Honors Senior Thesis

March 17, 2014

Measurement of the Higgs Spin and Parity in the Four-Lepton Channel with the ATLAS Detector

Kareem Hegazy

Abstract

The Spin and Parity (J^P) of the newly discovered Higgs boson is studied in the $H \rightarrow ZZ^* \rightarrow 4l$ decay channel with the ATLAS experiment at the Large Hadron Collider. This study uses the entire proton-proton collision data set: corresponding to an integrated luminosity of 4.6 fb^{-1} at 7 TeV and 20.3 fb^{-1} at 8 TeV. Comparisons with data of the predicted Standard Model, $J^P = 0^+$, hypothesis and alternative hypotheses ($J^P = 0^-$, 1^+, 1^-, and 2^+) are made through a multivariable analysis based on lepton kinematic, Z boson invariant masses, and angular observables. The results show that data is in favor of the Standard Model 0^+ hypothesis, while other alternative J^P states are excluded at a 83.1% to 99.8% confidence level, independent of assumptions on the coupling strengths between the Standard Model Higgs boson to other particle.

Contents

1 Introduction

On July 4th 2012 CERN (European Organization for Nuclear Research) announced the discovery of a new boson with a mass around 125 GeV , which was observed by both the ATLAS and CMS experiments in their search for the Standard Model (SM) Higgs boson at the Large Hadron Collider (LHC). Later that year CERN published the discovery of a particle consistent with the Higgs boson [1]. The discovery of this new particle was experimentally observed in three different final states: $\gamma\gamma$, ZZ^*, and WW. Both experiments have since collected more data and studied the properties of this newly discovered particle. The measured properties (couplings, spin, and parity) of the new particle by both experiments are consistent with the predictions of the SM Higgs boson. Confirming this particle as the SM Higgs boson will help to unlock the mystery of the electroweak symmetry breaking mechanism of the SM.

The SM describes the building blocks of the universe through elementary particles and their interactions. This theory is based on two principles: (1) gauge invariance, from which all the interactions between particles are naturally introduced in a theoretical framework; (2) Higgs Mechanism, from which the electroweak symmetry is broken and all the massive elementary particles acquire their mass through this symmetry breaking mechanism. The aspects of gauge invariance (interactions) in the SM has been successfully tested by many experiments over the past half century. However, the electroweak symmetry breaking mechanism has remained a mystery for almost 50 years. The discovery the Higgs field's quanta, the Higgs boson, at the LHC in 2012 provides evidence of and a means of further studying this SM symmetry breaking process. It is therefore crucial to measure the newly discovered Higgs boson's properties to advance our understanding of electroweak symmetry breaking mechanism and search for new physics beyond the SM at the LHC. This thesis reports the current measurements of the spin (J) and parity (P) quantum numbers of the Higgs boson in the $H \to ZZ^* \to 4l$ channel from reconstructed data events collected by the ATLAS experiment. This analysis uses the full ATLAS proton-proton collision data set, with an integrated luminosity of 4.6 fb^{-1} at a center-of-mass energy of $\sqrt{s} = 7$ TeV, and 20.3 fb^{-1} at $\sqrt{s} = 8$ TeV. In this thesis the predicted Standard Model, $J^P = 0^+$, hypothesis and alternative hypotheses corresponding to other J^P states ($jcp = 0^-$, 1^+, 1^-, and 2^+) are compared with a multivariable analysis dependent on discriminating variables constructed from lepton kinematics, Z boson invariant masses, and angular decay observables.

This thesis is organized as follows: Section 2 describes the LHC and ATLAS detector; Section 3 describes the event reconstruction and modeling of the ATLAS experiment; Section 4 briefly reports the $H \to ZZ^* \to 4l$ event selection and background estimation; Section **??** details the Spin and Parity measurement with four-lepton final state events and presents the results; the ongoing work and conclusion can be found in Section 6 and 7, respectively.

2 The Large Hadron Collider and ATLAS

2.1 The Large Hadron Collider

The LHC is the largest and highest energy particle collider in the world. The LHC is located on the France-Swiss boarder at CERN, and was built from 1998 to 2008. In 2010, the LHC physics run started with proton-proton collisions at a center of mass energy of $\sqrt{s} = 7$ TeV. In 2012, the energy was increased to $\sqrt{s} = 8$ TeV. The peak luminosity of the LHC reached $7 \times 10^{33} cm^{-2} s^{-1}$ in 2012, with a total integrated luminosity of 26 fb^{-1} delivered to each experiment at the LHC during its 2010 to 2012 operation. The designed LHC energy, $\sqrt{s} = 14$ TeV, is expected to be reached in 2015 after the current upgrade efforts are finished.

The LHC is a largest accelerator complex at CERN, which accelerates particles to various energies before finally being injected into the LHC ring where collisions take place. The proton beams consist of many proton bunches that are taken from a tank of hydrogen atoms stripped of their electrons by an electric field. The protons first enter Linac 2, which accelerates the protons to an energy of 50 MeV. The bunch is then injected into the Proton Synchrotron Booster, which accelerates the protons to 1.4 GeV. Next, the bunch enters the Proton Synchrotron and the Super Proton Synchrotron where the particles are accelerated to 25 GeV and 450 GeV, respectively. Finally, the bunch is injected into the Large Hadron Collider where each bunch is currently accelerated to 4 TeV and collide with a total energy of 8 TeV. A schematic of the accelerator complex can be seen in Figure 1. The two beams, comprised of thousands of bunches, travel in opposite directions and circulate the LHC for hours as they are contained and collimated using superconducting magnets to keep the bunches sufficiently dense. The beams collide inside four detectors: ALICE, ATLAS, CMS, and LHCb. Each bunch consists of about 10^{11} protons with bunch collisions occurring every 50 nanoseconds.

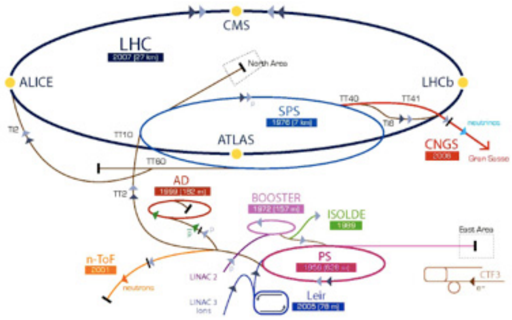

Figure 1: The CERN accelerator complex for the LHC.

2.2 The ATLAS Detector

The ATLAS detector (A Toroidal LHC Apparatus) is a multipurpose particle detector in the LHC complex with forward-backward symmetric cylindrical geometry as shown Figure 2. ATLAS uses a right-handed coordinate system with its origin at the interaction point (IP), where the beams collide in the center of the detector. The z-axis is along the beam line, while the x-axis points from the IP to the center of the LHC, and the y-axis extends upwards. Cylindrical coordinates (r, ϕ) are used in the transverse plane, where ϕ is the azimuthal angle around the beam line. The pseudorapidity (η) describes the final dimension analogous to the polar angle, and is defined in terms of the polar angle θ as $\eta = ln[tan(\theta/2)]$. Observables labeled ”transverse” are projected into the x-y plane.

The ATLAS detector consists of a magnet system and four major specialized detectors: the inner detector, electromagnetic calorimeter, hadronic calorimeter, and the muon spectrometer. The detector components important to this study include the inner tracking detector (ID), the liquid argon (LAr) electromagnetic calorimeter, and the muon spectrometer (MS). The ID is located in the center of ATLAS and consists of silicon pixel and microstrip trackers, as well as straw-tube transition-radiation trackers which in total cover the region of $|\eta| < 2.5$. The ID tracks charged particles using discrete measurements to produce precision measurements of the particle's trajectory, as well as its momentum. The ID is surrounded by a superconducting solenoid that produces a 2T magnetic field which bends the particle's trajectory. The LAr calorimeter surrounds the solenoidal magnetic field, and is divided into a central barrel covering $|\eta| < 1.475$ and two endcaps covering $|\eta| < 3.2$. The LAr calorimeter provides precision location and energy measurements of particles that interact electromagnetically (charged particles and photons) by absorbing the particles. The EM calorimeter is made of sheets of lead separated by layers of liquid argon. When a charged particle or a photon hits the lead it produces many more particles, creating

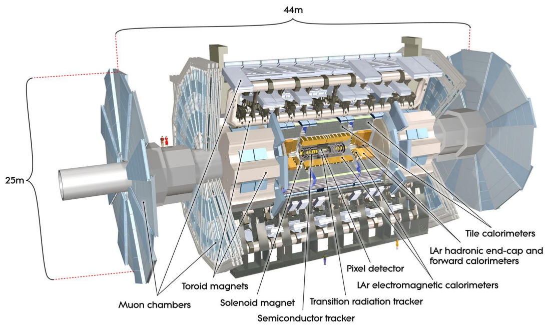

Figure 2: The ATLAS detector.

a shower of mostly electrons and some protons, that then ionize electrons in the liquid argon creating an avalanche effect. The shower of electrons strikes the next lead sheet and produces more particles. This process continues until the avalanched particles do not have enough energy to ionize electrons. The ionized electrons in the liquid argon are collected and used to measure the original particle's total energy. Additionally, the LAr calorimeter measures the position of the particle. When combining the momentum and tracking measurements from the inner detector with the measurements from the LAr calorimeter, and taking the magnetic field into account, one can measure the charge of the particle by the direction its track bended in the magnetic field. Additionally, one can use the curvature of the reconstructed track in the magnetic field to measure the particle's momentum. The MS surrounds the entire detector with a barrel region ($|\eta| < 1.7$), transition region ($1.7 < |\eta| < 1.9$), and endcaps which in total cover the region of $|\eta| < 2.7$. The MS contains a system of precision tracking chambers embedded in a toroidal magnetic field produced by three large air-core superconducting magnets. The MS has three measuring stations to make additional precision measurements of muons' tracks and momentum, since muons are the only particle that can penetrate through to the MS. The designed precision of the MS is to measure the momentum of a 100 GeV muon within 3% and a 1 TeV muon to within 10% accuracy.

In addition to the detector subsystems described above, the ATLAS experiment also contains a trigger system with three levels. Triggers are vital to the experiment and allow the experiment to decide which events should be recorded for physics analysis. The trigger system makes these decisions based upon certain event selection criteria. This procedure is critical for the experiment at the LHC since the 40 MHz collision rate in the LHC proton-proton collisions would produce terabytes of data per minute if each event were recorded. The first level of triggers (LV1) is hardware based and is completely located in the calorimeter and MS. LV1 triggers have course granularity and primarily determine regions of interest based upon particle flight path in the MS or the calorimeter. LV1 triggers have low granularity because they must make a decision before the next bunch crossing to avoid a large pileup of data. If an event passes the LV1 triggers it is sent to the higher level trigger system which includes the level two trigger (LV2) and event filter (EF or LV3). LV2 is a software based trigger which partially reconstructs the event by looking at the regions of interest indicated by the LV1 trigger in the MS and calorimeter. Those events that pass the LV2 trigger are sent to the EF where the event is fully reconstructed and the final trigger decisions are made After passing the EF, events are recorded at a rate of 300 MB/s. Because of the trigger system, the event rate falls from 40 MHz in the LHC to 75 KHz after LV1, then 3 KHz after LV2, and finally 300 Hz after EF for event collection.

Due to the 50 ns bunch-crossing there is a large "pileup" of collisions, since each bunch-crossing produces around 20, and up to 40, proton-proton collisions. This creates the challenging problem of correctly recreating events and correctly tracing particles back to the correct primary vertex of the interaction point. A graph of the number of proton-proton collisions per bunch crossing, as well as a reconstructed bunch crossing with 25 collisions can be seen in Figure 3 for data collected in 2011 and 2012. This rate of collisions allows the LHC to deliver data at a record breaking rate. The amount of data delivered by the LHC, as well as the amount that as suitable for physics analyses can be seen in Figure 4 for both 2011 and 2012.

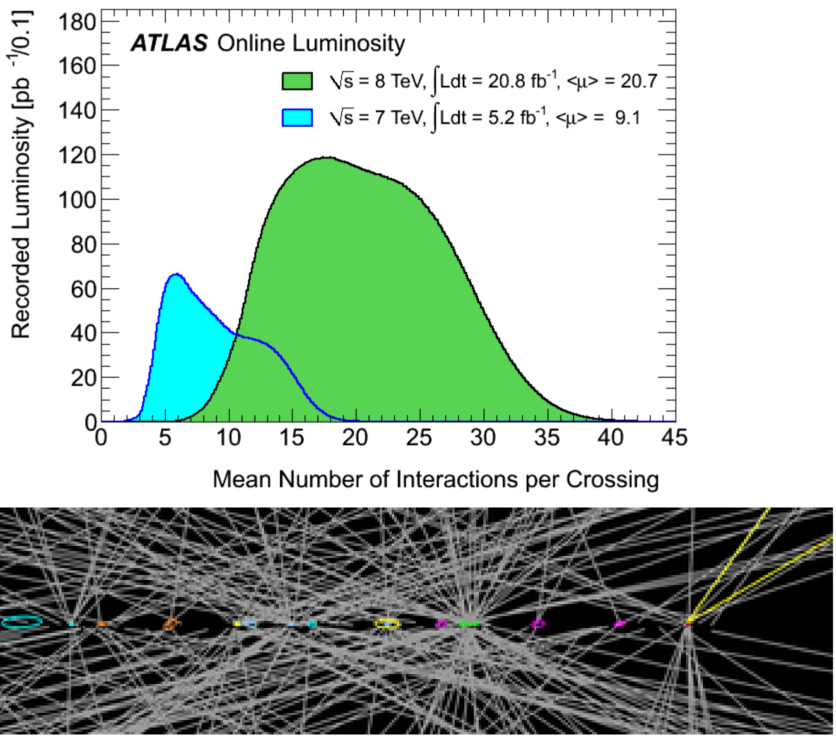

Figure 3: The mean number of proton-proton collisions per bunch crossing for 2011 and 2012 collected in ATLAS (top) [2] and a reconstruction bunch crossing with 20 interaction vertices (bottom).

Since protons are predominantly made of gluons with three small quarks (uud) ”floating” in a sea of gluons, the vast majority of the proton-proton collisions result in the collision of two gluons. It is extremely rare to see a quark-quark collision. For this reason our predominant model is based upon gluon-gluon fusion. Additionally, since the gluons that collide do not necessarily have the same velocity along the beam pipe (z-axis), the center of mass of the collision will have velocity along the z-axis. For this reason our measurements are done in the transverse plane perpendicular to the beam pipe to avoid any center of mass velocity.

2.3 Grid Computing for Data Processing

The ATLAS data processing (data quality checks, prompt event reconstruction and calibrations, as well as Monte Carlo event productions) for physics analyses are carried out with a grid computing system [13]. Data analysis for individual physics topics are done at individual institutes with local computing system. The Worldwide LHC Computing Grid (WLCG) is the world’s largest computing grid and serves more than 5000 physicists around the world. The WLCG contains more than 150 computing centers (from Tier-0 to Tier-4) around the world in 40 different countries to store and analyze multi-terabytes of data. Tier computing systems are ordered depending upon their responsibilities and can be accessed from

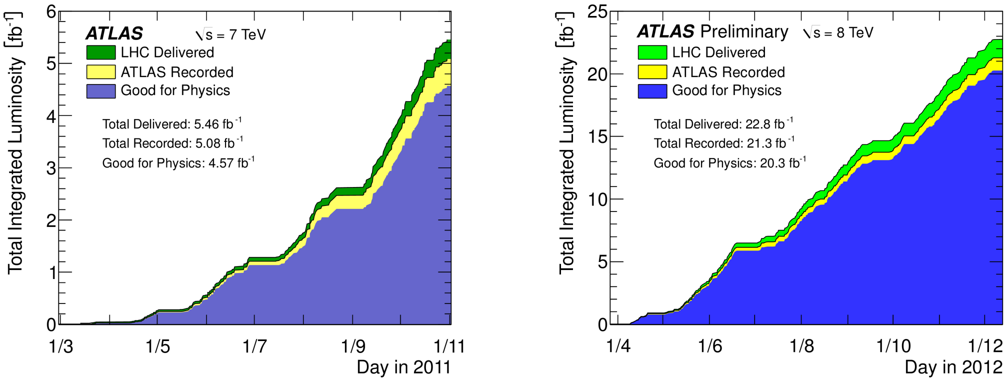

Figure 4: Total integrated luminosity from the LHC and ATLAS for 2011 at 7 TeV (left) and for 2012 at 8 TeV (right) [2].

around the world. CERN is Tier-0, where the data is produced, promptly reconstructed and distributed. Tier-1 centers are national centers that are responsible for holding the most updated raw data from CERN, in addition to doing data analysis. Tier-2 facilities are regional computer clusters used primarily for Monte Carlo event productions, detector calibrations, and data analyses. To run computing jobs on the the grid facility: physicists can upload the analysis code to the grid and specify which data files they wish to use. The grid sends the jobs to various tiers where the data is available and then splits the job among various computer cores at that tier according to how the splitting was specified by the maker of the code. The results of finished jobs are then sent back to the physicist. Most the computational jobs in the final stages of a physics analysis are done at local Tier-3 computing system. Michigan is responsible for a Tier-2 and Tier-3 computing system, where the Tier-3 is dedicated for Michigan physicists and the Teir-2 system stores the majority of the LHC dataset. The analysis presented in this thesis was carried out using Michigan Tier-3 computing cluster.

3 Event Reconstruction and Monte Carlo Simulations

Physics objects from the event reconstructions of raw data collected by ATLAS are the basis of data analysis, as well as physics modeling and detector simulations with Monte Carlo (MC) techniques used to simulate data. The important physics objects for this analysis are electrons and muons.

3.1 Lepton Reconstruction and Identification

Electrons are reconstructed from energy deposits in the electromagnetic calorimeter and matched to a track in the inner detector. For the 2012 data at 8 TeV, improved electron discrimination from jets is obtained using a likelihood function from parameters characterizing the electromagnetic shower shape and track association, which significantly increased electron identification efficiency and decreased mis-identification rate compared to 7 TeV data analyses. Electron candidates are required to have $p_{\mathrm{T}} > 7$ GeV and $|\eta| < 2.47$. The electron reconstruction and identification efficiency is greater than 75%.

Muons are identified by tracks (or track segments) reconstructed in the MS, which are matched to tracks reconstructed in the ID. The muon momentum is calculated by combining the information from the two subsystems and correcting for the energy lost in the calorimeter. Additionally, one muon in each event is allowed to be a stand-alone muon or a calorimeter-tagged muon, where stand-alone muon is identified by only having a muon spectrometer track in $2.5 < |\eta| < 2.7$, and the calorimeter-tagged muon is identified by an inner detector track with $p_{\mathrm{T}} > 15$ GeV associated with a compatible calorimeter energy deposit in $|\eta| < 0.1$. All muon candidates are required to have $p_{\mathrm{T}} > 4$ GeV and $|\eta| < 2.7$. The muon reconstruction and identification efficiency is greater than 94%.

3.2 Physics Modeling

The Higgs production and decays at the LHC are modeled by the ***Powheg*** MC generator, which is used to calculate the signal cross sections and includes perturbative QCD corrections to next-to-leading order (NLO) [6]. The CT10 set of parton distribution functions (pdfs), QCD renormalization, factorization scales, μ_R, and $\mu_F = M_{4l}$ are used in the calculations. To generate MC events with detector simulations to determine the signal acceptance, ***Powheg*** is interfaced with the ***Pythia*** MC Program and ***Photos***. ***Pythia*** is used to simulate parton showering and hadronization [12], while ***Photos*** is responsible for simulating radiated photons from charged leptons [9].

The ***JHU*** generator is used to generate Higgs events with alternative spin and parity states. These calculations are done with leading order corrections for perturbative QCD corrections [8]. Each event from the ***JHU*** generator must be reweighted to make sure the leading order Higgs p_{T} distribution is consistent with the Higgs p_{T} distribution generated by the NLO ***Powheg*** generated events, which generates the SM Higgs events. The ***JHU*** MC calculates the correlations of the leptons decay from the Higgs boson that are sensitive to the Higgs Spin and Parity quantum numbers.

Background events are modeled and produced using the following MC Generators: ***MC@NLO*** (for top events) [7], ***Alpgen*** (for $Z/W + jets$ events) [10], and ***Powheg*** for diboson (ZZ, Zg, and ZW) events.

MC generators produce events from proton-proton collisions based on theoretically predicted production cross sections and kinematic distributions. These events are input for the detector simulation and reconstruction programs which produce MC datasets used to simulate the data collected from the proton-proton collisions at the LHC and reconstructed by ATLAS experiment.

3.3 Detector Simulations

The detector response simulation is based on the GEANT4 program [5]. Additional inelastic proton-proton collisions in each bunch crossing (referred to as pile-up) are included in this simulation. Additionally, the MC events are reweighted to reproduce the observed pile-up distribution, the average number of collisions per bunch-crossing in the data. The distribution of the number of proton-proton collisions per bunch crossing for both 7 TeV and 8 TeV data can be seen in Figure 4 .

Furthermore, GEANT4 contains detailed descriptions of the detectors' materials and how they react to particles passing through them, such as gas ionization from a particle passing through a gaseous muon detector. With this information, a particle's interactions with the detector can be simulated. However, the results of all these procedures follow distributions governed by the laws of physics for particles interacting with matter. Therefore, Monte Carlo programs are used to generate random numbers from these distributions that are based on underlying physics to simulate the detector response.

MC simulated events for physics analysis must be calibrated and corrected to precisely simulate real detector responses and reproduce data distributions. The corrections and calibrations are done with control data samples that are very well understood and produce a very clean and strong signal, such as $Z \to \mu\mu$ and $Z \to ee$ events. In this analysis, the MC lepton identification and trigger efficiencies are corrected based on studies performed in data control regions. The energy and momentum scales, as well as their resolutions, of the MC events are calibrated to reproduce data from $Z \to ll$ and $J/\Psi \to ll$ decays. The uncertainties of the $H \to ZZ^* \to 4l$ signal detection efficiencies are determined by varying the nominal calibrations: lepton energy resolution, momentum resolutions, the lepton trigger, as well as reconstruction and identification efficiencies. The overall uncertainties in the Higgs signal efficiencies range from 2.7% to 9.8%, depending on final state lepton flavors, where the major uncertainty contributions are from the lepton reconstruction and identification efficiencies. These uncertainties are considered in the final J^P analysis and will be described in Section 5.4.

4 Higgs Event Selection

The $H \to ZZ^* \to 4l$ channel is characterized by a four-lepton final state with two pairs of same flavor opposite charge leptons. This channel is referred to as the "golden channel" for the Higgs boson discovery since the four leptons can be measured very well and the final state signature is easy to detect. Additionally, this channel has very low background contamination due to its clean final state. Figure 5 shows reconstructed Higgs candidates in the $4e$ and 4μ channel from data in ATLAS, where the tracks in the muon spectrometer and the showers in the LAr calorimeter are highlighted in the $4e$ and 4μ channel decays, respectively.

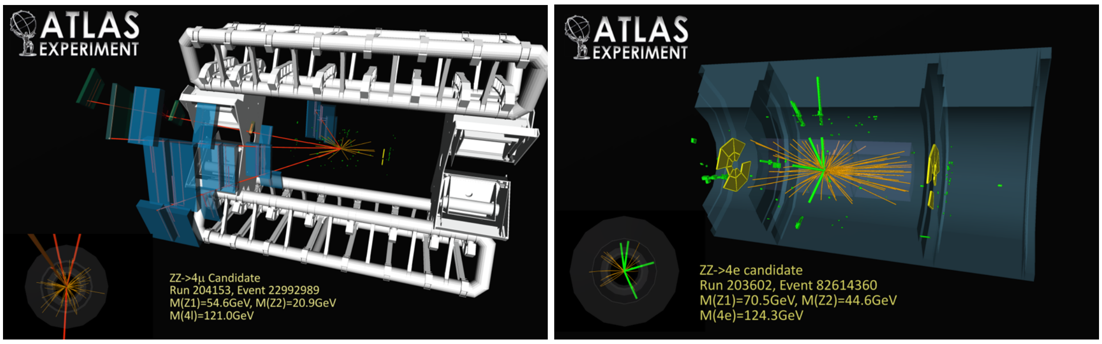

Figure 5: Higgs candidate in the 4μ channel (left) and $4e$ channel (right) in the ATLAS detector.

This decay channel contains one real (on-shell) Z boson and one virtual $Z*$ boson with a mass of 30 GeV. The cross section of the Higgs boson production is dominated by the gluon-gluon fusion process (19.5 fb at 8 TeV). Additionally, the Higgs can be produced through vector-boson fusion (1.57 fb at 8 TeV) and the association production mechanism with W and Z vector bosons (1.08 fb at 8 TeV). The association production with top pair accounts for less than 1% of the Higgs production cross section (0.13 fb). The decay branching ratio for $H \to ZZ^* \to 4l$ is only 0.012%. For these reasons, we expect the number of four-lepton events found at the Higgs resonance to be statistically limited, which requires great effort to retain high efficiency in the event selection.

4.1 $H \to ZZ^* \to 4l$ Event Selection

Events are selected from proton-proton collisions by requiring at least one reconstructed vertex with at least three charged particle tracks with $p_{\mathrm{T}} > 0.4$ GeV. If more than one vertex satisfies the selection requirement, the primary vertex is chosen as the one with the highest Σp_{T}^2, where the sum runs over all tracks associated with this vertex.

The $H \to ZZ^* \to 4l$ events must contain at least four identified leptons, selected with criteria described in Section 3.1. In order to reject electrons and muons faked by jets, only isolated leptons are selected, requiring the scalar sum of the transverse momentum, Σp_{T}, of other tracks inside a cone size of $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} = 0.2$ around the lepton to be less than 15% of the lepton p_{T}. In addition, the ΣE_{T} deposited in calorimeter cells inside a cone of $\Delta R = 0.2$ around the lepton track, excluding the transverse energy due to the lepton and corrected for the expected pile-up contributions, is required to be less than 30% of the lepton $p_{\mathrm{T}}$for 7 TeV data, less than 2% for electrons from 8 TeV data, and less than 15%

for stand-alone muons. To further reject leptons from heavy-flavor jets, the impact parameter relative to the primary vertex is required to be less than 3.5 (6.0) standard deviations for all muons (electrons). This looser electron requirement allows for the tails in the electron impact parameter distribution due to bremsstralung radiation in the inner detector.

The Higgs candidate quadruplets are formed by selecting two opposite-sign, same-flavor dilepton pairs in an event. The four leptons of the quadruplets are required to be well separated, $\Delta R > 0.1$ for same-flavor lepton pairs and $\Delta R > 0.2$ for $e\mu$ pairs. The two leading leptons, determined as the two leptons with the highest p_{T}, must have $p_{\mathrm{T}} > 20$ GeV and $p_{\mathrm{T}} > 15$ GeV. The third lepton must have $p_{\mathrm{T}} > 10$ (8) GeV if it is an electron (muon). The lepton pair with greatest invariant mass (denoted M_{12}) and closest to the Z pole mass is called the leading lepton pair, while the sub-leading lepton pair is chosen to have the largest invariant mass (denoted M_{34}) among the remaining possible pairs. The dilepton masses must satisfy $50 < M_{12} < 115$ GeV, and $M_{34} > 12$ GeV. In the $4e$ and 4μ channels all same-flavor, opposite-sign lepton pairs are required to have invariant masses greater than 5 GeV. This helps reject event contamination from $J/\Psi \to ll$ decays. The final cut on the quadruplet requires M_{4l} to be within the Higgs signal region of 115 GeV to 130 GeV. A total of 31 Higgs candidate events are selected from the 2011 and 2012 datasets: 13, 8, 9, and 7 events from the 4μ, $2\mu 2e$, $2e2\mu$, and $4e$ final state, respectively, where $2\mu 2e$ ($2e2\mu$) indicates the leading lepton pair is comprised of 2μ ($2e$).

From MC simulation studies, we expect to select total 15.1 Higgs signal events and 8.7 irreducible background events from the SM $qq \to ZZ^* \to 4l$ process. The reducible background from $Z + Jet$ and top events are estimated using a data-driven method to contribute 2.24 events, which will be described in the next section. Figure 6 shows the inclusive $4l$ invariant mass distribution without applying the final Higgs mass cut, from which we see that the Higgs resonance mass is around 125 GeV. This peak is clearly observed over the background. The Higgs signal significance is computed to be 6.6σ, exceeding the 5σ criteria to claim a discovery of a new particle in high energy physics.

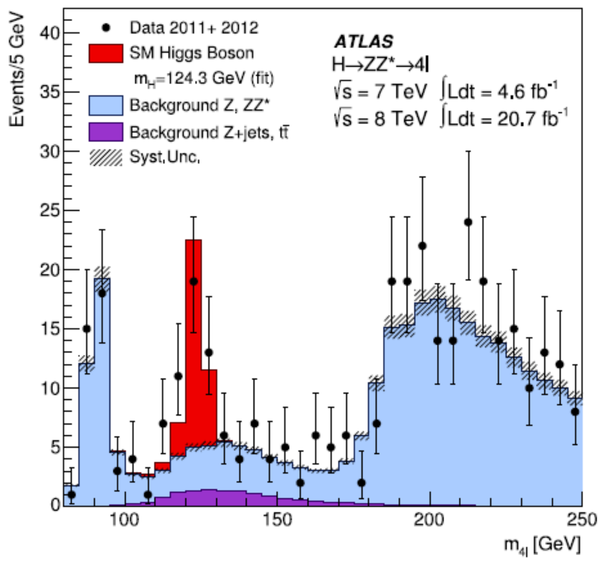

Figure 6: M_{4l} distribution with combined 7 TeV and 8 TeV data.

4.2 Background Estimation

$Z+Jet$ background consists of events that contain an on-shell Z boson, as well as a jet. This process can be confused as a $H \rightarrow ZZ^* \rightarrow 4l$ decay if the jet produces a pair of same flavor oppositely charged leptons. The leptons produced in jets will often have lower energies, which is similar to the low energy leptons produced from the decay of the virtual Z boson at 30 GeV. Therefore, since both decay processes involve an on-shell Z boson, if the jet produces a pair of same flavor oppositely charged leptons with a combined mass of 30 GeV, then a $Z+Jet$ event may be mistaken as a $H \rightarrow ZZ^* \rightarrow 4l$ event. There are however more cuts to avoid this mis-identification, such as the impact parameter cuts described above.

The background contamination from the $t\bar{t}$ production comes from its decay into a pair of opposite sign W bosons, as well as two b quarks. The opposite sign W bosons can in turn decay into leptons of the same flavor, producing a same flavor oppositely charged pair of leptons that could be mistaken for an on-shell Z boson. The two b quarks would produce jets that could produce another pair of leptons to mimic the virtual Z. In this manner a Higgs event could be mis-identified.

Both $Z+Jet$ and top background are estimated from data. As described above these background events may contain two isolated leptons from Z decays or W decays in top events, together with additional activities such as heavy flavor jets or misidentified components of jets yielding reconstructed leptons. These background estimations are done using background-enriched control data samples containing two isolated leptons (ll) and two lepton-like jets ($j_l j_l$) in each event. The control samples are selected with the standard signal requirement except that the lepton-like jets are selected in place of two of the signal leptons. The total reducible background in the signal sample is estimated by scaling each event in the ll and $j_l j_l$ control sample by $f_1 \cdot f_2$, where the factor $f_i(i=1,2)$ for each of the two lepton-like jets depends on lepton flavor and p_T. The factor f is the ratio of the probability for a jet to satisfy the signal lepton selection criteria to the probability for a jet to satisfy the lepton-like jet selection criteria, and is obtained from independent jet-enriched data samples dominated by $Z+Jet$ or top pair events. The total estimated reducible background is 2.24 events in the selected Higgs signal sample.

4.3 Observed and Expected Number of Events at the Higgs Resonance

Table 4.3 gives the number of observed and expected events from the combined 7 TeV and 8 TeV datasets. The expected numbers of events include the Higgs signal, the irreducible ZZ^* background, and the reducible background from $Z+Jet$ and top events.

Channel	Higgs	ZZ*	Reducible Background	Expected	Observed
$\mu\mu\mu\mu$	5.68	3.35	0.75	9.78	13
$\mu\mu ee$	2.94	1.58	0.52	4.04	8
$ee\mu\mu$	3.77	2.31	0.69	6.77	9
$eeee$	2.67	1.42	0.29	4.38	7
Total	15.06	8.65	2.24	25.95	31

These estimates for the expected number of events will be used to calculate the expected and observed significance of a specific J^P state for the spin and parity studies.

5 Spin-Parity Measurement

5.1 Multivariable Analysis

The discovery of the Higgs boson opened a new chapter in the history of particle physics. The measurements of the newly discovered boson's properties, including mass, couplings, spin, and parity quantum numbers, will play a central role in confirming the SM Higgs and aid in the search for unknown physics in the Higgs sector beyond SM. This chapter will provide a detailed description of the spin and parity measurement with the four-lepton final states using events selected by the criteria described in Section 4.1. The SM Higgs boson is predicted to have spin 0 and even parity, $J^P = 0^+$. However, there are other theoretical models that predict the Higgs-like boson could have different spin and parity states, or mixing states of even and odd parities. The method used to test the likelihood or find the exclusion of a specific J^P hypothesis involves a multivariable analysis which utilizes a boosted-decision-tree (BDT). This is necessary because the observed number of events is only 31, including the estimated background contribution of 10.9 events. The BDT output is calculated based on several sensitive variables and used to be the final discriminating variable to separate different spin and parity states. In this study, parity-even and parity-odd resonances of spin 0, 1, and 2 (denoted as $J^P = 0^+, 0^-, 1^+, 1^-$, and 2^+) are considered. In this study, spin and parity hypotheses are tested in pairs. In each individual test, a hypothesis is assumed for the spin and parity of the observed resonance, and the exclusion significance is calculated with respect to other modes. The goal is therefore to find a model for which the observed exclusion with respect to all other hypothesis is comparable to the expected sensitivity given by the observed data. To confirm the new boson is indeed the SM Higgs boson, the $J^P = 0^+$ state is compared with all other J^P states.

The $H \to ZZ^* \to 4l$ reconstructed events, with the decay of Z bosons, provides full information of the Z decay planes. The spin and parity sensitive variables used for this measurement include the five decay angles as shown in Figure 7, and the dilepton invariant masses: M_{12} and M_{34}. The BDT analysis will use these variables to create a final discriminate variable.

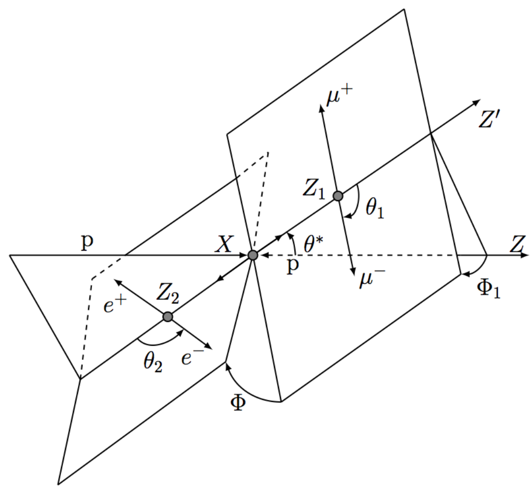

Figure 7: Illustration of the Higgs decay angular variables sensitive to the Higgs' spin and parity [4].

5.1.1 Boosted Decision Trees

The Boosted-Decision-Tree (BDT) is a multivariable analysis technique that was developed at the University of Michigan, and is used in this analysis to separate different spin and parity states in hypothesis tests. BDTs are often used with statistically limited datasets. Since no single spin sensitive variable can be used to completely determine the spin and parity, a BDT has been employed. BDTs utilize the separation power of each variable to separate two samples. Additionally, the BDT produces a single BDT output (score) for each event, which reduces the multivariable input into a single discriminating variable. Details of the BDT algorithm can be found in [11], and a brief description is given below.

The BDT technique involves a 'training' procedure for event pattern recognition. The BDT requires two data samples to separate, such as a *signal* and *background* sample. The data is represented by a set of physics variable distributions. These physics variables are given to the BDT as input. A decision-tree splits data recursively based upon cuts on the input variables until a stopping criterion is reached (e.g. *signal* purity, minimum number of event, and designed number of decision-tree nodes). After these recursive splittings, every event ends up in a signal (score=1) or a background (score=-1) leaf of the decision-tree. Misclassified events will be given larger weights in the training of the next decision-tree (boosting). This procedure is repeated several hundreds to thousands of times until the performance is optimized. The discriminator produced by the BDT training is the sum of the weighted scores from all the decision-trees. If the total score for a given event is relatively high this event is most likely a signal event, and if the score is low it is likely a background event.

Applying this technique to separate different spin and parity hypotheses, the $J^P = 0^+$ hypothesis is treated as *signal* and alternative J^P state as *background* in the BDT training. The BDT output (score) distribution is used as the final discriminant. The BDT output distributions are later used as the *signal* (0^+ state) and *background* (alternative J^P state) probability density functions for toy MC experiments produced to perform the statistic analysis for final spin and parity state measurement, as described in Section 5.3.

The systematic uncertainties of the BDT discriminants are evaluated by varying the input variable calibration uncertainties. These uncertainties are taken into account when producing the toy MC experiments for hypothesis testing and are described in Section 5.4. Additionally, BDTs suffer from overtraining systematics caused by insufficiently large samples. Overtraining occurs when a certain type of event is over sampled by the BDT and therefore over represented in the training, the statistical fluctuations in small samples can be large. BDTs can also be overtrained if the depth of the tree is too large. Overtraining effects can be seen when comparing the BDT output distribution from the sample used for training and the BDT output distributions from a "test" sample, often these training and testing samples are each half of the original input sample.

5.1.2 Spin and Parity Sensitive Variables

The $H \to ZZ^* \to 4l$ channel, where $l = e$ or μ, benefits from the full reconstruction of four-lepton final state, which produces many spin and parity sensitive variables. The observables sensitive to the Higgs spin and parity are the reconstructed Z boson masses, M_{12} and M_{34}, as well as five decay angles that describe the decay of the Higgs boson in its rest frame (Figure 7): the production angle θ^* and the decay angles Φ, Φ_1, θ_1, and θ_2.

The production angle θ^* defined in the four-lepton rest frame is the angle between Z_1 and the beam pipe. Φ_1 is the azimuthal angle of the Z_1 decay plane. The decay angle Φ is the azimuthal angle between the decay planes of Z_1 and Z_2. The decay angle θ_1 (θ_2) is the angle between the decay vector of the negative reconstructed lepton and the direction of flight of Z_1 (Z_2). The expected (via MC) and observed distributions of each variable, as well as a comparison between the expected results for 0^+ and 0^-, can be seen in Figure 8, 9, and 10. From these distributions it is evident that no single variable can be used to determine if data favors 0^+ or 0^-. For this reason the BDT multivariable analysis must be used. Additionally, the BDT output distribution for 0^+ versus 0^- can be seen in Figure 11

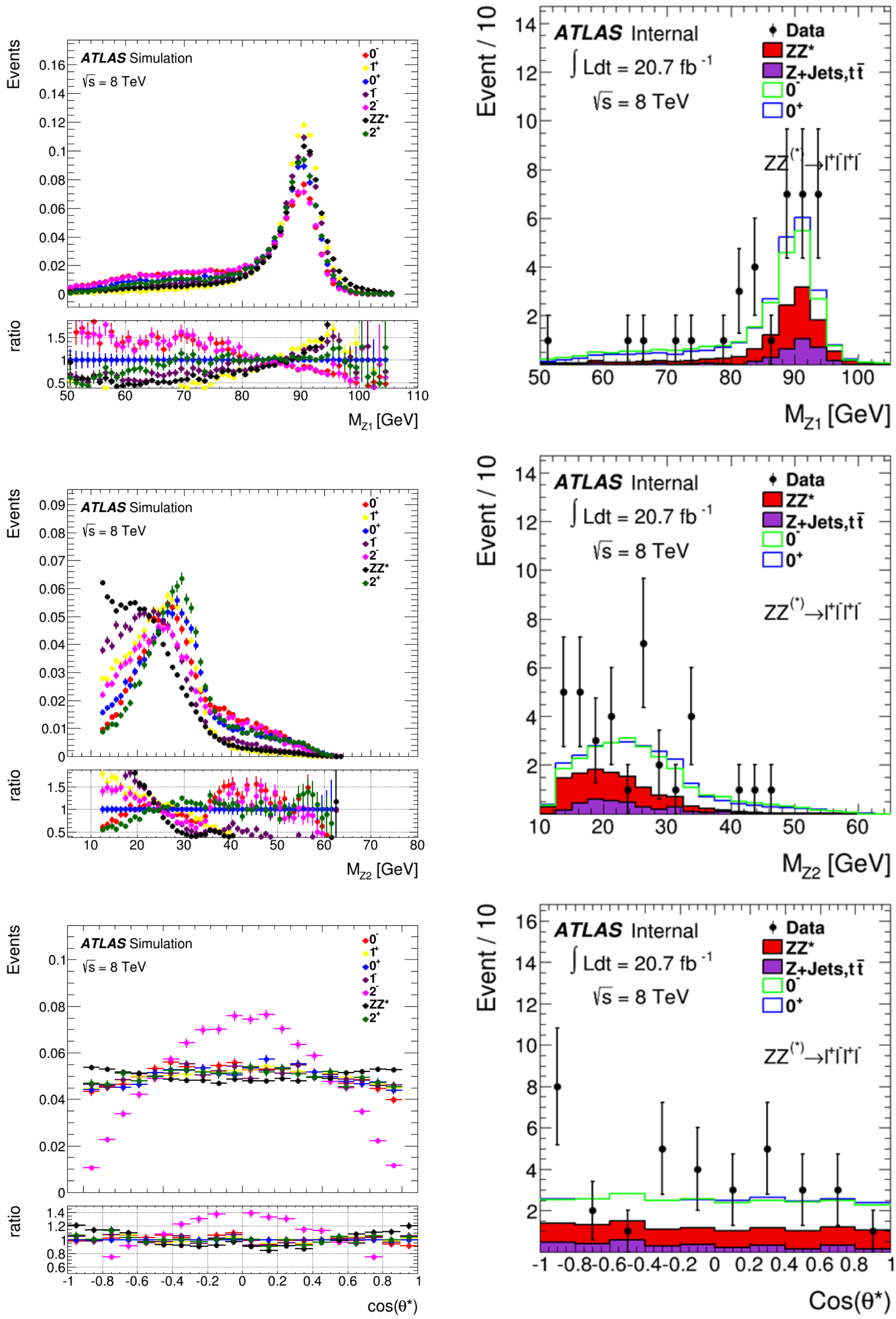

Figure 8: Comparison of kinematic J^P sensitive observables and the production angle that describes the decay of Z_1 for different J^P states and with data compared to 0^+ and 0^-.

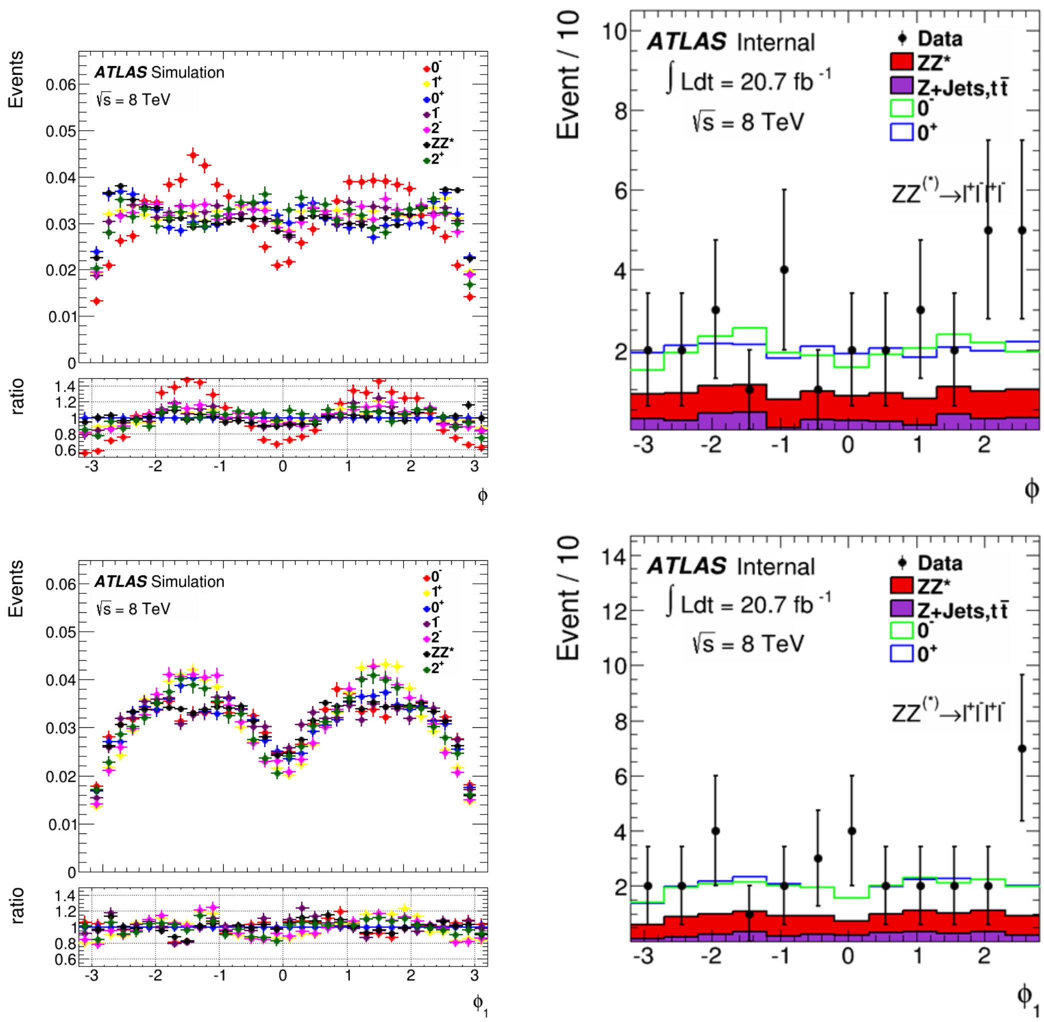

Figure 9: Comparisons of angular J^P sensitive observables which describe the Z decay planes for different J^P states and with data compared to 0^+ and 0^-.

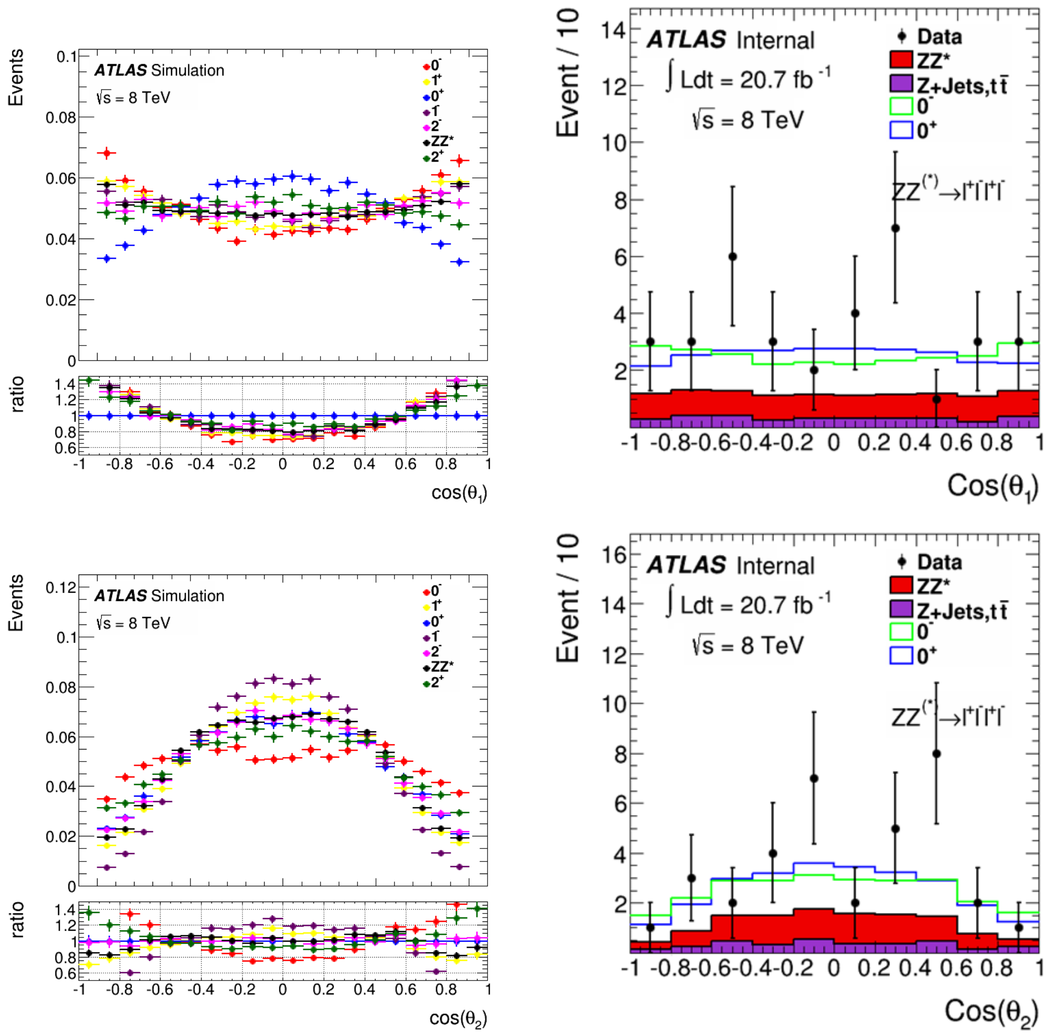

Figure 10: Comparisons of angular J^P sensitive observables which describe the Z decay products for different J^P states and with data compared to 0^+ and 0^-.

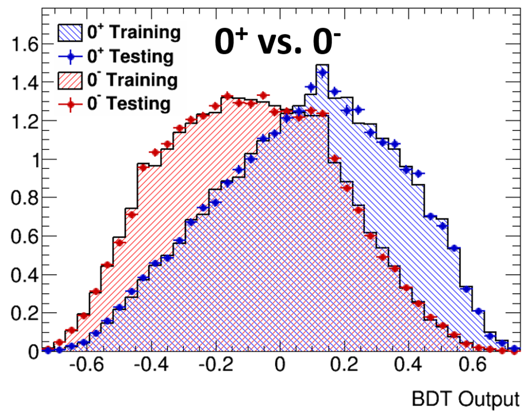

Figure 11: BDT output from training 0^+ versus 0^-.

5.2 Probability Density Function Creation

The hypothesis test used to obtain the final significance or exclusion of J^P states utilizes many conditional Poisson probabilities based upon the given and expected number of events, which is detailed in Section 5.3. To model the expected number of events, 1 dimensional probability density functions (PDFs) are made from the BDT output (D_{J^P} discriminator) and normalized to the expected number of events for the corresponding PDF. Furthermore, to better separate the Higgs signal from the background, these PDFs are made in *high* and *low* pairs defined by either a high or low signal to background ratio. The high signal to background region is defined as $121 < M_{4l} < 127$ GeV, while the low signal to background region covers $M_{4l} > 127$ GeV and $M_{4l} < 121$ GeV. A PDF is made for each decay channel in the two J^P states being considered, as well as for each background. This set of plots is also made for each systematic, as will be discussed in Section 5.4. Since these PDFs will be used to calculate the expected number of events, we do not expect there to be much fluctuation between neighboring bins. Additionally, the PDFs cannot have empty bins, as they have a large negative effect on the hypothesis test results since the probability of finding an event in this region should not be 0. To account for the empty bins at the tails of the distributions and to make sure bins do not vary too much from their neighbors, each PDF is smoothed using a kernal density estimator (kde) with a Gaussian kernal of varying width. An example of a PDF can be seen in Figure 12, where the PDFs for 0^+ and 0^- are compared to each other and data. From this distribution it is evident the separation power between the two hypotheses is still small given the amount of data we currently have. However, we can still get good exclusion confidence levels utilizing a hypothesis test method that samples these PDFs, as described in 5.3.

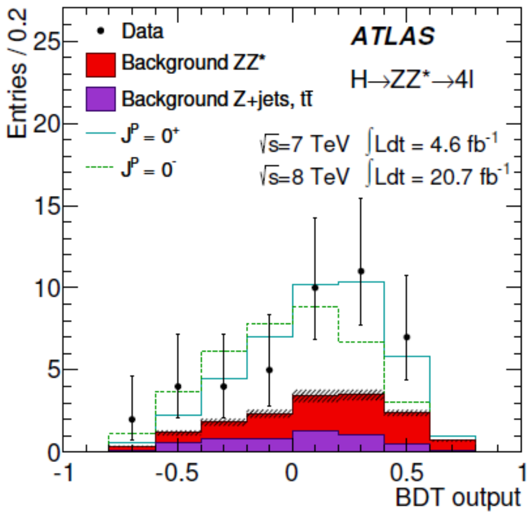

Figure 12: Distribution of the 0^+ versus 0^- BDT output for data and MC, where the J^P=0^+ hypothesis (solid line) and the J^P=0^- hypothesis are plotted together for comparison.

The kde is a way to estimate the PDF of a random variable using finite data points by treating entries as a PDF themselves, instead of a single point, by placing a normalized distribution, such as a Gaussian in this case, at the point of each data entry. The final PDF is given by summing individual distributions placed at each data entry. The formal definition of a kde with a Gaussian kernel is

$$f(x) = \frac{1}{N}\sum_{i=1}^{N} K\left(\frac{x - x_i}{h}\right) \text{ where } K(u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} \tag{1}$$

where h is referred to as the bandwidth. In the Gaussian case, h acts similar to the standard deviation.

The kde used in this measurement has a variable bandwidth based upon the number of events in a given region. In regions well populated by data points, the bandwidth is small so the Gaussians resemble points, while the bandwidth is increased in regions with fewer events. This process protects the original shape of the distribution from being over-smeared in areas with higher populations, while still providing a smooth distribution with no empty bins, which is expected for this analysis.

5.3 Statistical Method

To measure the separation between different J^P states, the PDFs created from BTD outputs, described above, are used to create a likelihood function $\mathcal{L}(J^P, \mu, \theta)$ that depends on the J^P assumptions. This likelihood is the production of conditional probabilities over the binned PDFs in each channel, and is calculated twice under the assumption of the signal and null hypotheses (H_0 and H_1). These two likelihoods are then used to create a final test-statistic, q. This process of calculating the test statistic is done thousands of times using MC toy experiments, where the expected number of events for signal and background are slightly varied based upon a Poisson Distribution. These thousands of toy MCs are used to create a distribution of test statistics for the case of assuming H_0 and H_1, which are then used to determine the corresponding p_0 value for both hypotheses.

As mentioned, the likelihood is the production of conditional probabilities over each bin of the PDFs:

$$\mathcal{L}\left(J^P, \mu, \theta\right) = \prod_{n}^{N_{Chann.}} \prod_{j}^{N_{Bins}} P\left(N_{n,j} | N_{n,j}^{Exp.}\right). \tag{2}$$

Here N is the given number of events either from data or simulated data, μ is the signal strength, and θ represents the nuisance parameters. $P\left(N_{n,j} | N_{n,j}^{Exp.}\right)$ is the Poisson probability of getting N events from data, given an expected number of events N^{Exp}.

The expected number of events, $N^{Exp.}$, is determined using the signal and background PDFs:

$$N_{n,j}^{Exp.} = \mu_n \mathscr{L} N_n^{signal} \left[\varepsilon \cdot PDF_{n,j}^{H_1} + (1-\varepsilon) \cdot PDF_{n,j}^{H_0}\right] + \sum_{b}^{N^{Background}} N_{n,b}^{Background} PDF_{n,b,j}^{Background} \tag{3}$$

Here ε is the parameter of interest, which is either 0 or 1 when looking at hypothesis H_0 ($\varepsilon = 0$) or H_1 ($\varepsilon = 1$). Additionally, $PDF_n^{Background}$ and PDF_n^{Signal} are the normalized PDFs of the nth channel created from the BDT output for background and signal, respectively, with the index j summing over the bins. $N_{n,b}^{Background}$ and N_n^{Signal} are the expected number of signal and background b events, as described in Section 4.2. $\mathscr{L}$ is a normalized variable that accounts for the uncertainty in the total integrated luminosity. The two likelihoods, $\mathcal{L}(H_1, \mu, \theta)$ and $\mathcal{L}(H_0, \mu, \theta)$ are calculated in the same way, except ε is 1 (0) for H_1 (H_0). $N_{n,b}^{Background}$ and N_n^{Signal} represent the number of events in the bth background and signal, respectively, in the nth channel. The parameters $\mathscr{L}$, $N_{n,b}^{Background}$, N_n^{Signal}, and the signal strength (μ) are nuisance parameters whose average values and uncertainties are found from the nominal event selection. The nuisance parameters are constrained by Gaussian terms with a standard deviation of the uncertainty. The nuisance parameters in the likelihood are profiled, such that each parameter is fitted to a value that maximizes the likelihood. The final test statistic used to distinguish between the two signal J^P states is based on the ratio of the two profiled likelihoods:

$$q = log\left(\frac{\mathcal{L}\left(J^P = 0^+, \hat{\hat{\mu}}_{0^+}, \hat{\hat{\theta}}_{0^+}\right)}{\mathcal{L}\left(J^P_{Alt}, \hat{\hat{\mu}}_{J^P_{Alt}}, \hat{\hat{\theta}}_{J^P_{Alt}}\right)}\right), \tag{4}$$

where $\mathcal{L}\left(J^P, \hat{\hat{\mu}}_{J^P}, \hat{\hat{\theta}}_{J^P}\right)$ is the profiled maximum likelihood, and $\hat{\hat{\mu}}_{J^P}$ and $\hat{\hat{\theta}}_{J^P}$ represent the values of the signal strength and nuisance parameters fit to the data under each J^P hypothesis.

In order to create distributions of the test statistic for each hypothesis, the test statistic is calculated thousands of times using MC toy experiments. In each toy experiment the number of expected signal and background events are smeared by a Poisson random number to create a distribution of q statistics. This is done by multiplying the expected number of signal events and the expected number of events from each background by a different Poisson random number. By smearing the number of expected events in each bin by this process, the Poisson statistic $P\left(N_{n,j}|N_{n,j}^{Exp.}\right)$ is changed, resulting in a different likelihood and subsequently a different test statistic. Test statistics are calculated thousands of times using this method to create PDFs that will then be compared to the test statistic calculated using data. When comparing the value of the test statistic from real data to the distributions made from the MC toy experiments, the p_0 value can be calculated by integrating the tail of the distributions that do not include the data test statistic in a simple hypothesis test fashion. The exclusion level of a specific hypothesis while assuming another J^P hypothesis can be calculated using the p_0 value.

5.4 Systematic Uncertainties

Most of the systematic uncertainties come from the event selection used to find Higgs candidates. Additionally, theoretical uncertainties, background normalization uncertainties, lepton reconstruction uncertainties, and luminosity uncertainties are considered. These uncertainties consequently affect the shapes and normalization of the PDFs. These affects are taken into account in various ways depending on how the uncertainty affects the measurement.

5.4.1 PDF Shape Uncertainty Treatment

The systematics described below will change the total number of events, in addition to the shape of the PDF. To study the effects of a systematic, the systematic is shifted both up and down by one standard deviation. The first sample (σ_+) is produced by shifting the systematic upwards by one standard deviation, while the second sample (σ_-) is the results of the event selection run with the systematic shifted down by one standard deviation. This is often made possible by the analysis packages used which often have a built in option to add or subtract one standard deviation to the systematic in question. The results from the event selection using these samples are then used to produce PDFs from the BDTs that were trained using the nominal sample. In the hypothesis test, each PDF from the nominal, σ_+, and σ_- sample are fit to find the mean. The means from the σ_+ and σ_- PDFs are used to determine the standard deviation of the mean value in the PDF. These numbers are used to create a normalized distribution with standard deviations determined by the mean of the σ_+ and σ_- PDF means. During each MC pseudo-experiment a random number for each systematic is chosen from the respective distribution and then multiplied by the expected number of events, Equation 3, used in the Poisson probability to calculate the likelihood in Equation 2. The expected number of events is smeared in this way for each bin independently.

MC Event Modeling

Lepton reconstruction and identification introduces the largest systematic uncertainties. The systematic errors introduced by the electron energy corrections and momentum smearing, as well as the muon momentum smearing, are considered in this analysis by producing separate samples with these corrections individually shifted up and down by 1 standard deviation. Additionally, the uncertainties introduced from the muon and electron scale factors are considered in the same manner.

Mis-Pairing Effects

Same flavor channels are affected by incorrectly pairing same flavor opposite sign leptons when forming Z_1 and Z_2. For the 0^+ MC sample the mis-pairing fraction is around 4.2% in the mass region of 115 and 130 GeV. This effect has been shown to be negligible on BDT shapes when the fraction of mis-paired events is changed by 10%. For this reason the mis-pairing effects on the BDT shapes has been neglected.

Mass Resolution

The M_{4l} resolution has been taken into consideration in the same way as the lepton reconstruction and identification uncertainties. This was done by shifting the Higgs mass both up and down by 1 GeV to create a σ_+ and σ_- PDF.

BDT Overtraining

Limited statistics in the BDT training can result in specific types of events having more weight in the training than others and result in a BDT that is not trained to an sample that adequately represents data. Overtraining can lead to inefficiencies in the separation of the two samples, but does not represent a source of systematic errors. To assess the magnitude of this affect, the BDT output from the training sample were compared to the BDT output of the test sample. By comparing the separations J^P states with the two independent samples (training and testing samples), the overtraining is seen to be negligable. This comparison can be seen in Figure 11 for 0^+ versus 0^-. These comparisons illustrate there are no significant differences between the testing and training sample, and thus the BDT was adequately trained to correctly reflect two input samples.

Theoretical Uncertainties

As described in Section 3.2, events from the ***JHU*** generator are calculated with leading order perturbative QCD corrections and must be reweighted based upon their p_{T}. This is done so the lepton p_{T} distributions from the ***JHU*** generator agree with the lepton p_{T} distributions from the next to leading order SM Higgs event generator: ***PowHeg***. The systematics from this reweighting procedure have been considered and treated in the same manner as the other reweighting systematics introduced by MC corrections.

5.5 Results

The results are obtained from running 100,000 pseudo-experiments to obtain a final p_0 value in order to exclude or validate J^P states. The p_0 values for different J^P states compared to 0^+ can be seen in Table 1. The resulting distributions of the test statistic for 0^+ and 0^- formed by the toy MC experiments, as well as the test statistic of data can be seen in Figure 13. Additionally, the expected and observed p_0 values and exclusion limits for 0^+ versus 2^+ with different amounts of $q\bar{q}$ production ($f_{q\bar{q}}$) can be seen in Table 2. The combined results from different Higgs decay channels for the expected and observed confidence levels when assuming $J^P = 0^+$, as well as for 2^+ for various $f_{q\bar{q}}$, can be seen in Figure 14. These results indicate that data is in favor of a $J^P = 0^+$ Standard Model Higgs [3].

Alternative J^P	$p_0(J^P=0^+)$ for assumed J^P_{Alt} Expected	Observed	$p_0(J^P=J^P_{Alt})$ for assumed 0^+ Expected	Observed	Exclusion CL
0^-	1.5×10^{-3}	0.015	3.7×10^{-3}	0.31	97.8%
1^+	4.6×10^{-3}	0.001	1.6×10^{-3}	0.55	99.8%
1^-	0.9×10^{-3}	0.051	3.8×10^{-3}	0.15	94%
2^+	0.099	0.532	0.092	0.079	83.1%

Table 1: p_0 values and exclusion confidence levels given for various J^P states when compared to 0^+.

$f_{q\bar{q}}$	$p_0(J^P=0^+)$ for assumed 2^+ Expected	Observed	$p_0(J^P=2^+)$ for assumed 0^+ Expected	Observed	Exclusion CL
100%	0.102	0.962	0.082	0.001	97.4%
75%	0.117	0.923	0.099	0.003	96.1%
50%	0.129	0.943	0.113	0.002	96.5%
25%	0.125	0.944	0.107	0.002	96.4%
0%	0.099	0.532	0.092	0.079	83.1%

Table 2: Expected and observed p_0 values and exclusion confidence levels for the J^P=0^+ and J^P=2^+ hypotheses with different fractions of $q\bar{q}$ production.

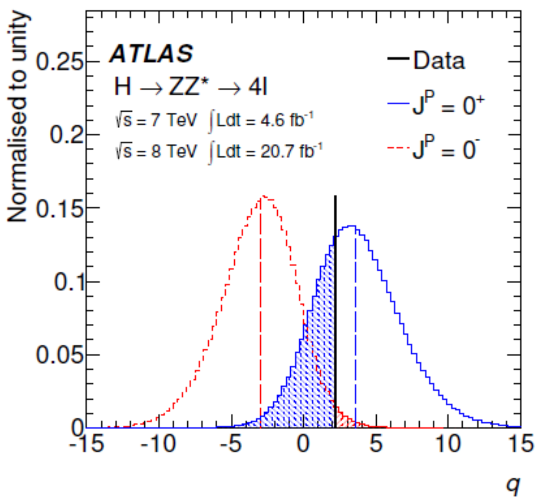

Figure 13: Hypothesis test result for 0^+ versus 0^-, with the respective p_0 values shaded in.

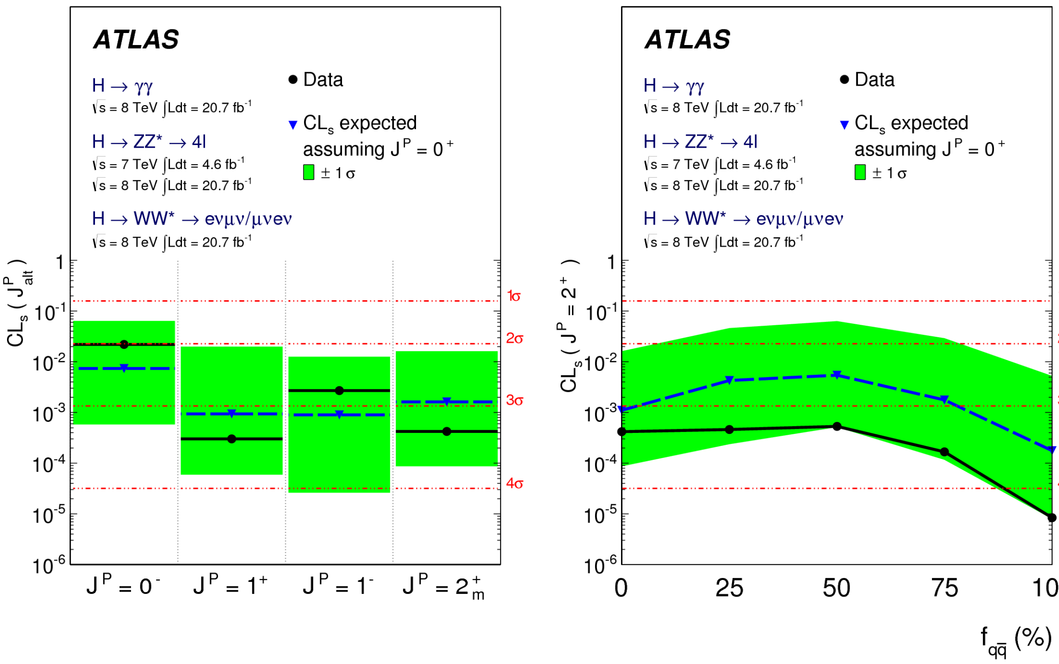

Figure 14: Combined expected (blue dashed lines) and observed (black solid lines) confidence levels, as well as Gaussian standard deviations, from other Higgs decay channels when comparing various J^P states with 0^+ (left). As well as the expected and observed results for comparing 0^+ versus 2^+ as a function of the percentage of events from $q\bar{q}$ processes (right). The green bands represent the 68% expected exclusion range for a signal with assumed J^P=0^+.

6 Ongoing Work

6.1 2 Dimensional PDF

To improve the sensitivity of this study through better separation of the Higgs signal and the background, we are currently altering this measurement to include 2 dimensional PDFs made from two BDTs. The second BDT will be trained with the Standard Model Higgs sample and ZZ^* sample. The BDT output from training 0^+ versus ZZ^* is plotted against the the BDT output from the 0^+ versus J^P_{Alt} BDT output. The two BDT discriminants used to form the 2 dimensional PDF are referred to as D_{J^P}, for separating J^P states, and D_{ZZ}, for separating signal and background. The training of the D_{ZZ} variable is described in Section 6.2. Again, a PDF is made for each spin state and in each individual decay channel, as well as for each background. However, there are no *high* or *low* PDFs, since the D_{ZZ} variable is used to separate Higgs events from background. This set of PDFs is also made for each systematic and cannot have empty bins and should be smooth as well. Therefore, the 2D PDFs are smoothed through the kde method. An example of 2D PDF before and after smoothing can be seen in Figure 15. These 2 dimensional PDFs can be seen in Figure 16, where the 0^+, 0^-, and ZZ^* PDF distributions are overlayed, as well as projected into two 1 dimensional PDFs, to illustrate the separation power of this new PDF.

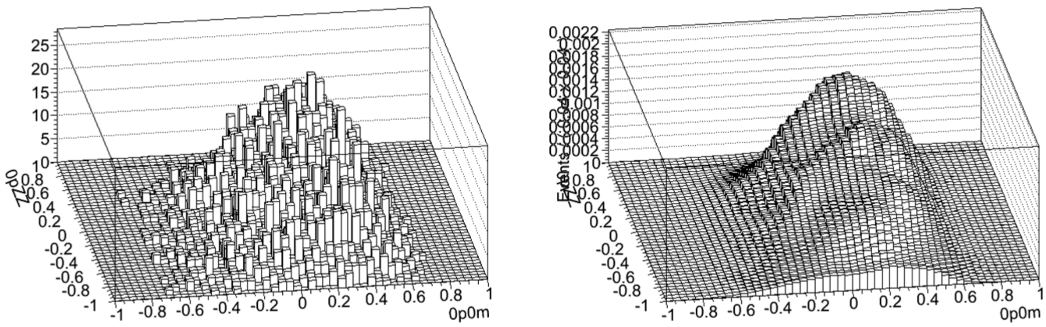

Figure 15: A 2 Dimensional PDF of the 0^+ vs. 0^- BDT output plotted against the 0^+ vs. ZZ^* BDT output both before smoothing (left) and after smoothing (right).

6.2 Discriminants for Higgs Signal Versus Background Separation

To separate the Higgs signal from the ZZ^* background after the event selection, a BDT is trained using kinematic variables. Only variables that are not sensitive to the Higgs spin are chosen in order to reduce biases towards a specific J^P state since this BDT is trained using a SM Higgs ($J^P = 0^+$) sample and SM ZZ^* sample. The observables chosen are M_{ZZ}, η_{4l}, and p_{T}^{4l}. The expected distributions of each observable for the SM Higgs and SM ZZ^* for using MC, and the observed distributions in data, can be seen in Figure 17. The BDT output, along with the ROC curve can be seen in Figure 18. This BDT is well trained, as is evident from the agreement between the testing and training samples' BDT output in Figure 18. Additionally, this BDT has a strong separation power between the Higgs signal and ZZ^* background, as shown by the large ROC integral of 0.814.

6.2.1 Matrix Element Kinematic Discriminant

Currently we are investigating the addition of a Matrix Element Discriminating variable (MEKD) to the current observables used to separate the Higgs signal from ZZ^*. MEKD is a discriminating variable for

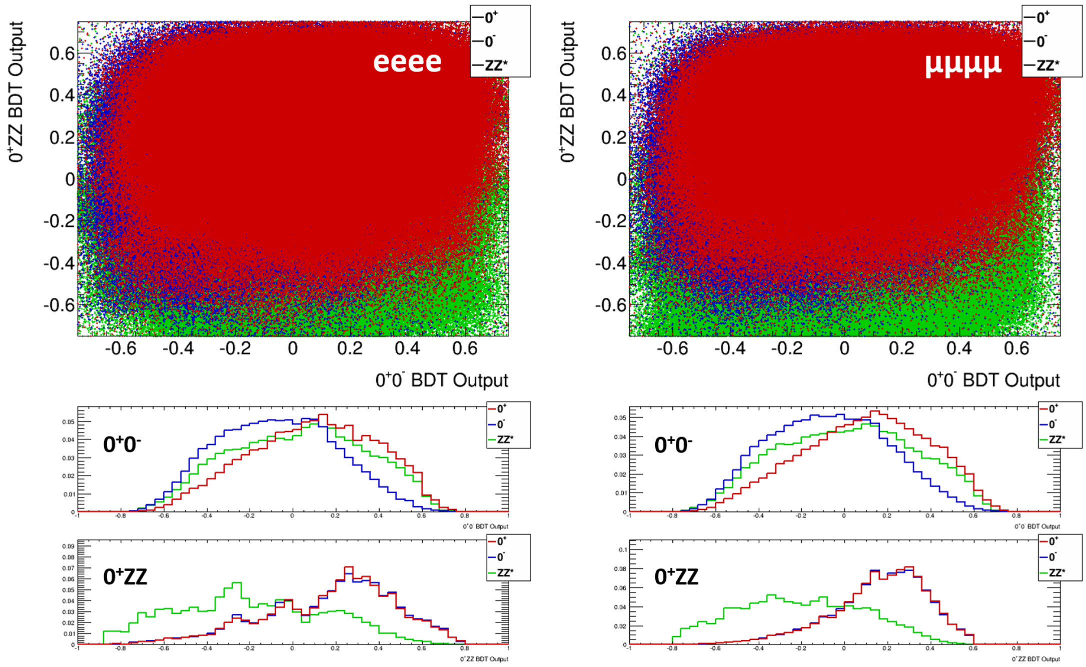

Figure 16: 2 Dimensional PDFs separating 0^+ and 0^-, as well as the Higgs signal and ZZ^* background through the D_{ZZ} and D_{J^P} discriminants built from BDTs in the $4e$ channel (left) and 4μ channel (right).

four lepton processes involving two Zs ($X \to ZZ \to 4l$) that is based upon lepton kinematics $\mathbf{p}$. The MEKD discriminates between two hypothesis (A and B) via a ratio of the probability of observing either event given the alternative production hypotheses:

$$D(A;B) = \frac{P(\mathbf{p}|A)}{P(\mathbf{p}|B)}, \tag{5}$$

where $P(\mathbf{p}|A)$ and $P(\mathbf{p}|B)$ are the probability density functions for observing the event in the case of either hypothesis. However, the probabilities can be represented by the matrix elements of the corresponding process ($\mathscr{M}_A$ and $\mathscr{M}_B$) derived from their Feynman diagrams. The MEKD variable is thus defined as

$$KD(A;B) = ln\left(\frac{|\mathscr{M}_A\,(a+b \to 4l)\,|^2}{|\mathscr{M}_B\,(a'+b' \to 4l)\,|^2}\right). \tag{6}$$

Here a and b (a' and b') stand for different types of partons that can produce a four lepton final state via process A (B). The log of the ratio is used for technical convenience due to the large dynamic range in the ratio. The addition of MEKD to the D_{ZZ} discriminator boosts its separation power, which can be seen in Figure 19, where the ROC integral increases from 0.814 to 0.873. Furthermore, the separation power of the hypothesis test is increased by 15%.

To reduce our baises towards a particular J^P state, the MEKD's correlation to spin sensitive variables was investigated. Density plots of the MEKD versus the spin sensitive variables can be seen in Figure 20. From these plots we can see that MEKD is slightly correlated to M_{34} and has no correlation to any other spin sensitive variables.

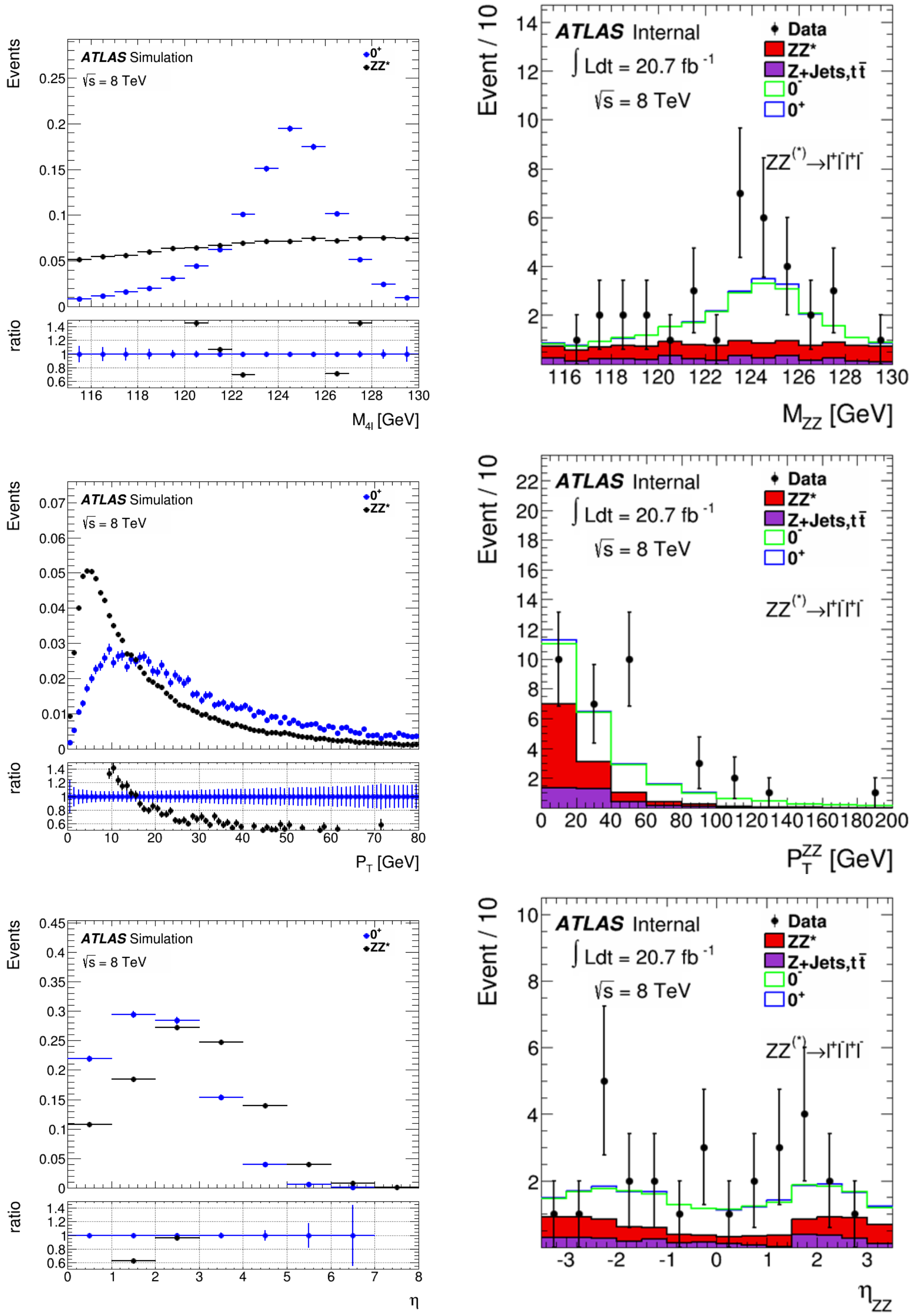

Figure 17: Discriminant variables used in the training of Higgs vs. ZZ^* BDT and the comparison of 0^+ and 0^- with data.

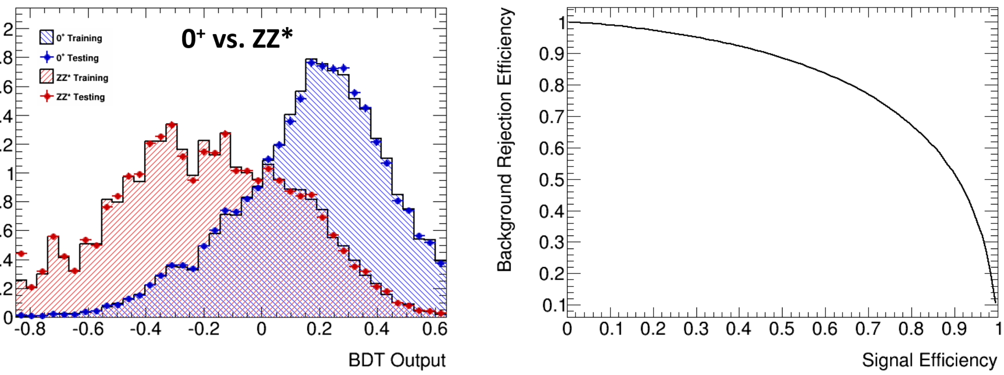

Figure 18: BDT output from both the testing and training sample for 0^+ vs. ZZ^* (left) and the ROC curve from this training (right) with an ROC integral of 0.814

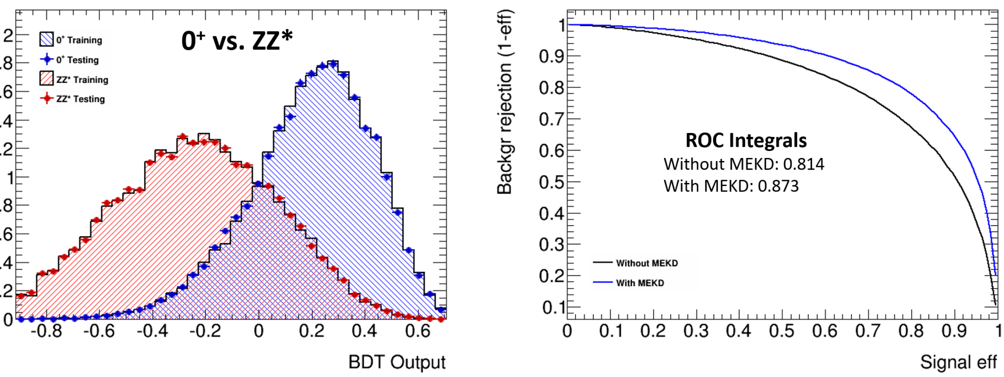

Figure 19: BDT output distribution for D_{ZZ} with MEKD (left), and the ROC curves for training D_{ZZ} with and without MEKD.

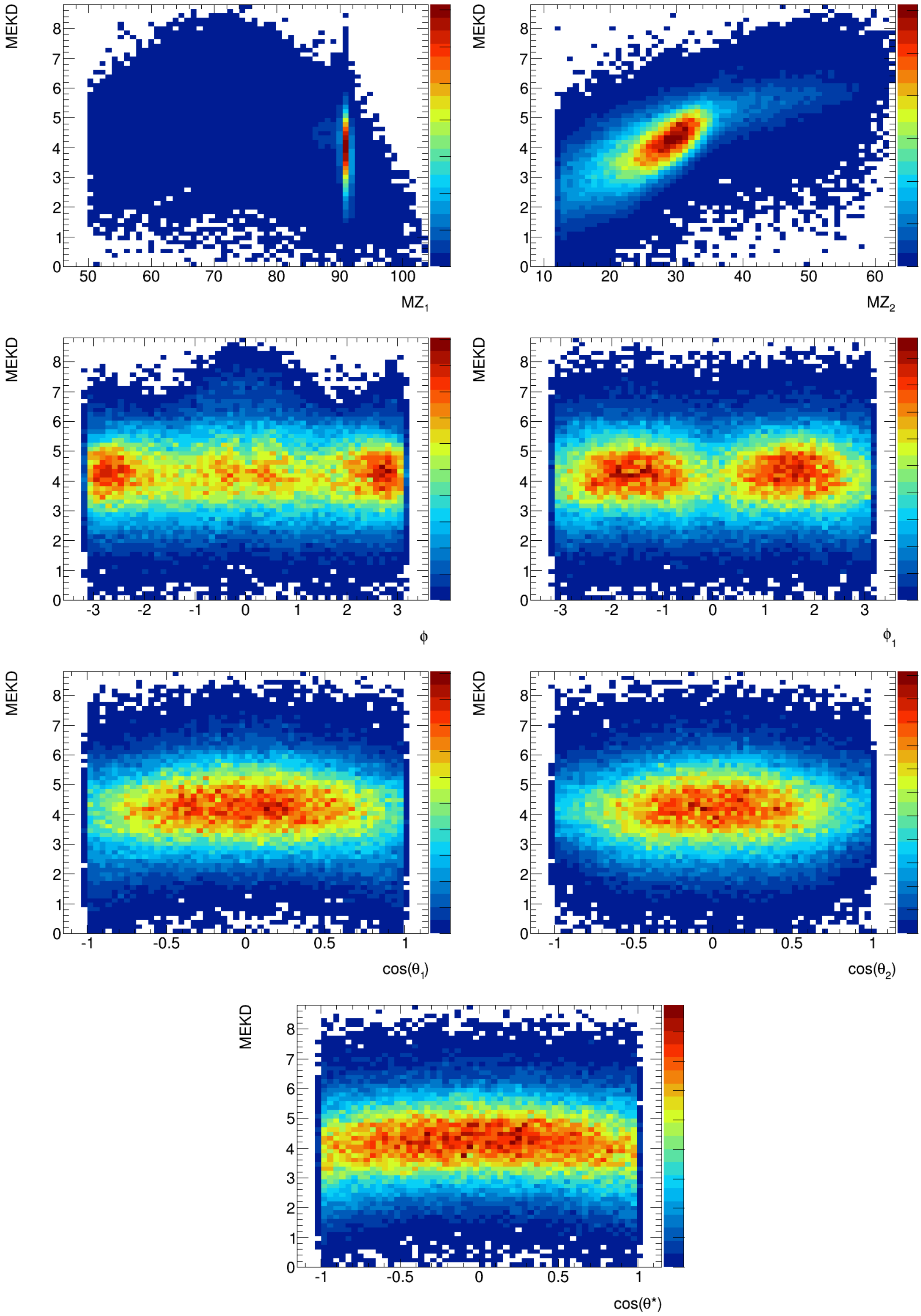

Figure 20: Correlation plots of the MEKD discriminant versus spin sensitive variables.

7 Conclusion

The SM Higgs spin and parity hypothesis has been compared to multiple spin and parity hypotheses through the use of BDTs and toy MCs using ATLAS data collected at the LHC. The results indicate the SM Higgs spin and parity ($J^P = 0^+$) is favored and 0^-, 1^+, 1^-, and 2^+ have been excluded with 97.8%, 99.8%, 94%, and 83.1% confidence, respectively. Current work is being done to improve these exclusions, including the addition of 2 dimensional PDFs where the second dimension is used to discriminate against the SM ZZ^* background. Additionally, it has been shown that the addition of the MEKD variable significantly improves the separation power of the D_{ZZ} discriminant. These additions are nearing their final stages and will produce unblinded results along with the rest of ATLAS with a new publication in 2014.

References

[1] A particle consistent with the higgs boson observed with the atlas detector at the large hadron collider. *Science*, 338(6114):1576–1582, Dec 2012.

[2] ATLAS Public Results, 2013.

[3] Evidence for the spin-0 nature of the higgs boson using atlas data. *Physics Letters B*, Jul 2013.

[4] Measurements of the properties of the Higgs-like boson in the four lepton decay channel with the ATLAS detector using $25fb^{-1}$ of proton-proton collision data. Report No. ATLAS-CONF-2013-013, 2013.

[5] Agostinelli S.; Allison J.; Amako K.; Apostolakis J.; et al. Geant4 - a simulation toolkit. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors, and Associated Equipment*, 506, 2003.

[6] Alioli S.; Nason P.; Oleari C.; Re E.; et al. Nlo higgs boson production via gluon fusion matched with shower in powheg. *Journal of High Energy Physics*, 904, Dec 2008.

[7] Frixione S.; Stoeckli F.; Torrielli; et al. The mc@nlo 4.0 event generator. Oct 2010.

[8] Gao Y.; Gritsan A.; Guo Z.; Melnikov K.; et al. Spin determination of single-produced resonances at hadron colliders. *Physical Review D*, 81, Apr 2010.

[9] Golonka P.; Was Z.; et al. A bried introduction to pythia 8.1. *European Physical Journal*, C45:97–107, 2006.

[10] Mangano M.; Piccinini F.; et al. Alpgen, a generator for hard multiparton processes in hadronic collisions. 2003.

[11] Roe B.; Yang H.; Zhu J.; Liu Y.; et al. Boosted decision trees as an alternative to artificial neural networks for particle identification. *Nuclear Instruments and Methods A*, Nov 2004.

[12] Sjostrand T.; Mrenna S.; et al. A brief introduction to pythia 8.1. *Computational Physics Communications*, 178:852–867, Jan 2008.

[13] Brumfiel G. Down the petabyte highway. *Nature*, 469, Dec 2011.