape : The central value template used is the average of the EWK MC and the γ+jets template from §5.2.2, with those templates used as the shape uncertainties. We assume no correlation in the shape between fitting channels.

QCD shape : As described in §5.2.1 and assuming no correlation between fitting channels.

JES : The shape and rate systematic uncertainty on WZ/ZZ and WW and rate systematic uncertainty on the top physics templates. The shapes are fit and parameterized as a function of JES so that it can be continuously varied in the fit, Gaussian constrained. The fits are shown in Figs. 5.9 and 5.10, and an example of the difference in signal template shapes due to this effect is shown in Fig. 5.11. The additional correlated rate uncertainty is taken to be 6–8% for all relevant samples.

B-tag efficiency/mistag rate : The uncertainty of the $b$ness cuts (see §5.2.3) is used as a nuisance parameter that can essentially move events between the no-tag and two-tag channels. This uncertainty is applied to our signal and WW templates. In the two-tag region, lowering the $b$ness cuts increases acceptance of signal (WW) by 14.5% (25.9%); raising the $b$ness cuts decreases acceptance of signal (WW) by 13.0% (24.2%).

In addition to these systematics, we apply the following systematic uncertainties on the acceptance to our signal, WW, and top physics templates. Many of these come from the measurement of diboson production with $\not{E}_\mathrm{T}$+jets in [52, 87].

Jet Energy Resolution : Smearing the dijet mass due to energy resolution effects results in a 0.7% uncertainty in the measured cross section.

$\not{E}_\mathrm{T}$ model : There is a 1.0% uncertainty associated with the cuts on $\Delta\phi(\vec{\not{E}}_\mathrm{T}, \mathrm{jet})$ and $\not{E}_\mathrm{T}$-significance.

ISR/FSR : We assign a 2.5% systematic uncertainty due to more/less ISR and FSR [52, 87].

PDF : We assign a 2% uncertainty due to variations in PDFs [52, 87].

Luminosity and trigger efficiency : We assign a 6.4% uncertainty on the cross section due to the uncertainty on the luminosity and trigger efficiency (see §5.3).

All systematic errors are summarized in Tab. 5.6.

5.4.3 SENSITIVITY & OPTIMIZATION

The MCLIMIT code provides a means of obtaining an estimate for the probability of obtaining a measurement at a significance of a given number of σ by generating pseudo-experiments (PEs) and constructing $\Delta\chi^2 = \chi^2_\mathrm{S+B} - \chi^2_\mathrm{B}$ distributions for signal+background and background-only hypotheses, for each of the signal and null hypothesis pseudo-data, just as in the CL_s method described in §2.7.5. To obtain acceptable accuracy, we generate 100 000 PEs. Due to the nature of the fit, optimizing in $S/\sqrt{B}$ does not necessarily translate into a greater sensitivity. For that reason, we optimize based on the probability
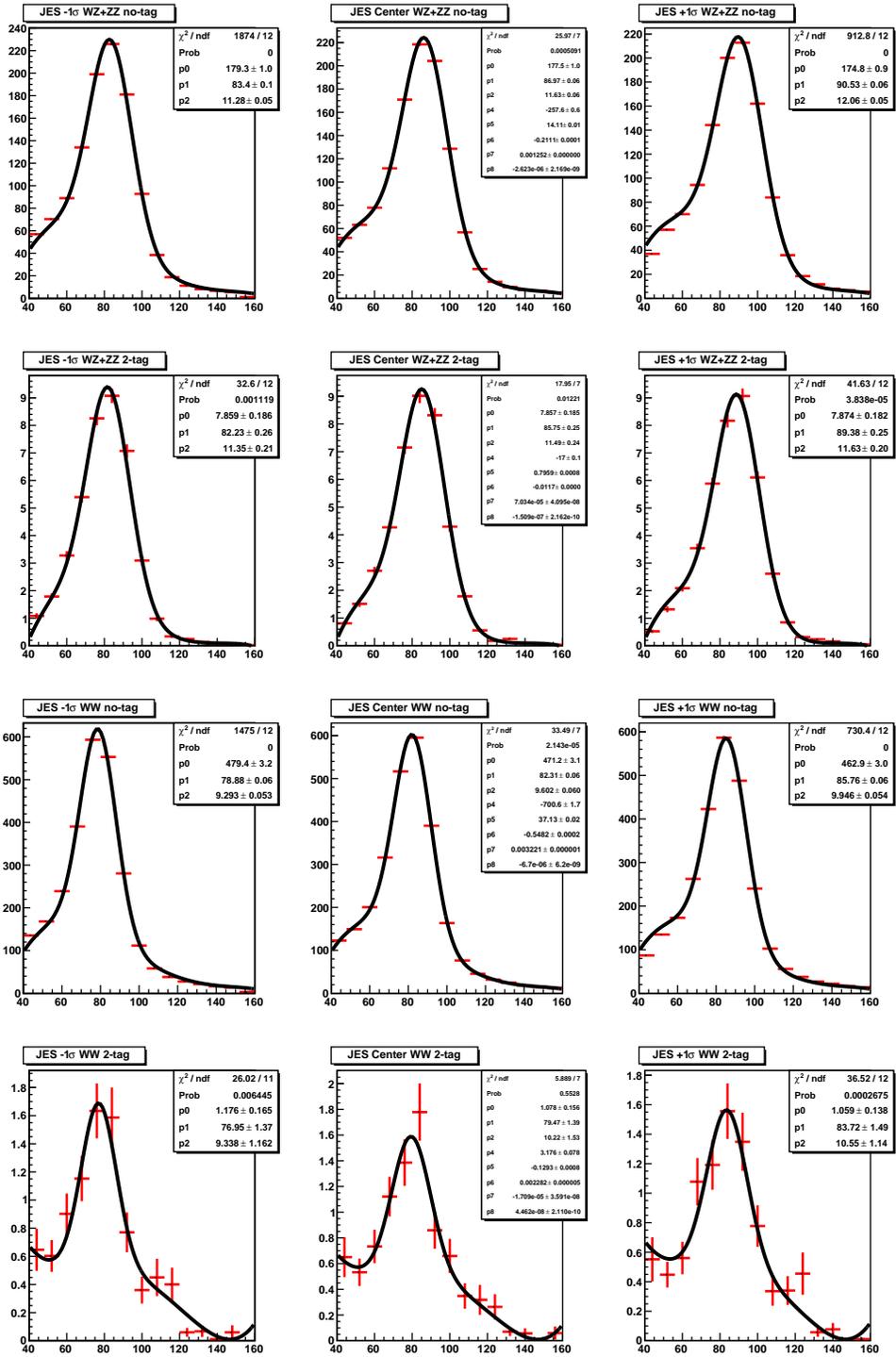
Figure 5.9: Fits to the m_{jj} distribution for calculating a JES shape systematic. The central (0σ) shape is fit to Gaussian distribution on top of a fourth degree polynomial, and the polynomial parameters are then fixed. Next the $\pm 1\sigma$ shapes are fit by allowing the parameters of the Gaussian to vary. The fit parameters are shown in Fig. 5.10.
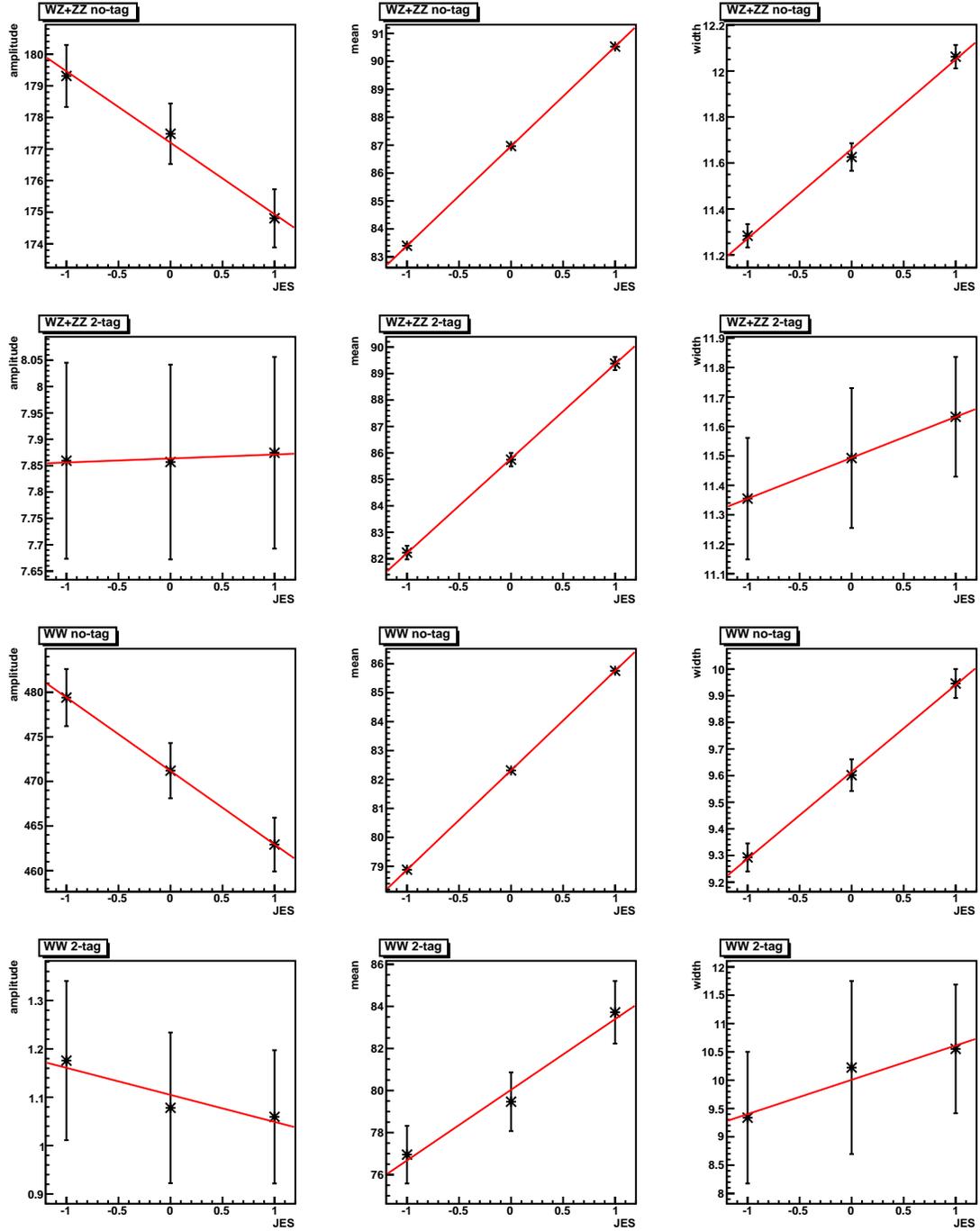
Figure 5.10: The three parameters of the Gaussian function in the fits shown in Fig. 5.9, as a function of JES. We fit these parameters with a straight line, allowing us to parameterize the dijet mass shapes for WZ/ZZ and WW, in both channels, as a function of JES. The error bars are the uncertainties from the fits in Fig. 5.9.
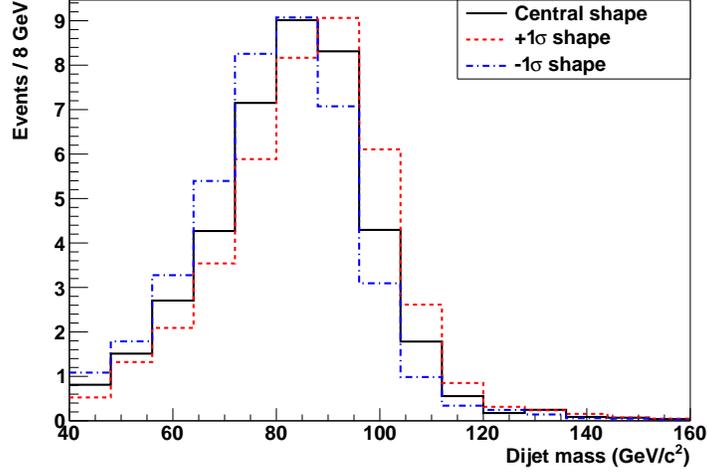
Figure 5.11: Signal template in the two-tag channel with $\pm 1\sigma$ JES corrections applied.

Systematic	channel	WZ/ZZ	WW	$t\bar{t}$ & single t	EWK	Multi-jet
Cross Section (Norm.)	no-tag	float	$\pm 6\%$	$\pm 5.8\%$	float	$\pm 7\%$
	two-tag	float	$\pm 6\%$	$\pm 5.4\%$	float	$\pm 11\%$
EWK Shape	both				yes	
Multi-jet Shape	both					yes
JES Shape/Rate	no-tag	yes/$\pm 7.1\%$	yes/$\pm 7.6\%$	no/$\pm 2.2\%$		
	two-tag	yes/$\pm 6.9\%$	yes/$\pm 7.6\%$	no/$\pm 1.7\%$		
$b$ness cuts (up)	no-tag	$+0.46\%$	$+0.08\%$	$+3.0\%$		
	two-tag	-13.0%	-24.2%	-11.8%		
$b$ness cuts (down)	no-tag	-0.51%	-0.08%	-3.6%		
	two-tag	$+14.5\%$	$+25.9\%$	$+13.8\%$		
Acceptance						
Jet Resolution		$\pm 0.7\%$	$\pm 0.7\%$	$\pm 0.7\%$		
$\not{E}_\mathrm{T}$ Model		$\pm 1.0\%$	$\pm 1.0\%$	$\pm 1.0\%$		
ISR/FSR		$\pm 2.5\%$	$\pm 2.5\%$	$\pm 2.5\%$		
PDF		$\pm 2.0\%$	$\pm 2.0\%$	$\pm 2.0\%$		
Luminosity/Trigger e		$\pm 6.4\%$	$\pm 6.4\%$	$\pm 6.4\%$		

Table 5.6: Summary of systematic uncertainties considered in our analysis. In the two-tag channel fit, where we treat the $WZ/ZZ \rightarrow \not{E}_\mathrm{T}$+light flavor jets as background, we impose a systematic uncertainty of 6% on its cross section.
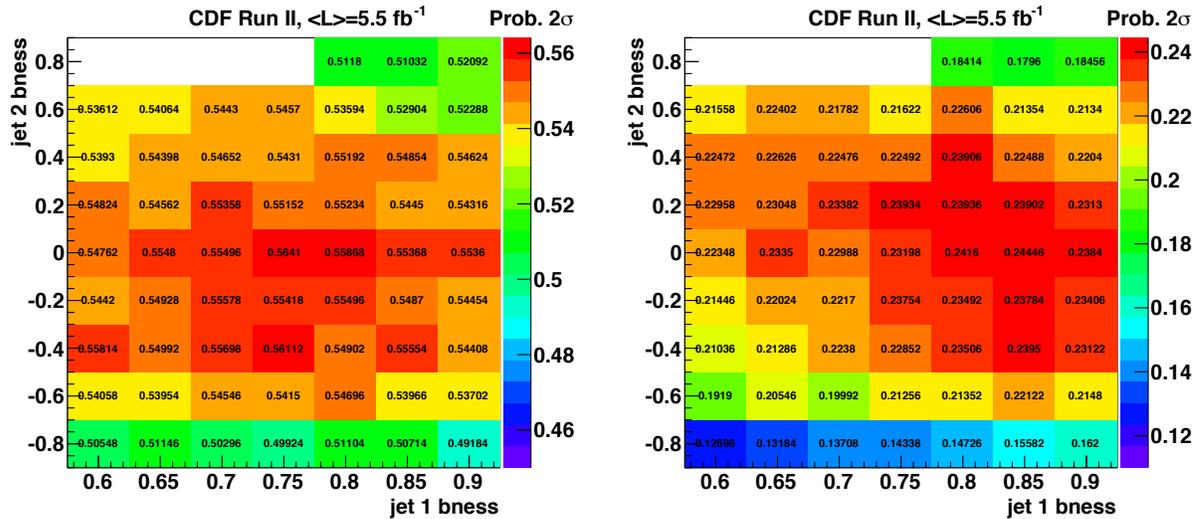
Figure 5.12: Sensitivity scan for optimizing the *b*ness cuts, for the fits to WZ/ZZ in both the no-tag and two-tag channel (left), and for only the two-tag channel (right). Both are roughly maximized for jet 1 *b*ness > 0.85 and jet 2 *b*ness > 0.0.

of 2σ, in particular for the jet *b*ness thresholds. We chose to use the probability of a 2σ measurement since it is more accurate than 3σ for the same number of PEs. Scans of the *b*ness parameter space are shown in Fig. 5.12. Based on this we chose to place the cuts on the highest jet *b*ness and 2nd highest jet *b*ness at 0.85 and 0.0, respectively. The no-tag channel was initially a one-tag channel, containing the events which failed the two-tag requirements but had jet 1 *b*ness > x. We then ran the optimization varying x and found that the maximum sensitivity was obtained with $x = -1$, *i.e.*, no additional *b*ness cut. Lastly, the sensitivity studies show that a fit in two channels is better than using only one channel with no *b* tagging at all. These numbers are shown in Tab. 5.7.

Fit method	Probability of 3σ	Probability of 2σ
2-channels (no-tag and two-tag, all of WZ/ZZ)	8.7%	34%
1-channel (no b tagging, all of WZ/ZZ)	3.0%	19%
1-channel (two-tag only, $WZ/ZZ \to b\bar{b}$)	2.9%	17%

Table 5.7: Sensitivity for the different methods of performing the fits. Note in the first two lines the gain in sensitivity by using b tagging and two channels. Based on this, we only measure the results for the first and the last methods.

114

CHAPTER 6

RESULTS & CONCLUSION

The procedure described in §5 yielded the final fits to the dijet mass distribution shown in Fig. 6.1 for the main measurement ($WZ/ZZ \rightarrow \not{E}_{\mathrm{T}} + \mathrm{jets}$) and in Fig. 6.2 for the measurement in the $b\bar{b}$ channel only. The fit parameters are shown in Tab. 6.1 and 6.2, and the extracted number of events are shown in Tab. 6.3 and 6.4, for the two measurements respectively.

The significance of the measurement of all of WZ/ZZ in the double-channel fit is 1.9σ, and the $\Delta\chi^2$ distributions for our null (background only) and test (signal+background) hypotheses are shown in Fig. 6.3. For the measurement of WZ/ZZ with $b\bar{b}$ in the two-tag channel only, we essentially have no sensitivity, and the $\Delta\chi^2$ distributions are shown in Fig. 6.4.

We construct Feldman-Cousins bands [88] in order to obtain limits on the true cross sections. We perform fits to PEs with the signal cross sections ranging from 0.1 to 3.0 times the standard model value, with a step size of 0.1. For each step value of input cross section, we analyze the distribution of measured cross sections to find the ranges which give 68% and 95% coverage. The resulting confidence bands along with the measured result in data are shown in Fig. 6.5 for the double-channel fit. Using the 1σ bands, the measured result is then $\sigma_{\mathrm{measured}} = 1.15^{+0.7}_{-0.6} \times \sigma_{\mathrm{SM}}$. We set a limit on $\sigma_{\mathrm{measured}}$ at $2.6 \times \sigma_{\mathrm{SM}}$ at 95% CL. Using

$$\sigma_{\mathrm{SM}} = \sigma_{WZ} + \sigma_{ZZ} = 5.08\,\mathrm{pb},$$

we then have

$$\sigma(p\bar{p} \to WZ, ZZ) = 5.8^{+3.6}_{-3.0}\,\text{pb},$$

with a 95% CL limit of $\sigma < 13\,\text{pb}$.

Similarly, the Feldman-Cousins bands for the measurement with $b\bar{b}$ performing the fit only in the two-tag channel is shown in Fig. 6.6. Our measured result is $0.63^{+1.0}_{-0.5} \times \sigma_{\text{SM}}$, which corresponds to a limit of $2.6 \times \sigma_{\text{SM}}$. However, here our signal is composed of only those events in which a Z boson decays to a $b\bar{b}$ pair, thus

$$\sigma_{\text{SM}} \times \text{Br}(Z \to b\bar{b}) = \sigma_{WZ} \times \text{Br}(Z \to b\bar{b}) + 2 \times \sigma_{ZZ} \times \text{Br}(Z \to b\bar{b}) = 0.975\,\text{pb},$$

where we assume the branching ratio of $Z \to b\bar{b}$ is 15.1% [14], and the factor of two on the ZZ term comes from the possibility of *either* Z decaying to $b\bar{b}$. Thus, our measured cross section is

$$\sigma(p\bar{p} \to WZ, ZZ) \times \text{Br}(Z \to b\bar{b}) = 0.61^{+1.0}_{-0.5}\,\text{pb},$$
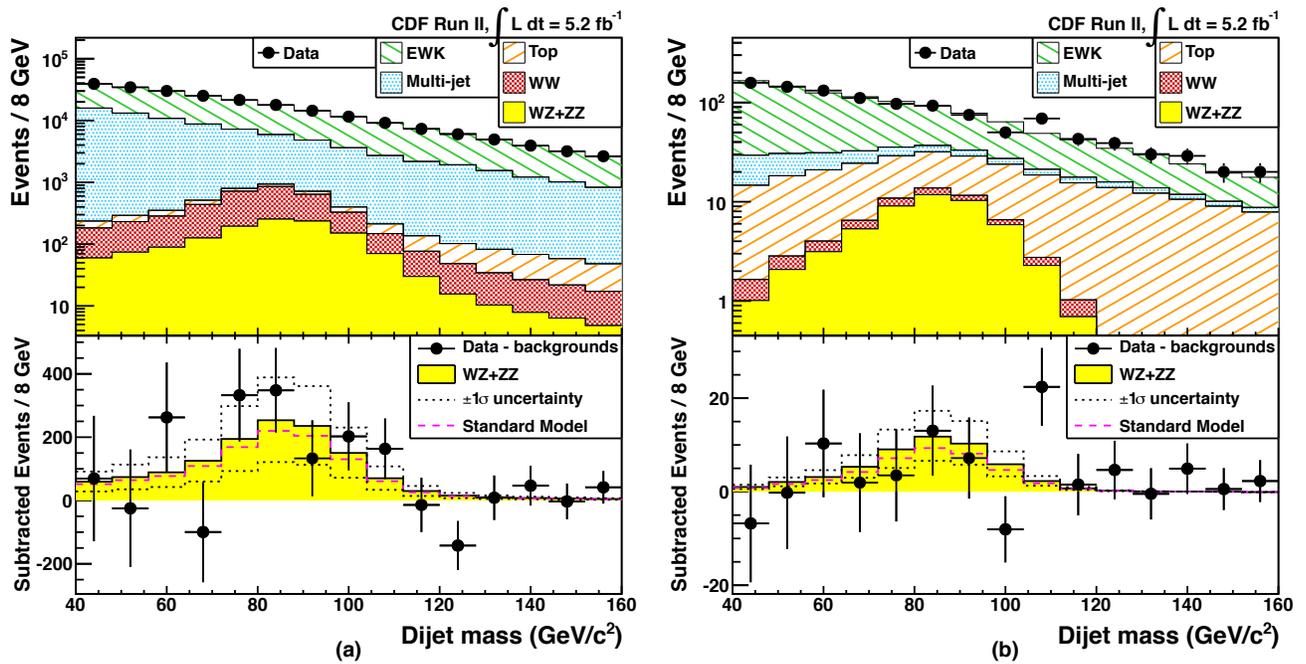
with a 95% CL limit of $\sigma < 2.5\,\text{pb}$.

Figure 6.1: Result of the fit to data for the double fit to all of WZ/ZZ. Left column is the no-tag channel; right column is the two-tag channel. The bottom row shows the data after subtracting all backgrounds.
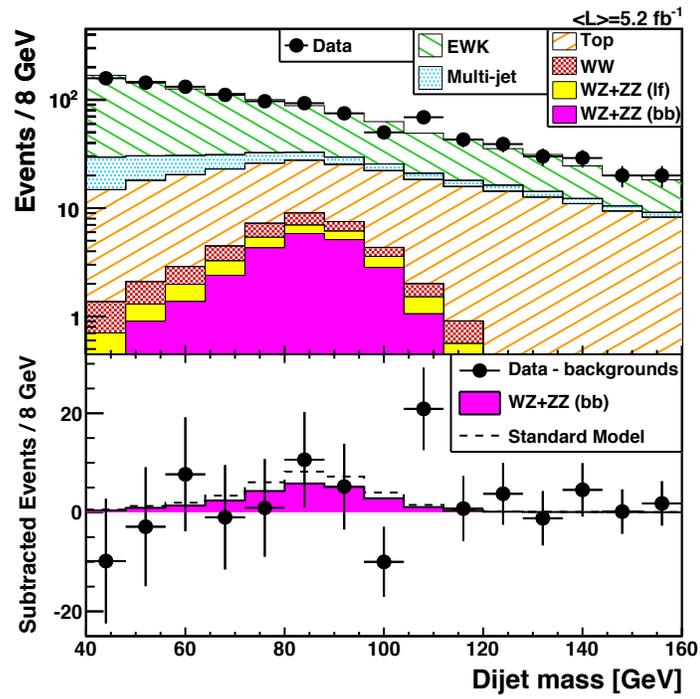
Figure 6.2: Result of the fit to data for the two-tag channel fit to WZ/ZZ with $b\bar{b}$. The bottom row shows the data after subtracting all backgrounds.

Nuisance parameter	Fit value
BG1UNCONSTRAINED	$+0.015926^{+0.03729}_{-0.34613}$
BG2UNCONSTRAINED	$+0.047054^{+0.06927}_{-0.06949}$
BGSHAPE	$-0.266970^{+0.60062}_{-0.60025}$
TOP	$+0.325154^{+1.00925}_{-1.00711}$
JES	$+0.097589^{+0.52163}_{-0.51052}$
WW	$+0.143408^{+1.01847}_{-1.02120}$
BNESS	$-0.653461^{+0.92895}_{-0.90401}$
QCD1	$-0.366006^{+0.88968}_{-0.96205}$
QCD2	$+0.004328^{+1.03209}_{-1.03234}$
QCD1SHAPE	$+0.254887^{+0.14526}_{-0.12045}$
QCD2SHAPE	$-0.254403^{+1.04061}_{-1.01795}$
SIGUNCONSTRAINED	$+0.147487^{+0.62060}_{-0.59754}$

Table 6.1: Nuisance parameter values, in units of standard deviations, from the double-channel fit for WZ/ZZ. Note that parameters with UNCONSTRAINED in the name are fractional changes; *e.g.*, signal is scaled by $(+15^{+62}_{-60})\%$. The 1's and 2's in the names correspond to the no-tag channel (1) and the two-tag channel (2).

Nuisance parameter	Fit value
BG2UNCONSTRAINED	$+0.063335^{+0.07079}_{-0.07116}$
BGSHAPE	$-0.046426^{+0.94905}_{-0.94147}$
TOP	$+0.334496^{+0.97738}_{-0.97748}$
JES	$+0.171451^{+0.98251}_{-0.96598}$
WW	$+0.004356^{+1.00038}_{-0.99960}$
BNESS	$-0.871867^{+0.92547}_{-0.89239}$
QCD2	$-0.003565^{+1.00114}_{-0.99974}$
QCD2SHAPE	$-0.271414^{+1.00560}_{-0.98650}$
SIG_LF	$+0.001477^{+1.00178}_{-0.99822}$
SIGUNCONSTRAINED	$-0.384927^{+0.88378}_{\text{at limit}}$

Table 6.2: Nuisance parameter values, in units of standard deviations, from the two-tag channel fit for WZ/ZZ with $b\bar{b}$. Note that parameters with UNCONSTRAINED in the name are fractional changes; *e.g.*, signal is scaled by -38%. The 1's and 2's in the names correspond to the no-tag channel (1) and the two-tag channel (2).

Process	Fit N_{events} (no-tag)	Fit N_{events} (two-tag)
EWK	149900^{+5600}_{-5200}	749 ± 48
$t\bar{t}$ and single t	898^{+59}_{-61}	217^{+23}_{-27}
Multi-jet	76600^{+4900}_{-5300}	76.3 ± 9.0
WW	2720 ± 200	$10.5^{+2.1}_{-2.3}$
WZ/ZZ	1330^{+710}_{-690}	52^{+24}_{-23}

Table 6.3: Extracted number of events from the double-channel fit for WZ/ZZ, with all systematic uncertainties applied. Each uncertainty is reported to two significant figures, and all event totals are reported to the precision reflected in the uncertainty.

Process(es)	Fit # events (two-tag)
EWK	$761\ ^{+50}_{-51}$
$t\bar{t}$ and single t	$224\ ^{+23}_{-27}$
Multi-jet	76.3 ± 8.7
WW	$11.0\ ^{+2.0}_{-2.2}$
WZ/ZZ (l.f.)	$7.1\ ^{+1.2}_{-1.3}$
WZ/ZZ ($b\bar{b}$)	$24\ ^{+30}_{-24}$

Table 6.4: Fit number of events from the two-tag channel fit for WZ/ZZ with $b\bar{b}$, with all systematics applied.
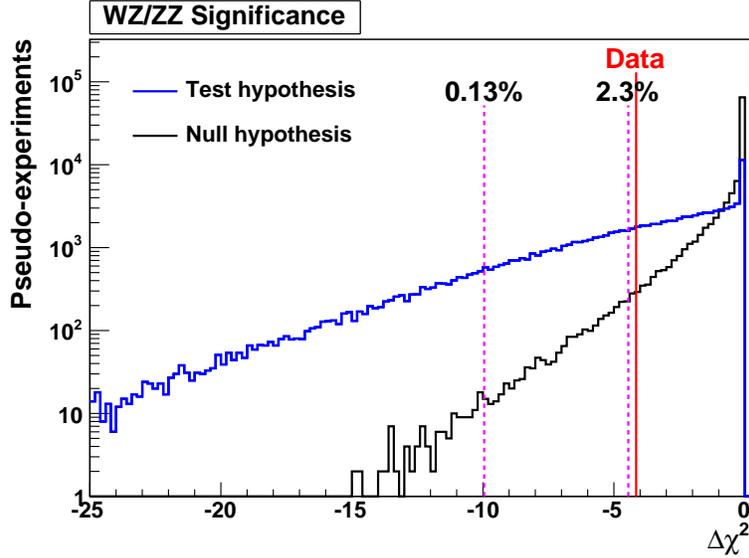

Figure 6.3: The $\Delta\chi^2$ distributions for null and test hypotheses for the double-channel fit to all of WZ/ZZ. In data, $\Delta\chi^2 = -4.15$, and 2.7% of the null hypothesis PEs have a $\Delta\chi^2$ less than this. That is, we observe a p-value of 2.7%, corresponding to a signal significance of 1.9σ where 1.7σ is expected. The dashed lines show the $\Delta\chi^2$ values for which the fraction of the null hypothesis distribution with smaller $\Delta\chi^2$ is 0.135% and 2.28%.
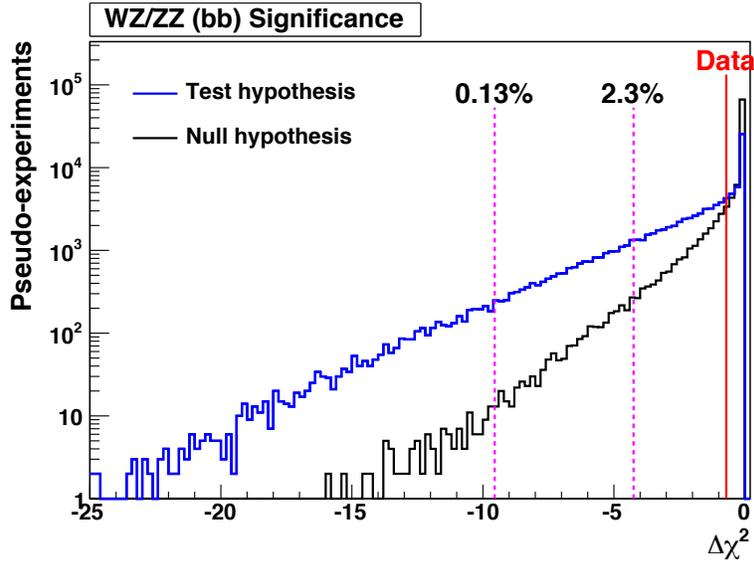
Figure 6.4: The $\Delta\chi^2$ distributions for null and test hypotheses for the two-tag channel fit for WZ/ZZ with $b\bar{b}$. In data, $\Delta\chi^2 = -0.71$, and 21% of the null hypothesis PEs have a $\Delta\chi^2$ less than this. We essentially have no sensitivity in this channel.
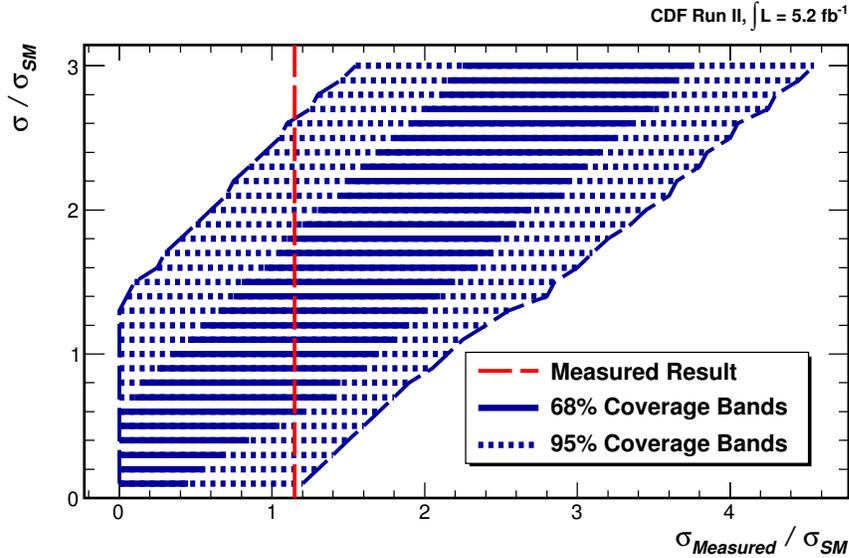

Figure 6.5: Confidence bands showing the expected range of measured cross sections as a function of the true cross section, with 68% CL (blue solid region) and 95% CL (blue dotted region). Our measured result of $1.15^{+0.7}_{-0.6} \times \sigma_{\mathrm{SM}}$ (red dashed vertical line) corresponds to a 95% CL limit at 13 pb ($2.6 \times \sigma_{\mathrm{SM}}$).
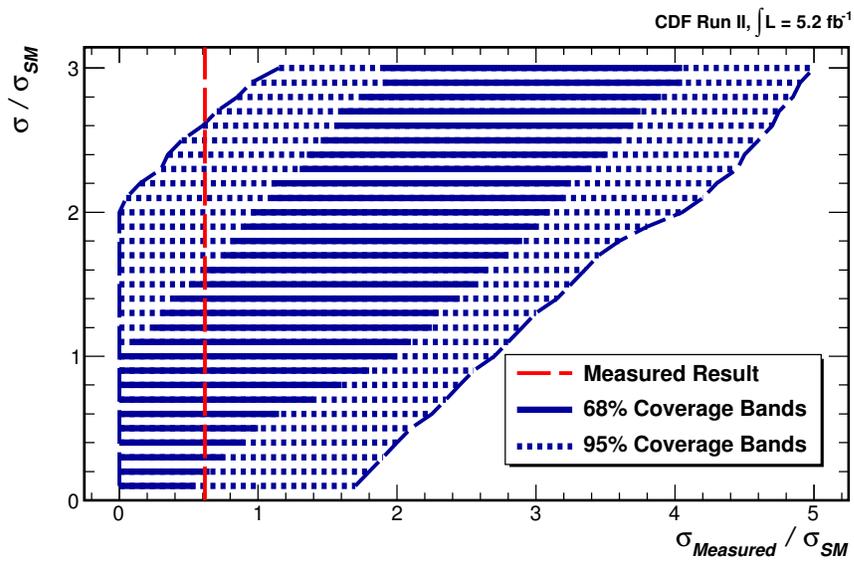
Figure 6.6: Confidence bands for the two-tag channel fit for WZ/ZZ with $b\bar{b}$, showing the expected range of measured cross sections as a function of the true cross section, with 68% CL (blue solid region) and 95% CL (blue dotted region). Our measured result of $0.63^{+1.0}_{-0.5} \times \sigma_{\mathrm{SM}}$ (red dashed vertical line) corresponds to a 95% CL limit at $2.6 \times \sigma_{\mathrm{SM}}$.

6.1 Conclusion

We have measured the cross section for $WZ + ZZ$ production with $\not{E}_\mathrm{T}$ + jets in $5.2\,\mathrm{fb}^{-1}$ of data at a center-of-mass energy of $\sqrt{s} = 1.96\,\mathrm{TeV}$ using the CDF II detector at the Fermilab Tevatron.

We developed a new neural network based b tagger and used it to suppress the peaking WW background. By examining all the tracks associated with jets, this tagger has a larger acceptance than previous neural network based taggers at CDF. Furthermore, the tagger is calibrated using data from Z boson decays and events containing top quark pair production—a novel method which yields small systematic uncertainties on the tagging efficiency and mistag rate. The $WZ + ZZ \to \not{E}_\mathrm{T}$ + jets analysis in this thesis presents the first demonstration of the utility of this tagger. The tagger has also been used in a search for $WZ + ZZ \to \ell^+\ell^- + jj$ [66], and in at least one more diboson analysis currently underway at CDF. Lastly, our tagger has been used in Higgs searches at CDF through its incorporation into the HOBIT tagger [67].

Using this b tagger, we applied the selection cuts and split the data into two categories: b tagged and non-b tagged events. With input from careful modeling of all backgrounds and systematic uncertainties, a simultaneous fit in both channels was performed to extract the number of $WZ + ZZ$ signal events: 1330^{+710}_{-690} in the untagged channel and 52^{+24}_{-23} in the tagged channel. This resulted in a cross section of $\sigma(p\bar{p} \to WZ, ZZ) = 5.8^{+3.6}_{-3.0}\,\mathrm{pb}$, consistent with the Standard Model prediction ($5.1\,\mathrm{pb}$). We also set a 95% CL upper limit on the cross section at $\sigma < 13\,\mathrm{pb}$.

This work has been published in [54] for the b tagger, and [65] for the diboson analysis.

APPENDIX A

CONTROL REGION PLOTS

Comparisons of data and Monte Carlo are shown in both the no-tag channel and the two-tag channel in Figs. A.1, A.2, A.3. The *b*ness distributions are shown in Fig. A.4. Comparisons for $\Delta\phi(\vec{\not{E}}_{\mathrm{T}}, \mathrm{jet})$ and $\not{E}_{\mathrm{T}}$-significance were shown in Fig. 5.1.
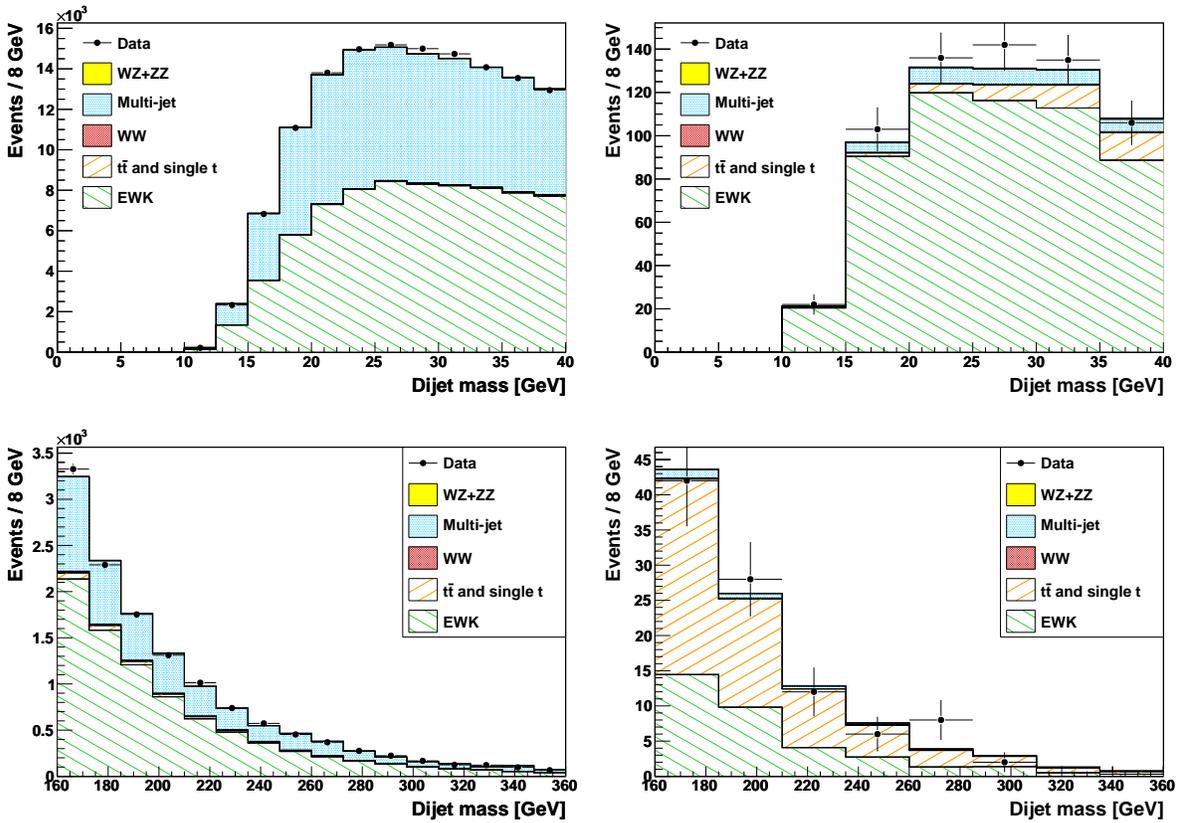

Figure A.1: Dijet mass plots for the no-tag channel (left column) and two-tag channel (right column), outside the signal region. No fit is performed.
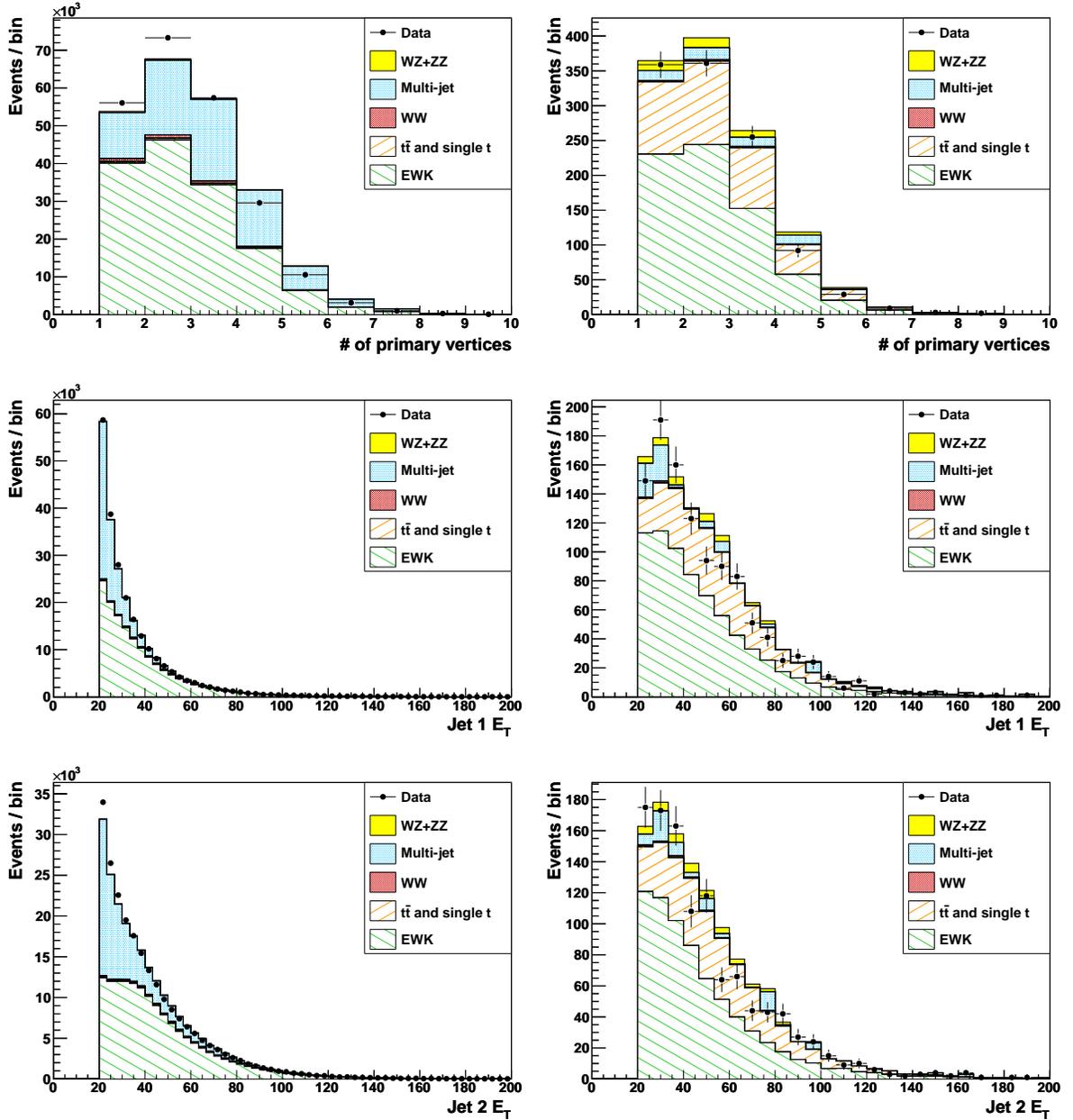
Figure A.2: Comparisons of predictions and data for the no-tag channel (left column) and two-tag channel (right column). The top row shows the N_{vert} distributions, which we have re-weighted for. The middle row is jet 1 E_T and the bottom row is jet 2 E_T.
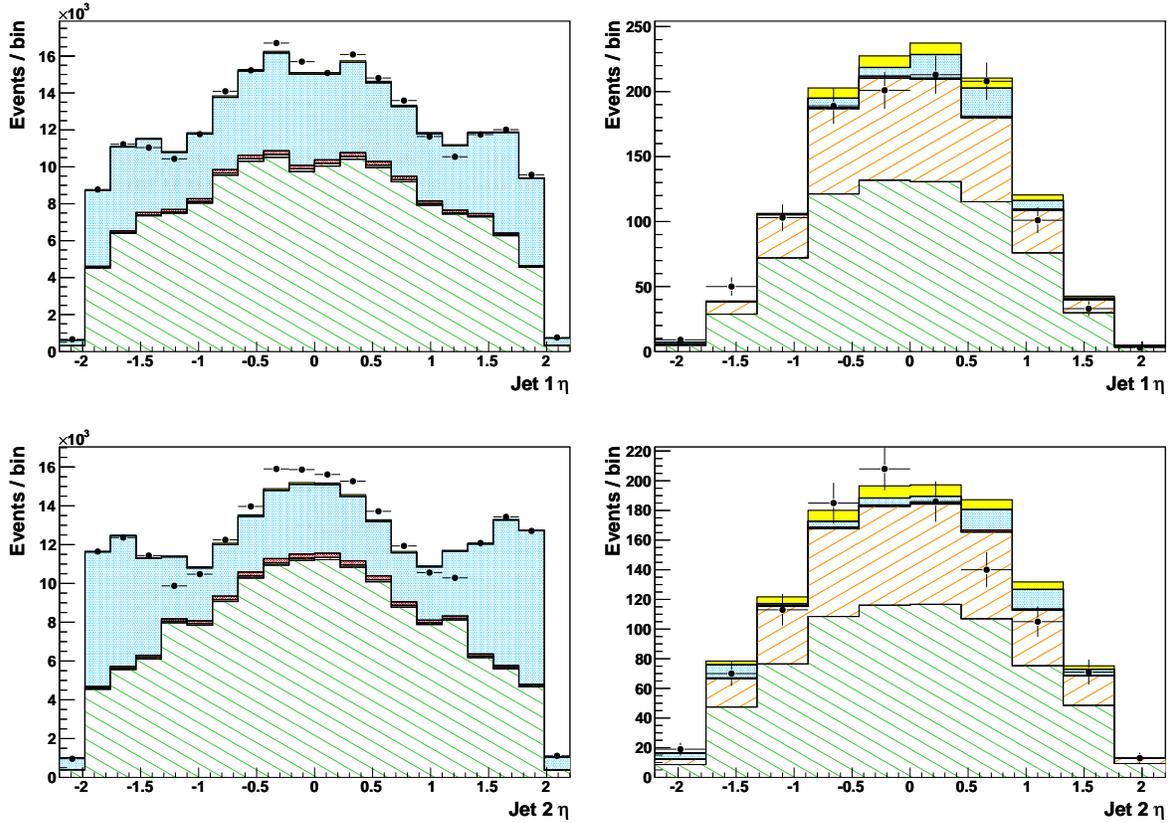
Figure A.3: Comparisons of predictions and data for the no-tag channel (left column) and two-tag channel (right column). The top row is jet 1 η and the bottom row is jet 2 η.
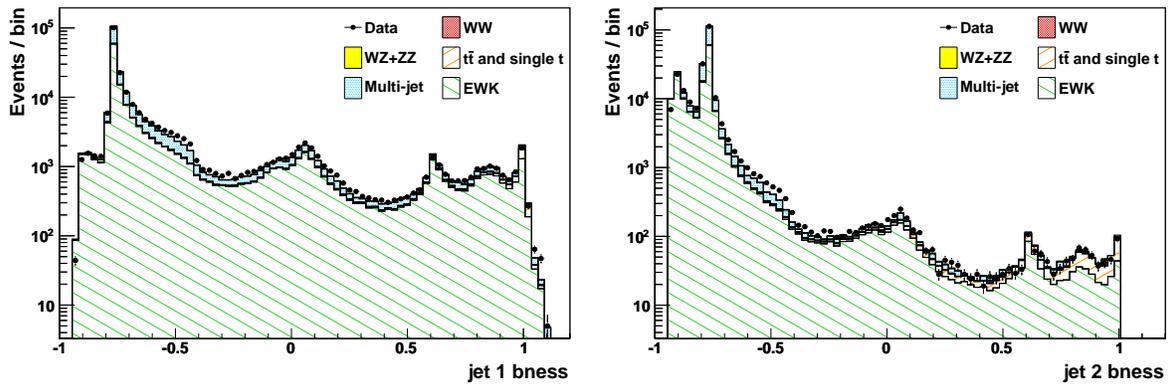

Figure A.4: Comparisons of predictions and data, excluding $b$ness cuts, for jet 1 $b$ness (left) and jet 2 $b$ness (right).

B

LIST OF MC SAMPLES USED

Table B.1: List of MC samples used. The k-factor which accounts for the NLO corrections is included in the cross section whenever necessary (e.g., for all diboson and Z samples). All Pythia samples with Z's have a virtual photon component as well. There are several instances where we list several samples for one process. This comes about for two reasons; firstly, Alpgen requires several n-parton subsamples to describe a process, and secondly, in the case of some Pythia samples we combine all the available samples.

Sample	High Lumi Sample	Cross section	Description	Process
it0sww	ht0sww	12.4	Pythia	WW
it0swz	ht0swz	3.7	Pythia	WZ
it0szz	ht0szz	1.38	Pythia	ZZ
	dhhs1a	977.28	Alpgen (Z+0p)	Z$\to \nu\nu$
	dhhs1b	131.90	Alpgen (Z+1p)	Z$\to \nu\nu$
	dhhs1c	21.25	Alpgen (Z+2p)	Z$\to \nu\nu$
	dhhs1d	3.350	Alpgen (Z+3p)	Z$\to \nu\nu$
	dhhs1e	0.603	Alpgen (Z+$\geq$4p)	Z$\to \nu\nu$
	dhhs1f	6.674	Alpgen (Z+0p)	Z$\to \nu\nu$ (+cc)
	dhhs1g	2.032	Alpgen (Z+1p)	Z$\to \nu\nu$ (+cc)
	dhhs1h	0.652	Alpgen (Z+$\geq$2p)	Z$\to \nu\nu$ (+cc)
	dhhs1i	3.162	Alpgen (Z+0p)	Z$\to \nu\nu$ (+bb)
	dhhs1j	0.822	Alpgen (Z+1p)	Z$\to \nu\nu$ (+bb)
	dhhs1k	0.236	Alpgen (Z+$\geq$2p)	Z$\to \nu\nu$ (+bb)
	tt1s25	7.04	Pythia	ttbar (172.5 GeV)
st0s26	st0s23	0.29	MadEvent+Pythia	single top s-channel
st0s28	st0s25	0.64	MadEvent+Pythia	single top t-channel (NLO)
pt0sw0	ut0s00	1800	Alpgen (W+0p)	W$\to$eν
pt0sw1	ut0s01	225	Alpgen (W+1p)	W$\to$eν
pt0s2w	ut0s02	35.3	Alpgen (W+2p)	W$\to$eν
pt0s3w	ut0s03	5.59	Alpgen (W+3p)	W$\to$eν
				Continued on next page

Sample	High Lumi Sample	Cross section	Description	Process
pt0s4w	ut0s04	1.03	Alpgen (W+≥4p)	W→eν
pt0sw5	ut0s05	1800	Alpgen (W+0p)	W→ $\mu\nu$
pt0sw6	ut0s06	225	Alpgen (W+1p)	W→ $\mu\nu$
pt0s7w	ut0s07	35.3	Alpgen (W+2p)	W→ $\mu\nu$
pt0s8w	ut0s08	5.59	Alpgen (W+3p)	W→ $\mu\nu$
pt0s9w	ut0s09	1.03	Alpgen (W+≥4p)	W→ $\mu\nu$
ut0sw0	ut0s10	1800	Alpgen (W+0p)	W→ $\tau\nu$
ut0sw1	ut0s11	225	Alpgen (W+1p)	W→ $\tau\nu$
ut0s2w	ut0s12	35.3	Alpgen (W+2p)	W→ $\tau\nu$
ut0s3w	ut0s13	5.59	Alpgen (W+3p)	W→ $\tau\nu$
ut0s4w	ut0s14	1.03	Alpgen (W+≥4p)	W→ $\tau\nu$
bt0s0w	bt0s00	2.98	Alpgen (W+0p)	W→eν (+bb)
bt0s1w	bt0s01	0.888	Alpgen (W+1p)	W→eν (+bb)
bt0s2w	bt0s02	0.287	Alpgen (W+≥2p)	W→eν (+bb)
bt0s5w	bt0s05	2.98	Alpgen (W+0p)	W→ $\mu\nu$ (+bb)
bt0s6w	bt0s06	0.888	Alpgen (W+1p)	W→ $\mu\nu$ (+bb)
bt0s7w	bt0s07	0.287	Alpgen (W+≥2p)	W→ $\mu\nu$ (+bb)
dt0s0w	bt0s10	2.98	Alpgen (W+0p)	W→ $\tau\nu$ (+bb)
dt0s1w	bt0s11	0.888	Alpgen (W+1p)	W→ $\tau\nu$ (+bb)
dt0s2w	bt0s12	0.287	Alpgen (W+≥2p)	W→ $\tau\nu$ (+bb)
ct0s0w		5.00	Alpgen (W+0p)	W→eν (+cc)
ct0s1w		1.79	Alpgen (W+1p)	W→eν (+cc)
ct0s2w		0.628	Alpgen (W+≥2p)	W→eν (+cc)
ct0s5w		5.00	Alpgen (W+0p)	W→ $\mu\nu$ (+cc)
ct0s6w		1.79	Alpgen (W+1p)	W→ $\mu\nu$ (+cc)
ct0s7w		0.628	Alpgen (W+≥2p)	W→ $\mu\nu$ (+cc)
ct0sw0		5.00	Alpgen (W+0p)	W→ $\tau\nu$ (+cc)
ct0sw1		1.79	Alpgen (W+1p)	W→ $\tau\nu$ (+cc)
ct0sw2		0.628	Alpgen (W+≥2p)	W→ $\tau\nu$ (+cc)
ze1sad		497	Pythia	Z→ee
ze0sdd		497	Pythia	Z→ee
ze0scd		497	Pythia	Z→ee
ze0sed		497	Pythia	Z→ee
ze0see		497	Pythia	Z→ee
ze0seh		497	Pythia	Z→ee
ze1s9m		497	Pythia	Z→$\mu\mu$
ze0sbm		497	Pythia	Z→$\mu\mu$

Sample	High Lumi Sample	Cross section	Description	Process
ze0sdm		497	Pythia	Z→$\mu\mu$
ze0sem		497	Pythia	Z→$\mu\mu$
ze0sfm		497	Pythia	Z→$\mu\mu$
ze0scm		497	Pythia	Z→$\mu\mu$
ze0sat		497	Pythia	Z→$\tau\tau$
ze0sbt		497	Pythia	Z→$\tau\tau$
zt0sb0		0.511	Alpgen (Z+0p)	Z→ee (+bb)
zt0sb1		0.134	Alpgen (Z+1p)	Z→ee (+bb)
zt0sb2		0.0385	Alpgen (Z+≥2p)	Z→ee (+bb)
zt0sb5		0.511	Alpgen (Z+0p)	Z→ $\mu\mu$ (+bb)
zt0sb6		0.134	Alpgen (Z+1p)	Z→ $\mu\mu$ (+bb)
zt0sb7		0.0385	Alpgen (Z+≥2p)	Z→ $\mu\mu$ (+bb)
zt0sbt		0.625	Alpgen (Z+≥0p)	Z→ $\tau\tau$ (+bb)
zt0sc0		1.08	Alpgen (Z+0p)	Z→ee (+cc)
zt0sc1		0.331	Alpgen (Z+1p)	Z→ee (+cc)
zt0sc2		0.107	Alpgen (Z+≥2p)	Z→ee (+cc)
zt0sc5		1.08	Alpgen (Z+0p)	Z→ $\mu\mu$ (+cc)
zt0sc6		0.331	Alpgen (Z+1p)	Z→ $\mu\mu$ (+cc)
zt0sc7		0.107	Alpgen (Z+≥2p)	Z→ $\mu\mu$ (+cc)
zt0sct		1.28	Alpgen (Z+≥0p)	Z→ $\tau\tau$ (+cc)

Sample	High Lumi Sample	Cross section (pb)	Description	Process
zt0sp0	bt0sz0	221.2	Alpgen (Z+0p)	Z→ee
zt0sp1	bt0sz1	30.24	Alpgen (Z+1p)	Z→ee
zt0szb	bt0sz2	4.844	Alpgen (Z+2p)	Z→ee
zt0s3p	bt0sz3	0.77	Alpgen (Z+3p)	Z→ee
zt0s4p	bt0sz4	0.13888	Alpgen (Z+≥4p)	Z→ee
zt0sp5	bt0sz5	221.2	Alpgen (Z+0p)	Z→$\mu\mu$
zt0sp6	bt0sz6	30.24	Alpgen (Z+1p)	Z→$\mu\mu$
zt0szt	bt0sz7	4.844	Alpgen (Z+2p)	Z→$\mu\mu$
zt0s8p	bt0sz8	0.7672	Alpgen (Z+3p)	Z→$\mu\mu$
zt0s9p	bt0sz9	0.13888	Alpgen (Z+≥4p)	Z→$\mu\mu$
zt0st3	bt0sza	221.2	Alpgen (Z+0p)	Z→$\tau\tau$
zt0st4	bt0szb	30.24	Alpgen (Z+1p)	Z→$\tau\tau$
zt0st2	bt0szc	5.796	Alpgen (Z+≥2p)	Z→$\tau\tau$
xt0s0p	zt0so6	224	Alpgen (Z+0p)	DY→ee (low mass)
xt0s1p	zt0so7	11.746	Alpgen (Z+1p)	DY→ee (low mass)
xt0s2p	zt0so9	2.254	Alpgen (Z+2p)	DY→ee (low mass)
xt0s3p	zt0soa	0.3262	Alpgen (Z+3p)	DY→ee (low mass)
xt0s4p	zt0sob	0.05572	Alpgen (Z+≥4p)	DY→ee (low mass)
xt0s5p	zt0soc	224	Alpgen (Z+0p)	DY→$\mu\mu$ (low mass)
xt0s6p	zt0sod	11.746	Alpgen (Z+1p)	DY→$\mu\mu$ (low mass)
xt0s7p	zt0sof	2.254	Alpgen (Z+2p)	DY→$\mu\mu$ (low mass)
xt0s8p	zt0sog	0.3262	Alpgen (Z+3p)	DY→$\mu\mu$ (low mass)
xt0s9p	zt0soh	0.05572	Alpgen (Z+≥4p)	DY→$\mu\mu$ (low mass)
xt0st0	zt0soi	224	Alpgen (Z+0p)	DY→$\tau\tau$ (low mass)
xt0st1	zt0soj	11.732	Alpgen (Z+1p)	DY→$\tau\tau$ (low mass)
xt0st2	zt0sok	2.548	Alpgen (Z+≥2p)	DY→$\tau\tau$ (low mass)
yt0s0p	zt0sol	5.698	Alpgen (Z+0p)	DY→ee (high mass)
yt0s1p	zt0som	0.9884	Alpgen (Z+1p)	DY→ee (high mass)
yt0s2p	zt0son	0.1638	Alpgen (Z+2p)	DY→ee (high mass)
yt0s3p		0.0259	Alpgen (Z+3p)	DY→ee (high mass)
	zt0sop	0.004662	Alpgen (Z+≥4p)	DY→ee (high mass)
yt0s5p	zt0soq	5.698	Alpgen (Z+0p)	DY→$\mu\mu$ (high mass)
yt0s6p	zt0sor	0.9884	Alpgen (Z+1p)	DY→$\mu\mu$ (high mass)
yt0s7p	zt0sos	0.1638	Alpgen (Z+2p)	DY→$\mu\mu$ (high mass)
yt0s8p	zt0sot	0.0259	Alpgen (Z+3p)	DY→$\mu\mu$ (high mass)
	zt0sou	0.004648	Alpgen (Z+≥4p)	DY→$\mu\mu$ (high mass)
zt0s0h	zt0sov	5.698	Alpgen (Z+0p)	DY→$\tau\tau$ (high mass)
zt0s1h	zt0sow	0.9884	Alpgen (Z+1p)	DY→$\tau\tau$ (high mass)
zt0s2h	zt0sox	0.1638	Alpgen (Z+2p)	DY→$\tau\tau$ (high mass)
zt0s3h	zt0soy	0.0259	Alpgen (Z+3p)	DY→$\tau\tau$ (high mass)
zt0s4h	zt0soz	0.00462	Alpgen (Z+≥4p)	DY→$\tau\tau$ (high mass)

Table B.2: List of Alpgen Z MC samples used in b tagger validation.

Sample	High Lumi Sample	Cross section (pb)	Description	Process
zt0sb0	bt0szd	1.4308	Alpgen (Z+$b\bar{b}$+0p)	Z→ee+$b\bar{b}$
zt0sb1	bt0sze	0.3752	Alpgen (Z+$b\bar{b}$+1p)	Z→ee+$b\bar{b}$
zt0sb2	bt0szf	0.1078	Alpgen (Z+$b\bar{b}$+≥2p)	Z→ee+$b\bar{b}$
zt0sb5	bt0szg	1.4308	Alpgen (Z+$b\bar{b}$+0p)	Z→$\mu\mu$+$b\bar{b}$
zt0sb6	bt0szh	0.3752	Alpgen (Z+$b\bar{b}$+1p)	Z→$\mu\mu$+$b\bar{b}$
zt0sb7	bt0szi	0.1078	Alpgen (Z+$b\bar{b}$+≥2p)	Z→$\mu\mu$+$b\bar{b}$
zt0sbt	bt0szj	1.75	Alpgen (Z+$b\bar{b}$+≥0p)	Z→$\tau\tau$+$b\bar{b}$
xt0sb0		0.8204	Alpgen (Z+$b\bar{b}$+0p)	DY→ee (low mass)
xt0sb1		0.1638	Alpgen (Z+$b\bar{b}$+1p)	DY→ee (low mass)
xt0sb2		0.04424	Alpgen (Z+$b\bar{b}$+≥2p)	DY→ee (low mass)
xt0sb5		0.8204	Alpgen (Z+$b\bar{b}$+0p)	DY→$\mu\mu$ (low mass)
xt0sb6		0.1638	Alpgen (Z+$b\bar{b}$+1p)	DY→$\mu\mu$ (low mass)
xt0sb7		0.04424	Alpgen (Z+$b\bar{b}$+≥2p)	DY→$\mu\mu$ (low mass)
xt0sbt		0.8764	Alpgen (Z+$b\bar{b}$+≥0p)	DY→$\tau\tau$ (low mass)
yt0sb0		0.04032	Alpgen (Z+$b\bar{b}$+0p)	DY→ee (high mass)
yt0sb1		0.01176	Alpgen (Z+$b\bar{b}$+1p)	DY→ee (high mass)
yt0sb2		0.00336	Alpgen (Z+$b\bar{b}$+≥2p)	DY→ee (high mass)
yt0sb5		0.04032	Alpgen (Z+$b\bar{b}$+0p)	DY→$\mu\mu$ (high mass)
yt0sb6		0.01176	Alpgen (Z+$b\bar{b}$+1p)	DY→$\mu\mu$ (high mass)
yt0sb7		0.00336	Alpgen (Z+$b\bar{b}$+≥2p)	DY→$\mu\mu$ (high mass)
yt0sbt		0.05068	Alpgen (Z+$b\bar{b}$+≥0p)	DY→$\tau\tau$ (high mass)

Table B.3: List of Alpgen $Z + b\bar{b}$ MC samples used in b tagger validation. Cross sections are doubled from their Alpgen calculations to match the measured Z+b cross section [79].

APPENDIX C

VICTR TRACK TRIGGER TESTING

In this appendix we describe the current status (Aug. 2012) of our testing of the **V**ertically **I**ntegrated **C**MS **Tr**acker ASIC (VICTR). This track-trigger chip is a prototype for a proposed upgrade to the tracking system in order to incorporate tracking information into the level one trigger at the Compact Muon Solenoid (CMS) experiment at the Large Hadron Collider (LHC) at CERN. Such an upgrade, or a similar one, will be required in order to handle the increased number of tracks when the instantaneous luminosity of the LHC is increased in a future run. The chip is manufactured using vertically integrated circuit (3D-IC) technology, and is designed to attach to silicon strip sensors on both sides in order to veto low p_T tracks. The testing activities documented here focus on verifying and characterizing chip behavior, discovering problems in the chip design or manufacturing process, and to serve as a proof of principle.

C.1 THE CMS DETECTOR AND THE ROLE OF VICTR

The LHC operated at an instantaneous luminosity of about $10^{32}\,\mathrm{cm^{-2}s^{-1}}$ in 2010, and has since increased to $10^{33}\,\mathrm{cm^{-2}s^{-1}}$ while the design peak luminosity is $10^{34}\,\mathrm{cm^{-2}s^{-1}}$. The Super-LHC upgrade [89] is a proposal to upgrade the LHC after about 10 years of running, and which would increase the instantaneous luminosity up to $10^{35}\,\mathrm{cm^{-2}s^{-1}}$. One side-effect of the higher luminosity is a larger number of pp interactions, resulting

in many more tracks, as demonstrated in Fig. C.1. At this rate the current triggering system would become saturated. The most promising way to deal with this is to bring the low p_T track filtering into the detector itself. Additionally the sensors and other components must be able to withstand the increased radiation load.
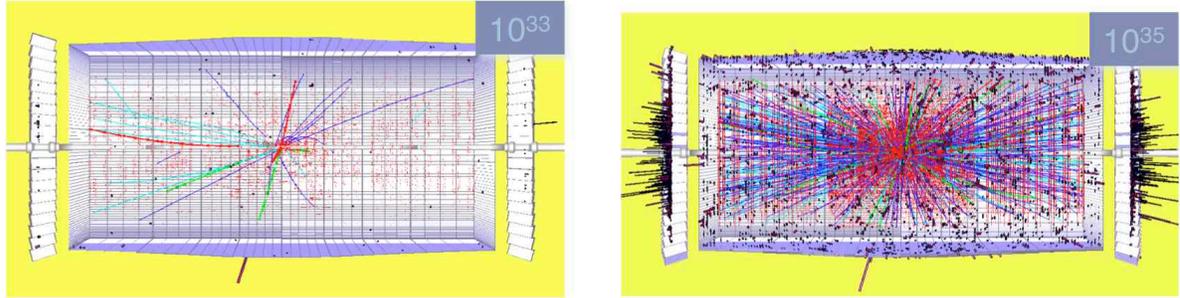

Figure C.1: An example of an event display for an event from a beam with an instantaneous luminosity of $10^{33}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ (left) and the same event with the luminosity increased to $10^{35}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ (right). The additional pp interactions, known as pile-up, results in many more tracks, saturating the triggering system. Figure from [89].

Unlike CDF which has a three level trigger system, the CMS trigger has just two tiers. The first level takes the design bunch-crossing rate of about $30\,\mathrm{MHz}$ down to 50–$75\,\mathrm{kHz}$ using dedicated hardware processors selecting events above p_T and $\not{E}_\mathrm{T}$ requirements based on fast approximate calculations using calorimeter and muon information. The second level then includes all detector information to do the full event reconstruction in software and selecting the potentially interesting events filtered down to a rate of a few hundred hertz sent to storage. VICTR would bring tracking information into the Level 1 trigger.

The present tracking system at CMS consists of an inner silicon pixel detector surrounded by layers of silicon microstrip detectors. VICTR would replace some or all of the microstrip detectors with chips sandwiched between sensor layers, with the chips combining hit information between local and adjacent sensors to veto low p_T tracks.

C.2 VICTR DESIGN

The design of the VICTR chip is described in [90]. The basic structure of a single strip is the sensor and chip sandwich shown in Fig. C.2. The sensor layout is shown in Fig. C.3. The interposer, described in [91], serves to make the electrical connections between the long strip tier and the chip, while also holding the stack together with the sensors separated by the appropriate distance (~ 1 mm).
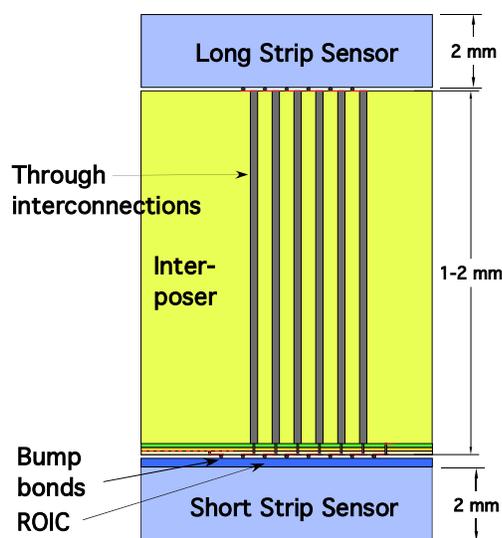

Figure C.2: Illustration of the VICTR stack: the long strip sensors is connected to the chip through a ~ 1 mm interposer bump bonded to the chip which is connected on the other side to the short strip sensors. Figure from [90].

The chip makes use of vertically integrated circuit technology, described in [92, 93], specifically through-silicon-vias (TSVs). These TSVs serve to make the vertical connections between the different layers of the chip. The current design features two chip layers which are bonded together into the 3D chip with Ziptronix Direct Bond Interconnect (DBI) technology. 3D technologies are useful for increasing chip speed (we need to process events within a 20 ns window), lowering capacitance and inductance and hence requiring less power and cooling, and for reducing mass.
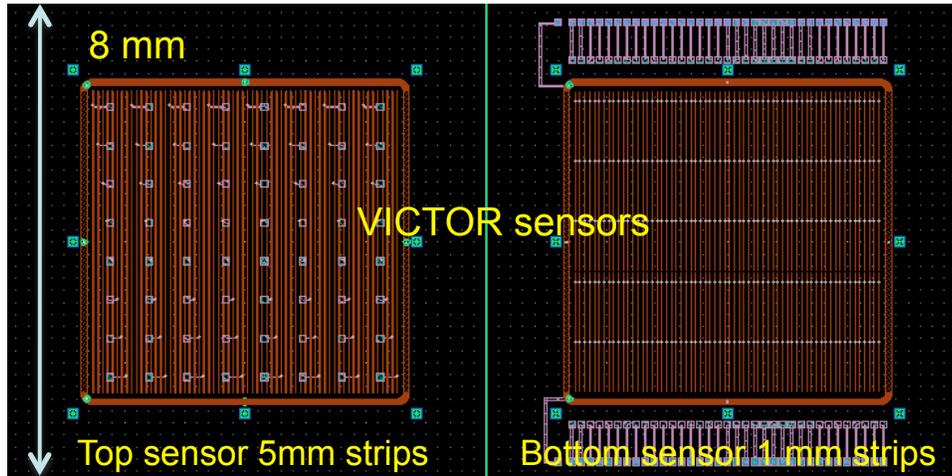
Figure C.3: Layout of the sensors on the two tiers of the VICTR chip. The long strip tier has 64 5 mm strips while the short strip tier has 64×5 1 mm strips.

The current chips available for testing have the two chip tiers bonded together into the 3D chip, but do not have sensors or an interposer connected, though this is planned for the future. A photograph of the chip is shown in Fig. C.4. In order to test the chip without sensors attached, we inject charge by applying a voltage pulse across capacitors connected to the sensor input line in order to mimic a sensor hit. This, along with the programming and readout of the chip, is done with the test stand described below.

C.3 TEST STAND

We initially tested the chip at the test stand at Fermilab. Later, to carry out testing locally at Cornell, we assembled a similar, compatible test stand. The heart of the test stand is a National Instruments FlexRIO system consisting of a NI PXIe-1075 chassis and NI PXIe-8133 controller running LabView 2010 SP1 and housing a NI PXI-4110 programmable power supply and a NI PXIe-7962R FPGA module (Virtex-5, 512 MB RAM) with the NI 6585 LVDS I/O adapter. The LVDS I/O cables connect to a custom
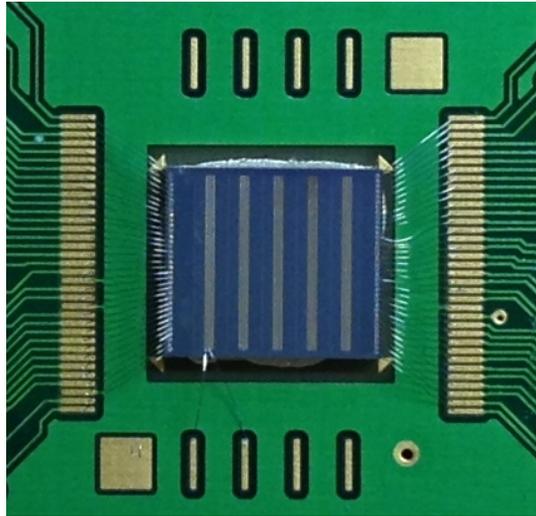
Figure C.4: Photograph of the VICTR chip mounted and wire bonded to the daughterboard. The top surface of the chip shows the sensor input and debug pads for the short strip tier.

PCB motherboard, to which custom PCB daughterboards containing mounted VICTR chips can be plugged in for testing. We also use an Agilent Technologies MSO7104B 1 GHz oscilloscope with 16 digital and 4 analog inputs for measuring various signals, a Tektronix AFG3021B arbitrary function generator for the charge injection, a probe station with passive and active (GGB Picoprobe Model 12C) probes for probing various pads on the chip, and an HP 4192A impedance analyzer for measuring capacitance. The VICTR chip and custom PCB boards were acquired through Fermilab. A photograph of the equipment is shown in Fig. C.5. To ensure stability of the boards while probing, a custom vacuum mount was made and shown in Fig. C.6.

The programming of the chip and I/O for testing is implemented in LabView. The I/O routines to communicate on a FIFO between the controller and chip are run directly on the FPGA. Chip programming commands are constructed in LabView and sent onto the FIFO, and readout is done by sending a readout command and parsing the output
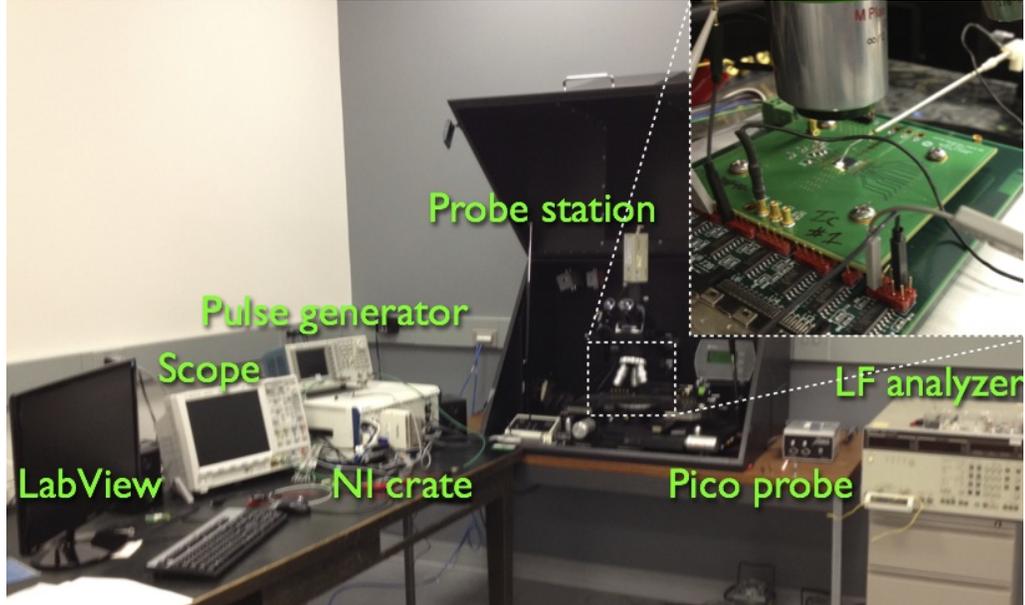
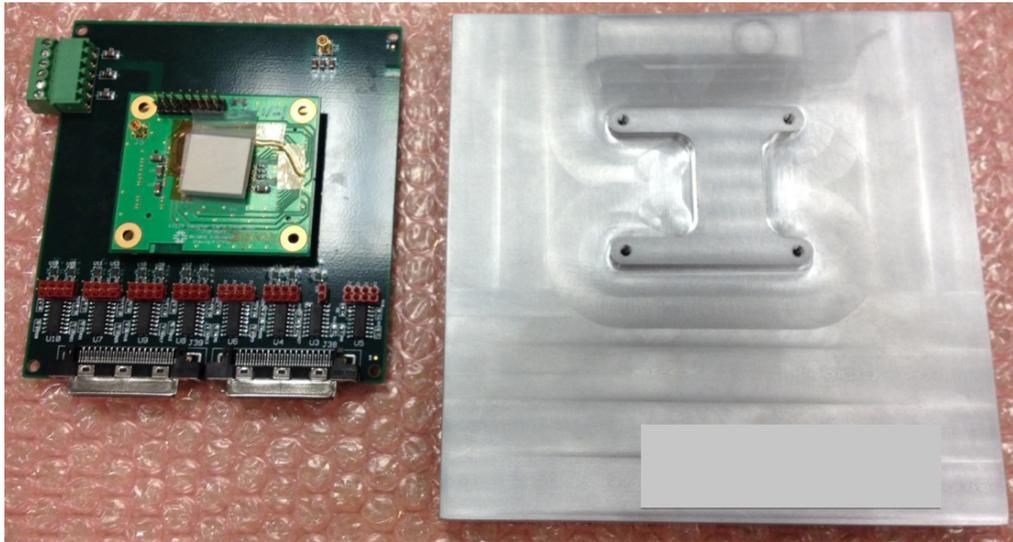**Figure C.5:** Photograph of the VICTR test stand at Cornell.

Figure C.6: Photograph of the PCB boards (left) and vacuum mount (right) for holding the PCB boards in place on the probe station. The PCB boards show the VICTR chip with a protective covering mounted onto the daughterboard, plugged into the motherboard which shows the LVDS I/O connectors on the bottom. The mounting platform has been designed with extra room on the bottom for stress release of the LVDS cables. Screws mount the boards to the aluminum plate which has a smooth and flat bottom surface.

from the FIFO. Various testing modules in LabView have been developed using this architecture.

C.4 Results

In this section we summarize the main results of our contribution to the testing program.

C.4.1 Time walk

The time walk curves show the variation of the delay time between injection of a charge pulse and the subsequent discriminator output (fast OR) from the chip, as a function of the amount of injected charge. The charge is injected as a single square pulse with height a given voltage, and the amount of charge injected is then obtained assuming the design capacitance (4.1 fF) of the charge injection capacitors and $Q = CV$. Such time walk curves are shown in Fig. C.7 for three different operating threshold settings and for a range of front-end currents.

The time walk curves obtained while having only a single strip's output active, for a sampling of strips in extreme geometric locations on the chip are shown in Fig. C.8. We do see a systematic effect which was expected.

Lastly, since the previous data shown were tediously taken by hand, we developed a routine in LabView to automate the data acquisition by looping over injected charges and querying the oscilloscope for the time delay measurement. A comparison of data taken by hand and subsequently taken with the automated acquisition routine is shown in Fig. C.9. The methods are in agreement. This will make future time walk studies on different chips more easily obtained.
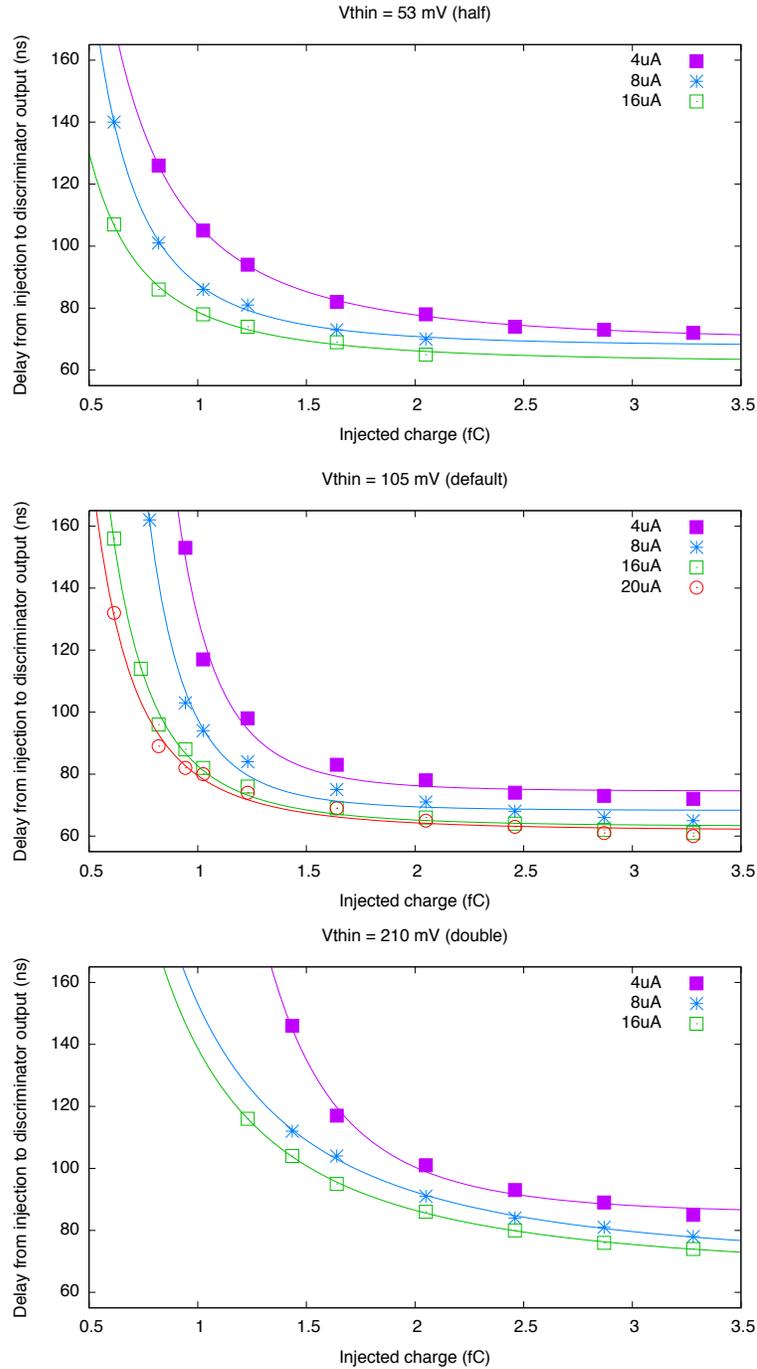
139

Figure C.7: Time walk curves for three threshold voltage settings. This threshold ("Vthin") is the voltage above which a signal is counted as a hit. The different curves in each plot are for different values of front-end bias current. The higher the current, the better the performance (shorter delay). The increasing threshold values shift the curves to the right. Note that in the bottom plot, no more data at smaller charges could be obtained since they are below threshold. The data are fit to curves parameterized by $a + bx^{-c}$.
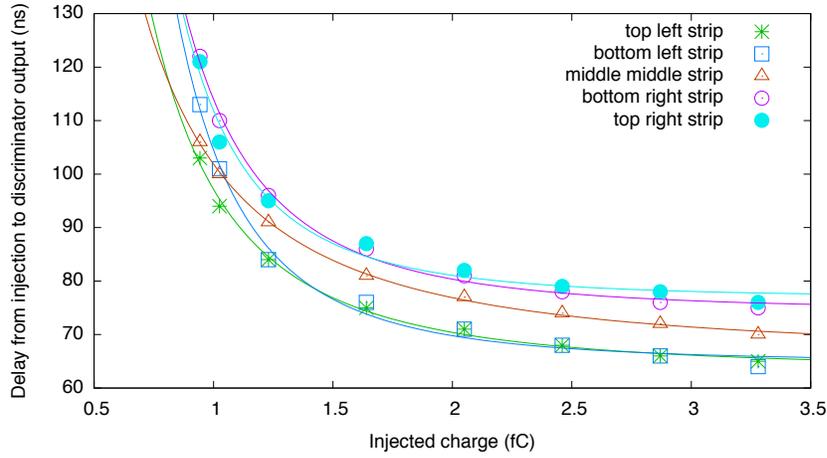
Figure C.8: The time walk curves for different strips on the same chip. The "top left" and "bottom left" correspond to strips in the same strip set (group of five on the short strip tier), whereas "left", "middle", and "right" refer to the long strip direction. We see that curves are consistent in the short strip direction but have a systematic effect in the long strip direction. This is expected since the fast OR output is constructed via a serial chain of OR gates, and it takes longer for the signal to propagate to the end of the chain the further it starts.
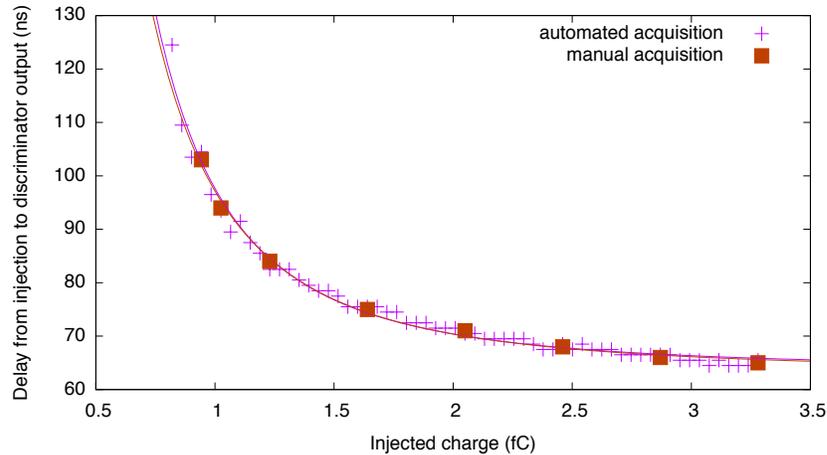
Figure C.9: Comparison of time walk curves obtained manually and by the automated acquisition routine. Fits are done to both sets of data and agree very well.

C.4.2 Turn-on curves

An automated routine for acquiring turn-on curves for individual strips was developed. This routine scans over a range of injection voltages and for each value, injects the charge many times and counts the fraction of hits. Plotting the hit fraction versus voltage gives a turn-on curve as in Fig. C.10. The data are fit to a cumulative Gaussian function with the center giving the measured threshold and the width giving a measure of the amount of noise.
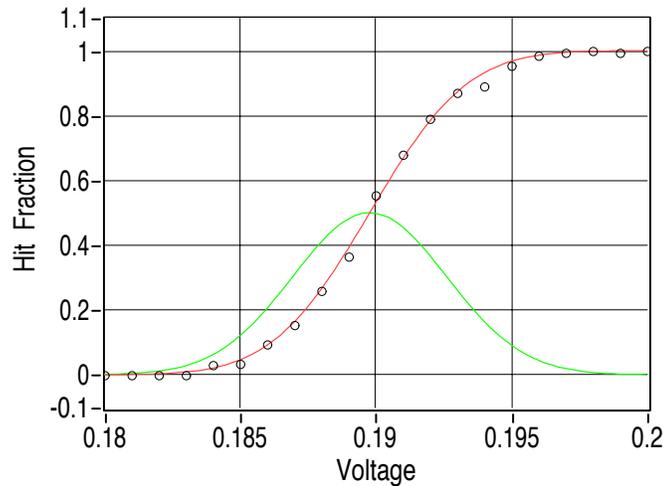

Figure C.10: Example turn-on curve and fit for 200 injections per voltage step.

C.4.3 Threshold scans

After collecting the turn-on curves for all strips, various informative and diagnostic plots can be obtained. These are collected in the "Threshold Scan" module shown in Fig. C.11.
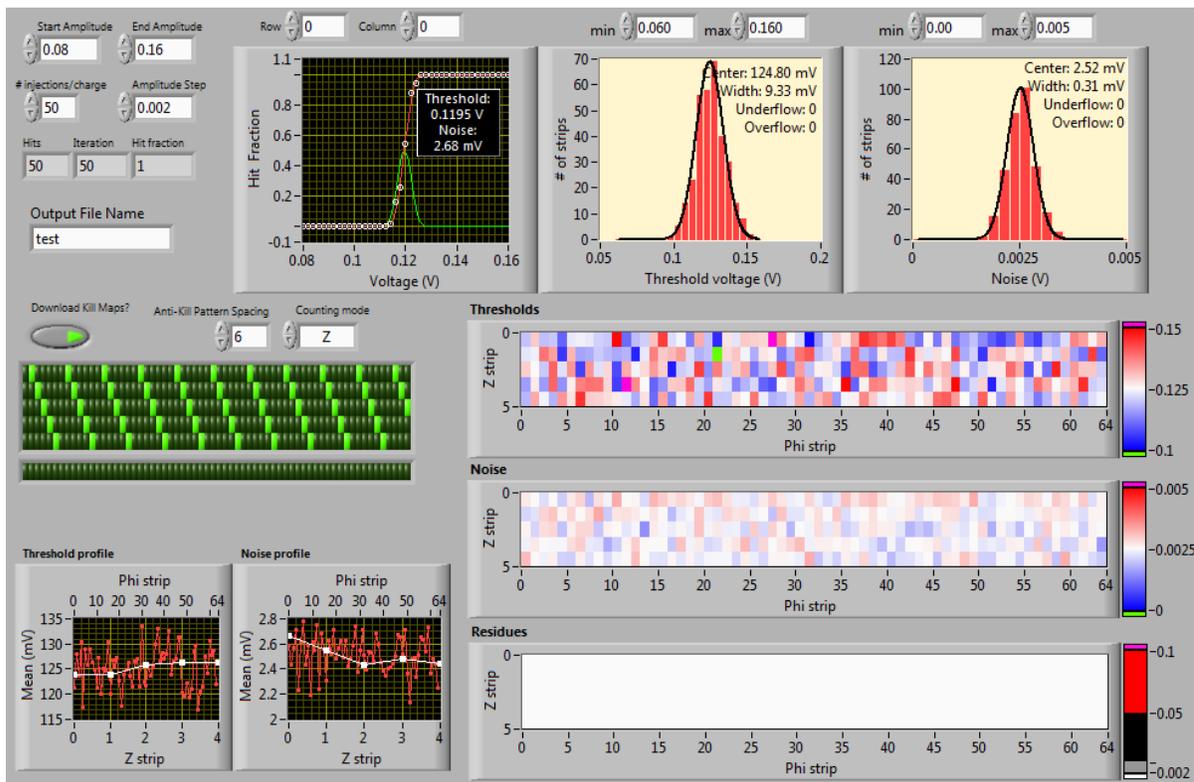
Figure C.11: Screenshot of the "Threshold Scan" module. A range and step size of injected voltage as well as the number of injections per voltage are input. Next the turn-on curves for all strips are collected, and histograms are made for the threshold and noise values, with Gaussians fit to the data. 2D plots of the threshold and noise values are also shown to see if there is any directional systematic effect. The "residues" 2D plot shows essentially the χ^2 of the fit for each strip in order to flag any poor fits (none in this example). The bottom left plots give another way of looking at those possible systematic effects by plotting the mean threshold and mean noise values of the projections along the two axes of the 2D plots. Finally, the matrix of green indicators shows which strips are active at a given time. In order to study crosstalk effects, the user can control the spacing amount between active strips.

C.4.4 Threshold tuning

Due to small differences in during the manufacturing process, the measured thresholds of all the strips will not be the same. In a working detector however, we desire they all fire above the same threshold. In order to achieve this, each strip front-end has a 5 bit trim byte (0-31). Fig. C.12 shows the variation in measured threshold as a function of this trim parameter. It is linear with a slope of about $3\,\mathrm{mV}$ per trim step.
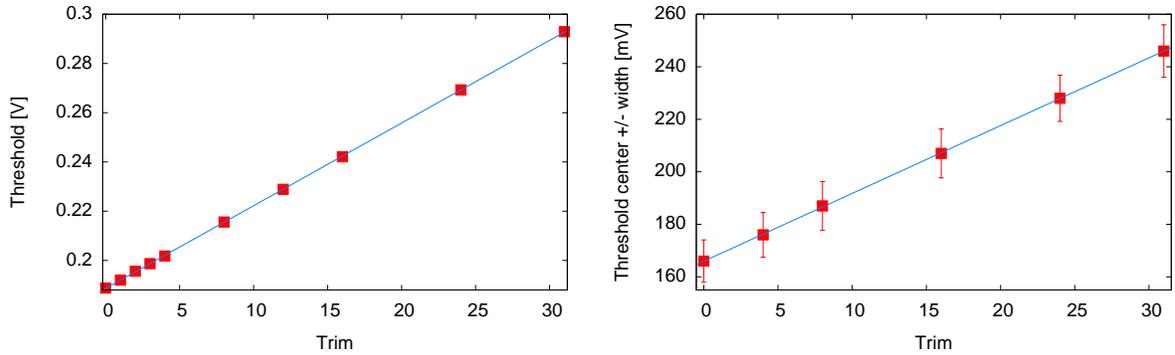

Figure C.12: *Left*: Measured threshold versus threshold trim setting for a single strip. *Right*: The same thing for all strips at once.

Given the trim response and the measured thresholds, we can calculate the appropriate trim settings for the individual strips in order to set them all to the same measured threshold. The target threshold is chosen to be the maximum measured threshold since the default trim setting is 0. The result of the threshold tuning is shown in Fig. C.13 where we see that the width of the thresholds distribution has gone from $\sim 12\,\mathrm{mV}$ to $\sim 3\,\mathrm{mV}$. In order to do better, we could either measure the trim response for each strip rather than just one or all, or we can apply the tuning procedure iteratively. Applying it a second time we go from a threshold distribution of width $\sim 3\,\mathrm{mV}$ to $\sim 1\,\mathrm{mV}$. This is good given the trimming resolution is $\sim 3\,\mathrm{mV}$.
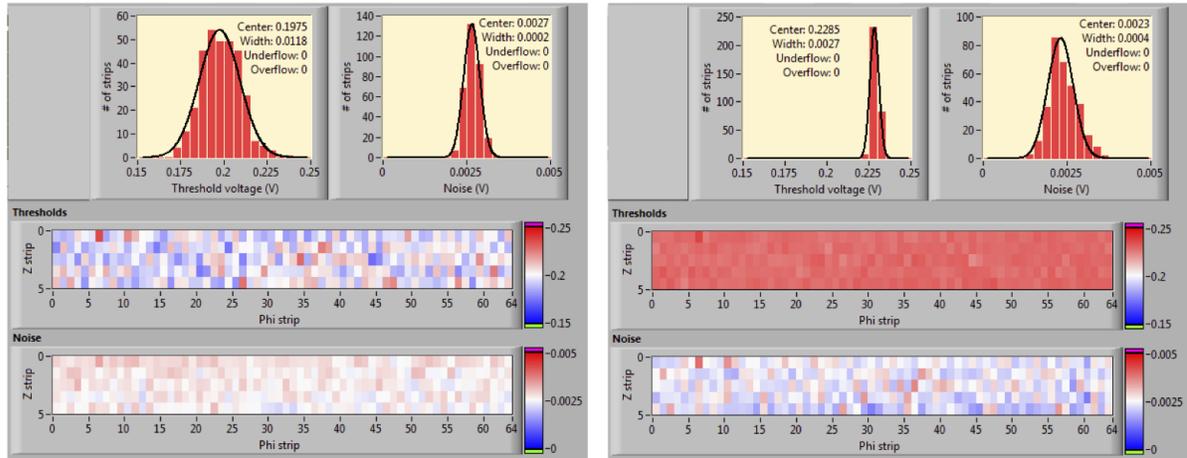
Figure C.13: Result of the threshold tuning procedure. The width of the thresholds distribution went from $\sim 12\,\mathrm{mV}$ to $\sim 3\,\mathrm{mV}$. Applying the procedure again resulted in a threshold width of $\sim 1\,\mathrm{mV}$.

C.4.5 CROSSTALK PROBLEM

Using the Threshold Scan module, we studied the effect of crosstalk between neighboring strips. This uncovered a flaw in the design of the chip. We saw a difference, shown in Fig. C.14, between the measured threshold maps when all strips are activated versus when they are activated in isolation. We see a large difference wherein the main feature is an entire strip set (the five strips in the short strip direction) having the same low threshold, equal to the value of the $\mathrm{Z}\langle 0 \rangle$ strip's threshold. Fig. C.15 shows the turn-on curves for one of the problematic $\mathrm{Z}\langle 1\text{–}4 \rangle$ strips as a function of the front-end bias current. Increasing the current decreases crosstalk effects and makes the true turn-on curve start to appear.

The evidence points to this issue coming from the geometrical layout of the chip, shown in Fig. C.16. The issue will be resolved in a future iteration of the chip, most likely by shielding the digital discriminator lines from the sensor input pads.
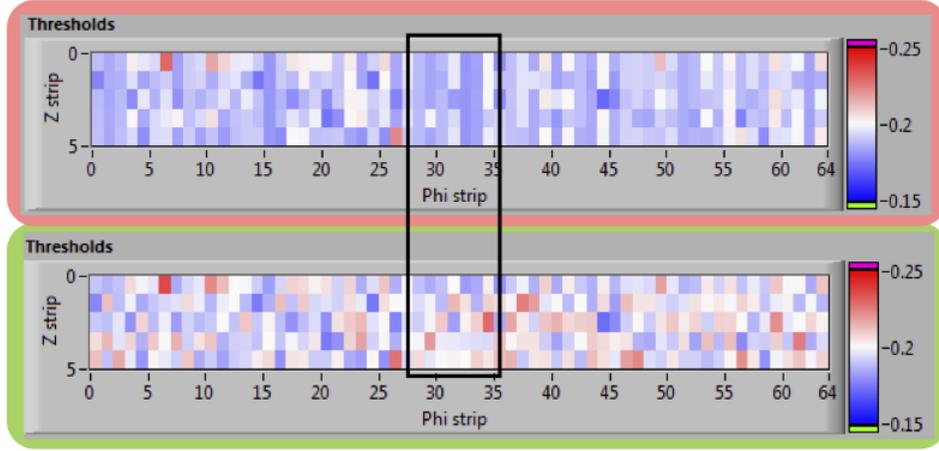
145

Figure C.14: The threshold maps obtained for all strips activated at once (top) and for each strip activated in isolation (bottom). We see a large difference wherein the main feature is an entire strip set (the five strips in the short strip direction) having the same low threshold, equal to the value of the $Z\langle 0 \rangle$ strip's threshold. The boxed area highlights a concentrated area of this issue.
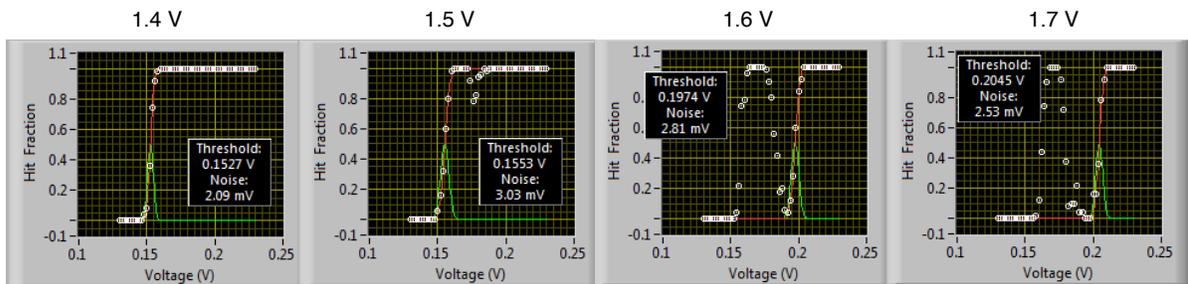

Figure C.15: Turn-on curves for one of the problematic $Z\langle 1\text{–}4 \rangle$ strips as a function of the front-end bias current. At low currents the measured turn-on curve is the same as that for the $Z\langle 0 \rangle$ strip, while at higher currents the true turn-on curve emerges due to lower crosstalk effects.
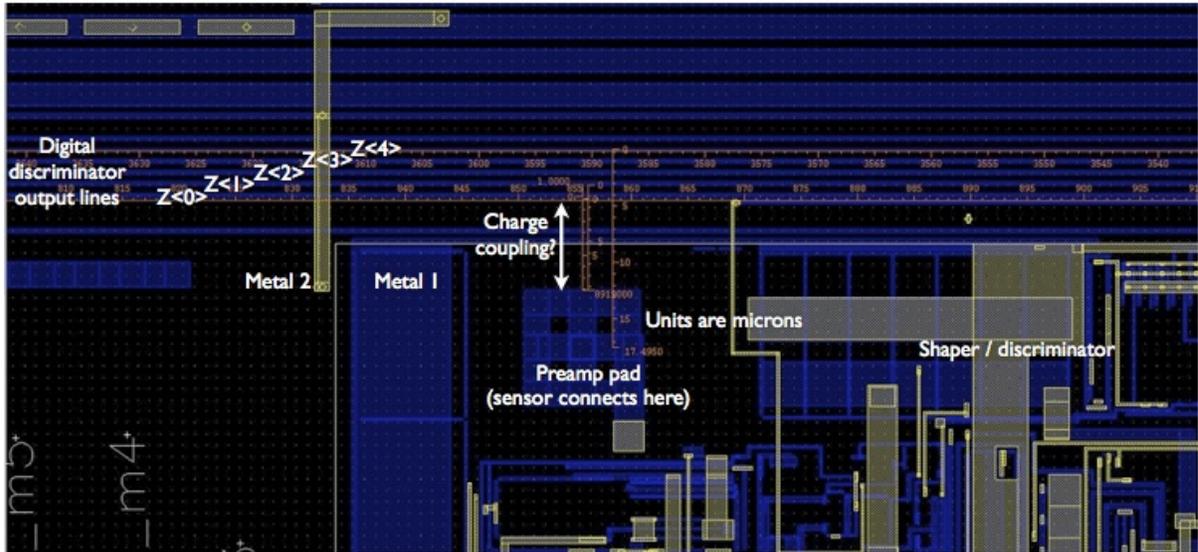
Figure C.16: Portion of the chip layout which shows that the digital discriminator output lines run near the sensor input pads for all of the strips in the strip set. These lines were designed to run the same distance in order to have the same capacitance. However, the $Z\langle 0\rangle$ line is the closest to the input pads (about $10\,\mu$m away) and, like a capacitor, can evidently induce some extra charge in their front-ends causing them to fire.

C.4.6 POWER REQUIREMENTS

In order to assess how much power must be supplied to the chip in order to operate at the required sensitivity, we began by studying the variation of the mean and width of the thresholds distribution as a function of the preamp bias current. This is shown in Fig. C.17. This is just one step in a longer study that should be done at some point in order to demonstrate the efficacy of this design (on-detector, 3D) in operating at lower power than conventional methods.

C.4.7 CAPACITIVE LOADING

Since the current chip is not connected to sensors, we would like to add some capacitance to the sensor input pads to mimic sensors. This would give more realistic measurements
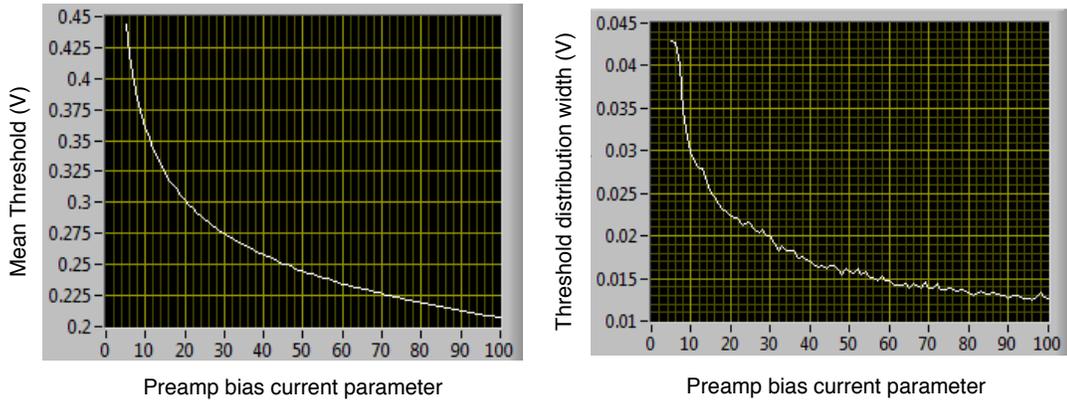
Figure C.17: The mean and width of the thresholds distribution as a function of the preamp bias current (in units of the DAC parameter controlling this).

and allow us to see how things may change when we do connect sensors. One method was to attach a regular probe tip to the sensor input (see Fig. C.18), and connect the probe to capacitor(s) connected to ground. Three 1 pF capacitors were mounted on the board in parallel, with standard probe plugs on one end. This allows for connecting them together with jumper cables to vary the added capacitance between 1–3 pF. We expect the added capacitance to have an effect on the turn-on curve, but in fact we were unable to measure any significant difference. One possible cause is that the probe itself has more inductance and capacitance, making the added capacitors negligible.

Figure C.18: Photograph under microscope of the probe making contact with a sensor input pad.

A second approach underway is to wire bond the sensor input pads directly to a series of dummy pads on the chip. These pads are separated from the ground plane by a layer of silicon, and hence will act as some added capacitance. The wire bonds on the sensor inputs are shown in Fig. C.19. Due to an issue with the chip these bonds were done on, more work is required to understand what is going on.
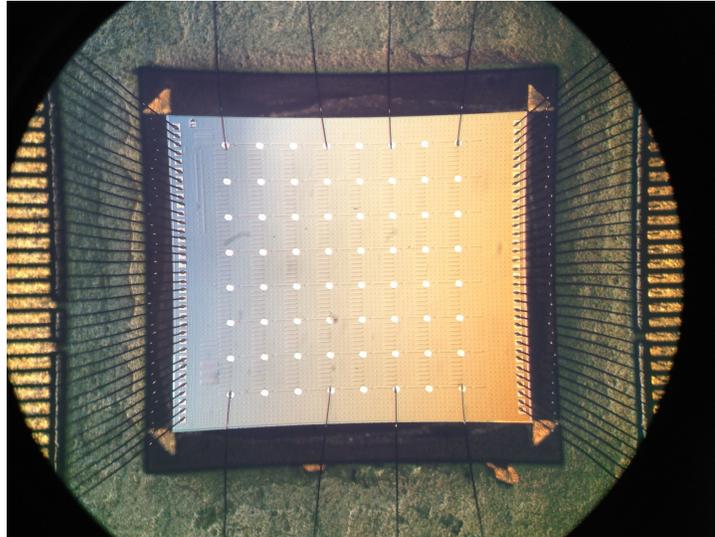

Figure C.19: Wire bonds connected to sensor inputs on the long strip tier.

C.4.8 FRONT-END CHARACTERISTICS

In addition to the sensor inputs on the short strip tier, there are also three debug pads for each front-end. These are related to the amplifier-shaper-discriminator (ASD) and are the amplifier, preamplifier, and discriminator outputs. We study the output of these pads to understand the behavior of the ASD, but also to make use of them in understanding the capacitive loading tests. The pads are shown in Fig. C.20. We use the Picoprobe to read the outputs out on the oscilloscope. These results are shown in Fig. C.21. However, note that these plots were obtained without pull-up and pull-down resistors connected

to the pads, as the design requires, and so may not reflect the proper measurements. One attempt at attaching these resistors is shown in Fig. C.22 wherein we wire bond the debug pads to dummy pads on the board and solder those pads to the resistors mounted to ground and power. Unfortunately the board was damaged in carrying out the modifications, but another attempt is planned.
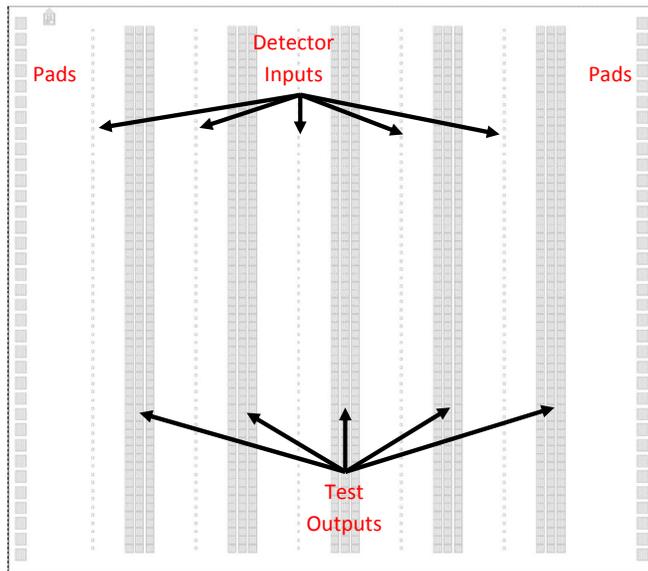

Figure C.20: Layout of the short strip tier input pads and three debug pads per strip front-end.
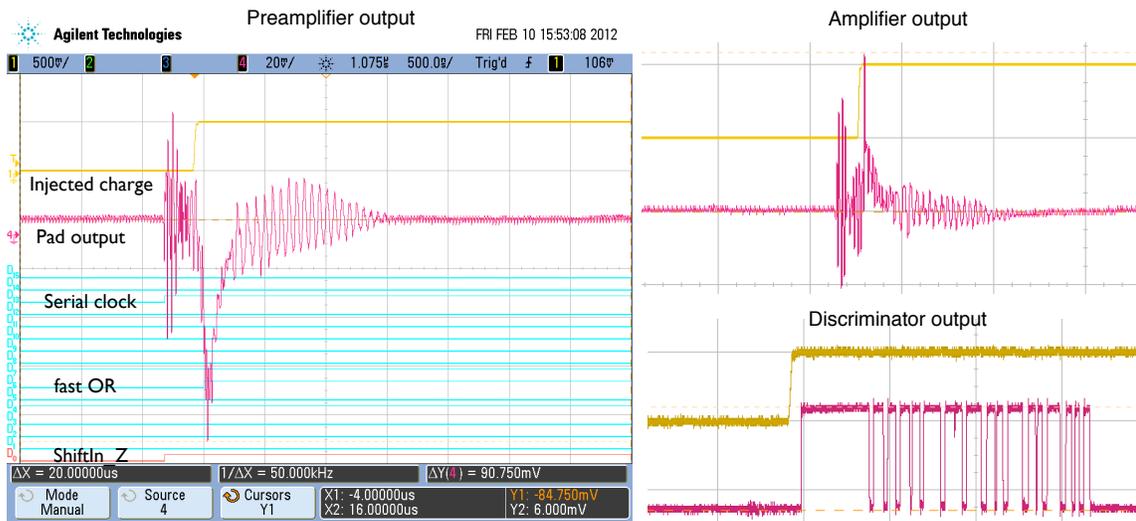
Figure C.21: Measured output of the ASD debug pads as a charge is injected into a single strip. These measurements were done without the required pull-up and pull-down resistors and hence may not make sense. The top left preamplifier output plot shows some pickup before the charge is injected, most likely from the serial clock or shift in Z lines, followed by the response in which ringing is apparent. The top-right shows the amplifier output which is so small it is likely *only* pickup. The bottom-right shows the output of the discriminator and the effect of the ringing.
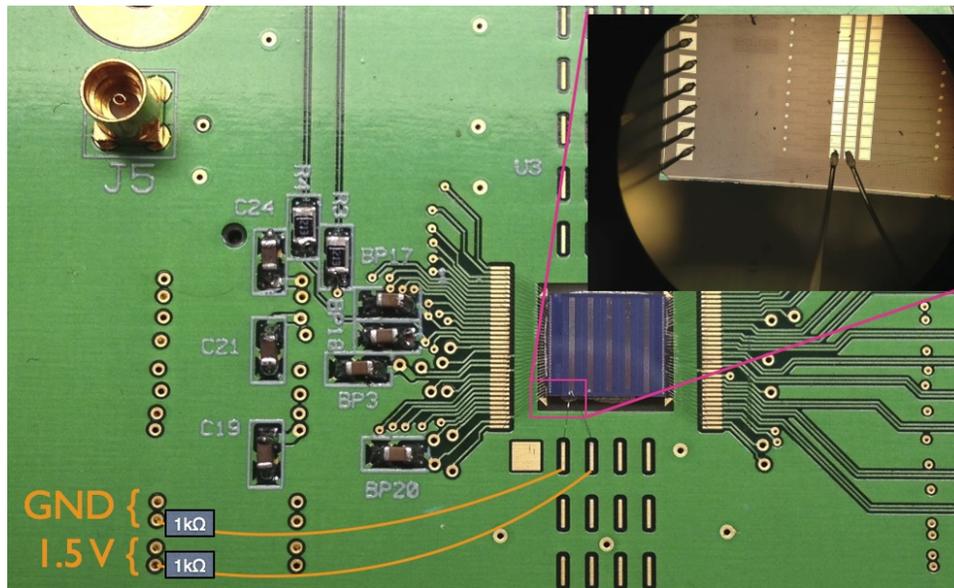

Figure C.22: Schematic of the pull-up and pull-down resistor modification. The wire bonds attach the debug pads to dummy pads on the board to which wires would be soldered to the resistors mounted to ground and power.

REFERENCES

1. Particle Data Group, "History of particle physics."
   http://www.particleadventure.org/other/history/. Accessed Jan. 2012. 6, 57

2. CDF and DØ Collaborations, T. Aaltonen *et al.*, "Evidence for a particle produced
   in association with weak bosons and decaying to a bottom-antibottom quark pair in
   Higgs boson searches at the Tevatron," arXiv:1207.6436 [hep-ex]. 7

3. K. Hagiwara *et al.*, "Probing the weak boson sector in $e^+e^- \to W^+W^-$," *Nucl.
   Phys. B* **282** (1987) 253. 8

4. M. Kober, B. Koch, and M. Bleicher, "First order calculation of the inclusive cross
   section $pp \to ZZ$ by graviton exchange in large extra dimensions," *Phys. Rev. D* **76**
   (2007) 125001. 8

5. E. J. Eichten, K. Lane, and A. Martin, "Technicolor explanation for the CDF Wjj
   excess," *Phys. Rev. Lett.* **106** (Jun, 2011) 251803. 8

6. Wikipedia contributors, "Standard model."
   http://en.wikipedia.org/wiki/Standard_Model. Accessed Jan. 13 2012. 11

7. T. Sjöstrand *et al.*, "PYTHIA 6 physics and manual," *J. High Energy Phys.* **05** (2006)
   026. 15, 60, 68, 93

8. J. M. Campbell and R. K. Ellis, "Update on vector boson pair production at
   hadron colliders," *Phys. Rev. D* **60** (1999) 113006. 22, 79, 93, 107

9. P. Tanedo, "Why do we expect a Higgs boson? Part I: Electroweak Symmetry
   Breaking." Quantum Diaries, 2011. http://www.quantumdiaries.org/2011/11/21/
   why-do-we-expect-a-higgs-boson-part-i-electroweak-symmetry-breaking/.
   Accessed May 2012. 21

10. F. Halzen and A. Martin, *Quarks & Leptons: An introductory course in modern
    particle physics.* John Wiley & Sons, New York, USA, 1984. 24, 26

11. M. E. Peskin and D. V. Schroeder, *An Introduction to Quantum Field Theory.*
    Addison-Wesley, Reading, 1995. 26

12. A. Djouadi, "The Anatomy of electro-weak symmetry breaking. I: The Higgs boson
    in the standard model," *Phys. Rept.* **457** (2008) 1–216. 31, 33, 34, 35

13. U. Aglietti *et al.*, "Tevatron for LHC report: Higgs," arXiv:hep-ph/0612172. 32

14. Particle Data Group, K. Nakamura, "Review of particle physics," *J. Phys. G* **37**
    (2010) 075021. 33, 116

15. LEP Electroweak Working Group, "LEP/TEV Electroweak Working Group Results
    for Winter 2012." http://lepewwg.web.cern.ch/LEPEWWG/. Accessed Mar. 2012.
    37, 38

16. J. Alcaraz, "Precision Electroweak Measurements and Constraints on the Standard

Model," `arXiv:0911.2604 [hep-ex]`. 37

17. Gfitter Group, "Gfitter - Results for the Global Electroweak Standard Model Fit," 2012. `http://gfitter.desy.de/Standard_Model/`. Accessed Mar. 2012. 37

18. H. Flacher *et al.*, "Gfitter - Revisiting the Global Electroweak Fit of the Standard Model and Beyond," *Eur. Phys. J. C* **60** (2009) 543–583. 37

19. T. Junk, "Confidence Level Computation for Combining Searches with Small Statistics," *Nucl. Instrum. Methods A* **434** (1999) 435–443. 38

20. A. L. Read, "Presentation of search results: The CL(s) technique," *J. Phys. G* **28** (2002) 2693–2704. 38, 39

21. LEP Working Group for Higgs boson searches, R. Barate *et al.*, "Search for the standard model Higgs boson at LEP," *Phys. Lett. B* **565** (2003) 61–75. 39, 40

22. TEVNPH (Tevatron New Phenomina and Higgs Working Group), CDF and DØ Collaboration, "Combined CDF and DØ Search for Standard Model Higgs Boson Production with up to $10.0\,\text{fb}^{-1}$ of Data," `arXiv:1203.3774 [hep-ex]`. Preliminary results prepared for the Winter 2012 Conferences. 40, 41

23. CMS Collaboration, S. Chatrchyan *et al.*, "Combined results of searches for the standard model Higgs boson in pp collisions at sqrt(s) = 7 TeV," *Phys. Lett. B* **710** (2012) 26–48. 40

24. ATLAS Collaboration, "An update to the combined search for the Standard Model Higgs boson with the ATLAS detector at the LHC using up to $4.9\,\text{fb}^{-1}$ of pp collision data at $\sqrt{s}$ = 7 TeV," *ATLAS-CONF-2012-019* (Mar, 2012) . 40

25. CMS Collaboration, S. Chatrchyan *et al.*, "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC," `arXiv:1207.7235 [hep-ex]`. 42

26. ATLAS Collaboration, G. Aad *et al.*, "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC," `arXiv:1207.7214 [hep-ex]`. 42

27. Fermilab, "R.R. Wilson's Congressional Testimony, April 1969." `http://history.fnal.gov/testimony.html`. Accessed May 2012. 43

28. D. Lincoln, *Understanding the Universe: From Quarks to the Cosmos.* World Scientific, Singapore, 2004. 44

29. Fermilab, "Accelerators — Fermilab's Chain of Accelerators," 2002. `http://www.fnal.gov/pub/inquiring/physics/accelerators/chainaccel.html`. Accessed May 2012. 44

30. Fermilab, "Fermilab Run II Handbook," 1998. `http://www-bd.fnal.gov/lug/runII_handbook/RunII_index.html`.

31. Wikipedia contributors, "Fermilab." `http://en.wikipedia.org/wiki/Fermilab`. Accessed May 10 2012. 44

32. K. Wille, *The Physics of Particle Accelerators: An Introduction.* Oxford University Press, USA, May, 2001. 47

33. H. Wiedemann, *Particle Accelerator Physics.* Springer-Verlag, Berlin, third ed., May, 2007. 47

34. E. Bruning, Oliver S., E. Collier, P., E. Lebrun, P., E. Myers, S., E. Ostojic, R., *et al.*, "LHC Design Report. Vol. 1: The LHC Main Ring — Ch. 3: Layout and Performance,". http://lhc.web.cern.ch/lhc/lhc-designreport.html. 47

35. CDF Collaboration, "CDF Summary of Stores." http://www-cdfonline.fnal.gov/ops/opshelp/stores/. Accessed May 2012. 47, 48

36. CDF Collaboration, A. Abulencia *et al.*, "Measurements of inclusive W and Z cross sections in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV," *J. Phys. G* **34** (2007) 2457. 48

37. CDF Collaboration, R. Blair *et al.*, "The CDF-II detector: Technical design report,". http://www-cdf.fnal.gov/upgrades/tdr/tdr.html. 48, 49

38. Symmetry Magazine, "Fermilab collider experiments discover rare single top quark," 2009. http://www.symmetrymagazine.org/breaking/2009/03/09/fermilab-collider-experiments-discover-rare-single-top-quark/. Accessed May 2012. 48

39. CDF Collaboration, "CDF Run I Experiment - Online Luminosity." http://www-cdf.fnal.gov/experiment/lumi.html. Accessed May 2012. 48

40. CDF Collaboration, T. Aaltonen *et al.*, "Observation of single top quark production and measurement of $|V_{tb}|$ with CDF," *Phys. Rev. D* **82** (2010) 112005. 49, 63, 106, 109

41. A. Sill, "CDF Run II silicon tracking projects," *Nucl. Instrum. Methods A* **447** (June, 2000) 1–8. 50

42. A. Mitr, "The CDF Run II Silicon Detector," in *Nuclear Science Symposium Conference Record, 2006. IEEE*, vol. 2, pp. 705–708. 29 2006-Nov. 1, 2006. 51

43. T. Affolder *et al.*, "CDF Central Outer Tracker," *Nucl. Instrum. Methods A* **526** no. 3, (2004) 249. 51

44. CDF Collaboration, "A Brief Description of the CDF Detector in Run II,". Unpublished. 52, 53

45. L. Balka *et al.*, "The CDF central electromagnetic calorimeter," *Nucl. Instrum. Methods A* **267** no. 2-3, (1988) 272 – 279. 53

46. S. Bertolucci *et al.*, "The CDF central and endwall hadron calorimeter," *Nucl. Instrum. Methods A* **267** no. 2-3, (1988) 301 – 314. 53

47. CDF Collaboration, A. Bhatti *et al.*, "Determination of the jet energy scale at the Collider Detector at Fermilab," *Nucl. Instrum. Methods A* **566** (2006) 375–412. 53, 54

48. CDF Collaboration, F. Abe *et al.*, "Topology of three-jet events in $p\bar{p}$ collisions at $\sqrt{s} = 1.8$ TeV," *Phys. Rev. D* **45** (1992) 1448. 54, 90

49. CDF Collaboration, F. Abe *et al.*, "Inclusive jet cross section in $p\bar{p}$ collisions at $\sqrt{s} = 1.8$ TeV," *Phys. Rev. Lett.* **68** (Feb, 1992) 1104–1108. 54

50. CDF Collaboration, D. Acosta *et al.*, "Measurement of the $t\bar{t}$ production cross section in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV using lepton + jets events with secondary vertex b-tagging," *Phys. Rev. D* **71** no. 5, (Mar, 2005) 052003. 54, 62

51. R. Downing *et al.*, "Track extrapolation and distribution for the CDF-II trigger system," *Nucl. Instrum. Methods A* **570** (Jan., 2007) 36–50. 55, 56

52. CDF Collaboration, T. Aaltonen *et al.*, "First observation of vector boson pairs in a hadronic final state at the Tevatron Collider," *Phys. Rev. Lett.* **103** (2009) 091803. 59, 79, 100, 110

53. M. Goncharov *et al.*, "The timing system for the CDF electromagnetic calorimeters," *Nucl. Instrum. Methods A* **565** (2006) 543. 59, 79

54. J. Freeman, W. Ketchum, J. Lewis, S. Poprocki, A. Pronko, V. Rusu, and P. Wittich, "An artificial neural network based b jet identification algorithm at the CDF Experiment," *Nucl. Instrum. Methods A* **663** (2011) 37. 60, 64, 82, 124

55. J. D. Bjorken, "Properties of hadron distributions in reactions containing very heavy quarks," *Phys. Rev. D* **17** (Jan, 1978) 171–173. 60

56. CDF Collaboration, F. Abe *et al.*, "Evidence for top quark production in $p\bar{p}$ collisions at $\sqrt{s} = 1.8$ TeV," *Phys. Rev. D* **50** (1994) 2966. 61

57. DØ Collaboration, V. M. Abazov *et al.*, "b-jet identification in the DØ experiment," *Nucl. Instrum. Methods A* **620** no. 2-3, (2010) 490. 61

58. CMS Collaboration, "Performance of the b-jet identification in CMS," *CMS-PAS-BTV-11-001* (2011) . 61

59. ATLAS Collaboration, "Commissioning of the ATLAS high-performance b-tagging algorithms in the 7 TeV collision data," *ATLAS-CONF-2011-102* (2011) . 61

60. CDF Collaboration, A. Abulencia *et al.*, "Measurement of the $t\bar{t}$ production cross section in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV using lepton+jets events with jet probability b-tagging," *Phys. Rev. D* **74** (2006) 072006. 62

61. CDF Collaboration, D. Acosta *et al.*, "Measurement of the $t\bar{t}$ production cross section in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV using lepton plus jets events with semileptonic B decays to muons," *Phys. Rev. D* **72** (2005) 032002. 62, 73

62. S. Richter, "Search for electroweak single top-quark production with the CDF II experiment," *FERMILAB-THESIS-2007-35* (2007) . 63

63. C. Ferrazza, "Identificazione di quark pesanti in getti adronici in interazioni $p\bar{p}$ con il rivelatore CDF al Tevatron," Master's thesis, Università "La Sapienza" Roma, 2006. 63

64. P. Mastrandrea, "Study of the heavy flavour fractions in Z+jets events from $p\bar{p}$ collisions at energy = 1.96 TeV with the CDF II detector at the Tevatron collider," *FERMILAB-THESIS-2008-63* (2008) . 63

65. CDF Collaboration, T. Aaltonen *et al.*, "Search for $WZ + ZZ$ production with missing transverse energy + jets with b enhancement at $\sqrt{s} = 1.96$ TeV," *Phys. Rev. D* **85** (Jan, 2012) 012002. 64, 89, 124

66. CDF Collaboration, "Search for $ZW + ZZ \rightarrow \ell^+\ell^- + jj$ production at $\sqrt{s} = 1.96$ TeV," CDF Internal Note #10864 (unpublished). 64, 124

67. J. Freeman, T. Junk, M. Kirby, Y. Oksuzian, T. Phillips, *et al.*, "Introduction to HOBIT, a b-Jet Identification Tagger at the CDF Experiment Optimized for Light Higgs Boson Searches," `arXiv:1205.1812 [hep-ex]`. 65, 124

68. CDF Collaboration, "Search for Standard Model Higgs Boson Production in Association with a $W^\pm$ Boson with 9.45 fb^{-1} of CDF Data," CDF Internal Note #10796 (unpublished). 65

69. CDF Collaboration, "A Search for the Standard Model Higgs Boson in the Process $ZH \rightarrow \ell^+\ell^- b\bar{b}$ Using 9.45 fb^{-1} of CDF II Data," CDF Internal Note #10799 (unpublished). 65

70. A. Hoecker, P. Speckmayer, J. Stelzer, J. Therhaag, E. von Toerne, and H. Voss, "TMVA: Toolkit for Multivariate Data Analysis," *PoS(ACAT)* (2007) 040, `arXiv:physics/0703039`. 65, 68

71. Wikipedia contributors, "Artificial neural network." `http://en.wikipedia.org/wiki/Artificial_neural_network`. Accessed July 9 2012. 65

72. J. Teeter and M.-Y. Chow, "Application of functional link neural network to HVAC thermal dynamic system identification," *IEEE Trans. Ind. Electron.* **45** no. 1, (1998) 170–176.

73. K. Asakawa and H. Takagi, "Neural networks in Japan," *Commun. ACM* **37** no. 3, (Mar., 1994) 106–112.

74. M. Thorson, F. Warthman, and M. Holler, "A neural-network audio synthesizer," *Dr. Dobb's J.* **18** no. 2, (Feb., 1993) 50–64. 65

75. G. D. Garson, "A comparison of neural network and expert systems algorithms with common multivariate procedures for analysis of social science data," *Soc. Sci. Comput. Rev.* **9** no. 3, (1991) 399–434. 67

76. R. Brun *et al.*, "GEANT3 manual," CERN Report CERN-DD-78-2-REV (unpublished). 69, 93

77. M. L. Mangano *et al.*, "ALPGEN, a generator for hard multiparton processes in hadronic collisions," *J. High Energy Phys.* **2003** no. 07, (2003) 001. 79, 93

78. J. Pumplin *et al.*, "New generation of parton distributions with uncertainties from

global QCD analysis," *J. High Energy Phys.* **0207** (2002) 012. 79, 93

79. CDF Collaboration, T. Aaltonen *et al.*, "Measurement of cross sections for b jet production in events with a Z boson in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV," *Phys. Rev. D* **79** (2009) 052008. 82, 132

80. CDF Collaboration, Z. Alawi, W. Ketchum, *et al.*, "Search for diboson production in the ll+jets channel," CDF Internal Note #10061 (unpublished). 91

81. J. Alwall *et al.*, "MadGraph/MadEvent v4: the new web generation," *J. High Energy Phys.* **09** (2007) 028. 93

82. S. Klimenko, J. Konigsberg, and T. M. Liss, "Averaging of the inelastic cross sections measured by the CDF and the E811 experiments," FERMILAB-FN-0741. 105

83. CDF Collaboration, T. Junk, "Sensitivity, exclusion and discovery with small signals, large backgrounds, and large systematics," CDF Internal Note #8128 (unpublished). 106, 109

84. U. Langenfeld, S. Moch, and P. Uwer, "Measuring the running top-quark mass," *Phys. Rev. D* **80** (Sep, 2009) 054009. 107

85. N. Kidonakis, "Single top quark production at the Fermilab Tevatron: Threshold resummation and finite-order soft gluon corrections," *Phys. Rev. D* **74** (2006) 114012. 107

86. N. Kidonakis, "Single top quark production cross section at hadron colliders," *PoS(DIS2010)* (2010) 196, `arXiv:1005.3330v1`. 107

87. CDF Collaboration, G. Flanagan *et al.*, "Observation of diboson production in $\not{E}_{\mathrm{T}}$ + jets channel," CDF Internal Note #9736 (unpublished). 110

88. G. J. Feldman and R. D. Cousins, "Unified approach to the classical statistical analysis of small signals," *Phys. Rev. D* **57** (1998) 3873. 115

89. M. L. Mangano, "The super-LHC," `arXiv:0910.0030 [hep-ph]`. 133, 134

90. R. Lipton, "3D technology for intelligent trackers," *J. Inst.* **5** no. 10, (2010) C10006. 135

91. J. P. Alexander, T. Lutz, M. Suri, and J. Thom, "Interposer development for 3D trackers," *J. Inst.* **5** no. 08, (2010) C08008. 135

92. G. Deptuch, M. Demarteau, J. Hoff, R. Lipton, A. Shenai, M. Trimpl, R. Yarema, and T. Zimmerman, "Vertically integrated circuits at fermilab," *IEEE Trans. Nucl. Sci.* **57** no. 4, (Aug., 2010) 2178 –2186. 135

93. R. Lipton, "3D-vertical integration of sensors and electronics," *Nucl. Instrum. Methods A* **579** no. 2, (2007) 690 – 694. 135