

SEARCH FOR THE STANDARD MODEL HIGGS BOSON
IN THE DECAY CHANNEL $H \rightarrow ZZ \rightarrow 2L2Q$ AT CMS

A Preliminary Report

Submitted to the Faculty

of

Purdue University

by

Matthew K. Kress

In Partial Fulfillment of the

Requirements for the Degree

of

Doctor of Philosophy

October 2012

Purdue University

West Lafayette, Indiana

TABLE OF CONTENTS

	Page
LIST OF TABLES	iv
LIST OF FIGURES	vi
ABBREVIATIONS	ix
ABSTRACT	x
1 Introduction	1
2 The Standard Model of Particle Physics	3
2.1 Forces and Particles	4
2.2 The Standard Model	5
2.3 The Higgs Mechanism	6
2.4 $H \rightarrow ZZ \rightarrow l^+l^-q\bar{q}$ Channel	6
2.5 Background Processes	7
3 Experimental Apparatus	9
3.1 The Large Hadron Collider	9
3.2 The Compact Muon Solenoid	11
3.3 The CMS Trigger System	12
4 Event Selection	15
4.1 Datasets	15
4.1.1 Data	15
4.1.2 Monte Carlo	15
4.2 Preselection	16
4.2.1 Lepton Selection	18
4.2.2 Jet Selection	20
4.2.3 Higgs candidates reconstruction	21
4.2.4 Kinematic fit	22
4.2.5 B-tagging of jets	23
5 Correlated and Uncorrelated Systematic Uncertainties	27
5.1 Trigger and Lepton Selection Efficiency	27
6 Signal Region Optimization	37
6.1 Tagging Classification	37
6.2 $H \rightarrow ZZ \rightarrow 2l2j$ decay	37
6.3 Signal Optimization Based on Helicity Neural Network	39

	Page
6.3.1 Neural Network Architecture	40
6.3.2 Results Based on Helicity Neural Network	41
6.3.3 Potential Improvement to to MVA	43
7 Conclusion	49
7.1 Future Plans	49
LIST OF REFERENCES	50

LIST OF TABLES

Table	Page
2.1 The four forces and their associated gauge bosons. Charge is in units of the proton charge.	4
2.2 The three generations of spin $\frac{1}{2}$ particles.	5
2.3 The main Standard Model background processes for the $H \rightarrow ZZ \rightarrow l^+l^-q\bar{q}$ channel and their cross sections at 8 TeV.	7
4.1 2011 data samples.	16
4.2 2012 data samples.	17
4.3 2011 background MC samples.	17
4.4 2012 background MC samples.	18
4.5 2011 signal MC samples.	19
4.6 POWHEG signal MC samples $H \rightarrow ZZ \rightarrow 2l2q$, $l = e, \mu, \tau$ 2012. . .	24
4.7 Event trigger requirements for data.	25
4.8 Electron ID requirements for the Loose ID working point.	25
4.9 Muon ID requirements for the Tight ID working point.	26
5.1 The Legs for the Electron HLT.	29
5.2 These are the event requirements for the tag and probe method of calculating trigger and lepton selection efficiency.	29
5.3 These are the official working points provided by the EGamma POG for the barrel.	30
5.4 These are the official working points provided by the EGamma POG for the endcaps.	31
5.5 ϵ_{RECO} efficiency values from 2012 data and Monte Carlo simulations using the Tag and Probe method by fitting the signal and background. . . .	32
5.6 $\epsilon_{ID} \times \epsilon_{ISO}$ efficiency values from 2012 data and Monte Carlo simulations using the Tag and Probe method by fitting the signal and background. . .	33

Table	Page
5.7 $\epsilon_{trigger}$ values for Ele8 Leg from 2012 data and Monte Carlo simulations using the Tag and Probe method by counting passing and failing probes.	34
5.8 $\epsilon_{trigger}$ values for Ele17 Leg from 2012 data and Monte Carlo simulations using the Tag and Probe method by counting passing and failing probes.	35

LIST OF FIGURES

Figure	Page
3.1 LHC injection scheme.	10
3.2 Integrated luminosity delivered by the LHC to each experiment in 2010	11
3.3 Integrated luminosity delivered by the LHC to each experiment in 2011	12
3.4 Integrated luminosity delivered by the LHC to each experiment in 2012	13
3.5 The CMS experiment. HCAL stands for hadron calorimeter and ECAL stands for Electromagnetic calorimeter.	13
3.6 A graphical slice of the CMS experiment	14
4.1 Di-lepton invariant mass in data and MC of $Z \rightarrow \ell\ell$ candidates after lepton selection. Left: Muon channel. Right: Electron channel. The selection described in preselection is applied except for the cut on m_{jj}	21
4.2 Di-jet invariant mass in data and MC for $\ell\ell jj$ candidates which have passed the pre-selection. Left: Muon channel. Right: Electron channel.	22
4.3 Kinematic fit results. Di-jet invariant mass vs. di-boson invariant mass for Higgs candidates (signal MC, $m_H = 325$ GeV) after loose selection requirements. Left: before kinematic fit; right: after kinematic fit. . .	23
5.1 Fitting of Supercluster to GSF Electrons for ϵ_{RECO} calculation.	29
5.2 Fitting of GSF Electrons for to WP Medium for ϵ_{ISO} calculation.	30
6.1 Diagram depicting the decay $X \rightarrow ZZ \rightarrow 2l2q$ and the 5 decay angles which describe it.	38
6.2 Five angular distributions of $\cos\theta_1, \cos\theta_2, \cos\theta^*, \Phi, \Phi_1$ and the helicity likelihood discriminant for 2012 electron data (points) and Summer 12 Monte Carlo samples (histogram). Open histograms indicate the expected distribution for a Higgs boson with mass 350 GeV, multiplied by a factor of 1000 for illustration. The selection described in preselection is applied except for the cut on m_{jj}	39

Figure	Page
6.3 Five angular distributions of $\cos \theta_1, \cos \theta_2, \cos \theta^*, \Phi, \Phi_1$ and the helicity likelihood discriminant for 2012 muon data (points) and Summer 12 Monte Carlo samples (histogram). Open histograms indicate the expected distribution for a Higgs boson with mass 350 GeV, multiplied by a factor of 1000 for illustration. The selection described in preselection is applied except for the cut on m_{jj}	40
6.4 Neural Network architecture used for the training.	41
6.5 The trainings are done after preselection and additionally require at least one TCHEM jet (Track Counting High Efficiency > 1.9) Left: Training 400 GeV Higgs boson with a MLP neural network. Right: Training 400 GeV Higgs boson with a Likelihood.	42
6.6 Signal and background Neural Network output in the 1tag region after preselection, the Z boson mass window cuts, cut on MET significance, and a -6%/+10% Higgs mass window. The signal is scaled to the sum of the Monte Carlo background.	43
6.7 Signal and background Neural Network output in the 2 tag region after preselection, the Z boson mass window cuts, cut on MET significance, and a -6%/+10% Higgs mass window. The signal is scaled to the sum of the Monte Carlo background.	44
6.8 Background Rejection Versus Signal Efficiency in the 1tag region comparing the Multi Variant Analysis output to the the Helicity Likelihood Discriminant for a Higgs mass of 200, 300, 400, and 500 GeV. This is calculated after preselection cuts, Z boson mass cuts, cut on MET significance, in a -6%/+10% Higgs mass window.	45
6.9 Background Rejection Versus Signal Efficiency in the 2tag region comparing the Multi Variant Analysis output to the the Helicity Likelihood Discriminant for a Higgs mass of 200, 300, 400, and 500 GeV. This is calculated after preselection cuts, Z boson mass cuts, cut on MET significance, in a -6%/+10% Higgs mass window.	45
6.10 Signal samples are normalized to background. Left: $\frac{Z_{upt}}{\sum pt}$ after preselection. Right: $\frac{Z_{upt}}{\sum pt}$ after preselection and $376 < m_{ZZ} < 440$ GeV. . . .	46
6.11 Applying a MLP training to preselection and at least one TCHEM jet. The working point is the equivalent performance of current analysis that we apply in the two tag region. For comparison the helyLDrefit variable is also shown. Left: ROC curves after preselection. Right: ROC curves after preselection and $376 < m_{ZZ} < 440$ GeV.	46

6.12	This training is for a MLP neural network trained on the 5 decay angles of a Higgs 400 GeV sample. This training is applied after preselection and requiring at least one TCHEM jet to samples with a Higgs mass of 300,400, and 500 GeV. The working point is the background rejection point that we currently achieve in the two tag region in our analysis. The helyLDrefit variable is shown for comparison.	47
------	--	----

ABBREVIATIONS

LHC	Large Hadron Collider
CMS	Compact Muon Solenoid
SM	Standard Model
fb^{-1}	inverse femtobarn
HLT	High Level Trigger
L1	Level 1 Trigger
WP	Working Point
POG	Physics Object Group
PAG	Physics Analysis Group
PDF	Probability Density Function

ABSTRACT

Kress, Matthew K. Ph.D., Purdue University, October 2012. Search for the Standard Model Higgs Boson in the decay channel $H \rightarrow ZZ \rightarrow 2l2q$ at CMS . Major Professor: Daniela Bortoletto.

A search for the Standard Model Higgs boson through the decay to two Z bosons to the final state of two leptons and two quark-jets, $H \rightarrow ZZ \rightarrow l^+l^-q\bar{q}$, is presented. The CMS experiment collected and analyzed data both in 2011 and 2012 of pp collisions at the LHC which are used in this search. The analysis uses the kinematics of the final state to select the Higgs boson signal in the mass range between 200GeV and 1000GeV. Using the probability of the jets to originate from light or heavy flavor quarks the events are classified and separately treated.

1. Introduction

The Standard Model of particles is one of the most successful theories that has been proposed. Since its inception, the predictions given by this theory have been confirmed numerous times with an amazing amount of precision. This has happened in various fields and with a number of experimental techniques. Despite all of the successes of this theory, its theoretical foundation rests on the observable quantum of the Higgs field called the Higgs boson. This particle lies as the capstone of the Standard Model.

The Standard Model does not constrain the mass of the Higgs boson. In the past many experiments have searched for the Higgs boson, and while they were not successful in discovery, they were able to exclude its existence in a number of mass ranges. A lower bound of 114.4 GeV was set at the Large Electron Positron (LEP) collider at CERN, and the Tevatron at Fermilab excluded the range of 156 - 177 GeV as well.

The Large Hadron Collider (LHC) was built at CERN to replace the LEP experiment. It is a superconducting proton-proton collider that ran at a center of mass energy of 7 TeV before 2012 and is running at 8 TeV currently. A large part of its design has been to produce proof of the existence of the Higgs boson.

The Compact Muon Solenoid (CMS) experiment is one of the four major experiments that analyzes the collisions produced at the LHC. The detector is sensitive to a wide range of processes but has been created to be particularly successful in Higgs boson searches.

The Higgs boson is an unstable particle and so its detection depends on the Higgs boson decay products. If it has a sufficiently large mass, the decay of Higgs to vector bosons will dominate the branching ratio of the decay channels. Of particular importance is the production of a Z-boson pair. If one of the Z bosons decay into a

lepton pair (e or μ) then the possible backgrounds produced at hadron colliders can be significantly diminished by cuts on the sharp Z mass peak.

In this preliminary report we examine the feasibility of searching for the Higgs boson in the $H \rightarrow ZZ \rightarrow l^+l^-q\bar{q}$ decay channel with data collected by the CMS experiment. Jets in the final state create a both advantages and disadvantages. Advantageously, the branching ratio of $Z \rightarrow q\bar{q}$ is significantly larger than $Z \rightarrow l^+l^-$ thus improving the number of Higgs evens we can see. The disadvantages of jets stem from complex reconstruction techniques and dealing with the added necessity of extremely effective background rejection.

The second chapter of this preliminary report gives an introduction to the Standard Model and the theory behind the need for the Higgs boson.

The third chapter describes the experiment, including both an introduction to the LHC and the CMS. The various sub-detectors of the CMS experiment are also described.

Chapter four describes the data-sets used for both data and Monte Carlo simulations. Afterword a description of preselection cuts for the analysis is given.

Trigger and Lepton Selection Efficiency for electrons are calculated and scale factors for data and Monte Carlo are given in Chapter five.

Chapter six gives the optimization for this signal channel.

2. The Standard Model of Particle Physics

Scientific knowledge grew in spurts and pockets until 1543 with the beginning of the Scientific Revolution. This period saw the work of many famous scientists like Nicolaus Copernicus, Galileo Galilei, Blaise Pascal, and more, until the publication of Isaac Newton’s “Mathematical Principles of Natural Philosophy” in 1687. The scientific method was further developed emphasizing experimentation and reason to explain the world. With the foundation work of previous scientists, the 18th century “Age of Enlightenment” broke upon Europe. Based firmly on the work of Newton, Pascal, and others the developments of modern mathematics, physics, and technology was accomplished.

In particular interest to physics, Isaac Newton proposed two successful physical theories: the laws of motion and the law of gravitation. These theories laid the backbone for classical mechanics and the theory of gravity. Further in the 19th century scientists like Michael Faraday and Georg Ohm studied electricity and magnetism with the eventual unification of the two into Maxwell’s equations, creating the theory of electromagnetism.

In 1960 Sheldon Glashow combined electromagnetism and weak interactions together to form the basis of the electroweak theory of interactions. [1] This allowed Steven Weinberg[2] and Abdus Salam[3] to incorporate the Higgs mechanism[4][5][6] into Glashow’s electroweak theory in 1967, giving it its modern form.

This Chapter will give an introduction to the current theoretical framework in elementary particle physics. We will describe the Standard Model and the motivation behind the Higgs mechanism. Additionally we will give a summary of the experimental limits placed on the Higgs boson by previous experiments. Finally we will present the $H \rightarrow ZZ \rightarrow l^+ l^- q \bar{q}$ decay channel and motivate its use in the search for the Standard Model Higgs boson.

2.1 Forces and Particles

Table 2.1

The four forces and their associated gauge bosons. Charge is in units of the proton charge.

Force	Boson	Charge	Mass
Gravitational	graviton(G)	0	0
Electromagnetic	photon(γ)	0	0
Weak	W boson(W^\pm)	± 1	81 GeV
	Z boson(Z)	0	92GeV
Strong	gluon(g)	0	0

There are four forces that govern the universe. These are the gravitational, electromagnetic, weak, and strong forces. The gravitational force is the weakest force and is primarily important in the interactions of large bodies, but does not play a significant roll in the interactions of particles. Attractions between charged particles are provided by the electromagnetic force. This is the primary force responsible for building atoms and molecules. The weak force is the motivator of radioactivity and fusion, while the strong force binds quarks into nucleons and nucleons to nuclei.

Our current understanding is that quarks and leptons are the the basic constituents of all matter. Leptons are spin 1/2 particles with a charge of -1. They do not participate in the strong interaction and additionally each one has a corresponding neutrino. These leptons are the electron, muon, and tau.

Quarks are the building blocks of hadrons and are held together by the strong force. There are three generations of quarks and leptons which are listed in Table 2.2.

Table 2.2
The three generations of spin $\frac{1}{2}$ particles.

charge	-1	-2/3	-1/3	0	+1/3	+2/3	+1
generation 1	e^-	\bar{u}	d	$\nu_e, \bar{\nu}_e$	\bar{d}	u	e^+
generation 2	μ^-	\bar{c}	s	$\nu_\mu, \bar{\nu}_\mu$	\bar{s}	c	μ^+
generation 3	τ^-	\bar{t}	b	$\nu_\tau, \bar{\nu}_\tau$	\bar{b}	t	μ^+

2.2 The Standard Model

For each of the four fundamental forces their interactions can be described both by fields or by exchanging particles. The Standard Model (SM) provides a quantitative description of three of the four interactions: electromagnetism, weak interactions, and the strong force. By the end of the 1960's the theory had been largely formed by Glashow, Weinberg, and Salam. [1–3] While it is desirable for a unified treatment of all the known forces, gravity plays a negligible roll in current particle physics experiments. Despite this, SM predictions have been confirmed by many experiments in the past decades.

Within the framework of relativistic quantum field theory the Standard Model is a description of the microscopic world in terms of interaction particles and fields. [4] It can be divided into two parts: quantum chromodynamics (QCD) and quantum electrodynamics (QED). This allows us to write the Lagrangian as the summation of two separate parts:

$$\mathcal{L}_{SM} = \mathcal{L}_{QED} + \mathcal{L}_{QCD} \quad (2.1)$$

In QED the unphysical infinite contributions can always be eliminated. This means that the theory is renormalizable. The Standard Model is renormalizable and also is compatible with special relativity. In 1960 Sheldon Glashow combined electromagnetism and weak interactions together to form the basis of the electroweak theory of interactions. [1]

2.3 The Higgs Mechanism

If a doublet of scalar fields is introduced then its self-interactions provide spontaneous symmetry breaking and give masses to the gauge and fermion fields. This addition to the Lagrangian is \mathcal{L}_Φ and \mathcal{L}_Φ^F which are described in Equation 2.2. It also introduces the Higgs boson, a new neutral scalar particle.

$$\mathcal{L}_\Phi = |D_\mu \Phi|^2 - V(|\Phi|^2) \quad (2.2)$$

For the scalar potential V the most general form that is renormalizable is Equation 2.3.

$$V = \mu^2 |\Phi|^2 + \lambda |\Phi|^4 \quad (2.3)$$

From this the Higgs mechanism allows the vacuum to emit or absorb a Higgs boson. The coupling of the W and Z bosons to the Higgs field effectively gives them mass, while because the photon and gluon cannot couple to it they remain mass-less.

2.4 $H \rightarrow ZZ \rightarrow l^+ l^- q \bar{q}$ Channel

In the high mass region for a Higgs boson discovery ($m_H > 200 \text{ GeV}$) the predominant Higgs decay is to vector boson pairs, $H \rightarrow WW$ and $H \rightarrow ZZ$. The signatures from the fully leptonic decay modes for these two processes are easily re-constructable and can be distinguished from the background processes. The fully leptonic decay modes are $H \rightarrow WW \rightarrow l^+ \nu l^- \bar{\nu}$ and $H \rightarrow ZZ \rightarrow l^+ l^- l^+ l^-$. The $H \rightarrow ZZ \rightarrow l^+ l^- l^+ l^-$ channel is particularly interesting because the decay chain can be fully reconstructed with a narrow invariant mass peak and almost no Standard Model background. The rate of decay of $Z \rightarrow l^+ l^-$ is only 3.37% [5] so for $l = e, \mu$ there is only a 0.45% chance of $ZZ \rightarrow l^+ l^- l^+ l^-$.

The branching ratio of $Z \rightarrow q \bar{q}$ is 69.9% [5], and has the largest branching ratio of any Z decay mode. Furthermore, $ZZ \rightarrow q \bar{q} q \bar{q}$ has a decay mode of just under 50%. The problem is that while this maximizes the branching ratio, this final state is indistinguishable from Standard Model background processes, like QCD.

The semi-leptonic final state in $H \rightarrow ZZ \rightarrow l^+l^-q\bar{q}$ produces a number of benefits over the previous two decay channels. The branching ratio of $ZZ \rightarrow l^+l^-q\bar{q}$ is 9.4%, or more than 20 times as large as $ZZ \rightarrow l^+l^-l^+l^-$. In comparison to $ZZ \rightarrow q\bar{q}q\bar{q}$ the leptonic decay of one of the Z bosons significantly reduces the Standard Model background in the final state.

Some of the difficulties in this channel compared to the fully leptonic decay are increased difficulty in event selection, the final reconstruction efficiency is approximately half of the fully leptonic final state, the backgrounds are significantly higher and dominated by Z+jets final states, and the resolution is worse. Despite these difficulties, this channel has improved sensitivity over the fully leptonic final state at high masses because of the higher branching ratio that the semi-leptonic final state enjoys and at high masses the Z+jets background is not as significant.

2.5 Background Processes

The background processes for this channel are processes which have a pair of opposite signed leptons with high transverse momentum that are associated with two jets. The main backgrounds are listed in Table 2.3.

Table 2.3

The main Standard Model background processes for the $H \rightarrow ZZ \rightarrow l^+l^-q\bar{q}$ channel and their cross sections at 8 TeV.

Background Process	Cross Section [pb] at 8 TeV
$Z + jets$	3503.71
$t\bar{t}$	225.197
ZZ	17.654
WZ	22.88
WW	57.1097

The dominant background in the $H \rightarrow ZZ \rightarrow l^+l^-q\bar{q}$ analysis is the Z+jets background, or more specifically, the inclusive Z production with QCD jets. The cross section of Z production at the LHC is more than 10^4 larger than the Higgs signal.

Events containing top quarks are another significant source of background. There are two top processes that result in the same final state as the signal we are studying. They include $t\bar{t}$ pair production and top quark associated production with a W boson.

$$t\bar{t} \rightarrow (W^+ \rightarrow \ell^+\nu)b(W^- \rightarrow \ell^-\bar{\nu})\bar{b}$$

3. Experimental Apparatus

This chapter describes the LHC and the CMS detector.

3.1 The Large Hadron Collider

The LHC accelerator is 26.7 km long and occupies the tunnel where the LEP collider was originally built. This tunnel is located approximately 100 meters underground and spans the French and Swiss borders near Geneva, Switzerland. It is a proton-proton collider. The LHC is the world's largest and most powerful particle accelerator. [6]

The LHC injection starts with the Linac2 to bring protons to 50 MeV. After the Proton Synchrotron (PS) accelerates them to 1.4 GeV, the Super Proton Synchrotron (SPS) then injects the protons into the LHC at an energy of 450 GeV. To finish, the LHC increases the energy 0.5 MeV per cycle until they are at 4.0 TeV. This can be seen schematically in Figure 3.1. The LHC can also accelerate lead ions.

Because the protons being collided have the same electric charge, there are two vacuum chambers for acceleration with two different magnetic fields. There are 1232 superconducting Niobium-Titanium magnets. Each one is 14.2 meters long and is cooled to 1.9 K with liquid Helium. The magnetic field produced is approximately 8.3 Tesla. These magnets are placed in 8 separate curved sections.

There are 4 interaction points in the LHC. Two are high luminosity points at the general purpose experiments of A Toroidal LHC Apparatus (ATLAS) [7] and CMS. The other two are for the A Large Ion Collider Experiment (ALICE) [8] and LHCb [9] experiments.

On November 23, 2009, the first proton-proton collisions were generated. After a few runs at 450 GeV and 1.18 TeV beam energies, on March 30, 2010 the first 7

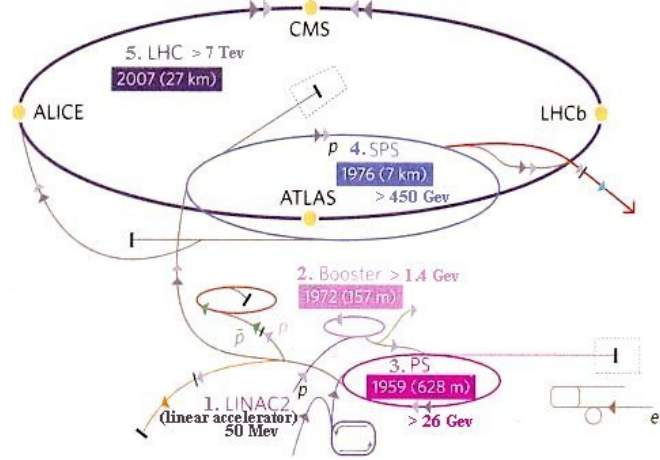


Figure 3.1. LHC injection scheme.

TeV center-of-mass energies were created. This was the highest ever reached at a man made particle collider. At 7 TeV approximately 47 pb^{-1} of integrated luminosity were delivered in 2010 (see Figure 3.2). The maximum instantaneous luminosity was $2 \times 10^{32} \text{ cm}^{-2} \text{ s}^{-1}$

During 2011 the machine was only offline for short amounts of time. On October 26, 2011 the highest ever instantaneous luminosity was reached with a peak value of $3.5 \times 10^{33} \text{ cm}^{-2} \text{ s}^{-1}$. This was 1331 bunches per beam with collisions every 50 ns. The total integrated luminosity delivered was 5.73 fb^{-1} (see Figure 3.3). In 2012 the beam energy was increased to 4 TeV and the LHC has continued to perform amazingly (see Figure 3.4).

In the coming years the LHC will continue to increase it's energy and instantaneous luminosity. Eventually it will reach the designed energy of 14 TeV collisions and an instantaneous luminosity of $10^{34} \text{ cm}^{-2} \text{ s}^{-1}$. This is seven times the energy of the highest Tevatron energy and two orders of magnitude more luminosity of any previous experiment.

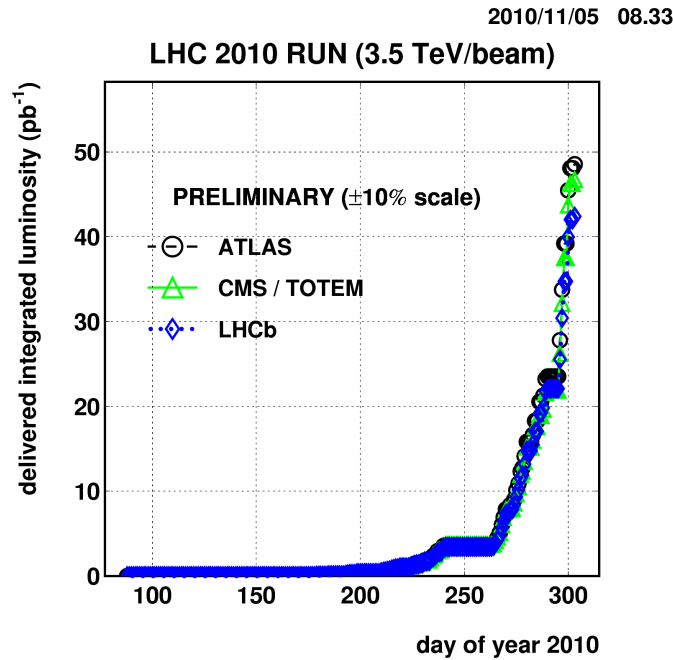


Figure 3.2. Integrated luminosity delivered by the LHC to each experiment in 2010

3.2 The Compact Muon Solenoid

As one of the two general purpose detectors at the LHC, the Compact Muon Solenoid (CMS) has a wide range of physics goals. These include but are not limited to: Higgs boson searches, physics beyond the SM, and precision measurement of known physics processes. The main parts of the CMS detector include the 3.8 T superconducting solenoid, muon system, electromagnetic calorimeter, and the tracking system [10]. The CMS has a number of cylindrical layers which are perpendicular to the beam axis (these are referred to as the barrel), and at both ends there are detector disks which are orthogonal to the beam axis (referred to as the end-caps). A more detailed picture of the detector can be found in Figures 3.5 and 3.6.

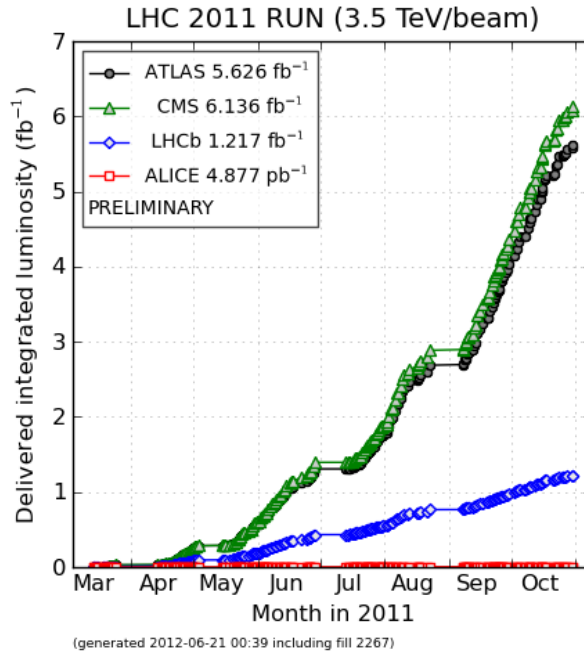


Figure 3.3. Integrated luminosity delivered by the LHC to each experiment in 2011

3.3 The CMS Trigger System

At the high luminosity achieved at the LHC, there is the drawback that several of the interactions overlap in the same bunch crossing. Also there is overlap from different bunch crossing because of the limited speed of the detector response and data read-out. These effects are known as pile-up. With the current luminosity, approximately 10 reconstructed primary vertexes per collision are observed at CMS. In addition to pile-up difficulties there is an extreme amount of data produced from the collisions at CMS. There are technical difficulties in the handling, processing, and storing of this data. These difficulties make the real time selection and recording of important events (the Trigger) very important.

The CMS Trigger has two levels. The Level 1 (L1) trigger is a hardware trigger. The full data is stored in the pipelines for processing while waiting for the trigger decision. If the L1 accepts the decision then the data is moved to the software based

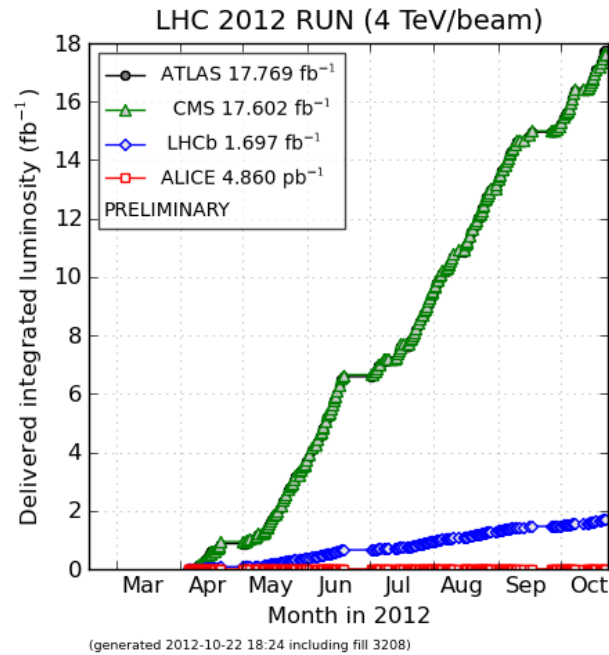


Figure 3.4. Integrated luminosity delivered by the LHC to each experiment in 2012

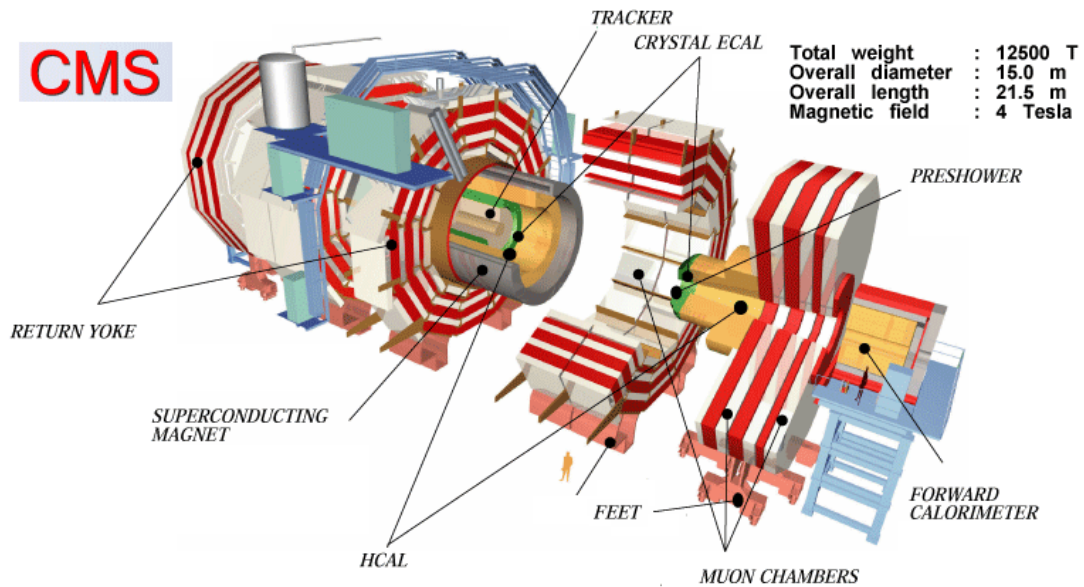


Figure 3.5. The CMS experiment. HCAL stands for hadron calorimeter and ECAL stands for Electromagnetic calorimeter.

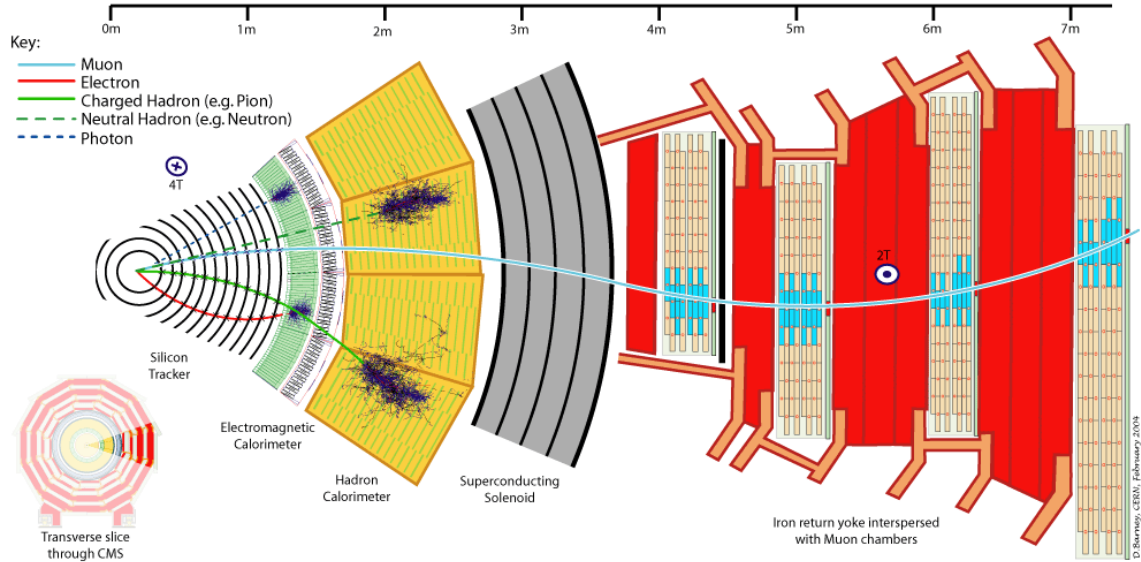


Figure 3.6. A graphical slice of the CMS experiment

High Level Trigger (HLT). The HLT reduces the output rate to around 100 Hz. The HLT software has a set of algorithms which give complete freedom to deciding which data to access. To free up processing there are three virtual trigger levels. The first only uses the muon and calorimeter data, the second adds the pixel seeds, and the final step uses the full event information.

4. Event Selection

This Chapter outlines the event selection that we use for the analysis. A list and brief description is given of both the data and the Monte Carlo generated events. Also the cuts and motivation for the analysis preselection is given.

4.1 Datasets

4.1.1 Data

The focus of this analysis is on a massive Higgs boson which is above the ZZ production threshold of 200 GeV. Because of this high energy the decay of one of the Z bosons will produce a pair of high p_T leptons. This will primarily fire the double-lepton HLT paths and be stored in one of the two Primary Datasets: DoubleElectron or DoubleMuon.

For 2011, we are using the double lepton primary datasets corresponding to 4.6 fb^{-1} of data collected in 2011 at 7 TeV. The data samples for 2011 are listed in Table 4.1. For 2012 we are using the double lepton primary datasets corresponding to 5 fb^{-1} at 8TeV. The data samples for 2012 are listed in Table 4.2.

4.1.2 Monte Carlo

The dominant background in this analysis is the inclusive Z production with jets. Particularly those jets coming from b-quarks. The Z+Jets sample uses the MADGRAPH generator [11]. Other backgrounds are $t\bar{t}$, ZZ, WW and WZ. The Monte Carlo samples for these processes are listed in Table 4.4 for 2012 and Table 4.3 for 2011.

Table 4.1
2011 data samples.

Dataset Name
/DoubleMu/Run2011A-May10ReReco-v1/AOD
/DoubleMu/Run2011A-PromptReco-v4/AOD
/DoubleMu/Run2011A-05Aug2011-v1/AOD
/DoubleMu/Run2011A-PromptReco-v6/AOD
/DoubleMu/Run2011B-PromptReco-v1/AOD
/DoubleElectron/Run2011A-May10ReReco-v1/AOD
/DoubleElectron/Run2011A-PromptReco-v4/AOD
/DoubleElectron/Run2011A-05Aug2011-v1/AOD
/DoubleElectron/Run2011A-PromptReco-v6/AOD
/DoubleElectron/Run2011B-PromptReco-v1/AOD

The signal Monte Carlo samples used in this analysis were generated using POWHEG [12–14]. Currently only samples with gluon fusion production mechanism are considered, but vector boson fusion may be considered in the future. These samples can be seen in Table 4.5 for 2011 and Table 4.5 for 2012.

4.2 Preselection

The analysis is over events in the DoubleMu and DoubleElectron datasets. In each of these datasets there is at least one un-prescaled trigger with looser requirements than the offline selections. Only events which satisfy the lowest threshold un-prescaled trigger for the dataset are considered for the analysis. The trigger requirements are summarized in Table 4.7. Since the level of precision of the trigger emulation in simulation is not well known, no trigger is applied on MC samples. Instead, proper event weights are assigned to MC events according to the probabilities of lepton

Table 4.2
2012 data samples.

Dataset Name	Luminosity [pb^{-1}]
/DoubleMu/Run2012A-13Jul2012-v1/AOD	808
/DoubleMu/Run2012A-recover-06Aug2012-v1	82
/DoubleMu/Run2012B-13Jul2012-v4/AOD	4429
/DoubleMu/Run2012C-24Aug2012-v1/AOD	495
/DoubleMu/Run2012C-PromptReco-v2/AOD	6394
/DoubleMu/Run2012D-PromptReco-v1/AOD	4580
/DoubleElectron/Run2012A-13Jul2012-v1/AOD	808
/DoubleElectron/Run2012A-recover-06Aug2012-v1	82
/DoubleElectron/Run2012B-13Jul2012-v4/AOD	4429
/DoubleElectron/Run2012C-24Aug2012-v1/AOD	495
/DoubleElectron/Run2012C-PromptReco-v2/AOD	6394
/DoubleElectron/Run2012D-PromptReco-v1/AOD	4580

Table 4.3
2011 background MC samples.

Dataset Name	Luminosity [pb^{-1}]
DYJetsToLL_TuneZ2_M-50.7TeV-madgraph-tauola	36277961
TTJets_TuneZ2_7TeV-madgraph-tauola	3701947
WW_TuneZ2_7TeV_pythia6-tauola	4225916
WZ_TuneZ2_7TeV_pythia6-tauola	4265243
ZZ_TuneZ2_7TeV_pythia6-tauola	4187885
ZZJetsTo2L2Q_TuneZ2_7TeV-madgraph-tauola	1013369

Table 4.4
2012 background MC samples.

Dataset Name	Number of Events
DYJetsToLL_M-50_TuneZ2Star_8TeV-madgraph-tarball	30461028
TTJets_TuneZ2star_8TeV-madgraph-tauola	6736135
WW_TuneZ2star_8TeV_pythia6_tauola	10000431
WZ_TuneZ2star_8TeV_pythia6_tauola	10000283
ZZ_TuneZ2star_8TeV_pythia6_tauola	9799908

candidates to pass the trigger. The trigger efficiency tables for leptons satisfying the same identification criteria as in the analysis are computed in bins of (pt, η) from data using tag & probe techniques [15].

4.2.1 Lepton Selection

$Z \rightarrow ee$ and $Z \rightarrow \mu\mu$ candidates are constructed from pairs of same-flavor, opposite-charge lepton candidates, which satisfy kinematic and identification criteria. Electron candidates are reconstructed with the GSF algorithm and in order to assure good electron reconstruction a cut is applied: the η of the electron super-cluster must be inside the ECAL acceptance volume ($|\eta| < 2.5$) but outside the ECAL barrel-end-cap overlap region ($1.4442 < |\eta| < 1.566$). Muon candidates must have been reconstructed by both the GlobalMuon and the PF muon reconstruction algorithms and must satisfy the acceptance cut $|\eta| < 2.4$.

Electron candidates must satisfy the standard “Loose” working point of the cut-based electron ID for 2012 analysis. The cuts are listed in Table 4.8 and comprise proper electron identification requirements, an isolation cut, and conversion rejection criteria. Muon candidates must satisfy the standard “Tight” working point of the cut-based muon ID for 2012 analysis. The cuts are listed in Table 4.9 and comprise proper muon identification requirements plus an isolation cut.

Table 4.5
2011 signal MC samples.

Dataset Name	σ NLO [fb]	Luminosity [fb^{-1}]
GluGluToHToZZTo2L2Q_M-130.7TeV-powheg-pythia6	86.31	3436.46
GluGluToHToZZTo2L2Q_M-150.7TeV-powheg-pythia6	133.57	2143.72
GluGluToHToZZTo2L2Q_M-190.7TeV-powheg-pythia6	194.04	1467.4
GluGluToHToZZTo2L2Q_M-200.7TeV-powheg-pythia6	211.51	1346.25
GluGluToHToZZTo2L2Q_M-210.7TeV-powheg-pythia6	205.02	1384.06
GluGluToHToZZTo2L2Q_M-220.7TeV-powheg-pythia6	193.3	1451.42
GluGluToHToZZTo2L2Q_M-230.7TeV-powheg-pythia6	179.56	1621.86
GluGluToHToZZTo2L2Q_M-250.7TeV-powheg-pythia6	156.63	1915.27
GluGluToHToZZTo2L2Q_M-275.7TeV-powheg-pythia6	134.23	2193.46
GluGluToHToZZTo2L2Q_M-300.7TeV-powheg-pythia6	117.81	2514.96
GluGluToHToZZTo2L2Q_M-325.7TeV-powheg-pythia6	108.12	2624.88
GluGluToHToZZTo2L2Q_M-350.7TeV-powheg-pythia6	108.99	2699.44
GluGluToHToZZTo2L2Q_M-375.7TeV-powheg-pythia6	102.4	2779.99
GluGluToHToZZTo2L2Q_M-400.7TeV-powheg-pythia6	83.17	3606.95
GluGluToHToZZTo2L2Q_M-425.7TeV-powheg-pythia6	70.46	4139.36
GluGluToHToZZTo2L2Q_M-450.7TeV-powheg-pythia6	49.42	5837.91
GluGluToHToZZTo2L2Q_M-475.7TeV-powheg-pythia6	41.29	7136.59
GluGluToHToZZTo2L2Q_M-500.7TeV-powheg-pythia6	34.72	8473.64
GluGluToHToZZTo2L2Q_M-525.7TeV-powheg-pythia6	29.03	9872.98
GluGluToHToZZTo2L2Q_M-550.7TeV-powheg-pythia6	24.48	11950.92
GluGluToHToZZTo2L2Q_M-575.7TeV-powheg-pythia6	17.42	17214.3
GluGluToHToZZTo2L2Q_M-600.7TeV-powheg-pythia6	14.73	19637.09

Given a lepton candidate, the PF isolation is defined as the sum of “isodeposits” (i.e. p_T or E_T) of charged hadrons (I_{ch}), neutral hadrons (I_{nh}), and photons (I_{ph}), computed in a ΔR cone around the lepton direction. In order to assure independence

of the isolation from the number of PU interactions, a corrected PF isolation definition is used as follows:

$$I_{PF,corr} = I_{ch}(PFnoPU) + \max(I_{nh} + I_{ph} - \rho \cdot A_{eff}, 0) \quad (4.1)$$

In the above corrected definition only isodeposits from charged hadrons not coming from PU vertexes ($I_{ch}(PFnoPU)$) are considered, and then subtract an overall PU energy contribution which is estimated as the average energy density in the event (ρ) multiplied by an effective area A_{eff} . This is strictly following the recommendations of egamma and muon POGs. The isolation cone for electrons is defined as $\Delta R < 0.3$, while for muons is defined as $\Delta R < 0.4$. PF muon isolation employs for PU correction the energy density ρ' computed in the central region of the detector after charged PU subtraction “RhoCentralNeutral”. The proper A_{eff} values are provided by the POGs in bins of lepton η . A cut is applied on the relative PF isolation ($I_{PF,corr}/p_T$) as reported in Tables 4.8 and 4.9.

$Z \rightarrow ee$ and $Z \rightarrow \mu\mu$ candidates are constructed from pairs of opposite-charge leptons. The leading lepton of the pair must have $p_T > 40GeV$, while the next-to-leading lepton must have $p_T > 20GeV$. The invariant mass of the pair must be $70 < m_{\ell\ell} < 110GeV$. The di-lepton invariant mass for the selected $Z \rightarrow \ell\ell$ candidates is shown in Fig. 4.1.

4.2.2 Jet Selection

The PF jets are reconstructed with the **anti- k_T** algorithm [16] with radius parameter set to $R = 0.5$. Jets are required to be inside the tracker acceptance ($|\eta| < 2.4$) thus allowing high reconstruction efficiency and precise energy measurements using PF techniques. Jet-energy corrections are applied to data and MC [17]. Correction for PU energy is applied at the L1-correction level, by using Fastjet algorithm. In order to remove jets which originate from PU interactions, only jets with $\beta \geq 0.2$ are selected, where β is defined as the sum of transverse momenta of all charged particles in the jet coming from the primary vertex, normalized to the total sum of transverse

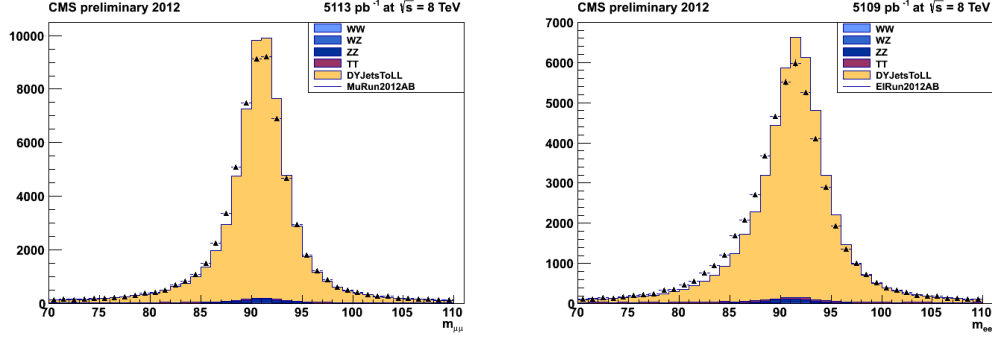


Figure 4.1. Di-lepton invariant mass in data and MC of $Z \rightarrow \ell\ell$ candidates after lepton selection. Left: Muon channel. Right: Electron channel. The selection described in preselection is applied except for the cut on m_{jj} .

momenta of all charged particles in the jet. $Z \rightarrow q\bar{q}$ candidates are reconstructed from jet-jet pairs. In order to reject fake candidates made by low- p_T jets from QCD background, both jets of the pair must have $p_T > 30 \text{ GeV}$.

4.2.3 Higgs candidates reconstruction

All the di-lepton and di-jet pairs satisfying the above requirements are combined to form $\ell\ell jj$ candidates. In addition, a $\Delta R > 0.5$ cut is applied between each lepton and jet within a candidate in order to avoid double counting of the same object reconstructed in different collections (for instance leptons inside a jet). In the following, the entire selection procedure described above is referred to as “pre-selection” of $\ell\ell jj$ candidates. The distribution of the di-jet invariant mass m_{jj} for $\ell\ell jj$ candidates reconstructed in the electron and muon channels is shown in Fig. 4.2. In order to suppress the dominant Z +jets background, only candidates in a signal region of m_{jj} , defined as $75 \text{ GeV} < m_{jj} < 105 \text{ GeV}$, are selected.

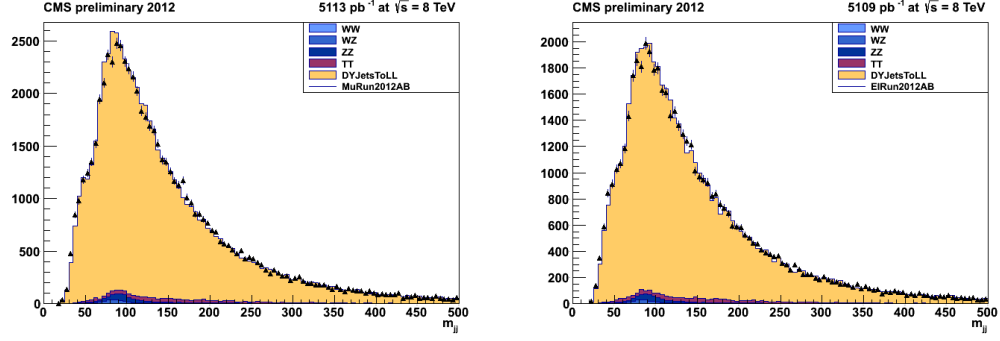


Figure 4.2. Di-jet invariant mass in data and MC for $\ell\ell jj$ candidates which have passed the pre-selection. Left: Muon channel. Right: Electron channel.

4.2.4 Kinematic fit

Finite resolution of the jet energy is the dominant source of uncertainty in both the di-jet invariant mass m_{jj} and di-boson invariant mass m_{ZZ} for Higgs candidates. Therefore, the two variables become highly correlated, as can be seen in results from 2011 shown in Fig. 4.3 (left). In order to optimally scale the di-jet quadrimomentum to the Z boson mass, a kinematic fit to the two jets is used [18,19]. The fit is provided with parametrization of jet transverse momentum and angular resolutions as functions of transverse momentum and pseudo-rapidity, and therefore constrains the mass of the di-jet system to the value of the Z boson mass by modifying the jet four-momenta in accordance to their expected resolutions. This brings a further improvement in the resolution on the signal invariant mass [18,19]. The kinematic fit to the di-jet system also removes the correlation between the di-jet and di-boson invariant mass in signal as can be seen in Fig. 4.3 (right). This allows a straightforward definition of signal and sideband regions, through simple addition of one-variable interval cuts.

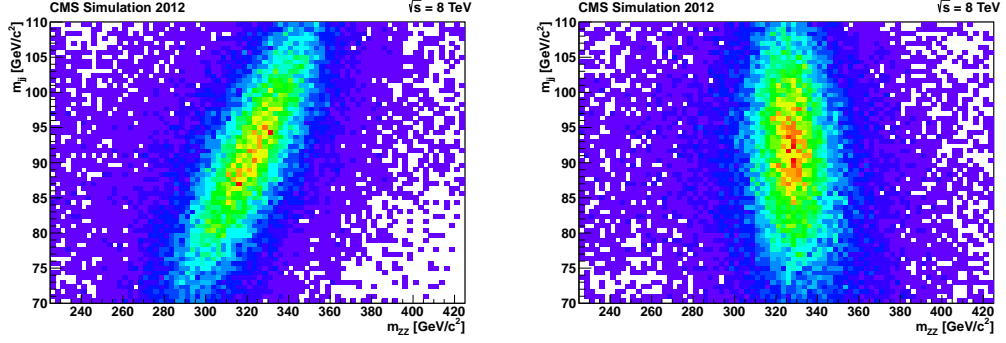


Figure 4.3. Kinematic fit results. Di-jet invariant mass vs. di-boson invariant mass for Higgs candidates (signal MC, $m_H = 325$ GeV) after loose selection requirements. Left: before kinematic fit; right: after kinematic fit.

4.2.5 B-tagging of jets

Due to the relatively large branching fraction of the Z -bosons decaying into a pair of bottom-anti-bottom quarks, compared to the abundance of light-quark or gluon jets in Z +jets background events, we use a b-tagging algorithm in order to identify jets originating from heavy-flavor quarks. However, no selection of candidates based on b-probabilities is employed in the analysis. Instead, the b-tagging information is used to classify $\ell\ell jj$ candidates into three categories, each one characterized by a certain signal-to-background ratio, and an optimized selection is applied on candidates belonging to different categories [15]. This exploits as much information as possible from the data increasing the analysis sensitivity.

Table 4.6
POWHEG signal MC samples $H \rightarrow ZZ \rightarrow 2l2q$, $l = e, \mu, \tau$ 2012.

M_H (GeV)	Name	$\sigma(H \rightarrow 2l2q)$ [pb]
200	/GluGluToHToZZTo2L2Q_M-200.8TeV-powheg-pythia6/ Summer12-PU_S7_START52_V9-v1/AODSIM	0.2566
210	/GluGluToHToZZTo2L2Q_M-210.8TeV-powheg-pythia6/ Summer12-PU_S7_START52_V9-v1/AODSIM	0.2538
220	/GluGluToHToZZTo2L2Q_M-220.8TeV-powheg-pythia6/ Summer12-PU_S7_START52_V9-v1/AODSIM	0.2416
230	/GluGluToHToZZTo2L2Q_M-230.8TeV-powheg-pythia6/ Summer12-PU_S7_START52_V9-v1/AODSIM	0.2278
250	/GluGluToHToZZTo2L2Q_M-250.8TeV-powheg-pythia6/ Summer12-PU_S7_START52_V9-v1/AODSIM	0.2022
275	/GluGluToHToZZTo2L2Q_M-275.8TeV-powheg-pythia6/ Summer12-PU_S7_START52_V9-v1/AODSIM	0.1751
300	/GluGluToHToZZTo2L2Q_M-300.8TeV-powheg-pythia6/ Summer12-PU_S7_START52_V9-v1/AODSIM	0.1563
325	/GluGluToHToZZTo2L2Q_M-325.8TeV-powheg-pythia6/ Summer12-PU_S7_START52_V9-v1/AODSIM	0.1478
350	/GluGluToHToZZTo2L2Q_M-350.8TeV-powheg-pythia6/ Summer12-PU_S7_START52_V9-v1/AODSIM	0.1482
375	/GluGluToHToZZTo2L2Q_M-375.8TeV-powheg-pythia6/ Summer12-PU_S7_START52_V9-v1/AODSIM	0.1360
400	/GluGluToHToZZTo2L2Q_M-400.8TeV-powheg-pythia6/ Summer12-PU_S7_START52_V9-v1/AODSIM	0.1111
425	/GluGluToHToZZTo2L2Q_M-425.8TeV-powheg-pythia6/ Summer12-PU_S7_START52_V9-v1/AODSIM	0.0914
450	/GluGluToHToZZTo2L2Q_M-450.8TeV-powheg-pythia6/ Summer12-PU_S7_START52_V9-v1/AODSIM	0.7311
475	/GluGluToHToZZTo2L2Q_M-475.8TeV-powheg-pythia6/ Summer12-PU_S7_START52_V9-v1/AODSIM	0.6
500	/GluGluToHToZZTo2L2Q_M-500.8TeV-powheg-pythia6/ Summer12-PU_S7_START52_V9-v1/AODSIM	0.4719
525	/GluGluToHToZZTo2L2Q_M-525.8TeV-powheg-pythia6/ Summer12-PU_S7_START52_V9-v1/AODSIM	0.0380
550	/GluGluToHToZZTo2L2Q_M-550.8TeV-powheg-pythia6/ Summer12-PU_S7_START52_V9-v1/AODSIM	0.0305
575	/GluGluToHToZZTo2L2Q_M-575.8TeV-powheg-pythia6/ Summer12-PU_S7_START52_V9-v1/AODSIM	0.025
600	/GluGluToHToZZTo2L2Q_M-600.8TeV-powheg-pythia6/ Summer12-PU_S7_START52_V9-v1/AODSIM	0.0201

Table 4.7
Event trigger requirements for data.

Dataset	trigger requirement
DoubleMu	HLT_Mu17_Mu8 HLT_Mu17_TkMu8
DoubleElectron	HLT_Ele17_CaloIdT_TrkIdVL_CaloIsoVL_TrkIsoVL_ Ele8_CaloIdT_TrkIdVL_CaloIsoVL_TrkIsoVL

Table 4.8
Electron ID requirements for the Loose ID working point.

Variable	Barrel cut	Endcap cut
$\Delta\eta_{trk,supercluster}$	< 0.007	< 0.009
$\Delta\phi_{trk,supercluster}$	< 0.15	< 0.1
$\sigma_{i\eta,i\eta}$	< 0.01	< 0.03
H/E	< 0.12	< 0.10
d_0 (wrt primary vertex)	$< 0.2mm$	$< 0.2mm$
d_z (wrt primary vertex)	$< 2mm$	$< 2mm$
$ 1/E - 1/p $	< 0.05	< 0.05
$I_{PF,corr}/p_T$	< 0.15	< 0.15
Missing hits	≤ 1	≤ 1
Conversion vertex fit prob.	$< 10^{-6}$	$< 10^{-6}$

Table 4.9
 Muon ID requirements for the Tight ID working point.

Variable	Cut
isGlobalMuon	True
isPFMuon	True
$\chi^2/ndof$ (global fit)	< 10
Muon chamber hits in global fit	> 0
Muon stations with muon segments	> 1
d_{xy} (from tracker, wrt primary vertex)	$< 2mm$
d_z (from tracker, wrt primary vertex)	$< 5mm$
Valid pixel hits (tracker track)	> 0
Tracker layers with hits	> 5
$I_{PF, corr}/p_T$	< 0.12

5. Correlated and Uncorrelated Systematic Uncertainties

5.1 Trigger and Lepton Selection Efficiency

To calculate the lepton efficiency we use the tag and probe method. This method is a way to use a data sample of pure leptons from Z boson decays to find efficiencies. We require our events to contain a minimum of two leptons from a Z decay as well as a minimum of one jet. This matches the event topology that we use in the analysis. The requirements for tag and probe events are shown in Table 5.2. After measuring efficiency in data by applying the same method to measure simulation efficiency, we are able to calculate the scale factors for data to simulation. The complete efficiency measurement can be split into five relative measurements: tracking efficiency, reconstruction efficiency, identification efficiency, isolation efficiency, and the total trigger efficiency. This is expressed in equation (5.1).

$$\epsilon_{lepton} = \epsilon_{tracking} \times \epsilon_{RECO/tracking} \times \epsilon_{ID/RECO} \times \epsilon_{ISO/ID} \times \epsilon_{trigger/ISO} \quad (5.1)$$

Each term in the total efficiency is calculated separately and is found using the official tag-and-probe method. This is a well established efficiency extraction technique. The efficiencies are measured in p_T and η bins. The trigger used for the 2012 data is the HLT_Ele17_CaloIdVT_CaloIsoVT_TrkIdT_TrkIsoVT_Ele8_SC8_Mass50. The trigger needs to be selected to yield unbiased efficiencies for each type. Also the cuts for the L1 and HLT triggers must be looser than those being measured. The tag is required to pass the WP Medium which is given in tables 5.3 and 5.4. Also the tag is spatially matched to a triggering object and has an energy above 32 GeV. For the tag-probe pair it is required that there is opposite charge for the tag and probe and that the event is within the window of 60 - 120 GeV. When there are multiple tag-probe pairs all of the pairs are used in order to avoid biasing the measurement. $\epsilon_{tracking}$ is assumed to be essentially 100% [20].

For ϵ_{RECO} the probes are super clusters with the standard η cuts and an energy above 10 GeV. At least one jet has to have a $p_T > 5$ GeV and the Hadronic energy fraction must be greater than 0.15. If there are photons in the events the requirements are $hadronicOverEm < 0.15$, standard η cuts, for the barrel $sigmaIetaIeta < 0.01$ while for end cap $sigmaIetaIeta < 0.03$, and $superCluster.energy * \sin(superCluster.position.theta) > 20$ GeV. For the electrons, passing probes are those that are reconstructed with the GSF algorithm giving a GSF electron. When calculating these efficiencies a fit is used as seen in Figure 5.1. The fit range is 60 -120 GeV. The functional form of the signal PDF is formed from a Breit-Wigner convoluted with a “crystal ball” function. The background PDF is an exponential multiplied by an error function modeling the kinematic turn on of CMS. The values are reported in table 5.5.

ϵ_{ISO} is computed with the same fits as described for calculating ϵ_{RECO} , as well as the same tags. The fits can be seen in Figure 5.2. This time the probes are required to be GSF Electrons and the passing probes are those GSF Electrons that also pass WP Medium (as described in tables 5.3 and 5.4). For electrons, the electron identification efficiency is factorized together with isolation and the efficiencies are shown in table 5.6.

Counting passing vs failing probes is used instead of fitting because the backgrounds are predicted from Monte Carlo to be negligible. The tags are again the same as described before. The probes are those GSF Electrons that pass WP medium and the passing probes pass one the the legs of the HLT for the DoubleElectron dataset. For those electrons passing the HLT we calculate the efficiency in two parts. The “legs” and the corresponding filters are given in Table 5.1. These efficiencies are shown in Tables 5.7 and 5.8.

Table 5.1
The Legs for the Electron HLT.

Leg	Filter
Ele8	hltEle17TightIdLooseIsoEle8TightIdLooseIsoTrackIsoDoubleFilter
Ele17	hltEle17TightIdLooseIsoEle8TightIdLooseIsoTrackIsoFilter

Table 5.2
These are the event requirements for the tag and probe method of calculating trigger and lepton selection efficiency.

Lepton Energy	$> 10 \text{ GeV}$
gsfElectron	> 1
Super Cluster $ \eta $	< 2.5
ak5PFJet	> 1
Jet $ \eta $	< 2.4
Jet pt	$> 30 \text{ GeV}$

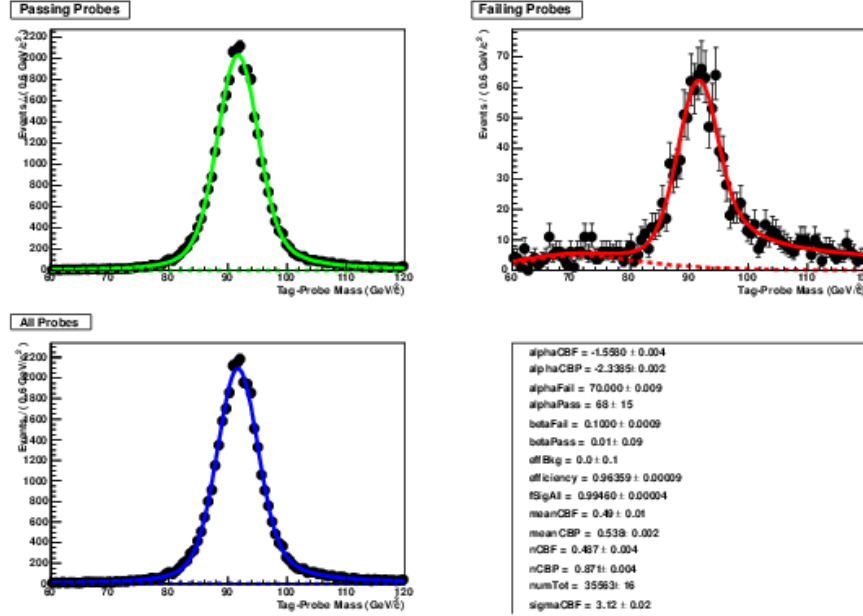


Figure 5.1. Fitting of Supercluster to GSF Electrons for ϵ_{RECO} calculation.

Table 5.3

These are the official working points provided by the EGamma POG for the barrel.

	Veto	Loose	Medium	Tight
dEtaIn	0.007	0.007	0.004	0.004
dPhiIn	0.8	0.15	0.06	0.03
sigmaIEtaIEta	0.01	0.01	0.01	0.01
H/E	0.15	0.12	0.12	0.12
d0 (vtx)	0.04	0.02	0.02	0.02
dZ (vtx)	0.2	0.2	0.1	0.1
fabs(1/E - 1/p)	N/A	0.05	0.05	0.05
PF isolation / pT	0.15	0.15	0.15	0.10
vertex fit probability	N/A	1e-6	1e-6	1e-6
missing hits	N/A	1	1	0

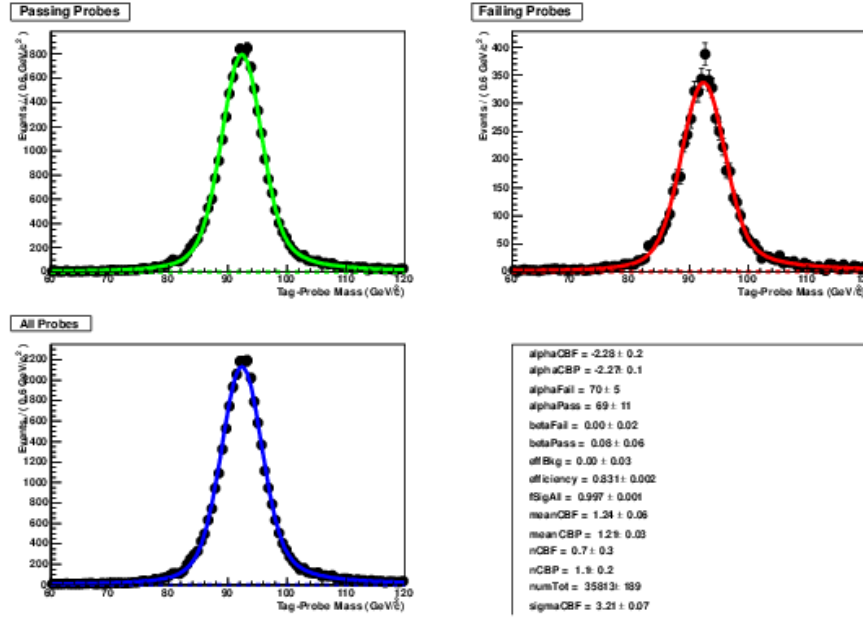


Figure 5.2. Fitting of GSF Electrons for to WP Medium for ϵ_{ISO} calculation.

Table 5.4

These are the official working points provided by the EGamma POG for the endcaps.

	Veto	Loose	Medium	Tight
dEtaIn	0.01	0.009	0.007	0.005
dPhiIn	0.7	0.10	0.03	0.02
sigmaIEtaIEta	0.03	0.03	0.03	0.03
H/E	N/A	0.10	0.10	0.10
d0 (vtx)	0.04	0.02	0.02	0.02
dZ (vtx)	0.2	0.2	0.1	0.1
fabs(1/E - 1/p)	N/A	0.05	0.05	0.05
PF isolation / pT	0.15	0.15(0.10)	0.15(0.10)	0.10(0.07)
vertex fit probability	N/A	1e-6	1e-6	1e-6
missing hits	N/A	1	1	0

Table 5.5
 ϵ_{RECO} efficiency values from 2012 data and Monte Carlo simulations
 using the Tag and Probe method by fitting the signal and background.

η coverage	p_T range (GeV)	Efficiency (data)	Efficiency (MC)	Data/MC ratio
$0.0 < \eta < 0.8$	$10 < p_T < 20$	$63.69\% \pm 0.77\%$	$68.23\% \pm 3.23\%$	0.933 ± 0.046
$0.8 < \eta < 1.4$	$10 < p_T < 20$	$72.56\% \pm 1.73\%$	$72.97\% \pm 1.02\%$	0.994 ± 0.028
$1.6 < \eta < 2.0$	$10 < p_T < 20$	$47.39\% \pm 10.46\%$	$52.68\% \pm 1.08\%$	0.899 ± 0.199
$2.0 < \eta < 2.5$	$10 < p_T < 20$	$43.49\% \pm 1.17\%$	$52.53\% \pm 1.34\%$	0.828 ± 0.031
$0.0 < \eta < 0.8$	$20 < p_T < 40$	$83.66\% \pm 8.23\%$	$85.06\% \pm 0.12\%$	0.983 ± 0.097
$0.8 < \eta < 1.4$	$20 < p_T < 40$	$84.02\% \pm 8.06\%$	$86.40\% \pm 1.06\%$	0.972 ± 0.094
$1.6 < \eta < 2.0$	$20 < p_T < 40$	$75.09\% \pm 0.34\%$	$77.06\% \pm 1.60\%$	0.974 ± 0.021
$2.0 < \eta < 2.5$	$20 < p_T < 40$	$68.29\% \pm 0.32\%$	$72.20\% \pm 0.31\%$	0.946 ± 0.006
$0.0 < \eta < 0.8$	$40 < p_T < 200$	$87.34\% \pm 0.12\%$	$88.26\% \pm 0.10\%$	0.990 ± 0.002
$0.8 < \eta < 1.4$	$40 < p_T < 200$	$88.73\% \pm 5.79\%$	$90.62\% \pm 0.11\%$	0.979 ± 0.064
$1.6 < \eta < 2.0$	$40 < p_T < 200$	$83.14\% \pm 0.22\%$	$83.48\% \pm 0.19\%$	0.996 ± 0.003
$2.0 < \eta < 2.5$	$40 < p_T < 200$	$77.31\% \pm 0.34\%$	$76.90\% \pm 0.24\%$	1.005 ± 0.005

Table 5.6
 $\epsilon_{ID} \times \epsilon_{ISO}$ efficiency values from 2012 data and Monte Carlo simulations using the Tag and Probe method by fitting the signal and background.

η coverage	p_T range (GeV)	Efficiency (data)	Efficiency (MC)	Data/MC ratio
$0.0 < \eta < 0.8$	$10 < p_T < 20$	$22.84\% \pm 10.13\%$	$54.53\% \pm 13.17\%$	0.419 ± 0.212
$0.8 < \eta < 1.4$	$10 < p_T < 20$	$37.14\% \pm 0.54\%$	$87.12\% \pm 1.68\%$	0.426 ± 0.010
$1.6 < \eta < 2.0$	$10 < p_T < 20$	$56.11\% \pm 0.98\%$	$93.84\% \pm 4.98\%$	0.598 ± 0.033
$2.0 < \eta < 2.5$	$10 < p_T < 20$	$56.39\% \pm 1.43\%$	$65.03\% \pm 1.81\%$	0.867 ± 0.033
$0.0 < \eta < 0.8$	$20 < p_T < 40$	$96.15\% \pm 0.44\%$	$95.88\% \pm 0.30\%$	1.003 ± 0.006
$0.8 < \eta < 1.4$	$20 < p_T < 40$	$94.67\% \pm 3.00\%$	$96.75\% \pm 1.88\%$	0.979 ± 0.036
$1.6 < \eta < 2.0$	$20 < p_T < 40$	$95.31\% \pm 0.79\%$	$96.29\% \pm 0.24\%$	0.990 ± 0.009
$2.0 < \eta < 2.5$	$20 < p_T < 40$	$95.02\% \pm 0.31\%$	$92.90\% \pm 3.64\%$	1.023 ± 0.040
$0.0 < \eta < 0.8$	$40 < p_T < 200$	$97.53\% \pm 0.16\%$	$98.22\% \pm 0.05\%$	0.993 ± 0.002
$0.8 < \eta < 1.4$	$40 < p_T < 200$	$98.08\% \pm 0.07\%$	$98.63\% \pm 0.05\%$	0.994 ± 0.001
$1.6 < \eta < 2.0$	$40 < p_T < 200$	$96.36\% \pm 0.11\%$	$97.47\% \pm 1.31\%$	0.989 ± 0.013
$2.0 < \eta < 2.5$	$40 < p_T < 200$	$95.61\% \pm 0.30\%$	$95.67\% \pm 0.12\%$	0.999 ± 0.003

Table 5.7

$\epsilon_{trigger}$ values for Ele8 Leg from 2012 data and Monte Carlo simulations using the Tag and Probe method by counting passing and failing probes.

η coverage	p_T range (GeV)	Efficiency (data)	Efficiency (MC)	Data/MC ratio
$0.0 < \eta < 0.8$	$10 < p_T < 20$	$96.03\% \pm 0.37\%$	$96.62\% \pm 0.34\%$	0.994 ± 0.005
$0.8 < \eta < 1.4$	$10 < p_T < 20$	$85.50\% \pm 0.64\%$	$88.63\% \pm 0.56\%$	0.965 ± 0.009
$1.6 < \eta < 2.0$	$10 < p_T < 20$	$84.15\% \pm 0.99\%$	$83.76\% \pm 1.02\%$	1.005 ± 0.017
$2.0 < \eta < 2.5$	$10 < p_T < 20$	$84.80\% \pm 1.04\%$	$88.53\% \pm 0.98\%$	0.958 ± 0.016
$0.0 < \eta < 0.8$	$20 < p_T < 40$	$98.19\% \pm 0.05\%$	$98.42\% \pm 0.04\%$	0.998 ± 0.001
$0.8 < \eta < 1.4$	$20 < p_T < 40$	$92.99\% \pm 0.11\%$	$94.14\% \pm 0.10\%$	0.988 ± 0.002
$1.6 < \eta < 2.0$	$20 < p_T < 40$	$89.07\% \pm 0.20\%$	$89.47\% \pm 0.20\%$	0.996 ± 0.003
$2.0 < \eta < 2.5$	$20 < p_T < 40$	$92.59\% \pm 0.19\%$	$92.53\% \pm 0.20\%$	1.001 ± 0.003
$0.0 < \eta < 0.8$	$40 < p_T < 200$	$98.81\% \pm 0.04\%$	$99.11\% \pm 0.03\%$	0.997 ± 0.001
$0.8 < \eta < 1.4$	$40 < p_T < 200$	$96.92\% \pm 0.07\%$	$97.42\% \pm 0.06\%$	0.995 ± 0.001
$1.6 < \eta < 2.0$	$40 < p_T < 200$	$93.22\% \pm 0.15\%$	$93.73\% \pm 0.13\%$	0.995 ± 0.002
$2.0 < \eta < 2.5$	$40 < p_T < 200$	$94.74\% \pm 0.15\%$	$95.04\% \pm 0.14\%$	0.997 ± 0.002

Table 5.8
 $\epsilon_{trigger}$ values for Ele17 Leg from 2012 data and Monte Carlo simulations using the Tag and Probe method by counting passing and failing probes.

η coverage	p_T range (GeV)	Efficiency (data)	Efficiency (MC)	Data/MC ratio
$0.0 < \eta < 0.8$	$10 < p_T < 20$	$48.80\% \pm 0.92\%$	$57.97\% \pm 0.91\%$	0.842 ± 0.021
$0.8 < \eta < 1.4$	$10 < p_T < 20$	$33.01\% \pm 0.85\%$	$46.75\% \pm 0.87\%$	0.706 ± 0.022
$1.6 < \eta < 2.0$	$10 < p_T < 20$	$47.63\% \pm 1.34\%$	$47.23\% \pm 1.37\%$	1.008 ± 0.041
$2.0 < \eta < 2.5$	$10 < p_T < 20$	$40.66\% \pm 1.41\%$	$50.52\% \pm 1.52\%$	0.805 ± 0.037
$0.0 < \eta < 0.8$	$20 < p_T < 40$	$98.49\% \pm 0.04\%$	$98.86\% \pm 0.04\%$	0.996 ± 0.001
$0.8 < \eta < 1.4$	$20 < p_T < 40$	$93.45\% \pm 0.11\%$	$94.62\% \pm 0.10\%$	0.988 ± 0.002
$1.6 < \eta < 2.0$	$20 < p_T < 40$	$89.71\% \pm 0.19\%$	$89.98\% \pm 0.19\%$	0.997 ± 0.003
$2.0 < \eta < 2.5$	$20 < p_T < 40$	$93.67\% \pm 0.18\%$	$93.59\% \pm 0.18\%$	1.001 ± 0.003
$0.0 < \eta < 0.8$	$40 < p_T < 200$	$99.05\% \pm 0.03\%$	$99.41\% \pm 0.03\%$	0.996 ± 0.000
$0.8 < \eta < 1.4$	$40 < p_T < 200$	$97.41\% \pm 0.07\%$	$97.92\% \pm 0.06\%$	0.995 ± 0.001
$1.6 < \eta < 2.0$	$40 < p_T < 200$	$93.90\% \pm 0.14\%$	$94.38\% \pm 0.13\%$	0.995 ± 0.002
$2.0 < \eta < 2.5$	$40 < p_T < 200$	$95.93\% \pm 0.13\%$	$96.17\% \pm 0.12\%$	0.998 ± 0.002

6. Signal Region Optimization

After the basic selection and the specific selections aiming to enhance the sensitivity for events with b jets and for VBF production, a final selection is performed to optimize the discrimination power for a Higgs signal.

6.1 Tagging Classification

Once the events are preselected as described in the previous sections, they are classified according to the tagging-content of the jets associated to the Higgs candidate. Events are classified as “two tags” if the two jets are b-tagged with at least one medium-tag and one loose-tag. Events failing that requirement are classified as “one tag” if at least one of the two jets satisfy the loose-tag condition. Events failing this requirement are finally classified as “no-tag”.

As these three categories present very different signal-to-background ratio and display some differences in the object description, they are treated separately in the final optimization. In addition, some differences in the cuts are also needed due to the different background composition, especially in the case of the “two tags” sample, in which the $t\bar{t}$ background starts to be noticeable.

On the other hand, since the optimization is based mostly on the intrinsic properties of the final state under investigation, there are also parts that are shared by the three categories.

6.2 $H \rightarrow ZZ \rightarrow 2l2j$ decay

The decay kinematics in the signal $H \rightarrow ZZ \rightarrow 2l2j$ have several distinct features that can be used to discriminate against background. Particularly useful are the

decay angles. Five angular observables fully describe kinematics in the decay $2 \rightarrow 1 \rightarrow 2 \rightarrow 4$ as in $ab \rightarrow X \rightarrow ZZ \rightarrow 2\ell 2j$ [21, 22]. Additionally they are orthogonal to the three invariant masses of the X and the two Z and to the longitudinal and transverse momenta of the X . The orthogonal observables are largely uncorrelated and are useful for event selection beyond using them as raw kinematic observables.

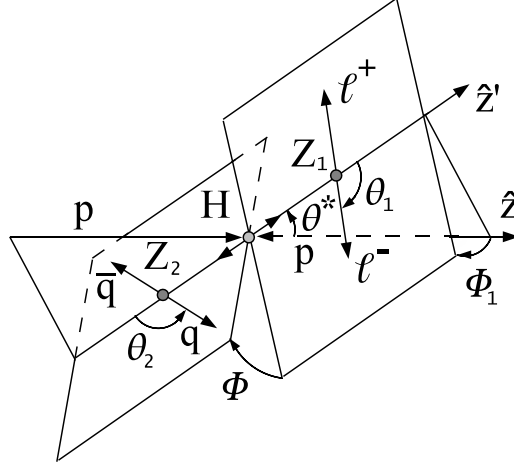


Figure 6.1. Diagram depicting the decay $X \rightarrow ZZ \rightarrow 2\ell 2q$ and the 5 decay angles which describe it.

Fig. 6.1 illustrates the angular distribution in the production and decay chain $ab \rightarrow X \rightarrow P_1 P_2 \rightarrow p_{11} p_{12} p_{21} p_{22}$ with an example of the $ab \rightarrow X \rightarrow ZZ \rightarrow 4\ell$ or $2\ell 2q$ (where quarks q hadronize to jets, which we refer to as $2\ell 2j$ channel later) chain with two partons a and b , such as gg or $q\bar{q}$. The angular distribution can be expressed as a function of three helicity angles θ_1 , θ_2 , and Φ , and two production angles θ^* and Φ_1 , as shown in Fig. 6.1.

Here θ_i is the angle between the direction of the ℓ^- or q from the $Z \rightarrow \ell^+ \ell^-$ or $q\bar{q}$ (where the quark-anti-quark pair produces two jets) and the direction opposite the X in the Z rest frame, and Φ is the angle between the decay planes of the two Z systems. The two Z 's are distinguished by their decay type or, in case their daughters are the same type of particles, by an arbitrary convention. The production angle θ^* is defined as the angle between the parton collision axis z and the X decay axis in the

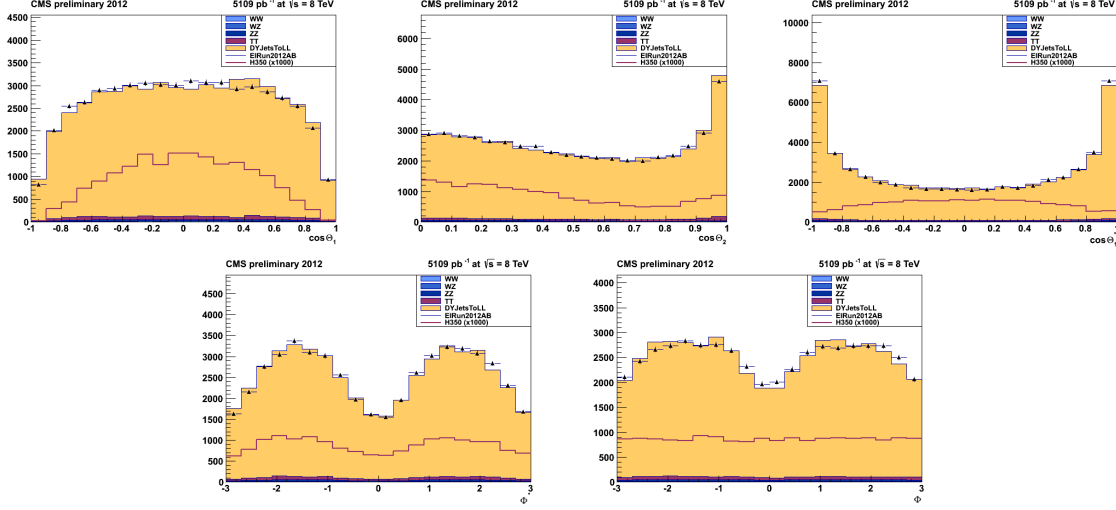


Figure 6.2. Five angular distributions of $\cos \theta_1$, $\cos \theta_2$, $\cos \theta^*$, Φ , Φ_1 and the helicity likelihood discriminant for 2012 electron data (points) and Summer 12 Monte Carlo samples (histogram). Open histograms indicate the expected distribution for a Higgs boson with mass 350 GeV, multiplied by a factor of 1000 for illustration. The selection described in preselection is applied except for the cut on m_{jj} .

X rest frame. The fifth angle can be defined as Φ_1 , the angle between the production plane and the first Z decay plane.

A comparison of angular distribution in data and MC for electron and muon events can be found in Figs. 6.2 and 6.3, where there is good agreement in samples dominated by background.

6.3 Signal Optimization Based on Helicity Neural Network

A Neural Network is applied to separate Higgs signal from background using the five helicity angles of the final objects in the analysis as inputs. A training and test evaluation has been performed with the framework of the TMVA package [23] using the real mixture of MC processes as background and Higgs MC. Both the training and testing samples for both background and signal are constructed and events randomly mixed outside of the Neural Network. There are 37,874 signal events and 38,378

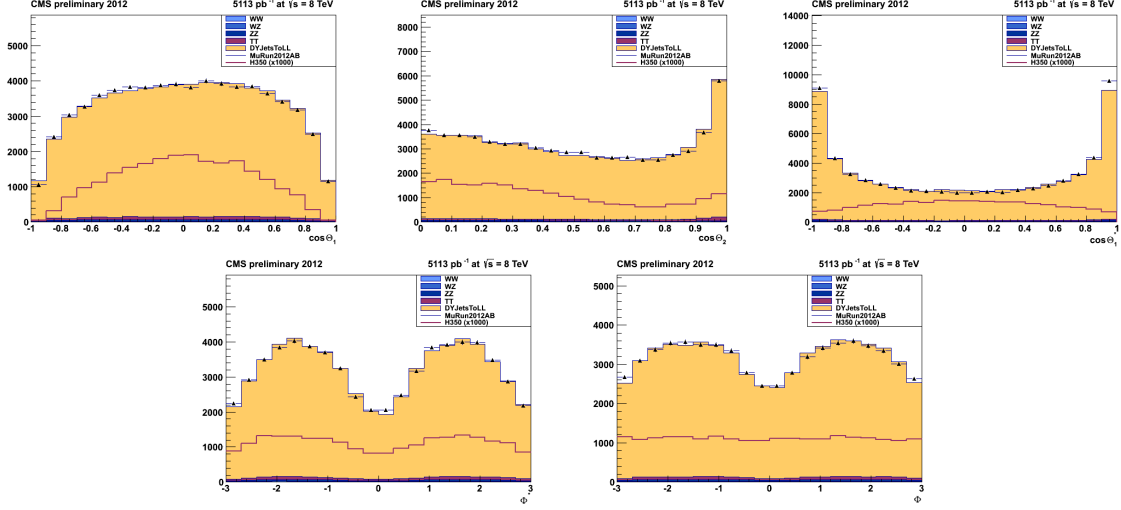


Figure 6.3. Five angular distributions of $\cos \theta_1$, $\cos \theta_2$, $\cos \theta^*$, Φ , Φ_1 and the helicity likelihood discriminant for 2012 muon data (points) and Summer 12 Monte Carlo samples (histogram). Open histograms indicate the expected distribution for a Higgs boson with mass 350 GeV, multiplied by a factor of 1000 for illustration. The selection described in preselection is applied except for the cut on m_{jj} .

background events evenly split between testing and training. The Neural Network was trained on the Monte Carlo generated for a hypothetical Higgs mass of 400 GeV. This training was on events that passed the previously explained preselection, the Z boson mass window cuts previously explained, but before the cut on MET significance.

6.3.1 Neural Network Architecture

The architecture of the Neural Network consists of two hidden layers with N and N neurons respectively, where N is the number of variables, and one output node as shown in Figure 6.4.

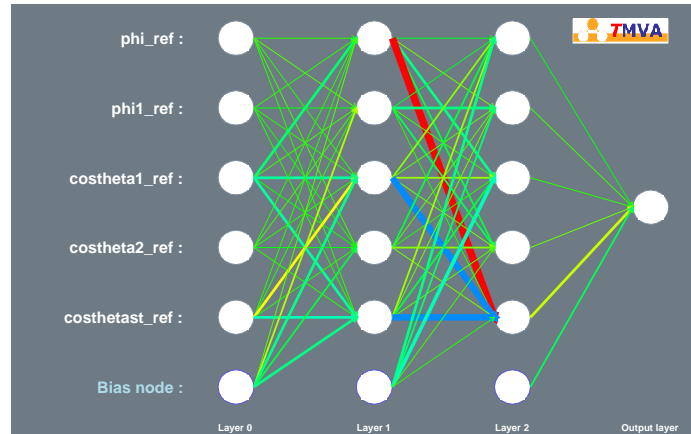


Figure 6.4. Neural Network architecture used for the training.

6.3.2 Results Based on Helicity Neural Network

The training and testing of the Neural Network can be seen in Figure 6.5. While there is separation between the signal and background, and good agreement between testing and training, there is a definite spike in background at the same place as the signal spikes in the MLP training, with a similar but less noticeable effect for the Likelihood training. The training was done with Monte Carlo generated for a hypothetical Higgs mass of 400 GeV but then applied to the Monte Carlo for hypothetical Higgs masses of 200, 300, 400, and 500 GeV. This is because the angular components should not depend on Higgs mass. In addition to the MLP neural network done, each training and testing was also performed with a Likelihood. The performance of the two analysis is almost identical in all cases, especially when looking in tight hypothetical Higgs mass regions.

In the 1-tag region, after applying the additional MET significance cut, the separation between signal and background looks similar to the training for Higgs 300, 400, and 500 GeV, but does not have much discrimination power for a Higgs of 200 GeV. The discrimination power is virtually the same after applying an additional cut

of of $-6\%/+10\%$ of the Higgs mass for all four cases as seen in Figure 6.6. When comparing background rejection versus signal efficiency between the Neural Network and Likelihood Discriminate the performance is for all practical purposes the same. See Figure 6.8.

In the 2-tag region, after applying the additional MET significance cut and a cut of of $-6\%/+10\%$ of the Higgs mass, the discriminating power of the Neural Network is similar to the 1tag case with poor ability for a Higgs of 200 GeV, but good separation for Higgs of 300, 400, and 500 GeV. This is shown in Figure 6.7. When comparing background rejection versus signal efficiency between the Neural Network and Likelihood Discriminate the performance is roughly the same, except for the Higgs of 200 GeV case where the Neural Network is almost consistantly better than the Likelihood Discriminate. This is shown in Figure 6.9.

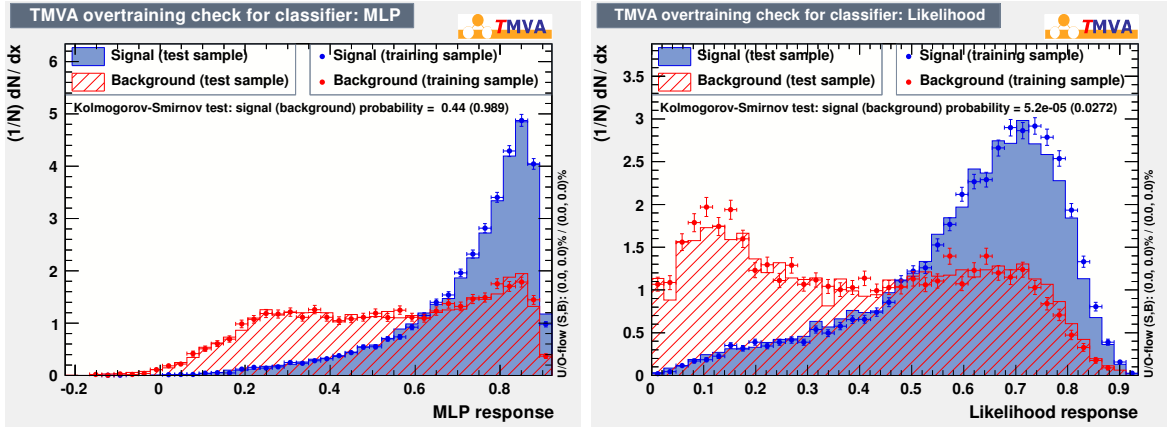


Figure 6.5. The trainings are done after preselection and additionally require at least one TCHEM jet (Track Counting High Efficiency > 1.9) Left: Training 400 GeV Higgs boson with a MLP neural network. Right: Training 400 GeV Higgs boson with a Likelihood.

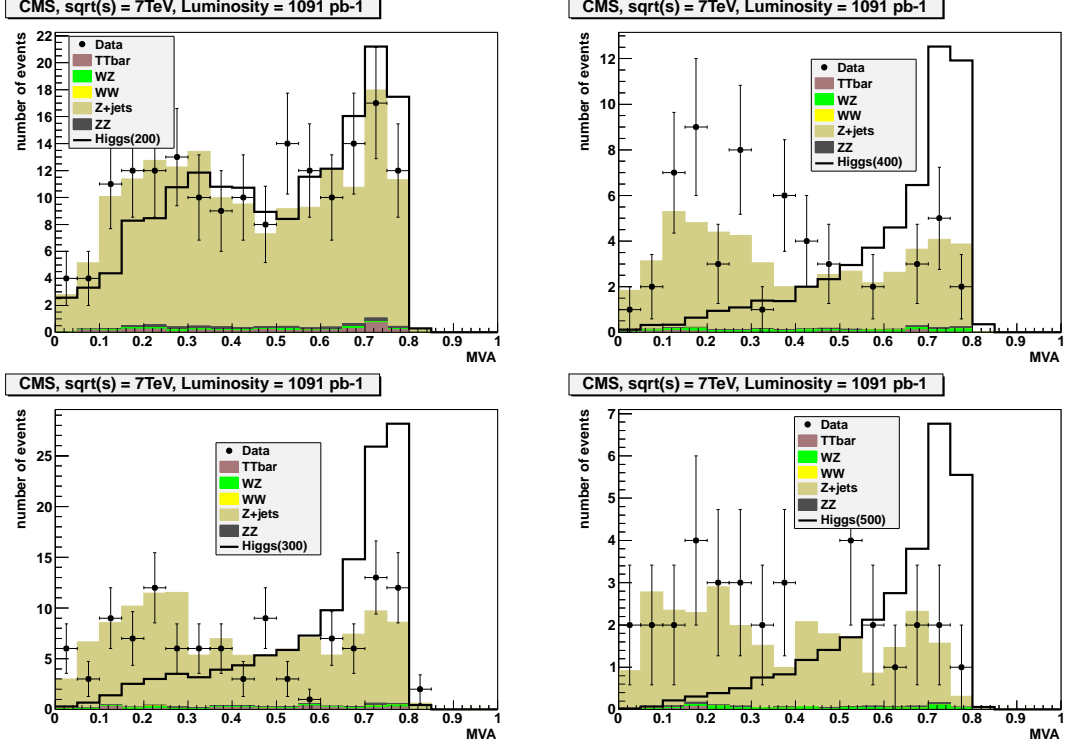


Figure 6.6. Signal and background Neural Network output in the 1tag region after preselection, the Z boson mass window cuts, cut on MET significance, and a $-6\%/+10\%$ Higgs mass window. The signal is scaled to the sum of the Monte Carlo background.

6.3.3 Potential Improvement to to MVA

A neural network or likelihood should be able to take advantage of more information than simply using the 5 decay angles as previously shown. One such variable that offers good discrimination power in the preselection region is $Z_{upt} / \sum pt$, where $\sum pt = l_0pt + l_1pt + j_0pt + j_1pt + met$, and Z_{upt} and $\sum pt$ are scalar quantities. This variable distribution is shown in Figure 6.11. While adding this variable to a MLP training improves the discrimination power in the inclusive pretag region, once we look in a mass window of $-6\%/+10\%$ around Higgs 400 adding $Z_{upt} / \sum pt$ to the training actually lowers the discrimination power of the neural network. This same

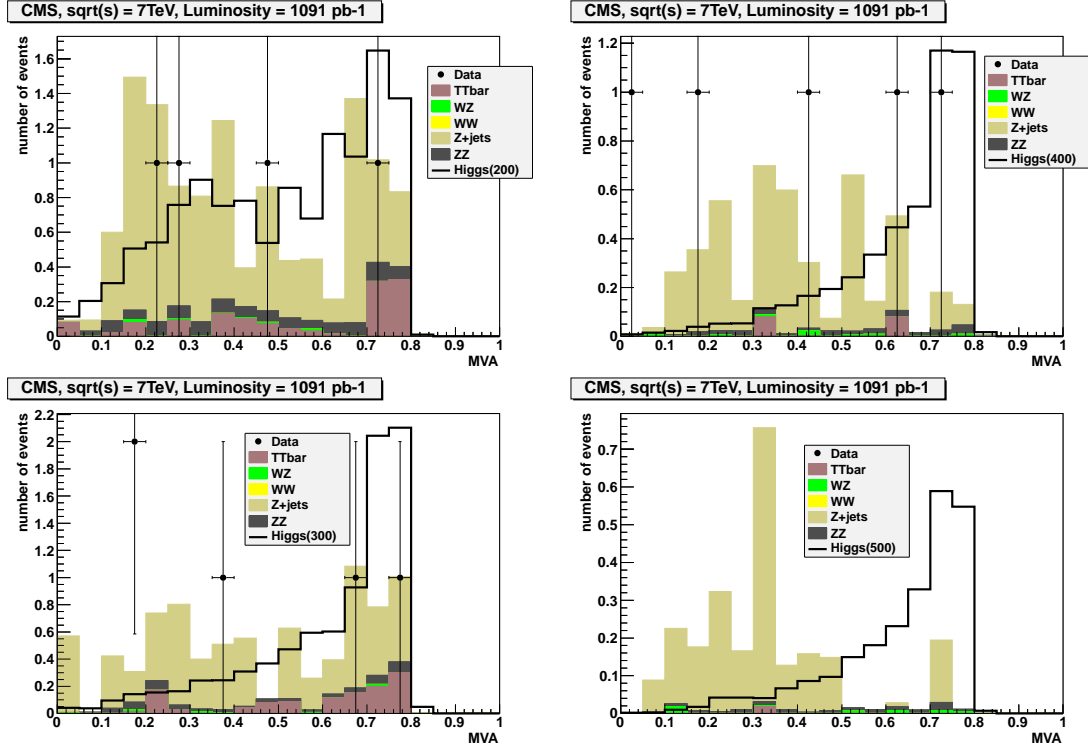


Figure 6.7. Signal and background Neural Network output in the 2 tag region after preselection, the Z boson mass window cuts, cut on MET significance, and a $-6\%/+10\%$ Higgs mass window. The signal is scaled to the sum of the Monte Carlo background.

performance drop is seen when training is done with the addition of the reconstructed higgs mass as well.

While training on each Higgs mass can give additional discrimination power over training on one Higgs mass and applying it to multiple hypothetical Higgs masses there are some difficulties. It is more convenient to be able to train on just one higgs mass and then apply this training to a range of hypothetical higgs masses. An example of training on a Higgs 400 GeV sample and then applying this training to various hypothetical Higgs masses is show in Figure 6.12. This training is for a MLP neural network tranined on the 5 decay angles. This training is applied after preselection and requiring at least one TCHEM jet. This shows that good performance can be acheived using the MLP by only training on one Higgs mass sample, without the

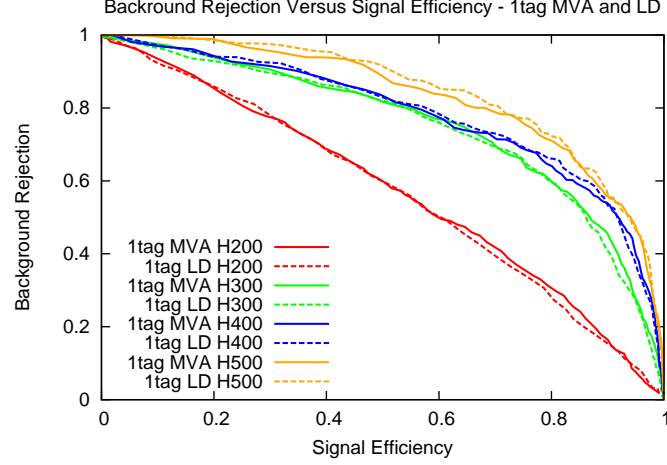


Figure 6.8. Background Rejection Versus Signal Efficiency in the 1tag region comparing the Multi Variant Analysis output to the the Helicity Likelihood Discriminant for a Higgs mass of 200, 300, 400, and 500 GeV. This is calculated after preselection cuts, Z boson mass cuts, cut on MET significance, in a $-6\%/+10\%$ Higgs mass window.

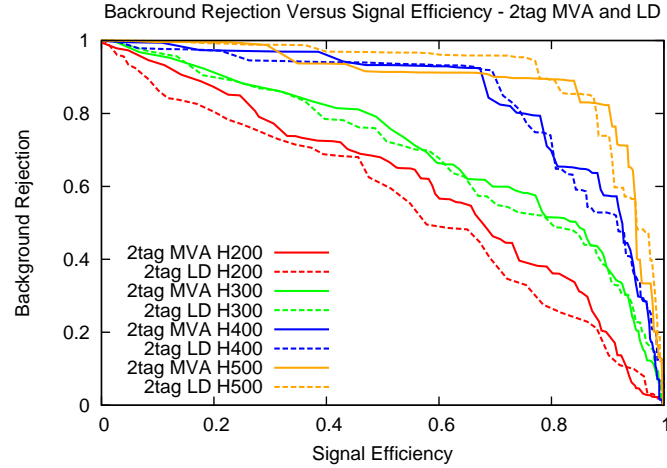


Figure 6.9. Background Rejection Versus Signal Efficiency in the 2tag region comparing the Multi Variant Analysis output to the the Helicity Likelihood Discriminant for a Higgs mass of 200, 300, 400, and 500 GeV. This is calculated after preselection cuts, Z boson mass cuts, cut on MET significance, in a $-6\%/+10\%$ Higgs mass window.

need for additional trainings, or parameterising a discriminate as a function of the reconstructed Higgs mass.

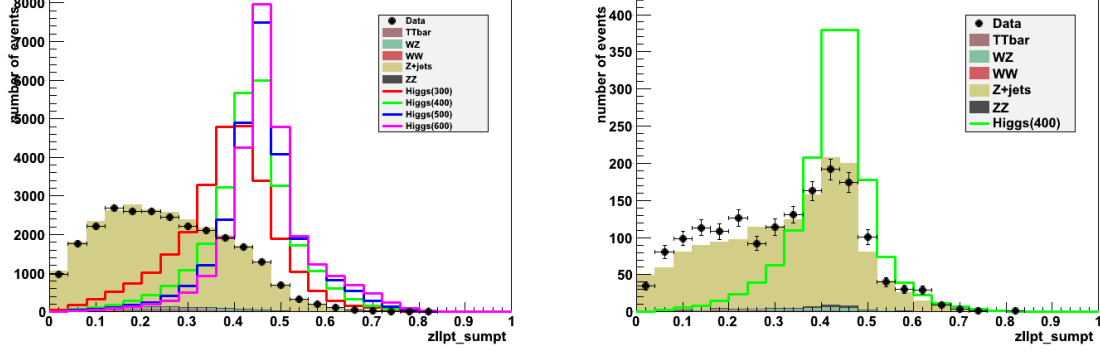


Figure 6.10. Signal samples are normalized to background. Left: $\frac{Z_{upt}}{\sum pt}$ after preselection. Right: $\frac{Z_{upt}}{\sum pt}$ after preselection and $376 < m_{ZZ} < 440$ GeV.

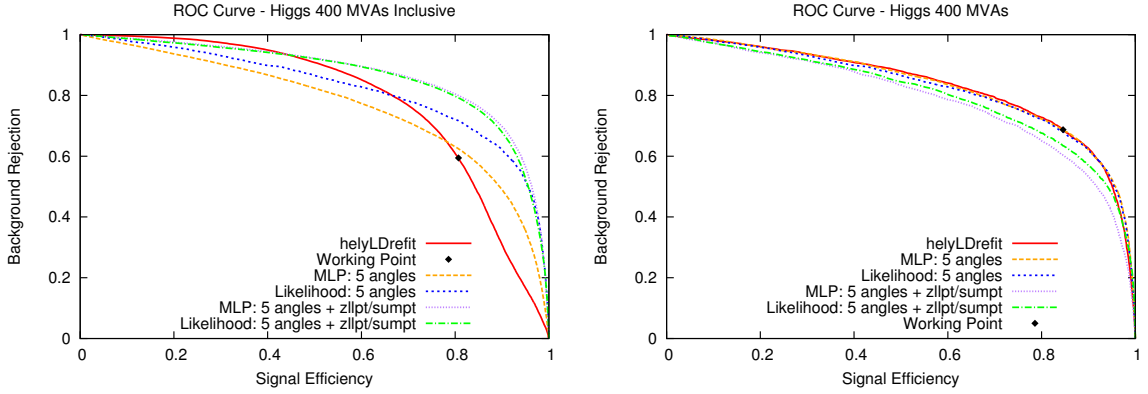


Figure 6.11. Applying a MLP training to preselection and at least one TCHEM jet. The working point is the equivalent performance of current analysis that we apply in the two tag region. For comparison the helyLDrefit variable is also shown. Left: ROC curves after preselection. Right: ROC curves after preselection and $376 < m_{ZZ} < 440$ GeV.

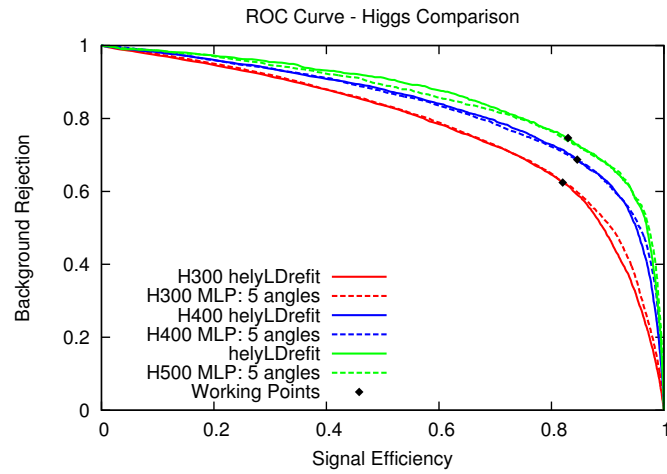


Figure 6.12. This training is for a MLP neural network trained on the 5 decay angles of a Higgs 400 GeV sample. This training is applied after preselection and requiring at least one TCHEM jet to samples with a Higgs mass of 300,400, and 500 GeV. The working point is the background rejection point that we currently achieve in the two tag region in our analysis. The helyLDrefit variable is shown for comparison.

7. Conclusion

There are many improvements that can be made to the analysis for the 2012 data to both improve sensitivity and extend the analysis range. There is also more work that needs to be done to with the new data that is constantly arriving from the CMS detector.

7.1 Future Plans

- Train analysis for high mass range (from 600 GeV to 1000 GeV).
- Run limit calculations on the MVA analysis.
- Run tag and probe efficiency calculations for muons.
- Optimize Z_{ll} and Z_{jj} cuts for preselection.
- Improve reconstruction of high mass jets.
- Run analysis over full 2012 data sample at the end of the run.

LIST OF REFERENCES

LIST OF REFERENCES

- [1] S.L. Glashow. Partial Symmetries of Weak Interactions. *Nucl.Phys.*, 22:579–588, 1961.
- [2] A. Salam. Elementary Particle Physics: Relativistic Groups and Analyticity. *Val. Proceeding to the Eighth Nobel Symposium. Almqvist and Wiksell*, 1968.
- [3] S. Weinberg. A Model of Leptons. *Phys.Rev.Lett.*, 19:1264, 1967.
- [4] L. H. Ryder. *Quantum Field Theory*. Cambridge University Press, 1996.
- [5] J. Beringer et al. (Particle Data Group). 2012 Review of Particle Physics. *Phys. Rev. D*, 86, 2012.
- [6] CERN 2004-003. *LHC Design Report*. 2004.
- [7] ATLAS Collaboration. The ATLAS Experiment at the CERN Large Hadron Collider. *Journal of Instrumentation*, 3:S08003, 2008.
- [8] ALICE Collaboration. The ALICE Experiment at the CERN LHC. *Journal of Instrumentation*, 3:S08002, 2008.
- [9] LHCb Collaboration. The LHCb Detector at the LHC. *Journal of Instrumentation*, 3:S08005, 2008.
- [10] The CMS Collaboration. The CMS experiment at the CERN LHC. *Journal of Instrumentation*, 3:S08004, 2008.
- [11] S. de Visscher et al. J. Alwall, P. Demin. MadGraph/MadEvent v4: the new webgeneration. *JHEP*, 09:028, 2007.
- [12] P.Nason. A New Method for Combining NLO QCD with Shower Monte Carlo Algorithms. *JHEP*, 11:040, 2004.
- [13] . Oleari S. Frixione, P. Nason. Matching NLO QCD Computations with Parton Shower Simulations: the POWHEG method. *JHEP*, 11:070, 2007.
- [14] C. Oleari et al. S. Alioli, P. Nason. NLO Vector-Boson Production Matched with Shower in POWHEG. *JHEP*, 07:06, 2008.
- [15] L. Borrello et al. "search for the standard model higgs boson in the decay channel $h \rightarrow zz \rightarrow 2l2b$ ". CMS Analysis Note CMS-AN-2011-399, 2011.
- [16] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. The Anti-k(t) jet clustering algorithm. *JHEP*, 0804:063, 2008.
- [17] CMS Collaboration. "jet energy corrections determination at 7 tev". CMS Physics Analysis Summary CMS-PAS-JME-10-010, 2010.

- [18] S. Bolognesi, A. Bonato, D. Del Re, A.V. Gritsan, M. Manelli, M. Mozer, F. Pandolfi, N.V Tran, and A. Whitbeck. Search for the standard model higgs boson in the decay channel $h \rightarrow zz \rightarrow 2\ell 2q$ at cms. *CMS Note*, CMS-AN-2011/388, 2011.
- [19] S. Bolognesi, A. Bonato, A.V. D. Del Re, A.and Gritsan, F. Pandolfi, N.V Tran, and A. Whitbeck. Search for a semileptonic decay of a sm higgs or bsm boson $H \rightarrow ZZ \rightarrow 2l 2j$. *CMS Note*, CMS-AN-2011/100, 2011.
- [20] The CMS Collaboration. Measurements of the Inclusive W and Z Cross Sections in pp Collisions at $\sqrt{s}=7$ TeV. *JHEP*, 1101:080, 2011.
- [21] Y. Gao et al. Spin determination of single-produced resonances at hadron collider. *Phys. Rev.*, D81:075022, 2010.
- [22] A. De Rujula et al. Higgs look-alikes at the LHC. *Phys. Rev.*, D82:013003, 2010.
- [23] et al. A. Hoecher. TMVA - Toolkit for Multivariate Data Analysis. *arXiv:physics/0703039*, 2007.