THE PENNSYLVANIA STATE UNIVERSITY
SCHREYER HONORS COLLEGE

DEPARTMENT OF PHYSICS

A SEARCH FOR DARK MATTER VIA HIGGS DECAY
USING QUARK JET SUBSTRUCTURE

DANIEL ROBERT ABERCROMBIE
SPRING 2014

A thesis
submitted in partial fulfillment
of the requirements
for baccalaureate degrees
in Physics and Nuclear Engineering
with honors in Physics

Reviewed and approved* by the following:

Sarah Shandera
Assistant Professor of Physics
Thesis Supervisor

Richard Robinett
Professor of Physics
Honors Advisor

* Signatures are on file in the Schreyer Honors College

ABSTRACT

The discovery of a Higgs-like boson has allowed the possibility to probe many models of physics beyond the Standard Model. In this thesis, the possibility of the Higgs decaying into invisible dark matter through $WH \rightarrow q\bar{q}\chi\chi$ is explored. There are two main goals of this thesis. First, reconstruction of the hadronic W is optimized. The main figures of merit for this were the consistency of mass measurements for different collision conditions, and the resolution of this mass measurement. Consistent values allows for the missing energy from the invisible Higgs to be more accurately and quickly measured. Final results in suggest that one of the smaller cone sizes in the study, $\Delta R = 0.4$, will give the best mass resolution without losing too much p_T from radiated particles outside the jet cone. The trimming and filtering grooming algorithms, compete for the best mass resolution, while pruning, the recent CMS default, appears to be an overly aggressive grooming algorithm.

The second goal of this thesis is to recommend cuts on the recently proposed jet substructure descriptor N-Subjettiness, to be used in future searches for dark matter at high energy collider experiments. Using cone sizes $\Delta R > 0.9$ with cuts on the N-Subjettiness variables of $\tau_2/\tau_1 < 0.44$ appears to give the best signal significance in this preliminary analysis. Using this jet selection has a clearly visible signal region for Higgs to invisible particles. The excess number of events with the current integrated luminosity is not great enough to claim detection of an invisible Higgs. However, increasing the luminosity in the upcoming run of the LHC will allow this decay channel to be investigated much more conclusively.

TABLE OF CONTENTS
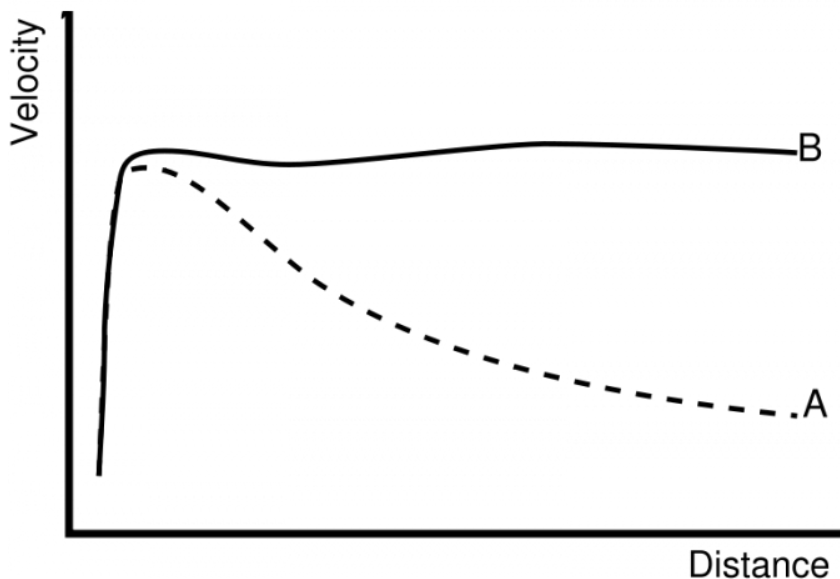
LIST OF FIGURES

ACKNOWLEDGEMENTS

Chapter 1

Introduction

With the discovery of a Higgs-like particle in 2012, it is likely all of the particles of the Standard Model have been detected and characterized [2]. This may lead some people to think, as William Thompson (Lord Kelvin) stated in 1900, "There is nothing new to be discovered in physics now. All that remains is more and more precise measurement." However, we are already aware that there is much about the universe that we do not know about. It has become widely accepted that the majority of the universe is made up of matter that we are unable to observe, dubbed dark matter. Dark matter is a required component of accurate gravitational models of galaxy rotation because visible matter should not behave in the manner observed. However, there is no dark matter particle encompassed in the Standard Model. Any additional characterization of this material would lead to new physics models. Towards this end, there has been a fervent search for any signal of dark matter, aside from gravitational evidence, to make any tie into our theories about the matter that we can observe. With the discovery of the Higgs boson, a particle that gives mass to some particles through electroweak symmetry breaking, we may have found a link to dark matter, which makes up a large portion of the universe's mass.

The rest of this chapter will describe in more detail attempts at detecting dark matter. I will then describe some of the points in a model where the Higgs could decay into dark matter. If this model is true, then we may be able to detect a dark matter signal during the operation of the Large Hadron Collider (LHC) at the European Organization for Nuclear Research (CERN). Chapter 2 describes the Compact Muon Solenoid (CMS) experiment, which gathered the data used to search for new particles. The remaining chapters of this thesis will outline steps conducted by me to help search for this signal at the LHC.
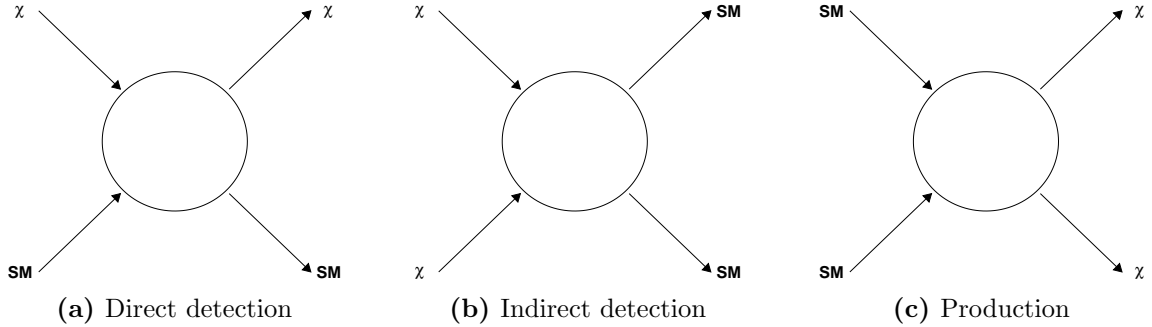
1.1 Status of Dark Matter Searches

In 1933, Fritz Zwicky first postulated dark matter when he tried to calculate the rotation of galaxies based on the amount of mass that he was able to see. The calculation disagreed with the observed rotation of our galaxy. Figure 1.1 shows the difference, which qualitatively says the outer edges of the galaxy rotate faster than they would if only visible matter was present. Visible matter in a galaxy is clumped closer together at the center, but distributing additional matter farther from the galactic center would give the observed velocity profile shown. Despite a need for this additional matter, none could be observed, leading Zwicky to call it "dark matter."

Figure 1.1: Some of the first evidence of dark matter is through what is called a galaxy rotation curve. This curve shows the predicted rotation, assuming only visible mass is present (A) and the actual measured rotation (B) of the Milky Way galaxy.

Since Zwicky's discovery, there have been many attempts to observe this dark matter through other means. We know that it interacts with the particles of the Standard Model through gravity, and through this interaction alone, we have determined that dark matter makes up over 80% of the mass in the universe. However, we know nothing else about it. One popular model of dark matter is that dark matter is made of Weakly Interacting Matter Particles (WIMPs). In this model, dark matter particles may interact in ways besides gravity. Individual particles are currently impossible to detect through gravitational forces, so there are three primary approaches to detecting

WIMPs using these possible alternative forces. These three methods, compared in Figure 1.2, are called direct detection, indirect detection, and production. In direct detection, experiments are searching for dark matter that scatters off of regular matter, leaving energy in large scintillating detectors. These experiments are located deep underground to reduce background scattering caused by cosmic rays. In indirect detection, experiments search for energetic Standard Model particles, usually photons or electrons, that are generated when dark matter annihilates. These experiments are usually satellite telescopes. In production, experiments search for a dark matter signal that appears when energetic particles collide. This method of dark matter detection will be the focus of subsequent chapters. Production experiments take place at large collider facilities.

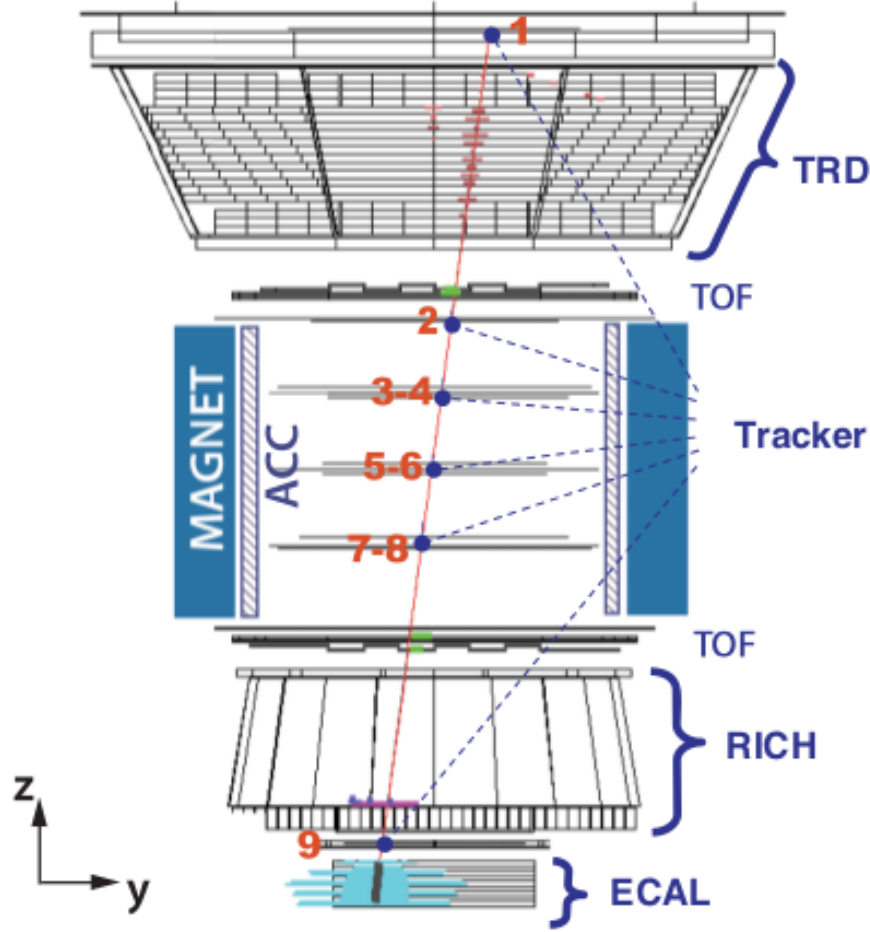| (a) Direct detection	(b) Indirect detection	(c) Production

Figure 1.2: There are three main schemes to search for dark matter. In direct detection, experiments attempt to identify dark matter scattering off of Standard Model particles, meaning there is a Standard Model particle (**SM**) and a WIMP (χ) in both the initial and final states. In indirect detection, experiments are designed to see Standard Model products that appear after dark matter annihilates with itself. In production, standard model particles are collided at high energy in an attempt to generate the heavy WIMPs.

Of all the experiments in the search for WIMPs, the most promising results have come out of the Alpha Muon Spectrometer (AMS) collaboration [5]. The AMS-02 detector is installed on the International Space Station (ISS), where it can measure incoming electron and positron cosmic rays. The components of the detector are shown in Figure 1.3. A magnet causes a charged particle to curve as it passes through the detector. Trackers and a calorimeter measure the charge to mass ratio and the energy of the particle, allowing the detector to distinguish different particles and ignore ions.
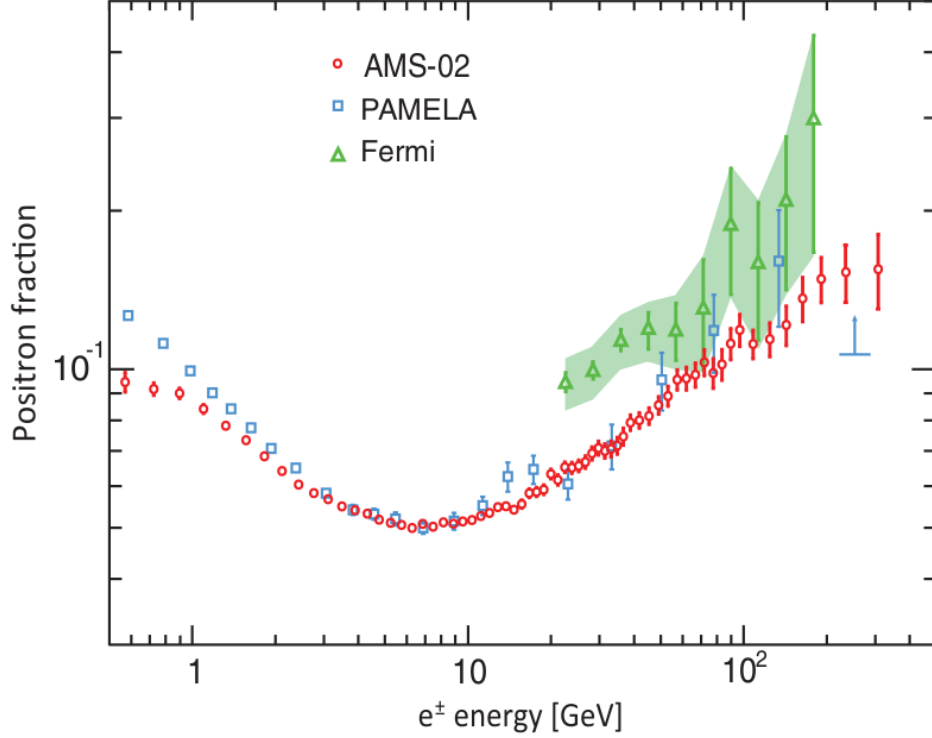
The motivation behind the AMS-02 experiment is the possibility that dark matter annihilates, ultimately resulting in an electron-positron pair. As a result, there will be a larger fraction of positrons at the energies up to the mass of the dark matter. Results from the first set of data released

Figure 1.3: Above is the basic layout of the Alpha Magnetic Spectrometer experiment. Electrons and positrons travel through detector trackers in a magnetic field. The energy of the electrons and positrons are deposited into the Electromagnetic Calorimeter. The image shows a simulated event of a 1.03 TeV electron passing through the detector. The red numbers show the number of planes that the particles pass through that measure momentum and charge [5].

by AMS-02 are shown in Figure 1.4. They confirm this trend found by previous experiments. Full evidence of production through annihilation will only come from locating a sharp drop in the positron fraction at some higher energy, corresponding to the mass of dark matter particles. Current detected positrons could have just been produced by Standard Model processes and accelerated via the strong electromagnetic fields surrounding pulsars. However, some current speculations predict dark matter to have a mass on the order of 1 TeV [12]. This should be well within the range of what current colliders, such as the Large Hadron Collider (LHC), are able to produce.
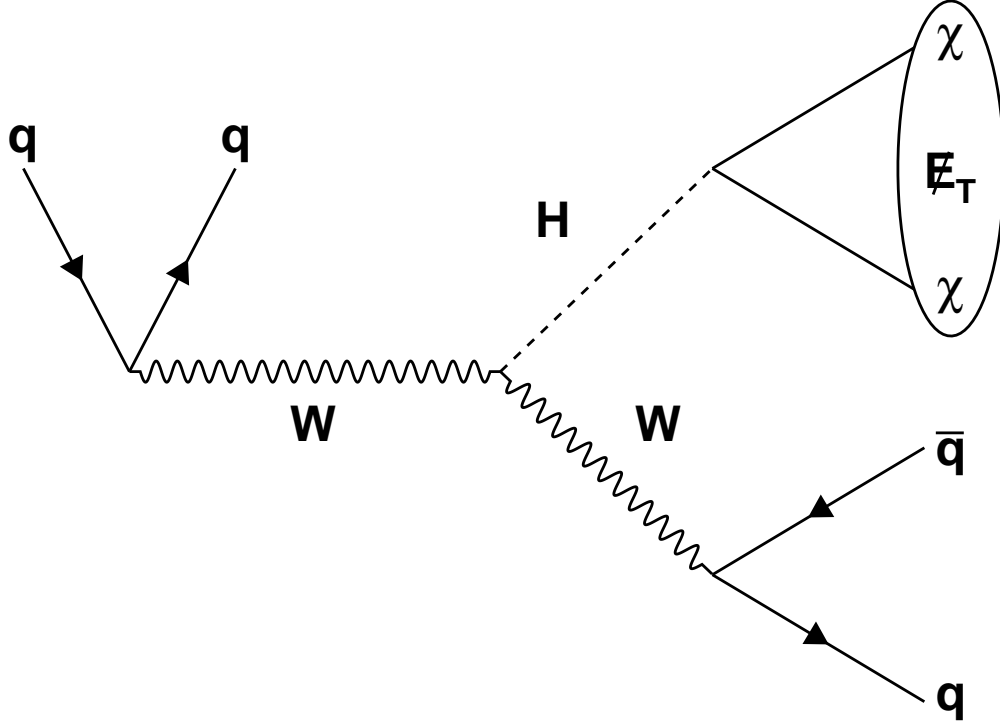
Figure 1.4: The first set of AMS results are compared to previous experiments [5]. The increase in positron fraction at higher energies suggests a mechanism that produces higher energy positrons, such as annihilation of massive particles.

1.2   The Processes in the Desired Event

One of the potential ways of producing dark matter is through the Higgs boson. Due to its role in electroweak symmetry breaking, a number of theories predict that the Higgs will couple with massive dark matter, leading to a decay mode with a final state of dark matter particles. Supersymmetric, extra dimensional, and dark matter singlet models all include variations making this prediction [4]. However, if dark matter has a mass on the $1\,\mathrm{TeV}$ scale, then this decay mode would be highly suppressed, compared to the case where the mass of the dark matter particle is less than half of the mass of an on-shell Higgs. Fortunately, given the amount of dark matter observed in the universe, the mass of the dark matter is not expected to be too much greater than that. A higher mass would have lead to much less production during the big bang [15]. This means, dark matter is still accessible to our experiments. Either way, current experimental results constrained the Higgs to invisibles branching ratio to $\mathcal{B}_{H\to inv} < 0.35$ with 3σ significance. Only a small part of this potential branching ratio is made up of Higgs to 4 neutrinos, with $\mathcal{B}_{H\to 4\nu} \approx 0.001$ [4].

There are still many channels that can be improved though. The particular process that will be considered in this thesis is shown in Figure 1.5. The Higgs boson can be radiated away from a W boson through a process known as Higgsstrahlung. It is then possible for the W to decay hadronically, $W \to q\bar{q}$. At the same time, the Higgs could decay into dark matter. Since dark matter consists of invisible particles, they would only be detectable in the form of missing transverse energy, also known as MET or $\not{E}_T$. Kinematics in collider experiments are measured using the transverse momentum, p_T, instead of total momentum because the velocity of the center of mass of a given collision is not known. Quarks within protons have various values of the total energy of the proton, so there is likely still net momentum along the beam line when two quarks collide. However, there is no momentum in the transverse plane. This is why missing energy is only measured in the transverse plane. That is the only region where it is certain to be hidden if not observed.

Figure 1.5: This is a Feynman diagram of the particular event that will be searched for using the methods of this thesis. Higgsstrahlung will radiate a Higgs that may decay into dark matter. This dark matter appears as MET. The W that radiated the Higgs would then decay hadronically, resulting in jets.

In the desired process, there will be two characteristics in the signal. First, events with large MET are considered. The most significant contribution to these events include the decay of $Z \to \nu\bar{\nu}$, which will be discussed in more detail in Section 4.2. Another process with large cross section and

high MET is $W \to \ell\nu$, where ℓ represents any charged lepton. This background process motivates vetoing every event that that contains a charged lepton.

Of of the remaining events, quark jets are analyzed and recombined to see if they could have been produced by a W boson. In particular, boosted Ws were identified in merged jets. If a W boson has a high enough momentum when it decays, the decay products will remain close together. Jets are recombined based on physical distance within the detector, so sometimes the two quarks from a hadronic W are recombined into the same jet. Other studies have been done searching for dark matter with identical final states [11]. However, this study includes a new variable called N-Subjettiness, which will be described in Section 4.4, a fat jet with two hard quarks can be distinguished from other massive jets. One of the final results of this study will be to recommend a cut for the N-Subjettiness variable.

Another important feature of this event is that the measured MET depends on the transverse momentum, or p_T, of the quark jets. Due to the multiple collisions, or pileup, during each event, there are stray particles that can be included inside of the jet that did not come from the proper collision vertex. Various grooming algorithms, which are described in Section 3.3, can be used in an attempt to reduce the presence of excessive particles. This is desirable because it will lead to more accurate characterization of the MET in an event, so the kinematics of the Higgs can be more accurately determined.

Chapter 2

The Compact Muon Solenoid Experiment

The work in this thesis was done as part of the CMS collaboration. CMS is a detector located at the Large Hadron Collider (LHC), which is designed to look for new particles. The data used in this analysis was collected by CMS during the LHC's proton-proton runs taking place from 2010 to 2012. All of the Monte Carlo simulations of collisions modeled CMS's response to reconstruct the events.

Figure 2.1 shows a small cross section of the cylindrical CMS detector. The barrel section of the components in this figure all cover the region $|\eta| < 1.2$ where η is the pseudorapidity, with some of the components covering a greater range. The endcap parts of these components detect particles in regions up to $|\eta| < 2.4$ [6]. More precise ranges will be given in the separate sections for each detector component. The pseudorapidity can be related to the angle from the plane perpendicular to the beam line, θ, by the following equation:

$$\eta = -\ln\left[\tan\left(\frac{\theta}{2}\right)\right] \tag{2.1}$$

Particles resulting from a collision have a uniform distribution through the azimuthal angle, ϕ. To more accurately reconstruct an event, all of the components form a full cylinder around the beam line. Only particles that travel nearly parallel to the beam line escape detection.

An overview of the detector's functionality can be given by describing the trajectory of different kinds of stable particles that are emitted from the collision point. Figure 2.1 also includes images of these trajectories. The first particle considered is the muon. As a relatively massive, compared to the electron, charged lepton, the muon is highly penetrative. It only interacts with matter through the weak and electromagnetic forces, reducing the stopping power of the detector when comparing the muon to hadrons. The magnetic field from the superconducting solenoid sends the muon on a curved trajectory through the central silicon tracker, the direction of the curvature depending on

Figure 2.1: The cross section of CMS shown above shows how different fundamental particles travel through the detector. It also shows the main components and detectors in the low rapidity region.

the muon's charge. The muon then passes unimpeded through the electromagnetic calorimeter and the hadronic calorimeter, losing little energy compared to its total energy. Then it passes through the out muon chambers of the detector, where the return yoke causes the muon to curve in the opposite direction. Being the only particle to consistently reach the muon chambers, its presence at this location makes the muon easy to identify.

The other lepton that is detected and identified directly is the electron. Since it is also charged, the electron leaves a curved path through the silicon tracker. However, due to its lower mass, which is comparable to the electrons present in the detector, electrons are quickly stopped in the electromagnetic calorimeter (ECAL). Most of the electron's energy is therefore deposited and measured in the ECAL. Coupled with the momentum to charge ratio measured in the silicon tracker, this allows for identification of electrons.

The last of the charged particles detected by CMS are grouped together as charged hadrons. Made of quarks and heavier than the electron, charged hadrons also leave energy behind as they curve through the silicon tracker and the ECAL, but travel through the ECAL to the hadronic calorimeter (HCAL). The rest of the energy in the charged hadrons is deposited in the HCAL.

Finally, there are neutral particles. Other hadrons have no net charge, which means that they

are not curved by the magnetic field, and they leave very little energy deposited in the silicon tracker and the ECAL. However, they still act through the strong force, so they deposit there energy in the denser HCAL. Photons also have no charge, so they do not leave a path through the silicon tracker. They do still interact through the electromagnetic interactions, so their energy is deposited into the ECAL. Depending which calorimeter detects the neutral particle, the type of neutral particle can be detected.

The following sections will examine separate components of the experiment in more detail. First a brief overview of the LHC is given. Then, each of the detector components in Figure 2.1 are discussed separately. The forward detectors, which are used for η as high as $\eta \leq 5$ [10], are not pictured and will not be covered in this report since the majority of the data used in this analysis is from the lower η regions.

2.1 The Large Hadron Collider

The Large Hadron Collider (LHC) is an accelerator designed to collide protons and lead ions. All of these are matter particles, so that the production of anti-matter would not constrain the luminosity of the experiment, which is designed to be approximately $10^{34}\text{cm}^{-2}\text{ s}^{-1}$ [8]. A full expression of the luminosity, L, is based on accelerator design depends on a number of beam parameters [14], but the following simpler expression may be more enlightening.

$$N_{event} = \mathcal{L}\sigma_{event} \tag{2.2}$$

N_{event} is the number of times a certain event with cross section σ_{event} should have occurred when the LHC has run for time T, giving the integrated luminosity, $\mathcal{L}$, related to the luminosity as follows.

$$\mathcal{L} = \int_0^T L \, dt \tag{2.3}$$

Since the LHC is designed to collide two particles of the same charge, two different beam pipes with opposite magnetic fields are required. Figure 2.2 shows a cross section of the collider. The two beam pipes have a vacuum of approximated 10^{-10}mbar after being pumped down and cooled [14].

10

Figure 2.2: A cross section of the LHC beam pipe is shown above [8]. Note there are two different beam pipes, each with its own set of superconducting magnets. Channels are present for coolant and instrumentation wires, but a large part of the accelerator mass is the steel and iron designed to resist the large forces generated by the magnetic field during operation.

This minimizes any interaction between the proton beam and any gas in the beam pipe. The operating temperature of the LHC is $1.9\,\text{K}$. This is done by using liquid helium, which has the advantage of being a super-fluid with zero viscosity at that temperature. In addition to lowering the pressure in the beam pipes, this low temperature is required for the superconducting magnets. These magnets are made out of niobium-titanium (NiTi) alloy, which has a critical temperature of $9.25\,\text{K}$, in the absence of a magnetic field, and they produced a magnetic field of $8.3\,\text{T}$ during the previous LHC run [10].

The LHC has a circumference of $27\,\text{km}$, which is designed to reach a collision energy of $\sqrt{s} = 14\,\text{TeV}$. For two colliding particles of the same mass, each particle has an energy of $E = \sqrt{s}/2$ in the center of mass frame [16]. The highest energy used in collisions at this point is $\sqrt{s} = 7\,\text{TeV}$.

In addition, the frequency of bunch crossings will be increased in subsequent runs, from a crossing every 50 ns to a crossing every 25 ns.

2.2   Super Conducting Solenoid

A major feature of the Compact Muon Solenoid is the superconducting solenoid itself. Figure 2.3 shows a representation of the solenoid. This solenoid, like the superconductor of the LHC, is also made out of NbTi alloy. During operation, the solenoid generates a magnetic field of 4 T in the center of the detector. This field lowers the critical temperature of NbTi to 7.30 K, but the solenoid is still cooled to 6.44 K at that time [10].

Along the outer edge of CMS, there is a return yoke for the current that travels through the solenoid. This return yoke also generates a magnetic field, which has a strength of 2 T.

Figure 2.3: Above is an artistic image of the NbTi superconducting solenoid. It is located at the center of CMS, and produces a magnetic field of 4 T during operation [10].

2.3   Silicon Tracker

The detecting component of the silicon tracker is the silicon wafers. An electric potential is applied to the silicon so that electrons removed by ionizing radiation drift to one end, creating a pulse when reaching an electrode. There are three layers of pixel detectors closest to the interaction

Figure 2.4: The geometric layout is shown above. Below that is the efficiency of the tracker at various values of η [10].

point, surrounded by ten layers of silicon strips. Figure 2.4 shows the range and layout of the three layers of pixel detectors.

Thanks to the strong magnetic field produced by the solenoid, the central tracker has a high resolution of the charge to mass ratio of particles passing through. This can be done by measuring the radius of curvature, r, of the particle's path and using the following expression.

$$\frac{p}{q} = Br \qquad (2.4)$$

The inner pixel detector, which is shown in Figure 2.5, has a resolution of $10\,\mu\mathrm{m}$ by $20\,\mu\mathrm{m}$ [6]. In addition to measuring the curvature, this allows the pixel detector to determine the point of origin

13

of charged particles with an accuracy on the order of millimeters. This is useful in removing some additional particles that are produced from pileup since the length along which collisions occur is often a few centimeters.

Figure 2.5: A 3-D view of the pixel detector in Figure 2.4 is shown here [6]. Note the endcap components that are meant to capture particles with large η.

2.4   Electromagnetic and Hadronic Calorimeters

There are two sets of calorimeters that are used to directly measure the energy of all of the visible particles, excluding muons. Both make use of scintillators, which give off a number of photons proportional to the amount of energy deposited, to make this measurement [17]. The Electromagnetic Calorimeter (ECAL) measures all of the energy of the photons and electrons generated in an event. The ECAL is made up of $PbWO_4$. This scintillating crystal has a density of $8.28\,g/cm^3$, which allows it to capture the majority of the energy of high-energy electrons and photons. There are approximately four photons produced in the scintillator for every MeV deposited in the detector, and 85% of that energy is read by photodetectors within 25 ns of when the first energy is deposited [10]. An image of the crystal is given in Figure 2.6. While the detector is operating, the ECAL is calibrated by injecting laser light in between bunch crossings.

The Hadronic Calorimeter (HCAL) needs to stop much heavier particles in a comparable amount of space. Towards this end, the HCAL is made up of mostly brass absorber material. This causes

Figure 2.6: A PbWO$_4$ crystal that is used in the ECAL of CMS is shown. A photo detector is attached to the end [10].

showers of lower energy particles that interact with scintillators and cables that are placed throughout the HCAL, but highly localized near the outer edge. The HCAL is calibrated via and test beam and radioactive sources prior to installation [6]. The energy resolution of the HCAL is not as important as the ECAL's resolution because CMS is designed to perform jet reconstruction using particle flow, where each individual particle is reconstructed and added to the jet [1]. As long as the particles' energies are good enough to distinguish different species of particles, the momentum of each individual particle can be determined in the central tracker.

2.5   Muon Chambers

The final main component of the CMS detector is the outer muon chambers. Like the silicon tracker, the muon chambers are only designed to track particles, not measure their energy directly. The muon chambers are particularly important for the detector because of the Higgs to ZZ or $ZZ*$ (where $Z*$ is a virtual Z) to four leptons decay mode. In the event that all four leptons are muons, the Higgs has relatively little background since only muons reach the outer chamber, and it is easy to reconstruct the invariant mass correctly.

The barrel muon system is made up of drift tube (DT) chambers that follow the design shown in Figure 2.7. There are approximately 172000 wires in that portion of the detector. Each DT chamber is filled with a gas that can be ionized by the passage of muons that consists of 15% Ar and

85% CO_2 gas. The ratios were chosen in order to reduce drift time, which is 380 ns in the current configuration, but have adequate sensitivity. There is a voltage of +3600 V applied to the anode wire, +1800 V applied to the cathode strips, and -1200 V applied to the anode wires [6]. When ionization occurs inside the DT chamber, the electric field causes a cascade of additional ionization to give enough charge to be easily measured [17]. This makes it difficult to measure the amount of energy deposited in the DT, but the location is more important.

Figure 2.7: The location of a muon is determined using a simple drift tube with 15% Ar and 85% CO_2 gas. Muons passing through ionize the gas, and charged particles drift towards their appropriate electrodes. This causes a drop in voltage, indicating the presence of ionizing radiation [6].

The DT chambers are long strips, so in order to locate the muons, the DT chambers are crossed with sections perpendicular to each other to find the muon location with more accuracy. A side view of this arrangement is shown in Figure 2.8. In addition to measuring the curvature of the muons' paths to determine momentum and charge, this system allows for the elimination of background. Muons from cosmic rays are constantly penetrating the detector. However, by combining the silicon tracker and the muon chambers, it is possible to see if a muon did not originate from the beam line. In addition, there are frequently neutrons and photons escape the calorimeters and absorbers and cause ionization in the muon chambers. This is the reason why many of the layers in the muon chambers are redundant, as can be seen in Figure 2.8. The dead time in the DT chambers is

relatively long compared to event frequency, but there is rarely an entire row of muon chambers out of the detector that are unresponsive at any one time, unless a real muon just passed through. This helps distinguish muons from the collision from any other particles that might cause ionization in the DT chambers.

Figure 2.8: Layers of drift tube are crossed in the muon chambers in order to get a better resolution of the muon location [6].

Chapter 3

Characterizing the W Kinematics

This chapter marks the beginning of the original work done for this thesis. Referring back to Figure 1.5, there are two main characteristics of the desired WH to $q\bar{q}+$ invisibles processes. One is that there are quark jets that can be used to reconstruct a W boson, and the other is a high value for MET. Note that there could be a ZH process with a similar final state. However, emphasis on the W process is due to the subject of this chapter. In order to fully reconstruct the event, it is important to get accurate values for the W kinematics. There are different ways to cluster jets that can be compared in order to achieve this. Getting more accurate jet characteristics will ultimately lead to a better measurement of the MET from the Higgs.

In high energy collisions, quarks are often separated or produced in a manner that leaves them relatively isolated. Since they are colored particles, these quarks often conjure other particles from the vacuum to satisfy the natural requirement of detected particles being colorless [16]. This results in a spray of hadrons traveling in generally the same direction, which is called a jet. Jet reconstruction algorithms focus on identifying which group of particles came from the same isolated quark. Accurately including the correct particles allows for more accurate kinematics of the isolated quarks, and for more accurate characterization of the quarks' parent particle. However, this process is made more complicated by the fact that events of interest are not isolated. Within a single crossing of the proton bunches within the LHC, there are many collisions occurring that can create hadrons or gluons. In high pileup events, this can cause overlap between two different jets. It is important to be able to identify which of these particles came from the primary vertex of interest and which particles are created at other vertices and should be removed from the jet.

This study seeks to identify an optimal method of selecting hadrons in jet reconstruction. In particular, the dijet spectra resulting from hadronic W boson decays were compared for different jet algorithms. The primary figure of merit for this study was the resolution, or the width, of the W

resonance peak in the dijet system's invariant mass spectrum. For each reconstruction algorithm, the mass distribution was fitted with a function that was a combination of a quadratic background and a Gaussian signal, such as the fit shown in Fig. 3.1. The parameters of these fits were compared for different jet cone sizes, ΔR, which is defined by the following equation.

$$\Delta R = \sqrt{(\Delta \eta)^2 + (\Delta \phi)^2} \tag{3.1}$$

Using this cone size, particles are first clustered into jets using the anti-k_T algorithm, which has been shown to give well-defined boundaries for jets, even in the presence of lots of soft radiation [9]. This is done by essentially clustering particles around the highest energy quarks first. This means that quarks from a boosted W will not have their measured momentum affected drastically by the pileup that occurs in the vicinity.

Figure 3.1: Part of the mass spectrum for the anti-kT algorithm is using $\Delta R = 0.5$. The W resonance is simply modeled as a Gaussian peak on top of a quadratic background.

3.1   Jet Selection Criteria

The event studied was the following decay.

$$t\bar{t} \to b\bar{b}W^+W^- \to b\bar{b}\mu\nu_\mu q\bar{q} \tag{3.2}$$

This decay is also shown in Figure 3.2. This event differs from the one shown in Figure 1.5 but is often used as a benchmark for studying W bosons since its cross section is relatively high, and there are many events that can be used in this study. The signature used to identify this process is also distinct, requiring the presence of at least two b-tagged jets, and exactly one isolated muon. Jets can be b-tagged by identifying a secondary vertex that is displaced from the beam line due to the relatively long lifetime of the b quark. A muon is considered isolated if it is a distance away from all other particles that is on the order of the jet cone size. Potential quark jets from one W boson were selected. One important thing to look for was the presence of "fat jets." The two quarks in Equation (3.2) do not always create two jets. If the W they originate from is highly boosted, the opening angle between the quarks is small. If this opening angle is below the cone size, ΔR, used in the jet reconstruction, the two quarks are grouped into one fat jet [3].

Figure 3.2: By studying the production and decay of $t\bar{t}$, the presence of W bosons is ensured. Optimizing the resolution of the hadronic W, with its decay products circled, has led to new jet parameters for CMS.

After picking out event candidates for the $t\bar{t}$ decay and identifying the two b-tagged jets, the remaining jets were required to pass a series of cuts. First, the jets needed to be located in the hadronic hemisphere. The two top quarks would be boosted in opposite directions, so their decay products would be as well. To enforce this directional dependence, the following equation had to

be satisfied, where the vectors are the momenta three vectors.

$$\vec{p}_{jet} \cdot (\vec{p}_\mu + \vec{p}_{MET} + \vec{p}_b) < 0 \tag{3.3}$$

Of the jets that satisfy Eq. (3.3), the jets were first scanned for potential fat jet candidates. An invariant mass cut of $60\,\mathrm{GeV}$ and a MET cut of $30\,\mathrm{GeV}$ needed to be passed for each fat jet. This mass cut is the reason for the lower limit of mass used for the fit in Fig. 3.1. Since there are no fat jets with a mass of $60\,\mathrm{GeV}$, the contour of the plot is less smooth below the cut. Therefore, the mass window used in each fit was $60\,\mathrm{GeV}$ to $140\,\mathrm{GeV}$. A p_T cut of $250\,\mathrm{GeV}$ was used for fat jet candidates.

If none of the jets passed these criteria, each possible combination of dijets satisfying Eq. (3.3) was compared. The pair of jets that had the greatest transverse momentum, p_T, falling within the mass window of $30\,\mathrm{GeV} < m < 250\,\mathrm{GeV}$ were selected as the potential dijet system from the W decay.

3.2   Comparing Cone Sizes

First, different cone sizes using the anti-k$_T$ algorithm were fitted and compared. Figures 3.3 through 3.6 feature cone sizes from $\Delta R = 0.3$ to $\Delta R = 1.0$ with a resolution of 0.1. In Figure 3.3, the apparent mass of the W boson increases with p_T. As the system becomes more highly boosted, the particles radiated by the system have smaller opening angles and are less divergent. As a result, more particles fall within the fixed cone size, raising their masses. Note that there is also a jump down in the mass spectrum for the larger cone sizes at $p_T = 250\,\mathrm{GeV}$. This is caused by the cuts for selecting fat jets, which were optimized for $\Delta R = 0.5$. For larger cone sizes, fat jets are formed at lower p_T. Looking at Figure 3.4, it can be seen that the best mass resolution is achieved by the smallest cone sizes, $\Delta R < 0.4$.

In Figure 3.5, the mean for each cone size as a function of the Number of Primary Vertices (NPV) are plotted. Each mean is relatively constant, making these means good characteristic values for the each cone size. This gives the response and correction factor that will need to be applied for each cone, which for the smaller cone sizes, is only a few percent of the W mass. Figure 3.6 also

Figure 3.3: The apparent W mass as a function of p_T is shown.

Figure 3.4: The W mass resolution as a function of p_T is shown.

supports the conclusion made using Fig. 3.4 that the best mass resolution is acquired using smaller cone sizes. In this case it could be recommended to use $\Delta R < 0.5$.

Figure 3.5: The apparent W mass as a function of NPV is shown.

Figure 3.6: The W mass resolution as a function of NPV is shown.

3.3 Comparison of Grooming Algorithms

Accurately including the correct particles in a jet allows for more accurate kinematics of the isolated quarks, and for more accurate characterization of the quarks' parent particle. However, this process is made more complicated by the fact that events of interest are not isolated. Within a single crossing of the proton bunches within the LHC, there are many events occurring simultaneously that can create hadrons or gluons. In high pileup events, this can cause overlap between multiple jets, as seen in Section 4.1. It is important to be able to accurately identify which of these particles came from the process of interest and which particles are created at other interaction points and

22

should be removed from the jet.

The other comparison made was between grooming algorithms. The three algorithms compared were pruning, filtering, and trimming. After clustering particles in a jet, grooming is applied. Pruning, the most aggressive of the three grooming algorithms, looks at each recombination that has been made. If the opening angle between two particles that were combined is too large or if the fraction of combined p_T that was given to one of the particles in an apparent decay is too small, the softer particle is discarded from the recombination [13]. Filtering and trimming both cluster a jet's particles into subjets with a lower ΔR [18]. Filtering keeps a fixed number of these subjets, discarding the rest. Trimming discards all of the subjets that have less than a certain fraction of the total jet's p_T. Figures 3.7 through 3.10 compare two different parameter sets for pruning and filtering with the ungroomed jet, with four parameter sets for trimming. The parameters are as follows.

- Pruning:

  $p_1 \rightarrow n_{subjets} = 2; z_{cut} = 0.1; d_{cut\_factor} = 0.5$

  $p_2 \rightarrow n_{subjets} = 2; z_{cut} = 0.1; d_{cut\_factor} = 0.2$

- Filtering:

  $f_1 \rightarrow r_{filt} = 0.2; n_{filt} = 3$

  $f_2 \rightarrow r_{filt} = 0.3; n_{filt} = 3$

- Trimming:

  $t_1 \rightarrow r_{trim} = 0.2; p_{T,frac} = 0.05$

  $t_2 \rightarrow r_{trim} = 0.2; p_{T,frac} = 0.03$

  $t_3 \rightarrow r_{trim} = 0.1; p_{T,frac} = 0.03$

  $t_4 \rightarrow r_{trim} = 0.05; p_{T,frac} = 0.03$

In Figure 3.7, pruning keeps the mass of the dijets at a constant rate. The other grooming algorithms follow the same trend seen in the anti-kT algorithm of increasing mass with increasing p_T. However, when looking at the widths, shown in Figure 3.8, the filtering algorithm, followed by the trimming algorithm, seem to give the best mass resolution.

Figure 3.7: The apparent W mass as a function of p_T is shown.

Figure 3.8: The W mass resolution as a function of p_T is shown.

Figure 3.9: The apparent W mass as a function of NPV is shown.

Figure 3.10: The W mass resolution as a function of NPV is shown.

When looking at the trend in Fig. 3.9, pruning once again follows a different trend than what is expected from the results in Section 3.2. This time, the pruning gives a mass value that increases at low numbers of NPV, and decreases at higher pileup. While Fig. 3.7 shows favorable characteristics in the aggressiveness of pruning, Fig. 3.9 shows that pruning may be overly aggressive. Another important point against pruning is that in Fig. 3.10, the best mass resolution still appears to be either filtering or trimming.

3.4 Jet Reconstruction Effects on p_T

Another important thing to consider is just how aggressive each grooming algorithm is being. This can be measured by the difference between the simulated measured value of p_T and the $p_{T,gen}$ of the GenJet. In this case, all of the individual jets in a sample were used to calculate the difference

$p_T - p_{T,gen}$. These points were plotted in a histogram with a difference range of $-50\,\mathrm{GeV}$ to $50\,\mathrm{GeV}$. The mean and RMS was calculated for each of these histograms, with the results shown below.

Figures 3.11 and 3.12 show the mean and RMS as a function of p_T. They show that the pruning algorithms as well as the new trimming parameters are the most aggressive at high p_T. They also give the least consistent differences, having the highest RMS. This will make them the most difficult to correct.

Figure 3.11: The $p_T - p_{T,gen}$ for $\Delta R = 0.5$ cones for different p_T is shown.

Figure 3.12: The RMS of $p_T - p_{T,gen}$ for $\Delta R = 0.5$ cones for different p_T is shown.

Figures 3.13 and 3.14 show the mean and RMS as a function of NPV. These figures actually show that the pruning algorithms center more around the correct p_T when the pileup is greater. However, the comparison of RMS still does not look favorable to pruning or the new trimming parameters. Comparisons of cone sizes on the lost p_T are given in Appendix D.

Figure 3.13: The $p_T - p_{T,gen}$ for $\Delta R = 0.5$ cones for different NPV is shown.

Figure 3.14: The RMS of $p_T - p_{T,gen}$ for $\Delta R = 0.5$ cones for different NPV is shown.

Chapter 4

Optimization of the Higgs Signal

It is not straightforward to detect dark matter because it is largely a search for missing transverse energy (MET). From this alone, it is difficult to tell what portion of the MET is from poor reconstruction, which is the focus of Chapter 3, or what kind of particles might be missing. For example, neutrinos also show up as missing energy, meaning the generation of W bosons or τ leptons also causes a high MET. However, this metric is still discriminatory against most backgrounds, so a trigger is used to filter out events with MET $< 150\,\mathrm{GeV}$. The efficiency of simulated events passing a simulated trigger as a function of reconstructed MET is shown in Figure 4.1. This reveals that a hard cut removing events with MET $< 220\,\mathrm{GeV}$ will prevent any bias caused by the trigger.

Figure 4.1: The trigger efficiency increases at higher $\not{E}_T$, suggesting a cut of $220\,\mathrm{GeV}$.

The most dominant background that we will have when looking at high MET events is the following simple Z decay into neutrinos.

$$Z \rightarrow \nu\bar{\nu} \tag{4.1}$$

Since neutrinos only interact via weak interaction, they are not detected by any of the components of CMS. A boosted Z undergoing this decay could then result in a significant amount of MET. The only way to distinguish this process from a Higgs decay is to search for a hadronic W. This requirement will cut down severely on the background events, allowing a reasonable signal to be predicted and measured. We know that a W will more frequently be present when the Higgs is produced in a WH event, such as the one described in Section 1.2, with an event resulting in ZW being less common when considering all events with a Z. However, events with both a Z and H will have a large number of jets present from other particles and pileup. This chapter will describe how differences in jets were characterized for each case, resulting in a recommended procedure for to optimize the Higgs signal over the Z background. First, some of the characteristics of pileup jets will be described. These pileup jets account for many of the jets that could be mistaken for W daughters. Then, the process of identifying Z events using muons will be described. Using these muon events will allow for accurate reconstruction of the Z kinematics and associated jets during background events. At the end of the chapter, the jet characteristic N-Subjettiness will be described and be used to cut down on the background events to optimize the signal to background.

4.1 Monte Carlo Model Characterizing Pileup Jets

Soft gluon jets can easily be mistaken for quark jets in high pileup events. While a single gluon jet would normally be eliminated by a p_T cut, when these jets overlap, their p_T, which could result in a p_T higher than the prescribed cut.

A simple Monte Carlo model was created where individual jets were created, and the number of jets overlapping was weighted by the appropriate binomial coefficient.

$$\frac{d\sigma}{dp_T} = \begin{cases} Cp_T^{-5}, & \text{if } p_T > 3\,\text{GeV} \\ mp_T + b, & \text{otherwise} \end{cases} \tag{4.2}$$

C, m, and b are all constants that satisfy the conditions that the distribution and its slope is continuous at $p_T = 3\,\text{GeV}$ and that the entire distribution is normalized.

The weighting of the overlap was determined by the ratio between a jet area and the full area

of interest in the detector. This was approximated by the following expression.

$$\frac{A_{jet}}{A_{CMS}} \approx 0.0125 \qquad (4.3)$$

Here A_{jet} and A_{CMS} are the areas of a jet and the CMS detector. This ratio gave the following binomial coefficients.

$$w(NPU, n_{jets}) = \frac{NPU!}{(NPU - n_{jets})!n_{jets}!}(0.0125)^{n_{jets}}(0.9875)^{NPU-n_{jets}} \qquad (4.4)$$

Here, NPU is the number of pileup events and $n_{jets} + 1$ is the number of overlapping gluon jets. The weighting w is the probability that for a given jet in an event with pileup NPU, n_{jets} additional jets are overlapping.

Using Eq. (4.2) to generate jet p_T and Eq. (4.4) to determine the relative weights of multi-jet events, an overall p_T spectrum was simulated. Figure 4.2 shows the contributions of events with up to five jets to the spectrum for events with various NPU. As the pileup increases, it is easy to see that the number of pileup jets occurring at high p_T becomes more significant. This demonstrates one source of jets during background events such as $Z \to \nu\bar{\nu}$.

4.2 The Dominant Z Background

It is through these jets that we can tell the difference between a Z and a WH event. In order to know which jet characteristics are associated with Z production, the process $Z \to \mu\bar{\mu}$ can be studied. Searching for a dimuon signal ensures that the Z boson is present. Each of the muons must have opposite charge. That is, one of the muons must be an anti-muon. From there, the invariant mass of the dimuon system is required to be between $60\,\text{GeV}$ and $120\,\text{GeV}$. The reconstruction of part of the data is compared to Monte Carlo simulation in Figure 4.3. Then the muons are removed from the system, and their p_T is added to the MET. Figure 4.4 shows this same partial data set compared to Monte Carlo. In both cases, though the number of events differ, the profiles match quite well. This is likely due to some incorrect assumptions about the integrated luminosity and detector efficiency when reconstructing events. These events can then be treated as simulated $Z \to \nu\bar{\nu}$ by multiplying by the ratio of the branching ratios. The Z boson decays to neutrinos 5.942

(a) NPU = 10.

(b) NPU = 20.

(c) NPU = 30.

(d) NPU = 40.

Figure 4.2: The simulated p_T distributions for pileup jets are shown for various numbers of pileup events. Only the contribution to p_T of the first five overlapping jets are shown, but the simulation included additional overlapping.

times more often than it decays to muons [7].

4.3   Event Selection

As mentioned in Section 1.2, there are several characteristics that we look for in a WH event. First, a trigger is used that have the characteristic efficiency as a function of MET shown in Figure 4.1. Then, any leptons in an event cause the event to be vetoed, except in cases where we used a simulated $V \to \mu\bar{\mu}$.

Figure 4.3: The plots above compare the $Z \to \mu\bar{\mu}$ peak in part of the data to relative contributions to this peak by different processes in Monte Carlo simulations.

Figure 4.4: By taking out the muons in the $Z \to \mu\bar{\mu}$ events, the distribution of MET for $Z \to \nu\bar{\nu}$ can be characterized.

30

4.4   Description and Use of N-Subjettiness

N-Subjettiness is a set of variables that is used to determine a likely number of subjets within a larger jet. The definition is given as follows [19].

$$\tau_N = \frac{1}{d_0} \sum_k p_{T,k} \min\{\Delta R_{1,k}, \Delta R_{2,k}...\Delta R_{N,k}\} \tag{4.5}$$

The normalization d_0 is defined as follows.

$$d_0 = \sum_k p_{T,k} R_0 \tag{4.6}$$

The summation takes place over the k particles that make up a jet. $\Delta R_{i,k}$ is the distance defined in Equation (3.1) between a candidate subjet and the particle in question. R_0 is the cone size of the larger jet.

By examining Equation (4.5), it is apparent that τ_N for a jet that is made up of N subjets will be significantly less than τ_{N-1}. However, the particles should not be distributed in such a way that τ_{N+1} will be much less. Therefore, a good metric to determine if a large jet has N subjets is to require τ_N/τ_{N-1} to be low. Figure 4.5 shows the differences between a merged jet from a hadronic W and pileup jets like those simulated in Section 4.1.

Applying cuts of τ_2/τ_1 on jets should then eliminate fat jets that are actually randomly distributed pileup jets, while keeping jets that originated from a hadronic W. Figure 4.6 shows a ROC curve scanning over τ_2/τ_1 cuts. On one axis is the efficiency of the W signal making it past the N-Subjettiness cut, and the other axis has $1 - \epsilon_{bkgd}$, where ϵ_{bkgd} is the background efficiency for passing the cut. The signal is taken only to be VH events, where V, a vector boson, is either a W or a Z. It is desirable to have both values to be close to one. Figure 4.7 shows the ROC curve counting only dijet events, and Figure 4.8 considers only monojet events. From these three figures, it is confirmed that N-Subjettiness works best with merged jets, as opposed to dijets. Also, this suggests that it may be best to use the substructure of single jets with large cone sizes to select for W bosons.

Another way to see the effectiveness of utilizing the N-Subjettiness cuts is through the signal

Figure 4.5: N-Subjettiness distinguishes between quark jets and pileup, or QCD, jets by searching for the two distinct locations of high p_T. This is due to the two quark jets from a boosted hadronic W being merged into a single jet [19].

significance. This is defined by the following expression.

$$\text{Significance} = \frac{S}{\sqrt{B}} \tag{4.7}$$

S is the number of signal, or Higgs, events that passed the cut, and B is the number of background events, which mostly consist of Z production. Figure 4.9 shows the signal significance for all events that pass a MET cut of $200\,\text{GeV}$ and an N-Subjettiness cut of $(\tau_2/\tau_1)_{jets} < (\tau_2/\tau_1)_{cut}$. Figures 4.10 and 4.11 shows the signal significance for events that pass the same cuts for events that appear to have a W that decays into two jets and a single jet, respectively. The values of lower τ_2/τ_1 should be ignored because they have poor statistics. In general, the larger cone sizes are once again better for selecting events due to their tendency to form merged jets.

As an example of what the MET distribution looks like after cuts, Figure 4.12 shows the distribution using a cone size of $\Delta R = 0.9$ and $(\tau_2/\tau_1)_{cut} < 0.44$. The data has not yet been

Figure 4.6: Above is the ROC curve for events that pass a MET cut of $200\,\mathrm{GeV}$ and a varied τ_2/τ_1 cut. This is considering all events.

Figure 4.7: Above is the ROC curve for events that pass a MET cut of $200\,\mathrm{GeV}$ and a varied τ_2/τ_1 cut for apparent W to dijet events.

Figure 4.8: Above is the ROC curve for events that pass a MET cut of $200\,\mathrm{GeV}$ and a varied τ_2/τ_1 cut for apparent W to single jet events.

Figure 4.9: Above is the signal significance of the signal for events that pass a MET cut of $200\,\text{GeV}$ and a varied τ_2/τ_1 cut. This is considering all events.

Figure 4.10: Above is the signal significance of the signal for events that pass a MET cut of $200\,\text{GeV}$ and a varied τ_2/τ_1 cut for apparent W to dijet events.

Figure 4.11: Above is the signal significance of the signal for events that pass a MET cut of $200\,\text{GeV}$ and a varied τ_2/τ_1 cut for apparent W to single jet events.

unblinded due to incomplete investigation of the grooming algorithms in jet selection. However, this figure shows the excess missing energy that should be seen in the event that the Higgs decays to invisible particles over the background events. Also, using the mass calculated by grooming algorithms, such as t_1 shown in in Figure 4.13, can give more accurate mass of the parent particle and better signal significance.

Figure 4.12: Above is the simulated MET distribution in the event that the Higgs decays to invisible products. The jet selection is based off of merged jets with $\Delta R = 0.9$ and $(\tau_2/\tau_1)_{cut} < 0.44$. The duller background colors are all predictions made by the standard model. Any excess data in the brightly colored Higgs region would indicate a significant branching ratio of Higgs to invisibles.

Figure 4.13: Above is the simulated MET distribution in the event that the Higgs decays to invisible products. The jet selection is based off of merged jets with $\Delta R = 0.9$, $(\tau_2/\tau_1)_{cut} < 0.44$, and the jet mass is determined by the t_1 grooming algorithm. After other grooming algorithms are investigated, and results seem consistent, the data will be unblinded. The amount of Higgs excess is significant. However, this weighting is based on the full cross section, and will be significantly reduced by the branching ratio.

Chapter 5

Conclusions

Based on studies of $t\bar{t} \to b\bar{b}W^+W^- \to b\bar{b}\mu\nu_\mu q\bar{q}$, resolution of the W characteristics is optimized when using smaller cone sizes and dijet systems. They consistently give the best resolution as a function of the Number of Primary Vertices (NPV). However, it appears to be advantageous to use slightly larger cone sizes for greater p_T. An interesting result from this study is that smaller substructure size in t_4 led to very stable jet masses. (See Appendix A.) Unfortunately, this grooming algorithm has poor resolution. It also tends to lose a fair amount of p_T from aggressive trimming of particles, as shown in Figure 3.13. An algorithm that suffered from even worse resolution and aggressive reduction of p_T was pruning. Pruning has been the CMS default due to its resilience in change of mass as p_T increases. However, this would lead to less accurate measurements of MET.

In contrast to using small cones to measure the W mass, using N-Subjettiness on large cones appears to be a promising method of selecting signal events. While statistics are poor for the data from the lower energy runs of the LHC, this study would suggest using cone sizes $\Delta R > 0.6$ with $\tau_2/\tau_1 < 0.5$.

5.1   Future Work

Obvious future work would be to run simulations with addition events and improve the di-muon Z selection efficiency. This will improve the uncertainty in the predicted distribution, allowing for unblinding of the data shortly afterwards. Of course, the upcoming runs of the LHC will also add data for this search. The higher energy collisions will result in much greater MET, allowing for interesting events to be better distinguished from poorly reconstructed events. Also, the integrated luminosity of the upcoming runs will be over an order of magnitude higher than the current $19.7\,\text{fb}^{-1}$, meaning it will be much easier to observe any partial signal shown in Figure 4.13 in the actual data.

Other future work could focus on improving the implementation of grooming and N-Subjettiness together. Grooming is already shown to be a superior method of determining the di-jet mass spectrum. In addition, the value of τ_2 could be constrained during high MET events for the same reason that hadronic Ws coalesce into single jets at higher p_T. The subjets of each quark will be highly localized. In addition, jet grooming will affect the number and the distribution of particles in a jet, also affecting N-Subjettiness. This effect will be significant in higher pileup events, requiring extensive study in future LHC runs.

Appendix A

Comparison of Cone Sizes

The following plots are all of the comparisons of the W mass for cone sizes using different grooming algorithms. To see the individual fits for each point in the plots, visit `http://personal.psu.edu/dra5110/Summer13/cone-fitu4-quad/` or download the directory at `http://personal.psu.edu/dra5110/Summer13/cone-fitu4-quad.tar`

Figure A.1:
cone-app/f1˙mjj˙npv˙Nwidth.png

Figure A.2:
cone-app/f1˙mjj˙npv˙Nwidth˙line.png

Figure A.3:
cone-app/f1˙mjj˙npv˙mean.png

Figure A.4:
cone-app/f1˙mjj˙npv˙width.png

Figure A.5:
cone-app/f1˙mjj˙ptjj˙Nwidth.png

Figure A.6:
cone-
app/f1˙mjj˙ptjj˙Nwidth˙line.png

Figure A.7:
cone-app/f1˙mjj˙ptjj˙mean.png

Figure A.8:
cone-app/f1˙mjj˙ptjj˙width.png

Figure A.9:
cone-app/f2˙mjj˙npv˙Nwidth.png

Figure A.10:
cone-
app/f2˙mjj˙npv˙Nwidth˙line.png

Figure A.11:
cone-app/f2˙mjj˙npv˙mean.png

Figure A.12:
cone-app/f2˙mjj˙npv˙width.png

Figure A.13:
cone-app/f2˙mjj˙ptjj˙Nwidth.png

Figure A.14:
cone-
app/f2˙mjj˙ptjj˙Nwidth˙line.png

41

Figure A.15:
cone-app/f2˙mjj˙ptjj˙mean.png

Figure A.16:
cone-app/f2˙mjj˙ptjj˙width.png

Figure A.17:
cone-app/mjj˙npv˙Nwidth.png

Figure A.18:
cone-app/mjj˙npv˙Nwidth˙line.png

Figure A.19:
cone-app/mjj˙npv˙mean.png

Figure A.20:
cone-app/mjj˙npv˙width.png

Figure A.21:
cone-app/mjj˙ptjj˙Nwidth.png

Figure A.22:
cone-app/mjj˙ptjj˙Nwidth˙line.png

Figure A.23:
cone-app/mjj˙ptjj˙mean.png

Figure A.24:
cone-app/mjj˙ptjj˙width.png

Figure A.25:
cone-app/p1˙mjj˙npv˙Nwidth.png

Figure A.26:
cone-
app/p1˙mjj˙npv˙Nwidth˙line.png

43

Figure A.27:
cone-app/p1˙mjj˙npv˙mean.png

Figure A.28:
cone-app/p1˙mjj˙npv˙width.png

Figure A.29:
cone-app/p1˙mjj˙ptjj˙Nwidth.png

Figure A.30:
cone-
app/p1˙mjj˙ptjj˙Nwidth˙line.png

Figure A.31:
cone-app/p1˙mjj˙ptjj˙mean.png

Figure A.32:
cone-app/p1˙mjj˙ptjj˙width.png

Figure A.33:
cone-app/p2˙mjj˙npv˙Nwidth.png

Figure A.34:
cone-app/p2˙mjj˙npv˙Nwidth˙line.png

Figure A.35:
cone-app/p2˙mjj˙npv˙mean.png

Figure A.36:
cone-app/p2˙mjj˙npv˙width.png

Figure A.37:
cone-app/p2˙mjj˙ptjj˙Nwidth.png

Figure A.38:
cone-app/p2˙mjj˙ptjj˙Nwidth˙line.png

45

Figure A.39:
cone-app/p2˙mjj˙ptjj˙mean.png

Figure A.40:
cone-app/p2˙mjj˙ptjj˙width.png

Figure A.41:
cone-app/t1˙mjj˙npv˙Nwidth.png

Figure A.42:
cone-
app/t1˙mjj˙npv˙Nwidth˙line.png

Figure A.43:
cone-app/t1˙mjj˙npv˙mean.png

Figure A.44:
cone-app/t1˙mjj˙npv˙width.png

Figure A.45:
cone-app/t1˙mjj˙ptjj˙Nwidth.png

Figure A.46:
cone-
app/t1˙mjj˙ptjj˙Nwidth˙line.png

Figure A.47:
cone-app/t1˙mjj˙ptjj˙mean.png

Figure A.48:
cone-app/t1˙mjj˙ptjj˙width.png

Figure A.49:
cone-app/t2˙mjj˙npv˙Nwidth.png

Figure A.50:
cone-
app/t2˙mjj˙npv˙Nwidth˙line.png

47

Figure A.51:
cone-app/t2˙mjj˙npv˙mean.png

Figure A.52:
cone-app/t2˙mjj˙npv˙width.png

Figure A.53:
cone-app/t2˙mjj˙ptjj˙Nwidth.png

Figure A.54:
cone-
app/t2˙mjj˙ptjj˙Nwidth˙line.png

Figure A.55:
cone-app/t2˙mjj˙ptjj˙mean.png

Figure A.56:
cone-app/t2˙mjj˙ptjj˙width.png

Figure A.57:
cone-app/t3˙mjj˙npv˙Nwidth.png

Figure A.58:
cone-
app/t3˙mjj˙npv˙Nwidth˙line.png

Figure A.59:
cone-app/t3˙mjj˙npv˙mean.png

Figure A.60:
cone-app/t3˙mjj˙npv˙width.png

Figure A.61:
cone-app/t3˙mjj˙ptjj˙Nwidth.png

Figure A.62:
cone-
app/t3˙mjj˙ptjj˙Nwidth˙line.png

Figure A.63:
cone-app/t3˙mjj˙ptjj˙mean.png

Figure A.64:
cone-app/t3˙mjj˙ptjj˙width.png

Figure A.65:
cone-app/t4˙mjj˙npv˙Nwidth.png

Figure A.66:
cone-app/t4˙mjj˙npv˙Nwidth˙line.png

Figure A.67:
cone-app/t4˙mjj˙npv˙mean.png

Figure A.68:
cone-app/t4˙mjj˙npv˙width.png

Figure A.69:
cone-app/t4˙mjj˙ptjj˙Nwidth.png

Figure A.70:
cone-
app/t4˙mjj˙ptjj˙Nwidth˙line.png

Figure A.71:
cone-app/t4˙mjj˙ptjj˙mean.png

Figure A.72:
cone-app/t4˙mjj˙ptjj˙width.png

Appendix B

Comparison of Algorithms

The following plots are all of the comparisons of the W mass for algorithms at different cone sizes. To see the individual fits for each point in the plots, visit `http://personal.psu.edu/dra5110/ Summer13/algo-fitu4-quad/` or download the directory at `http://personal.psu.edu/dra5110/ Summer13/algo-fitu4-quad.tar`

Figure B.1:
algo-app/R03˙mjj˙npv˙Nwidth.png

Figure B.2:
algo-
app/R03˙mjj˙npv˙Nwidth˙line.png

Figure B.3:
algo-app/R03˙mjj˙npv˙mean.png
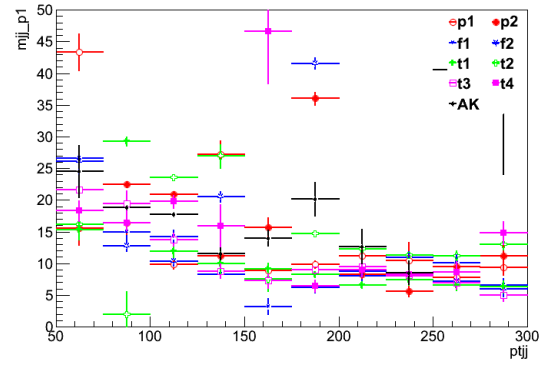

Figure B.4:
algo-app/R03˙mjj˙npv˙width.png

Figure B.5:
algo-app/R03˙mjj˙ptjj˙Nwidth.png

Figure B.6:
algo-
app/R03˙mjj˙ptjj˙Nwidth˙line.png
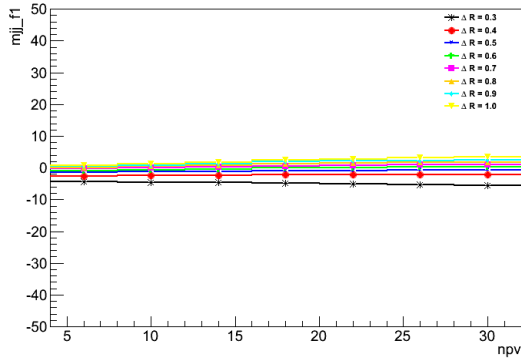

Figure B.7:
algo-app/R03˙mjj˙ptjj˙mean.png

Figure B.8:
algo-app/R03˙mjj˙ptjj˙width.png
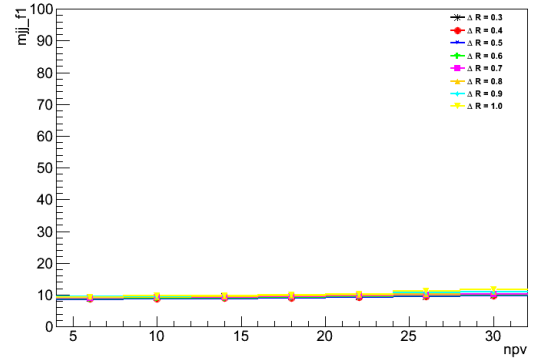
Figure B.9:
algo-app/R04˙mjj˙npv˙Nwidth.png

Figure B.10:
algo-
app/R04˙mjj˙npv˙Nwidth˙line.png

Figure B.11:
algo-app/R04˙mjj˙npv˙mean.png

Figure B.12:
algo-app/R04˙mjj˙npv˙width.png

Figure B.13:
algo-app/R04˙mjj˙ptjj˙Nwidth.png

Figure B.14:
algo-
app/R04˙mjj˙ptjj˙Nwidth˙line.png

Figure B.15:
algo-app/R04˙mjj˙ptjj˙mean.png

Figure B.16:
algo-app/R04˙mjj˙ptjj˙width.png

Figure B.17:
algo-app/R05˙mjj˙npv˙Nwidth.png

Figure B.18:
algo-
app/R05˙mjj˙npv˙Nwidth˙line.png

Figure B.19:
algo-app/R05˙mjj˙npv˙mean.png

Figure B.20:
algo-app/R05˙mjj˙npv˙width.png

Figure B.21:
algo-app/R05˙mjj˙ptjj˙Nwidth.png

Figure B.22:
algo-
app/R05˙mjj˙ptjj˙Nwidth˙line.png

Figure B.23:
algo-app/R05˙mjj˙ptjj˙mean.png

Figure B.24:
algo-app/R05˙mjj˙ptjj˙width.png

Figure B.25:
algo-app/R06˙mjj˙npv˙Nwidth.png

Figure B.26:
algo-
app/R06˙mjj˙npv˙Nwidth˙line.png

Figure B.27:
algo-app/R06˙mjj˙npv˙mean.png

Figure B.28:
algo-app/R06˙mjj˙npv˙width.png

Figure B.29:
algo-app/R06˙mjj˙ptjj˙Nwidth.png

Figure B.30:
algo-
app/R06˙mjj˙ptjj˙Nwidth˙line.png

Figure B.31:
algo-app/R06˙mjj˙ptjj˙mean.png

Figure B.32:
algo-app/R06˙mjj˙ptjj˙width.png

57

Figure B.33:
algo-app/R07˙mjj˙npv˙Nwidth.png

Figure B.34:
algo-
app/R07˙mjj˙npv˙Nwidth˙line.png

Figure B.35:
algo-app/R07˙mjj˙npv˙mean.png

Figure B.36:
algo-app/R07˙mjj˙npv˙width.png

Figure B.37:
algo-app/R07˙mjj˙ptjj˙Nwidth.png

Figure B.38:
algo-
app/R07˙mjj˙ptjj˙Nwidth˙line.png

Figure B.39:
algo-app/R07˙mjj˙ptjj˙mean.png

Figure B.40:
algo-app/R07˙mjj˙ptjj˙width.png

Figure B.41:
algo-app/R08˙mjj˙npv˙Nwidth.png

Figure B.42:
algo-
app/R08˙mjj˙npv˙Nwidth˙line.png

Figure B.43:
algo-app/R08˙mjj˙npv˙mean.png

Figure B.44:
algo-app/R08˙mjj˙npv˙width.png

Figure B.45:
algo-app/R08˙mjj˙ptjj˙Nwidth.png

Figure B.46:
algo-
app/R08˙mjj˙ptjj˙Nwidth˙line.png

Figure B.47:
algo-app/R08˙mjj˙ptjj˙mean.png

Figure B.48:
algo-app/R08˙mjj˙ptjj˙width.png

Figure B.49:
algo-app/R09˙mjj˙npv˙Nwidth.png

Figure B.50:
algo-
app/R09˙mjj˙npv˙Nwidth˙line.png

60

Figure B.51:
algo-app/R09˙mjj˙npv˙mean.png

Figure B.52:
algo-app/R09˙mjj˙npv˙width.png

Figure B.53:
algo-app/R09˙mjj˙ptjj˙Nwidth.png

Figure B.54:
algo-
app/R09˙mjj˙ptjj˙Nwidth˙line.png

Figure B.55:
algo-app/R09˙mjj˙ptjj˙mean.png

Figure B.56:
algo-app/R09˙mjj˙ptjj˙width.png

Figure B.57:
algo-app/R10˙mjj˙npv˙Nwidth.png

Figure B.58:
algo-
app/R10˙mjj˙npv˙Nwidth˙line.png

Figure B.59:
algo-app/R10˙mjj˙npv˙mean.png

Figure B.60:
algo-app/R10˙mjj˙npv˙width.png

Figure B.61:
algo-app/R10˙mjj˙ptjj˙Nwidth.png

Figure B.62:
algo-
app/R10˙mjj˙ptjj˙Nwidth˙line.png

Figure B.63:
algo-app/R10˙mjj˙ptjj˙mean.png

Figure B.64:
algo-app/R10˙mjj˙ptjj˙width.png

Appendix C

Comparison of Cone Sizes for p_T Drop

The following plots are all of the comparisons of the p_T difference for cone sizes using different grooming algorithms. To see the individual fits for each point in the plots, visit `http://personal.psu.edu/dra5110/Summer13/cone-ptdiff2/` or download the directory at `http://personal.psu.edu/dra5110/Summer13/cone-ptdiff2.tar`

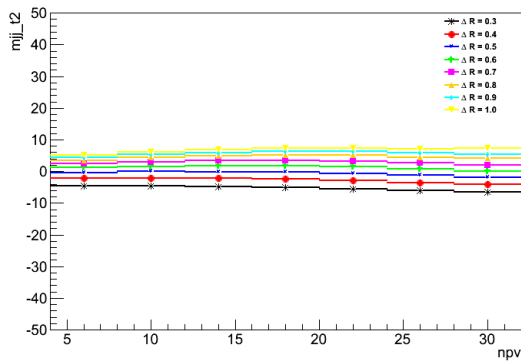

Figure C.1:
cone-ptdiff-
app/f1˙mjj˙npv˙mean.png

Figure C.2:
cone-ptdiff-
app/f1˙mjj˙npv˙width.png
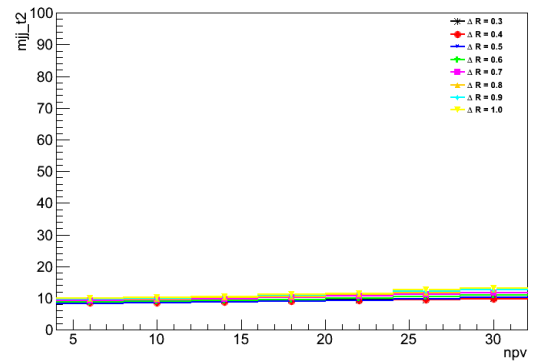
Figure C.3:
cone-ptdiff-
app/f1˙mjj˙ptjj˙mean.png
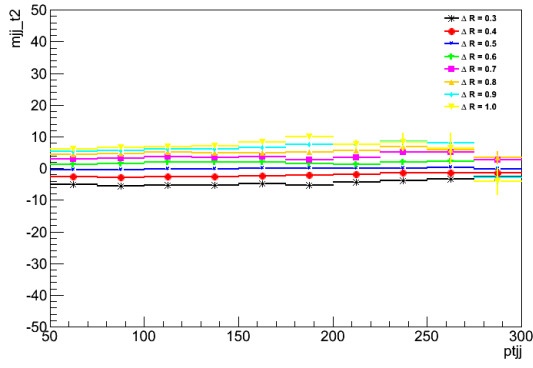

Figure C.4:
cone-ptdiff-
app/f1˙mjj˙ptjj˙width.png
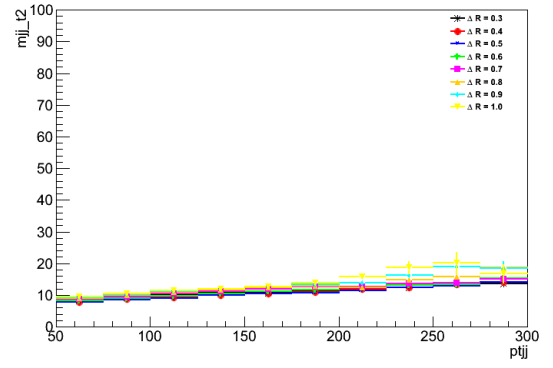

Figure C.5:
cone-ptdiff-
app/f2˙mjj˙npv˙mean.png

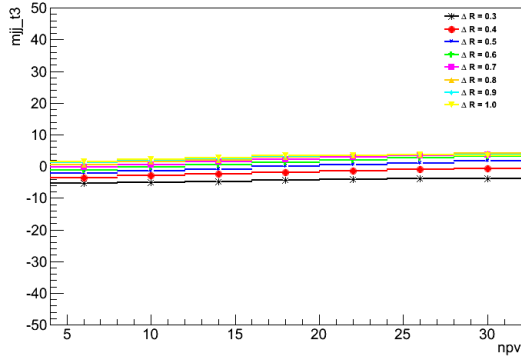**Figure C.6:**
cone-ptdiff-
app/f2˙mjj˙npv˙width.png

Figure C.7:
cone-ptdiff-
app/f2˙mjj˙ptjj˙mean.png

Figure C.8:
cone-ptdiff-
app/f2˙mjj˙ptjj˙width.png

65

Figure C.9:
cone-ptdiff-app/mjj˙npv˙mean.png
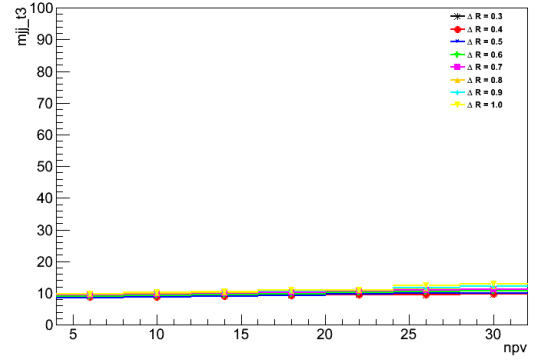

Figure C.10:
cone-ptdiff-app/mjj˙npv˙width.png
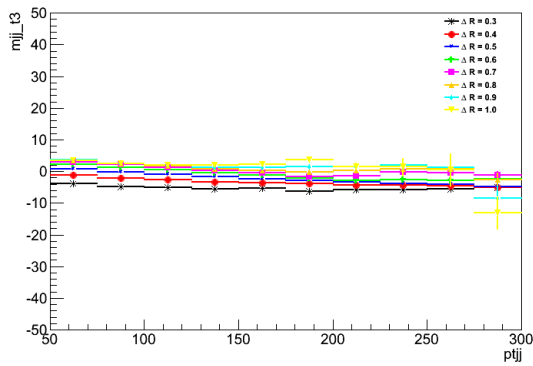

Figure C.11:
cone-ptdiff-app/mjj˙ptjj˙mean.png

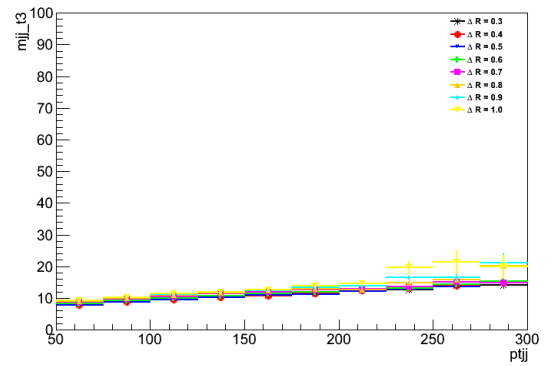
Figure C.12:
cone-ptdiff-app/mjj˙ptjj˙width.png

Figure C.13:
cone-ptdiff-
app/p1˙mjj˙npv˙mean.png

Figure C.14:
cone-ptdiff-
app/p1˙mjj˙npv˙width.png

Figure C.15:
cone-ptdiff-
app/p1˙mjj˙ptjj˙mean.png

Figure C.16:
cone-ptdiff-
app/p1˙mjj˙ptjj˙width.png

Figure C.17:
cone-ptdiff-
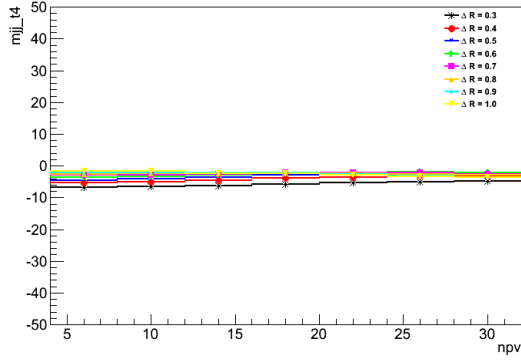app/p2˙mjj˙npv˙mean.png

Figure C.18:
cone-ptdiff-
app/p2˙mjj˙npv˙width.png

Figure C.19:
cone-ptdiff-
app/p2˙mjj˙ptjj˙mean.png

Figure C.20:
cone-ptdiff-
app/p2˙mjj˙ptjj˙width.png

Figure C.21:
cone-ptdiff-
app/t1˙mjj˙npv˙mean.png

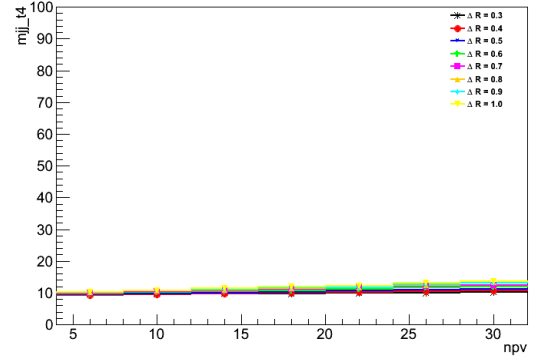**Figure C.22:**
cone-ptdiff-
app/t1˙mjj˙npv˙width.png
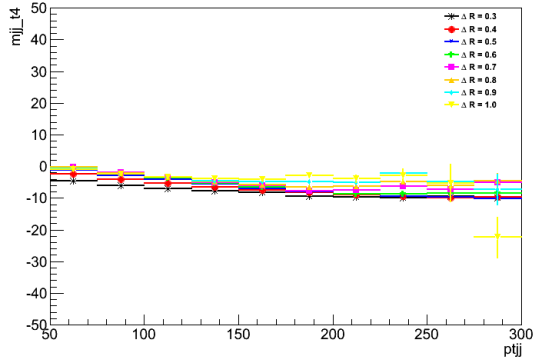

Figure C.23:
cone-ptdiff-
app/t1˙mjj˙ptjj˙mean.png

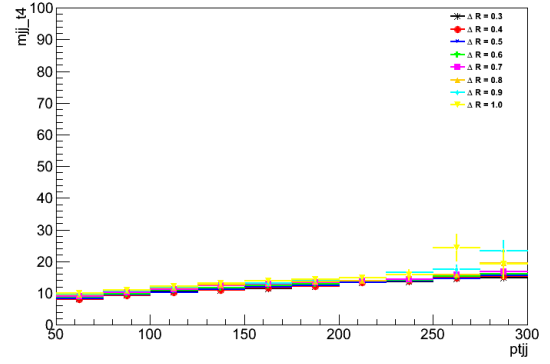

Figure C.24:
cone-ptdiff-
app/t1˙mjj˙ptjj˙width.png

Figure C.25:
cone-ptdiff-
app/t2˙mjj˙npv˙mean.png

Figure C.26:
cone-ptdiff-
app/t2˙mjj˙npv˙width.png

68

Figure C.27:
cone-ptdiff-
app/t2˙mjj˙ptjj˙mean.png

Figure C.28:
cone-ptdiff-
app/t2˙mjj˙ptjj˙width.png

Figure C.29:
cone-ptdiff-
app/t3˙mjj˙npv˙mean.png

Figure C.30:
cone-ptdiff-
app/t3˙mjj˙npv˙width.png

Figure C.31:
cone-ptdiff-
app/t3˙mjj˙ptjj˙mean.png

Figure C.32:
cone-ptdiff-
app/t3˙mjj˙ptjj˙width.png

69

Figure C.33:
cone-ptdiff-
app/t4˙mjj˙npv˙mean.png

Figure C.34:
cone-ptdiff-
app/t4˙mjj˙npv˙width.png

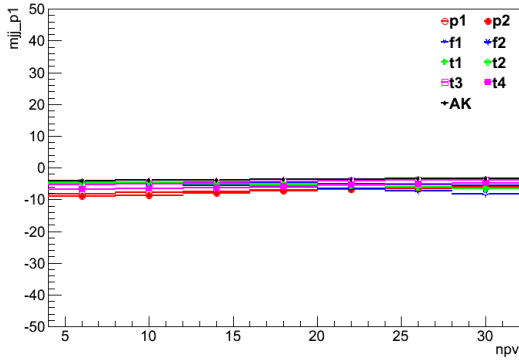**Figure C.35:**
cone-ptdiff-
app/t4˙mjj˙ptjj˙mean.png

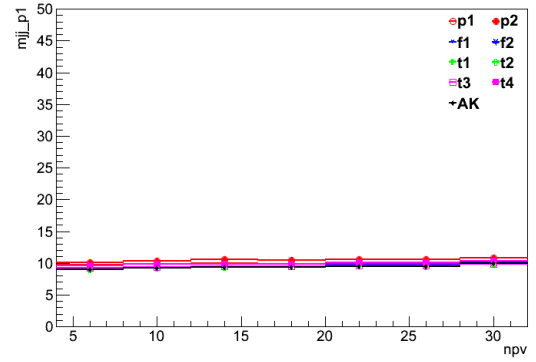
Figure C.36:
cone-ptdiff-
app/t4˙mjj˙ptjj˙width.png

Appendix D
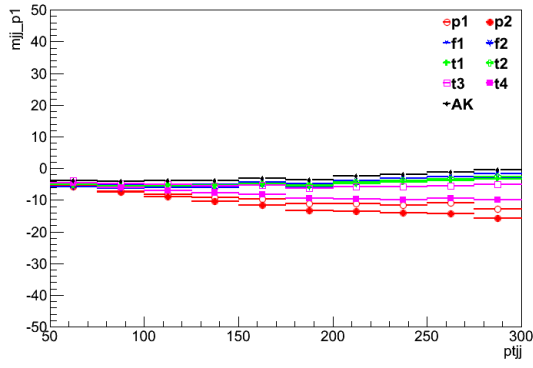
Comparison of Algorithms for p_T Drop

The following plots are all of the comparisons of the p_T difference for algorithms at different cone sizes. To see the individual fits for each point in the plots, visit `http://personal.psu.edu/dra5110/Summer13/algo-ptdiff2/` or download the directory at `http://personal.psu.edu/dra5110/Summer13/algo-ptdiff2.tar`
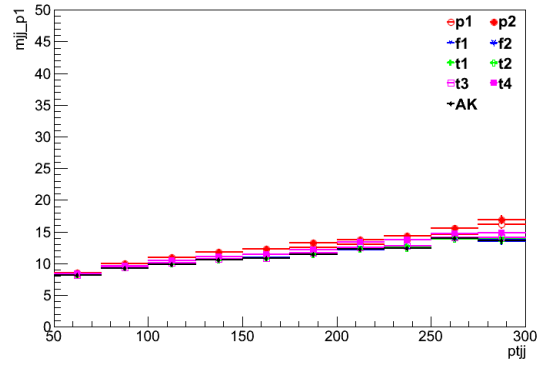

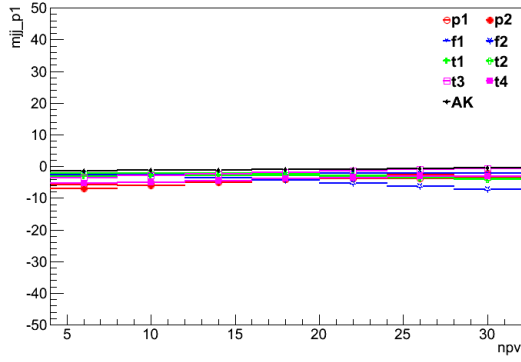**Figure D.1:**
algo-ptdiff-
app/R03˙mjj˙npv˙mean.png

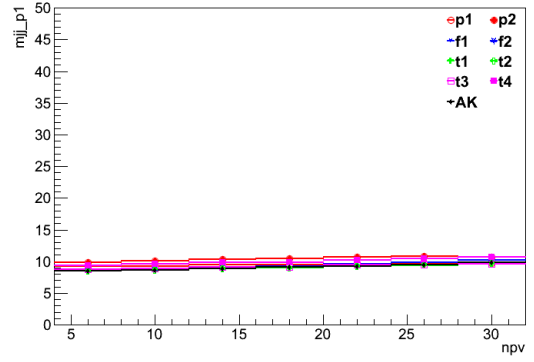**Figure D.2:**
algo-ptdiff-
app/R03˙mjj˙npv˙width.png

Figure D.3:
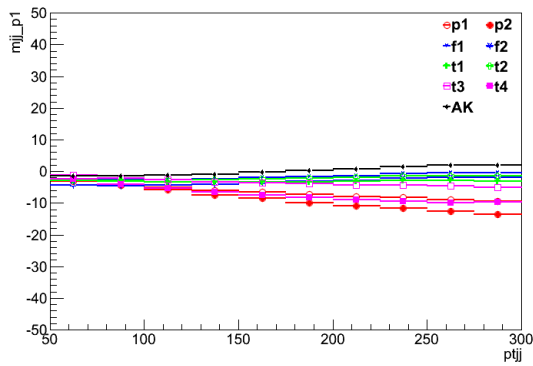algo-ptdiff-
app/R03˙mjj˙ptjj˙mean.png

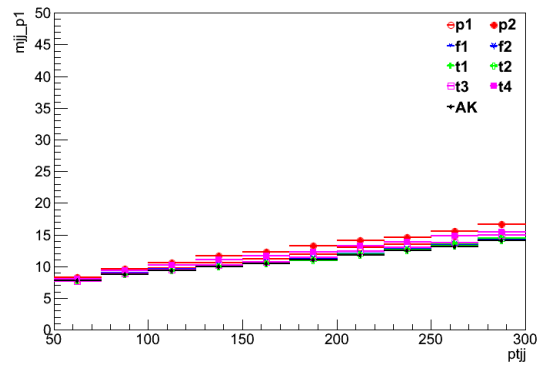**Figure D.4:**
algo-ptdiff-
app/R03˙mjj˙ptjj˙width.png

Figure D.5:
algo-ptdiff-
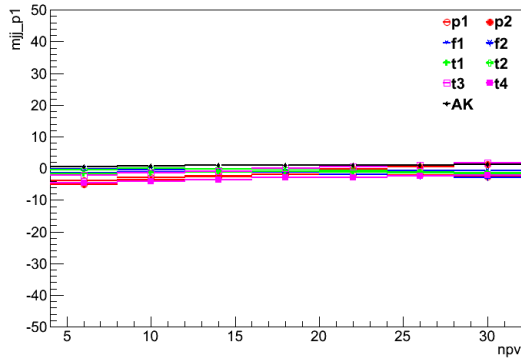app/R04˙mjj˙npv˙mean.png

Figure D.6:
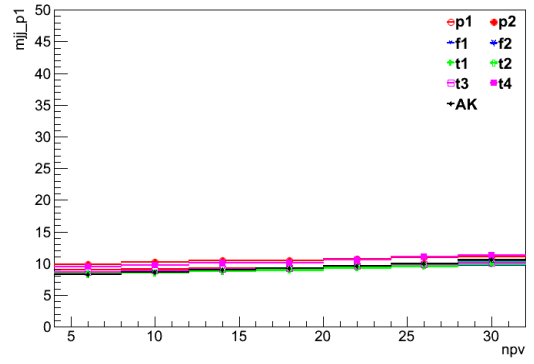algo-ptdiff-
app/R04˙mjj˙npv˙width.png

Figure D.7:
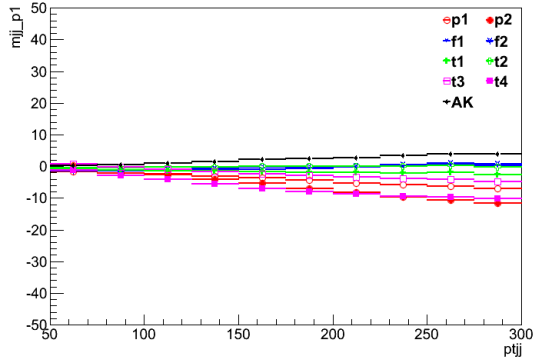algo-ptdiff-
app/R04˙mjj˙ptjj˙mean.png

Figure D.8:
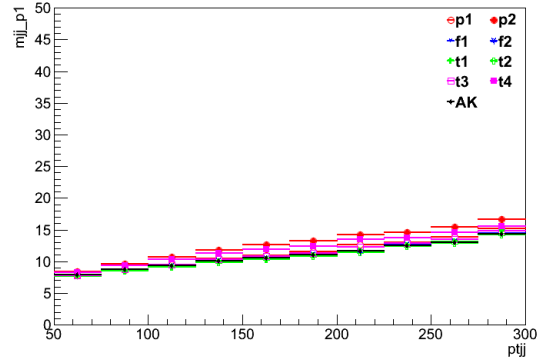algo-ptdiff-
app/R04˙mjj˙ptjj˙width.png

72

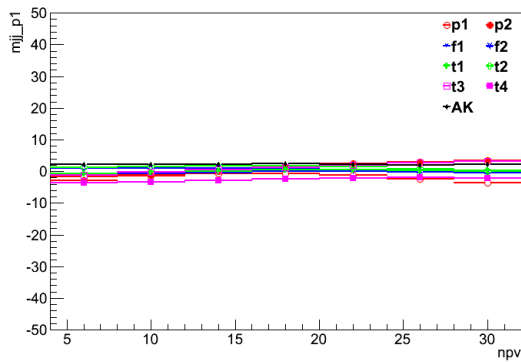**Figure D.9:**
algo-ptdiff-
app/R05˙mjj˙npv˙mean.png

Figure D.10:
algo-ptdiff-
app/R05˙mjj˙npv˙width.png

Figure D.11:
algo-ptdiff-
app/R05˙mjj˙ptjj˙mean.png

Figure D.12:
algo-ptdiff-
app/R05˙mjj˙ptjj˙width.png

Figure D.13:
algo-ptdiff-
app/R06˙mjj˙npv˙mean.png

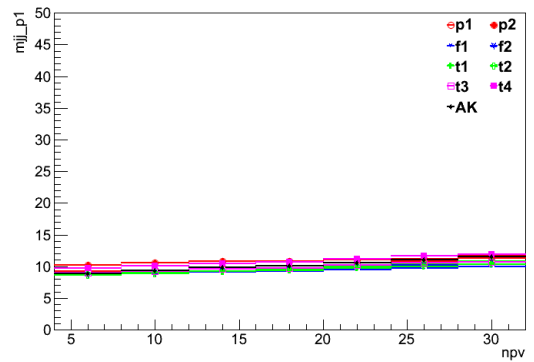**Figure D.14:**
algo-ptdiff-
app/R06˙mjj˙npv˙width.png

Figure D.15:
algo-ptdiff-
app/R06˙mjj˙ptjj˙mean.png

Figure D.16:
algo-ptdiff-
app/R06˙mjj˙ptjj˙width.png

Figure D.17:
algo-ptdiff-
app/R07˙mjj˙npv˙mean.png

Figure D.18:
algo-ptdiff-
app/R07˙mjj˙npv˙width.png

Figure D.19:
algo-ptdiff-
app/R07˙mjj˙ptjj˙mean.png

Figure D.20:
algo-ptdiff-
app/R07˙mjj˙ptjj˙width.png

74

Figure D.21:
algo-ptdiff-
app/R08˙mjj˙npv˙mean.png

Figure D.22:
algo-ptdiff-
app/R08˙mjj˙npv˙width.png

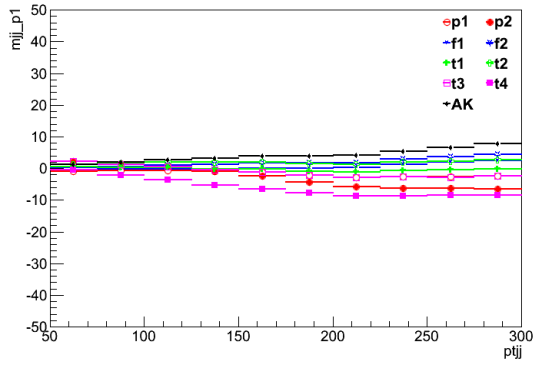**Figure D.23:**
algo-ptdiff-
app/R08˙mjj˙ptjj˙mean.png

Figure D.24:
algo-ptdiff-
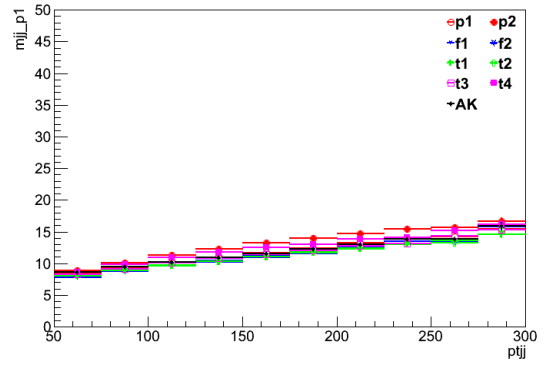app/R08˙mjj˙ptjj˙width.png

Figure D.25:
algo-ptdiff-
app/R09˙mjj˙npv˙mean.png

Figure D.26:
algo-ptdiff-
app/R09˙mjj˙npv˙width.png

75

Figure D.27:
algo-ptdiff-
app/R09˙mjj˙ptjj˙mean.png

Figure D.28:
algo-ptdiff-
app/R09˙mjj˙ptjj˙width.png
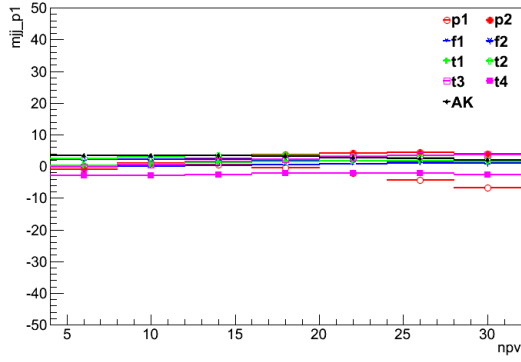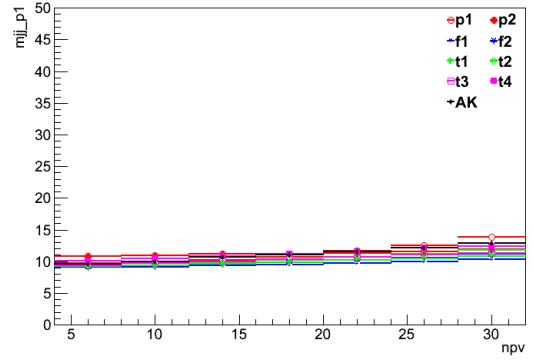

Figure D.29:
algo-ptdiff-
app/R10˙mjj˙npv˙mean.png

Figure D.30:
algo-ptdiff-
app/R10˙mjj˙npv˙width.png

Figure D.31:
algo-ptdiff-
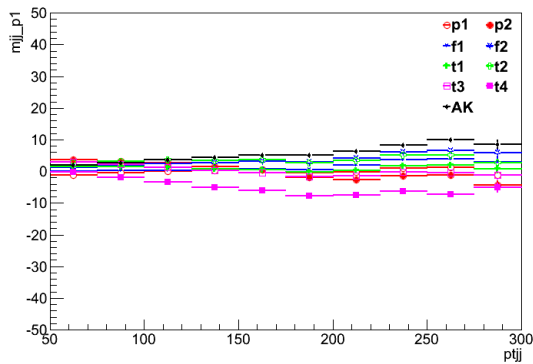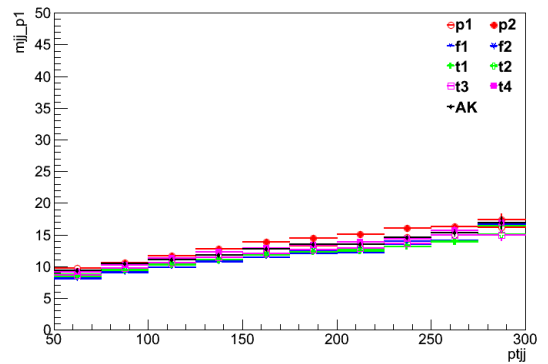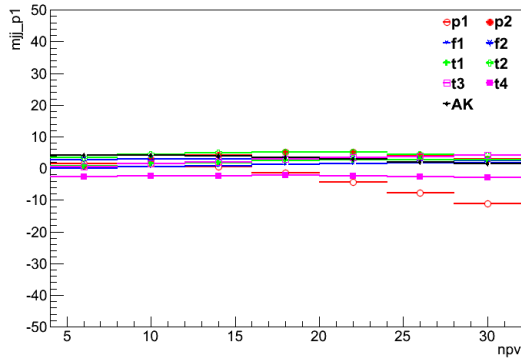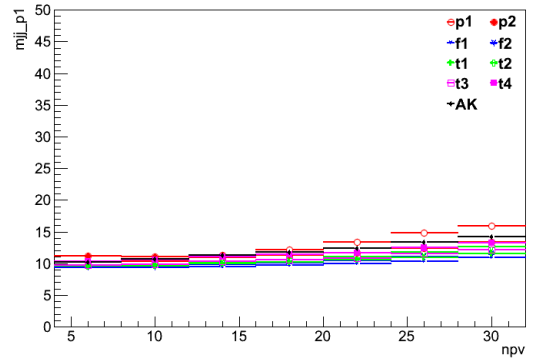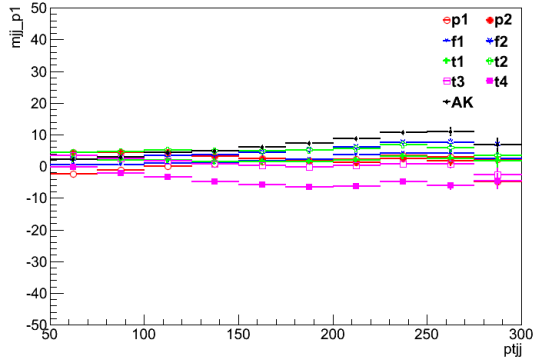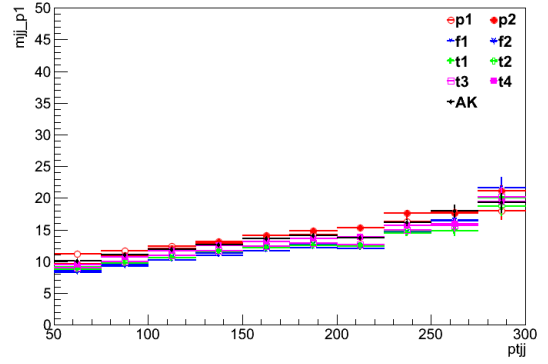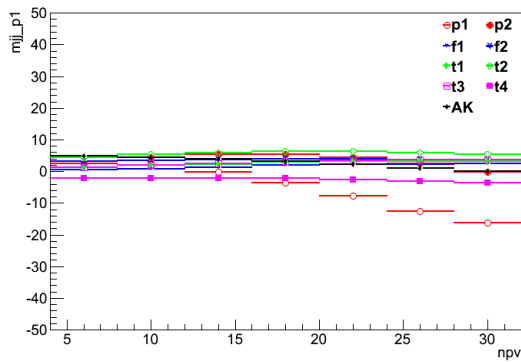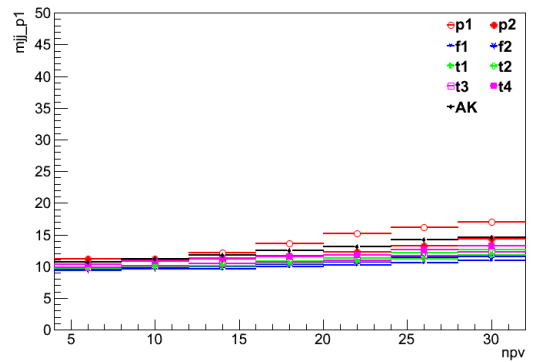app/R10˙mjj˙ptjj˙mean.png

Figure D.32:
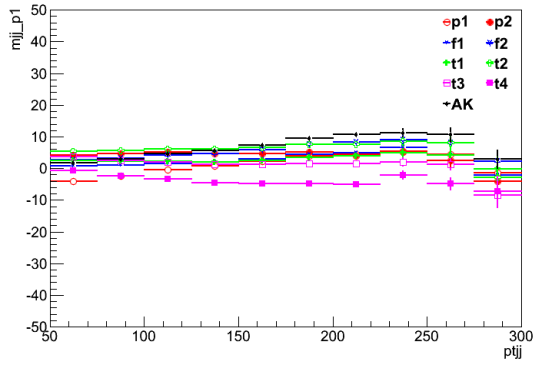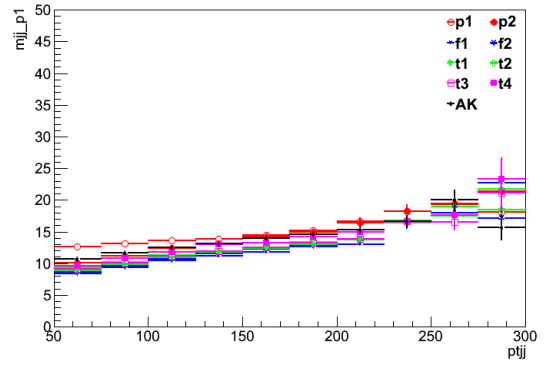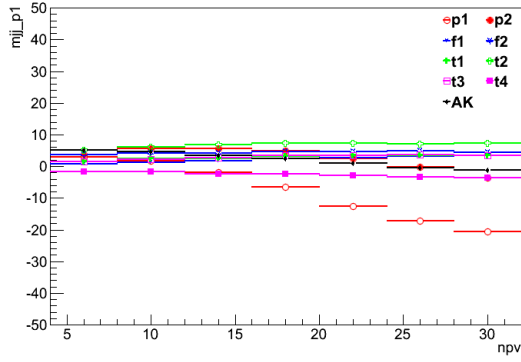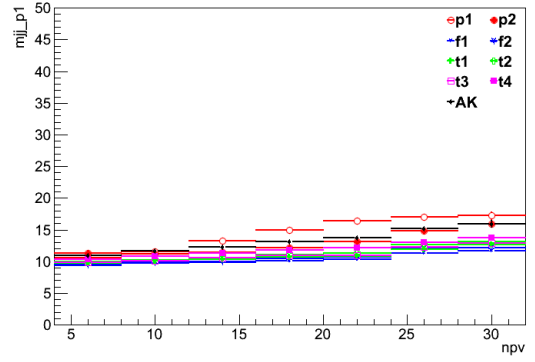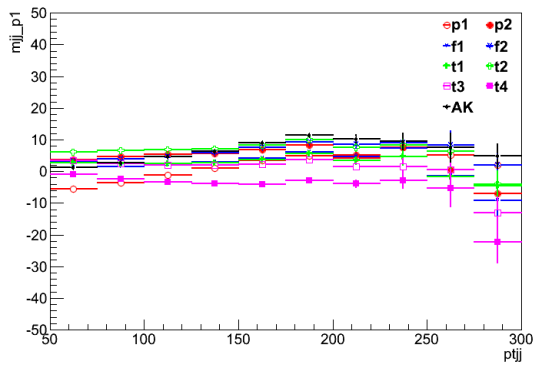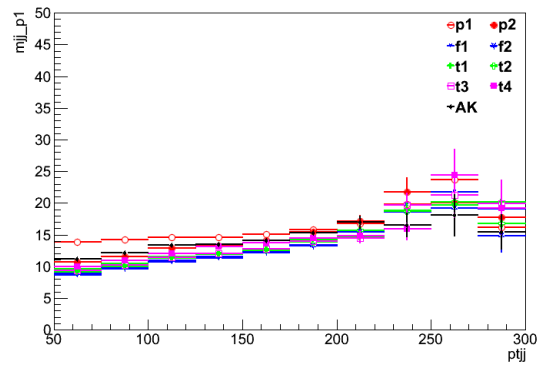algo-ptdiff-
app/R10˙mjj˙ptjj˙width.png

BIBLIOGRAPHY

[1] Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and MET. Tech. Rep. CMS-PAS-PFT-09-001, CERN, 2009. Geneva, Apr 2009.

[2] Combination of standard model Higgs boson searches and measurements of the properties of the new boson with a mass near 125 GeV. Tech. Rep. CMS-PAS-HIG-13-005, CERN, Geneva, 2013.

[3] Identifying Hadronically Decaying Vector Bosons Merged into a Single Jet. Tech. Rep. CMS-PAS-JME-13-006, CERN, Geneva, 2013.

[4] Search for the Higgs boson decaying to invisible particles produced in association with Z bosons decaying to bottom quarks. Tech. Rep. CMS-PAS-HIG-13-028, CERN, Geneva, 2013.

[5] AGUILAR, M. E. First Result from the Alpha Magnetic Spectrometer on the International Space Station: Precision Measurement of the Positron Fraction in Primary Cosmic Rays of 0.5350 GeV. *Phys. Rev. Lett. 110* (Apr 2013), 141102.

[6] BAYATIAN, G. L. *et al. CMS Physics: Technical Design Report Volume 1: Detector Performance and Software.* Technical Design Report CMS. CERN, Geneva, 2006.

[7] BERINGER, J. E. Review of Particle Physics. *Phys. Rev. D 86* (Jul 2012), 010001.

[8] BRUNING, O., AND COLLIER, P. Building a behemoth. *Nature 448*, 7151 (July 2007), 285–289.

[9] CACCIARI, M., SALAM, G. P., AND SOYEZ, G. The Anti-k(t) jet clustering algorithm. *JHEP 0804* (2008), 063, 0802.1189.

[10] CHATRCHYAN, S. *et al.* The CMS experiment at the CERN LHC. The Compact Muon Solenoid experiment. *J. Instrum. 3* (2008), S08004. 361 p.

[11] COLLABORATION, A. Search for dark matter in events with a hadronically decaying W or Z boson and missing transverse momentum in pp collisions at sqrt(s)=8 TeV with the ATLAS detector. arXiv/1309.4017.

[12] DEV, P. S. B., GHOSH, D. K., OKADA, N., AND SAHA, I. Neutrino Mass and Dark Matter in light of recent AMS-02 results. arXiv/1307.6204.

[13] ELLIS, S. D., VERMILION, C. K., AND WALSH, J. R. Recombination Algorithms and Jet Substructure: Pruning as a Tool for Heavy Particle Searches. arXiv/0912.0033.

[14] EVANS, L., AND BRYANT, P. LHC Machine. *Journal of Instrumentation 3*, 08 (2008), S08001.

[15] GRELJO, A., JULIO, J., KAMENIK, J. F., SMITH, C., AND ZUPAN, J. Constraining Higgs mediated dark matter interactions. arXiv/1309.3561.

[16] GRIFFITHS, D. *Introduction to Elementary Particles.* John Wiley & Sons, New York, USA, 1987.

[17] KNOLL, G. *Radiation Detection and Measurement.* John Wiley & Sons, 2010.

[18] KROHN, D., THALER, J., AND WANG, L.-T. Jet Trimming. arXiv/0912.1342.

[19] THALER, J., AND TILBURG, K. V. Identifying Boosted Objects with N-subjettiness. arXiv/1011.2268.

Academic Vita

Daniel Abercrombie

1528 Daugherty's Run Rd, Linden, PA 17744
(570) 777-0685, abercrombie@psu.edu

EDUCATION

B.S., Physics (with Honors); *B.S.,* Nuclear Engineering
Minor in Mathematics
The Pennsylvania State University, University Park, PA
Expected Graduation: May 2014

EXPERIENCE

European Organization for Nuclear Research   Meyrin, Switzerland
Summer Student   Summer 2013
- Conducted research under the JetMET Algorithms and Reconstruction group of the CMS detector
- Compared resolution and stability of jet grooming algorithms and cone sizes
- Simulated p_T distribution of gluon jets resulting from event pileup

Penn State Dept. of Mech. and Nuclear Engineering   University Park, PA
Grader, NucE 310W: Issues in Nuclear Engineering   Fall 2013
- Evaluated students' understanding of challenges in nuclear energy industry
- Provided writing instruction

Research Assistant, Intense Laser Lab   Spring 2012 - Summer 2013
- Designed and constructed a laser compressor
- Modeled radiation transport using Geant4 and MCNP
- Optimized high energy X-ray generation for active interrogation
- Built interface to control lab equipment using LabVIEW

Research Assistant, Advanced Multiphase Flow Lab   Fall 2011
- Learned theory and construction techniques of conductivity probes
- Acquired data used to validate three dimensional CFD code

PSUKnowHow   State College, PA
Private Tutor   Fall 2011
- Taught physics and engineering concepts to engineering students
- Presented topics in multiple ways to help student comprehension

COMPUTER SKILLS

C++	ROOT	MATLAB	SolidWorks
Fortran	Geant4	LabVIEW	HTML
Python	MCNP	Mathematica	Bash

ACTIVITIES

American Nuclear Society: Penn State Student Chapter   2011 - Present
Technical Lead for ANS 2014 Student Conference at Penn State
Vice President, Fall 2012 - Spring 2013
Outreach Volunteer, Spring 2011 - Present

HONORS & AWARDS

2014 Student Marshal for the Eberly College of Science
2014 Student Marshal for the Department of Nuclear Engineering
2014 John J. Brennan Excellence in Nuclear Engineering Award
2013 Astronaut Scholarship Foundation Award
2013 American Nuclear Society National Student Conference
    Best Undergraduate Presentation in Reactor Physics
2013 College of Engineering Research Symposium Poster Award
2013 Evan Pugh Scholar Award
2012 DOE's Nuclear Energy University Programs Scholarship
2012 American Nuclear Society's Accelerator Applications Division Scholarship