Search for the Standard Model Higgs Boson Decaying to Two W Bosons at CMS

by

Si Xie

Submitted to the Department of Physics
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2012

Author .....
Department of Physics
May 08, 2012

Certified by .....
Christoph M. E. Paus
Professor
Thesis Supervisor

Accepted by .....
Prof. Krishna Rajagopal
Associate Department Head for Education

Search for the Standard Model Higgs Boson Decaying to Two W Bosons at CMS

by

Si Xie

Submitted to the Department of Physics
on May 08, 2012, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

In this thesis, we search for the production of the Standard Model Higgs boson at the Large Hadron Collider, through its decay mode to two W bosons, which each in turn decay into a charged lepton and a neutrino. The Higgs boson is the only elementary particle predicted by the Standard Model of particle physics which has not yet been experimentally observed. The question of its existence has been the subject of a wide body of experimental and theoretical work for the past half-century. Using 4.9 fb^{-1} of integrated luminosity, this search is expected, on average, to exclude the Standard Model predicted Higgs boson production cross section for masses between 126 GeV/c^2 and 260 GeV/c^2 under the background-only hypothesis. We observe no statistically significant excesses in the data, and we exclude the Standard Model Higgs boson production for masses between 129 GeV/c^2 and 260 GeV/c^2. This result represents a significant reduction of the mass region in which the Standard Model Higgs boson is allowed to exist.

Thesis Supervisor: Christoph M. E. Paus
Title: Professor

Acknowledgments

I would like to begin by thanking my Ph.D. supervisor, Christoph Paus, whose insistence on expressing problems in the most simplistic terms, helped to guide me towards the core and fundamental issues of any problem. I am extremely grateful for the help, both fundamental and technical, that I received from Guillelmo Gomez-Ceballos, who has served as a second advisor for me during my time at CERN. Additionally, I want to thank the two other faculty members of the MIT CMS group, Steve Nahn and Markus Klute, who have given numerous sharp and critical suggestions during my Ph.D. period.

I want to thank, in particular, the three people I have shared my office with for the last four years. Gerry Bauer and Sham Sumorok have given me countless, invaluable advice and I thank them for transferring some bits of their vast experience and wisdom to me. I enjoyed, in particular, the many stories from past experiments and discoveries, and stories about certain fabled characters in particle physics. Although I have shared my office with Fabian Stoeckli only for the last year of my Ph.D., he has left a strong impression on me, both for his vast knowledge about theoretical and experimental aspects of particle physics in general, and for his extreme insistence on precision and correctness.

I want to thank all the people who worked with me on this analysis for all of their help and support: Dima Kovalskyi, Marco Zanetti, Kevin Sung, Dave Evans, Gao Yanyan, and Giuseppe Cerati. I thank, in particular, Avi Yagil, Claudio Campagnari, and Frank Wuerthwein for offering their unique perspectives. Finally, I thank all of my fellow students, postdocs, and other colleagues in the MIT and friends group for all of their support through the years: Bruce Knuteson, Georgios Choudalakis, Conor Henderson, Khaldoun Makhoul, Phil Harris, Pieter Everaerts, Matt Rudolph, Kristian Hahn, Ilya Kravchenko, Erik Butz, Slawek Tkaczyk, Josh Bendavid, Max Goncharov, Matthew Chan, Valentina Dutta, Roger Wolf, Duncan Ralph, Yang Mingming, Aram Apyan, Andrew Lewin. Last but not least I thank Olivier Raginel for

always being available whenever one of my desktops misbehaved or died.

I thank my mother and my father for giving me life and guiding me to where I am today. I thank them and my two twin brothers for their unconditional love and support.

Contents

Chapter 1

Introduction

Physics is a scientific discipline whose primary aim is to describe all naturally occurring phenomena in terms of the matter content of the universe and how this matter interacts with each other. Elementary particle physics addresses these goals at the most fundamental level, attempting to enumerate the most basic building blocks of matter, and to describe the most basic interactions between them. Historically, progress in particle physics has followed a reductionist path, whereby layers of complexity have been successfully explained in terms of ever more basic building blocks. Atoms were reduced to electrons, protons and neutrons, and protons and neutrons were in turn reduced to quarks and gluons. Similarly, the electric and magnetic forces were combined into the electromagnetic force, which was in turn combined with the weak force to give us the combined description of the electroweak force. These successes have converged, at the current state of our understanding, to the theory of elementary particles and interactions known as the Standard Model of particle physics. This theory has been tested experimentally to unprecedented precision and so far no evidence has been found that contradict its predictions [1]. Despite this success, three fundamental questions, currently, remain:

- What is the origin of dark energy and dark matter, accounting for more than 95% of the energy density of the universe [2] and more than 80% of its matter [3], respectively?

- What are the masses of the neutrinos, and what is the mechanism that generates them [4]?

- What is the mechanism for electroweak gauge symmetry breaking, which generates the masses of the W and Z bosons?

The work presented in this thesis attempts to address the third, and the oldest, of these questions.

1.1 The Standard Model

The Standard Model of elementary particle physics is a theory that attempts to describe all observed particles and their mutual interactions. We begin by providing a summary of all known elementary particles that any theory of particle physics must describe [4]. The four fundamental forces that we observe in nature are the electromagnetic force, the weak force, the strong force, and the gravitational force, each affecting the fundamental particles in specific ways. The Standard Model describes all of these forces with the exception of the gravitational force.

Certain patterns are observed among the known elementary particles that allow us to group them according to the symmetries that are obeyed. All elementary particles that have been observed are either spin 1/2 fermions, which obey the Pauli exclusion principle, or spin 1 bosons. The fermions exist in two varieties: leptons, which do not interact via the strong interactions, and quarks, which do interact via strong interactions. For each fermion there exists also a corresponding anti-fermion which carry same mass but opposite electric charge. There are six currently known leptons, of which there are two types. The electron, muon, and tau particle are electrically charged, while the three corresponding neutrinos are electrically neutral. Other than the difference in mass, the three charged leptons are identical with respect to how they interact under all of the known fundamental forces. Due to the way in which they interact through the weak force, the leptons are classified into three generations

and are represented as the following three pairs

$$\begin{pmatrix}\nu_e\\ e\end{pmatrix} \quad , \quad \begin{pmatrix}\nu_\mu\\ \mu\end{pmatrix} \quad , \quad \begin{pmatrix}\nu_\tau\\ \tau\end{pmatrix} \tag{1.1}$$

where the neutrinos are first in the pair and the charged leptons are the second in the pair. Similarly, there are six currently known quarks, of which there are also two types. The up, charm, and top quarks all have an electric charge of $+2/3$, while the down, strange, and bottom quarks all have an electric charge of $-1/3$. They are analogously classified into three generations and represented as the following three pairs

$$\begin{pmatrix}u\\ d\end{pmatrix} \quad , \quad \begin{pmatrix}c\\ s\end{pmatrix} \quad , \quad \begin{pmatrix}t\\ b\end{pmatrix}. \tag{1.2}$$

There are four fundamental bosons that have been observed so far. The photon (γ) is massless, electrically neutral, and does not interact via the weak or the strong force. The W bosons are massive, have $+1$ or -1 electric charge, and interacts via the weak force, but does not interact via the strong force. The Z boson is massive, electrically neutral, and interacts via the weak force, but does not interact via the electromagnetic force or the strong force. Finally, the gluon is massless, electrically neutral, and does not interact via the weak force, but does interact via the strong force. More precisely, there are eight different types of gluons carrying the eight different types of strong force color charges.

The standard model provides a description of all of these elementary particles and their mutual interactions, with the exception of the gravitational interaction.

1.1.1 Gauge Symmetry

Formally, the Standard Model is a quantum field theory, in which every particle is described by a dynamical field $\Psi(x)$ residing inside four dimensional space-time (x), and globally respecting the symmetries of special relativity: spatial translation, spatial rotations, and boosts of the reference frame. The kinematics and dynamics of all particles are described by a Lagrangian function constructed by postulating the set of

fundamental symmetries that exist in nature. Any interactions between the particles are fully described by the individual terms of the Standard Model Lagrangian [5].

The fundamental postulate underlying the Standard Model is that the dynamics of all particles are determined by an underlying local gauge symmetry. A local gauge symmetry refers to the invariance of the Lagrangian under a local gauge group transformation, where the fields describing particles transform in the following manner:

$$\Psi(x) \rightarrow e^{i\alpha^A(x)T^A}\Psi(x) \tag{1.3}$$

where T^A are the generating matrices of a particular continuous group, α^A are parameters which can be physically interpreted as coupling strenghts, and the indices A run over all of the generators of the fundamental representation of the particular group. By imposing local gauge invariance, it is necessary to introduce additional gauge fields which describe the particles mediating the forces that determine the dynamics.

1.1.2 The U(1) Gauge Symmetry Group: Quantum Electrodynamics

As a simple illustration of this principle, we consider the example of Quantum Electrodynamics [5], which describes the dynamics of electromagnetically charged particles. Charged particles of mass m and charge Q are described by the Dirac field $\Psi(x)$, formally a four component spinor. We begin with a theory of free charged particles described by the Dirac free Lagrangian:

$$L_\Psi = \bar{\Psi}(i\partial\!\!\!/ - m)\Psi. \tag{1.4}$$

For convenience, we introduce the notation $\partial\!\!\!/ = \sum_\mu \gamma^\mu \partial_\mu$, where γ^μ are the Dirac matrices, and μ are indices running from 0 to 3, representing the four space-time coordinates. We illustrate the way in which postulating a local U(1) gauge invariance on the phase of the charged particle fields Ψ, automatically requires the addition of

a gauge field A describing the photon, as well as a term in the Lagrangian describing the interaction between the photon and the charged Dirac particles [5].

We begin by performing the local U(1) gauge transformation $\Psi \rightarrow \Psi' = e^{iQ\alpha(x)}\Psi(x)$, and find that the Lagrangian transforms as:

$$L_\Psi \rightarrow L'_\Psi = L_\Psi + \bar{\Psi}\gamma^\mu\Psi(\partial_\mu\alpha), \tag{1.5}$$

and is therefore not invariant under the local gauge transformation. We are, thus, forced to introduce the additional gauge field A_μ through the minimal coupling prescription by replacing the partial derivative ∂_μ by the covariant derivative $D_\mu = \partial_\mu + ieA_\mu$. Now if A_μ transforms as :

$$A_\mu \rightarrow A'_\mu = A_\mu + \frac{1}{e}\partial_\mu\alpha, \tag{1.6}$$

then we find that under the local gauge symmetry, the new Lagrangian transforms as:

$$\begin{aligned}
L_\Psi \rightarrow L'_\Psi &= \bar{\Psi}'(i\not{D}_\mu - m)'\Psi' \\
&= \bar{\Psi}'[(i\not{\partial} - e\not{A}') - m]\Psi' \\
&= \bar{\Psi}e^{i\alpha}[i\not{\partial} - e(\not{A} + \frac{1}{e}\not{\partial}\alpha) - m]e^{-i\alpha}\Psi \\
&= \bar{\Psi}[i\not{\partial} + \not{\partial}\alpha - e(\not{A} + \frac{1}{e}\not{\partial}\alpha) - m]\Psi \\
&= \bar{\Psi}[i\not{\partial} - e\not{A} - m]\Psi \\
&= L_\Psi,
\end{aligned} \tag{1.7}$$

and is now invariant.

The kinetic term of the new gauge field is formed from the electromagnetic strength tensor:

$$F_{\mu\nu} = [D_\mu, D_\nu]A = \partial_\mu A_\nu - \partial_\nu A_\mu. \tag{1.8}$$

This field strength tensor is invariant under the local gauge transformation and there-

fore the kinetic term for the gauge field A,

$$\mathcal{L}_A = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} \tag{1.9}$$

is also gauge invariant. A hypothetical mass term for the gauge field A,

$$\mathcal{L}_A^{mass} = -\frac{1}{2} m_A^2 A_\mu A^\mu \tag{1.10}$$

is not gauge invariant, and therefore the gauge field A must be massless in this theory.

Thus, simply by postulating a local U(1) gauge invariance, the theory of free charged particles is transformed into the theory of Quantum Electrodynamics, which describes interactions between charged particles mediated by the massless photon. The final Lagrangian describing Quantum Electrodynamics is

$$\mathcal{L}_{QED} = \bar{\Psi}(i\partial_\mu \gamma^\mu - m)\Psi - \frac{1}{4} F_{\mu\nu} F^{\mu\nu} - e\bar{\Psi}\gamma^\mu A_\mu \Psi. \tag{1.11}$$

1.1.3 The Standard Model Gauge Group

The gauge symmetry group of the Standard Model is $SU(3)_{\text{color}} \otimes SU(2)_L \otimes U(1)_Y$. The subgroup $SU(3)_{\text{color}}$ describes the strong interactions between quarks and gluons, while the subgroup $SU(2)_L \otimes U(1)_Y$ describes the electroweak interactions. We extend the concept of charge from the electric charge described by the U(1) local gauge symmetry to all of the above subgroups of the Standard Model gauge group.

The strong force in the Standard Model is determined by the particular properties of the $SU(3)$ part of the Standard Model gauge group [6,7]. The only particles which are charged under the strong force are quarks, which transform as a triplet under the $SU(3)$ gauge transformation, and the gluons transform as an octet under this transformation. The particular mathematical properties of the $SU(3)$ group, the color charges of the quarks and gluons, and the number of fermion generations observed are responsible for the fact that the strong coupling constant decreases as the energy

scale of the interaction increases.

This property has two very important phenomenological consequences. The first consequence is asymptotic freedom, the property whereby the strong coupling constant vanishes as the energy scale of a particular interaction increases. This property allows for definitive predictions to be made from perturbative calculations at high energy scales. The second consequence is that free quarks and gluons only exist at very short distance scales, typically referred to as confinement. As a result, at distance scales typical of observations made in particle physics experiments, only bound states of quarks and gluons are observed. Thus, confinement is responsible for the rich jet structures and hadron spectra observed at collider experiments.

The electromagnetic force and the weak force are described by the $SU(2) \otimes U(1)$ subgroup of the Standard Model gauge group [8–10]. The observation that parity is maximally broken in the weak interaction requires that the pairs of leptons given in Equation 1.1 form the doublets of the $SU(2)$ weak gauge group only if they are left handed. The right handed leptons are uncharged and form $SU(2)$ singlets. The remaining $U(1)$ symmetry, called hypercharge and denoted by Y, is not the usual symmetry of electromagnetism, because the $SU(2)$ doublets describing the leptons and neutrinos would need to have the same hypercharge. The $U(1)$ symmetry group associated with electromagnetism is hidden inside the $SU(2)_L \otimes U(1)_Y$ gauge group.

The gauge fields associated with the generators of the $SU(2)_L \otimes U(1)_Y$ gauge group are denoted by W^1_μ, W^2_μ, W^3_μ, and B_μ. They can be transformed into the fields describing the observed electroweak bosons as follows:

$$W^\pm_\mu = \frac{1}{\sqrt{2}}(W^1_\mu \mp W^2_\mu) \tag{1.12}$$

$$Z_\mu = \cos\theta_W W^3_\mu - \sin\theta_W B_\mu \tag{1.13}$$

$$A_\mu = \sin\theta_W W^3_\mu + \cos\theta_W B_\mu, \tag{1.14}$$

where θ_W is the weak mixing angle defined by the following relations between the $SU(2)$ coupling constant g and the $U(1)$ coupling constant g' and quantifies the

degree of mixing of the $SU(2)$ gauge fields and the $U(1)_Y$ gauge field in the observed photon and the Z boson:

$$\sin\theta_W = \frac{g'}{\sqrt{g^2 + g'^2}} \tag{1.15}$$

$$\cos\theta_W = \frac{g}{\sqrt{g^2 + g'^2}}. \tag{1.16}$$

The weak mixing angle is a free parameter of the Standard Model and is measured to be approximately 30° [4].

The usual electromagnetic charge is identified from the term coupling the photon field A_μ with the W^+ and W^- fields as follows:

$$e = g\sin\theta_W. \tag{1.17}$$

We also obtain a specific prediction for the vector (V) and axial-vector (A) couplings of the Z boson to the fermions, in terms of the $SU(2)$ charge, T_3, and the electric charge Q:

$$g_V = T_3 - 2Q_i \sin^2\theta_W \tag{1.18}$$

$$g_A = T_3. \tag{1.19}$$

We make two critical observations. First, an explicit mass term for the electroweak bosons would violate gauge symmetry. Second, an explicit mass term for the fermions:

$$\mathcal{L}^{\text{mass}}_{\text{fermion}} = m\bar{\Psi}\Psi = m(\bar{\Psi_R}\Psi_L + \bar{\Psi_L}\Psi_R), \tag{1.20}$$

would couple left and right handed fermions and therefore would similarly violate the $SU(2)_L$ gauge symmetry. Therefore, we are presented with a troubling problem that explicit mass terms in the Lagrangian are disallowed on the one hand, while W and Z bosons as well as fermions are experimentally measured to be massive, on the other hand. We refer to this problem as the mass problem.

1.1.4 Mass Generation via Spontaneous Symmetry Breaking

In the Standard Model, the mass problem is solved by the spontaneous breaking of the gauge symmetry through what is known as the Higgs mechanism [11–16]. A symmetry of a particular Lagrangian is spontaneously broken if the vacuum of the theory is not invariant under the given symmetry transformation. Goldstone's theorem [17] states that for every spontaneously broken global symmetry, there must exist a massless boson. If the global symmetry that is spontaneously broken, is a local gauge symmetry, then the additional degrees of freedom carried by the Goldstone bosons are absorbed by the gauge bosons into new longitudinally polarized states as they acquire mass.

In the Standard Model, we introduce an additional Higgs field $\phi = \binom{h_1}{h_2}$, which transforms as a doublet under the weak $SU(2)$ gauge group, and as a singlet under the color $SU(3)$ gauge group, with a non-zero hypercharge Y_H. The Lagrangian describing this Higgs sector is the following:

$$\mathcal{L}_{\text{Higgs}} = (D^\mu \phi)^\dagger D_\mu \phi - \lambda \left(\phi^\dagger \phi - \frac{v^2}{2} \right)^2 . \tag{1.21}$$

The vacuum state of this Lagrangian lies on a circle in a complex two dimensional space described by $h_1^* h_1 + h_2^* h_2 = v^2/2$. A particular choice of the vacuum state on this hypersphere will spontaneously break the original gauge symmetry. By a choice of a particular gauge, called the unitary gauge, the vacuum state can be written as $\phi_0 = \frac{1}{\sqrt{2}}\binom{0}{v}$. Reparameterizing the Higgs fields by expanding around this chosen vacuum $\phi(x) = \frac{1}{\sqrt{2}}\binom{0}{v+h(x)}$, we find:

$$\begin{aligned}
&\mathcal{L}_{\text{Higgs}}^{\text{kinetic}} \\
&\quad = \tfrac{1}{\sqrt{2}} \begin{pmatrix} 0 & v+h \end{pmatrix} (\partial^\mu + igt^a W^{a\ \mu} + ig' \tfrac{1}{2} B^\mu)(\partial_\mu + igt^a W_\mu^a + ig' \tfrac{1}{2} B_\mu) \tfrac{1}{\sqrt{2}} \binom{0}{v+h} \\
&\quad = \tfrac{1}{2}\partial_\mu h \partial^\mu h + \tfrac{1}{8}(v+h)^2 g^2 (W_\mu^1 + iW_\mu^2)(W^{1\ \mu} - iW^{2\ \mu}) \\
&\qquad\qquad + \tfrac{1}{8}(v+h)^2 (gW_\mu^3 - g'B_\mu)(gW^{3\ \mu} - g'B^\mu) \\
&\quad = \tfrac{1}{2}(\tfrac{vg}{2})^2 W_\mu^+ W^{+\ \mu} + \tfrac{1}{2}(\tfrac{vg}{2})^2 W_\mu^- W^{-\ \mu} + \tfrac{1}{2}(\tfrac{v}{2}\sqrt{g^2+g'^2})^2 Z_\mu Z^\mu \\
&\qquad\qquad + \text{terms involving } h
\end{aligned} \tag{1.22}$$

where the $t^a = \sigma^a/2$ are the Pauli spin matrices. From equation 1.22, it can be seen that the spontaneous breaking of the symmetry of the gauge group $SU(2)_L \otimes U(1)_Y$ to the symmetry of the gauge group of electromagnetism, $U(1)_{\rm EM}$, results in the gauge bosons $W^{\pm}$ acquiring a mass of $M_{\rm W} = vg/2$ and the Z boson acquiring a mass of $M_{\rm Z} = v\sqrt{g^2 + g'^2}/2$.

In contrast to the way the gauge boson masses are generated, the masses of quarks and leptons are generated by Yukawa couplings of the fermions with the Higgs boson. Consequently, the strength of the coupling of all fermions to the Higgs boson is directly proportional to its mass. Defining the quark and lepton weak eigenstate $SU(2)$ doublet as $Q'_L = \binom{u'_L}{d'_L}$, the Yukawa interaction terms are :

$$\mathcal{L}_{\rm Yukawa} = -(G^{ij}_{\rm d}\bar{\rm d'_R}^{\,i}\phi^{\dagger}Q'^{j}_{\rm L} + G^{ij}_{\rm u'}\bar{\rm u'_R}^{\,i}(i\sigma_2\phi)^{\dagger}Q'^{j}_{\rm L}) + \text{hermitian conjugate}, \tag{1.23}$$

where the $G^{ij}_{\rm d}$ and $G^{ij}_{\rm u}$ are the matrices of Yukawa coupling constants between the fermions and the Higgs boson. If the Higgs field is expanded about its vacuum expectation value, $\phi = (1/\sqrt{2})\binom{0}{v+h}$, the terms from $\mathcal{L}_{\rm Yukawa}$ involving the vacuum expectation value will become fermion mass terms of the form :

$$\mathcal{L}_{\text{Fermion Mass}} = \bar{\rm d'_R}^{\,i} M^{ij}_d d'^{j}_{\rm L} + \bar{\rm u'_R}^{\,i} M^{ij}_u u'^{j}_{\rm L} + \text{hermitian conjugate}, \tag{1.24}$$

where $M^{ij}_{u,d} = (v\sqrt{2})G^{ij}_{u,d}$. These Yukawa coupling matrices may be diagonalized by unitary transformations $U_{L,R}$ and $D_{L,R}$ from the weak eigenstates u' and d' to the mass eigenstates u and d:

$$\begin{aligned}
\begin{pmatrix} u' \\ c' \\ t' \end{pmatrix}_{L,R} &= U_{L,R} \begin{pmatrix} u \\ c \\ t \end{pmatrix}, \\
\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix}_{L,R} &= D_{L,R} \begin{pmatrix} d \\ s \\ b \end{pmatrix},
\end{aligned} \tag{1.25}$$

such that

$$
\begin{aligned}
U_R^{-1} M_u U_L &= \begin{pmatrix} m_u & 0 & 0 \\ 0 & m_c & 0 \\ 0 & 0 & m_t \end{pmatrix}, \\
D_R^{-1} M_d D_L &= \begin{pmatrix} m_d & 0 & 0 \\ 0 & m_s & 0 \\ 0 & 0 & m_b \end{pmatrix}.
\end{aligned} \tag{1.26}
$$

The fact that the weak eigenstates are not exactly the same as the mass eigenstates leads to the rich structure of the CKM matrix and flavor changing interactions. The analogous structure exists also for the neutrino sector.

1.1.5 A Brief History of the Higgs Boson

At this stage, it is appropriate to place the current search for the Higgs boson within a global context. We do this by taking a short diversion through a discussion of the history of the Higgs boson and the Standard Model. Throughout the early 1960's, particle physicists were busy at work trying to construct a theory describing the interactions of leptons, some of which attempted to unify the electromagnetic and weak interactions. The central problem that one repeatedly returned to was how to incorporate a description of the masses of particle in a gauge invariant way.

The mechanism of generating masses through spontaneous symmetry breaking was first proposed by Phillip W. Anderson in 1962 [18], but without any discussion on an explicit relativistic model. In the spring of 1963, the relativistic model was first worked out by three groups working independently: Peter Higgs [13]; Robert Brout and Francois Englert [11]; and Gerald Guralnik, C. R. Hagen, and Tom Kibble [14]. Despite being given a name, "The Higgs mechanism", this work was largely ignored at the time and did not generate significant interest. The idea was picked up independently by Steven Weinberg and Abdus Salam in 1967 for their construction of

the Standard Model electroweak theory [9, 10]. They realized that the $SU(2) \otimes U(1)$ group used by Sheldon Glashow in 1961 to describe the leptonic electroweak interactions [8] can be spontaneously broken to generate the interactions needed for a full electroweak theory. Theoretical interest ramped up by 1971, after Gerard t'Hooft showed that Yang-Mills gauge theories, on which the electroweak theory was based, can be renormalized [19]. This proved that the electroweak theory could produce a viable and physically meaningful description of nature. By this time, theorists we engaged in exploring various alternative spontaneously broken gauge theories for the electroweak interactions and also other interactions.

In 1973, the entire picture changed through the discovery of neutral current interactions by the Gargamelle bubble chamber experiment at CERN [20], providing tantalizing evidence in favor of the electroweak theory. Parity violation measurements in inelastic electron scattering on deuterium and hydrogren targets at the Stanford Linear Accelerator in 1978 measured the weak mixing angle θ_W to an accuracy of 15% [21], which essentially predicted the masses of the W and Z bosons. Guided by these predictions, the UA1 and UA2 experiments unequivocally verified the Standard Model through the observation of the W and Z bosons [22, 23] in 1983.

The final missing piece of the puzzle is the Higgs boson itself. The Large Electron-Positron Collider (LEP) at CERN was constructed in 1989 and was a circular collider with an unprecedented circumference of 27 km. In the initial operation at a collision center-of-mass energy around the mass of the Z boson, precision measurements were made giving various indirect tests of the electroweak theory and constraints on the Higgs boson [1]. In the second stage of operation (LEP2), the center of mass energy was increased up to 209 GeV with the explicit goal of direct experimental observation of the Higgs boson. At the end of the LEP2 run, despite some suggestive hints, no definitive evidence was establish, and a lower limit on the Higgs boson mass at 114.4 GeV was extracted [24].

Finally, this historical interlude ends at the Large Hadron Collider (LHC) [25]. The LHC is a circular collider of protons on protons, and reuses the underground

tunnels formerly inhabited by the LEP collider. Its proposal dates back to 1977, only two years after the plans for the LEP collider were put forward, and had significant impact on the eventual size of the LEP tunnel. The main construction of the LHC began in 2001, shortly after the LEP collider was decommissioned. It finally completed construction in 2009 and began full operation in 2010. One of the main purposes of the LHC is to provide the definitive answer on the existence or non-existence of the Higgs boson. This answer will finally complete the half-century story of the Higgs boson. The results of this thesis will contribute an important step towards this final answer.

1.1.6 Experimental Constraints on the Higgs Boson

The mass of the Higgs boson is not a prediction of the Standard Model, but rather a free parameter of the theory. However, there are a number of experimental constraints, both from direct searches as well as precision measurements of electroweak observables [1] sensitive to the presence of the Higgs boson through higher-order virtual-loop corrections. The status of these experimental constraints at the summer of 2010 are summarized in Figure 1-1. Direct searches at LEP [24] excluded Higgs boson production for masses below 114.4 GeV/c^2 at the 95% confidence level, and direct searches at the Tevatron [26–28] excluded Higgs boson production for masses between 158 GeV/c^2 and 175 GeV/c^2 at the 95% confidence level. Fits of the precision measurements of electroweak observables at the LEP experiments and SLD, as well as the precision measurements of the W mass and the top quark mass at the Tevatron provide additional constraints on the Higgs boson mass, illustrated by the $\Delta\chi^2$ curve shown in Figure 1-1.

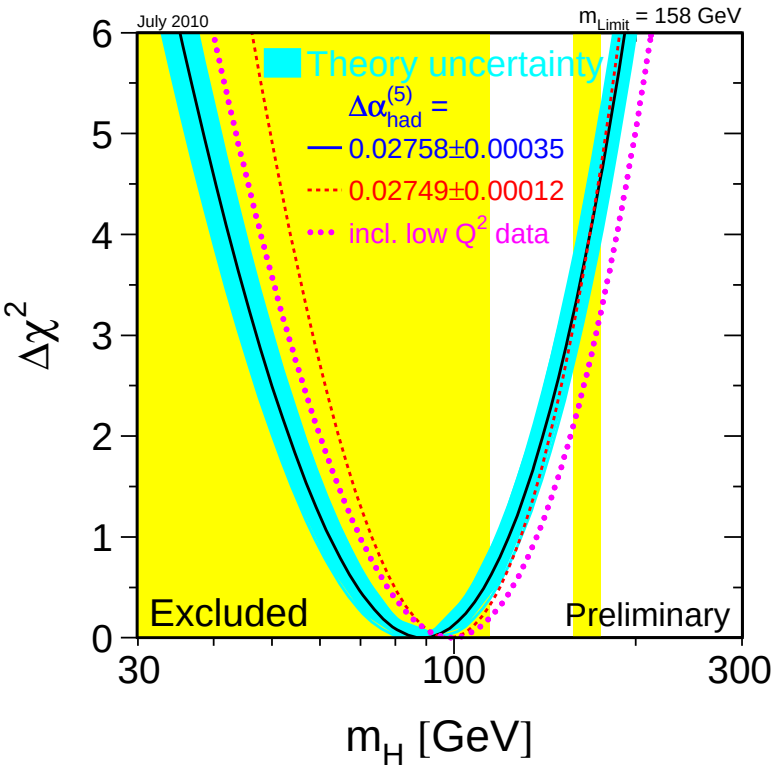


Figure 1-1: Combined results from the LEP Electroweak Working Group,the Tevatron Electroweak Working Group, and the SLD Electroweak and Heavy Flavor Groups, from the summer of 2010. The $\Delta\chi^2$ is plotted as a function of the hypothesized Higgs boson mass, resulting from a fit to all high-Q^2 data. The blue band represents an estimate of the theoretical systematic error due to missing higher order corrections. The vertical yellow bands represent the 95% confidence level limits from LEP and the Tevatron.

1.2 Higgs Boson Production at the LHC

The Higgs boson production mechanisms at the LHC are depicted by the Feynman diagrams shown in Figure 1-3. The main production mechanism is the gluon fusion process, where the Higgs boson is produced from two initial state gluons mediated by virtual heavy fermions that couple to the Higgs boson. A smaller contribution, with a rate that is roughly one order of magnitude lower, comes from the vector boson fusion (VBF) process, where the Higgs boson is produced at tree level by two vector bosons, leaving a distinct signature of two high energy quarks with a large gap in rapidity between them. Higgs bosons produced in association with a W or a Z boson via the Higgs-strahlung process has a rate that is roughly another factor of 2 smaller at Higgs boson masses between 140 GeV/$c^2$and 200 GeV/c^2. Finally, the Higgs boson can be produced in association with a top quark and an anti-top quark, with a rate that is roughly another order of magnitude smaller than Higgs-strahlung [29–43]. These LHC production cross sections at a center of mass energy of 7 TeV are shown in Figure 1-2 as a function of the Higgs boson mass.

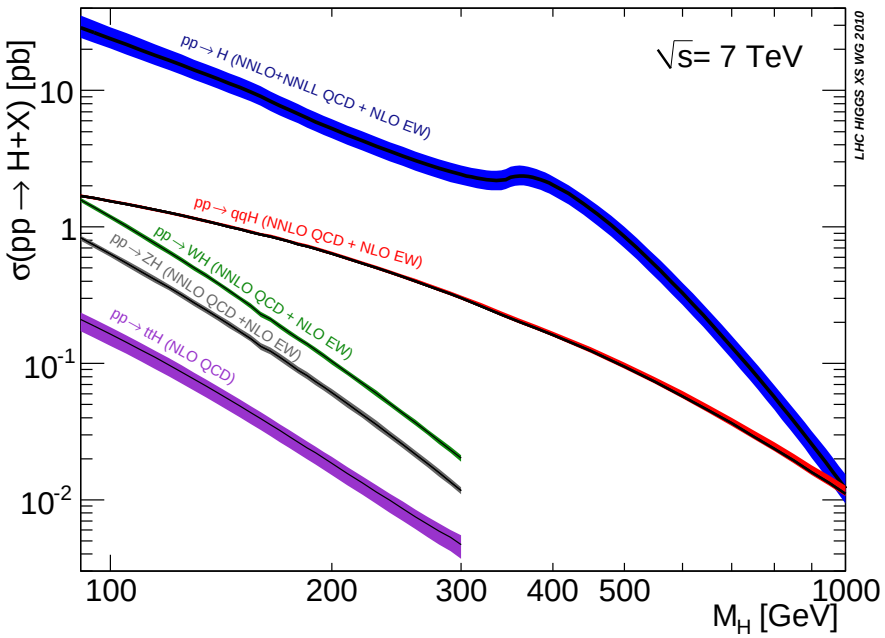
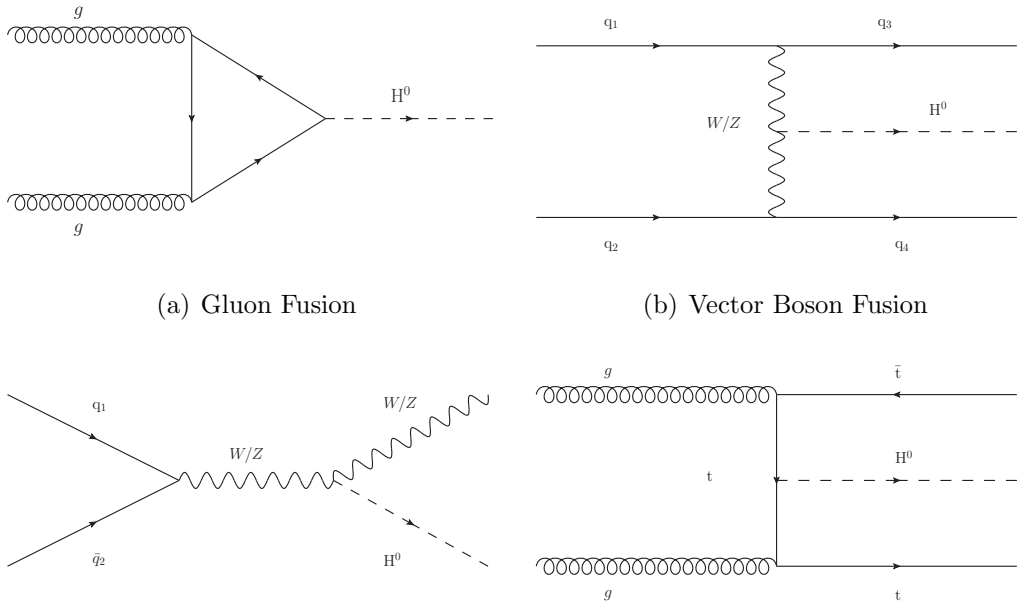


Figure 1-2: LHC Higgs boson production cross sections at a center of mass energy of 7 TeV [44]. The theoretical uncertainties are shown in the colored bands.



(a) Gluon Fusion

(b) Vector Boson Fusion

(c) Higgs-strahlung

(d) $t\bar{t}$ Higgs

Figure 1-3: Feynman diagrams for the main production mechanisms of the Higgs boson at the LHC.

Once the Higgs boson is produced it will decay to a pair of W bosons, a pair of Z bosons, or a pair of massive fermions at tree level. It can also decay, via radiative loop diagrams, to a pair of photons, a pair of gluons, or a photon and a Z boson [45–48]. The branching ratios to the most relevant decay modes are shown in Figure 1-4. From Equations 1.22 and 1.23, we observe the fact that the coupling of the Higgs boson to the electroweak bosons is quadratic in the mass of the electroweak bosons, while the coupling of the Higgs boson to the fermions is linear in the mass of the fermions. This property has the important consequence that while the available phase space for the Higgs decay to two electroweak bosons becomes smaller as the mass of the Higgs boson decreases below $2 \times m_{W \text{ or } Z}$, the larger coupling ensures that the branching ratio remains relatively high all the way down to known exclusion limit at about 114.4 GeV/c^2. We note in particular that the branching ratio to a pair of W bosons is the largest for all Higgs boson masses above about 135 GeV/c^2.

At the LHC, the Higgs boson is typically produced with a non-zero transverse

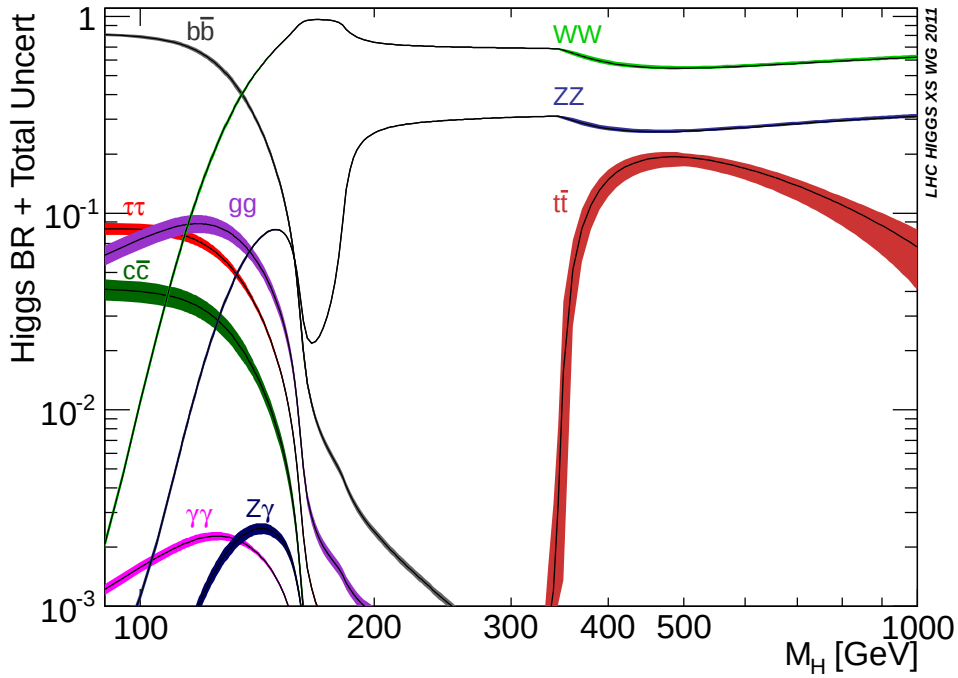


Figure 1-4: Higgs decay branching ratios as a function of its mass [44].

boost, a property that has important implications on the fraction of events where the Higgs boson is produced in association with one or more jets. Due to non-perturbative effects, the differential cross section calculations in fixed orders of α_s fail at low values of Higgs boson p_T. To model the Higgs signal processes, we use the next-to-leading-order (NLO) POWHEG Monte Carlo generator, and reweight the Higgs p_T spectrum to the one obtained from the next-to-next-to-leading order (NNLO) calculation with resummation to the next-to-next-to-leading log (NNLL) term described in reference [49]. The p_T distribution obtained through this procedure is shown in Figure 1-5 for a Higgs boson of mass 160 GeV/c^2 produced via the gluon fusion process.

1.2.1 The $H \rightarrow W^+W^-$ Decay Mode

In this thesis, we search for the Higgs boson in its decay mode to two W bosons which each subsequently decay to a charged lepton and a neutrino. This decay channel is

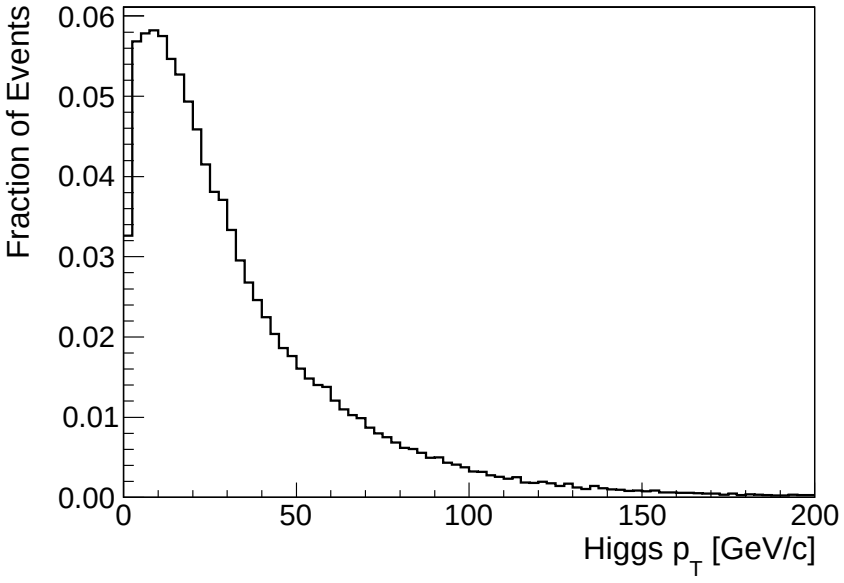
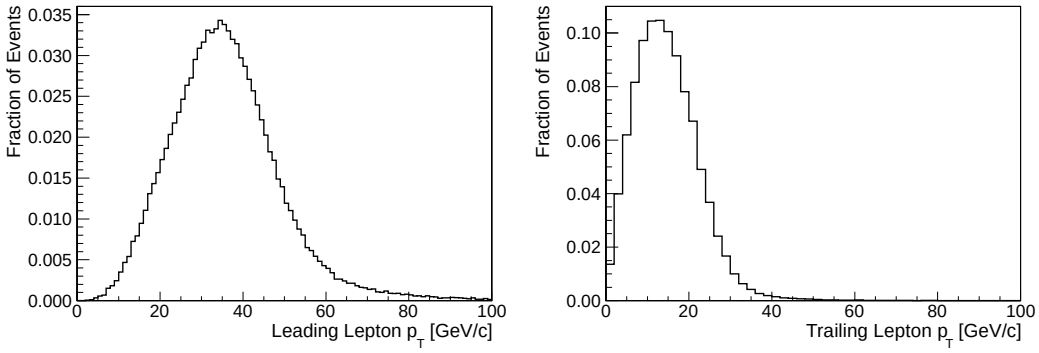


Figure 1-5: The transverse momentum distribution of a Higgs boson with mass of 160 GeV/c^2 produced in the gluon fusion process in shown. These Higgs boson events were generated using the POWHEG Monte Carlo generator [50], and the subsequent Higgs p_T spectrum has been reweighted to the calculation provided by reference [49].

characterized by a relatively large branching ratio over a large range of hypothesized Higgs boson masses. Therefore, it is one of the channels in which signals for the Higgs boson are first expected to be observed.

Higgs bosons with mass below $2M_W$ will decay to one on-shell W, and one off-shell W, typically producing one lepton with fairly large transverse momentum and one lepton with lower transverse momentum. From Figure 1-6, which shows the p_T distribution for the leading lepton and trailing lepton for a Higgs boson of mass 120 GeV/c^2, we observe that a significant fraction of signal events have a trailing lepton with very low transverse momenta. In roughly 30% of signal events, the trailing lepton p_T is below 10 GeV/c and in roughly 75% of signal events, the trailing lepton p_T is below 20 GeV/c. Therefore efficient and effective selection of leptons with transverse momentum as low as 10 GeV/c is very important for extracting low mass Higgs boson signal events.

The two neutrinos from the decay of the W bosons do not interact with any



(a) Leading Lepton (b) Trailing Lepton

Figure 1-6: The transverse momentum distribution of the leading and trailing charged lepton from a Higgs boson produced via the gluon fusion process with a mass of 120 GeV/c^2 and decaying to two W bosons

components of the detector and will simply escape undetected resulting in a large missing transverse energy ($E_{\mathrm{T}}^{\mathrm{miss}}$). The distribution of the missing transverse energy is shown in Figure 1-7 for Higgs boson signal events with a mass of 160 GeV/c^2, indicating typical $E_{\mathrm{T}}^{\mathrm{miss}}$ around 50 to 60 GeV/c. Finally, due to the fact that the Higgs boson is a spin 0 resonance, the angle between the two leptons from the decay of the W bosons in the transverse plane is preferentially small [51]. The distribution of this angle is shown in Figure 1-8. Later, these kinematical features will be used to discriminate Higgs boson signal against backgrounds.

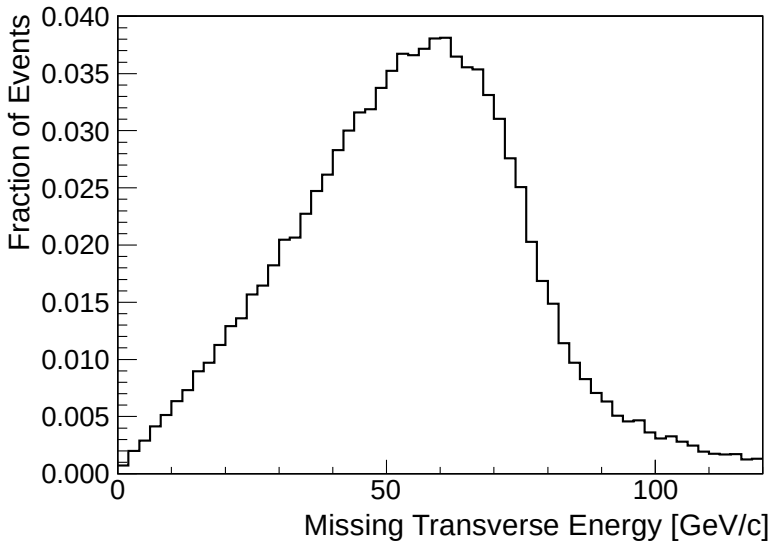


Figure 1-7: The missing transverse energy distribution from a Higgs boson produced via the gluon fusion process with a mass of 160 GeV/c^2 and decaying to two W bosons

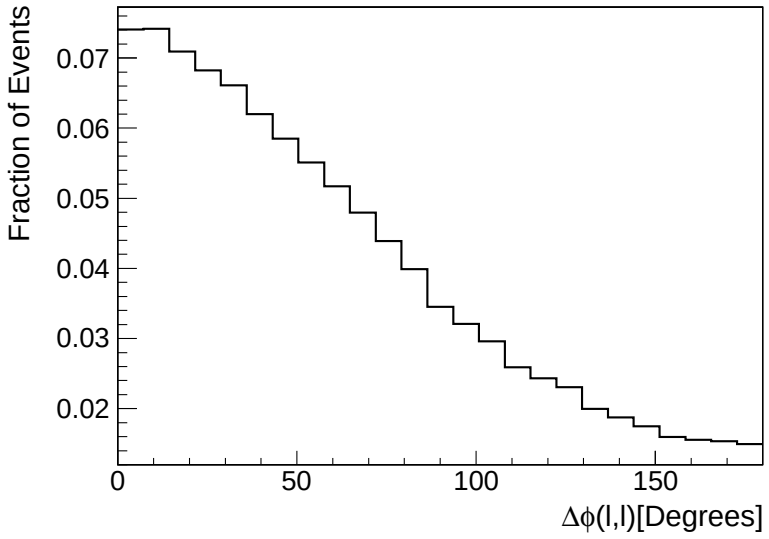


Figure 1-8: The distribution of the angle between the two leptons in the azimuthal plane for Higgs boson events produced via the gluon fusion process with a mass of 160 GeV/c^2 and decaying to two W bosons

Chapter 2

Experimental Apparatus

2.1 The Large Hadron Collider

The main technology that enables the study of particle collisions at the highest energies achieved is provided by the Large Hadron Collider (LHC) accelerator complex [25] located at CERN on the border between Switzerland and France. The accelerator complex is composed of various smaller accelerators that deliver two particle beams running in opposite directions at increasing energies until they reach the main Large Hadron Collider accelerator ring with a circumference of 26.7 km. A schematic diagram of the accelerator complex is shown in Figure 2-1. The main LHC accelerator is composed of 1232 superconducting dipole magnets whose design field reaches 8.33 T, allowing for a maximum center of mass collision energy of 14 TeV. Roughly 7000 additional magnets are employed for optical focusing and trajectory corrections. The instantaneous design luminosity of $10^{34}\text{cm}^{-2}\text{s}^{-1}$ is expected to be reached with beams consisting of 2808 bunches with 10^{11} protons per bunch. The instantaneous luminosity is given by

$$L = \frac{N_p^2 n_b f_{\text{rev}} \gamma_r}{4\pi \epsilon_n \beta^*} F, \tag{2.1}$$

where N_p is the number of protons per bunch, n_b is the number of bunches per beam, f_{rev} is the revolution frequency, γ_r is the relativistic gamma factor, ϵ_n is the normalized transverse beam emittance, and β^* is the value of the beta function at the collision point which relates to the transverse size of the beams at the interaction point, and F is the geometric factor due to the crossing angle of the two beams. The beams are approximately Gaussian in profile with a nominal width of about 16 μm, and are made to collide at the two primary interaction points where the two main experiments, ATLAS and CMS, are located. The nominal collision rate at the interaction point is once every 25 ns. Due to beam safety concerns, the center of mass energy of the LHC during most of its operation in 2010 and 2011 was 7 TeV. At the peak performance in 2011, the LHC delivered an instantaneous luminosity of about 3.5×10^{33} cm^{-2}s^{-1}, with a maximum of 1380 bunches colliding once every 50 ns, 1.4×10^{11} protons per bunch, and a beam width between 25 μm and 30 μm. The two beams collided at a crossing angle of roughly 200 μrad.

To create the two particle beams, protons are first accelerated in the Linac2 linear accelerator and then transferred to the Proton Synchrotron Booster where they are accelerated to 1.4 GeV/c. They are then transferred to the Proton Synchrotron (PS) ring, arranged into bunches spaced 25 ns or 50 ns apart, and accelerated to 25 GeV/c. The proton beams are then transported to the Super Proton Synchroton (SPS), where they are accelerated to 450 GeV/c and finally injected into the LHC. Eight radio frequency (RF) resonating cavities are responsible for accelerating the proton beam to the final center of mass energy through a field gradient of 5.5 MV/m increasing the energy of the beams by 16 MeV per turn.

To reach the design magnetic field of 8.33 T, a nominal current of 12 kA must be passed through the dipole magnet coils. To achieve this, the Nb-Ti alloy coil wires are cooled to 1.9 K using superfluid helium, at which point they enter a superconducting state. The magnet coils and the corresponding beam pipes are embedded in a common cryostat and enclosed in the same mechanical structure, shown in Figure 2-2. After the accident in 2008 [52], where a significant number of magnets were damaged due to

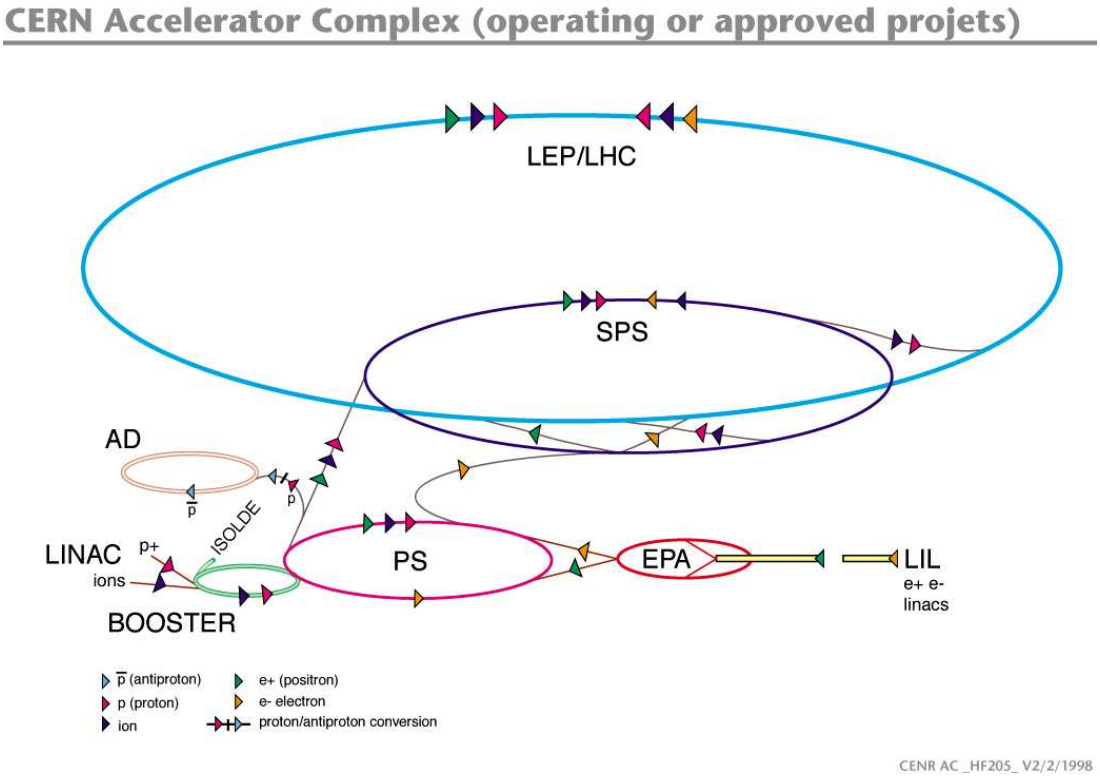


Figure 2-1: A schematic diagram of the LHC accelerator complex is shown. Protons are first accelerated in the linear accelerator (LINAC), and transferred to the Booster where they are accelerated to a kinetic energy of 1.4 GeV. Next, they enter the Proton Synchrotron (PS) ring, arranged into bunches, and accelerated to 25 GeV. Then, they are transported to the Super Proton Synchroton (SPS), where they are accelerated to 450 GeV and finally injected into the LHC.

faulty energy dissipation after a magnet quench, a state of the art quench protection system was installed, which is intended to protect the magnets under a similar re-occurrence.

The high proton bunch density combined with the strong focusing at the interaction point, results in a large probability to have multiple inelastic proton-proton interactions per bunch crossing. These multiple simultaneous inelastic collisions are referred to as "pileup". As the beam optics and proton bunch densities were adjusted during the 2011 run, the distribution of the number of pileup interactions per bunch crossing also evolved with time. This is illustrated in Figure 2-3, where a comparison is made between the first half of the 2011 run, the second half of the 2011 run, and

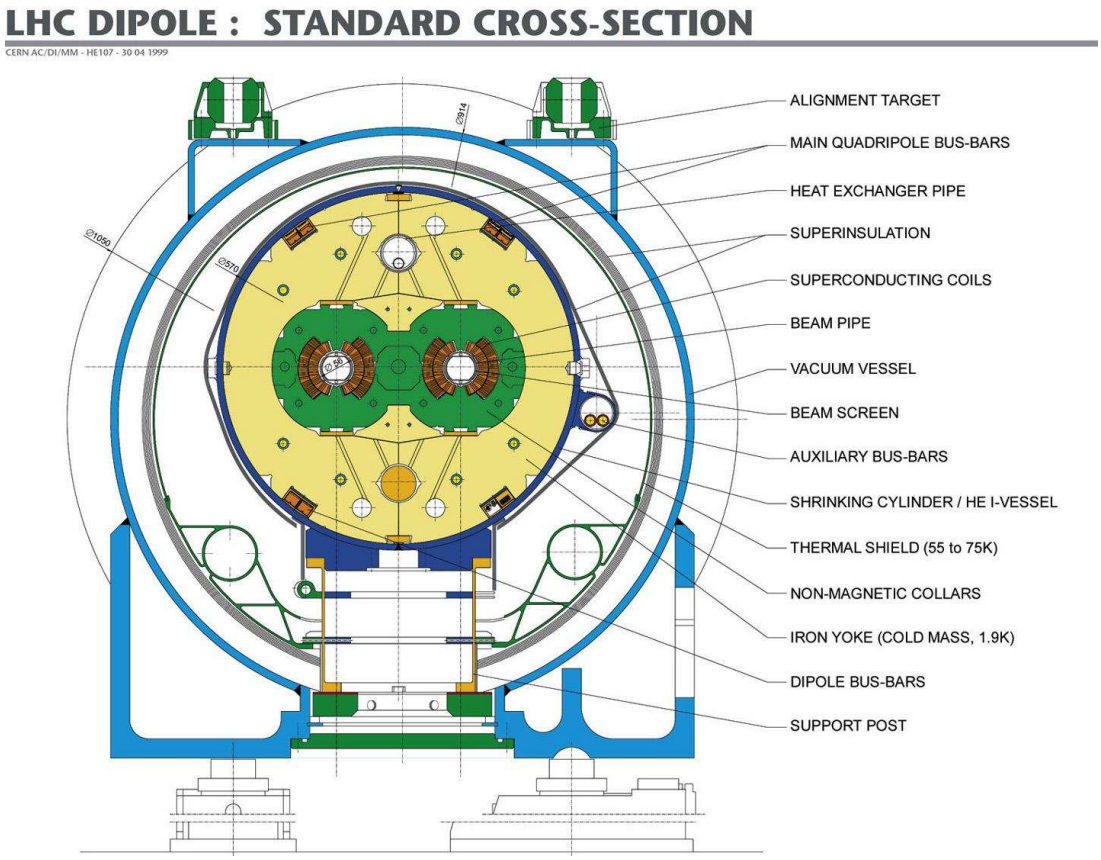


Figure 2-2: A schematic diagram of the LHC dipole magnets and cryostat structure

the average. It is important to properly model the effect of these pileup interactions, and the distribution of the number of pileup interactions.

2.2 The CMS Detector

The Compact Muon Solenoid (CMS) detector is one of the two large general purpose particle physics detectors installed at opposite sides of the LHC accelerator rings. CMS is located at interaction point 5, while ATLAS is located at interaction point 1. CMS consists of a silicon based tracking system, a homogeneous crystal electromagnetic calorimeter, a sampling hadronic calorimeter, and muon chambers behind the calorimeters [53]. For illustration, a transverse slice of the CMS detector is shown in Figure 2-4. A photograph of the detector during the installation phase [54] is

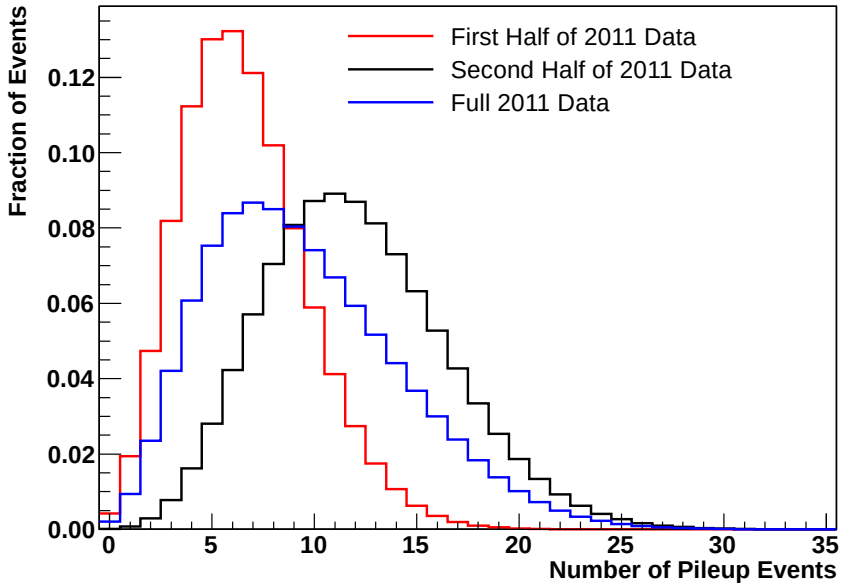


Figure 2-3: The distribution of the number of pileup interactions per bunch crossing for different collision periods in 2011

shown in Figure 2-5, illustrating the large physical size of the detector. The main distinctive feature of the CMS detector is the superconducting solenoid magnet, able to generate a nominal magnetic field of 3.8 T, allowing for high precision momentum measurements for high energy charged particles.

The geometric position at which the main interaction point is expected to occur is designated as the origin of the CMS coordinate system. The x coordinate axis points horizontally towards the center of the accelerator ring, and the y coordinate axis points straight up. The z coordinate axis is tangential to the beam, and the direction is the one consistent with a right handed coordinate system. This Cartesian coordinate system is transformed into a cylindrical coordinate system defined by the radial distance r, the azimuthal angle ϕ, and the z coordinate, or a spherical coordinate system defined by the radial distance ρ, an azimuthal angle ϕ and a polar angle θ which begins in the positive direction of the z axis and proceeds counter-clockwise. The plane transverse to the beam line is also referred to as the $r - \phi$

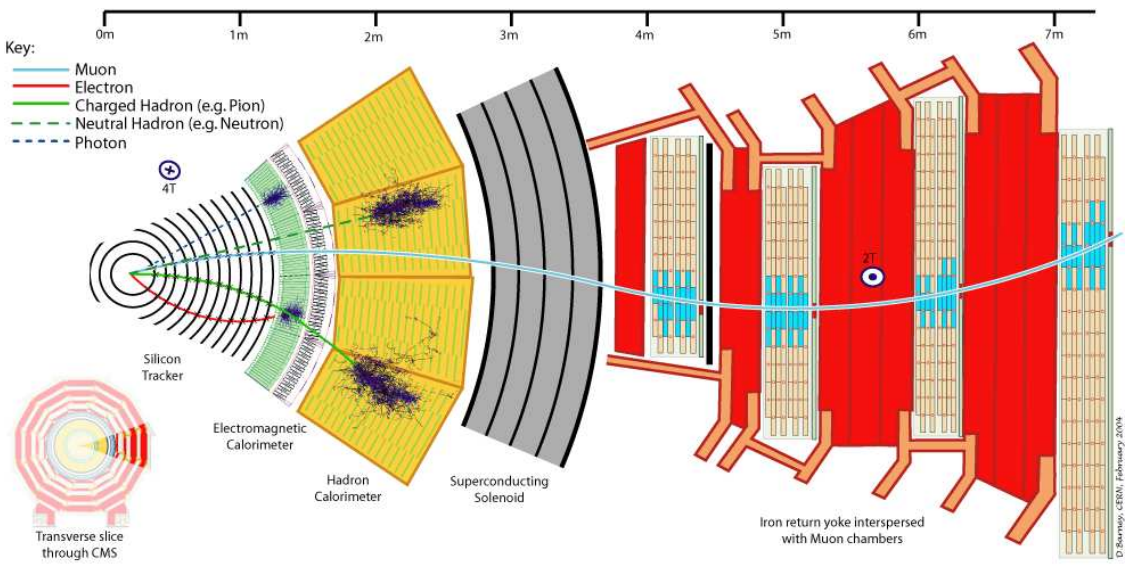


Figure 2-4: A transverse slice of the CMS detector is shown. Various components of the CMS detector are shown as well as graphical representations of the response of the detector to the passage of different types of particles.

plane.

We define the pseudorapidity η as an alternative to the polar angle by:

$$\eta = -\log\left(\tan\left(\frac{\theta}{2}\right)\right), \tag{2.2}$$

which is the massless limit of the rapidity y of a particle, defined by:

$$y = \log\left(\frac{E + p_Z}{E - p_z}\right), \tag{2.3}$$

where E and p_Z are the energy and z component of the momentum of the particle. The pseudorapidity is preferred over the polar angle due to the fact that the rate of particle production at hadron colliders are approximately constant as a function of η, and due to the Lorentz invariance of rapidity differences.

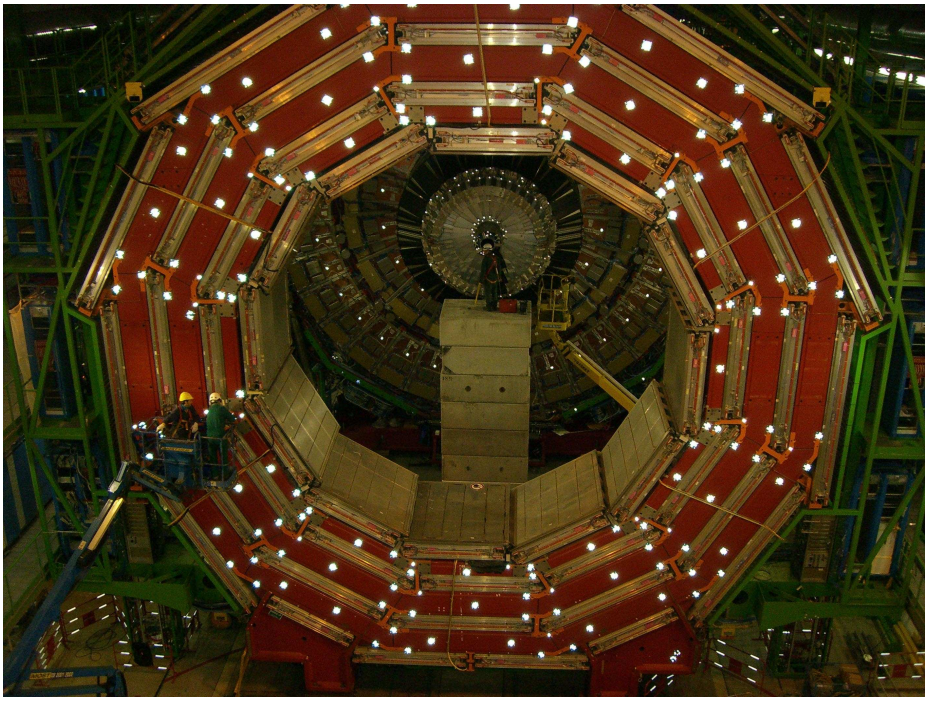


Figure 2-5: A photograph of the CMS detector during the installation phase

2.2.1 Tracking System

The tracking system at CMS is entirely based on silicon tracking detector technology. The tracking system has been divided into three main regions, motivated by the requirement of low occupancy for robust track reconstruction and the varying particle flux in the three different detector regions. A complete layout of the tracking system is shown in Figure 2-6.

In the region closest to the beam axis, within 10 cm of the beam line, silicon pixel detectors are used. Each pixel has dimensions of 100 μm by 150 μm, keeping occupancy low at 10^{-4} per LHC bunch crossing. In the barrel, they are arranged in 3 cylindrical layers situated at 4.4 cm, 7.3 cm, and 10.2 cm away from the beam line respectively. In the endcaps, they are arranged in four disks, two on each side, at a distance of 34.5 cm and 46.5 cm from the nominal vertex position, and cover the region between 6 cm and 15 cm from the beam axis. The pixel detectors are able to attain a spatial resolution of 10 μm in the $r - \phi$ coordinate measurement, and 20 μm

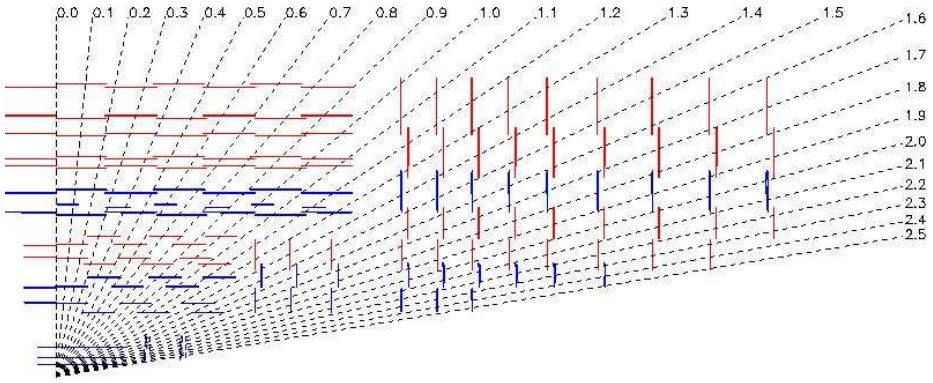


Figure 2-6: The tracker layout showing one quarter of the detector in the $r - z$ view. The horizontal lines represent the barrel layers of the silicon tracking system, while the vertical lines represent the endcap layers. The interaction point is located at the lower left edge of the diagram. The three horizontal and four vertical lines closest to the interaction point represent the pixel detector, while the remaining lines represent the silicon strip tracking system. The dotted lines and the associated number indicate the value of the pseudorapidity at the given angle.

in the z coordinate measurement.

The remaining parts of the tracking system consist of silicon strip detectors. In the region at a radial distance between 20 cm and 55 cm from the beamline, the particle flux remains high enough that smaller sized silicon strip modules must be used, while in the region beyond, larger sized modules can be used while keeping roughly the same occupancy. The silicon strip tracker is divided into four main parts. In the barrel region of the detector, it is divided into the Tracker Inner Barrel (TIB) and the Tracker Outer Barrel (TOB). The TIB contain four layers of sensors with a thickness of 320 μm and a strip pitch width of between 80 μm and 120 μm. It covers a region that ranges 65 cm on either side of the nominal collision vertex position. The TOB consists of six layers of silicon strip modules covering 110 cm on either side of the nominal vertex position. The sensors are thicker with a thickness of 500 μm, and have a wider pitch width that lie between 120 μm and 180 μm. The first two layers of the TIB and TOB are stereo layers that lie at an angle of 100 mrad with respect to the other layers and provide both $r - \phi$ and $r - z$ coordinate measurements. The single point resolution achievable for the TIB is $16 - 28$ μm in $r - \phi$ and 230 μm in

z, and for the TOB is $25 - 41$ μm in $r - \phi$ and 530 μm in z.

In the endcap regions, the silicon strip tracker is divided into the Tracker End Cap (TEC) and the Tracker Inner Disk (TID). The TEC consists of nine disks located at a distance from the nominal vertex between 120 cm and 280 cm. The two inner most rings and the fifth ring are stereo modules which provide measurements in the $r - \phi$ and z coordinates. The three innermost rings have sensors of thickness of 320 μm, while the rest have a thickness of 500 μm. The TID are three small disks that cover the gap region between the TIB and the TEC. They have a sensor thickness of 320 μm with the first two rings providing stereo measurements.

A total of 15400 modules of the silicon tracking system are finally mounted on carbon fiber structures inside of outer support tubes providing strict temperature control. In order to suppress the effects of radiation damage by limiting the mobility of structural defects in the silicon lattice, the entire system should ideally operate at a nominal temperature of $-20°$C. Due to the large amount of structural support as well as active detector material that is required by the silicon tracking system, the amount of material that a charge particle must traverse inside of the active volume of the tracker is large relative to previous experiments. The current best estimate of the material budget is depicted in Figure 2-7, showing an increase of the amount of material from $0.4X_0$, where X_0 is one radiation length, at the center of the detector to $1.9X_0$ at an $|\eta|$ of 1.6, subsequently decreasing to $1.0X_0$ at an $|\eta|$ of 2.5. An estimate of the tracking resolution for the transverse momentum, transverse impact parameter, and longitudinal impact parameter measurements are shown in Figure 2-8.

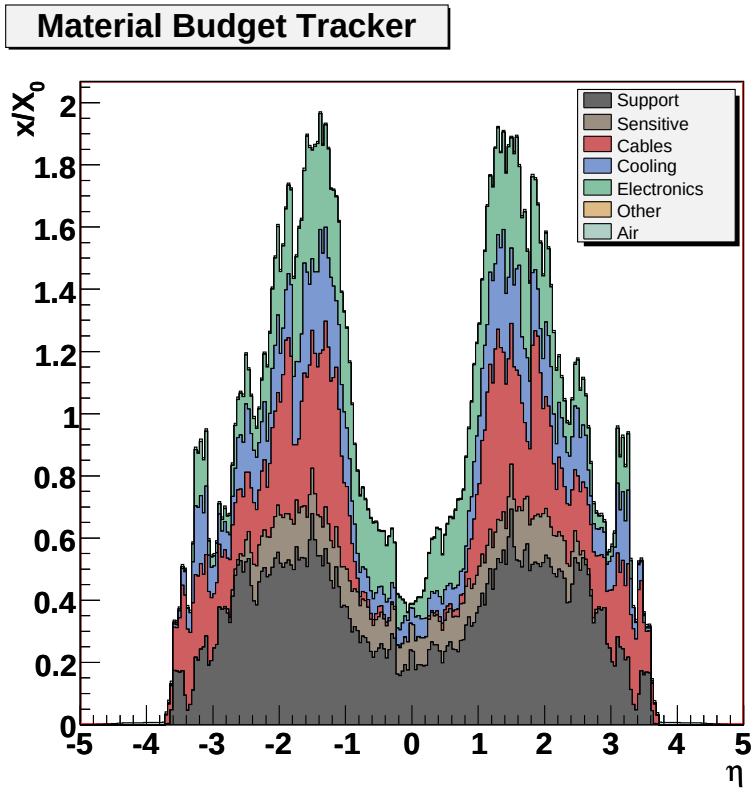
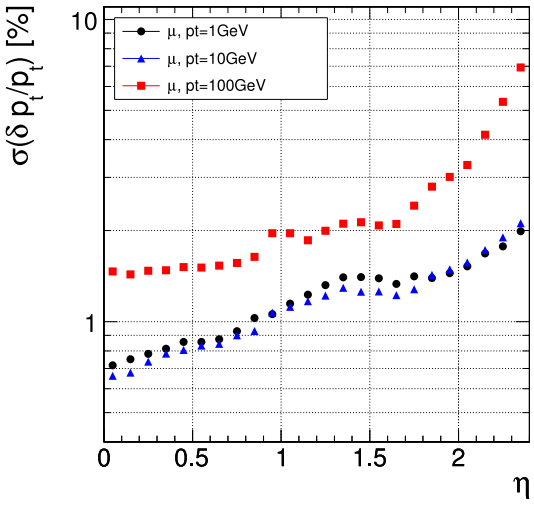
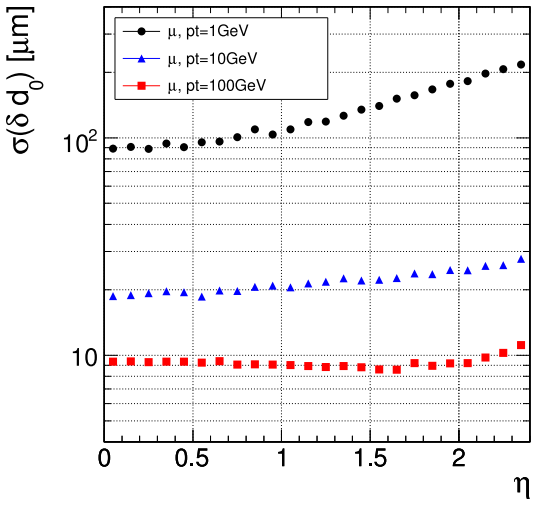


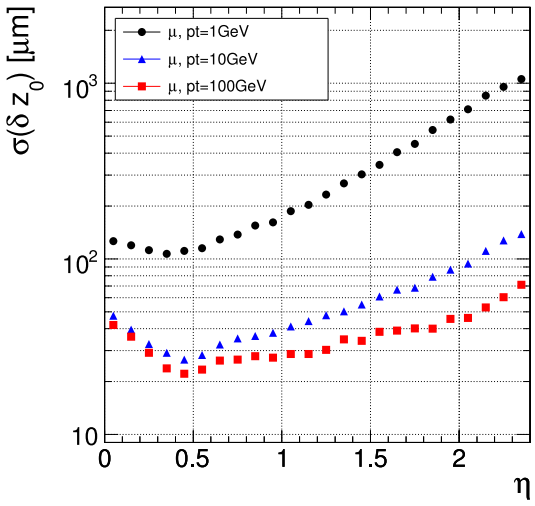
Figure 2-7: The tracker material budget as a function of pseudorapidity displayed in units of radiation lengths from simulation



(a) p_T Resolution



(b) d_0 Resolution



(c) d_z Resolution

Figure 2-8: Resolution of the transverse momentum (a), transverse impact parameter (b), and longitudinal impact parameter (c) measurement for single muons of transverse momentum 1, 10, and 100 GeV/c citeCMSTDR.

2.2.2 Calorimetry System

The calorimetry system at CMS has been designed with the goal of hermiticity and precision energy measurement in mind. The electromagnetic calorimeter (ECAL) is a fully hermetic homogeneous crystal calorimeter, while the hadronic calorimeter system (HCAL) is a sampling calorimeter consisting of brass absorber material and quartz fiber as active material. The combination yields a calorimetry system with superb electromagnetic energy resolution and fast response time that are demanded by the CMS physics goals and the LHC collision environment.

The electromagnetic calorimeter consists of 61200 lead tungstate ($PbWO_4$) crystals in the barrel region, and 7324 crystals in each of the two endcap regions. The crystals are scintillating with a short radiation length of 0.89 cm and small Moliere radius of 2.2 cm. They also have very fast response, with 80% of the scintillation light emitted within one bunch crossing period of 25 ns. The low light yield of the scintillating crystals requires the use of photo-detectors able to give high gain within an environment of strong magnetic fields. Silicon avalanche photodiodes (APDs) are used in the barrel region, while vacuum phototriodes (VPTs) are used in the endcap region. These detectors and the crystal scintillators demand temperature stability due to the sensitivity of their response to changes in operating temperature.

The front face of each crystal in the barrel has dimensions of 22×22 mm^2 and a length of 230 mm, corresponding to a total of 25.8 radiation lengths. Each covers 0.0174 in η and ϕ. The barrel system covers the region up to an $|\eta|$ of 1.479. The endcap system is located at a distance of 314 cm from the primary interaction point and covers a region in $|\eta|$ from 1.479 to 3.0. The crystals in the endcap have front face dimensions of 28.6×28.6 mm^2 and a length of 220 mm, corresponding to a total of 24.7 radiation lengths. Furthermore, a preshower detector is placed in front of the endcap electromagnetic calorimeter to help in the identification of neutral pions. It consists of two disks of lead absorber material located at $2X_0$ and $3X_0$ respectively, and two planes of silicon strip detectors as active layers. A schematic diagram showing all three components is given in Figure 2-9. The electromagnetic calorimeter has undergone

calibration using data collected from a high energy electron test beam, showing a resolution of

$$\sigma_E/E = \frac{2.8\%}{\sqrt{E[\text{GeV}]}} \oplus \frac{12\%}{E[\text{GeV}]} \oplus 0.3\%. \tag{2.4}$$

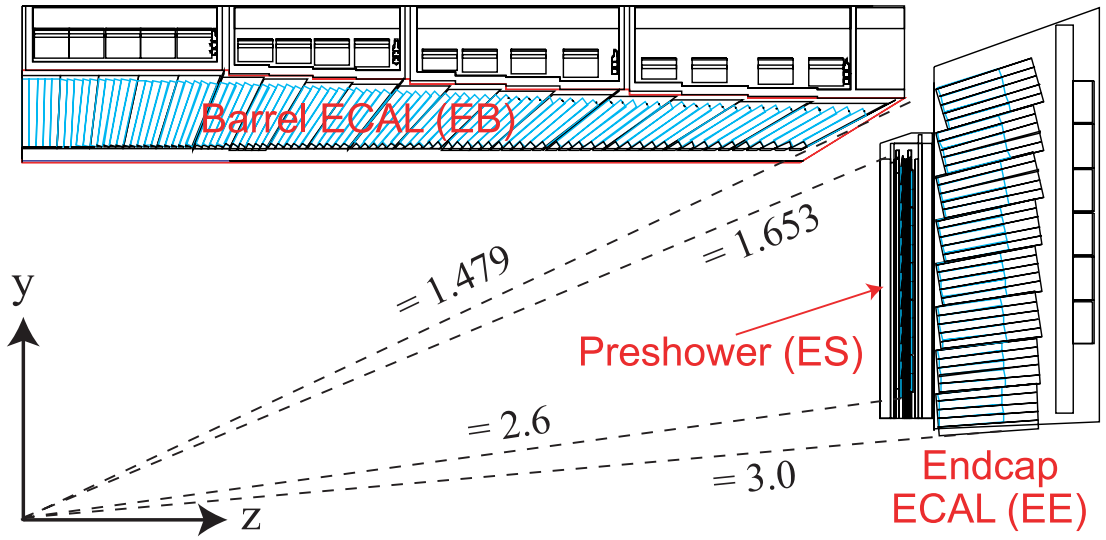


Figure 2-9: A schematic cross sectional diagram of the ECAL, showing the geometrical arrangement of the barrel ECAL, the endcap ECAL, and the preshower detector. The interaction point is located at the lower left edge of the diagram. The dotted lines show the values of the pseudorapidity at the given angle.

The majority of the hadronic calorimeter lies inside the solenoid magnet coils and therefore the ability to operate within a high magnetic field environment was an important design consideration. Brass was chosen as the absorber material due to its relatively short nuclear interaction length and non-magnetic properties. The active material consists of plastic scintillator tiles 3.7 mm thick. The light signal is transmitted via wavelength shifting fibers to hybrid photodiode detectors. A schematic diagram of the HCAL is shown in Figure 2-10.

The barrel region of the HCAL detector consists of 2304 individual towers arranged in a projective geometry lying between η of -1.4 and 1.4, each covering an area 0.087 $\times$ 0.087 in η and ϕ. There are a total of 15 brass absorber plates, each 5 cm thick, while the innermost and outermost layer is made of stainless steel for

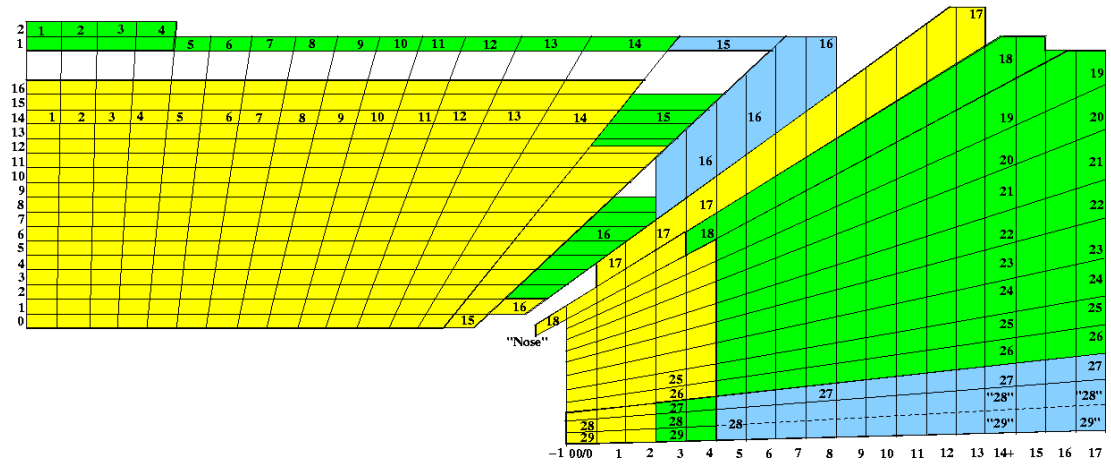


Figure 2-10: A schematic diagram showing the tower arrangements in the HCAL barrel and endcap regions. The barrel HB detector runs from tower 1 through tower 15, and overlap with the endcap HE detector in tower 16. The HE runs from tower 16 through tower 28. Towers 29 and above are part of the HF forward hadron calorimeter.

increased structural strength. All except the innermost and outermost layers of the 17 active scintillator layers are 3.7 mm thick. The outer active layers are 9 mm thick, optimized to sample low energy showering particles from the interaction with support material lying between the ECAL and HCAL. The endcap region of the HCAL covers the region in $|\eta|$ from 1.3 to 3.0. The towers are segmented to 5° in ϕ and 0.087 in eta for the five outermost towers that are smallest in η and to 10° in ϕ and between 0.09 and 0.35 in eta in the eight innermost towers at largest η. There are 19 active scintillator layers of thickness 3.7 mm, between absorber plates of thickness 78 mm.

In order to improve the energy and missing energy resolution, additional layers of scintillators comprising the hadron outer calorimeter system (HO) are installed just outside of the magnet coil in order to improve the measurement of the tails of the distributions. It is separated into five rings covering the eta region between -1.26 and 1.26. The central ring has two scintillator layers at a radial distance of 3.850 m and 4.097 m from the beam axis, on either side of a layer of iron absorber of thickness 18 cm, while the other four rings consists of only one scintillator layer at a radial distance of 4.097 m from the beam axis. These calorimeter systems combined give

between 11 and 12 nuclear interaction lengths (λ_I) depending on the particular region of the detector.

Finally, the very forward region between 3.0 and 5.0 in $|\eta|$ contains the Hadron Forward (HF) calorimeter system. It consists of steel absorbers of depth 1.65 m and active quartz fiber material. Cerenkov light is emitted as charged particles travel through the quartz fibers, which are channeled by light guiding fibers to photomultiplier tubes. The quartz fibers have a diameter of 0.6 mm and are inserted into 1 mm square grooves cut out of the steel plates, each placed 5 mm apart in a square grid. These quartz fibers have alternating lengths of 1.32 m and 1.65 m resulting in samplings of two different longitudinal lengths. The HF is segmented into towers of 0.175 in η and 10° in ϕ.

Results from test beam data using high energy pion, electron and muon beams show an energy resolution of 120% for the stochastic term and 6.9% for the constant term.

2.2.3 Muon System

The mass of the muon is 200 times larger than that of the electron, and therefore the rate of radiation by the bremsstrahlung process is a factor of 40000 smaller as the rate of bremsstrahlung is inversely proportional to the mass squared. As a result a muon is able to traverse the calorimeter material with minimal energy loss. This property implies that placing charged particle detectors behind the calorimeter volume is an effective method to detect muons.

In CMS, the muon spectrometer system can be categorized into three different components. In the barrel region of the detector, where the rate is relatively low, muons are identified by a system of drift tube chambers (DT) consisting of 4 stations positioned at different distances from the beam axis arranged into 5 wheels running along the length of the detector. In the endcap region, where the rate is relatively higher, there is a system of cathode strip chambers arranged in concentric disks centered on the beam axis. Finally, associated with each of these muon stations are a

set of resistive plate chambers which give excellent time resolution and are important for the operation of the online muon triggering system. The layout of one quarter of the CMS muon systems is shown in Figure 2-11.

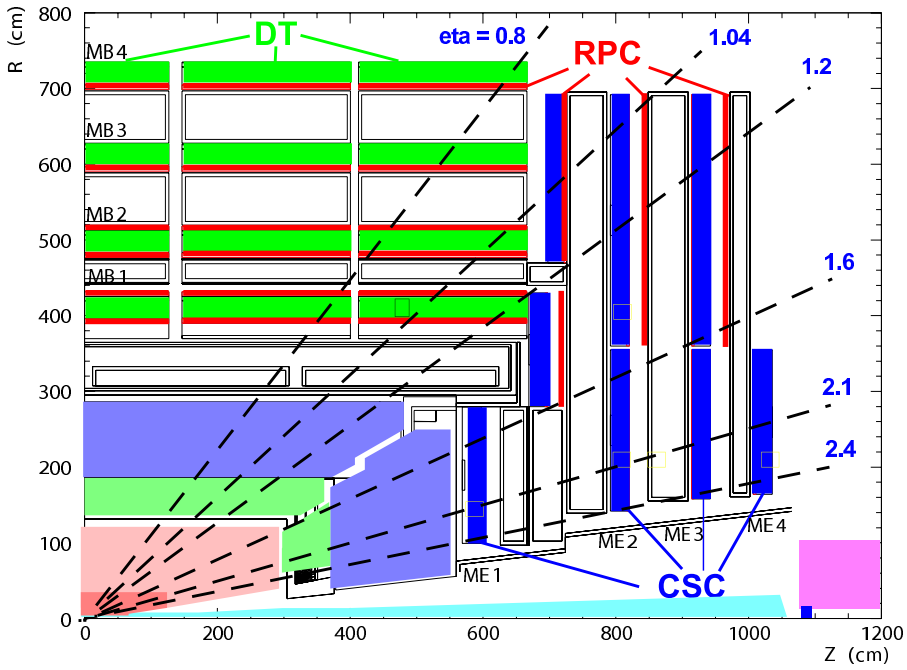


Figure 2-11: The layout of one quarter of the CMS muon systems in $r - z$ view. The drift tube (DT) muon stations are shown in green and the cathode strip chamber (CSC) muon stations are shown in blue. The resistive plate chambers (RPC) detectors are shown in red. The interaction point is located at the lower left edge of the diagram.

The drift tube system consists of a total of 250 individual chambers arranged in stations of four layers placed inside the return yoke of the magnet. The layers are arranged into five wheels placed longitudinally along the beam axis. Each wheel is separated into 12 sections each covering roughly 30 degrees in ϕ. The chambers are placed in a staggered position with respect to other layers to avoid gaps in the active region of the detector. Within the three stations closest to the beam axis, there are three superlayers each consisting of 4 planes of drift tubes. The outermost superlayers contain $r - \phi$ measuring planes, whose wires run parallel to the beamline, while the central superlayer contain $r - z$ measuring planes with wires running perpendicular

to the beamline. Figure 2-12 shows a cross sectional diagram of a drift tube chamber. The layer furthest away from the beam axis only includes the two $r-\phi$ measuring superlayers and does not include the $r-z$ measuring one. The maximum drift length within a single drift tube cell is 2.0 cm. The design resolution for single point measurements is 200 μm and 1 mrad for angular measurements.

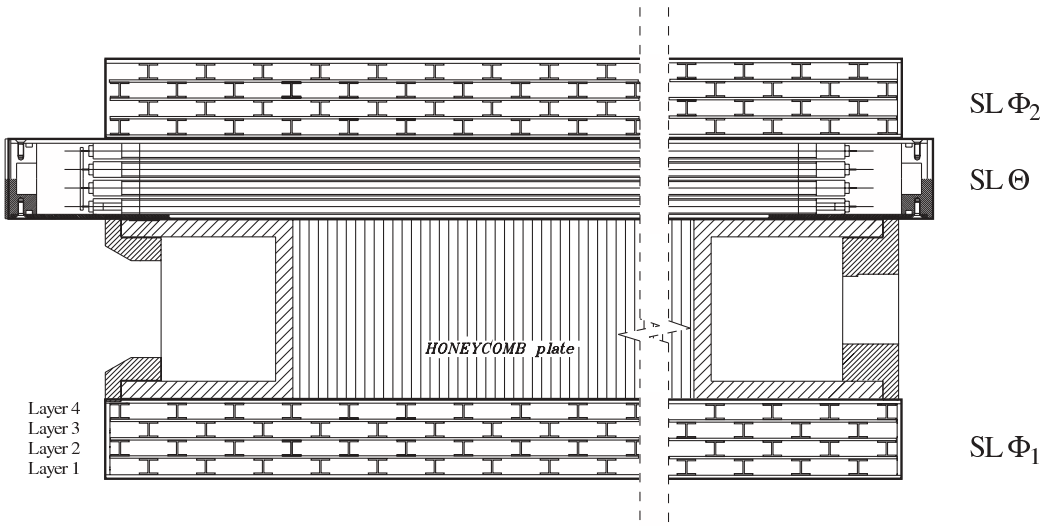


Figure 2-12: The layout of a drift tube (DT) chamber inside a muon station. Three superlayers are shown, each consisting of four layers of drift tubes. The first and third superlayer measure the $r-\phi$ coordinate with drift tube wires running parallel to the beamline. The middle superlayer measures the $r-z$ coordinate with drift tube wires running perpendicular to the beamline. A honeycomb support structure separates the first superlayer with the second and third superlayers, providing a lever arm of approximately 28 cm for local momentum measurements.

The endcap muon system consists of a total of 468 cathode strip chambers arranged in concentric rings around the beam axis. There are 36 chambers in each of the rings except for the inner most ring of the second to fourth disks which have only 18 chambers. Each chamber is trapezoidal in shape and consists of 6 gas gaps each, illustrated in Figure 2-13. The cathode strip chamber is a multiwire proportional chamber where one cathode plane has been segmented into strips that run perpendicular to the wires. A charged particle traveling through the chamber induces ionization electrons within the gas gap. The avalanche which subsequently develops on the wire

induces a well known charge distribution on the cathode plane from which it is possible to reconstruct the track's position along the wire. The procedure is graphically represented in Figure 2-14. This allows a simultaneous measurement of two coordinates with a single plane of chambers using the signal on the wire and the signal from the strips. The signal on the wire is fast and is used for the triggering decision, while the signal on the cathode strips is combined later to improve position measurement resolution. A spatial resolution of 200 μm and an angular resolution of 10 mrad is achieved. As for the barrel muon system, almost all of the cathode strip chambers are overlapping in ϕ to minimize gaps in the instrumented region.

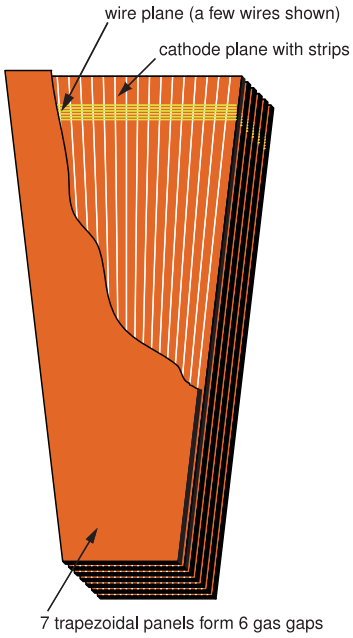


Figure 2-13: A schematic diagram of the CSC chamber is shown. Each chamber has six gas gaps containing one plane of radial cathode strips and one plane of anode wires running approximately perpendicular to the strips.

There is at least one resistive plate chamber associated with each of the muon stations in the barrel and endcap. Due to their fast response, they are essential for the operation of the muon triggering system in identifying which bunch crossing a particular measured muon originated from. The resistive plate chambers are double-gap Bakelite chambers with a gap width of 2 mm, which operates in avalanche mode.

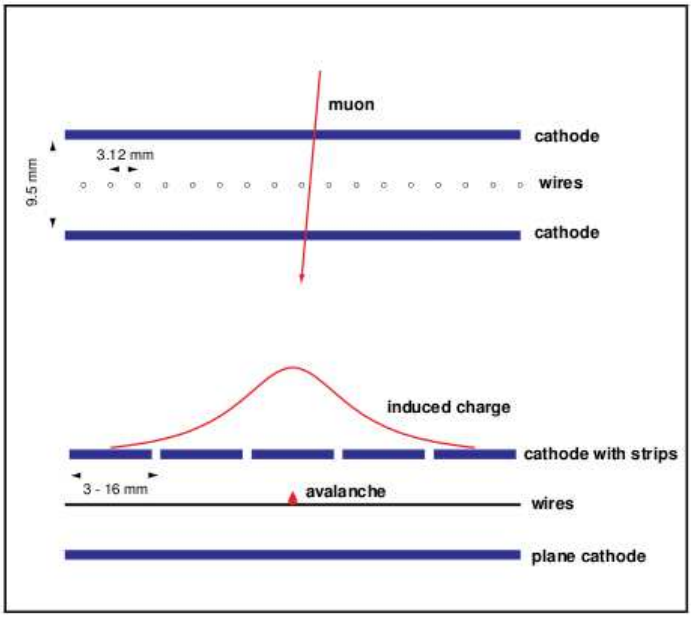


Figure 2-14: A diagram illustrating the operational principle of the CSC position measurement. The coordinate along the wire can be inferred from the induced charge distribution on the cathode.

A schematic diagram can be found in Figure 2-15. A charged particle traversing the gas gaps produces ionization electrons which induces an avalanche multiplication. This results in charge build up inside the gap. The charges drift towards the anode, inducing a signal charge in the strip between the two gaps. The use of a double gap enhances the efficiency of the detector and allows for a decrease in the high voltage due to decreased requirement on the amplification gain.

2.2.4 Data Acquisition and Triggering System

An integrated data acquisition (DAQ) and triggering system is in place to record data quickly and efficiently. The bunch crossing rate of the LHC in 2011 reached up to 15 MHz, requiring a rejection factor of about 10^5 in order to reduce the rate to a level consistent with the current highest possible rate of data archival. As a result, the triggering system is of critical importance to the experiment. The rejection is achieved in two stages. The first stage is called the Level-1 trigger (L1) system and consists of

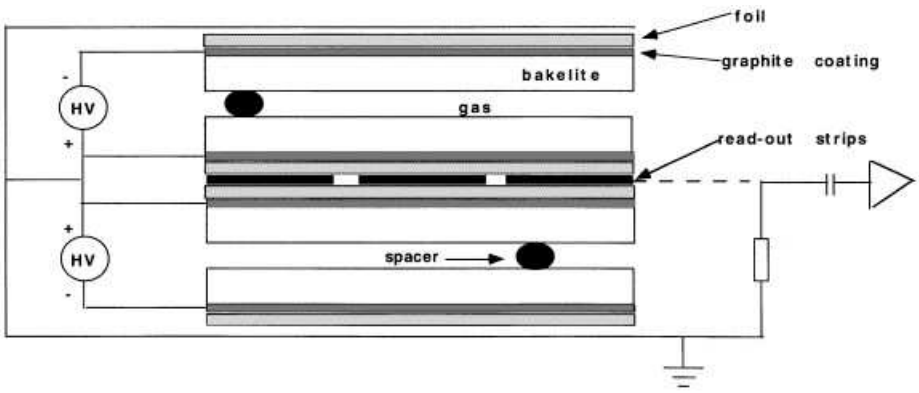


Figure 2-15: A schematic diagram of the RPC double gap chamber. A charged particle traversing the gas gaps produces ionization electrons that induce an avalanche multiplication under high voltage conditions. The build up of charges inside the gap induces a signal charge in the strip between the two gaps.

custom hardware processors whose decisions must be made within microseconds. The second stage is the High-Level trigger (HLT) system which runs more sophisticated reconstruction on a dedicated online computing farm.

The trigger and DAQ system consists of the detector electronics, the Level-1 trigger processing system, the readout network and event building system, and the online High-Level trigger computing farm. Data collected by the various sub-detector components are held in buffers for 3.2 μs, while the signals travel from the detector to the Level-1 trigger logic services and back. Because limiting the latency is critical, the Level-1 trigger decision logic is allocated a latency of 1 μs only. The L1 decisions are made by custom designed hardware processors, and involve the muon and calorimetry systems only. The decision is based on the presence of trigger primitive objects such as electrons, photons, muons, and jets that are formed by the calorimetry and muon system information. Global conditions such as the global summed transverse energy, and the missing transverse energy are used as well. The maximum design L1 trigger accept rate is 100 kHz, limited by the average time it takes to transfer the full detector information through the readout system.

After a L1 accept signal is received, data which were temporarily held in buffers are transferred to the front end readout system. After more signal processing, zero

suppression and data compression, the data are placed in dual port memories for access by the DAQ system. The data for each event are contained in several hundred front end readout buffers and total roughly 0.5 MB in size. The data fragments are collected via the event building network switch system to a single builder unit which combines the fragments into a single event. The event is then passed to one of the trigger filter units that make up the online HLT computing farm. The HLT decisions are based on more sophisticated algorithms with better measurement resolution, more closely resembling the offline event reconstruction. The strategy employed attempts to reject an event as early as possible, thus reducing the processing time. This leads to various internal trigger levels representing the levels of partially reconstructed events, ranging from muon and calorimeter only information, data including the pixel tracker information, and data including the full tracking information. In 2011, the average HLT accept rate was roughly 300 Hz.

Events accepted by the HLT are subsequently sent to the storage manager system. The storage manager serves primarily as an intermediate data buffering system, and monitors and manages transfers of data from the HLT computing farm located on the CMS detector site at interaction point 5 to the main CERN site at Meyrin. Once successfully transferred, the events are sent to the Tier-0 system for permanent storage and distribution.

2.2.5 CMS Computing Model

The CMS experiment is faced with a great challenge in both data storage and computational resources that are needed to achieve the physics goals of the experiment. The data volume that the experiment collects, and the computing resources that are needed to analyze the data are more than an order of magnitude larger than previous high energy physics experiments. The development of grid computing in the past ten years have allowed the choice of distributed computing as a solution to meet the challenge [55]. The CMS computing model partitions the aggregate computing resources into three different tiers that serve specific purposes.

The Tier-0 center consists of a single large computing farm along with a mass storage system situated at the CERN Meyrin site. It is responsible for receiving the raw data from the storage manager system and converting it into the "RAW" data format to be stored permanently on the mass storage system at CERN. It also performs a prompt reconstruction, using the best available information on calibration and alignments at the time. The first copy of the RAW data and the prompt reconstruction is stored at the Tier-0 for custodial storage. A second permanently stored copy is transferred to one pre-designated Tier-1 site for custodial archiving.

The eight Tier-1 centers currently in existence serve primarily two purposes. The first purpose is to archive the custodial copy of the data on their tape mass storage systems. The second purpose is to provide computing resources for reprocessing of the data. Reprocessing occurred fairly frequently in 2011 as improved calibrations were produced, and a fast turnaround was desired. As a result, the Tier-1 sites are almost exclusively used for this purpose and analysis activities are severely restricted.

The roughly fifty Tier-2 centers are primarily dedicated to performing physics analysis. They do not have any tape storage systems and have limited storage space. Roughly 50% of the computing resources of the Tier-2 centers are used for centralized Monte Carlo production, while the remaining resources are for user analysis. The centrally produced Monte Carlo data samples are subsequently transferred to a pre-designated Tier-1 center for custodial storage.

The system responsible for data transfers and data management in CMS is called "PhEDEx", an acronym for Physics Experiment Data Export. It is composed of a collection of collaborating agents, responsible for various aspects of the data replication and data management at each individual computing site, which communicate through a blackboard architecture. This set up ensures that the system as a whole is particularly robust against occurrence of local site failures, a situation that is unavoidable given the large number of independent sites involved. In 2012, the total volume of archived data in CMS is on the order of 100 PB, and the rate of data transfer is on the order of 300 TB per week, illustrated in Figure 2-16. The scale of

these data volumes presents a unique challenge for computing operations, in terms of the number of failures and errors that has to be overcome at any one time, as well as the amount of data that must be cataloged and managed. Throughout the operating period of the experiment, many campaigns were carried out, related to large volume deletion of obsolete data, systematic diagnosis of site-to-site transfer rate and quality, and systematic data file integrity checks, in order to ensure that day-to-day operations remain smooth.

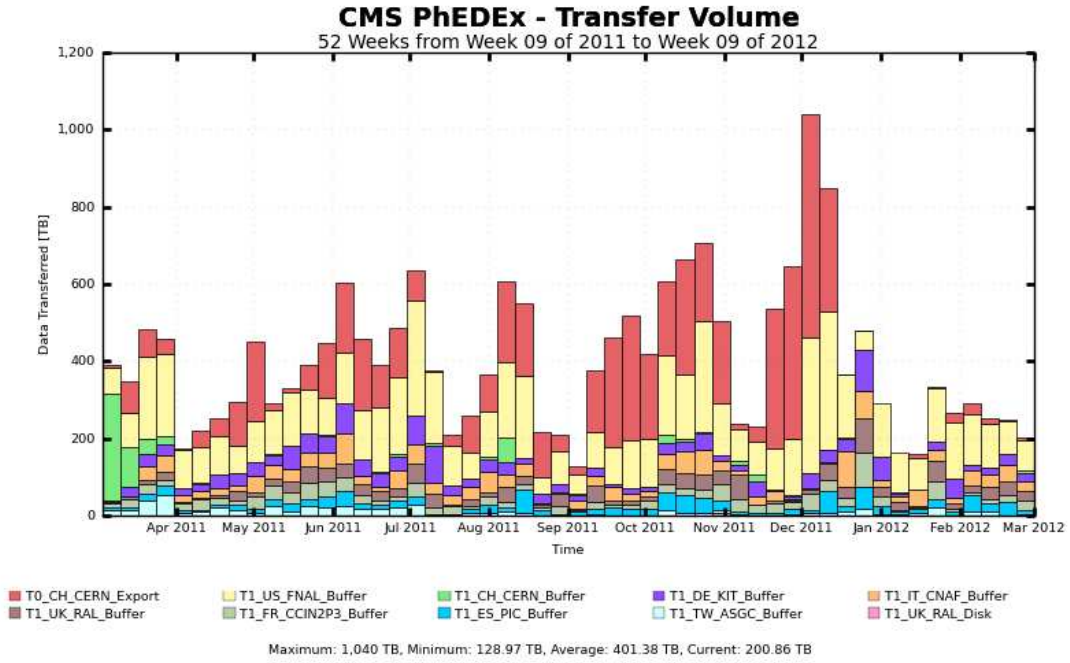


Figure 2-16: The data transfer volume per week in 2011 CMS operations. The transfer volume is categorized by different source sites.

Chapter 3

Trigger Selection and Event Reconstruction

In order to search for the Higgs boson, we must first trigger on and collect the data events containing the Higgs boson signal. For each event that we archive, we must be able to reconstruct all of the objects in order to form a global description of the event. An in-depth understanding of almost all components of the detector and all event objects is required for the Higgs boson search in the W^+W^- decay mode. Many of these analysis features present unique challenges, starting from selection of leptons, to jet selection and b-tagging, and finally to missing transverse energy. These challenges and the specific procedures used to address them are described in detail in this chapter.

3.1 Trigger Selection and Datasets

Events are collected through a sophisticated triggering system, described in Section 2.2.4, and are stored according to their particular properties into different datasets. The events of interest are collected through "trigger paths" defined by the basic objects identified in the event. A trigger path is defined by a sequence of requirements beginning from a seed Level-1 trigger and proceeding through the set of trigger filters

imposing various requirements. Particular "primary datasets" are defined typically by the types of objects that are identified in the events. In this analysis, we are primarily interested in identifying electrons and muons. Electrons are found by matching narrow electromagnetic energy clusters with a consistent track, while muons are found by matching a track with consistent track segments reconstructed in the muon chambers behind the calorimeter volume. We primarily collect events through trigger paths that require the presence of one or two electrons or muons. The number of identified electrons and muons also determine the primary datasets that the events are classified into.

This analysis uses data collected in 2011. The integrated luminosity delivered to CMS in 2011 is plotted as a function of time in Figure 3-1, showing significant improvements in the instantaneous luminosity as we approached the end of 2011. The LHC delivered 5.72 $\text{fb}^{-1}$of integrated luminosity, of which 5.2 fb^{-1} was recorded by the CMS detector under nominal detector conditions. About 6% of this integrated luminosity was eventually deemed not suitable for physics analysis, leaving a final dataset with integrated luminosity equal to about 4.9 fb^{-1}. These data are divided further into the first 2.1 fb^{-1}, referred to as Run2011A which has relatively smaller amount of pileup, and the last 2.8 fb^{-1}, referred to as Run2011B which has comparatively larger amount of pileup. Within the Run2011A dataset, we further split the data events into four different periods, which were reconstructed using different detector calibration conditions. The individual datasets used are summarized in Table 3.1 separated into the five different primary datasets.

3.1.1 Electron Trigger Paths

Level-1 electron trigger candidates are formed by considering regions composed of 3×3 grids of trigger towers, defined by energy measurements in geometrically matching ECAL crystals and HCAL cells. These L1 electron trigger primitive objects must satisfy electron-like shower profile requirements. The transverse energy of the candidate is computed from the sum of the transverse energy of the central trigger tower

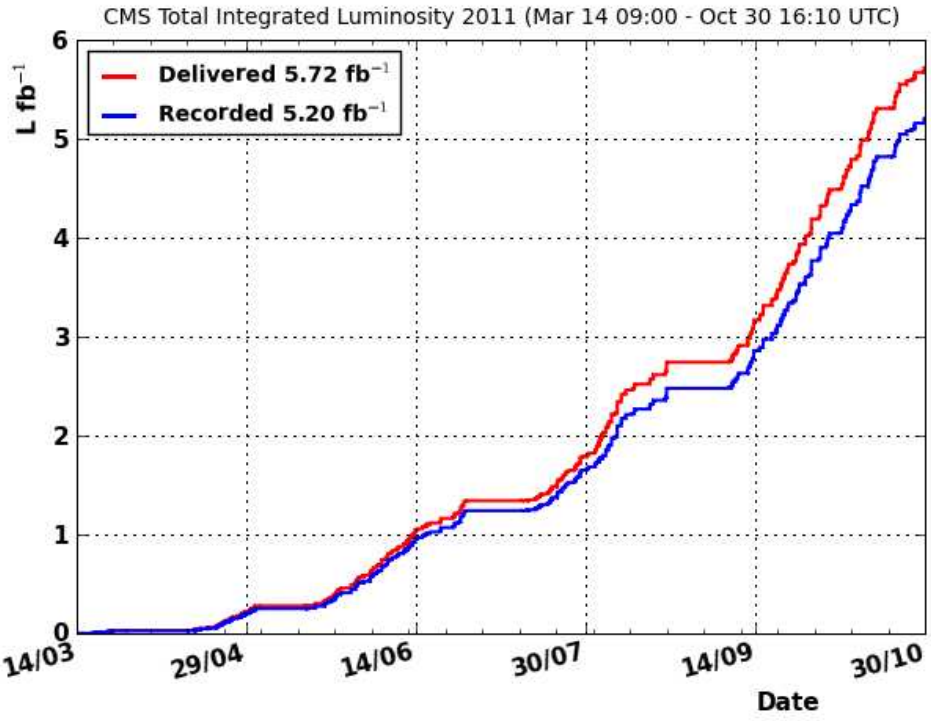


Figure 3-1: Integrated luminosity at CMS as a function of time. The red curve represents the integrated luminosity that was delivered by the LHC, and the blue curve represents the integrated luminosity that was recorded by CMS under nominal detector conditions.

Primary Dataset	Dataset Name
Double Electron Triggers	/DoubleElectron/Run2011A-May10ReReco-v1/AOD /DoubleElectron/Run2011A-PromptReco-v4/AOD /DoubleElectron/Run2011A-05Aug2011-v1/AOD /DoubleElectron/Run2011A-03Oct2011-v1/AOD /DoubleElectron/Run2011B-PromptReco-v1/AOD
Double Muon Triggers	/DoubleMu/Run2011A-May10ReReco-v1/AOD /DoubleMu/Run2011A-PromptReco-v4/AOD /DoubleMu/Run2011A-05Aug2011-v1/AOD /DoubleMu/Run2011A-03Oct2011-v1/AOD /DoubleMu/Run2011B-PromptReco-v1/AOD
Electron Muon Triggers	/MuEG/Run2011A-May10ReReco-v1/AOD /MuEG/Run2011A-PromptReco-v4/AOD /MuEG/Run2011A-05Aug2011-v1/AOD /MuEG/Run2011A-03Oct2011-v1/AOD /MuEG/Run2011B-PromptReco-v1/AOD
Single Electron Triggers	/SingleElectron/Run2011A-May10ReReco-v1/AOD /SingleElectron/Run2011A-PromptReco-v4/AOD /SingleElectron/Run2011A-05Aug2011-v1/AOD /SingleElectron/Run2011A-03Oct2011-v1/AOD /SingleElectron/Run2011B-PromptReco-v1/AOD
Single Muon Triggers	/SingleMu/Run2011A-May10ReReco-v1/AOD /SingleMu/Run2011A-PromptReco-v4/AOD /SingleMu/Run2011A-05Aug2011-v1/AOD /SingleMu/Run2011A-03Oct2011-v1/AOD /SingleMu/Run2011B-PromptReco-v1/AOD

Table 3.1: Summary of datasets used to collect the signal event sample. The datasets are separated into five primary datasets defined by the number of electrons or muons found by the high level trigger system. Each primary dataset is split into five different data taking periods reflecting changing beam and detector conditions. The datasets for the Run2011A period contain events from the first 2.1 fb^{-1} of integrated luminosity and contain relatively fewer amount of pileup interactions. The datasets for the Run2011B period contain events from the final 2.8 fb^{-1} of integrated luminosity and contain relatively larger amount of pileup interactions. The label "AOD" is a name that refers to the particular data structure format of the events in these datasets.

and the maximum transverse energy among the four adjacent towers. The four candidates with the highest transverse energies are passed to the global trigger system to be used as seeds for the high level trigger, from which high level trigger electron objects will be constructed.

At the high level trigger, clusters of ECAL energy deposits are formed near each L1 electron seed by collecting the energy depositions in nearby ECAL crystals. Due to the relatively large material budget of the tracker, electrons are likely to undergo bremsstrahlung in the tracking volume which produce secondary photons. These photons subsequently spread out in the ϕ direction relative to the electron due to the influence of the magnetic field on the electron. So called "superclusters" are formed by collecting different clusters with the same η coordinate but separated in ϕ, in order to collect all of these secondary photons resulting from bremsstrahlung. Starting from the supercluster centroid position corresponding to the energy weighted average of all cluster positions forming the supercluster, we propagate an electron and positron hypothesis through the magnetic field back to the vertex where we require that a match to measured hits in the pixel detector is found. Next, an electron track is constructed seeded by the matching hits in the pixel detector and adding hits in subsequent matching layers of the tracker. Finally, to further reduce the contribution of background to the trigger rate, cuts are made on :

- the ratio of energy deposition in the HCAL to the energy deposition in the ECAL (H/E),
- the width of the energy cluster in the η direction ($\sigma_{i\eta i\eta}$),
- the distance in the η and ϕ coordinate between the supercluster position and the position of the track extrapolated to the primary vertex ($\Delta\eta_{\text{in}}$ and $\Delta\phi_{\text{in}}$), and
- isolation observables computed from the sum of the energy of calorimeter deposits and the transverse momentum of tracks within a geometric cone around the supercluster centroid position defined by $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} < 0.3$.

These variables will be defined more precisely in Section 3.2.2.

The rapid improvement in the instantaneous luminosity delivered by the LHC in 2011 over short periods of time, induced large increases in the trigger rates. This

increase required a corresponding adjustment of the high level trigger menu in order to keep the total trigger rate below the maximum allowed rate threshold. As a result, the requirements in the electron triggers changed significantly over different run ranges defining the different data taking periods. The exact triggers used are shown for each CMS run range in Table 3.2. The exact cut values for each electron related trigger is summarized in Table 3.3, listed by the label name for the trigger path.

3.1.2 Muon Trigger Paths

The L1 muon trigger candidates are constructed by the global muon trigger system which builds muon trigger candidates based on matching tracks reconstructed locally in the DT and CSC trigger systems and hits from the RPC system. For the single muon triggers, we suppress the contribution of background muons to the total trigger rate by imposing the requirement that the muons are isolated. This isolation is determined by geometrically matching the candidate muon tracks with corresponding calorimeter towers. These L1 muon trigger primitive objects are used as seeds for the high level trigger muon reconstruction.

In the high level trigger, we reconstruct trajectories based on all associated muon chamber hits and the resulting trajectories are required to be matched to the L1 seed. Hits from the silicon tracker are then added to the muon trajectory to provide a more accurate momentum measurement. For single muon triggers, we reduce the rate from background by requiring that there is a minimal amount of calorimeter energy deposit and track momentum within a geometric cone around the direction of the muon candidate. The isolation will be defined more precisely in Section 3.2.4. These trigger paths have name labels containing "IsoMu" in Table 3.2.

3.1.3 Triggers for Control Samples

Control samples are event samples that are used to understand specific properties of the background. The production rate of background events is typically prohibitively

Double Electron Trigger Paths	
Run Range	Trigger Used
Run $\leq$ 170053	HLT_Ele17_CaloIdL_CaloIsoVL_Ele8_CaloIdL_CaloIsoVL
Run $\geq$ 170054	HLT_Ele17_CaloIdT_CaloIsoVL_TrkIdVL_TrkIsoVL_ Ele8_CaloIdT_CaloIsoVL_TrkIdVL_TrkIsoVL
Single Electron Trigger Paths	
Run Range	Trigger Used
Run $\leq$ 165888	HLT_Ele27_CaloIdVT_CaloIsoT_TrkIdT_TrkIsoT
165900 $\geq$ Run $\leq$ 166967	HLT_Ele32_CaloIdVT_CaloIsoT_TrkIdT_TrkIsoT
166968 $\geq$ Run $\leq$ 170053	HLT_Ele52_CaloIdVT_TrkIdT
170054 $\geq$ Run $\leq$ 178380	HLT_Ele65_CaloIdVT_TrkIdT
Run $\geq$ 178381	HLT_Ele80_CaloIdVT_TrkIdT
Electron Muon Trigger Paths	
Run Range	Trigger Used
Run $\leq$ 173198	HLT_Mu17_Ele8_CaloIdL
Run $\geq$ 173199	HLT_Mu17_Ele8_CaloIdT_CaloIsoVL
Run $\leq$ 170053	HLT_Mu8_Ele17_CaloIdL
Run $\geq$ 170054	HLT_Mu8_Ele17_CaloIdT_CaloIsoVL
Double Muon Trigger Paths	
Run Range	Trigger Used
Run $\leq$ 164237	HLT_DoubleMu7
165085 $\geq$ Run $\leq$ 178380	HLT_Mu13_Mu8
Run $\geq$ 178381	HLT_Mu17_Mu8
Single Muon Trigger Paths	
Run Range	Trigger Used
Run $\leq$ 163261	HLT_Mu15
163262 $\geq$ Run $\leq$ 170053	HLT_IsoMu17
Run $\geq$ 170054	HLT_IsoMu24

Table 3.2: Single and double electron and muon trigger paths used in the HLT to collect signal events, separated into different run ranges. The number in the trigger path name following the electron or muon labels refers to the p_T requirement for the electron or muon. The labels following the electrons in the trigger path names indicate specific cuts applied on the electron HLT objects. These cuts are listed in detail in Table 3.3.

large, and therefore only a small representative fraction of the event sample is triggered and archived for data analysis. We typically implement this procedure by imposing "prescales" on the particular trigger, where only one in every N_{prescale} events are allowed to trigger. The value of the trigger prescale, N_{prescale}, is adjusted such that a sufficiently large control sample may be collected to achieve an accurate un-

Trigger Name Label	Requirement for Barrel (Endcap) electrons
CaloId_L	H/E $<$ 0.15(0.10) $\sigma_{\eta\eta}<$ 0.014 (0.035)
CaloId_T	H/E $<$ 0.10(0.075) $\sigma_{\eta\eta}<$ 0.011 (0.031)
CaloId_VT	H/E $<$ 0.05(0.05) $\sigma_{\eta\eta}<$ 0.011 (0.031)

Trigger Name Label	Requirement
TrkId_VL	$\|\Delta\eta\|<$ 0.01 (0.01) $\Delta\phi<$ 0.15 (0.10)
TrkId_T	$\|\Delta\eta\|<$ 0.008 (0.008) $\Delta\phi<$ 0.07 (0.05)

Trigger Name Label	Requirement
CaloIso_VL	ECalIso/E_T $<$ 0.2 (0.2) HCalIso/E_T $<$ 0.2 (0.2)
CaloIso_T	ECalIso/E_T $<$ 0.125 (0.075) HCalIso/E_T $<$ 0.125 (0.075)
CaloIso_VT	ECalIso/E_T $<$ 0.05 (0.05) HCalIso/E_T $<$ 0.05 (0.05)

Trigger Name Label	Requirement
TrkIso_VL	TrkIso/E_T $<$ 0.2 (0.2)
TrkIso_T	TrkIso/E_T $<$ 0.125 (0.075)
TrkIso_VT	TrkIso/E_T $<$ 0.05 (0.05)

Table 3.3: Definitions of the electron trigger path labels. The abbreviations in the trigger path labels represent: L=Loose, VL=Very loose, T=Tight, VT=Very Tight. The variables on which we make these requirements will be defined more precisely in Section 3.2.2.

derstanding of the background.

For this analysis, it was critical to collect a set of multijet events containing isolated leptons resulting from hadronization of quarks and gluons. These events are collected from single electron and single muon triggers with the exact same requirements as those used for the double electron, double muon, and electron-muon signal triggers, to eliminate any trigger bias. The trigger prescales were tuned in a such a way that we were able to collect a sufficiently large sample of background electron and muons both to measure and control the background rate, and to obtain characteristic background samples for training multivariate discriminators. In total, using

these prescaled triggers, we collected an equivalent luminosity of about 5 pb^{-1} for background electrons and 60 pb^{-1} for background muons. The trigger paths used are summarized in Table 3.4.

Lepton Type	Trigger Path
Electrons	HLT_Ele8 HLT_Ele8_CaloIdL_CaloIsoVL HLT_Ele8_CaloIdL_CaloIsoVL_Jet40 HLT_Ele17_CaloIdL_CaloIsoVL HLT_Ele8_CaloIdT_CaloIsoVL_TrkIdVL_TrkIsoVL
Muon	HLT_Mu8 HLT_Mu15

Table 3.4: Triggers used to collect the background electron and muon control samples. The number in the trigger path name following the electron or muon labels refers to the p_T requirement for the electron or muon. The labels following the electrons in the trigger path names indicate specific cuts applied on the electron HLT objects. These cuts are listed in detail in Table 3.3.

In addition, special dielectron triggers were implemented in order to collect a set of $Z \rightarrow e^+e^-$ events where one electron was left unbiased by selection requirements. This is a very important sample of events that is used to measure the electron selection and triggering efficiency. In order for the trigger rate to remain sufficiently small, very tight requirements were necessary on one of the two electrons to reduce the contribution from backgrounds. The trigger path HLT_Ele17_CaloIdVT_CaloIsoVT_TrkIdT_TrkIsoVT_SC8_Mass30 is used to collect events to measure the efficiency of electrons with p_T between 10 GeV/c and 20 GeV/c, while the trigger path HLT_Ele32_CaloIdL_CaloIsoVL_SC17 is used to collect events to measure the efficiency of electrons with p_T above 20 GeV/c. The particular meaning of the labels in the trigger path name is explained in more detail in Tables 3.2 and 3.3. For muons, the single muon trigger rates were sufficiently low that such an analogous trigger was not required.

3.2 Lepton Reconstruction and Selection

Reconstructing and selecting electrons and muons with high purity is one of the most important aspects of this analysis. Jets are copiously produced at hadron colliders, and the identification of electrons and muons in this environment is the primary tool by which the largest backgrounds are suppressed. From here on, we will refer to electrons or muons from the decay of a W or Z boson as signal leptons or "real leptons", and we will refer to quarks or gluons which fragment and hadronize in a way that creates the experimental signature consistent with an electron or muon as background leptons or "fake leptons".

3.2.1 Electron Reconstruction

An electron in CMS produces a track as it traverses the layers of the silicon tracker, followed by a characteristic energy shower pattern when it interacts in the electromagnetic calorimeter. Electron reconstruction at CMS [56] is complicated by the large amount of material that it must traverse before it reaches the electromagnetic calorimeter resulting in large rates of bremsstrahlung. Due to this large amount of bremsstrahlung and the strong magnetic field the electron typically produces an energy pattern spread out in the ϕ direction as the trajectory of the electron is pulled by the magnetic field while photons from bremsstrahlung continue along the original trajectory.

The first step in electron reconstruction is the reconstruction of all of the energy clusters produced by the electron as well as the photons resulting from bremsstrahlung. Energy clusters in the ECAL are collected in a narrow strip in the ϕ direction and combined to form what is called a "supercluster". The supercluster is propagated back towards the pixel detector, accounting for the influence of the magnetic field, and matched with compatible track segments reconstructed in the pixel detector. Next, the electron track is built by proceeding through each layer of the silicon tracker starting from the pixel track segment, and adding measured hits in subsequent layer

that are consistent with the propagated trajectory. The propagation and track fitting algorithm used is based on the combinatorial Kalman filter [57]. To account for the high rate of bremsstrahlung and to maximize electron track finding efficiency, much looser track hit position compatibility requirements are imposed in the track building process.

The Kalman filter technique of track fitting is known to be optimal only if all energy loss distributions are Gaussian. The best reference model for electron energy loss is the Bethe-Heitler model [58] and the energy loss distribution described by this model is not particular well approximated by a Gaussian. We generalize the Kalman filter technique by representing all track trajectory states by a sum of Gaussians rather than a single Gaussian. This track fitting technique is called the Gaussian Sum Filter (GSF) technique [59]. This method results in an electron track trajectory that is composed of multiple weighted trajectories. The trajectory of the electron is taken as the trajectory with the largest weight. The electron momentum is measured as a weighted mean of the supercluster energy measurement and the track momentum measurement, whose weights are based on the uncertainties of each measurement. Three different charge measurements are used together to determine the charge of the electron: the charge of the GSF track, the charge of a track reconstructed from the algorithm used for generic minimum ionizing particles matched to the electron candidate, and the charge measured from the relative position of the inner-most track hit and the supercluster. The charge of the electron candidate is taken as the majority decision of these three methods. This choice reduces the rate of charge misidentification by a factor of two.

3.2.2 Electron Identification and Isolation

Standard electron selection at CMS typically proceeds by imposing simple cuts on various observables related to the isolation of the electron, the shape of the electromagnetic shower, the quality of the matching between the position of the shower and the trajectory of the track, the ratio of the energy deposited in the hadronic calorimeter

to the energy deposited in the electromagnetic calorimeter, and the impact parameter of the electron track. One typically also explicitly reconstructs photons converting to an electron-positron pair in material, and an electron candidate is rejected if its track matches to one leg of a reconstructed conversion. These types of selections are referred to as "cut-based" electron selection. This search is particularly sensitive to background processes where jets are misindentified as electrons. As a result, we significantly improve the sensitivity of this analysis by making use of multivariate methods to improve the discrimination of signal electrons and background electrons.

In this section, we begin by giving a general description of the typical methods that have been used to discriminate signal electrons from background electrons in the past, and allude to particular cut-based electron selections that have been used in past CMS results [60, 61]. This is followed by an in-depth description of the multivariate selection and a comparison of the performance with a standard cut-based selection.

Conversion Rejection

One of the major sources of background electrons are those produced by the conversion of a photon from the decay of a relatively isolated neutral pion. Those conversions for which one of the resulting electron or positron legs carries the majority of the photon momentum, leaving the remaining one fairly soft, are most likely to be misidentified as signal electrons. To suppress such sources of background, we perform vertex fits, pairing the track associated with the particular electron candidate under consideration and any other compatible and well measured track. A Gaussian constraint is imposed on the vertex fit under the hypothesis that the resulting mass is zero.

There is a non-zero chance to select conversion candidates that are not real conversions. These situations may result from the pairing of the electron track with a random geometrically matching track, referred to as combinatorial background, or they may result from an electron that emits a photon through radiation or bremsstrahlung which subsequently converts and is mistakenly paired with the original electron track. By rejecting such electron candidates as background, we potentially lose a significant

amount of signal efficiency. To suppress such occurrences, we require that the vertex fit probability is larger than 10^{-6} and that the conversion occurred only in or after the first layer of the pixel detector by requiring that the distance between the conversion vertex and the beamspot in the transverse plane is larger than 2 cm. We further require that none of the tracks which form the conversion have any hits in the tracker layers that are between the interaction point and the fitted vertex position, which reduces the case of combinatorial background.

Finally, an electron whose track is successfully paired and reconstructed as a conversion passing these requirements are rejected. The efficiency for a signal electron to survive the conversion rejection is roughly 99% and 93% for electrons in the barrel and endcap respectively. The conversion veto reduces background electrons from photon conversions by more than a factor of two.

An electron produced at the interaction point is expected to produce hits in the innermost layers of the pixel detector; an electron produced by a photon which converts in the material associated with the first or subsequent pixel layers is expected not to produce such a hit. Making use of the fact that the inefficiency for reconstructing a hit in the first pixel layer is only a few percent, we further reject a large fraction of fake electrons produced by conversions by requiring that there be no such missing hits in the electron track trajectory. The efficiency of this requirement is roughly 99% for electrons in the barrel, and about 97% for electrons in the endcap.

Isolation

Furthermore, electrons are required to be geometrically isolated from other charged particles and energy deposits. We match each track to calorimeter deposits with very fine granularity. This allows for individual tracks to be identified as electrons, muons, or charged hadrons, and for the unmatched calorimeter deposits to be identified as photons or neutral hadrons. This procedure results in a global description of the event with all particles reconstructed, and is referred to as particle flow (PF). The procedure will be described further in Section 3.4. Such reconstructed "PF candidates" are used

as the basis for defining isolation observables for an electron. A "cone" is defined in the η and ϕ coordinate space around the direction of the electron candidate by the requirement that $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} < R_0$, for some particular value of R_0. As mentioned already in Section 2.2, η is used instead of the polar angle θ due to the fact that the rate of particle production is approximately flat as a function of η. R_0 is typically chosen to be a number between 0.3 and 0.5. Next, the p_T of all reconstructed particles within this cone is summed to obtain the isolation observable.

We define three different isolation observables based on the type of the reconstructed particles: charged hadrons, photons, and neutral hadrons. The PF charged particle isolation is obtained by summing the momenta of all particles inside the cone that are charged, imposing the extra requirement that the z coordinate of the track trajectory at the point of closest approach to the beamline is within 0.1 cm of the event primary vertex, which is intended to reject particles produced from pileup events. The event primary vertex is defined as the primary vertex with the largest value of $\sum p_T^2$.

The PF photon isolation is computed by summing the momenta of all photons inside the cone, but vetoing those PF photons which lie within a narrow strip in the ϕ coordinate defined by $|\Delta\eta| < 0.025$. In the particle flow electron reconstruction procedure, an attempt is made to collect all PF photon candidates that are emitted from the electron undergoing bremsstrahlung and to combine their momenta with the momentum of the electron. Due to some inefficiency of this bremsstrahlung collection procedure, some fraction of such PF photons do not get collected. The veto of a narrow strip in the ϕ coordinate attempts to make the isolation definition more robust against this particular situation.

Finally the PF neutral hadron isolation is computed by summing the momenta of all neutral hadron PF candidates inside the cone, but vetoing those which fall within an inner cone defined by $\Delta R < 0.07$. This veto is intended to address the situation where the electromagnetic shower induced by the electron is not fully contained within the ECAL and spills into the HCAL. In the case that the particle flow reconstruction

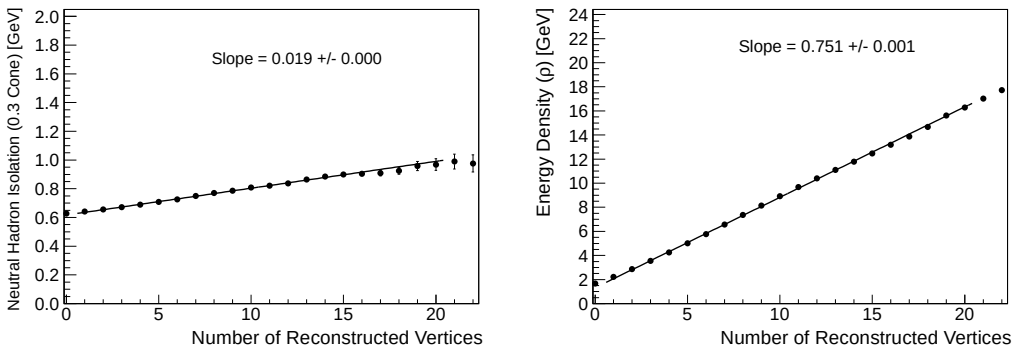
fails to identify such an object as an electron, the energy will be assigned to a neutral hadron, and included in the isolation sum. To ensure robustness against such failures we impose this inner cone veto.

For the PF photon isolation and the PF neutral hadron isolation, a minimum p_T cut is imposed on the particle flow candidates that are considered in order to increase robustness against the influence of particles from pileup events. The minimum p_T cut is typically chosen to be 0.5 GeV/c. Once the three isolation variables are defined, they are summed to form a combined isolation quantity and this is required to be less than some fraction of the electron p_T. Typical cut values for this fraction are between 10% and 15%.

Due to the presence of a large amount of pileup associated with collisions produced by the LHC, random particles from pileup interactions may enter into the isolation cone of the electron candidate under consideration. Such particles increase the measured value of the photon and neutral hadron isolation and as a result decrease the efficiency of the isolation requirement as the number of pileup interactions increases. In order to mitigate this effect, we perform a correction for the isolation observables which subtracts some amount of isolation energy or momentum based on the average energy density measured in the event. The energy density in the event, represented by the symbol ρ, is measured using the FastJet approach [62–64], where the full geometric phase space is subdivided into different areas, constructed using the anti-Kt jet clustering algorithm described in more detail in Section 3.3.2. A set of infinitely soft particles are randomly generated and used as seeds for the jet clustering algorithm. For each jet constructed in this way, we divide its energy by its area, and take the median over all such jets in the event as the representative energy density of the event.

Due to different instrumental effects, the average measured contribution of pileup to the isolation quantities is not equal to the simple product of the isolation cone area and the energy density. Instead, we compute an effective area constructed by comparing the dependence of the measured energy density and the isolation quantity

as a function of the number of pileup events. We use the unbiased probe leg in $Z \rightarrow e^+e^-$ events, imposing the tag-and-probe selection described above, to plot the mean of the energy density and the isolation observable as a function of the number of pileup events. As illustrated in Figure 3-2, a linear fit is performed both for the isolation variable and the energy density, and the ratio of the slope for the isolation observable to the slope of the energy density is defined as the effective area. In data, we use the number of reconstructed primary vertices instead because it is an easily measured quantity and the ratio is insensitive to the vertexing efficiency. Using this procedure, we obtain the effective areas for a number of different isolation quantities and summarize them in Table 3.5. Some non-trivial dependence on psuedorapidity is observed. The statistical uncertainties in these measured values are negligible, and any systematic errors that are made in this measurement will only make the correction suboptimal.



(a) Isolation Observable (b) Energy Density

Figure 3-2: The mean of the isolation observable and the energy density is plotted as a function of the number of reconstructed vertices. A linear fit is performed on both, resulting in a slope of 0.019 for the isolation observable and a slope of 0.751 for the energy density. The effective area is defined as the ratio of these slopes, and measured to be 0.025 in this case.

Electron η Bin	PF NeutralHadron Isolation 0.3 Cone	PF Gamma Isolation 0.3 Cone
$0.0 < \|\eta\| < 1.0$	0.017	0.031
$1.0 < \|\eta\| < 1.479$	0.025	0.022
$1.479 < \|\eta\| < 2.0$	0.030	0.036
$2.0 < \|\eta\| < 2.25$	0.022	0.107
$2.25 < \|\eta\| < 2.5$	0.018	0.149
Electron η Bin	PF NeutralHadron Isolation 0.4 Cone	PF Gamma Isolation 0.4 Cone
$0.0 < \|\eta\| < 1.0$	0.034	0.065
$1.0 < \|\eta\| < 1.479$	0.050	0.043
$1.479 < \|\eta\| < 2.0$	0.060	0.053
$2.0 < \|\eta\| < 2.25$	0.055	0.142
$2.25 < \|\eta\| < 2.5$	0.073	0.207

Table 3.5: The effective areas used to correct the electron isolation in different bins of pseudorapidity. The statistical uncertainties for these effective area values are negligible. Systematic errors in these measured values will only affect the optimality of the correction, and are not shown.

Multivariate Electron Selection

In the previous section we discussed all of the individual pieces that are used for typical "cut-based" approaches to electron identification and isolation. These approaches were used for various precision electroweak measurements in the early stages of the LHC run [60, 61], and provide a baseline definition of electron candidates. The sensitivity of this analysis is critically dependent on the degree to which the background type leptons are suppressed. Thus, to optimize the sensitivity, we attempt to increase the performance of the signal-to-background discrimination by making use of multivariate discrimination techniques. The observables mentioned above, and further observables with signal to background discrimination power, are used as input variables to a boosted decision tree (BDT) [65], and trained on data samples of signal type electrons and background type electrons. We use version 4.0.7 of the TMVA implementation [66] of a boosted decision tree. The sample of signal electrons is constructed from dielectron data events with dielectron mass between 75 GeV/c^2 and

105 GeV/c^2 and has a purity in $Z \to e^+e^-$ events that is better than 95%. The sample of background electrons is constructed from data events triggered by the dedicated fake electron triggers:

- HLT_Ele8_CaloIdL_CaloIsoVL,
- HLT_Ele17_CaloIdL_CaloIsoVL, or
- HLT_Ele8_CaloIdT_TrkIdVL_CaloIsoVL_TrkIsoVL,

and requiring that the event has one and only one electron candidate. Contamination from W events are suppressed by requiring that the missing transverse energy is less than 20 GeV/c, and that the p_T of the electron is less than 35 GeV/c. This results in a sample of multijet events, where the electron candidate is a quark or gluon that has been misidentified as a promptly produced electron, with a purity above 97%.

It is possible for the BDT to focus on particular statistical fluctuations of the training sample. The resulting discriminator may not be as discriminating when applied on a statistically independent but otherwise identical sample. This generic phenomenon is referred to as overtraining. If our particular BDT exhibits any degree of overtraining, then the measurements of the signal and background efficiency and therefore the performance for a particular selection will be biased if it is evaluated on the training sample itself. To mitigate this possible bias due to overtraining of the BDT the signal and background samples are explicitly divided in half. Events with an event number that is even are used to train the BDT, and events with an odd event number are used for efficiency measurements and to evaluate performance.

A set of preselection cuts are applied to the electron candidates before training. They are designed to be tighter than the electron selection requirements imposed by the HLT. These preselection cuts also allow the BDT to focus in on the more relevant regions of phase space, to improve its discrimination. The exact preselection cuts are as follows:

- $\sigma_{i\eta i\eta} < 0.01$ (0.03) for barrel (endcap),

- $|\Delta\phi_{in}| < 0.15$ (0.10),
- $|\Delta\eta_{in}| < 0.007$ (0.009),
- $H/E < 0.12$ (0.10),
- Track Isolation with ΔR= 0.3 cone / $p_T^{\text{ele}} < 0.2$,
- ECAL Isolation with ΔR= 0.3 cone / $p_T^{\text{ele}} < 0.2$,
- HCAL Isolation with ΔR= 0.3 cone / $p_T^{\text{ele}} < 0.2$,
- $|d_0| < 0.02$ cm,
- $|d_z| < 0.1$ cm, and
- the conversion rejection as described above,

where $\sigma_{i\eta i\eta}$ is a measure of the width of the electromagnetic shower in the η direction, $\Delta\eta_{in}$ and $\Delta\phi_{in}$ are the distances in the η and ϕ coordinates between the centroid position of the supercluster and the direction of the track trajectory extrapolated back to the primary vertex, H/E is the ratio of the hadronic energy deposition to the electromagnetic energy deposition, d_0 is the transverse impact parameter of the electron track, and d_z is the distance in the z coordinate between the point of closest approach of the electron track to the beamline and the primary event vertex. The shower width, $\sigma_{i\eta i\eta}$, is defined by:

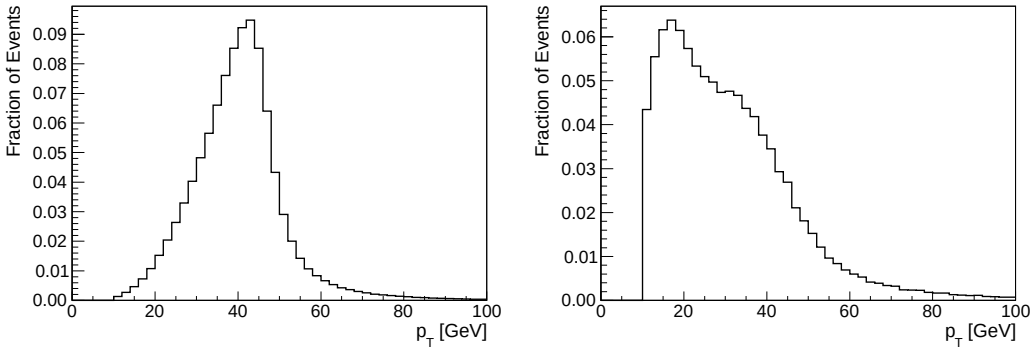
$$\sigma_{i\eta i\eta} = \sqrt{\frac{\sum_{i\in 5\times 5} w_i(\eta_i - \langle\eta\rangle_{5\times 5})^2}{\sum_{i\in 5\times 5} w_i}}, w_i = \max\left(0, 4.7 + \ln\frac{E_i}{E_{5\times 5}}\right), \tag{3.1}$$

where E_i and η_i are the energy and pseudorapidity of the i'th crystal in the 5×5 grid of crystals around the highest energy crystal, $E_{5\times 5}$ is the energy of summed over all crystals in the 5×5 grid, and $\langle\eta\rangle_{5\times 5}$ is the pseudorapidity coordinate of the centroid position computed over all crystals in the 5×5 grid. The particular form of the weight

for each crystal is intended to reduce systematic biases in the position measurement for showers with an exponential profile.

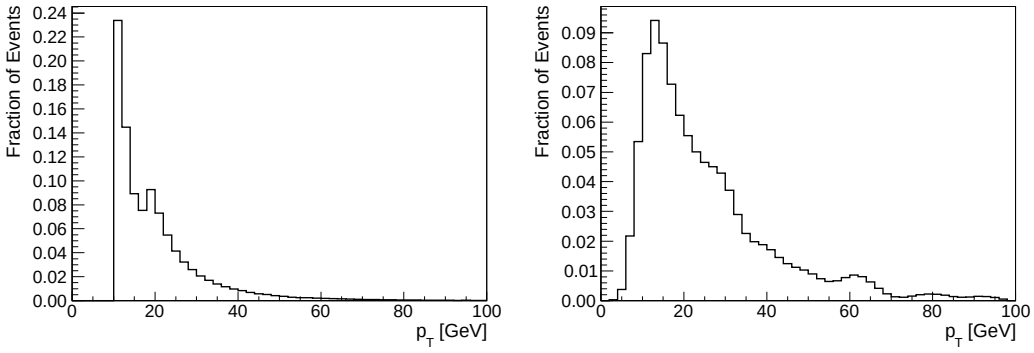
The exact properties of signal and background electrons may have a significant dependence on the p_T of the electron. In order for the trained BDT to be optimal for discriminating signal and background electrons for the Higgs boson search, we reweight the signal and background training samples such that the resulting p_T spectra resemble the characteristic spectra for the signal and the relevant background. The target p_T spectrum for signal electrons is taken from the lepton p_T spectrum in Higgs boson events from Monte Carlo simulation with a Higgs boson mass hypothesis of 115 GeV/c^2. This choice ensures that the BDT is more optimal for low mass Higgs bosons, which is preferred by the precision electroweak measurements and faces larger fake lepton backgrounds. The original and reweighted spectra for the signal electron sample are shown in Figure 3-3. The target p_T spectrum for the background sample is taken from the W+jets Monte Carlo, which is the main fake lepton background process. The original and reweighted spectra for the background electron sample are shown in Figure 3-4. Due to the limited size of the W+jets Monte Carlo sample, the target spectrum exhibits some amount of statistical fluctuations and has been artificially smoothed. Finally, due to the fact that the background sample is constructed from events collected through prescaled triggers, whose prescales were changing with time and therefore changing with the LHC beam conditions, the distribution of the amount of pileup for events in this sample is different from the events in the signal electron sample. We reweight the pileup distribution of the background electron sample to match that of the signal electron sample because the discriminating observables, in particular the isolation, has some nontrivial dependence on the amount of pileup.

The observables that are used as inputs to the BDT fall broadly into one of the following categories: electromagnetic shower shape observables, measures of the quality of the geometric matching between the calorimeter cluster and the track trajectory, measures of consistency between the electromagnetic calorimeter energy measurement



(a) Unweighted (b) Weighted

Figure 3-3: The unweighted and reweighted p_{T} spectra of the signal electron training sample.



(a) Unweighted (b) Weighted

Figure 3-4: The unweighted and reweighted p_{T} spectra of the background electron training sample.

and the track momentum measurement, impact parameter observables, and isolation observables.

The shower shape observables that enter as input into the BDT training are:

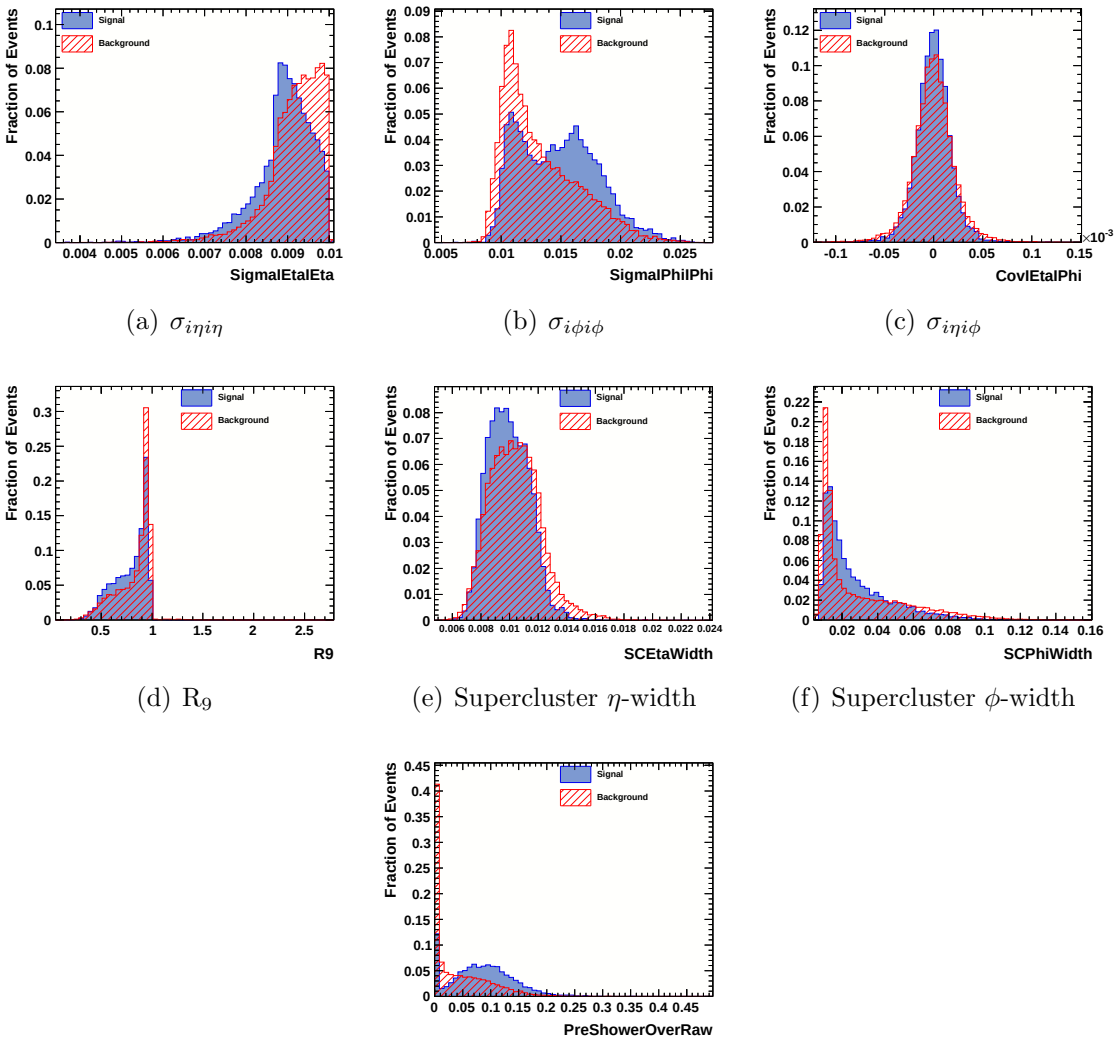
- $\sigma_{i\eta i\eta}$, a measure of the width of the energy deposition pattern in the seed cluster in the η direction,
- $\sigma_{i\phi i\phi}$, a measure of the width of the energy deposition pattern in the seed cluster in the ϕ direction,
- $\sigma_{i\eta i\phi}$, a measure of the correlation of the energy deposition pattern in the seed cluster in the η and ϕ directions,
- R_9, the ratio of the energy deposited in a three by three square of ECAL crystals around the most energetic crystal to the uncorrected energy of the supercluster,
- the supercluster η-width, a measure of the width of the entire supercluster in the η direction,
- the supercluster ϕ-width, a measure of the width of the entire supercluster in the ϕ direction, and
- $E_{\text{Preshower}}/E_{\text{Supercluster}}$, an observable characterizing the shape of the shower in the longitudinal direction.

The variables $\sigma_{i\phi i\phi}$, supercluster η-width, and supercluster ϕ-width are defined analogously to the definition of $\sigma_{i\eta i\eta}$ given in Equation 3.1. For the supercluster widths, the sum is performed over all energy clusters that make up the supercluster. The covariance variable $\sigma_{i\eta i\phi}$ is computed as follows:

$$\sigma_{i\eta i\phi} = \frac{\sum_{i\in 5\times 5} w_i(\eta_i - \langle\eta\rangle_{5\times 5})(\phi_i - \langle\phi\rangle_{5\times 5})}{\sum_{i\in 5\times 5} w_i}, \tag{3.2}$$

with the same weights w_i as in Equation 3.1. The longitudinal shower shape variable $E_{\text{Preshower}}/E_{\text{Supercluster}}$ is only used for electrons in the endcap region, where the

preshower detector is situated. For illustration, a comparison of the distribution of the shower shape observables for electrons in the signal and background training samples are shown in Figure 3-5.



(a) $\sigma_{i\eta i\eta}$ (b) $\sigma_{i\phi i\phi}$ (c) $\sigma_{i\eta i\phi}$

(d) R_9 (e) Supercluster η-width (f) Supercluster ϕ-width

(g) $E_{\text{Preshower}}/E_{\text{Supercluster}}$

Figure 3-5: For illustration, we compare the signal and background distributions of the MVA input observables related to the shower shape of the electron, for the kinematic bin with electrons of p_T between 10 GeV/c and 20 GeV/c and η between 0.0 and 1.0. The plot for $E_{\text{Preshower}}/E_{\text{Supercluster}}$ is shown for electrons with p_T between 10 GeV/c and 20 GeV/c and η between 1.479 and 2.5, as the variable is only defined in the endcap. We observe differences between signal and background in the distributions of all of these variables. The BDT combines these differences to give greater discrimination.

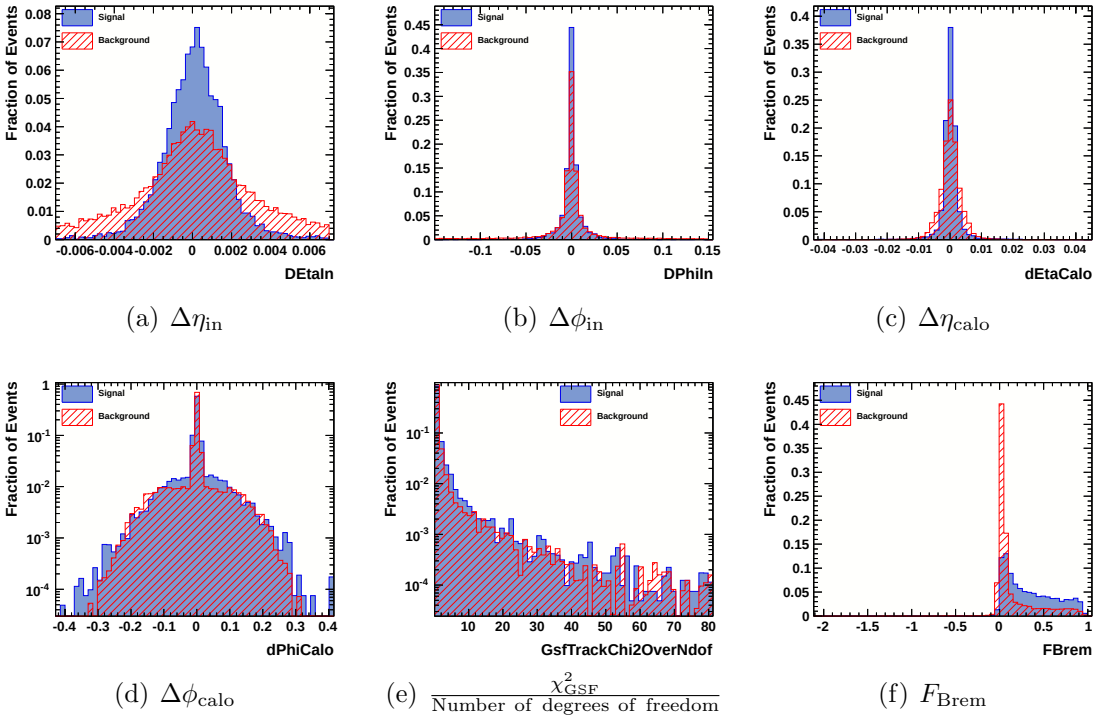
The observables related to the properties of the electron track are:

- $\Delta\eta_{\text{in}}$ and $\Delta\phi_{\text{in}}$, as defined above,
- $\Delta\eta_{\text{calo}}$ and $\Delta\phi_{\text{calo}}$, the distance in the η and ϕ coordinate between the centroid position of the supercluster and the direction of the track trajectory extrapolated to the front face of the ECAL,
- χ^2_{GSF}/Number of degrees of freedom, a measure of the fit quality of the electron GSF track, and
- F_{Brem}, a measure of the amount of bremsstrahlung based on the difference in the measured momenta of the track trajectory extrapolated to the primary vertex and extrapolated to the calorimeter

More precisely, the variable F_{Brem} is defined as the normalized difference in the momentum magnitude measured from the trajectory state at the primary vertex and the momentum magnitude measured from the trajectory state at the edge of the ECAL. The signal and background distributions of the track related observables are shown in Figure 3-6.

The observables related to the compatibility of the calorimeter energy measurement and the track momentum measurement are:

- $E_{\text{Supercluster}}/P_{\text{GSF}-\text{Mean}}$, where the track momentum is taken from the weighted mean of all of the trajectory states,
- $1/E_{\text{Supercluster}} - 1/P_{\text{GSF}-\text{Mode}}$, where the track momentum is taken from the trajectory state with the largest weight,
- $E_{\text{seed cluster}}/P_{\text{out}}$, where P_{out} is the momentum of the track extrapolated to the ECAL, and
- $E_{\text{seed cluster}}/P_{\text{in}}$, where P_{in} is the momentum of the track extrapolated to the primary vertex.

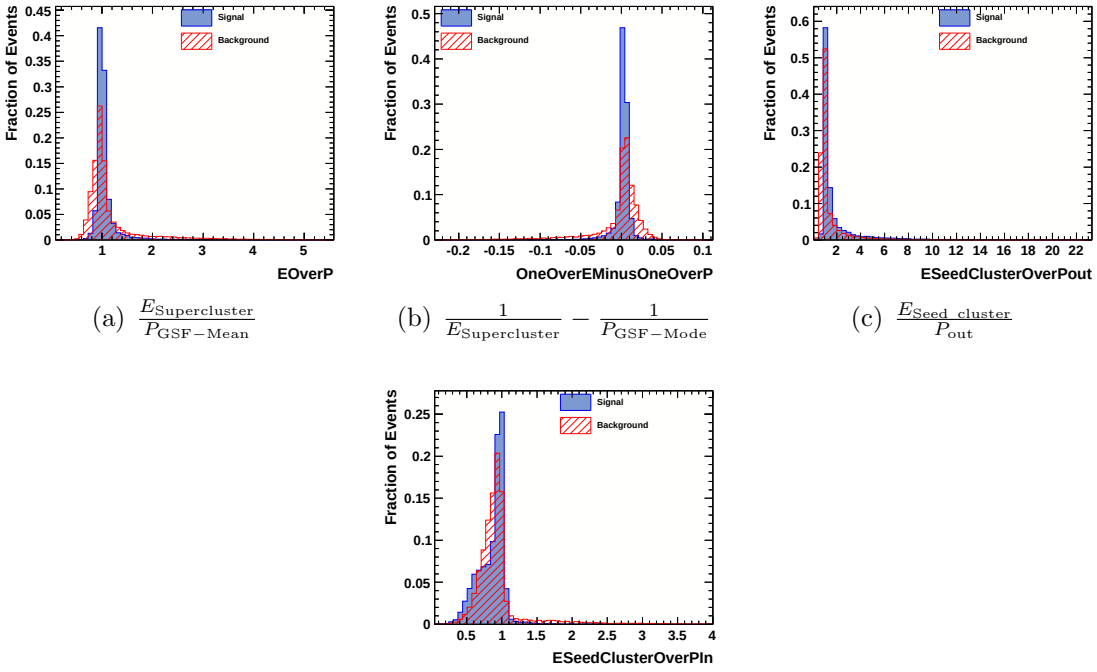


(a) $\Delta\eta_{\text{in}}$ (b) $\Delta\phi_{\text{in}}$ (c) $\Delta\eta_{\text{calo}}$

(d) $\Delta\phi_{\text{calo}}$ (e) $\frac{\chi^2_{\text{GSF}}}{\text{Number of degrees of freedom}}$ (f) F_{Brem}

Figure 3-6: For illustration, we compare the signal and background distributions of the electron track related MVA input observables, for the kinematic bin with electrons of p_T between 10 GeV/c and 20 GeV/c and η between 0.0 and 1.0. The most discriminating variables are $\Delta\eta_{\text{in}}$ and F_{Brem}.

The use of both P_{out} and P_{in} combined with the variable F_{Brem} allows us to better evaluate the compatibility of the energy cluster measurement with the track momentum measurement depending on whether the electron emitted a significant amount of its energy via bremsstrahlung or not. The use of both the mean and the mode of the GSF track trajectory state, and their correlation allows for a more accurate determination of the electron momentum and a more accurate evaluation of its compatibility with the calorimeter energy measurement. The signal and background distributions of these observables are shown in Figure 3-7.



(a) $\frac{E_{\text{Supercluster}}}{P_{\text{GSF-Mean}}}$ (b) $\frac{1}{E_{\text{Supercluster}}} - \frac{1}{P_{\text{GSF-Mode}}}$ (c) $\frac{E_{\text{Seed cluster}}}{P_{\text{out}}}$

(d) $\frac{E_{\text{Seed cluster}}}{P_{\text{in}}}$

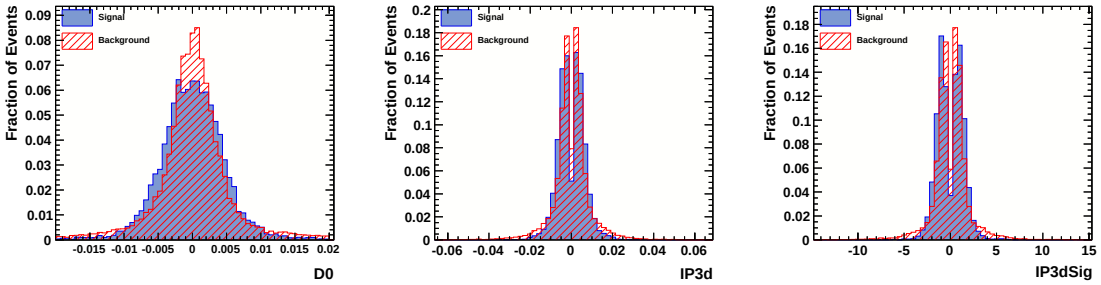
Figure 3-7: For illustration, we compare the signal and background distributions of the MVA input observables measuring the compatibility of the calorimeter energy measurement and the track momentum measurement with a muon hypothesis, for the kinematic bin with electrons of p_T between 10 GeV/c and 20 GeV/c.

The impact parameter observables are:

- d_0, the transverse impact parameter,
- $d_{3\text{D}}$, the three dimensional impact parameter, and

- $d_{3D}/\sigma(d_{3D})$, the significance of the three dimensional impact parameter.

The signal and background distributions of the impact parameter observables are shown in Figure 3-8.



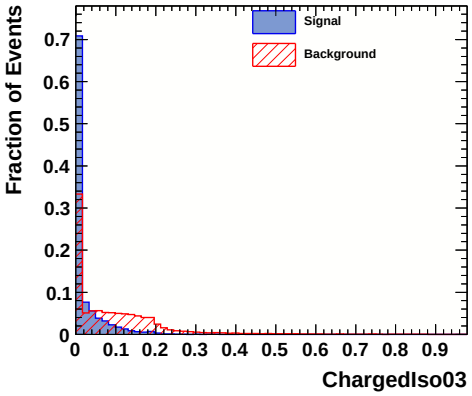
(a) d_0 (b) d_{3D} (c) $\frac{d_{3D}}{\sigma(d_{3D})}$

Figure 3-8: For illustration, we compare the signal and background distributions of the MVA input observables related to the impact parameter of the electron, for the kinematic bin with electrons of p_T between 10 GeV/c and 20 GeV/c.

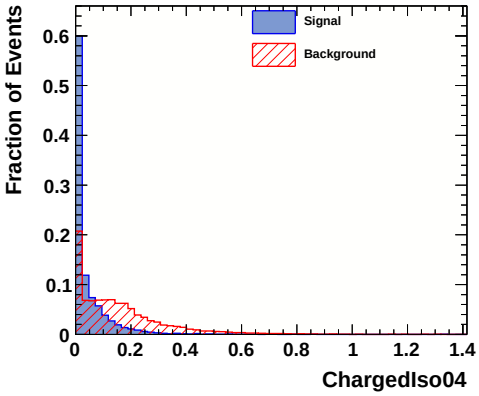
Finally, the isolation observables that we use are the pileup corrected charged particle isolation, photon isolation, and neutral hadron isolation, as defined above in Section 3.2.2, and with two different cone sizes of ΔR= 0.3 and ΔR= 0.4. The signal and background distributions of the isolation are shown in Figure 3-9.

These input observables are passed to a boosted decision tree training, using the gradient boost method [66], with 2000 trees and a shrinkage parameter of 0.10. The training is performed separately in two bins of p_T, one with p_T between 10 GeV/c and 20 GeV/c and one with p_T between 20 GeV/c and 35 GeV/c; and three bins of η: $0.0 \leq |\eta| < 1.0$, $1.0 \leq |\eta| < 1.479$, $1.479 \leq |\eta| < 2.5$. The number of training events differ from bin to bin but is always at least 10000. For illustration, an example of the linear correlations between all of the variables described above that were given as input to the BDT training for one particular kinematic bin is shown in Figure 3-10 for signal and for background. An example of the BDT output is shown in Figure 3-11.

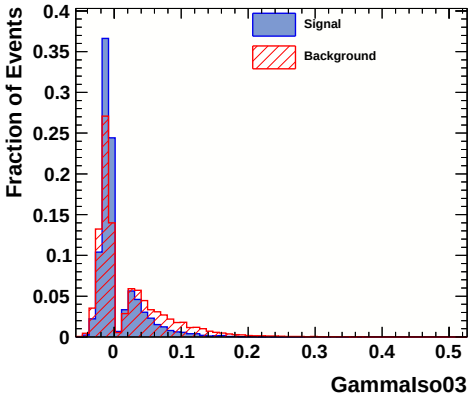
The performance in the signal-to-background discrimination is measured by plotting the receiver operating characteric or "ROC curve", which maps out the values of



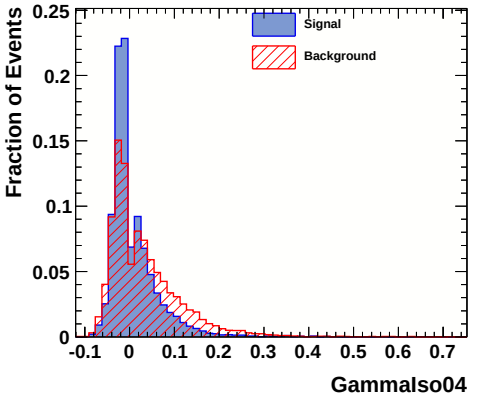
(a) Charged Particle Isolation 0.3 Cone



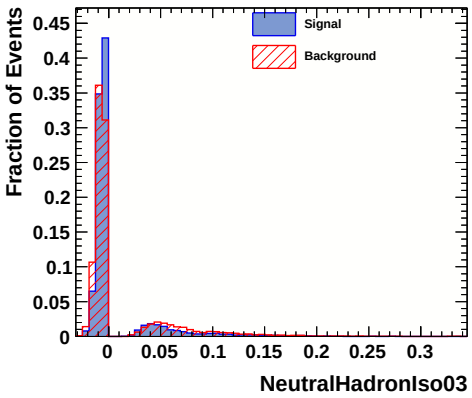
(b) Charged Particle Isolation 0.4 Cone



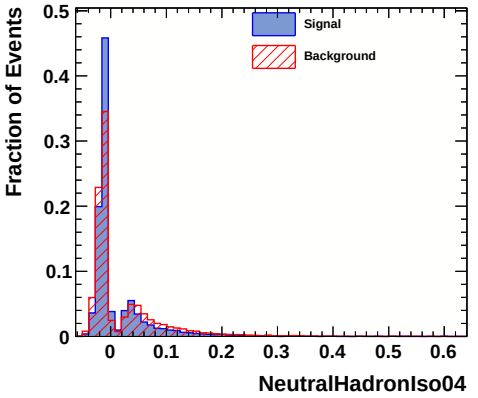
(c) Photon Isolation 0.3 Cone



(d) Photon Isolation 0.4 Cone

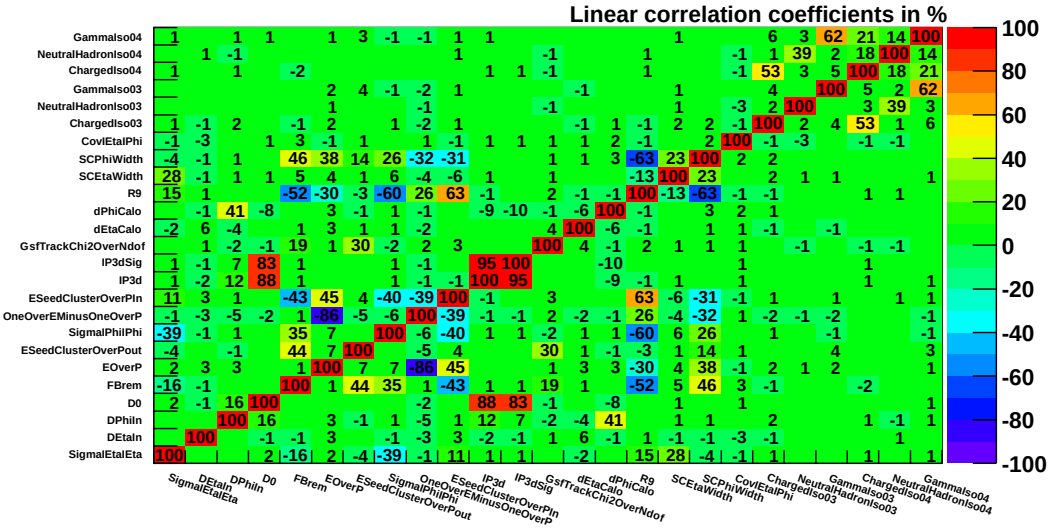


(e) Neutral Hadron Isolation 0.3 Cone

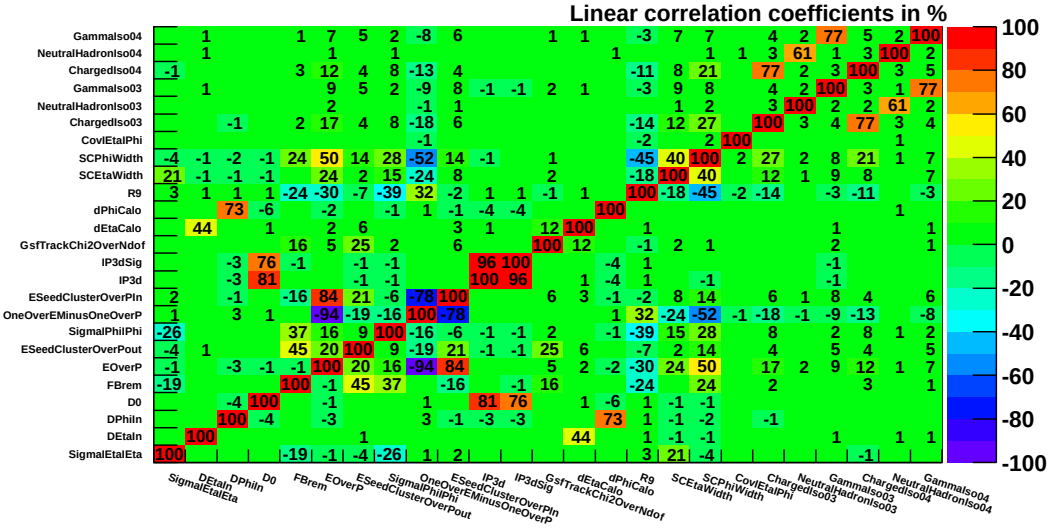


(f) Neutral Hadron Isolation 0.4 Cone

Figure 3-9: For illustration, we compare the signal and background distributions of the electron isolation MVA input observables, for the kinematic bin with electrons of p_T between 10 GeV/c and 20 GeV/c.



(a) Signal



(b) Background

Figure 3-10: The matrix of linear correlation coefficients between all electron selection MVA input observables in the signal and background training samples, for the kinematic bin with electrons of p_T between 10 GeV/c and 20 GeV/c. A few features are worth pointing out explicitly. We observe significant correlations between the various E/P-related observables, F_{Brem}, and the supercluster widths. These correlations appear to be signficantly different between signal and background. Furthermore, there are significant correlations between the isolation quantities of the same particle type but different cone sizes, and are also different between signal and background. Therefore, the BDT yields greater discrimination when it makes use of these correlations.

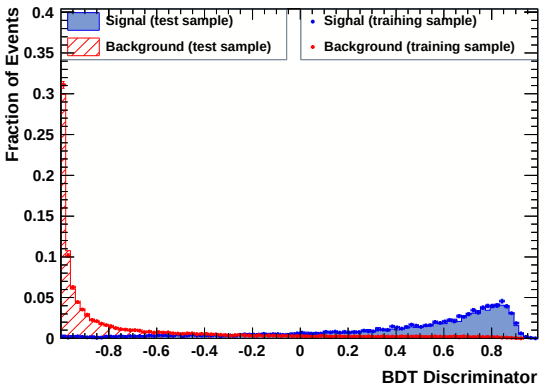


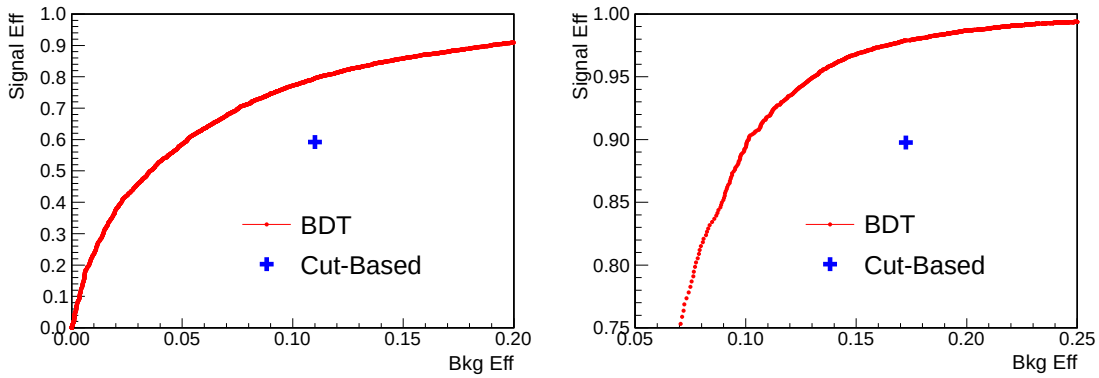
Figure 3-11: The electron BDT discriminator output for signal and background for the kinematic bin with electrons of p_T between 10 GeV/c and 20 GeV/cThe filled histograms show the distributions of the BDT discriminator in the training sample and the points show the distribution of the BDT discriminator in the statistically independent test sample. We observe no large differences between the training sample and test sample distributions indicating that the degree of overtraining is small.

the signal and background efficiencies for each cut in the discriminator variable. The performance for the BDT electron selection discriminator is shown in the ROC curves in Figures 3-12, 3-13, and 3-14. These are compared with the signal and background efficiency obtained using a standard cut-based electron selection used in a past version of this analysis [61, 67]. For the same signal efficiency, we achieve reductions in the background rate of a factor of two to three.

The electrons that are finally selected for the analysis must pass the preselection cuts defined above and the cuts on the BDT discriminator given in Table 3.6, whose choices are primarily motivated by the desire to increase rejection of background while keeping the same signal efficiency as the conventional cut-based electron selection.

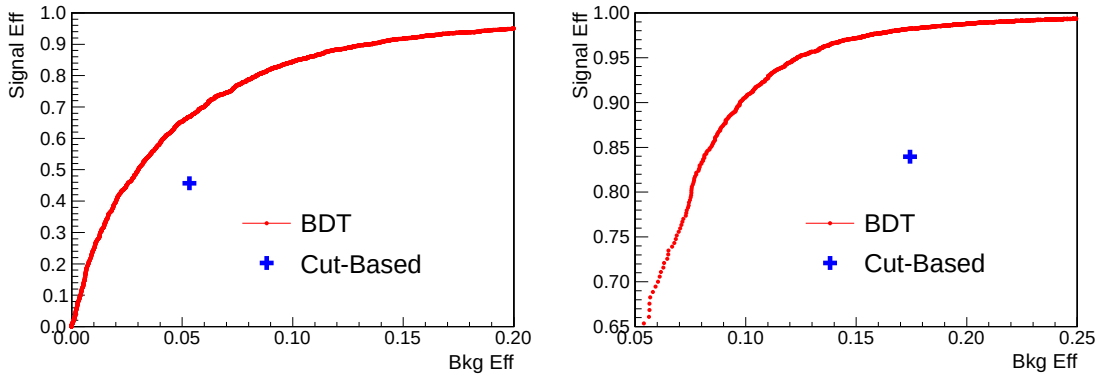
3.2.3 Muon Reconstruction

Muon reconstruction [68] begins with track segment reconstruction using hits in the DT and CSC muon detector systems. First, track segments are reconstructed within each muon chamber, which are loosely matched and combined to form seeds for muon



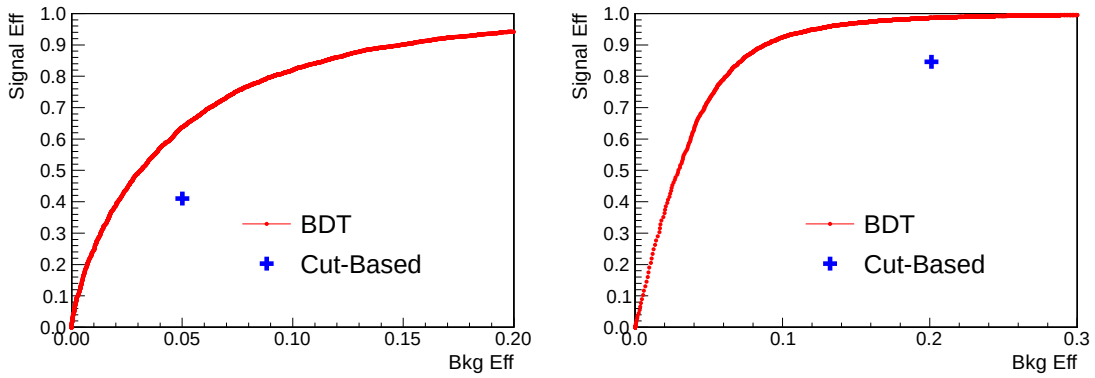
(a) 10 GeV/$c \leq p_T <$ 20 GeV/c (b) 20 GeV/$c \leq p_T <$ 35 GeV/c

Figure 3-12: Signal efficiency versus background efficiency curves obtained by varying the cut on the value of the BDT discriminator for electrons with $0.0 \leq |\eta| < 1.0$. The blue cross shows the signal and background efficiency for a cut-based electron selection used in past versions of this analysis. We observe that the BDT achieves reductions in the background rate by a factor of two to three relative to the cut-based electron selection at the same signal efficiency.



(a) 10 GeV/$c \leq p_T <$ 20 GeV/c (b) 20 GeV/$c \leq p_T <$ 35 GeV/c

Figure 3-13: Signal efficiency versus background efficiency curves obtained by varying the cut on the value of the BDT discriminator for electrons with $1.0 \leq |\eta| < 1.479$. The blue cross shows the signal and background efficiency for a cut-based electron selection used in past versions of this analysis.



(a) 10 GeV/$c \leq p_T <$ 20 GeV/c (b) 20 GeV/$c \leq p_T <$ 35 GeV/c

Figure 3-14: Signal efficiency versus background efficiency curves obtained by varying the cut on the value of the BDT discriminator for electrons with $1.479 \leq |\eta| < 2.5$. The blue cross shows the signal and background efficiency for a cut-based electron selection used in past versions of this analysis.

p_T Bin	η Bin	BDT Discriminator Cut Value
$10 < p_T \leq 20$	$0.0 < \|\eta\| < 1.0$	BDT > 0.420
$10 < p_T \leq 20$	$1.0 < \|\eta\| < 1.479$	BDT > 0.621
$10 < p_T \leq 20$	$1.479 < \|\eta\| < 2.5$	BDT > 0.619
$20 < p_T$	$0.0 < \|\eta\| < 1.0$	BDT > 0.959
$20 < p_T$	$1.0 < \|\eta\| < 1.479$	BDT > 0.959
$20 < p_T$	$1.479 < \|\eta\| < 2.5$	BDT > 0.928

Table 3.6: Summary of the BDT cut values used to select electrons.

reconstruction. Each seed track is extended using an iterative tracking procedure, where the candidate trajectory is propagated from one muon detector layer to the next, accounting for energy loss and multiple scattering in material. This procedure results in a set of muon candidates referred to as "standalone" muons, because only information from the muon detectors are used for their reconstruction.

Next, one attempts to match standalone muons with tracks reconstructed in the silicon tracker to form muon candidates referred to as "global" muons because they use the full information from the muon system and the tracking system. First, a loose rectangular region defined by the measured uncertainties in the direction of the standalone muon is used to find geometrically matching tracks. The trajectory defined by these tracks and the trajectory defined by the standalone muons are propagated to a common surface, typically chosen to be the tracker outer boundary or the muon system inner boundary, where various matching criteria are applied. A number of observables including spatial parameters, momentum parameters, and combined variables are used to match the standalone muons with the standard reconstructed tracks.

Finally, for each pair of standalone muon matched to a silicon track, a global fit is performed using the individual hits that comprise the standalone muon and the silicon track. This procedure results in a global muon candidate. For each standalone muon, a unique global muon is chosen with the best global fit χ^2 result, in the case of ambiguity.

Muons with badly measured segments in the muon chambers may fail the standalone reconstruction sequence entirely. In order to recover such muons we employ an alternate muon reconstruction sequence which begins by considering all well reconstructed silicon tracks. These tracks are propagated through the calorimeter and into the muon stations, accounting for possible energy loss and multiple scattering. Track segments reconstructed in each muon station are uniquely associated to candidate tracks by finding the nearest geometric match. Muons reconstructed by this alternative algorithm are referred to as "tracker muons". In order to reduce backgrounds,

requirements are placed on the compatibility of the measured calorimeter energy deposition with a muon hypothesis, as well as the quality of the geometric matching of the associated muon segments. Finally, to suppress background muons to sufficiently low levels we enforce that the muon penetrated through to the last muon station by requiring that there is a reconstructed track segment in the outermost muon station. Furthermore, the silicon track trajectory is required to match the segment in the outermost muon station geometrically to within three standard deviations in both transverse local coordinates. For this analysis, this alternative muon reconstruction adds roughly $2-3\%$ absolute efficiency to the muon reconstruction. As a side note, the tracker muon reconstruction is particularly useful for recovering muons with low transverse momentum, which fail the global muon reconstruction sequence more often because they are unable to penetrate fully through the material to the last muon stations. However, allowing tracker muons which do not penetrate through to the last muon station would introduce too much background for this analysis, and are therefore not used.

3.2.4 Muon Identification and Isolation

Standard muon identification at CMS imposes requirements on the quality of the global muon track and the track reconstructed using only the silicon tracker information. Discriminating variables include the number of hits, the normalized χ^2 measure of the goodness of fit, and the impact parameter. A particle must traverse at least 11.8 nuclear interaction lengths through the calorimeter system before reaching the muon chambers. As a result, the background from charged hadrons which traverse this material with minimal interaction, a process referred to as "punchthrough", is almost negligible. The primary sources of the background are secondary muons produced from semi-leptonic decay of heavy flavor hadrons, and decay in flight of pions and kaons. As a result, isolation is the single most important observable to suppress background muons.

As for electrons, we first give a general description of the methods used in typical

cut-based muon selection criteria. As an example, such cut based muon selections have been used in references [60, 61]. We follow with an in-depth description of the multivariate muon selection used in this analysis, which gives an increase in signal-to-background discrimination of more than 60%.

Isolation

Muons are required to be geometrically isolated from other charged particles and energy deposits. The procedure used is analogous to the case for electrons. For muons, we define the isolation in the more traditional way using reconstructed tracks, ECAL energy deposits, and HCAL energy deposits that fall within a geometric cone around the muon candidate direction. This differs from the treatment for electrons, for which it was observed that the use of particle flow candidates significantly improved the signal-to-background discrimination.

Tracks falling within an inner cone of $\Delta R < 0.01$ are not included to exclude tracks that were reconstructed from hits produced by the muon candidate itself. Tracks whose distance to the primary vertex at the point of closest approach to the beamline in the z coordinate is not within 0.1 cm are inferred to be produced from pileup interactions and therefore not included. In the isolation sum, we include only the electromagnetic energy deposits falling outside of the inner veto cone of $\Delta R < 0.07$ and hadronic energy deposits falling outside of the inner veto cone of $\Delta R < 0.10$. Muons will deposit some non-negligible amount of energy as it traverses through the material of the calorimeter. The inner veto cones are used to reduce the impact of this effect on the isolation. Finally, we include only electromagnetic energy deposits with $E_T > 0.2$ GeV, and hadronic calorimeter energy deposits with $E_T > 0.5$ GeV. These energy thresholds are imposed in order to reduce sensitivity to electronic noise.

Similarly to electrons, pileup interactions introduce energy into the isolation cone. The same procedure described for electrons is also used for muons to correct for this contamination, where $Z \rightarrow \mu^+\mu^-$ events are used to measure the effective areas summarized in Table 3.7. The statistical uncertainties in these measured values are

negligible, and any systematic errors that are made in this measurement will only make the correction suboptimal.

	EM Isolation 0.3 Cone	Hadronic Isolation 0.3 Cone
$0.0 < \|\eta\| < 1.0$	0.080	0.025
$1.0 < \|\eta\| < 1.479$	0.043	0.028
$1.479 < \|\eta\| < 2.0$	0.025	0.036
$2.0 < \|\eta\| < 2.25$	0.025	0.050
$2.25 < \|\eta\| < 2.5$	0.027	0.060
	EM Isolation 0.5 Cone	Hadronic Isolation 0.5 Cone
$0.0 < \|\eta\| < 1.0$	0.290	0.091
$1.0 < \|\eta\| < 1.479$	0.184	0.106
$1.479 < \|\eta\| < 2.0$	0.124	0.140
$2.0 < \|\eta\| < 2.25$	0.120	0.186
$2.25 < \|\eta\| < 2.5$	0.139	0.228
	Muon Hadronic Energy	Muon Hadronic S9 Energy
$0.0 < \|\eta\| < 1.0$	0.000	0.016
$1.0 < \|\eta\| < 1.479$	0.005	0.041
$1.479 < \|\eta\| < 2.0$	0.020	0.072
$2.0 < \|\eta\| < 2.25$	0.056	0.148
$2.25 < \|\eta\| < 2.5$	0.093	0.260

Table 3.7: Measurements of the effective area for muon isolation observables and variables that measure the energy deposited in the calorimeter towers crossed by the muon trajectory, in different bins of pseudorapidity. Significant dependence of the effective area on the pseudorapidity is observed near the edge of the tracking acceptance, in the region with η above 2.0. This dependence is the main motivation for the choice of η bins. For completeness, we also give the effective areas for the calorimeter energy measured in the tower crossed by the muon trajectory, and in the 3×3 square of towers centered on the muon trajectory (S9). These variables are described in more detail in the next section. The statistical uncertainties for these effective area values are negligible. Systematic errors in these measured values will only affect the optimality of the correction, and are not shown.

Multivariate Muon Selection

Analogous to electrons, we attempt to increase the sensitivity of the Higgs boson search by making use of multivariate methods to discriminate between signal and background muons. A BDT is trained using input observables related to the quality of the silicon track, the quality of the global muon track, the compatibility of track segments reconstructed in the muon stations with the extrapolated silicon track, the compatibility of the energy deposition in the ECAL and HCAL with a minimum ionizing particle, and the isolation.

The signal training sample is selected using dimuon events in data with dimuon mass between 75 GeV/c^2 and 105 GeV/c^2, very pure in $Z \rightarrow \mu^+\mu^-$ events. The background training sample is constructed from events triggered by the single muon triggers, HLT_Mu8 and HLT_Mu15, requiring that the event has one and only one muon candidate. Contamination from W events are suppressed by requiring that the missing transverse energy is less than 20 GeV/c, that the transverse mass formed using the muon candidate and the missing transverse energy is less than 20 GeV/c, and that the p_T of the muon candidate is less than 35 GeV/c. These samples are again explicitly divided in half based on the parity of the event number in order to allow for efficiency measurements that are unbiased by possible overtraining of the BDT.

Analogous to electrons, a set of preselection cuts are applied to the muon candidates before training. By explicitly removing regions of phase space that are clearly background dominated and sparsely populated by the training samples, the BDT is able to make more optimal use of the given information. The exact preselection cuts are as follows:

- more than 10 hits associated with the silicon track,
- at least one pixel detector hit associated with the muon,
- $|d_0| < 0.2$ cm,

- $|d_z| < 0.1$ cm,
- (Track Isolation + ECAL Isolation + HCAL Isolation)$/p_T < 0.4$,
- $\sigma(p_T)/p_T < 0.1$, and
- $\chi^2_{\text{kink}} < 20$,

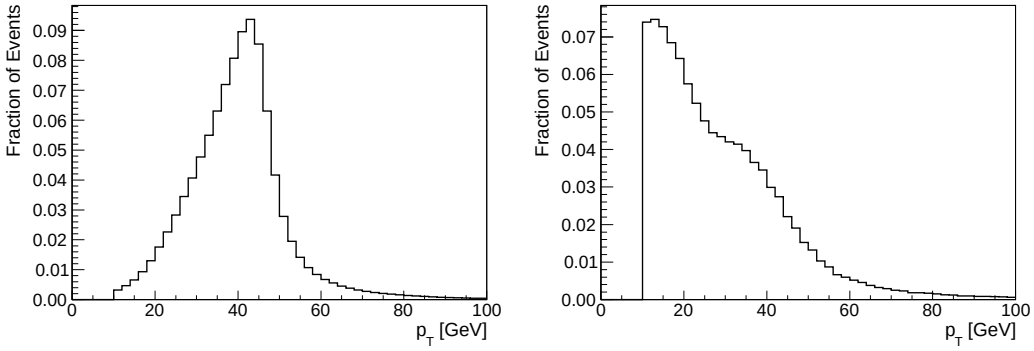
where d_0 is the transverse impact parameter of the electron track, d_z is the distance in the z coordinate between the point of closest approach of the electron track to the beamline and the primary event vertex, the isolation is computed using a cone size of 0.3, and the χ^2_{kink} observable is a measure that quantifies the probability that the muon was the result of a decay in flight. It is computed as the difference in the fit quality between the track built by propagating the trajectory starting from the innermost hits and the track built by propagating from the outermost hits. In addition, if the muon is reconstructed using the global muon algorithm, it is furthermore required to have at least two matching muon station segments, to have at least one muon station hit compatible with the global muon track fit, and to have a normalized χ^2 less than 10 for the global muon fit.

The p_T spectra of the training sample are reweighted to match the spectra of the signal and the background respectively. The original and reweighted spectra for the signal and background electron samples are shown in Figures 3-15 and 3-16. The pileup distribution of the background training sample are reweighted to the distribution for the signal, as for electrons.

We broadly classify the observables which enter as input to the BDT training into the following categories: track quality observables, impact parameter variables , calorimeter energy observables, and isolation.

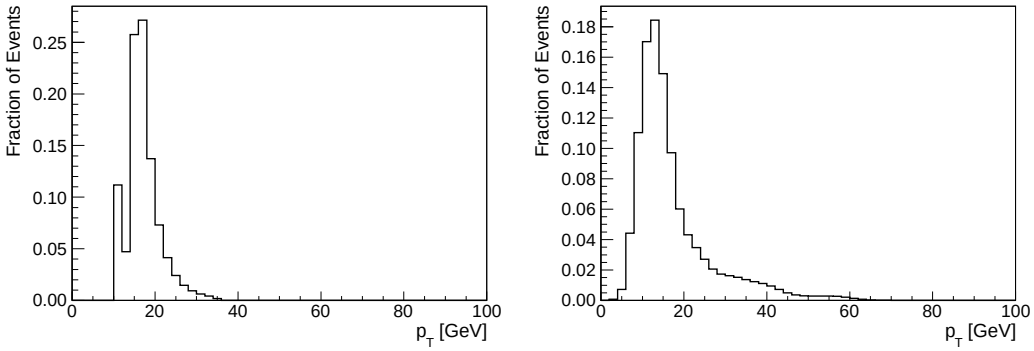
The track quality observables are:

- the number of hits associated with the track constructed using only the silicon and pixel detectors,
- χ^2/ Number of degrees of freedom for the silicon track,



(a) Unweighted (b) Weighted

Figure 3-15: The unweighted and reweighted p_T spectra of the signal muon training sample. The p_T spectrum of charged leptons resulting from the decay of a Z boson is peaked at around 45 GeV, while for the $H \rightarrow W^+W^-$ signal at low Higgs boson masses, there are a significant fraction of charged leptons that have p_T between 10 and 20 GeV/c. In order to be optimal for for a low mass Higgs boson, a significantly larger weight is required for lower p_T charged leptons.



(a) Unweighted (b) Weighted

Figure 3-16: The unweighted and reweighted p_T spectra of the background muon training sample. The p_T reweighting for the background training sample is needed primarily to smoothen the sharp turn on at 17 GeV due to the use of two different triggers with different p_T thresholds.

- the number of hits in the pixel detector associated with the track,
- χ^2/ Number of degrees of freedom for the global muon track fit,
- the number of muon station hits associated with the track,
- the number of track segments reconstructed in the muon stations that are matched to the muon candidate,
- the "SegmentCompatibility", a variable that quantifies the likelihood that the associated muon station segments are geometrically compatible with a muon candidate hypothesis, and
- χ^2_{kink} as defined above.

The signal and background distributions of the track related observables are shown in Figure 3-17.

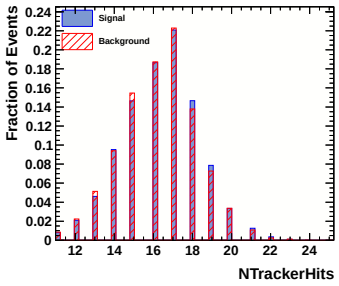
The impact parameter observables are:

- d_0, the transverse impact parameter,
- $d_{3\text{D}}$, the three dimensional impact parameter, and
- $d_{3\text{D}}/\sigma(d_{3\text{D}})$, the significance of the three dimensional impact parameter.

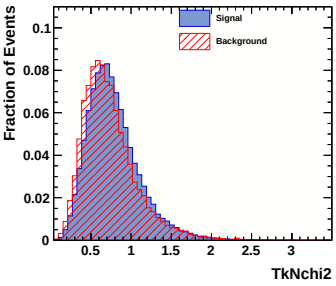
The signal and background distributions of the impact parameter observables are shown in Figure 3-18.

The observables characterizing the calorimeter energy deposition are:

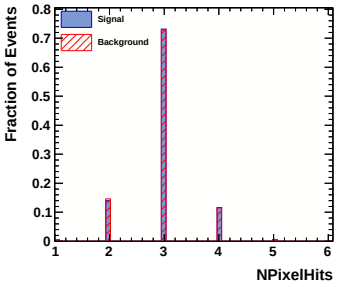
- Hadronic Energy $/p_T$, the amount of energy deposited within the HCAL towers crossed by the muon trajectory divided by the muon transverse momentum,
- EM Energy $/p_T$, the amount of energy deposited within ECAL crystals crossed by the muon trajectory divided by the muon transverse momentum,



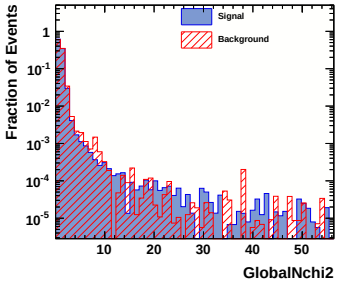
(a) Number silicon track hits



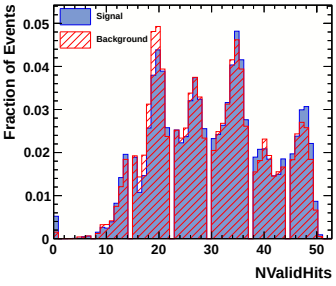
(b) Silicon track χ^2



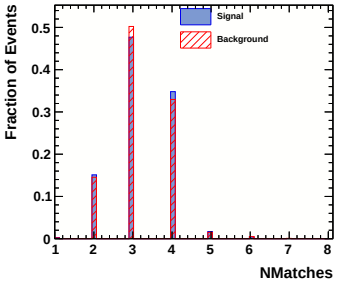
(c) Number of pixel hits



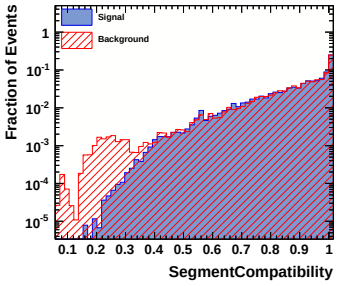
(d) Global muon fit χ^2



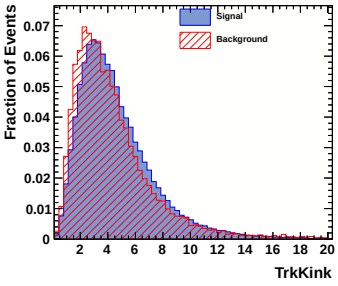
(e) Number of muon station hits



(f) Number of muon segments

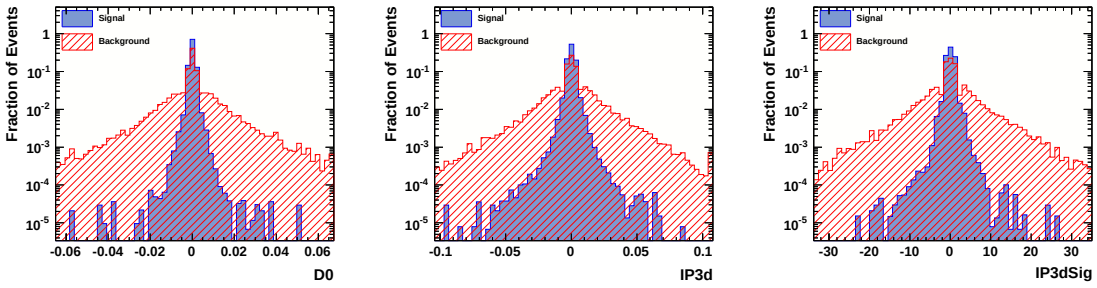


(g) SegmentCompatibility



(h) Kink

Figure 3-17: For illustration, we compare the signal and background distributions of the MVA input observables related to the muon track quality, for the kinematic bin with muons of p_T between 10 GeV/$c$and 14.5 GeV/c in the barrel region of the detector. All of these variables show rather weak discrimination, and will be combined by the BDT to yield a more powerful discriminator.



(a) d_0 (b) d_{3D} (c) $\frac{d_{3D}}{\sigma(d_{3D})}$

Figure 3-18: For illustration, we compare the signal and background distributions of the MVA input observables related to the impact parameter of the muon, for the kinematic bin with muons of p_T between 10 GeV/c and 14.5 GeV/c in the barrel region of the detector. The impact parameter variables are among the most discriminator variables for muons. Moreover, the correlations between them are larger for background than for signal.

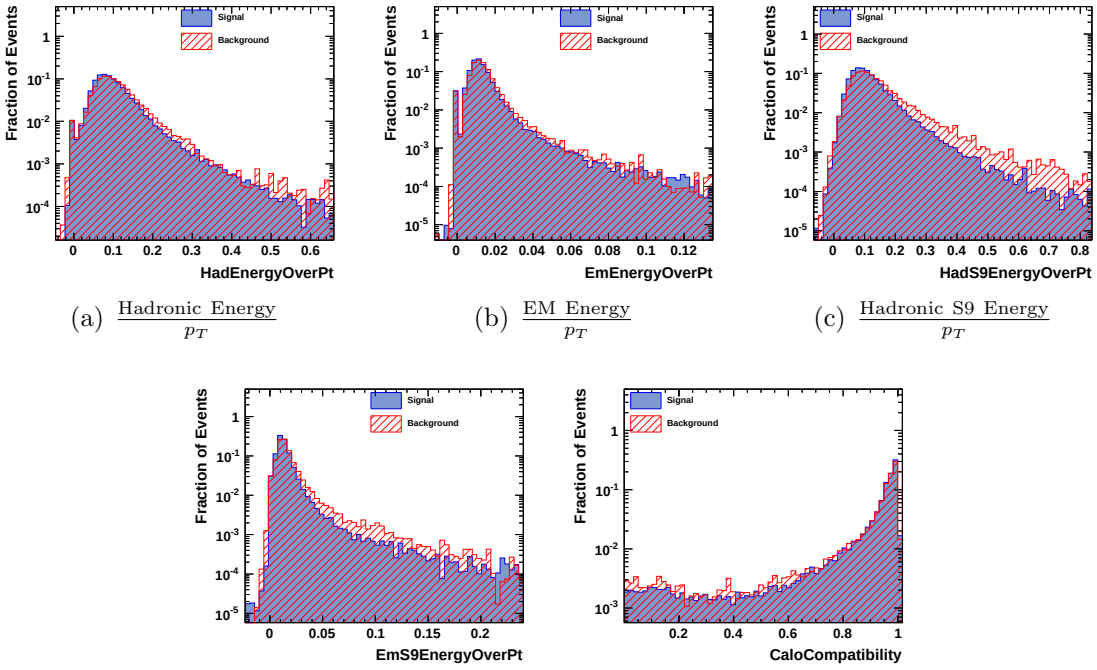
- Hadronic S9 Energy $/p_T$, the amount of energy deposited within the 3 by 3 square of HCAL towers closest to the muon trajectory divided by the muon transverse momentum,

- EM S9 Energy $/p_T$, the amount of energy deposited within the 3 by 3 square of ECAL crystals closest to the muon trajectory divided by the muon transverse momentum, and

- the "CaloCompatibility" observable, a variable that quantifies the likelihood that the observed calorimeter energy deposition in calorimeter towers crossed by the muon trajectory are compatible with the muon hypothesis.

These quantities have been corrected for the effect of pileup using the energy density based approach described in Sections 3.2.4 and 3.2.2, with the effective areas given by Table 3.7. The signal and background distributions of the calorimeter energy related observables are shown in Figure 3-19.

Finally, the isolation observables that we use are the pileup corrected track isolation, electromagnetic isolation, and hadronic isolation, as defined above, and with



(a) $\frac{\text{Hadronic Energy}}{p_T}$ (b) $\frac{\text{EM Energy}}{p_T}$ (c) $\frac{\text{Hadronic S9 Energy}}{p_T}$

(d) $\frac{\text{EM S9 Energy}}{p_T}$ (e) CaloCompatibility

Figure 3-19: For illustration, we compare the signal and background distributions of the MVA input observables related to the calorimeter energy deposit of the muon in calorimeter towers in near proximity to the muon trajectory, for the kinematic bin with muons of p_T between 10 GeV/c and 14.5 GeV/c in the barrel region of the detector. These variables exhibit fairly weak discrimination power, but yield additional information through their correlations with the isolation observables.

two different cone sizes of 0.3 and 0.5. The signal and background distributions of the isolation observables are shown in Figure 3-20.

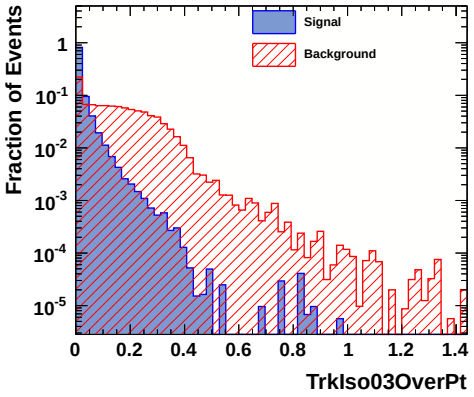
These input observables are passed to a boosted decision tree training, using the same training parameters as described above for the electrons. The training is performed separately in three bins of p_T: 10 GeV/c < p_T ≤ 14.5 GeV/c, 14.5 GeV/c < p_T ≤ 20 GeV/c, and p_T > 20 GeV/c. The bin boundary at 14.5 GeV/c is chosen just below the trigger turn on at 15 GeV/c to produce a fairly smooth p_T spectrum within each bin. Two bins of $|\eta|$, above and below 1.5, are chosen to distinguish the response of the muon in the barrel and endcap. For illustration, an example of the linear correlations for one particular bin is shown in Figure 3-21 for signal and for background, and an example of the BDT output is shown in Figure 3-22.

The performance gain in the signal-to-background discrimination is shown in Figures 3-23, 3-24, and 3-25, where we show curves of signal efficiency versus background efficiency obtained by varying the cut on the value of the BDT discriminator. These are compared with the signal and background efficiency obtained using a standard cut based muon selection, used in a past version of this analysis [61,67]. For the same signal efficiency, we achieve reductions in the background rate of roughly 40%.

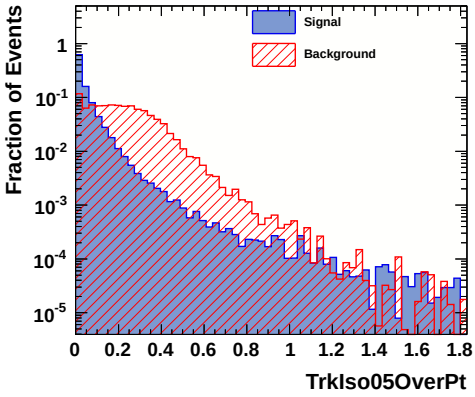
The muons that are finally selected for the analysis must pass the preselection cuts defined above and the cuts on the BDT discriminator given in Table 3.8. The cut value choices are motivated by the desire to increase rejection of background while keeping the same signal efficiency as the conventional cut-based muon selection.

Muon p_T/η Bin	BDT Discriminator Cut Value
$10 < p_T \leq 14.5$, $0.0 < \|\eta\| < 1.5$	BDT > −0.562
$10 < p_T \leq 14.5$, $1.5 < \|\eta\| < 2.4$	BDT > −0.300
$14.5 < p_T \leq 20$, $0.0 < \|\eta\| < 1.5$	BDT > −0.464
$14.5 < p_T \leq 20$, $1.5 < \|\eta\| < 2.4$	BDT > −0.248
$20 < p_T$, $0.0 < \|\eta\| < 1.5$	BDT > 0.171
$20 < p_T$, $1.5 < \|\eta\| < 2.4$	BDT > 0.815

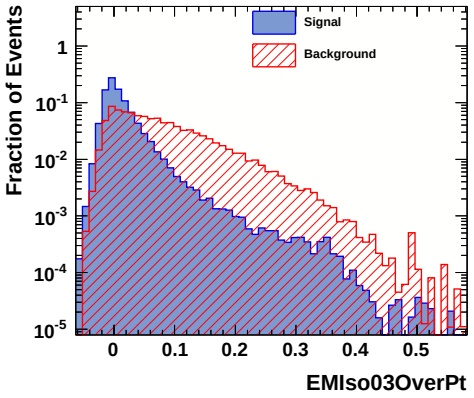
Table 3.8: Summary of the BDT cut values used to select muons.



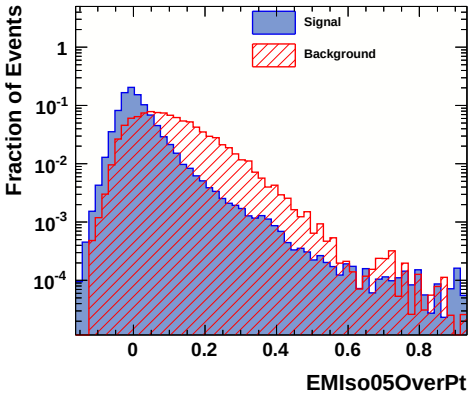
(a) Track Isolation 0.3 Cone



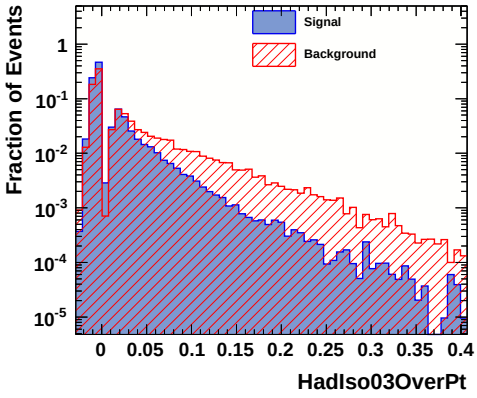
(b) Track Isolation 0.5 Cone



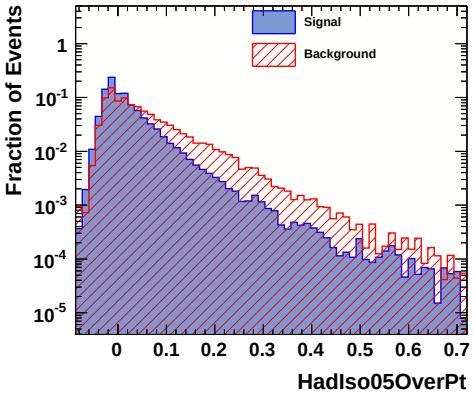
(c) EM Isolation 0.3 Cone



(d) EM Isolation 0.5 Cone

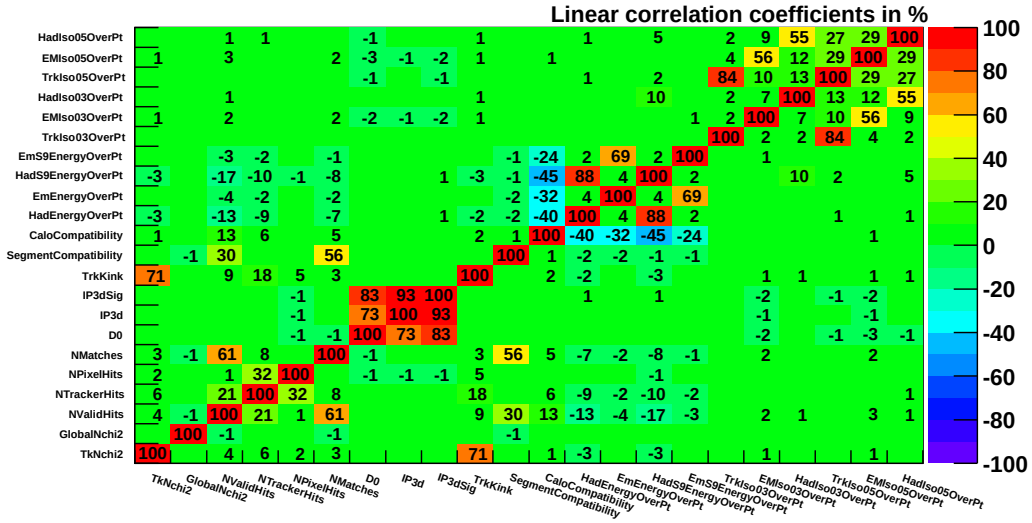


(e) Hadronic Isolation 0.3 Cone

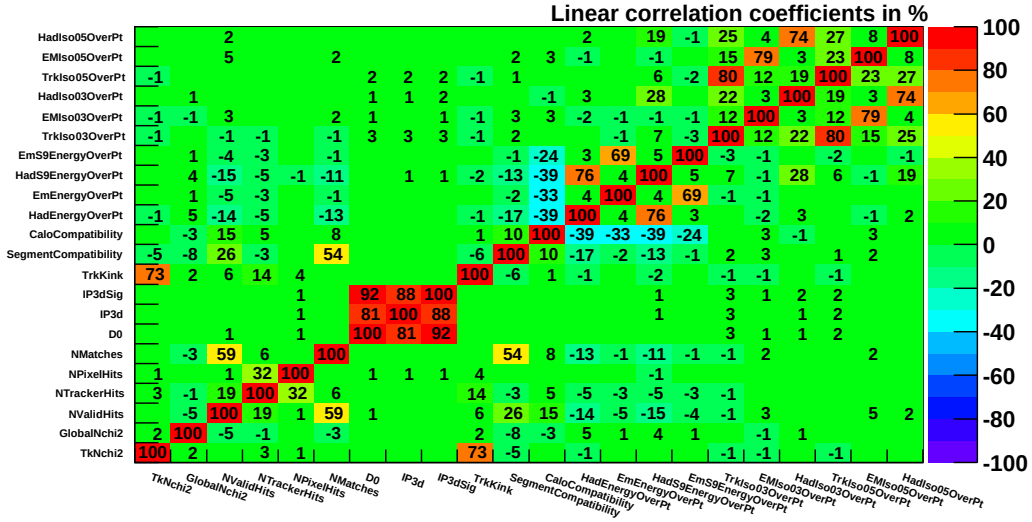


(f) Hadronic Isolation 0.5 Cone

Figure 3-20: For illustration, we compare the signal and background distributions of the muon isolation MVA input observables, for the kinematic bin with muons of p_T between 10 GeV/c and 14.5 GeV/c in the barrel region of the detector. The isolation variables are among the most discriminating ones. Making of the correlations between them also adds additional discrimination power.



(a) Signal



(b) Background

Figure 3-21: The matrix of linear correlation coefficients between all muon selection MVA input observables in the signal and background training samples, for the kinematic bin with muons of p_T between 10 and 14.5 GeV/c in the barrel region of the detector. A number of features are worth noting. We observe significant correlations between the three impact parameter variables, and the correlations are larger for background than for signal. The BDT yields better signal-to-background discrimination by making use of thses differences. Further discrimination power is gained from the correlations between the hadronic energy measured in the towers near the muon trajectory and the isolation, which are also different between signal and background.

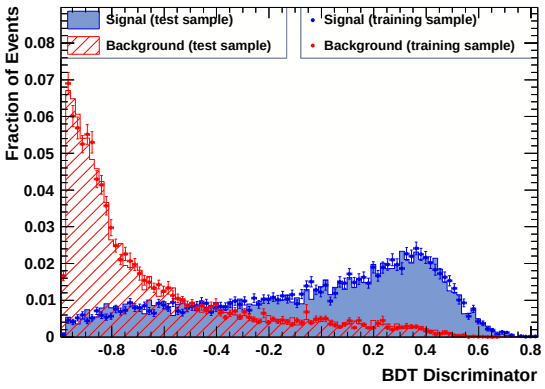
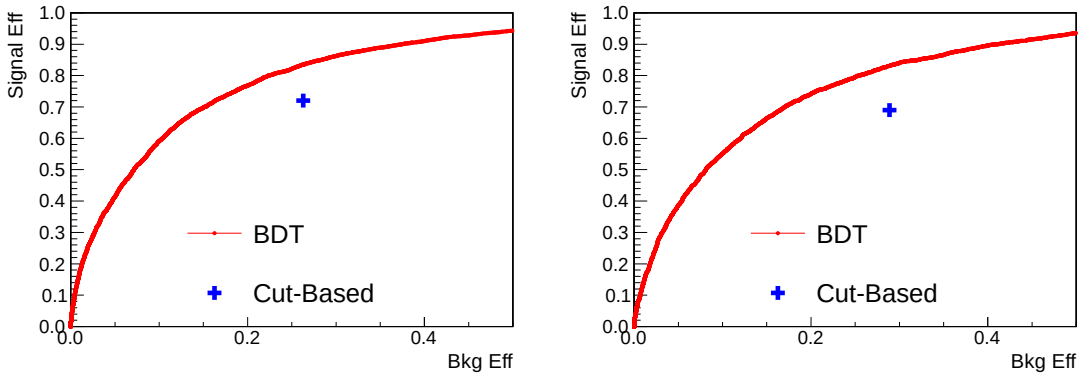
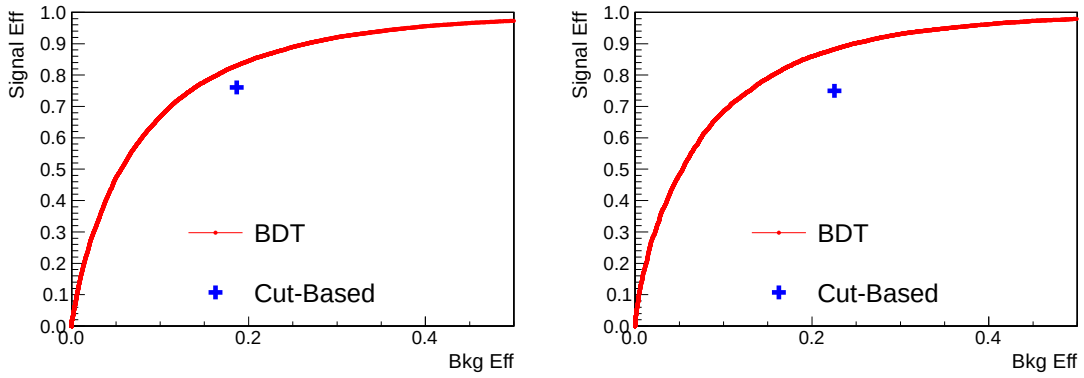


Figure 3-22: The muon BDT discriminator output for signal and background, for the kinematic bin with muons of p_T between 10 GeV/c and 14.5 GeV/c in the barrel region of the detector. The filled histograms show the distributions of the BDT discriminator in the training sample and the points show the distribution of the BDT discriminator in the statistically independent test sample. We observe no large differences between the training sample and test sample distributions indicating that the degree of overtraining is small.



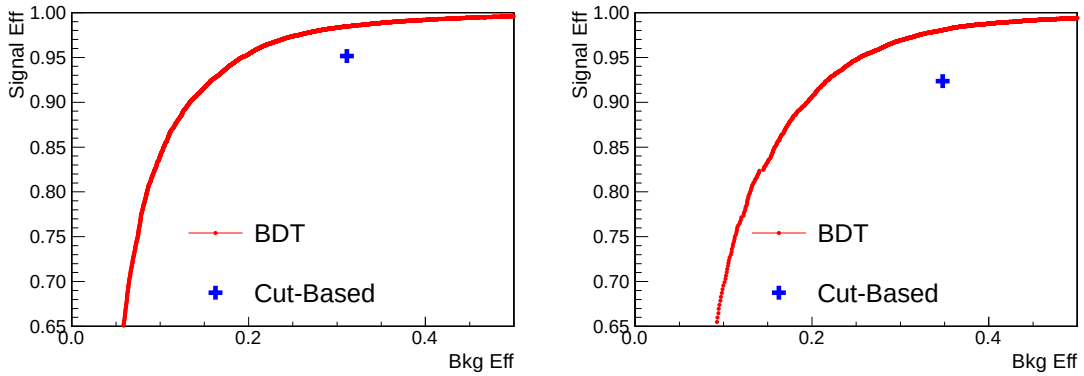
(a) $|\eta| \leq 1.5$ (b) $1.5 < |\eta| < 2.4$

Figure 3-23: Signal efficiency versus background efficiency curves obtained by varying the cut on the value of the BDT discriminator for muons with 10 GeV/c $< p_T \leq$ 14.5 GeV/c. The blue cross shows the signal and background efficiency for a cut-based muon selection used in past versions of this analysis. We observe reductions in the background rate by about 40% relative to the cut-based muon selection at the same signal efficiency.



(a) $|\eta| \leq 1.5$ (b) $1.5 < |\eta| < 2.4$

Figure 3-24: Signal efficiency versus background efficiency curves obtained by varying the cut on the value of the BDT discriminator for muons with 14.5 GeV/c $< p_T \leq$ 20 GeV/c. The blue cross shows the signal and background efficiency for a cut-based muon selection used in past versions of this analysis.



(a) $|\eta| \leq 1.5$ (b) $1.5 < |\eta| < 2.4$

Figure 3-25: Signal efficiency versus background efficiency curves obtained by varying the cut on the value of the BDT discriminator for muons with 20 GeV/c $< p_T \leq$ 35 GeV/c. The blue cross shows the signal and background efficiency for a cut-based muon selection used in past versions of this analysis.

3.3 Jet Selection

Jets are reconstructed in this analysis to be used to veto events for suppression of $t\bar{t}$ background. Therefore, our focus is to attempt to maintain as high a jet reconstruction efficiency as possible while keeping the rate of fake jets from noise and the influence from pileup interactions as low as possible.

3.3.1 Particle Flow Reconstruction

Particle flow refers to a method of reconstructing objects with very fine geometric granularity [69, 70], allowing for a global description of the event in terms of all particles that are produced. Conventionally, jets are constructed from energy deposits measured in calorimeter towers, defined by the angular segmentation of the hadronic calorimeter. In the particular flow reconstruction method, tracks are propagated from the last silicon tracking layer to the ECAL and the HCAL, to which energy clusters are associated if the track is found to cross within the energy cluster boundary. This track-to-energy-cluster association allows for a global description of the event and all of the particles produced in the event. The particles are classified into charged hadrons, neutral hadrons, photons, electrons, and muons, depending on properties of the track, the energy cluster, and the quality of their association. Various detailed noise cleaning algorithms are used to suppress its effect, employing both calorimeter shower shape and timing information. The combined use of the tracking and calorimeter measurements gives an overall improvement in the resolution of jet energy measurements of about 40% and an overall angular resolution improvement of more than a factor of two.

3.3.2 Jet Clustering

Such reconstructed "particle flow candidates" are combined to construct jets. The momentum of each particle flow candidate is used as input for jet clustering algorithms to produce jet candidates. We use the anti-Kt sequential jet clustering algorithm [71].

Pairs of particles are successively combined based on the distance between each other and the distance to the beam. For a given collection of particles, two sets of distances are computed corresponding to the distance between any two particles (d_{ij}) and the distance between any particle and the beam ($d_{i\mathrm{B}}$), respectively:

$$d_{ij} = \min(k_{T_i}^{2p}, k_{T_j}^{2p}) \frac{(y_i - y_j)^2 + (\phi_i - \phi_j)^2}{D^2}, \tag{3.3}$$

$$d_{i\mathrm{B}} = k_{T_i}^{2p}, \tag{3.4}$$

where k_{T_i},y_i,ϕ_i are the transverse momentum, the rapidity, and the azimuthal angle of particle i respectively. The parameter D characterizes the size of a jet in phase space, and p is a parameter that determines the relative power of energy and geometrical scale.

Next, one finds the smallest amongst these sets of distances. If the smallest distance is the distance between two particles, then one combines the two by adding their momenta. If the smallest distance is the distance between some particle and the beam, then one removes that particle from the set and promotes it to a jet. The procedure is repeated until no particles are left to be clustered. It has been demonstrated that under the addition of particles from soft and collinear radiation, the jets that are reconstructed by this algorithm do not change significantly, a property referred to as "infrared safety". Only infrared safe observables can be accurately predicted by perturbative calculations. For each choice of the parameter p, one obtains a different set of jets with different properties. The anti-Kt algorithm is the one defined by $p = -1$. It has been shown that the anti-Kt algorithm exhibits geometrical shapes that are very similar to the conventional cone-type algorithms, avoiding the problem of unusual jet shapes that have affected past sequential clustering algorithms [72, 73]. For this analysis, we use anti-Kt jets with the size parameter D equal to 0.5.

3.3.3 Jet Energy Corrections

Nonuniformities in the response of jets are corrected for by applying jet energy corrections. Due to the use of the particle flow reconstruction algorithm, the overall jet energy response is much improved relative to conventional jets, reducing the size of the jet energy correction by a factor of about three.

The jet energy corrections are applied in a factorized scheme. An offset correction, is applied to subtract energy originating from electronic noise and pileup. The energy introduced by particles produced in pileup interactions are subtracted using the FastJet technique [62–64], already described above for lepton isolation. The correction is computed from the product of the energy density of the event and the area of the particular jet. Next, we apply the relative correction which symmetrizes the jet response as a function of the pseudorapidity of the jet. Finally, the absolute correction is applied in bins of transverse momentum, correcting the energy response of the jet. These jet energy corrections are measured using dijet, photon-jet, and Z-jet balancing techniques [74].

For this analysis, we define a "counted" jet as a jet with corrected p_T larger than 30 GeV. The number of counted jets is used to categorize events into events with no counted jets (0-jet bin), events with one counted jet (1-jet bin), and events with two or more counted jets (2-jet bin). The counted jets are to be distinguished from the jets used for b-tagging, described later in Section 3.5, which may generally have lower p_T.

3.4 Missing Transverse Energy

Neutrinos produced in the collision event will, most of the time, traverse through all material of the CMS detector without interactions, producing no signal or energy deposition in the calorimeters. The vectorial momentum in the transverse plane is a conserved quantity. Therefore, the neutrino signature can be identified as an imbalance of the measured energy and momentum of all particles in the transverse

plane. We define the particle flow missing transverse energy (MET) observable as the negative vectorial sum of all particle flow candidates reconstructed using the algorithm described above in Section 3.3.1:

$$E_{\mathrm{T}}^{\mathrm{miss}} = \left| - \sum_{i \in \text{ PF Candidates}} \overrightarrow{p_T}_i \right|, \tag{3.5}$$

where $\overrightarrow{p_T}$ is the momentum vector of the particle flow candidate in the transverse plane [75]. The particle flow reconstruction combines the energy deposited by a muon traversing the calorimeter with the track momentum measurement of the muon, suppressing the need for additional missing transverse energy corrections that are typically required using conventional energy measurements alone.

If the energy of an electron or the momentum of a muon is mismeasured, a corresponding mismeasurement of the missing transverse energy will be observed. In the case of background, where there is no real neutrino escaping detection, the direction of the missing transverse energy induced by the mismeasured lepton momentum will tend to align in the direction of this lepton. Such a configuration is also observed in the case of a $Z \to \tau^+\tau^-$ event, where the tau lepton decays to an electron or a muon, and two neutrinos, because the neutrino momenta will tend to be in the same direction as the lepton momentum due to the large boost of the decaying tau lepton. In order to suppress these two types of background events, we define the projected $E_{\mathrm{T}}^{\mathrm{miss}}$ to be:

$$\text{Projected } E_{\mathrm{T}}^{\mathrm{miss}} = \begin{cases} E_{\mathrm{T}}^{\mathrm{miss}} & \text{if } \Delta\phi_{min} > \frac{\pi}{2}, \\ E_{\mathrm{T}}^{\mathrm{miss}} \sin(\Delta\phi_{min}) & \text{if } \Delta\phi_{\mathrm{min}} < \frac{\pi}{2} \end{cases} \tag{3.6}$$

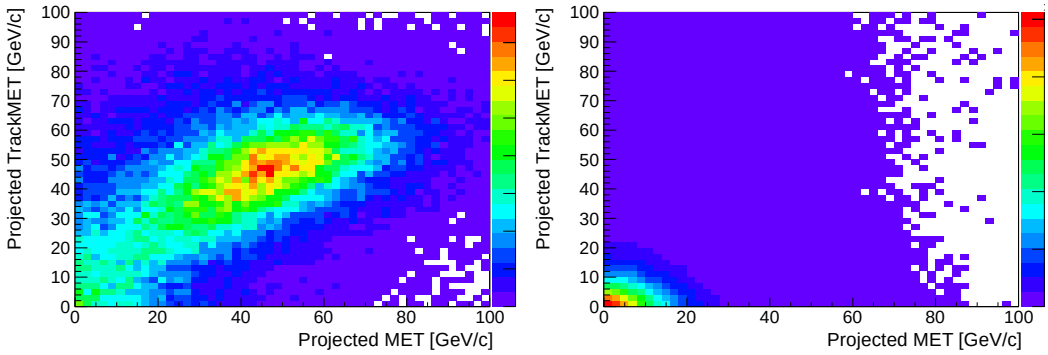
where $\Delta\phi_{min} = \min(\Delta\phi(\ell_1, E_{\mathrm{T}}^{\mathrm{miss}}), \Delta\phi(\ell_2, E_{\mathrm{T}}^{\mathrm{miss}}))$, and $\Delta\phi(\ell_i, E_{\mathrm{T}}^{\mathrm{miss}})$ is the angle between the missing transverse energy and the i'th lepton in the transverse plane. If the $E_{\mathrm{T}}^{\mathrm{miss}}$ points towards the same hemisphere as the closest lepton, we project the $E_{\mathrm{T}}^{\mathrm{miss}}$ perpendicular to the direction of the closest lepton in the transverse plane.

Due to the large amount of pileup interactions, the additional energy deposited in the calorimeters significantly degrades the missing energy resolution. To address this effect of pileup, we introduce an additional projected "trackMET" observable computed using Equations 3.5 and 3.6, where we only sum over charged particle flow candidates with $|d_{\mathrm{z}}| < 0.1$ cm, so that it is well associated with the primary vertex of the event. As a result, the resolution of the trackMET does not depend on the number of pileup interactions. However it does introduce a non-Gaussian tail because it does not account for neutral particles and charged particles with trajectories falling outside of the tracking acceptance.

The tail of the particle flow missing transverse energy distribution results from gaussian fluctuations introduced by the large amount of energy produced in the pileup interactions. The pileup does not significantly affect trackMET because charged particles from pileup may be removed with the cut on the d_z parameter. Instead, the primary source of the tail of the trackMET distribution are due to the fact that the neutral particles and the charged particles with trajectories falling outside of the tracking acceptance are not accounted for. Because the sources for the tails of the particle flow missing transverse energy and the trackMET are very different, one expects that they are largely uncorrelated for background. For signal they are largely correlated due to the presence of the neutrino. This is confirmed in Figure 3-26, where the two dimensional distribution is shown for signal and background. Thus, making a cut in the minimum of the two quantities, which essentially makes the requirement on both observables, will significantly suppress the background at a small loss of signal.

3.5 Top Tagging

Top production presents itself as a challenging background in the Higgs boson search decaying to W^+W^-, because it produces a W^+W^- boson pair. The only additional feature of top quark events is the presence of final state b-jets. We make use of two prominent features of b-jets in order to suppress the top background.



(a) Signal (b) Background

Figure 3-26: Two dimensional distributions of the projected particle flow missing transverse energy and the projected TrackMET is compared for signal and background. A significant positive correlation is observed between the $E_{\mathrm{T}}^{\mathrm{miss}}$ and the trackMET for the signal, while the two variables are largely uncorrelated for the background. We make use of this feature in the correlation to improve the signal-to-background discrimination.

In roughly 11% of cases, B flavored hadrons will decay semileptonically, resulting in a final state muon that is not necessarily isolated and fairly soft in p_T. We identify such events by searching for not necessarily isolated muons, in addition to the two primary selected leptons. Such a muon is required to have $p_T > 3$ GeV/c and the transverse impact parameter requirement is loosened to $|d_0| < 0.2$ cm. If such a "soft muon" is found, the event is rejected as a top event.

Furthermore, due to the longer lifetime of the B hadrons, b-jets tend to contain some number of tracks with impact parameters inconsistent with the hypothesis that they were produced at the primary vertex. This is quantified by what is called the "track counting" b-tagging variable, which is computed as the transverse impact parameter significance of the track with the second most significant transverse impact parameter [76, 77]. To suppress jets from pileup, we consider only jets with $p_T >$ 10 GeV/c, and we require that

$$\frac{|\sum^i d_z^i (p_T^i)^2|}{\sum^i (p_T^i)^2} < 2\mathrm{cm}, \tag{3.7}$$

where the sum is over all tracks that belong to the jet. This quantity is essentially a d_z observable for the jet. Any such jet with the value of the track counting observable larger than 2.1 is tagged, and rejected as a top event.

These two requirements are collectively referred to as top tagging. These top veto requirements suppress the top background by roughly a factor of two in the zero-jet bin and a factor of ten in the one-jet and two-jet bins. Top background events in the zero-jet bin only contain fairly soft b-jets, with p_T below 30 GeV, which have much lower b-tagging efficiency. As a result, it is more difficult to suppress top backgrounds in the zero-jet bin. Using $Z \to e^+e^-$ and $Z \to \mu^+\mu^-$ events, we measured the top tagging efficiency on non-top events, typically referred to as the mistag rate, to be less than 2%, demonstrating that the top veto requirements are fairly efficient on Higgs boson signal.

Chapter 4

Event Selection

A brief look at Figure 4-1 reveals that the rate for Higgs boson production at the LHC is ten orders of magnitude below the inclusive inelastic cross section for pp collisions and four orders of magnitude below the inclusive electroweak vector boson production cross section. Therefore, extracting the Higgs boson signal is generally extremely difficult. To suppress all of these background processes, we must make use of many different types of event features. This results in an analysis that relies on multiple detector components and reconstruction features, and one with a high degree of complexity. These features are described in greater detail below.

4.1 W^+W^--Preselection

The event selection is divided into two stages. In the first stage we select events with two leptons and relatively large missing transverse energy, characteristic of the presence of two W bosons. This W^+W^- event sample contains the signal in a highly diluted form, and is used as a control sample to extract important information that allow us to obtain an accurate estimate of the main backgrounds. In the second stage we apply further selection criteria in order to discriminate the Higgs boson signal from the W^+W^- background.

The first step in this analysis is to select events with two oppositely charged,

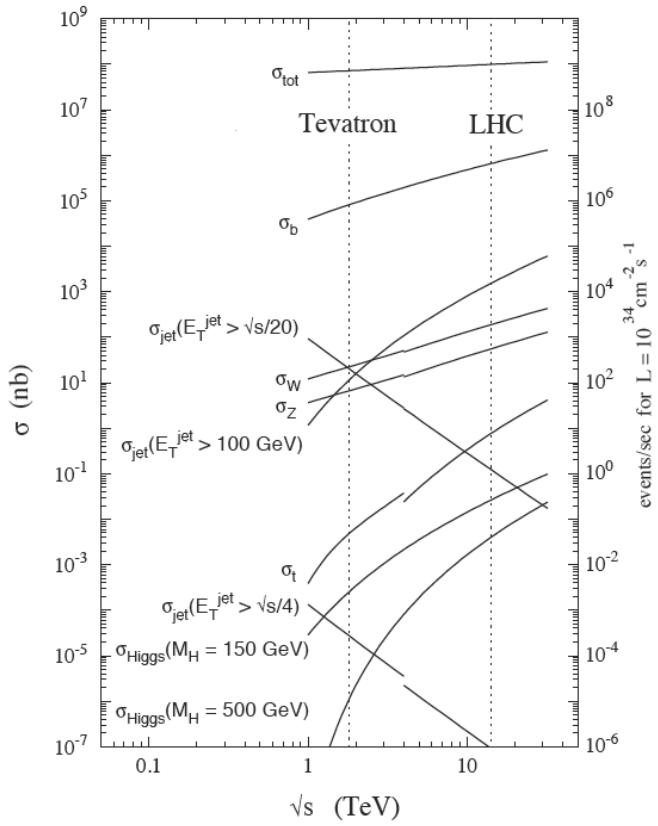


Figure 4-1: Cross-sections for various inelastic processes at proton-proton colliders as a function of the center of mass collision energy from the ATLAS Trigger and DAQ TDR [78]. The dotted line labelled LHC is at the original design center of mass energy of 14 TeV. The operating center of mass energy in 2011 was 7 TeV and is not explicitly marked.

prompt, and isolated leptons with high signal efficiency and high background rejection. The electron and muon selection criteria were described in detail in Section 3.2. After selecting two leptons with relatively high purity, the event sample is essentially completely dominated by the Drell-Yan process, the majority of which results from the Z boson resonance. To suppress the Z boson component, we reject any event with dilepton mass $m_{\ell\ell}$ within 15 GeV/c^2 of the Z boson pole mass (91.1876 GeV/c^2 [79]) in the final state with two electrons or two muons.

The missing transverse energy is typically small for Drell-Yan events because they do not contain neutrinos in the final state, and is primarily present due to energy

mismeasurements in the calorimeters. To suppress this background, we require that the minimum of the projected $E_{\mathrm{T}}^{\mathrm{miss}}$ and the projected trackMET to be less than 37.0 GeV/c $+0.5 \times N_{\mathrm{PV}}$, where N_{PV} is the number of reconstructed primary vertices, for the dielectron and dimuon final states. The requirement is loosened to 20 GeV/c for the electron-muon final state, where the Drell-Yan background is primarily due to $Z \to \tau^+\tau^-$. Because of the relatively small branching ratio of tau leptons to electrons and muons, the $Z \to \tau^+\tau^-$ background is much smaller. The distribution of the minimum projected missing transverse energy is shown in Figure 4-2. We observe that the $E_{\mathrm{T}}^{\mathrm{miss}}$ requirements described above suppresses the Drell-Yan background by more than three orders of magnitude. Due to the degradation of the missing transverse energy resolution as the number of pileup interactions increases, we introduced a dependence of the cut on the number of reconstructed primary vertices to make the efficiency of the cut for background more flat as a function of the number of pileup interactions. This effect is shown in Figure 4-3. Finally, to further suppress this background in events containing one or more jets with $p_T > 15$ GeV/c in the dielectron and dimuon final state, we require that the angle between the dilepton system and highest p_T jet in the transverse plane is larger than 165°. The effectiveness of this cut is illustrated by the comparison of this distribution after the $E_{\mathrm{T}}^{\mathrm{miss}}$ requirement for signal events and Drell-Yan background events shown in Figure 4-4.

After requiring large $E_{\mathrm{T}}^{\mathrm{miss}}$ in the event, the dominant background process is the production of a top and an anti-top quark pair and the production of a single top quark in association with a W boson. To suppress these top processes, we veto events which pass the top tagging criteria described in Section 3.5. The effectiveness of the top veto is illustrated in Figure 4-5, where we show the number of additional muons found in signal events compared to $t\bar{t}$ background events, and the distribution of the track counting b-tagging discriminator. We also use the fact that the distribution of the number of counted jets, shown in Figure 4-6, is different for signal and top background to divide the event sample into three exclusive classes with different signal-to-background ratios: events with no counted jets, events with one counted

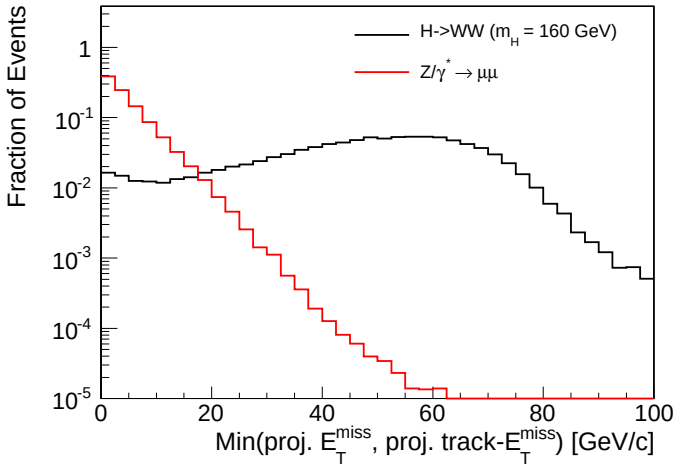


Figure 4-2: The normalized distributions of the minimum of the projected missing transverse energy and the projected charged particle missing transverse energy for $H \rightarrow W^+W^-$ events with a mass of 160 GeV/c^2 compared to $Z \rightarrow \ell^+\ell^-$ events on the left.

jet, and events with two or more counted jets. Further event selections are optimized in each jet bin.

Finally, we reject events with dilepton mass less than 12 GeV/c^2 in order to remove background from the Upsilon resonance, and events with the presence of a third lepton with $p_T > 10$ GeV/c in order to suppress the WZ and ZZ backgrounds.

4.2 Cut-Based Higgs Selection

In the second stage of the analysis selection we make use of specific signatures of the $H \rightarrow W^+W^-$ signal in order to separate it from the dominant standard model W^+W^- background. Two approaches are pursued. In the first approach, to establish a well controlled baseline analysis we use simple cuts on a small set of well discriminating observables. We begin by categorizing events into the zero-jet bin, the one-jet bin, or the two-jet bin depending on whether the event contains no counted jets, one counted jets, or two or more counted jets. In the two-jet bin, the $t\bar{t}$ background becomes prohibitively large for extracting the Higgs boson signal produced in the

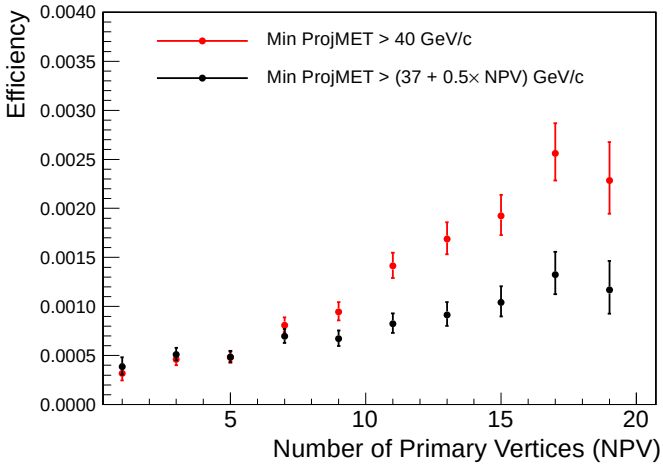


Figure 4-3: The efficiency of the minimum projected missing transverse energy cut for the Drell-Yan background obtained using Monte Carlo simulation is plotted as a function of the number of reconstructed primary vertices (NPV). We impose a cut that is dependent on the number of reconstructed primary vertices, shown in black, in order to flatten this efficiency versus the number of pileup interactions. A cut on the minimum projected missing transverse energy that does not depend on the number of reconstructed primary vertices, shown in red, is observed to increase significantly versus NPV because the missing energy resolution becomes significantly worse as the number of pileup interactions increases.

gluon fusion process. Instead, we focus on the VBF production mode of the Higgs boson by requiring the specific signature of two energetic forward jets with a large gap in rapidity to suppress the $t\bar{t}$ background. Thus, from here on the two-jet bin will be referred to as the VBF bin.

For the zero-jet and one-jet bins, the cuts that are applied are summarized in Table 4.1. We make use of the spin of the Higgs boson signal resulting in events with preferentially smaller angle between the two leptons in the transverse plane. We also make use of the transverse mass of the Higgs boson system defined by

$$m_T^{\ell\ell E_{\rm T}^{\rm miss}} = \sqrt{2 p_{\rm T}^{ll} E_{\rm T}^{\rm miss} (1 - cos(\Delta\phi_{\ell\ell - E_{\rm T}^{\rm miss}}))}, \tag{4.1}$$

where $\Delta\phi_{\ell\ell - E_{\rm T}^{\rm miss}}$ is the angle between dilepton direction and the $E_{\rm T}^{\rm miss}$ in the transverse plane, because the signal exhibits a broad peak structure in this observable.

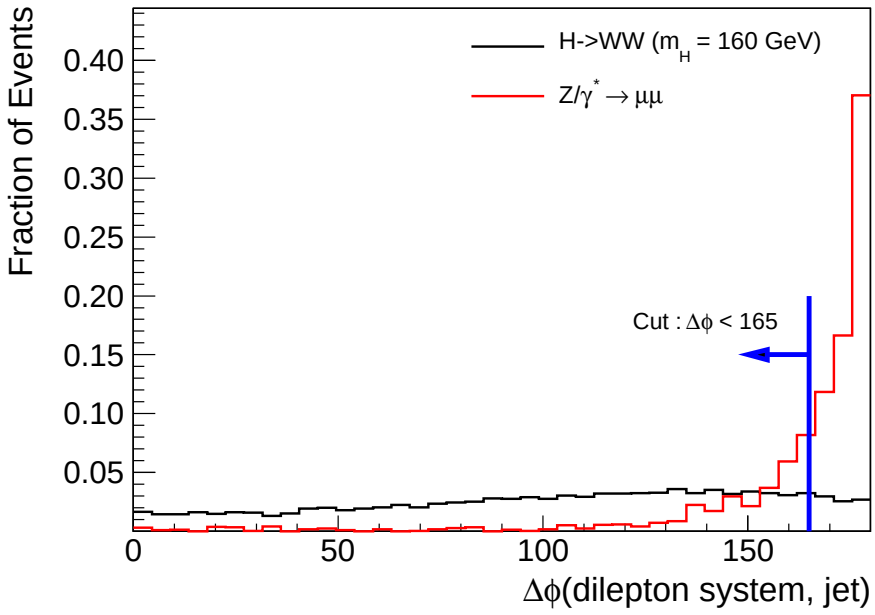
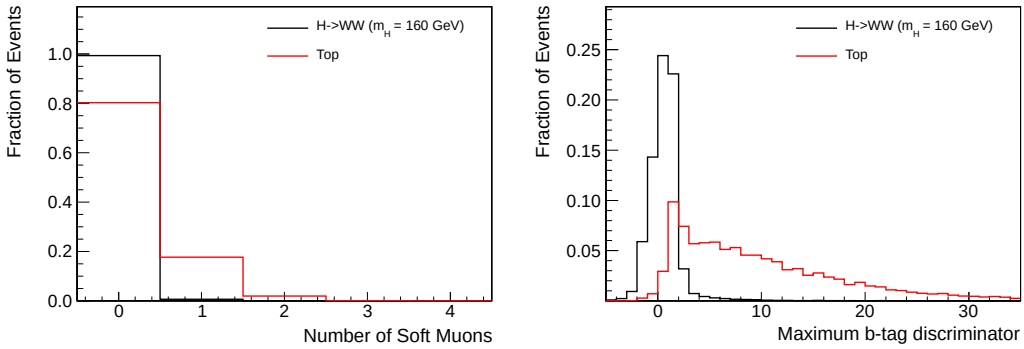


Figure 4-4: The normalized distributions of the angle between the momentum of the dilepton system and the direction of the highest p_T jet in the transverse plane for $H \rightarrow W^+W^-$ events with a mass of 160 GeV/c^2 compared to $Z \rightarrow \ell^+\ell^-$ events, after making the projected missing transverse energy requirement.

Finally, we make use of the transverse momentum of the dilepton system, p_{T}^{ll}, and the p_T of the two charged leptons. A comparison of the distributions of the observables used between the Higgs boson signal and the W^+W^- background are shown in Figure 4-7. The values of the cuts on each of the relevant variables are changing as a function of the hypothesized Higgs boson mass, and are the result of an optimization procedure. A two-sided window requirement is made on the transverse mass of the Higgs system to make use of the peaking structure. These cuts are imposed on events falling into the zero counted jet and one counted jet categories.



(a) Number of Additional Muons (b) Maximum b-tagging Discriminator

Figure 4-5: A comparison of two observables used to tag events as top background is made between signal and top background events. The normalized distribution of the number of additional, not necessarily isolated, muons is shown in (a) and of the b-tagging discriminator for the most significant b-tagged jet in (b).

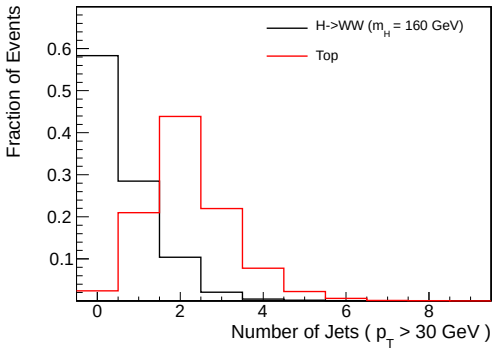
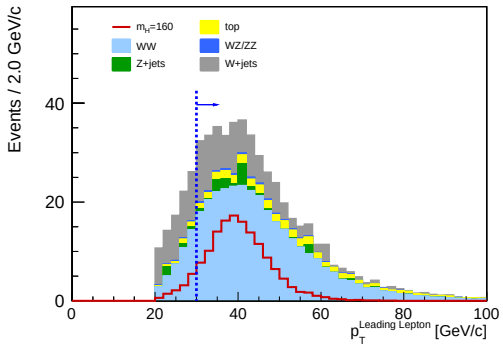


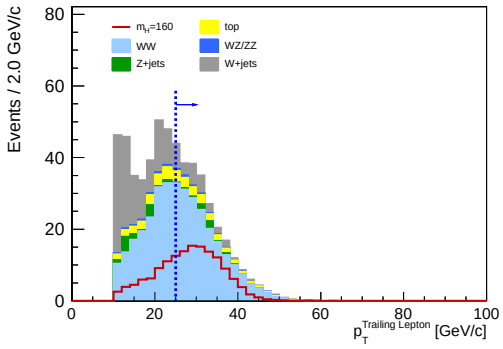
Figure 4-6: The normalized distributions of the number of jets with $p_T > 30$ GeV/c for signal and top background

m_{Higgs} [GeV/c^2]	$p_{\mathrm{T}}^{\ell,\mathrm{max}}$ [GeV/c]	$p_{\mathrm{T}}^{\ell,\mathrm{min}}$ [GeV/c]	$m_{\ell\ell}$ [GeV/c^2]	$\Delta\phi_{\ell\ell}$ [dg.]	$m_T^{\ell\ell E_{\mathrm{T}}^{\mathrm{miss}}}$ [GeV/c^2]
	$>$	$>$	$<$	$<$	[,]
115	20	10	40	115	[70,110]
120	20	10	40	115	[70,120]
130	25	10	45	90	[75,125]
140	25	15	45	90	[80,130]
150	27	25	50	90	[80,150]
160	30	25	50	60	[90,160]
170	34	25	50	60	[110,170]
180	36	25	60	70	[120,180]
190	38	25	80	90	[120,190]
200	40	25	90	100	[120,200]
250	55	25	150	140	[120,250]
300	70	25	200	175	[120,300]
350	80	25	250	175	[120,350]
400	90	25	300	175	[120,400]
450	110	25	350	175	[120,450]
500	120	25	400	175	[120,500]
550	130	25	450	175	[120,550]
600	140	25	500	175	[120,600]

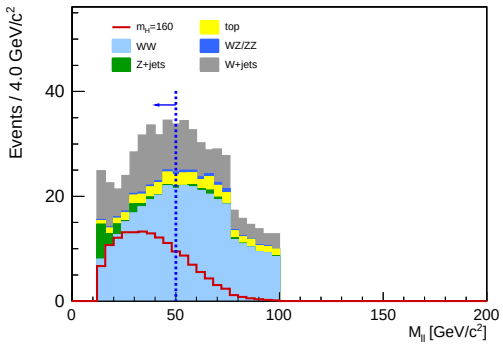
Table 4.1: Final event selection requirements for the cut-based analysis in the zero-jet and one-jet bins depending on the mass of the signal. One-sided cuts are imposed on the p_T of the two leptons, the dilepton mass, and the angle between the two leptons in the transverse plane. A two-sided cut is used for the transverse mass of the W^+W^- system.



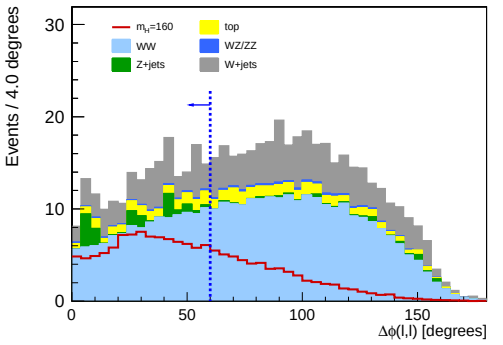
(a) Leading Lepton



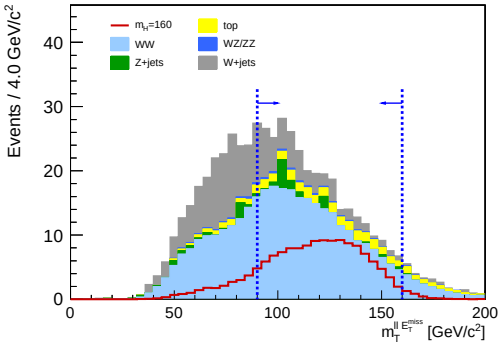
(b) Trailing Lepton



(c) Dilepton Mass



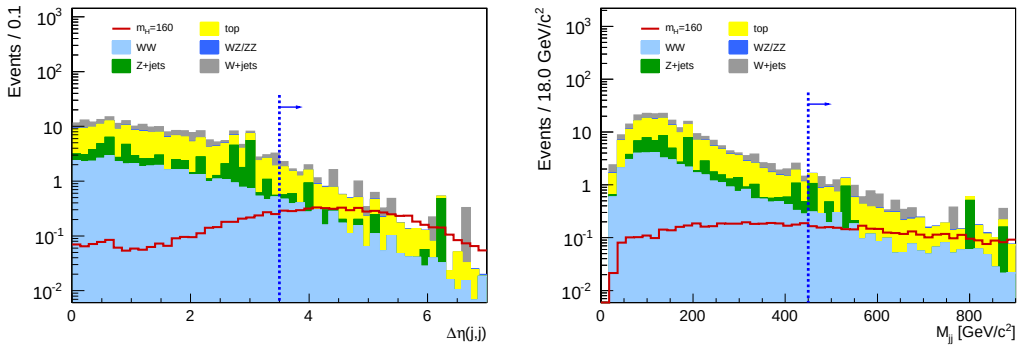
(d) $\Delta\phi_{\ell\ell}$



(e) $m_T^{\ell\ell E_{\mathrm{T}}^{\mathrm{miss}}}$

Figure 4-7: Distributions of observables used to distinguish $H \to W^+W^-$ signal from the W^+W^- background. The cuts used for the 160 GeV/c^2 Higgs cut-based selection are shown. The histograms are normalized to an arbitrary integrated luminosity.

In the VBF category, we require that the event contains two jets with $p_T > 30$ GeV/c. The $\Delta\eta(j_1, j_2)$ between the two jets is required to be larger than 3.5, and the dijet mass $m_{j_1 j_2}$ is required to be larger than 450 GeV/c^2. Finally we require that there are no other jets that lie between these two jets in pseudorapidity. The distribution of $\Delta\eta(j_1, j_2)$ and $m_{j_1 j_2}$ is shown in Figure 4-8 comparing the $H \to W^+W^-$ signal from VBF production mode with the $t\bar{t}$ background. These requirements suppresses the top background by more than one order of magnitude. Finally for hypothesized Higgs boson masses less than 200 GeV/c^2 we require that the dilepton mass is smaller than 100 GeV/c^2 in order to reduce the W^+W^- background.



(a) $\Delta\eta(j_1, j_2)$ (b) $m_{j_1 j_2}$

Figure 4-8: A comparison of the distributions of observables used to distinguish VBF $H \to W^+W^-$ signal from the $t\bar{t}$ background is made. The histograms are normalized to an arbitrary integrated luminosity.

4.3 Multivariate Selection

Once we have established a good understanding of the behavior of the backgrounds in the W^+W^- cntrol sample and the signal region defined by the cut-based analysis, we can make use of additional observables as well as correlations by employing multivariate discrimination methods. To establish an appropriate control region for the training of the MVA, we require that events pass the W^+W^- preselection, and that the dilepton mass satisfy the cuts given in Table 4.2 for each Higgs boson mass

m_{Higgs}	[GeV/c^2]	120	130	140	150	160	170	180	190	200	210
$m_{\ell\ell} <$	[GeV/c^2]	70	80	90	100	100	100	110	120	130	140
m_{Higgs}	[GeV/c^2]	220	230	250	300	350	400	450	500	550	600
$m_{\ell\ell} <$	[GeV/c^2]	150	230	250	300	350	400	450	500	550	600

Table 4.2: $m_{\ell\ell}$ upper limit requirement as a function of the Higgs boson mass used to enrich the background datasets of signal-like events. These samples are employed in the training of the multivariate classifier used for the signal extraction.

hypothesis.

The TMVA toolkit [66] is used to train a boosted decision tree to discriminate between events from the $H \rightarrow W^+W^-$ signal Monte Carlo simulation sample and events from the W^+W^- background Monte Carlo simulation sample. We do not train against other background components because we want the BDT to focus on the kinematic features that discriminate the signal against the dominant W^+W^- background. Furthermore, in the zero-jet bin, the contribution of other backgrounds are small after the W^+W^- preselection. The following observables were used as input to the BDT:

- leading lepton p_T,
- trailing lepton p_T,
- $\Delta\phi_{\ell\ell}$ the angle between the two leptons in the transverse plane,
- $\Delta\text{R}_{\ell\ell}$ the distance between the two leptons in η-ϕ space,
- the flavor of the leading and trailing lepton, and
- $m_T^{\ell\ell E_{\text{T}}^{\text{miss}}}$ the transverse mass of the W^+W^- system.

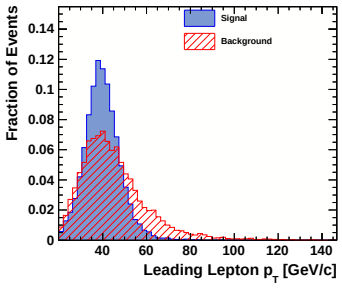
In the one-jet category, the top quark backgrounds are also included in the background training sample because they compose a significant fraction of the total background. We also add the following additional observables in the one-jet class to improve discrimination against top quark backgrounds:

- $\Delta\phi_{E_{\mathrm{T}}^{\mathrm{miss}},\ell\ell}$, the angle between the $E_{\mathrm{T}}^{\mathrm{miss}}$ and the dilepton system in the transverse plane, and

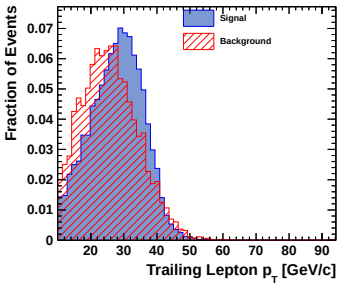
- $\Delta\phi_{E_{\mathrm{T}}^{\mathrm{miss}},\text{leading jet}}$, the angle between the $E_{\mathrm{T}}^{\mathrm{miss}}$ and the leading jet in the transverse plane.

We compare the signal and background distributions of the training sample for a Higgs boson with mass of 160 GeV in Figures 4-9 and 4-10, for the zero-jet and one-jet event classes respectively. We observe that all input variables contribute some amount of discrimination. In Figure 4-11 we compare the distribution of the BDT discriminator output for signal and background and explicitly observe the increase in discrimination power over any single observable. Due to limited expected signal yield with the current dataset in the VBF class of events, no significant gain is expected from a multivariate selection and is therefore not pursued.

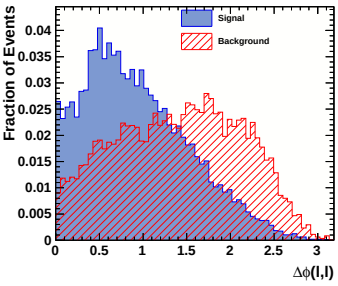
Finally, the shape of the output distributions of the MVA discriminator is used as the basis of the statistical analysis to extract an upper limit on the Higgs boson production cross section and the signal significance.



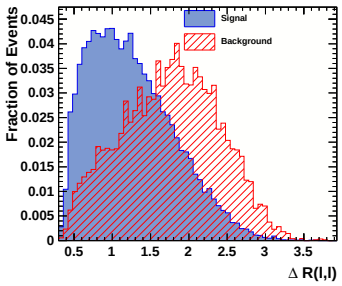
(a) Leading Lepton p_T



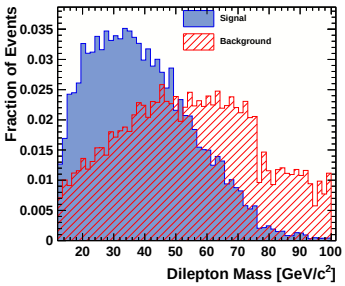
(b) Trailing Lepton p_T



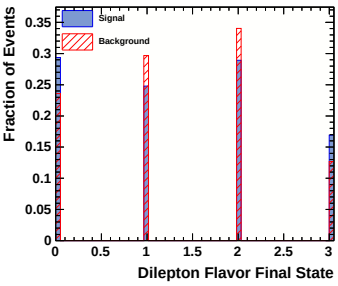
(c) $\Delta\phi(l,l)$



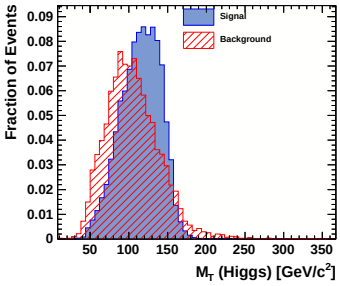
(d) ΔR(l,l)



(e) Dilepton Mass

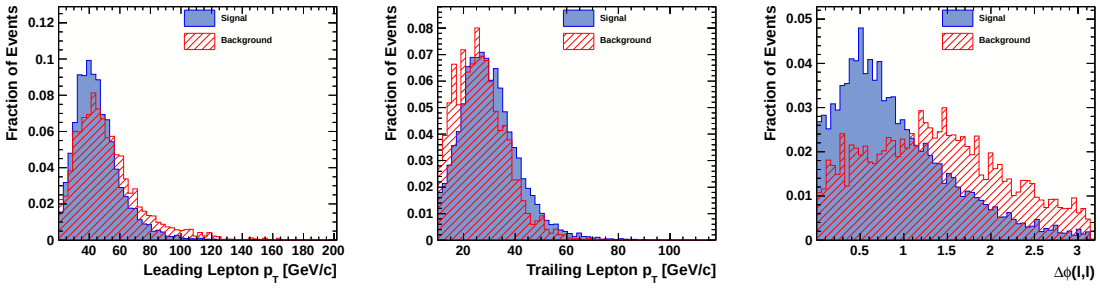


(f) Final State

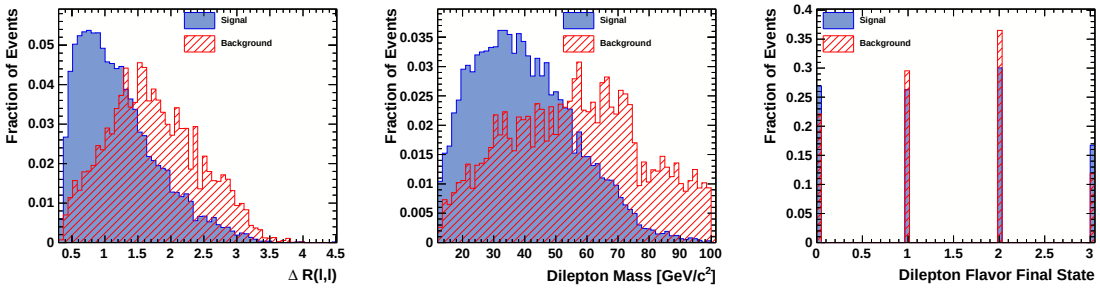


(g) $m_T^{\ell\ell E_{\mathrm{T}}^{\mathrm{miss}}}$

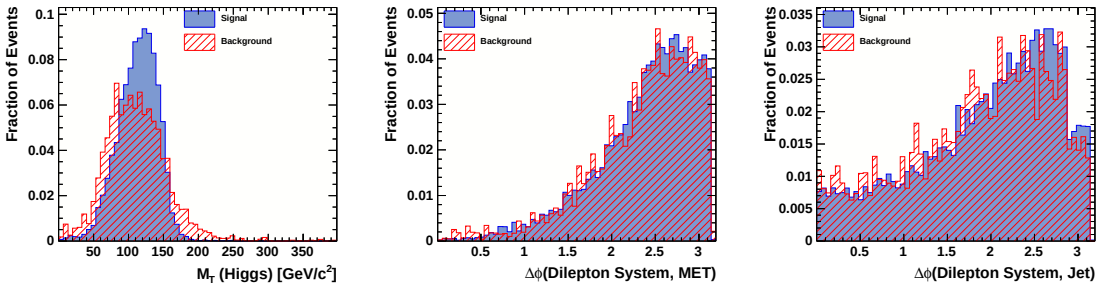
Figure 4-9: The distributions of the input observables to the BDT for the $H \rightarrow W^+W^-$ signal with a mass of 160 GeV/c^2 and the W^+W^- background, for events in the zero-jet bin



(a) Leading Lepton p_T (b) Trailing Lepton p_T (c) $\Delta\phi(l,l)$

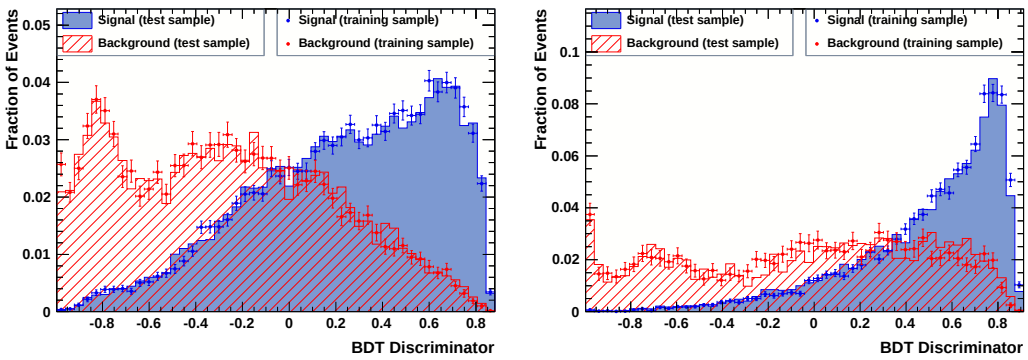


(d) ΔR(l,l) (e) Dilepton Mass (f) Final State



(g) $m_T^{\ell\ell E_{\mathrm{T}}^{\mathrm{miss}}}$ (h) $\Delta\phi$(Dilepton, MET) (i) $\Delta\phi$(Dilepton, Jet)

Figure 4-10: The distributions of the input observables to the BDT for a $H \rightarrow W^+W^-$ signal with a mass of 160 GeV/c^2 and the W^+W^- background, for events in the one-jet bin



(a) 0-Jet Bin

(b) 1-Jet Bin

Figure 4-11: The BDT discriminator distributions for a Higgs boson signal with a mass of 160 GeV/c^2 and the W^+W^- background are shown. The filled histograms show the distributions of the BDT discriminator in the training sample and the points show the distribution of the BDT discriminator in the statistically independent test sample. We observe no large differences between the training sample and test sample distributions indicating that the degree of overtraining is small.

Chapter 5

Event Selection Efficiency

In order to achieve an accurate prediction of the expected signal yield and the expected background contribution, the event selection efficiency must be determined precisely and accurately. The Monte Carlo simulation, which uses the GEANT4 program [80] to simulate interactions of particles with the detector material and components, provides an excellent approximation for the bulk of processes. However there may be some regions of phase space for which the simulation is not accurately tuned or modeled. Additionally, very rare effects may not be well simulated simply because it is practically impossible to simulate as many events as produced by LHC collisions.

For most event selection requirements, the Monte Carlo simulation is used to estimate its efficiency. Systematic uncertainties on these efficiencies are, then, estimated by comparing certain sensitive observables in various control regions. These estimates are described further in Chapter 7. One of the most important selection efficiencies is the lepton selection efficiency and these efficiencies can actually be measured fairly precisely in-situ using Z boson events which decay to a pair of electrons or a pair of muons. In this chapter, we discuss in detail the methods used to measure lepton selection efficiencies.

It is generally very difficult to measure the absolute lepton selection efficiency from the data. For this analysis, we nominally make use of the efficiency predicted by the Monte Carlo simulation for the most basic processes. For electrons, efficiency for re-

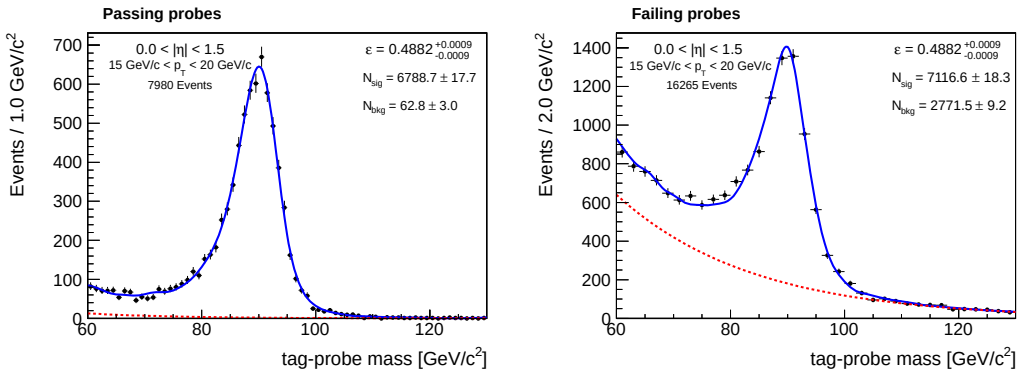
constructing energy clusters in the ECAL is predicted by the Monte Carlo simulation to be approximately 99.5%. The 0.5% inefficiency is entirely due to mismeasurement of the electron energy. The only way to have an actual energy clustering inefficiency is to have an electron which does not shower in the ECAL or to have some locally correlated hardware failure. These occurrences are negligible. For muons, the efficiency for reconstructing a track is predicted by the Monte Carlo simulation to be very near 100%. This efficiency has been cross checked using a sample of muons that are reconstructed using muon chamber measurements only [60].

Starting with an electromagnetic energy cluster for electrons and a track for muons, we measure the efficiency for these objects to be selected in two factorized steps, which is motivated primarily by convenience. In the first stage we measure the efficiency for one of these two basic objects to be successfully reconstructed as an electron or muon, and in the second stage we measure the efficiency for a reconstructed lepton to pass the specific lepton identification and isolation selection requirements. The measurements are performed in bins of p_T and η that are sufficiently small to capture significant dependencies.

5.1 Tag And Probe Method

For each step, the efficiency measurement is performed using the "tag and probe" technique. We use $Z \to e^+e^-$ or $Z \to \mu^+\mu^-$ events, where tight selection requirements are imposed on one leg of the Z decay in order to suppress backgrounds, and is denoted the "tag" leg. The other leg remains unbiased with no selection requirements, and is denoted the "probe" leg. This unbiased probe leg can be used to study the efficiency of any generic set of requirements. In the case where both legs satisfy the tight selection requirements, we consider both legs for the efficiency measurement. We measure the efficiency for the probe leg to pass a particular requirement. Depending on which step and which kinematic region we are measuring the efficiency for, the background contribution can be suppressed to negligible levels by requiring that the dilepton mass

is within a window around the Z boson pole mass. We take the range to be from 76 GeV/c^2 to 106 GeV/c^2. In the case where backgrounds are significant, we perform an unbinned likelihood fit simultaneously to the observed dilepton mass distributions of the sample for which the probe passed the selection requirements, called the "pass sample", and the sample for which the probe failed the selection requirements, called the "fail sample". For illustration, an example of the likelihood fit to the dilepton mass distributions for the pass and fail samples are shown in Figure 5-1.



(a) Probe Passes Selection (b) Probe Fails Selection

Figure 5-1: For illustration we show an example of the likelihood fit to the dilepton mass distribution for the pass and fail samples. The example shows the fits used to measure the electron selection efficiency for electrons in the barrel with p_T between 15 and 20 GeV/c. The signal model is obtained by a convolution of the mass shape template obtained from the Monte Carlo simulation with a Gaussian that accounts for possible mismodelling of the energy scale and resolution. The background model is the product of a linear function and an exponential. The red curve represents the fitted background. The blue curve represents the sum of the signal and the background.

The dilepton mass shape for the Z decay is taken from the simulation corrected by a convolution with a Gaussian function whose mean and width is allowed to float freely to account for possible differences in the scale and resolution of the lepton energy or momentum measurement. The dilepton mass shape for the background processes is taken to be an exponential times a linear function. Systematic uncertainties resulting from the inaccuracy of these shape parametrization have been evaluated in previous measurements to be less than 2% [60]. This modeling uncertainty is propagated as

an uncorrelated systematic uncertainty for each efficiency measurement.

The efficiencies measured using this tag and probe technique may be biased by the particular event selection requirements and by the typical kinematics of Z boson events. Therefore, rather than applying the measured efficiencies directly, we derive Monte-Carlo-to-data correction scale factors by dividing the measured efficiency in data by the efficiencies measured in Monte Carlo simulation using the same selection criteria. These correction scale factors are applied to any prediction of the yield that are derived from the simulation. This procedure ensures that the Monte Carlo simulation accounts for any event selection or kinematic biases, while the correction factor accounts for the differences between data and simulation.

5.2 Electron and Muon Reconstruction Efficiency

In the first step for electrons, we measure the efficiency for a reconstructed ECAL supercluster to be successfully reconstructed as an electron. This step primarily tests the efficiency for finding a pixel seed compatible in directionality with the supercluster and successfully building a GSF track candidate that loosely matches geometrically with the supercluster.

We measure this efficiency using the tag-and-probe method described above and verified that the efficiency estimated by the Monte Carlo simulation is consistent with the measurement in all kinematic bins. Thus, no correction factor is applied to the Monte Carlo estimate. The efficiency for electrons averaged over the full p_T range of the signal is about 97% in the barrel, and about 94% in the endcap. The systematic uncertainties intrinsic to the tag-and-probe method as well as the signal and background models used has been estimated to be about 2% [60]. The statistical uncertainties are negligible by comparison.

For muons, the reconstruction efficiency is factorized. We measure the efficiency for a reconstructed silicon track produced by a muon to be successfully reconstructed in the muon chambers as a standalone muon. Then, we measure the efficiency for

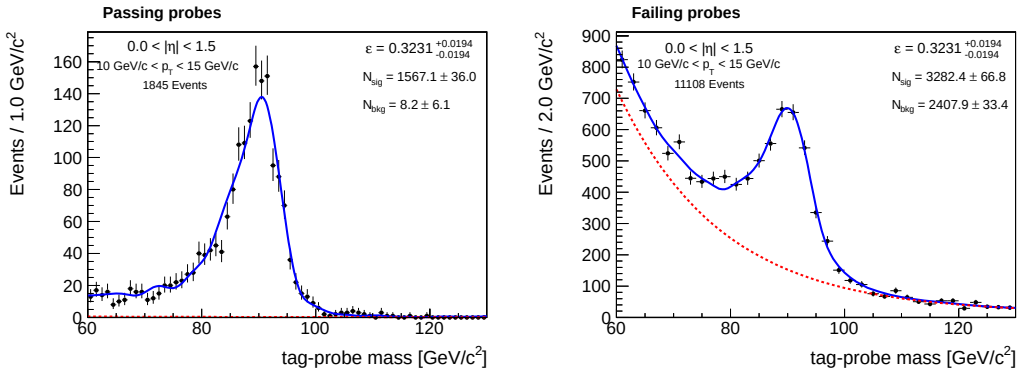
a reconstructed standalone muon to produce a track in the silicon tracker. The muon reconstruction efficiency is the product of these two efficiencies. Using the tag-and-probe method we measure efficiencies consistent with the estimate from the simulation, and no correction is applied to the Monte Carlo. Averaged over the full p_T and η range of the signal, the muon reconstruction efficiency is about 98%. The systematic uncertainty of 2% is progagted, while the statistical uncertainties are again negligible by comparison.

5.3 Electron Selection Efficiency

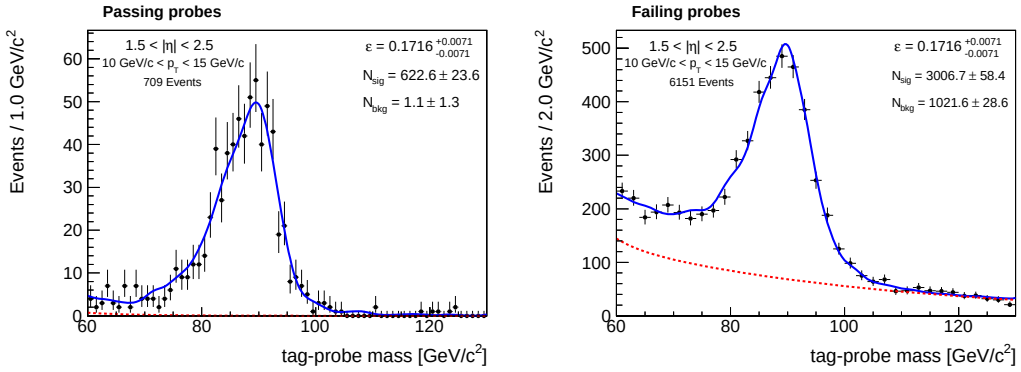
In this section, we describe the measurement of the efficiency for a reconstructed electron to satisfy the preselection requirements and the requirements on the BDT discriminator. The measurement is nominally performed in two bins of p_T and three bins of η. The three bins of pseudorapidity is motivated by the regions of different tracker and support structure material shown in Figure 2-7. The amount of material that an electron must cross before reaching the ECAL significantly affects the amount of bremsstrahlung that it undergoes and can have significant impact on the selection efficiency. Two p_T bins are used due to the fact that the selection efficiency is significantly lower for electrons with p_T less than 20 GeV due to the choice of the tightness of the selection. Electrons with p_T less than 20 GeV are selected with a very tight working point, which is needed to sufficiently suppress the large fake electron background.

The efficiencies are measured using the tag-and-probe technique described in Section 5.1. In order to keep the probe leg unbiased we use only events triggered by the set of specially designed electron tag and probe triggers described in Section 3.1.3, where the unbiased leg is only required to be reconstructed as a supercluster at the trigger level with no additional requirements. The fits that are performed to determine the amount of signal and background in the Z resonance peak region are shown in Figures 5-2, 5-3, and 5-4, for each kinematic bin. The measured efficiency

is then compared to the efficiency determined in the Monte Carlo simulation, and a Monte-Carlo-to-data scale factor is derived for each kinematic bin. These scale factors are then used to correct the efficiencies for the predictions derived from Monte Carlo simulation. As described in Section 5.1 we apply correction scale factors to efficiencies derived from Monte Carlo simulation rather than using the measured efficiencies themselves in order to account for any kinematic biases introduced by the event selection requirements imposed in the efficiency measurement procedure. The measurements are summarized in Table 5.1. Only the uncertainties accounted for by the fitting procedure are included. Systematic uncertainties will be discussed in Section 7.4.2.

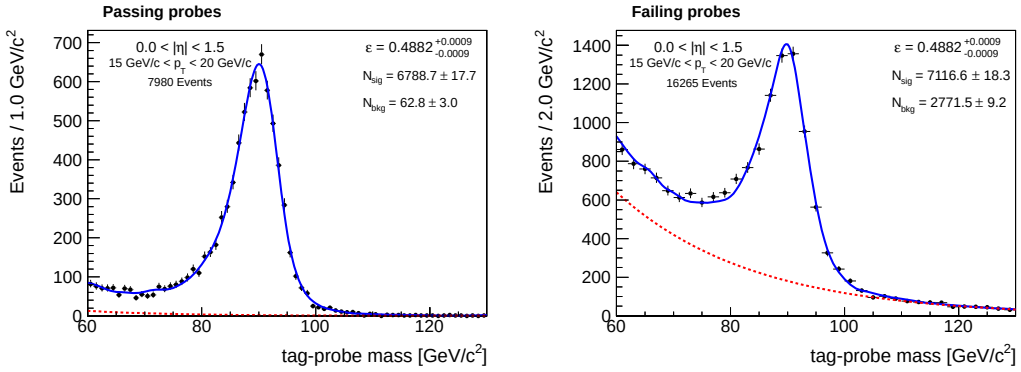


(a) $|\eta| < 1.5$

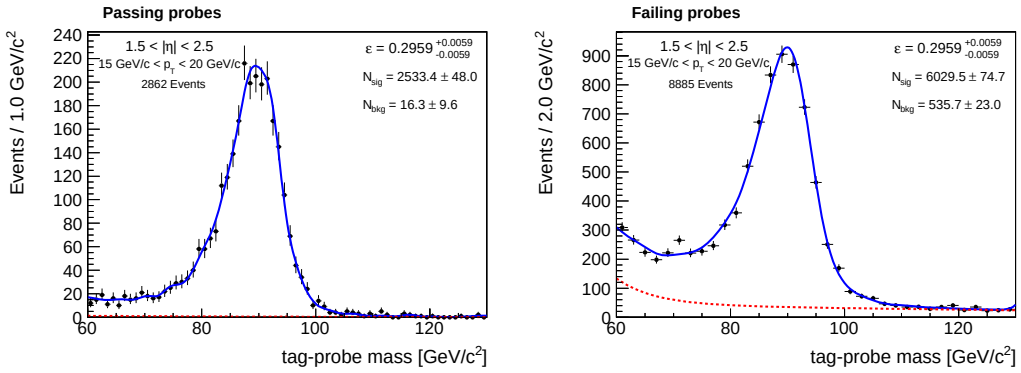


(b) $1.5 \leq |\eta| < 2.5$

Figure 5-2: Unbinned likelihood fits for the $Z \rightarrow e^+e^-$ signal and the background to the dielectron mass distribution for electrons with 10.0 GeV/c < $p_T \leq$ 15.0 GeV/c. The signal and background models are as described in the captions of Figure 5-1.

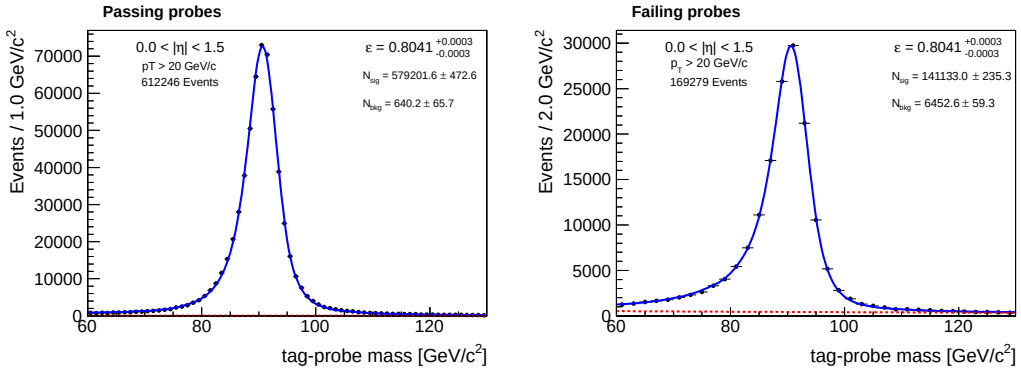


(a) $|\eta| < 1.5$

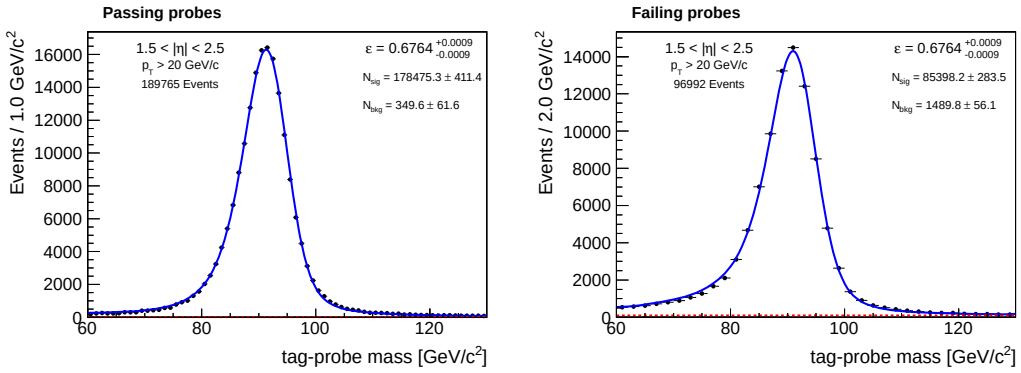


(b) $1.5 \leq |\eta| < 2.5$

Figure 5-3: Unbinned likelihood fits for the $Z \rightarrow e^+e^-$ signal and the background to the dielectron mass distribution for electrons with 15.0 GeV/c < $p_T \leq$ 20.0 GeV/c. The signal and background models are as described in the captions of Figure 5-1.



(a) $|\eta| < 1.5$



(b) $1.5 \leq |\eta| < 2.5$

Figure 5-4: Unbinned likelihood fits for the $Z \rightarrow e^+e^-$ signal and the background to the dielectron mass distribution for electrons with $p_T > 20.0$ GeV/c. The signal and background models are as described in the captions of Figure 5-1.

p_T Bin	η Bin	Data Efficiency	MC to Data Scale Factor
$10.0 < p_T \leq 15.0$	$0.0 \leq \|\eta\| < 1.5$	0.33 +/- 0.02	1.03 +/- 0.05
$10.0 < p_T \leq 15.0$	$1.5 \leq \|\eta\| < 2.5$	0.172 +/- 0.007	1.09 +/- 0.05
$15.0 < p_T \leq 20.0$	$0.0 \leq \|\eta\| < 1.5$	0.4882 +/- 0.0009	0.995 +/- 0.007
$15.0 < p_T \leq 20.0$	$1.5 \leq \|\eta\| < 2.5$	0.296 +/- 0.006	1.04 +/- 0.02
$20.0 < p_T$	$0.0 \leq \|\eta\| < 1.5$	0.8041 +/- 0.0003	0.9849 +/- 0.0006
$20.0 < p_T$	$1.5 \leq \|\eta\| < 2.5$	0.6764 +/- 0.0009	1.021 +/- 0.002

Table 5.1: Electron selection efficiencies and Monte-Carlo-to-data scale factors, in different p_T and η bin. The efficiency generally increases as the p_T increases, and is higher for the barrel than for the endcap. This is primarily by design as our choice of the tightness of the selection cut is required to be tighter towards lower p_T and in the endcap due to larger background rates. Only the uncertainties obtained from the dielectron mass fit are shown. Systematic uncertainties due to intrinsic biases of the method have not been included here.

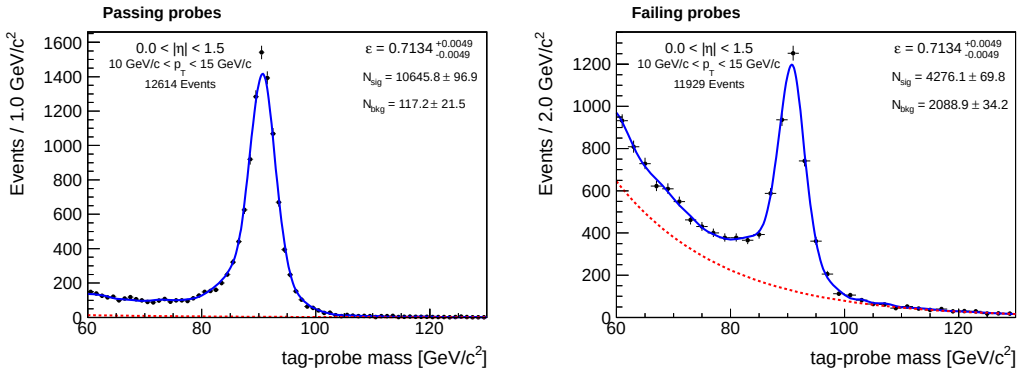
5.4 Muon Selection Efficiency

We measure the efficiency for reconstructed muons to pass the preselection and the cut on the BDT discriminator. The muon selection efficiency is measured separately in three bins of p_T and two bins of η. These bins are primarily motivated by the choice of bins for the BDT training. As for electrons, significantly tighter working points are needed for muons with p_T less than 20 GeV to suppress backgrounds.

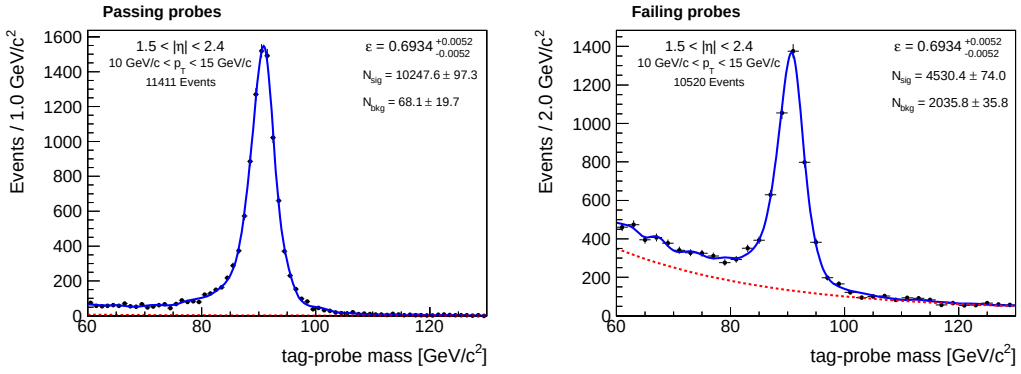
Events collected by the single muon triggers are used in order to select muon probes that remain unbiased by the trigger. Figures 5-5, 5-6, and 5-7 show the dimuon mass fits used to determine the amount of signal and background under the Z resonance and to extract the efficiency. We observe that for some kinematic bins, the signal model has a resolution that is worse than the data. We model the signal using the dimuon mass shape obtained from the simulation convoluted with a Gaussian. This signal model describes well the case where the resolution in data is worse than in the simulation, however does not cover the situation where the resolution in data is better. But, in all such cases, the background level is almost neglgible and the effect of this mismodeling of the signal is covered by the systematic uncertainties.

The efficiency measured in data are compared with the efficiency determined in

the Monte Carlo simulation to derive the Monte-Carlo-to-data scale factor for each kinematic bin. Table 5.2 summarizes the efficiency measurements and the scale factors. Only the uncertainties accounted for by the fitting procedure are included. Systematic uncertainties will be discussed in Section 7.4.2.

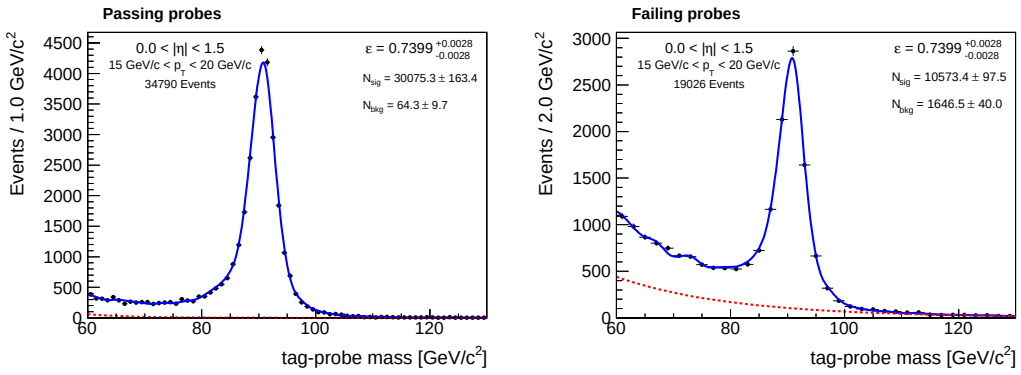


(a) $|\eta| < 1.5$

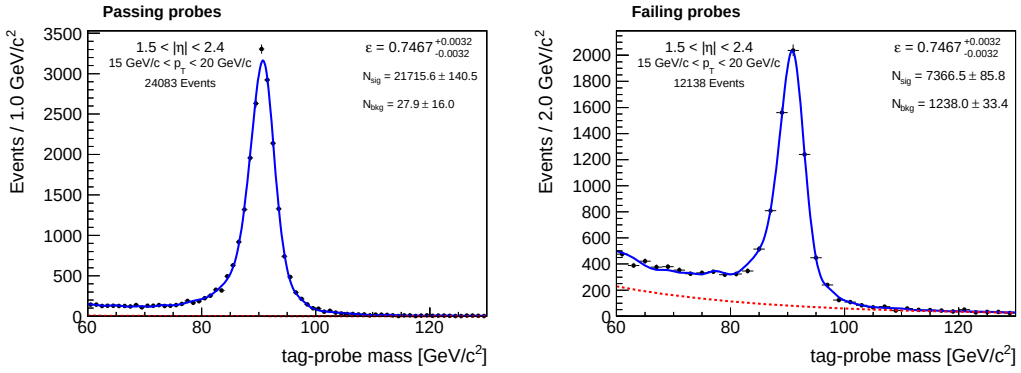


(b) $1.5 \leq |\eta| < 2.4$

Figure 5-5: Unbinned likelihood fits of the dimuon mass distribution to the $Z \rightarrow \mu^+\mu^-$ signal and the background for muons with 10.0 GeV/$c < p_T \leq$ 15.0 GeV/c. The signal and background models are as described in the captions of Figure 5-1. For some kinematic bins, the mass resolution in data is better than the resolution in Monte Carlo, leading to some mismodeling of the signal. Due to the fact that in such cases, the backgrounds are negligible, the signal mismodeling in the peak region has a small effect on the efficiency, and is covered by the systematic uncertainty.

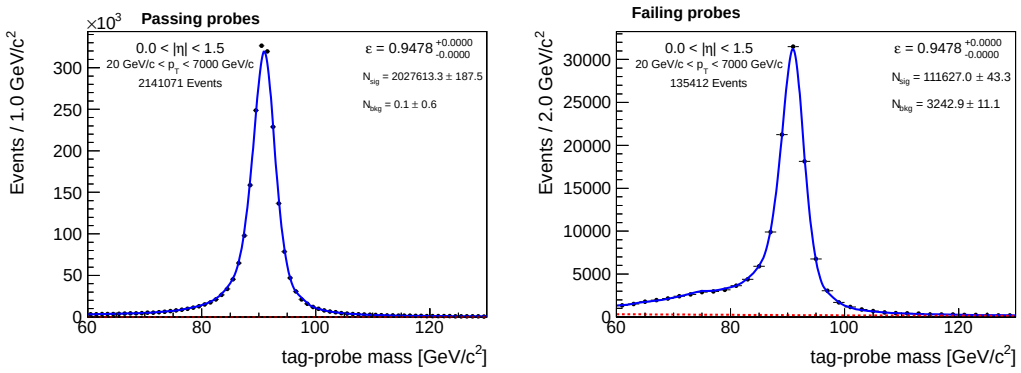


(a) $|\eta| < 1.5$

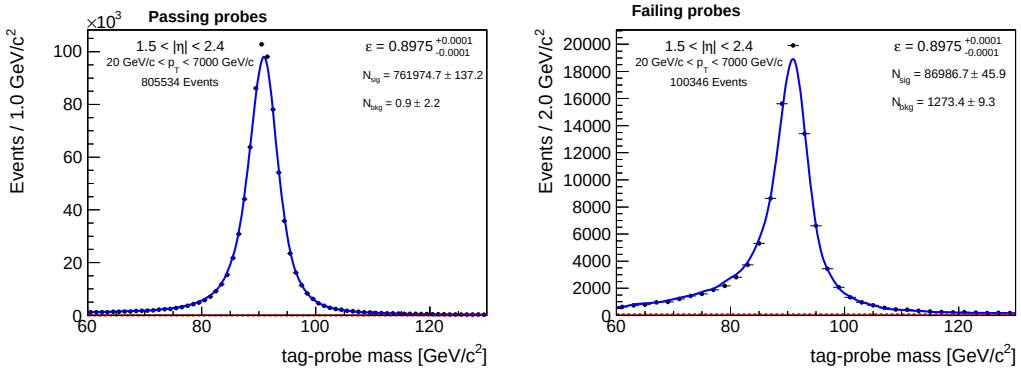


(b) $1.5 \leq |\eta| < 2.4$

Figure 5-6: Unbinned likelihood fits of the dimuon mass distribution to the $Z \rightarrow \mu^+\mu^-$ signal and the background for muons with 15.0 GeV/c < $p_T \leq$ 20.0 GeV/c. The signal and background models are as described in the captions of Figure 5-1. For some kinematic bins, the mass resolution in data is better than the resolution in Monte Carlo, leading to some mismodeling of the signal. Due to the fact that in such cases, the backgrounds are negligible, the signal mismodeling in the peak region has a small effect on the efficiency, and is covered by the systematic uncertainty.



(a) $|\eta| < 1.5$



(b) $1.5 \leq |\eta| < 2.4$

Figure 5-7: Unbinned likelihood fits of the dimuon mass distribution to the $Z \rightarrow \mu^+\mu^-$ signal and the background for muons with $p_T > 20.0$ GeV/c. The signal and background models are as described in the captions of Figure 5-1. For some kinematic bins, the mass resolution in data is better than the resolution in Monte Carlo, leading to some mismodeling of the signal. Due to the fact that in such cases, the backgrounds are negligible, the signal mismodeling in the peak region has a small effect on the efficiency, and is covered by the systematic uncertainty.

p_T Bin	η Bin	Data Efficiency	MC to Data Scale Factor
$10.0 < p_T \leq 15.0$	$0.0 \leq \|\eta\| < 1.5$	0.713 +/- 0.005	0.994 +/- 0.008
$10.0 < p_T \leq 15.0$	$1.5 \leq \|\eta\| < 2.4$	0.693 +/- 0.005	1.022 +/- 0.009
$15.0 < p_T \leq 20.0$	$0.0 \leq \|\eta\| < 1.5$	0.740 +/- 0.003	0.981 +/- 0.005
$15.0 < p_T \leq 20.0$	$1.5 \leq \|\eta\| < 2.4$	0.747 +/- 0.003	1.009 +/- 0.005
$20.0 < p_T$	$0.0 \leq \|\eta\| < 1.5$	0.9478 +/- 0.0001	0.9943 +/- 0.0002
$20.0 < p_T$	$1.5 \leq \|\eta\| < 2.4$	0.8975 +/- 0.0001	0.9814 +/- 0.0003

Table 5.2: Muon selection efficiencies and Monte-Carlo-to-data scale factors, in different p_T and η bins. These scale factors are very close to 1, indicating that the Monte Carlo simulation gives an accurate description of signal muons. Only the uncertainties obtained from the dielectron mass fit are shown. Systematic uncertainties due to intrinsic biases of the method have not been included here.

5.5 Trigger Efficiency Measurement

The per event triggering efficiency given that the event passes the signal selection, is derived from the efficiency per trigger object to pass the particular requirements of each leg of the dilepton triggers and single lepton triggers. The per event trigger efficiency is computed by taking into account the two ways an event can be selected, under the condition that both leptons were reconstructed in the instrumented region of the detector:

- the event passes the dilepton trigger
- the event fails the dilepton trigger because one leg fails the dilepton trigger requirements, but the other leg passes the single lepton trigger.

These are the only possibilities because the requirements on the single lepton triggers are strictly tighter than the requirements on each leg of the dilepton triggers. Thus taking into account the combinatorics, the event efficiency $\varepsilon_{\ell\ell'}(p_T,\, \eta,\, p'_T,\, \eta')$ is given by:

$$\begin{aligned} \varepsilon_{\ell^a\ell^b}(p^a_T,\, \eta^a,\, p^b_T,\, \eta^b) &= 1 - [\bar{\varepsilon}^a_{\mathrm{D-l}}\bar{\varepsilon}^b_{\mathrm{D-l}} + \varepsilon^a_{\mathrm{D-l}}\bar{\varepsilon}^b_{\mathrm{D-t}} + \bar{\varepsilon}^a_{\mathrm{D-t}}\varepsilon^b_{\mathrm{D-l}}] \\ &\quad + \bar{\varepsilon}^a_{\mathrm{D-t}}\varepsilon^b_{\mathrm{S}} + \varepsilon^a_{S}\bar{\varepsilon}^b_{\mathrm{D-t}}, \end{aligned} \tag{5.1}$$

where ε_S is the single lepton trigger efficiency, $\varepsilon_{\mathrm{D-l}}$ is the efficiency of the leading leg of the appropriate double lepton trigger, $\varepsilon_{\mathrm{D-t}}$ is the efficiency of the trailing leg of the appropriate double lepton trigger, and $\bar{\varepsilon} = 1 - \varepsilon$. The index a in ϵ^a indicates that the efficiency is evaluated at p_T^a and η^a, while the index b in ϵ^b indicates that the efficiency evaluated at p_T^b and η^b.

The efficiency computed from Equation 5.1 provides the weight for any Monte Carlo event used to obtain estimates of the signal or background yield. Using the values of the per lepton trigger efficiencies computed in subsequent Sections 5.5.1 and 5.5.2, we find an average per event trigger efficiency between 98% and 99%, depending on the particular process.

5.5.1 Electron Trigger Efficiency

In this section, we measure the efficiency for an electron that has passed the BDT selection to pass each of the following trigger requirements:

- the leading leg of the double electron and electron-muon triggers,
- the trailing leg of the double electron and electron-muon triggers, and
- any one of the single electron triggers.

We measure these efficiencies averaged over the changing conditions of the data collected in 2011, including the effect of changing prescales on the different single electron triggers. As triggers become prescaled for particular data-taking periods, the efficiency of the given trigger will decrease to reflect the size of the prescale. The requirements on the electron legs for the double electron and electron-muon triggers are identical and therefore it is sufficient to measure them using dielectron events. The efficiency is measured simply counting passing and failing probes under the zero background assumption, because the background is suppressed to negligible levels after requiring electrons passing the BDT selection. The effect of the remaining backgrounds have been evaluated to be negligible, and are neglected. The measured

trigger efficiencies for the leading and trailing legs of the double electron triggers are summarized in Table 5.3, for different p_T and η bins. In Table 5.4 we summarize the analogous measurements for the single electron triggers. These efficiencies are used to calculate the per event trigger efficiency using Equation 5.1.

Leading Leg of the Double Electron Trigger		
p_T/η Bin	$0.0 \leq \|\eta\| < 1.5$	$1.5 \leq \|\eta\| < 2.5$
$10.0 < p_T \leq 20.0$	0.534 +/- 0.004	0.331 +/- 0.006
$20.0 < p_T \leq 30.0$	0.9856 +/- 0.0003	0.975 +/- 0.001
$30.0 < p_T$	0.9928 +/- 0.0001	0.9935 +/- 0.0001
Trailing Leg of the Double Electron Trigger		
p_T/η Bin	$0.0 \leq \|\eta\| < 1.5$	$1.5 \leq \|\eta\| < 2.5$
$10.0 < p_T \leq 20.0$	0.988 +/- 0.001	0.993 +/- 0.001
$20.0 < p_T \leq 30.0$	0.9926 +/- 0.0002	0.9938 +/- 0.0004
$30.0 < p_T$	0.9948 +/- 0.0001	0.9953 +/- 0.0001

Table 5.3: The efficiency for a selected electron to pass the leading and trailing legs of the double electron trigger. The leading leg has the additional requirement that it fire the Level-1 seed trigger. The turn on for the efficiency of the leading leg, which is required to have $p_T > 17$ GeV/c at the HLT level, is reflected in the significantly lower measured efficiency for electrons with p_T below 20 GeV/c. Only the statistical uncertainties are shown. Systematic uncertainties due to intrinsic biases of the method have not been included here.

p_T/η Bin	$0.0 \leq \|\eta\| < 1.5$	$1.5 \leq \|\eta\| < 2.5$
$10.0 < p_T \leq 25.0$	0.0000 +/- 0.0001	0.0000 +/- 0.0003
$25.0 < p_T \leq 30.0$	0.0303 +/- 0.0006	0.013 +/- 0.001
$30.0 < p_T \leq 35.0$	0.1361 +/- 0.0009	0.116 +/- 0.001
$35.0 < p_T \leq 40.0$	0.2018 +/- 0.0008	0.214 +/- 0.002
$40.0 < p_T \leq 50.0$	0.2039 +/- 0.0006	0.198 +/- 0.001
$50.0 < p_T \leq 65.0$	0.255 +/- 0.001	0.212 +/- 0.002
$65.0 < p_T \leq 80.0$	0.733 +/- 0.003	0.448 +/- 0.006
$80.0 < p_T$	0.962 +/- 0.002	0.879 +/- 0.006

Table 5.4: The efficiency for a selected electron to pass the single electron trigger. The decrease in the efficiency as the p_T decreases reflects the fact that the single electron triggers with lower p_T thresholds were systematically prescaled as the instantaneous luminosity delivered by the LHC ramped up during 2011. Only the statistical uncertainties are shown. Systematic uncertainties due to intrinsic biases of the method have not been included here.

5.5.2 Muon Trigger Efficiency

Finally, we measure the efficiency for a muon that has passed the BDT selection to pass each of the following trigger requirements:

- the leading leg of the double muon and electron-muon triggers,
- the trailing leg of the double muon and electron-muon triggers, and
- any one of the single muon triggers.

The trigger efficiencies are measured over the integrated data collected in 2011, including the effect of changing prescales on the different single muon triggers. The measured trigger efficiencies for the leading and trailing legs of the double muon triggers are summarized in Tables 5.5 and 5.6, respectively. The single muon trigger efficiencies are summarized in Table 5.7. These efficiencies are used to calculate the per event trigger efficiency using Equation 5.1.

Leading Leg of the Double Muon Trigger		
p_T/η Bin	$0.0 \leq \|\eta\| < 0.8$	$0.8 \leq \|\eta\| < 1.2$
$10.0 < p_T \leq 13.0$	0.059 +/- 0.010	0.068 +/- 0.010
$13.0 < p_T \leq 15.0$	0.804 +/- 0.013	0.801 +/- 0.013
$15.0 < p_T \leq 17.0$	0.820 +/- 0.010	0.798 +/- 0.011
$17.0 < p_T \leq 20.0$	0.964 +/- 0.003	0.948 +/- 0.004
$20.0 < p_T \leq 30.0$	0.964 +/- 0.001	0.952 +/- 0.001
$30.0 < p_T$	0.9662 +/- 0.0002	0.9518 +/- 0.0004
p_T/η Bin	$1.2 \leq \|\eta\| < 2.1$	$2.1 \leq \|\eta\| < 2.4$
$10.0 < p_T \leq 13.0$	0.096 +/- 0.006	0.107 +/- 0.012
$13.0 < p_T \leq 15.0$	0.787 +/- 0.008	0.706 +/- 0.016
$15.0 < p_T \leq 17.0$	0.819 +/- 0.007	0.748 +/- 0.014
$17.0 < p_T \leq 20.0$	0.934 +/- 0.003	0.861 +/- 0.007
$20.0 < p_T \leq 30.0$	0.945 +/- 0.001	0.870 +/- 0.002
$30.0 < p_T$	0.9461 +/- 0.0004	0.874 +/- 0.001

Table 5.5: The efficiency for a selected muon to pass the leading leg of the double muon trigger. The p_T requirement at the HLT level increased from 13 GeV to 17 GeV, as the instantaneous luminosity delivered by the LHC increased. The effect of this evolution is reflected in the increase of the efficiency from 13 GeV to 20 GeV. The p_T bins are motivated by the turn-on region of the trigger, and the η bins are motivated by the detector geometry. Only the statistical uncertainties are shown. Systematic uncertainties due to intrinsic biases of the method have not been included here.

p_T/η Bin	$0.0 \leq \lvert\eta\rvert < 0.8$	$0.8 \leq \lvert\eta\rvert < 1.2$
$10.0 < p_T \leq 13.0$	0.967 +/- 0.008	0.942 +/- 0.009
$13.0 < p_T \leq 15.0$	0.957 +/- 0.007	0.954 +/- 0.008
$15.0 < p_T \leq 17.0$	0.969 +/- 0.005	0.940 +/- 0.007
$17.0 < p_T \leq 20.0$	0.968 +/- 0.003	0.954 +/- 0.004
$20.0 < p_T \leq 30.0$	0.965 +/- 0.001	0.954 +/- 0.001
$30.0 < p_T$	0.9667 +/- 0.0002	0.9536 +/- 0.0004
p_T/η Bin	$1.2 \leq \lvert\eta\rvert < 2.1$	$2.1 \leq \lvert\eta\rvert < 2.4$
$10.0 < p_T \leq 13.0$	0.952 +/- 0.005	0.896 +/- 0.012
$13.0 < p_T \leq 15.0$	0.941 +/- 0.005	0.884 +/- 0.012
$15.0 < p_T \leq 17.0$	0.957 +/- 0.004	0.892 +/- 0.010
$17.0 < p_T \leq 20.0$	0.951 +/- 0.003	0.907 +/- 0.006
$20.0 < p_T \leq 30.0$	0.953 +/- 0.001	0.898 +/- 0.002
$30.0 < p_T$	0.9513 +/- 0.0003	0.896 +/- 0.001

Table 5.6: The efficiency for a selected muon to pass the trailing leg of the double muon trigger. The efficiency is fairly constant at about 95% for $|\eta|$ below 2.1, and decreases to about 90% for $|\eta|$ above 2.1. Only the statistical uncertainties are shown. Systematic uncertainties due to intrinsic biases of the method have not been included here.

p_T/η Bin	$0.0 \leq \lvert\eta\rvert < 0.8$	$0.8 \leq \lvert\eta\rvert < 1.5$
$10.0 < p_T \leq 15.0$	0.0008 +/- 0.0008	0.001 +/- 0.001
$15.0 < p_T \leq 24.0$	0.241 +/- 0.002	0.234 +/- 0.002
$24.0 < p_T \leq 30.0$	0.398 +/- 0.002	0.385 +/- 0.002
$30.0 < p_T \leq 40.0$	0.893 +/- 0.001	0.827 +/- 0.001
$40.0 < p_T$	0.9562 +/- 0.0003	0.8932 +/- 0.0004
p_T/η Bin	$1.5 \leq \lvert\eta\rvert < 2.1$	$2.1 \leq \lvert\eta\rvert < 2.4$
$10.0 < p_T \leq 15.0$	0.001 +/- 0.001	0.001 +/- 0.001
$15.0 < p_T \leq 24.0$	0.231 +/- 0.002	0.188 +/- 0.003
$24.0 < p_T \leq 30.0$	0.394 +/- 0.002	0.275 +/- 0.003
$30.0 < p_T \leq 40.0$	0.817 +/- 0.001	0.281 +/- 0.002
$40.0 < p_T$	0.866 +/- 0.001	0.298 +/- 0.002

Table 5.7: The efficiency for a selected muon to pass the single muon trigger. The decrease in the efficiency as the p_T decreases reflects the fact that the single muon triggers with lower p_T thresholds were systematically prescaled as the instantaneous luminosity delivered by the LHC ramped up during 2011. Only the statistical uncertainties are shown. Systematic uncertainties due to intrinsic biases of the method have not been included here.

Chapter 6

Backgrounds

Despite a achieving a suppression of other particle production processes by several orders of magnitude with the selection criteria described in the previous chapter, there remain residual contributions from non-signal processes in the signal region. These processes are collectively referred to as background, and their contribution in the signal region must be estimated in order to infer the presence of a signal.

Due to the presence of two neutrinos in the final state for which the momenta cannot be individually measured, the Higgs boson cannot be reconstructed as a mass resonance in the decay mode to two leptonically decaying W bosons. Therefore, the only method to infer its presence is through an overall event count excess in the designated signal phase space region. Thus, providing a precise and accurate estimate of the backgrounds is crucial in extracting the Higgs boson signal. After the W^+W^--preselection, the main backgrounds are standard model diboson production including W^+W^- WZ, and ZZ, top production including $t\bar{t}$ and single top, Drell-Yan production of Z/γ^*, W+γ production, and W+jets production. We perform both a cut based counting experiment and a BDT shape analysis, and therefore an estimation of the rate as well as the shape of the BDT discriminator must be estimated for each background. Due to the fact that the phase space region from which we extract the Higgs signal lies in the tails of the distributions of various kinematic observables for the main backgrounds, one does not always obtain an accurate description from the

Monte Carlo simulation alone. Therefore, for most of the backgrounds, we derive estimates using auxiliary measurements in appropriately chosen control regions dominated by particular background processes. From these measurements of the behavior of background processes, we then extrapolate to its behavior in the signal region by making certain well motiated assumptions. These types of procedures for background estimates are generically referred to as data-driven background estimation techniques.

In the following sections we will describe, in greater detail, the methods used to estimate each of the backgrounds, beginning with the Drell-Yan background, followed by the W+jets background. Next, we describe the estimation method for the top background, and the W^+W^- background. Finally, we end with short descriptions of the estimates for the WZ ZZ and Drell-Yan $Z/\gamma^* \to \tau^+\tau^-$ backgrounds. In each section, we describe the methods that are used, the estimate of the background rate in the signal region, and the estimate of the BDT discriminator shape. Systematic uncertainties on these estimates are also discussed.

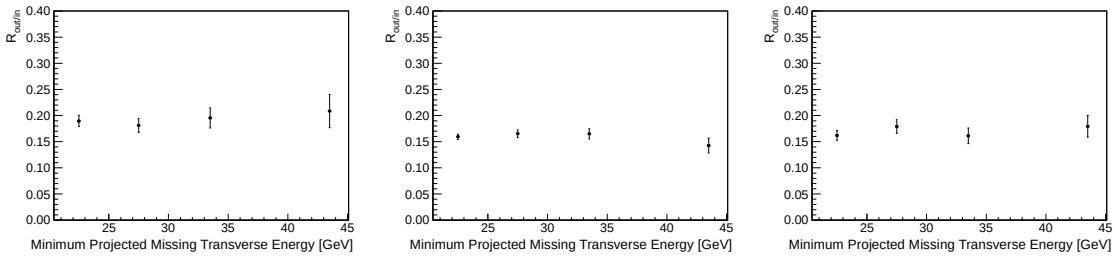
6.1 Drell-Yan $Z/\gamma^* \to e^+e^-/\mu^+\mu^-$ Background

Events from the Z/γ^* Drell-Yan process contribute as background if a large missing transverse energy is observed due to mismeasurement of the lepton momenta or mismeasurement of the hadronic recoil. We reject events with dilepton mass in the 76 GeV/c^2 to 106 GeV/c^2 range to suppress the bulk of the Z resonance. The main remaining contribution is due to the offshell γ^* process. In this section we consider only the decays of the Z/γ^* to an e^+e^- pair or a $\mu^+\mu^-$ pair. The decays to a $\tau^+\tau^-$ pair are discussion in Section 6.6.

Through the use of the projected missing transverse energy as well as imposing requirements on the quality of the momentum measurement for muons, the impact of lepton momentum mismeasurement on the missing transverse energy is reduced to negligible levels. Therefore, non-zero missing transverse energy measured in Z/γ^* events is predominantly a result of mismeasurement of the hadronic recoil. As a re-

sult, the dilepton mass line shape is not significantly altered by the presence of large missing transverse energy. Thus, it is possible to derive a data-driven estimate of this background by extrapolating the dilepton mass line shape from the region of low missing transverse energy to the region of high missing transverse energy. The extrapolation is made by using the Monte Carlo simulation to determine the ratio of the number of events in the mass range defined by the signal selection and the number of events within the narrow mass range around the Z pole mass (84 GeV/c^2 99 GeV/c^2). This is done after selecting events passing all analysis cuts except the projected missing transverse energy cut and the Z mass veto. We denote this ratio as $R_{\mathrm{out/in}}$. This ratio is only sensitive to the kinematic features of the two leptons, for example the p_T of the dilepton system or on the azimuthal angle between the two charged leptons, and is insensitive the features of the hadronic recoil. Due to the fact that the kinematic selection cuts are varying with hypothesis of the Higgs boson mass, we obtain different values of $R_{\mathrm{in/out}}$ depending on the Higgs boson mass hypothesis. Figure 6-1 shows the dependence of the $R_{\mathrm{out/in}}$ ratio as a function of the projected missing transverse energy in the event, as measured in the Monte Carlo simulation, for events passing the W^+W^--preselection. Due to a lack of Monte Carlo events at very high missing transverse energy, for the purpose of the dilepton mass lineshape extrapolation we use the value of $R_{\mathrm{out/in}}$ from the events in the second to last bin, with minimum projected missing transverse energy between 30 GeV/c and 37 GeV/c. This choice of $R_{\mathrm{out/in}}$ gives greater robustness against statistical fluctuations due to the limited size of the Monte Carlo simulation sample. The systematic uncertainty is estimated by taking the largest difference in the $R_{\mathrm{in/out}}$ values among all the bins of minimum projected missing transverse energy above 20 GeV/c, reflecting its possible dependence on the projected missing transverse energy. The values of the $R_{\mathrm{out/in}}$ ratio for the zero-jet and one-jet bins are summarized in Figure 6-2 for the W^+W^--preselection and the signal selection corresponding to each Higgs mass hypothesis. The $R_{\mathrm{out/in}}$ ratio for the events with two or more counted jets is summarized in Table 6.1.

Next we count the number of data events inside the narrow Z mass window, de-



(a) 0-jet bin (b) 1-jet bin (c) 2-jet bin

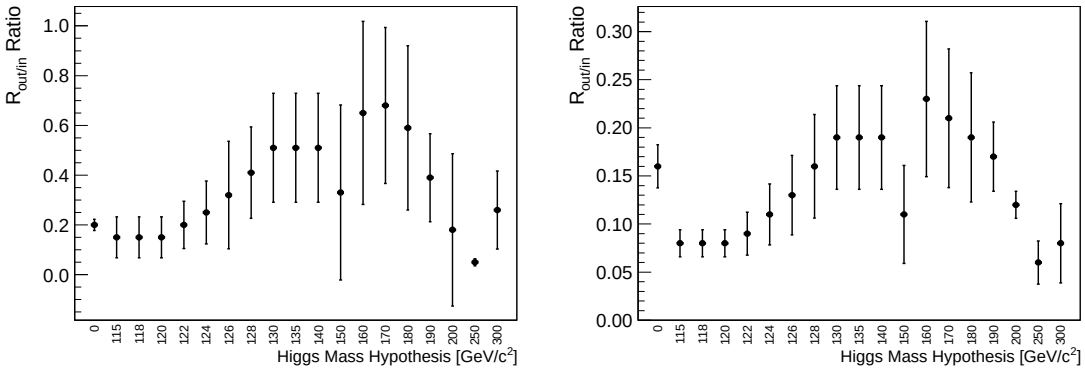
Figure 6-1: The $R_{\mathrm{out/in}}$ ratios determined from the Drell-Yan Monte Carlo simulation are shown in bins of the projected missing transverse energy, for events passing the W^+W^--preselection. The ratios are shown separately for events with zero counted jets, one counted jet, and two or more counted jets. We observe that the ratios are relatively constant as a function of the projected missing transverse energy requirement.

Selection	$R_{out/in}$
W^+W^-Preselection	0.17 ± 0.01

Table 6.1: Summary of the $R_{out/in}$ ratios used for the Drell-Yan background estimate for events with two counted jets. The uncertainty includes both statistical and systematic components.

noted the Z control region, for events passing all other selection cuts. The remaining flavor symmetric backgrounds, dominated by $t\bar{t}$ and W^+W^- production, are subtracted by counting the event yield in the final state with one electron and one muon after the same event selection, correcting for the difference in selection efficiency between electrons and muons. The background from WZ and ZZ events where the two charged leptons are the result of the decay of the same Z boson are not accounted for by the above estimate and are subtracted using Monte Carlo simulation. The background subtracted yield in the Z control region is multiplied by the appropriate $R_{\mathrm{out/in}}$ ratio to obtain the estimate of the Z/γ^* background yield in the signal regiond. This background estimate, $N(\ell\ell)^{\mathrm{DY}}$, is summarized by the equation below:

$$N(\ell\ell)^{\mathrm{DY}} = [N(\ell\ell)^{\mathrm{data}}_{\mathrm{control}} - 0.5 \times N(e\mu)^{\mathrm{data}}_{\mathrm{control}} \times k_{\ell\ell} - N^{WZ/ZZ}_{\mathrm{control}}] \times R(\ell\ell)^{DY}_{out/in}, \quad (6.1)$$



(a) 0-Jet Bin

(b) 1-Jet Bin

Figure 6-2: The $R_{\mathrm{out/in}}$ ratios are plotted for the event selections corresponding to each Higgs boson mass hypothesis. The mass hypothesis labelled "0" refers to the W^+W^--preselection. The $R_{\mathrm{out/in}}$ ratios are shown separately for the zero-jet and one-jet bins. The uncertainties include both statistical and systematic components. The change of the $R_{\mathrm{out/in}}$ ratio as a function of the Higgs boson mass primarily reflects the changes in the kinematic selection cuts imposed in the analysis versus the hypothesized mass.

where $N(\ell\ell)^{\mathrm{DY}}$ is the background estimate in the signal region, $N(\ell\ell)^{\mathrm{data}}_{\mathrm{control}}$ is the data yield in the Z window control region, $N^{WZ/ZZ}_{\mathrm{control}}$ is the expected yield from W^+W^- and WZ in the Z window control region estimated from Monte Carlo simulation, and $k_{\ell\ell}$ is the correction factor to normalize the electron and muon selection efficiencies. We measure $k_{\ell\ell}$ by comparing the event yields in the dielectron and dimuon Z control regions. The final predicted Drell-Yan background and the corresponding Monte-Carlo-to-data scale factor for the W^+W^--preselection and the signal selection corresponding to each Higgs mass hypothesis are summarized in Tables 6.2, 6.3, and 6.4, for events with zero, one, and two counted jets respectively. The systematic uncertainty of the Drell-Yan background estimate is dominated by the statistical uncertainty of the event yield in the Z window control region.

For the BDT shape analysis, we obtain the normalization of the Drell-Yan background using the same method described above but applying the BDT preselection (Section 4.3). To avoid large statistical fluctuations due to limited number of Monte Carlo events, we use the events with minimum projected transverse energy between

Selection / m_{Higgs} hypothesis	$N_{\text{Z window}}$(data)	$N_{\text{signal region}}$(data)	Scale Factor (Data/MC)
W^+W^--Preselection	97.0 ± 21.9	$19.0 \pm 4.7 \pm 1.4$	2.6 ± 0.7
115	42.6 ± 7.3	$6.5 \pm 1.5 \pm 3.2$	4.0 ± 2.2
118	49.7 ± 8.3	$7.5 \pm 1.8 \pm 3.8$	4.6 ± 2.6
120	53.8 ± 9.1	$8.2 \pm 1.9 \pm 4.1$	3.8 ± 2.1
122	46.2 ± 8.1	$9.2 \pm 2.2 \pm 4.1$	3.4 ± 1.7
124	42.1 ± 7.7	$10.7 \pm 2.6 \pm 5.0$	3.5 ± 1.8
126	32.3 ± 7.4	$10.3 \pm 2.9 \pm 6.7$	2.8 ± 2.0
128	23.1 ± 6.6	$9.6 \pm 3.2 \pm 4.0$	2.6 ± 1.4
130	20.9 ± 6.2	$10.6 \pm 3.6 \pm 4.2$	2.9 ± 1.5
135	22.4 ± 6.4	$11.4 \pm 3.8 \pm 4.6$	2.9 ± 1.5
140	22.0 ± 6.6	$11.2 \pm 3.8 \pm 4.5$	2.6 ± 1.4
150	23.7 ± 6.6	$7.7 \pm 3.0 \pm 8.0$	3.3 ± 3.7
160	4.0 ± 3.7	$2.6 \pm 2.6 \pm 1.0$	1.1 ± 1.2
170	3.7 ± 3.4	$2.5 \pm 2.5 \pm 0.3$	1.4 ± 1.4
180	2.6 ± 4.3	$1.6 \pm 2.6 \pm 0.5$	0.9 ± 1.5
190	12.9 ± 6.6	$5.0 \pm 3.1 \pm 1.5$	2.7 ± 1.9
200	11.9 ± 8.0	$2.1 \pm 1.6 \pm 3.6$	1.1 ± 2.1
250	6.6 ± 11.3	$0.3 \pm 0.5 \pm 0.1$	0.5 ± 0.9
300	2.3 ± 8.8	$0.6 \pm 2.3 \pm 0.3$	2.9 ± 11.0

Table 6.2: Summary of the background subtracted Drell-Yan yield inside the Z resonance window, the predicted Drell-Yan yield in the signal region, and Monte-Carlo-to-data scale factors for events with no counted jet. The uncertainty for the Drell-Yan yield inside the Z resonance window and the Monte-Carlo-to-data scale factors include statistical and systematic components, while they are separately reported for the predicted Drell-Yan yield in the signal region.

Selection / m_{Higgs} hypothesis	$N_{\text{Z window}}$(data)	$N_{\text{signal region}}$(data)	Scale Factor (Data/MC)
W^+W^--Preselection	254.2 ± 19.9	$41.9 \pm 4.1 \pm 5.7$	2.7 ± 0.4
115	54.0 ± 7.7	$4.3 \pm 0.7 \pm 0.7$	4.1 ± 1.0
118	70.8 ± 8.8	$5.6 \pm 0.9 \pm 0.9$	4.5 ± 1.0
120	88.3 ± 9.9	$7.0 \pm 1.0 \pm 1.2$	4.4 ± 1.0
122	74.9 ± 9.3	$6.8 \pm 1.1 \pm 1.3$	2.8 ± 0.7
124	62.5 ± 8.6	$6.9 \pm 1.2 \pm 1.6$	2.9 ± 0.8
126	54.3 ± 8.0	$7.3 \pm 1.3 \pm 1.9$	3.0 ± 1.0
128	43.8 ± 7.3	$6.8 \pm 1.3 \pm 2.0$	2.8 ± 1.0
130	38.1 ± 6.9	$7.2 \pm 1.5 \pm 2.1$	2.7 ± 1.0
135	45.7 ± 7.4	$8.6 \pm 1.6 \pm 2.5$	3.3 ± 1.1
140	48.4 ± 7.6	$9.1 \pm 1.7 \pm 2.6$	3.4 ± 1.2
150	61.7 ± 8.8	$6.5 \pm 1.3 \pm 2.9$	3.1 ± 1.6
160	21.5 ± 5.5	$5.0 \pm 1.5 \pm 1.4$	3.9 ± 1.6
170	26.1 ± 5.8	$5.5 \pm 1.5 \pm 1.5$	5.9 ± 2.4
180	29.7 ± 6.5	$5.8 \pm 1.5 \pm 1.7$	6.3 ± 2.5
190	61.6 ± 9.0	$10.5 \pm 2.0 \pm 1.6$	4.5 ± 1.1
200	85.0 ± 10.9	$10.5 \pm 1.8 \pm 1.0$	4.5 ± 0.9
250	105.0 ± 12.7	$6.0 \pm 1.0 \pm 2.6$	1.4 ± 0.7
300	83.8 ± 11.0	$7.0 \pm 1.4 \pm 3.1$	2.8 ± 1.3

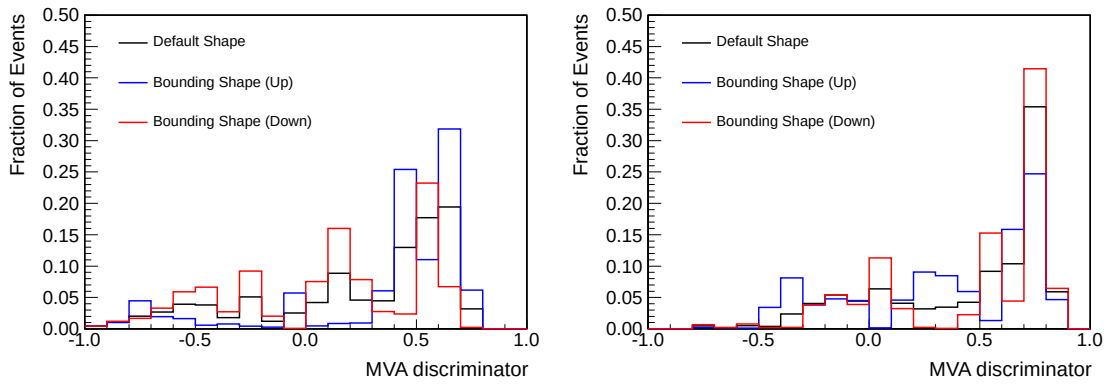
Table 6.3: Summary of the background subtracted Drell-Yan yield inside the Z resonance window, the predicted Drell-Yan yield in the signal region, and Monte-Carlo-to-data scale factors for events with one counted jet. The uncertainty for the Drell-Yan yield inside the Z resonance window and the Monte-Carlo-to-data scale factors include statistical and systematic components, while they are separately reported for the predicted Drell-Yan yield in the signal region.

Selection	$N_{\text{Z window}}$(data)	$N_{\text{signal region}}$(data)	Scale Factor(Data/MC)
W^+W^--Preselection	361.6 ± 20.9	$59.9 \pm 6.1 \pm 3.2$	3.3 ± 0.4

Table 6.4: Summary of the background subtracted Drell-Yan yield inside the Z resonance window, the predicted Drell-Yan yield in the signal region, and Monte-Carlo-to-data scale factors for events with two counted jets. The uncertainty for the Drell-Yan yield inside the Z resonance window and the Monte-Carlo-to-data scale factors include statistical and systematic components, while they are separately reported for the predicted Drell-Yan yield in the signal region.

20 GeV/c and 40 GeV/c to build the prediction of the shape of the BDT discriminator.

We model the systematic uncertainty for the shape by a vertical morphing technique that will be described in further detail in Section 7.2. Essentially, we construct two alternative shapes representing the upper and lower bounds on the true shape. We interpolate between these extreme bounds and the nominal shape using a linear parameterization. This technique will be used for all background shape systematic uncertainties. The alternative shapes for the Drell-Yan background are obtained from the shape given by the Monte Carlo events passing the full selection and the shape obtained by mirroring the difference between this shape and the default shape. The default and alternative shapes are shown in Figure 6-3 for the $m_{\text{Higgs}} = 130$ GeV/c^2 analysis for illustration.



(a) 0-Jet Bin (b) 1-Jet Bin

Figure 6-3: The default and alternative shapes of the BDT discriminator distribution for the Drell-Yan background in the $m_{\text{Higgs}} = 130$ GeV/c^2 analysis are shown. The alternative shapes are used to model and propagate the systematic uncertainty due to the possible dependence of the BDT discriminator shape on the missing transverse energy.

6.2 W+jets and $W + \gamma$ Background

Events from processes where one of the final state leptons is a misidentified lepton are a very important source of background. After imposing the missing transverse

energy requirement, the main sources of such backgrounds are W+jets, where a jet is misidentified as an electron or a muon, and $W + \gamma$ where the photon converts asymmetrically and is misidentified as an electron. Events with two misidentified leptons from multijet and γ+jet production processes comprise only about one percent of such backgrounds.

The selected leptons are required to be isolated. Therefore, only those jets resulting from the tail of the parton fragmentation function where the parton fragments into a leading hadron carrying the majority of the parton momentum may be misidentified as a selected lepton. The dominant physics processes for misidentified electrons are production of leading neutral pions decaying asymmetrically to two photons, one of which converts asymmetrically, production of leading charged pions which interact inelastically early in the electromagnetic calorimeter via the charge exchange process, and production of electrons resulting from the semi-leptonic decay of a leading B or D hadron. The dominant physics processes for misidentified muons are production of leading pions or kaons that decay in flight with a small decay angle, and production of a muon from the semi-leptonic decay of a leading B or D hadron. In order to arrive at an accurate prediction of the rate of such processes one needs to model parton fragmentation and the detector material distribution well. The former has large theoretical uncertainties and the latter is difficult to measure very precisely. As a result, the prediction of the rates of such processes from the Monte Carlo simulation are not particularly reliable. Therefore we use a data driven technique to estimate these backgrounds. Different variants have been used widely in the past [28, 61, 81].

We begin by defining a set of loosely selected lepton-like objects, subsequently referred to as "denominator" leptons. Tables 6.5 and 6.6 summarize the exact cuts that define the denominator electrons and denominator muons, respectively. Next, we measure the efficiency of the denominator lepton to pass the full lepton selection requirements in a data sample dominated by misidentified leptons. This efficiency is denoted as ϵ_{fake} and is typically referred to as the "fake rate". This background dominated sample is defined by events passing any of the single lepton triggers listed

in Table 3.4, and containing one denominator lepton and one reconstructed jet, well separated from the denominator lepton by a ΔR of at least 1.0. In order to reject real leptons from W production, we require that the missing transverse energy is less than 20 GeV/c and that the transverse mass computed from the denominator lepton and the missing transverse energy is less than 20 GeV/c. In order to reject real leptons from Z production, we require that the event does not have a second reconstructed lepton. In order to capture any dependence on the kinematics of the lepton candidate and any dependence on the detector region, we parametrize this efficiency in bins of the p_{T} and η of the denominator lepton.

Cut Observable	Cut Value
$\sigma_{i\eta i\eta}$	$< 0.01(0.03)$ for barrel (endcap)
$\lvert\Delta\phi_{in}\rvert$	$< 0.15(0.10)$ for barrel (endcap)
$\lvert\Delta\eta_{in}\rvert$	$< 0.007(0.009)$ for barrel (endcap)
H/E	$< 0.12(0.10)$ for barrel (endcap)
$\lvert d_0\rvert$	< 0.02 cm
Track Isolation $(\frac{\sum_{\mathrm{trk}} E_{\mathrm{T}}}{p_{\mathrm{T}}^{\mathrm{ele}}})$	< 0.2
ECAL Isolation $(\frac{\sum_{\mathrm{ECAL}} E_{\mathrm{T}}}{p_{\mathrm{T}}^{\mathrm{ele}}})$	< 0.2
HCAL Isolation $(\frac{\sum_{\mathrm{HCAL}} E_{\mathrm{T}}}{p_{\mathrm{T}}^{\mathrm{ele}}})$	< 0.2
full conversion rejection	

Table 6.5: Requirements that define a denominator electron are shown. The definition of the cut variables are given in Section 3.2.2.

Cut Observable	Cut Value
$\lvert d_0\rvert$	< 0.02 cm
Combined Isolation $(\frac{\sum_{\mathrm{trk}} E_{\mathrm{T}}}{p_{\mathrm{T}}^{\mathrm{mu}}} + \frac{\sum_{\mathrm{ECAL}} E_{\mathrm{T}}}{p_{\mathrm{T}}^{\mathrm{mu}}} + \frac{\sum_{\mathrm{HCAL}} E_{\mathrm{T}}}{p_{\mathrm{T}}^{\mathrm{mu}}})$	< 0.4

Table 6.6: Requirements defining a denominator muon. The definition of the cut variables are given in Section 3.2.4.

Next, we select data events passing the full analysis selection with the exception that one of the two lepton candidates is required to pass the denominator selection cuts but to fail the full lepton selection cuts. This lepton will be denoted the "failing" lepton. The other lepton is required to pass the full selection and is denoted the

"tight" lepton. The data sample selected in this way is denoted the "tight+fail" sample. Each of the events passing this selection is given a weight computed from the fake rate in the particular p_T and η bin of the failing lepton, as follows:

$$w = \frac{\epsilon_{\text{fake}}(p_{\text{T}}^{\text{fail}}, \eta^{\text{fail}})}{1 - \epsilon_{\text{fake}}(p_{\text{T}}^{\text{fail}}, \eta^{\text{fail}})} \tag{6.2}$$

where $p_{\text{T}}^{\text{fail}}$ and η^{fail} are the transverse momentum and pseudorapidity of the failing lepton. Summing the weights w over all such events in the tight + fail sample yields the prediction for the jet-induced background.

This tight + fail extrapolation prediction will in fact double count the QCD component of the background where both leptons are jet induced fakes. This is essentially a combinatorial artifact, because in the tight+fail selection, one is unable to uniquely distinguish which lepton is required to be the tight one and which lepton is required to be the failing one, and therefore one customarily selects both combinations. In order to estimate the amount of double counting, we perform the fake rate extrapolation on both lepton legs, selecting events which pass all event selection criteria, except that both leptons are required to pass the denominator selection and to fail the full lepton selection. This event sample is denoted as the "fail+fail" sample. Events in the fail + fail sample are then given weights as follows:

$$w_{i,j} = \frac{\epsilon_{\text{fake}}(p_{\text{Ti}}, \eta_{\text{i}})}{1 - \epsilon_{\text{fake}}(p_{\text{Ti}}, \eta_{\text{i}})} \times \frac{\epsilon_{\text{fake}}(p_{\text{Tj}}, \eta_{\text{j}})}{1 - \epsilon_{\text{fake}}(p_{\text{Tj}}, \eta_{\text{j}})} \tag{6.3}$$

where i and j denote separately the two failing leptons, and $p_{\text{Ti/j}}$ and $\eta_{\text{i/j}}$ are their transverse momenta and pseudorapidity. Summing the weights $w_{i,j}$ over all such events in the fail + fail sample yields the total QCD double fake background. Performing this estimate, we observe that the double fake background is roughly one percent of the total jet induced background. Compared with the total uncertainty of the jet induced background prediction, it is very small and therefore neglected.

Due to an inefficiency of the lepton selection for processes producing two real leptons, the tight+fail sample contains some amount of contamination from such sources. The contamination is dominated by W^+W^- production. We account for this contamination using the Monte Carlo prediction, correcting for the difference in the inefficiency measured in data and simulation. Moreover, events where a W boson is produced in association with an isolated photon ($W+\gamma$) can contaminate this event sample in the case where the photon undergoes conversion in the detector material. These contributions are also accounted for using Monte Carlo simulation.

Photons from $W + \gamma$ events may convert asymmetrically producing a misidentified electron that passes the full electron selection. The rate of such events are not estimated using the procedure described above to account for jet induced background because the fake rate for a jet induced fake electron is smaller than the fake rate for a photon induced fake electron by roughly one order of magnitude, due to the fact that such a photon is typically produced in isolation. Therefore we account for such backgrounds using the Monte Carlo simulation. The prediction is cross checked in the control sample defined by the W^+W^--preselection but requiring that the two leptons have the same charge. This cross check is described in greater detail in Section 6.2.3.

6.2.1 Fake Rates and Predicted Jet Induced Background

Using the procedure outlined above, we measure the electron and muon fake rates and summarize the results in Tables 6.7 and 6.8. The fake rates are plotted as a function of p_T and η in Figures 6-4 and 6-5 for electrons and muons respectively. We have different requirements on the multivariate lepton discriminators in different kinematic, and therefore this dependence on p_T and η primarily reflects the choice of the tightness of the cuts in the different kinematic bins. This choice is, in turn, motivated by the size of the background in each of the kinematic bins and the need to reduce them to a manageable level.

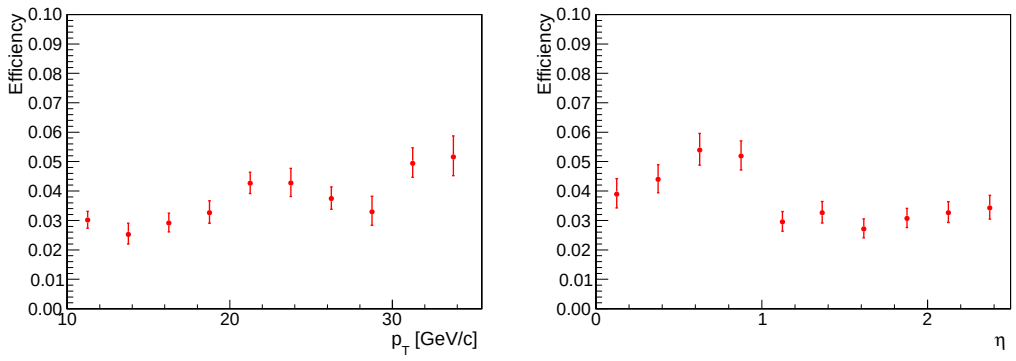
The fake rates are measured from events triggered by prescaled triggers whose prescales were evolving with the LHC beam conditions, and therefore also evolving

p_T and η Bins	$0 < \eta < 1.0$	$1.0 < \eta < 1.479$
$10 < p_T <= 15$	0.043 ± 0.005	0.027 ± 0.004
$15 < p_T <= 20$	0.047 ± 0.005	0.019 ± 0.004
$20 < p_T <= 25$	0.055 ± 0.006	0.040 ± 0.006
$25 < p_T <= 30$	0.040 ± 0.005	0.031 ± 0.006
$30 < p_T <= 35$	0.055 ± 0.007	0.047 ± 0.008
	$1.479 < \eta < 2.0$	$2.0 < \eta < 2.5$
$10 < p_T <= 15$	0.013 ± 0.003	0.022 ± 0.005
$15 < p_T <= 20$	0.021 ± 0.004	0.029 ± 0.005
$20 < p_T <= 25$	0.033 ± 0.005	0.038 ± 0.005
$25 < p_T <= 30$	0.034 ± 0.005	0.037 ± 0.006
$30 < p_T <= 35$	0.054 ± 0.008	0.044 ± 0.007

Table 6.7: The electron fake rates are shown in bins of p_T and η. Only statistical uncertainties are given. The systematic uncertainties are discussed further in Section 6.2.2. A combination of the electron triggers from Table 3.4 are used, with a p_T threshold on the leading jet in the event of 35 GeV/c.

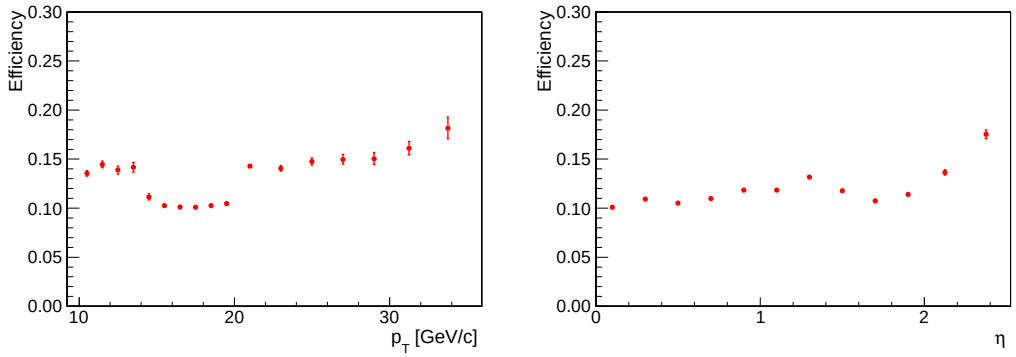
p_T and η Bins	$0 < \eta < 1.0$	$1.0 < \eta < 1.479$
$10 < p_T <= 15$	0.119 ± 0.002	0.168 ± 0.004
$15 < p_T <= 20$	0.095 ± 0.001	0.113 ± 0.001
$20 < p_T <= 25$	0.146 ± 0.002	0.138 ± 0.003
$25 < p_T <= 30$	0.146 ± 0.004	0.146 ± 0.006
$30 < p_T <= 35$	0.153 ± 0.008	0.174 ± 0.012
	$1.479 < \eta < 2.0$	$2.0 < \eta < 2.5$
$10 < p_T <= 15$	0.117 ± 0.003	0.156 ± 0.005
$15 < p_T <= 20$	0.097 ± 0.001	0.136 ± 0.003
$20 < p_T <= 25$	0.129 ± 0.003	0.170 ± 0.007
$25 < p_T <= 30$	0.155 ± 0.007	0.205 ± 0.014
$30 < p_T <= 35$	0.185 ± 0.013	0.213 ± 0.027

Table 6.8: The muon fake rates are shown in bins of p_T and η. Only statistical uncertainties are given. The systematic uncertainties are discussed further in Section 6.2.2. A combination of the muon triggers from Table 3.4 are used, with a p_T threshold on the leading jet in the event of 15 GeV/c.



(a) p_T (b) η

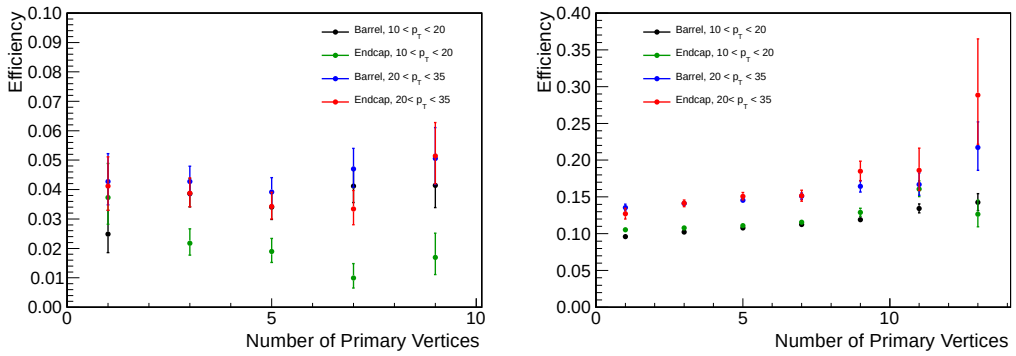
Figure 6-4: Electron fake rates are plotted as a function for p_T and η. The dependence on p_T and η are mainly a reflection of the choice of tightness on the BDT discriminator selection cut in each kinematic bin. The rate of background electrons are typically larger at lower p_T and in the endcap. Therefore, we have chosen tighter selection cuts on the BDT discriminator in those kinematic regions.



(a) p_T (b) η

Figure 6-5: Muon fake rates are plotted as a function for p_T and η. The dependence on p_T are mainly a reflection of the choice of tightness on the BDT discriminator selection cut in each kinematic bin. The selection cut was chosen such that the signal efficiency is maintained at the same level as the cut-based selection used in previous versions of this search [67] in each kinematic bin. This choice induces the discontinuity in the fake rate for the p_T bin between 14.5 GeV/c and 20 GeV/c. The rise of the fake rate for η above 2.1 reflect the fact that discrimination is more difficult in this region of the muon detector.

with the amount of pileup. As a result, the pileup distribution between this data sample and the data sample selected for the signal is slightly different. In Figure 6-6, we illustrate the size of the fake rate dependence on the number of reconstructed vertices. Using these dependencies and reweighting the pileup distribution of the fake rate measurement sample to the unprescaled signal sample, we verified that this bias is negligible compared with the total uncertainty.



(a) Electrons (b) Muons

Figure 6-6: Electron and Muon fake rates as a function of the number of reconstructed primary vertices in four different p_T and η bins. No significant dependence on pileup is observed for the electron fake rate. For muons, the effect of the increase in the fake rate over the range of pileup scenarios faced in 2011 is negligible compared with the systematic uncertainties, which are discussed later in Section 6.2.2.

Using these measured fake rates, we predict the fake lepton backgrounds in the signal region of the analysis corresponding to each Higgs mass hypothesis. The fake electron and fake muon background predictions for the sample defined by the W^+W^--preselection are summarized in Tables 6.9 and 6.10 respectively.

For the BDT shape analysis, each event in the tight+fail sample enters the distribution of the BDT discriminator with a weight given by Equation 6.2 computed using the p_T and η of the failing lepton. This procedure results in a data driven prediction of the shape of the BDT discriminator distribution. This data driven prediction for the $m_{\text{Higgs}} = 130$ GeV/c^2 analysis is shown in Figure 6-7 and compared to the prediction from the W+jets Monte Carlo simulation. We observe reasonably good agreement in

Final State	μe	eμ	ee	Total
0-Jet Bin	11.5 ± 0.8	16.2 ± 1.0	9.6 ± 0.8	37.3 ± 1.5
1-Jet Bin	8.5 ± 0.7	12.2 ± 0.9	4.2 ± 0.5	24.9 ± 1.3
2-Jet Bin	3.0 ± 0.5	6.0 ± 0.6	1.9 ± 0.4	10.9 ± 0.9

Table 6.9: Data driven predictions of the fake electron background after the W^+W^--preselection are shown. Only statistical uncertainties are reported. The order of the charged leptons in the final state label reflects the ordering of the charged leptons in p_T. For the 2-Jet bin, no VBF selection has been imposed.

Final State	$\mu\mu$	μe	eμ	Total
0-Jet Bin	5.3 ± 1.3	4.0 ± 1.0	19.1 ± 2.0	28.4 ± 2.7
1-Jet Bin	3.9 ± 1.2	1.9 ± 0.8	14.1 ± 1.8	19.9 ± 2.3
2-Jet Bin	1.3 ± 0.9	3.3 ± 0.9	6.8 ± 1.4	11.4 ± 1.9

Table 6.10: Data driven predictions of the fake muon background after the W^+W^--preselection are shown. Only statistical uncertainties are reported. The order of the charged leptons in the final state label reflects the ordering of the charged leptons in p_T. For the 2-Jet bin, no VBF selection has been imposed.

the shape between these two predictions.

6.2.2 Systematic Uncertainties

This data driven method for estimating fake lepton backgrounds, relies on the assumption that fake rates for lepton candidates induced by jets measured in an event sample dominated by multijet production is the same as fake rates measured in a sample dominated by W+jets events. The degree to which this assumption fails must be reflected in the systematic uncertainties of the fake lepton background prediction. In order to test the validity of the assumption and to extract some quantitative measure of the systematic uncertainties, we perform a closure test on the W+jets Monte Carlo sample by comparing the background yield predicted by the Monte Carlo simulation with the yield predicted using the fake rate procedure. To be consistent, we use the Monte Carlo sample of QCD multijet processes to measure the fake rates, and then apply them to the tight + fail sample selected in the W+Jet Monte Carlo sample. The degree of disagreement yields a quantitative measure of the systematic uncertainty of

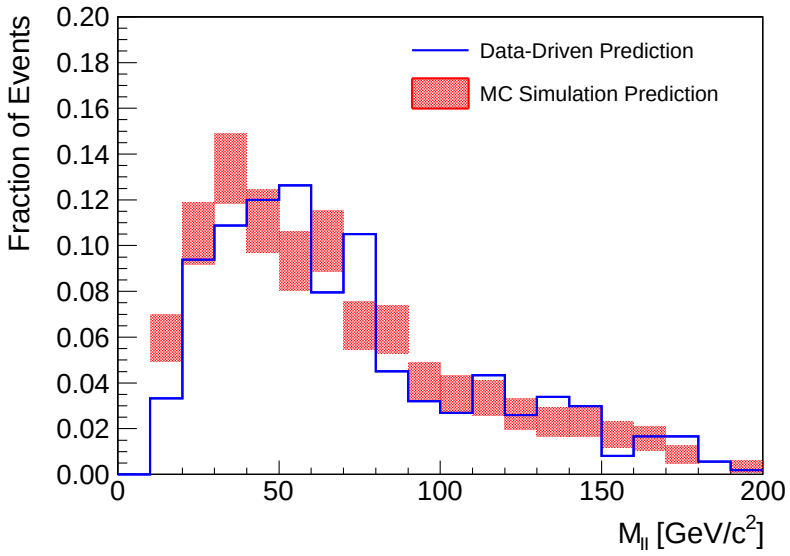


Figure 6-7: The shape of the BDT discriminator observable for the fake lepton background predicted from the data driven method is compared with the corresponding prediction from the Monte Carlo simulation. Only the statistical uncertainties are shown for the Monte Carlo prediction.

the method.

The systematic uncertainties can be factorized into two main sources. The first source is due to the difference in the p_T spectrum of the jets, or more accurately of the partons, in the measurement sample, dominated by QCD multijet events compared to the p_T spectrum in the tight+fail sample, dominated by W+jets events. We measure the fake rate in bins of denominator-object p_T. Thus, for a denominator object with a given p_T, the efficiency of the isolation cut varies greatly depending on whether the jet producing this denominator object has larger or smaller p_T. The second main source of systematic uncertainty is the composition of the origin of the fake lepton. Due to the difference in the quark content in W+jet events compared to QCD events, the fraction of fake leptons resulting from different sources and fake mechanisms may be different. Because these different sources can typically have different fake rates, the fake rate measured in the ensemble can be different as well.

To address the first source of systematic uncertainty in data, we perform the fake

rate measurement using different thresholds on the leading away-jet in the event, in order to capture the degree of uncertainty in the jet p_T spectrum. The away-jets are any jets that are separated from the charged lepton candidate by ΔR$>$ 1.0. To provide some reasonable guidance on the upper and lower bounds of the jet threshold to be used, we compare the p_T spectrum of the jet that a given denominator-object lies inside, referred to as the "lepton jet", for the W+jets Monte Carlo sample and for each of the fake rate measurement samples in data defined by different thresholds on the p_T of the leading jet. This spectrum is representative of the p_T spectrum of a particular parton which produces the fake lepton after fragmentation and hadronization. This comparison is shown in Figure 6-8 and 6-9 for electrons and muons, respectively. For electrons we observe that a threshold on the leading jet of 35 GeV/c matches the W+jets spectrum the best, while for muons, a threshold of 15 GeV/c matches the best. These are chosen for the nominal fake rate measurement. The systematic uncertainty is computed as the difference between the fake rates measured with thresholds of 20 GeV/c and 50 GeV/c for electrons, and with thresholds of 0 GeV/c and 30 GeV/c for muons, as these ranges are observed to cover the nominal W+jets spectrum in the Monte Carlo simulation. Tables 6.11 and 6.12 show the changes in the fake electron and fake muon background estimates after the W^+W^--preselection for different thresholds on the leading jet in the event. Taking the largest difference to the nominal central value as the systematic uncertainty gives an estimate of 28% for electrons and 31% for muons.

To address the second main source of systematic uncertainty, due to the uncertainty of the fake-lepton composition, we perform a closure test using the Monte Carlo simulation. Fake rates measured from the QCD multijet Monte Carlo samples, are applied on the W+jets Monte Carlo sample. The same selection that is used to measure the fake rate in data is used, with a p_T threshold of 35 GeV/c and 15 GeV/c on the leading jet in the event for electrons and muons respectively. To ensure that we compare only the relevant components of the fake lepton background, we do not consider fake muon events when evaluating the closure test for fake electrons, and

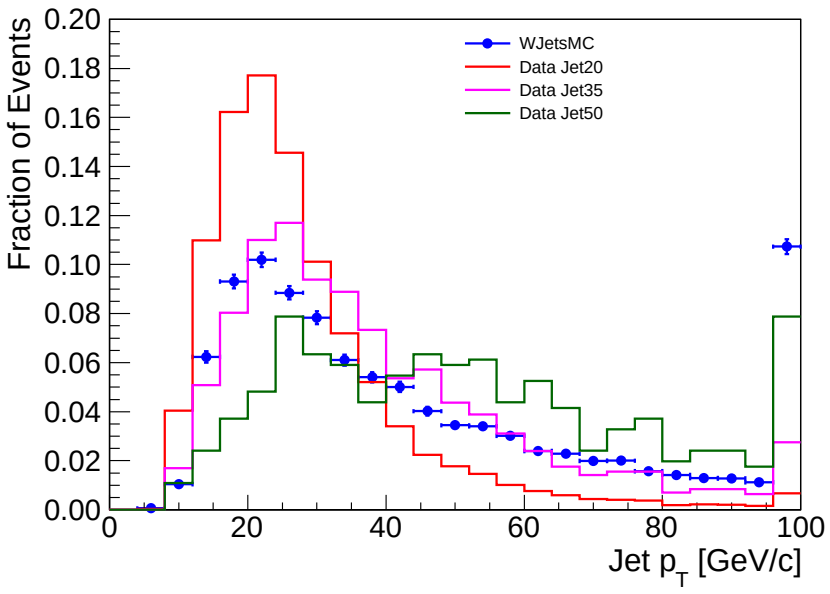


Figure 6-8: The p_T spectrum of the lepton jet is shown for the W+jets Monte Carlo and the data fake rate measurement sample, with different p_T thresholds placed on the leading away-jet in the event. This away-jet is required to be well separated from the charged lepton candidate. We observe that placing a threshold of 35 GeV/c on the leading away-jet, results in a p_T spectrum for the lepton jet that is closest in shape to the W+jets Monte Carlo. Therefore, this threshold is used for the nominal fake rate measurement. Thresholds on the leading away-jet of 20 GeV/c and 50 GeV/c, yield p_T spectra for the lepton jet that bound the W+jets spectrum, and are used as estimates for the systematic uncertainty on the electron fake rate.

Jet Threshold	% Change in fake electron background estimate after W^+W^--preselection
Jet20	+28%
Jet25	+20%
Jet30	+6%
Jet35 (Central Value)	0%
Jet40	−2%
Jet45	−15%
Jet50	−22%

Table 6.11: Relative changes in the estimate of the fake-electron background using different thresholds on the leading jet for computing the fake rate.

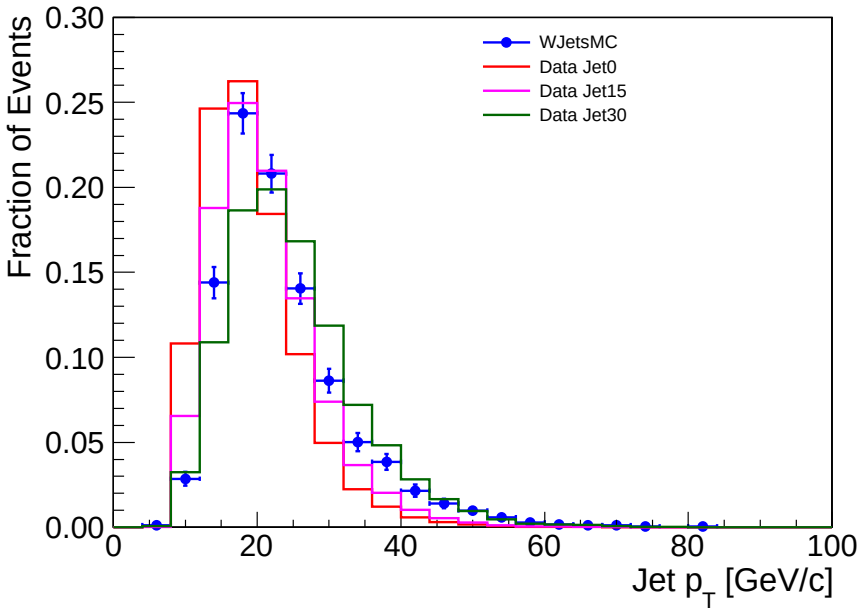


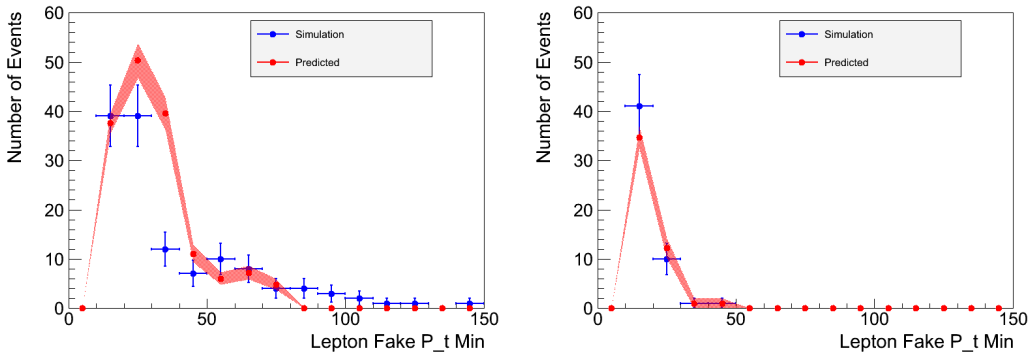
Figure 6-9: The p_T spectrum of the jet that a denominator muon lies inside is shown for the W+jets Monte Carlo and the data fake rate measurement sample, with different p_T thresholds placed on the leading away-jet in the event. A threshold of 15 GeV on the leading away-jet gives the closest match to the lepton jet p_T spectrum produced by the W+jets Monte Carlo, and is used for the nominal fake rate measurement. Placing no threshold and a threshold of 30 GeV on the leading away-jet, bound the W+jets spectrum for the lepton jet and are used to estimate the systematic uncertainties on the muon fake rate.

Jet Threshold	% Change in fake muon background estimate after W^+W^--preselection
Jet0	+19%
Jet5	+17%
Jet10	+13%
Jet15 (Central Value)	0%
Jet20	−8%
Jet25	−20%
Jet30	−31%

Table 6.12: Relative changes estimate of the fake-muon background using different thresholds on the leading jet for computing the fake rate.

vice versa for fake muons. For fake electrons we also remove any W+γ events with a generator level photon produced with $p_T > 10$ GeV/c. The predictions of yields and distributions are compared with the yields and distributions obtained from the simulation-based result after the W^+W^- selection.

We compare the p_T distribution of the fake lepton after the W^+W^--preselection predicted by the fake-rate method with that predicted by the simulation in Figure 6-10. The difference that is observed reflects the systematic uncertainty of the fake rate extrapolation between the different event samples.



(a) Fake Electrons (b) Fake Muons

Figure 6-10: The p_T distributions of the fake electron (a) and fake muon (b) after the W^+W^--preselection, predicted using the fake-rate method are compared with that predicted by the Monte Carlo simulation. The bands represent the statistical uncertainty in the prediction using the fake-rate method. The predicted shape agrees reasonably well with the simulation shape.

Tables 6.13 and 6.14 summarize the comparison between the yields predicted using the fake-rate method and the yields from the Monte Carlo simulation after the W^+W^--preselection for fake electron and fake muon backgrounds respectively. The largest difference among the different final states is used as an estimate of the systematic uncertainty. This results in a systematic uncertainty of 23% for electrons and 18% for muons.

Finally the systematic uncertainties from the two main sources are added in quadrature to give a total systematic uncertainty on the fake lepton background

After W^+W^--preselection			
Final State	Predicted Yield	Simulation Yield	Fractional Difference
ee	40.5 ± 2.9	33 ± 5.7	+23%
e μ	119.9 ± 5.3	105 ± 10.2	+14%
total	160.4 ± 6.1	138 ± 11.7	+16%

Table 6.13: Comparison of fake electron background yields after the W^+W^--preselection between the fake-rate prediction and the simulation prediction

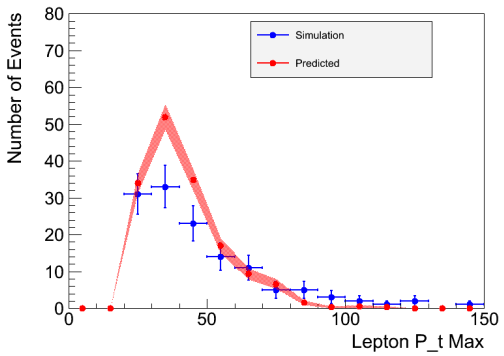
After W^+W^--preselection			
Final State	Predicted Yield	Simulation Yield	Fractional Difference
$\mu\mu$	16.5 ± 1.9	20 ± 4.5	−18%
e μ	33.8 ± 2.8	35 ± 5.9	−3%
total	50.3 ± 3.4	55 ± 7.4	−9%

Table 6.14: Comparison of fake muon background yields after the W^+W^- preselection between the fake-rate prediction and the simulation prediction

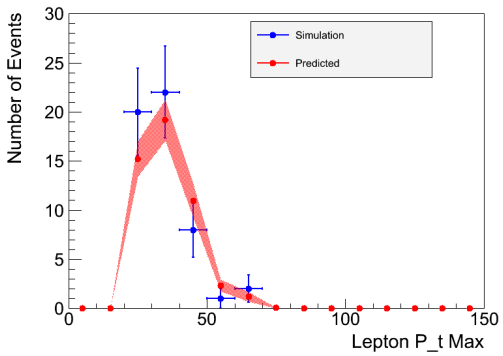
normalization of 36% for electrons and 36% for muons.

For the BDT shape analysis, it is important to estimate the systematic uncertainty in the shape of the BDT discriminator distribution. To gain some intuition and confidence in the ability for the fake rate method to predict the shapes of kinematic distributions, we first compare the distributions of important kinematic observables predicted by the fake rate method in Monte Carlo with the distributions predicted by the simulation. This comparison for the distribution of the p_T of the leading and trailing lepton are shown in Figure 6-11. The distribution for the missing transverse energy and the $\Delta\phi$ between the two leptons is shown in Figure 6-12, and the distribution for the dilepton mass and the transverse mass of the Higgs system is shown in Figure 6-13. Overall, we observe no statistically significant biases beyond 30% in any of the predicted distributions.

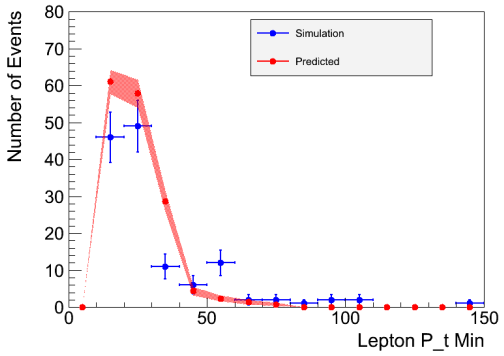
We model the systematic uncertainty of the BDT discriminator shape using the vertical morphing technique, where we linearly interpolate between two shapes that are designated as extreme bounds on the true shape. We consider two potential systematic biases, one due to the uncertainty in the jet spectrum and one due to the



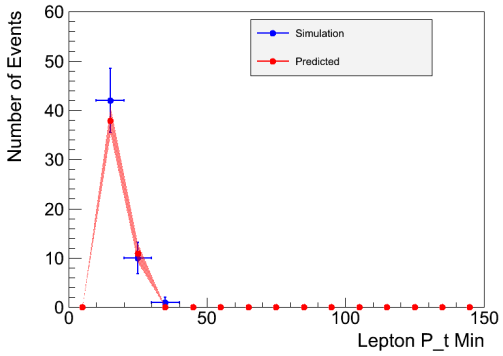
(a) Leading lepton p_T, fake electrons



(b) Leading lepton p_T, fake muons

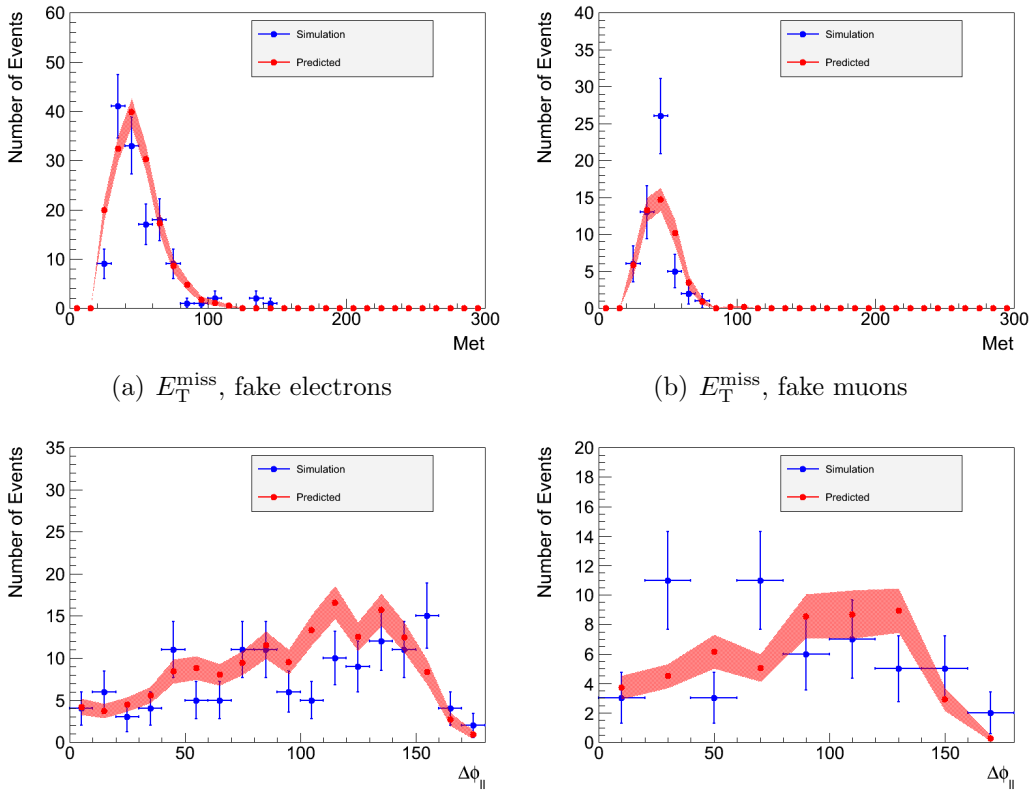


(c) Trailing lepton p_T, fake electrons



(d) Trailing lepton p_T, fake muons

Figure 6-11: A comparison of the p_T distribution of the leading and trailing lepton between the fake rate method prediction and the simulation prediction



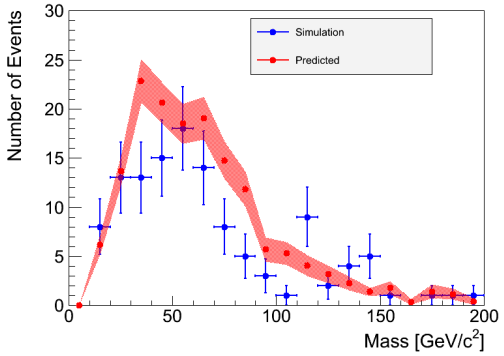
(a) $E_{\mathrm{T}}^{\mathrm{miss}}$, fake electrons

(b) $E_{\mathrm{T}}^{\mathrm{miss}}$, fake muons

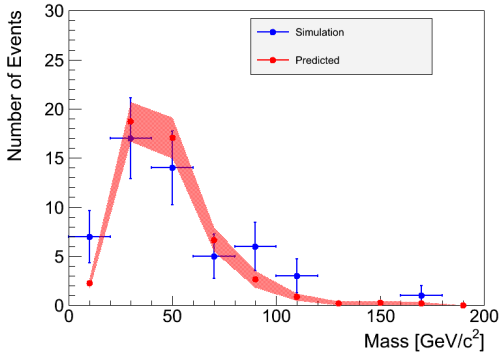
(c) $\Delta\phi$, fake electrons

(d) $\Delta\phi$, fake muons

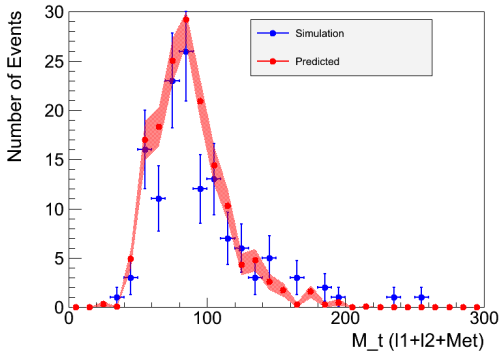
Figure 6-12: A comparison of the distribution of the missing transverse energy, and the $\Delta\phi$ between the two leptons predicted using the fake rate method and the full simulation



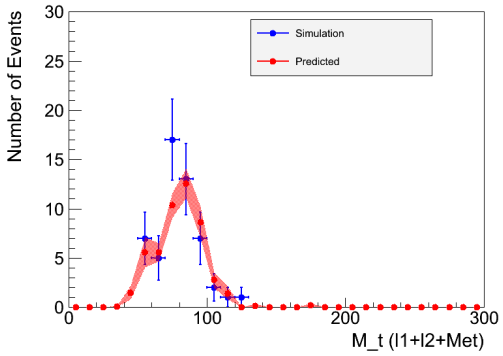
(a) Dilepton Mass, fake electrons



(b) Dilepton Mass, fake muons



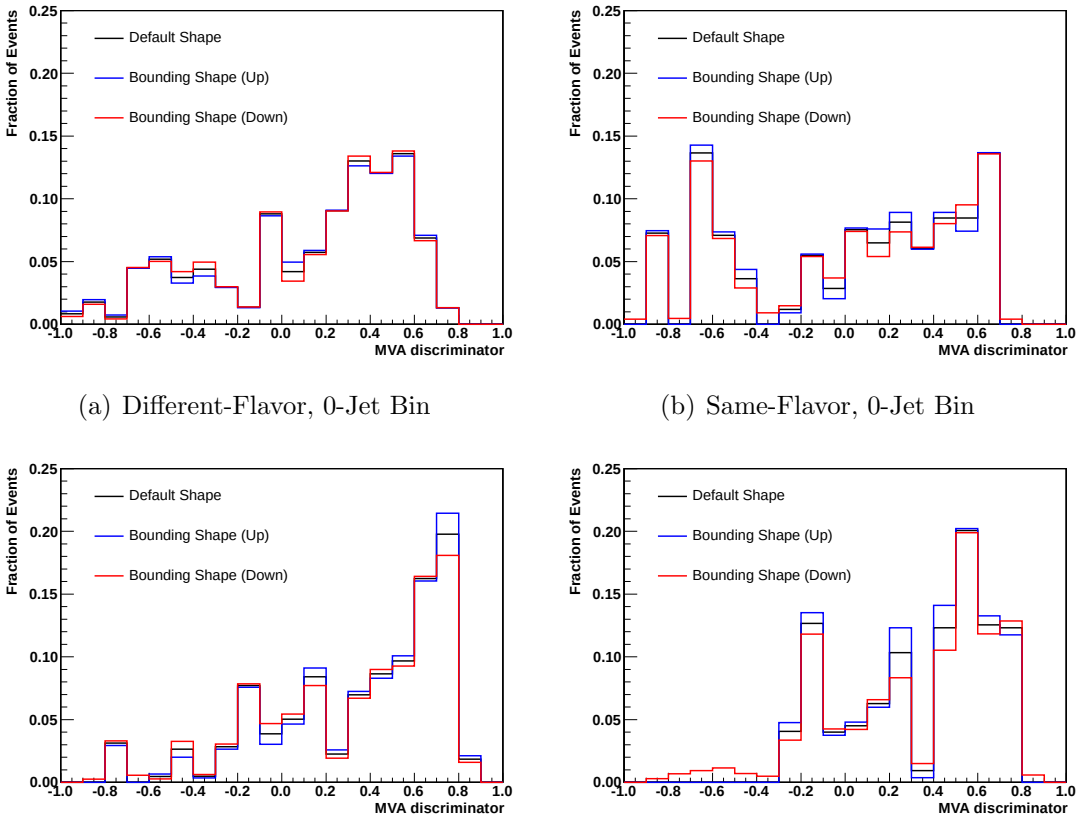
(c) Higgs M_T, fake electrons



(d) Higgs M_T, fake muons

Figure 6-13: A comparison of the distribution of the dilepton mass and Higgs transverse mass predicted using the fake rate method and the full simulation

sample dependence. For the uncertainty in the jet spectrum we define the two extreme shapes to be the BDT discriminator shapes predicted using fake rates measured with a leading jet p_T threshold of 20 GeV/c and 50 GeV/c for fake electrons. For fake muons, the extreme shapes are obtained by using fake rates measured with no cut on the leading jet p_T and a cut of 30 GeV/c. The extreme shapes corresponding to the jet spectrum uncertainty systematic for the $m_{\text{Higgs}} = 130$ GeV/c^2 analysis are shown in Figure 6-14 for illustration.



(a) Different-Flavor, 0-Jet Bin

(b) Same-Flavor, 0-Jet Bin

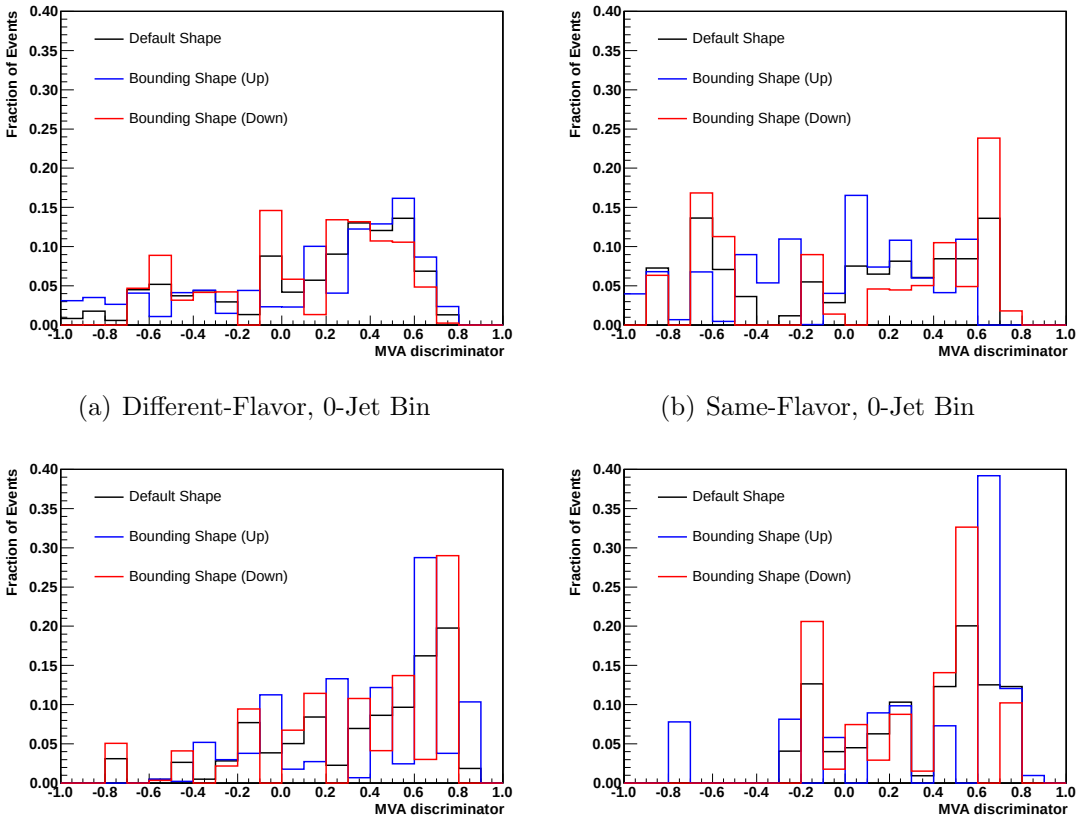
(c) Different-Flavor, 1-Jet Bin

(d) Same-Flavor, 1-Jet Bin

Figure 6-14: The default and alternative shapes used to model and propagate the systematic uncertainty due to the uncertainty in the p_T spectrum of the jet that produces the fake lepton are shown for the $m_{\text{Higgs}} = 130$ GeV/c^2 analysis.

For the systematic uncertainty due to the fake-lepton composition uncertainty, we calculate the relative difference between the fake rate method prediction and the

Monte Carlo simulation prediction in each bin of the BDT discriminator distribution, in the Monte Carlo closure test. One alternative shape are then obtained by multiplying the default shape prediction by these fractional differences. We mirror the difference between this alternative shape and the default shape to obtain the opposite extreme bound. These alternative shapes modeling the fake lepton composition uncertainty are shown in Figure 6-15 for the $m_{\text{Higgs}} = 130$ GeV/c^2 analysis.



(a) Different-Flavor, 0-Jet Bin

(b) Same-Flavor, 0-Jet Bin

(c) Different-Flavor, 1-Jet Bin

(d) Same-Flavor, 1-Jet Bin

Figure 6-15: The default and alternative shapes used to model and propagate the systematic uncertainty due to the uncertainty in the fake lepton composition are shown for the $m_{\text{Higgs}} = 130$ GeV/c^2 analysis.

6.2.3 Same Sign Control Region

Taking the event sample after the W^+W^--preselection but requiring two leptons with the same charge significantly enhances the contribution from fake lepton processes. By comparing the prediction with the observed data yield, the accuracy of the fake lepton background prediction is further cross checked.

Tables 6.15 and 6.16 summarize the estimates of all processes contributing to the same sign control region for events with zero and one counted jet. The W+jets process contributes a large fraction of the total yield in this control region. We observe that the predicted yields agree well with the observation, providing additional validation of the W+jets background estimation method.

Background Process	Yield
W+jets	14.4 ± 5.3
$W+\gamma$, $W+\gamma^*$	13.2 ± 4.9
WZ, ZZ	13.6 ± 1.4
Other	3.2 ± 0.2
Total	44.3 ± 7.4
Data	39

Table 6.15: Summary of the background yields from various processes in events with two same charge leptons and zero counted jets after W^+W^--preselection

Background Process	Yield
Fake Lepton Background	13.5 ± 5.0
$W+\gamma$, $W+\gamma^*$	9.6 ± 2.7
WZ, ZZ	15.3 ± 0.2
Other	3.8 ± 0.8
Total	42.1 ± 5.7
Data	37

Table 6.16: Summary of the background yields from various processes in events with two same charge leptons and one counted jets after W^+W^--preselection

6.3 Top Background

Events from top pair production and single top production can enter as background in the Higgs signal region when the top quarks decay in the leptonic channels and the associated b-quarks are left unidentified. Unidentified b-quarks are due to an inefficiency of the tagging procedure or due to the quark falling outside of the acceptance region in p_T and η.

Different data driven approaches are pursued for each of the jet bins, all relying on the assumption that the efficiency for a b-quark from the decay of a top quark to be identified as such is independent of the number of reconstructed jets. We give a general description of the method first, and follow with details for each jet bin in subsequent sections. In each jet bin, we define a top background dominated control region by inverting the top-veto cuts, requiring that the event has either a soft muon or a b-tagged jet (Section 4.1). We multiply the event yield from this top background control region by the top tagging efficiency, measured in a corresponding orthogonal data sample dominated by top production, to obtain the estimated top background in the signal region. Depending on the jet bin, there may be other non-top backgrounds present in the control region, which must be estimated and subtracted. We account for the W+jets background using the fake-rate method (Section 6.2), the Drell-Yan background using Monte Carlo simulation corrected by the Monte-Carlo-to-data scale factor (Section 6.1), and all other backgrounds using Monte Carlo simulation. In practice we compute a Monte-Carlo-to-data scale factor at the W^+W^--preselection level and use the simulation to extrapolate to each of the different signal regions. The precise procedure to obtain normalization scale factors for each jet bin is described in greater detail in following sections.

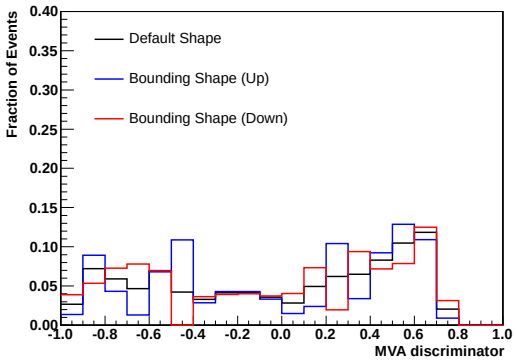
Although a data driven approach is used, there are still some second order dependencies on the Monte Carlo prediction. The most important dependence is in the prediction of the relative contribution of the $t\bar{t}$ and single top processes. We use the POWHEG next-to-leading order Monte Carlo generator [82] for this prediction, em-

ploying two different schemes for treating the overlap in next-to-leading order single top production diagrams and leading order $t\bar{t}$ production diagrams [83, 84]. In the "diagram removal" scheme the common diagrams are discarded at the matrix element level and in the "diagram subtraction" scheme the common diagrams are discarded at the cross section level. We observe a negligible difference in the normalization of the final prediction between the two schemes.

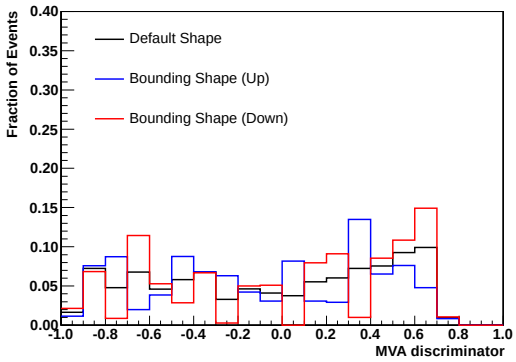
For the BDT shape analysis, the BDT discriminator distribution for top background is estimated from the Monte Carlo simulation. The systematic uncertainty for the $t\bar{t}$ shape is modeled using the vertical morphing technique. The alternative shapes are obtained from the Madgraph Monte Carlo generator [85], whose difference with the NLO POWHEG generator is taken as a conservative estimate of higher order corrections. The opposite extreme bound is obtained by mirroring the difference between the shape predicted by Madgraph and the shape predicted by POWHEG. The systematic uncertainty for the single top shape is modeled by the analogous morphing of the shape obtained from the diagram subtraction scheme and the shape obtained by mirroring the difference between this shape and the default shape obtained using the diagram removal scheme. These alternative shapes are shown in Figure 6-16 for the $m_{\text{Higgs}} = 130$ GeV/c^2 analysis.

6.3.1 Top Background in the 1-Jet Bin

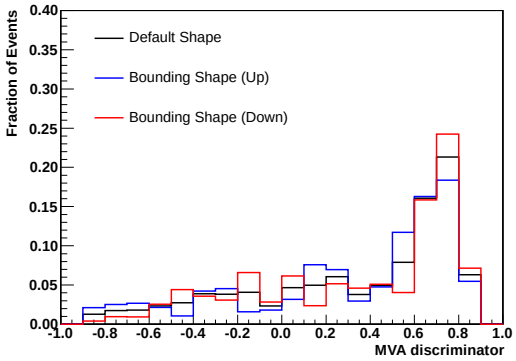
We first estimate the top background in the one-jet bin because the results of this estimate are needed for the estimate of the top background in the zero-jet bin. We extrapolate in the b-tagging efficiency for the highest p_T jet in the event. We take events with two counted jets, requiring that the lower p_T jet is b-tagged and that there are no other b-tagged jets with p_T below the jet counting threshold of 30 GeV/c, and measure the efficiency for the higher p_T to pass the b-tagging requirement. These requirements significantly increase the purity of $t\bar{t}$ events. For single top production, only a single b-quark is produced instead of two b-quarks in the case of $t\bar{t}$. This yields a potential bias, and therefore we subtract the contribution from single top using the



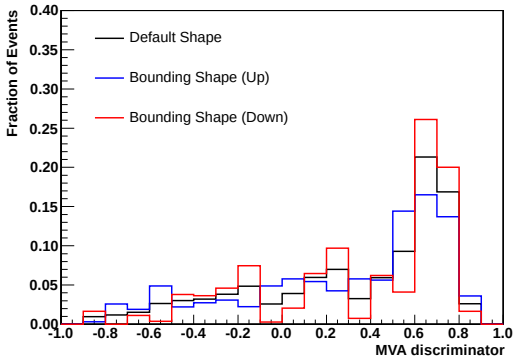
(a) Different-Flavor, 0-Jet Bin



(b) Same-Flavor, 0-Jet Bin



(c) Different-Flavor, 1-Jet Bin



(d) Same-Flavor, 1-Jet Bin

Figure 6-16: The default and alternative shapes for the top background in the $m_{\text{Higgs}} = 130$ GeV/c^2 analysis.

Monte Carlo prediction, and then compute the efficiency.

Next, we count the number of data events passing the W^+W^--preselection with one and only one counted jet. We require that the only counted jet in the event is tagged as a b jet, thus defining the top background control region in the one-jet bin. All non-top backgrounds are subtracted using the prediction from Monte Carlo, and the result is multiplied by the b-tagging efficiency computed above to obtain the estimate for the top background in the one-jet bin of the W^+W^--preselected sample. We then derive a scale factor, dividing the data driven prediction by the Monte Carlo prediction. This scale factor is used to correct the Monte Carlo prediction of the top background in the signal region corresponding to each Higgs mass hypothesis. A

	Data Measurement	$t\bar{t}$ MC	single top MC
0-Jet Bin Top-Tagging Efficiency	0.378 ± 0.014	0.355 ± 0.002	0.149 ± 0.010

Table 6.17: The efficiency for finding a soft muon or a b-tagged jet between 10 GeV/c and 30 GeV/c per b-quark, measured in data, in the $t\bar{t}$ and the single top Monte Carlo simulation samples.

closure test of the method performed on Monte Carlo simulation gives a statistically significant but small bias of 3% which is propagated as a systematic uncertainty on the top background estimate. The total systematic uncertainty is roughly 6%, dominated by the statistical uncertainty of the yield in the top control region. Using this method, we predict a top background in the W^+W^- preselected region of 352 ± 22. This yields a Monte-Carlo-to-data scale factor of 1.06 ± 0.07.

6.3.2 Top Background in the 0-Jet Bin

For top backgrounds in the zero-jet bin, both b-quarks must develop into jets below the jet counting threshold of 30 GeV/c. To estimate the yield of such background events, we extrapolate in the combined top tagging efficiency of finding either an additional soft muon or a jet with p_T between 10 GeV/c and 30 GeV/c that is b-tagged in the event. We first measure the efficiency (ϵ_b) of finding a soft muon or a b-tagged jet with p_T between 10 GeV/c and 30 GeV/c per b-quark. This is measured in events with a single b-tagged counted jet, dominated by top processes. We do not include any soft muons which are within a ΔR < 0.3 cone around the counted jet. The single top events, comprising about 20% of this sample, have lower top tagging efficiency due to the fact that some fraction of single top events produce only one final state b-quark. To remove the potential bias to the top tagging efficiency caused by the presence of such events, we subtract the contribution from the single top process using the Monte Carlo prediction corrected by the scale factor that we measured for the one-jet bin in the section above. Table 6.17 compares the top-tagging efficiency computed from data and the Monte Carlo simulation, showing reasonable agreement.

Using the per b-quark top-tagging efficiency, we compute the per event top-tagging efficiency ($\epsilon_{\text{top-tag}}$) using the following formula:

$$\epsilon_{\text{top-tag}} = f^{MC}_{2-b-\text{quarks}}(1-(1-\epsilon_b^{data})^2) + (1-f^{MC}_{2-b-\text{quarks}})\epsilon_b^{data}, \tag{6.4}$$

where $f^{MC}_{2-b-\text{quarks}}$ is the fraction of top events produced with two b-quarks inside the p_T and η acceptance. In practice we compute this fraction from the Monte Carlo simulation as follows:

$$f^{MC}_{2-b-\text{quarks}} = \frac{N^{MC}_{t\bar{t}} + N^{MC}_{\text{single top}} \times \epsilon^{\text{tagged,1jet}}_{\text{single top MC}}}{N^{MC}_{t\bar{t}} + N^{MC}_{\text{single top}}}, \tag{6.5}$$

where $\epsilon^{\text{tagged,1jet}}_{\text{single top MC}}$ is the efficiency of finding a soft muon or a b-tagged jet between 10 GeV/c and 30 GeV/c in single top Monte Carlo events with one counted jet. This formula ignores the inefficiency for tagging a generated b-quark, yielding a small systematic uncertainty on this fraction. Using this formula we measure $f^{MC}_{2-b-\text{quarks}}$ to be 0.705, and the per event top tagging efficiency to be 0.544 ± 0.044. From this efficiency we predict a top background after W^+W^--preselection of 120 ± 27 yielding a Monte-Carlo-to-data scale factor of 1.18 ± 0.22. A closure test of the method performed on the Monte Carlo simulation yields a bias of 5% which is propagated as a systematic uncertainty. The remaining systematic uncertainties result from the uncertainty on the background in the top-tagged control region, dominated by the uncertainty in the W+jets background.

6.3.3 Top Background in the Vector Boson Fusion Analysis

To estimate the top background in the vector boson fusion category, we extrapolate in the b-tagging efficiency of the most central jet among the two counted jets. We measure this efficiency for the most central jet in events with two counted jet and passing the W^+W^--preselection. Due to the large contribution of Drell-Yan background, only events with one electron and one muon are used. All non-top backgrounds and

η region	Data Measurement	$t\bar{t}$ MC
$0.0 \leq \|\eta\| < 0.5$	0.68 ± 0.02	0.741 ± 0.002
$0.5 \leq \|\eta\| < 1.0$	0.74 ± 0.02	0.745 ± 0.003
$1.0 \leq \|\eta\| < 1.5$	0.71 ± 0.03	0.685 ± 0.004
$1.5 \leq \|\eta\| < 2.0$	0.73 ± 0.04	0.623 ± 0.006
$2.0 \leq \|\eta\|$	0.40 ± 0.56	0.396 ± 0.009

Table 6.18: The efficiency for b-tagging the most central jet above 30 GeV/c in events with two counted jets measured in data and in the $t\bar{t}$ Monte Carlo simulation.

the single top background is subtracted. We measure this efficiency separately in five different bins of η to parametrize its dependence on pseudorapidity. Table 6.18 compares the b-tagging efficiency measured from data and from the $t\bar{t}$ Monte Carlo simulation. The b-tagging efficiency measured in data is used for the background prediction. A small difference between data and Monte Carlo is observed.

Next we define the VBF top control region to be events passing the VBF jet selection and the W^+W^--preselection except that the most central counted jet is required to be b-tagged. Events in this control region are then subdivided into η bins of the most central jet. In each bin we subtract the non-top background and multiply the remaining top yields by the corresponding b-tagging efficiencies from Table 6.18 to give the top background prediction in the VBF category passing the W^+W^--preselection. Summing the top background in each η bin yields a total background of 8.0 ± 3.4, and a Monte-Carlo-to-data scale factor of 0.66 ± 0.43. The closure test of the method performed on the Monte Carlo simulation yields a bias of 6% which is propagated as a systematic uncertainty. The total uncertainty is dominated by the statistical uncertainty of the yield in the top control region.

6.4 W^+W^- Background

The largest background for the Higgs search is standard model production of W^+W^-. This background is estimated using a combination of Monte Carlo simulation and control regions in data. From the dilepton mass distribution shown in Figure 6-17,

we observe that the region above 100 GeV/c^2 is relatively pure in W^+W^- background with no contamination from Higgs signal with mass hypotheses below 200 GeV/c^2. This "high mass region" is therefore used to provide a normalization of the W^+W^- background rate for any mass hypothesis below 200 GeV/c^2.

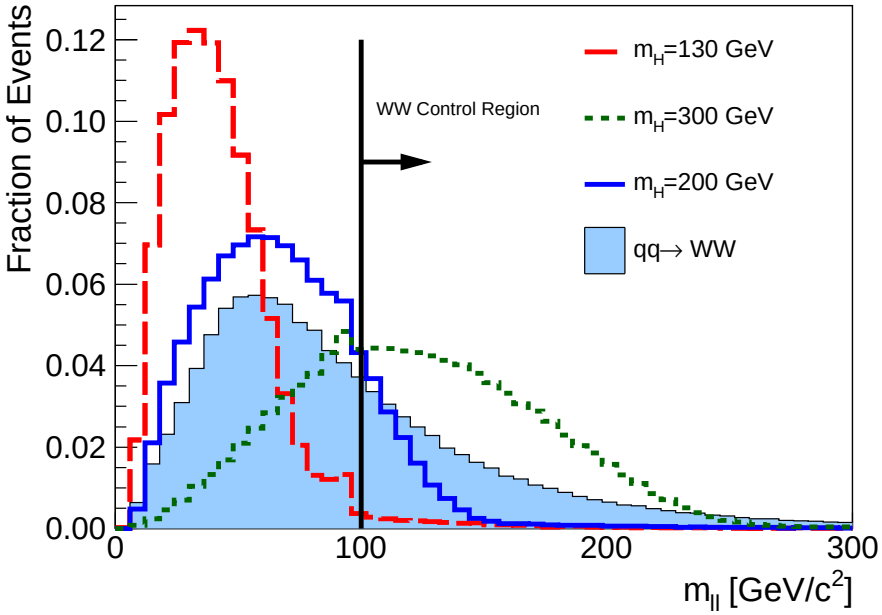


Figure 6-17: The distribution of the dilepton mass for the standard model W^+W^-process and three different signal hypothesis, illustrating the W^+W^- control region used for low mass Higgs hypotheses. Each distribution has been normalized to unit area.

For each mass hypothesis in the cut-based analysis, the W^+W^- control region is defined by events passing the W^+W^--preselection, the p_T requirements on the leading and trailing leptons from Table 4.1, and the requirement that $m_{\ell\ell}$ is larger than 100 GeV/c^2. For the W^+W^- control region in the BDT shape analysis, we do not require the p_T cuts on the two leptons, and add the requirement of inverting the $m_{\ell\ell}$ cuts defined in Table 4.2. Each such control region contains contributions from the Drell-Yan background, the W+jets background, and the top background. We predict these backgrounds using the methods described in the previous sections. Then for each control region we compare the background subtracted data yield with

the yield predicted by the Monte Carlo simulation and derive a Monte-Carlo-to-data scale factor. This scale factor is then applied in each corresponding signal region for each Higgs mass hypothesis. The relevant yields and scale factors are summarized in Tables 6.19 and 6.20 for the cut based and BDT shape analysis selections, respectively.

The statistical uncertainty in the control region yield contributes roughly half of the total systematic uncertainty that is propagated to the estimate of the W^+W^- background. The uncertainty on the top and W+jets background estimate in the control region contributes the remaining half. Theoretical systematic uncertainties on the extrapolation from the high mass to the low mass region due to missing higher order corrections and parton distribution function uncertainties are estimated to be 6% and are also propagated to the final result [86]. The cross section for gluon fusion production of W^+W^- has a large theoretical uncertainty. We propagate a systematic uncertainty of 30% on the gluon fusion component of the W^+W^- background [87], which account for about 5% of the total W^+W^- background.

For Higgs mass hypotheses above 200 GeV/c^2 the signal contamination in the control region becomes significant, and therefore we must estimate the W^+W^- background entirely using the Monte Carlo simulation and the theoretical cross section. As a result, a number of theoretical uncertainties must be propagated. The systematic uncertainty on the W^+W^- production cross section and the efficiency of the selection cuts due to the parton distribution function uncertainties are 4% [88]. For the systematic uncertainty due to missing higher order corrections, we factorize its total effect into its effect on the production cross section and the selection efficiency, and its effect on the migration of events between different jet bins.

The effect of missing higher order corrections on the production cross section and the selection efficiency is estimated to be 6% by varying the renormalization and factorization scales by factors of 1/2 and 2 using next to leading order calculations [88–90]. Its effect on jet bin migration are described in further detail in Section 7.5.3.

In the BDT shape analysis, the BDT discriminator distribution for the W^+W^- background is estimated from the Madgraph Monte Carlo simulation [85]. Two dif-

0-Jet Bin Analysis			
m_{Higgs} hypothesis	W^+W^- yields (data)	W^+W^- yields (MC)	Scale Factor (Data/MC)
W^+W^--Preselection	352.4 ± 26.9	322.9 ± 2.7	1.09 ± 0.08
118-130 GeV/c^2	352.4 ± 26.9	322.9 ± 2.7	1.09 ± 0.08
135 GeV/c^2	353.3 ± 26.9	322.3 ± 2.7	1.10 ± 0.08
140 GeV/c^2	353.9 ± 26.7	320.4 ± 2.7	1.10 ± 0.08
150 GeV/c^2	315.0 ± 25.0	290.4 ± 2.5	1.08 ± 0.09
160 GeV/c^2	315.4 ± 24.9	290.0 ± 2.5	1.09 ± 0.09
170 GeV/c^2	312.7 ± 24.9	288.3 ± 2.5	1.08 ± 0.09
180 GeV/c^2	307.9 ± 24.7	286.6 ± 2.5	1.07 ± 0.09
190 GeV/c^2	305.1 ± 24.7	284.0 ± 2.5	1.07 ± 0.09
1-Jet Bin Analysis			
m_{Higgs} hypothesis	W^+W^- yields (data)	W^+W^- yields (MC)	Scale Factor (Data/MC)
W^+W^--Preselection	169.1 ± 22.0	145.5 ± 1.8	1.16 ± 0.15
118-130 GeV/c^2	169.1 ± 22.0	145.5 ± 1.8	1.16 ± 0.15
135 GeV/c^2	169.9 ± 21.9	145.1 ± 1.8	1.17 ± 0.15
140 GeV/c^2	171.4 ± 21.8	144.3 ± 1.8	1.19 ± 0.15
150 GeV/c^2	155.9 ± 20.4	130.7 ± 1.7	1.19 ± 0.16
160 GeV/c^2	156.1 ± 20.4	130.6 ± 1.7	1.19 ± 0.16
170 GeV/c^2	154.4 ± 20.4	130.2 ± 1.7	1.19 ± 0.16
180 GeV/c^2	152.6 ± 20.3	129.8 ± 1.7	1.18 ± 0.16
190 GeV/c^2	151.9 ± 20.3	129.2 ± 1.7	1.18 ± 0.16

Table 6.19: The background subtracted W^+W^- yields from data, the Monte Carlo simulation W^+W^- yields, and the corresponding Monte-Carlo-to-data scale factors are shown for the W^+W^- control regions defined by the cut-based analysis selections corresponding to each Higgs mass hypothesis. In the W^+W^- control region, there is an excess of data over the Monte Carlo prediction of about 15% for the zero-jet bin, and about 25% for the one-jet bin. This disagreement likely reflects a mismodeling of the jet spectrum that is produced in association with the diboson pair, and is will be scrutinized further in the future. These factors are used to scale up the predicted W^+W^- background yields in the signal region.

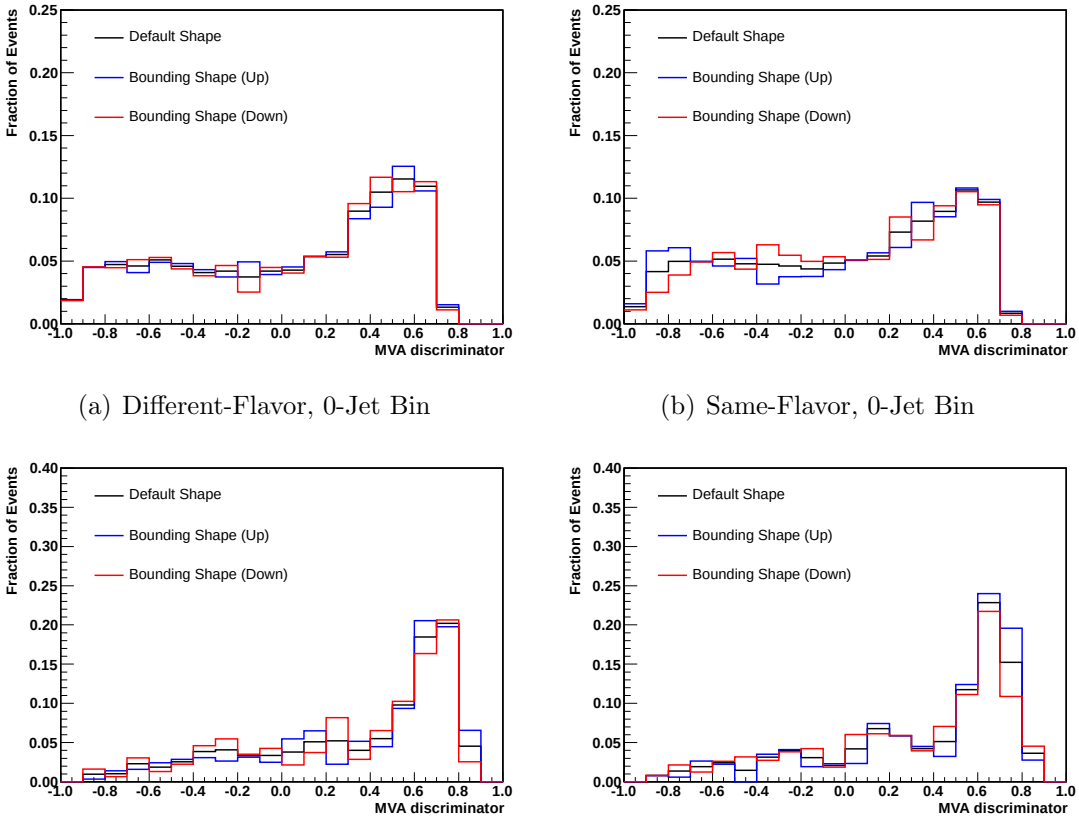
0-Jet Bin Analysis			
m_{Higgs} hypothesis	W^+W^- yields (data)	W^+W^- yields (MC)	Scale Factor (Data/MC)
W^+W^--Preselection	352.4 ± 26.9	322.9 ± 2.7	1.09 ± 0.08
118-170 GeV/c^2	352.4 ± 26.9	322.9 ± 2.7	1.09 ± 0.08
180 GeV/c^2	307.7 ± 25.0	285.7 ± 2.5	1.08 ± 0.09
190 GeV/c^2	262.0 ± 22.5	240.8 ± 2.3	1.09 ± 0.09
1-Jet Bin Analysis			
m_{Higgs} hypothesis	W^+W^- yields (data)	W^+W^- yields (MC)	Scale Factor (Data/MC)
W^+W^--Preselection	169.1 ± 22.0	145.5 ± 1.8	1.16 ± 0.15
118-170 GeV/c^2	169.1 ± 22.0	145.5 ± 1.8	1.16 ± 0.15
180 GeV/c^2	149.1 ± 20.2	129.8 ± 1.7	1.15 ± 0.16
190 GeV/c^2	131.4 ± 18.3	110.1 ± 1.6	1.19 ± 0.17

Table 6.20: The background subtracted W^+W^- yields from data, the Monte Carlo simulation W^+W^- yields, and the corresponding Monte-Carlo-to-data scale factors are shown for the W^+W^- control regions defined by the BDT shape analysis selection corresponding to each Higgs mass hypothesis. The scale factors are consistent with those observed for the cut-based selection in Table 6.19.

ferent shape systematic uncertainties are propagated in the analysis. In both cases we use the vertical morphing technique described before. The first systematic uncertainty covers the difference between the default prediction obtained from Madgraph, which gives a leading order prediction for each jet bin, and the inclusive next-to-leading order prediction obtained from MC@NLO [90]. One extreme bound is obtained from the shape predicted by the MC@NLO Monte Carlo generator, and the opposite extreme bound is obtained by mirroring the difference between this shape and the default shape obtained from the Madgraph Monte Carlo generator. The default and alternative shapes for this effect are shown in Figure 6-18.

The second systematic uncertainty accounts for the missing higher order corrections of the next-to-leading order shape prediction. We obtain the BDT discriminator shapes from the MC@NLO generator where the renormalization scale and factorization scale are multiplied by 1/2 and 2. The relative difference of these shapes to the nominal MC@NLO shape prediction is computed,Ṫhese fractional differences are

multiplied by the default Madgraph prediction, bin-by-bin, to derive the shapes corresponding to the extreme bounds. These alternative shapes are shown in Figure 6-19.



(a) Different-Flavor, 0-Jet Bin

(b) Same-Flavor, 0-Jet Bin

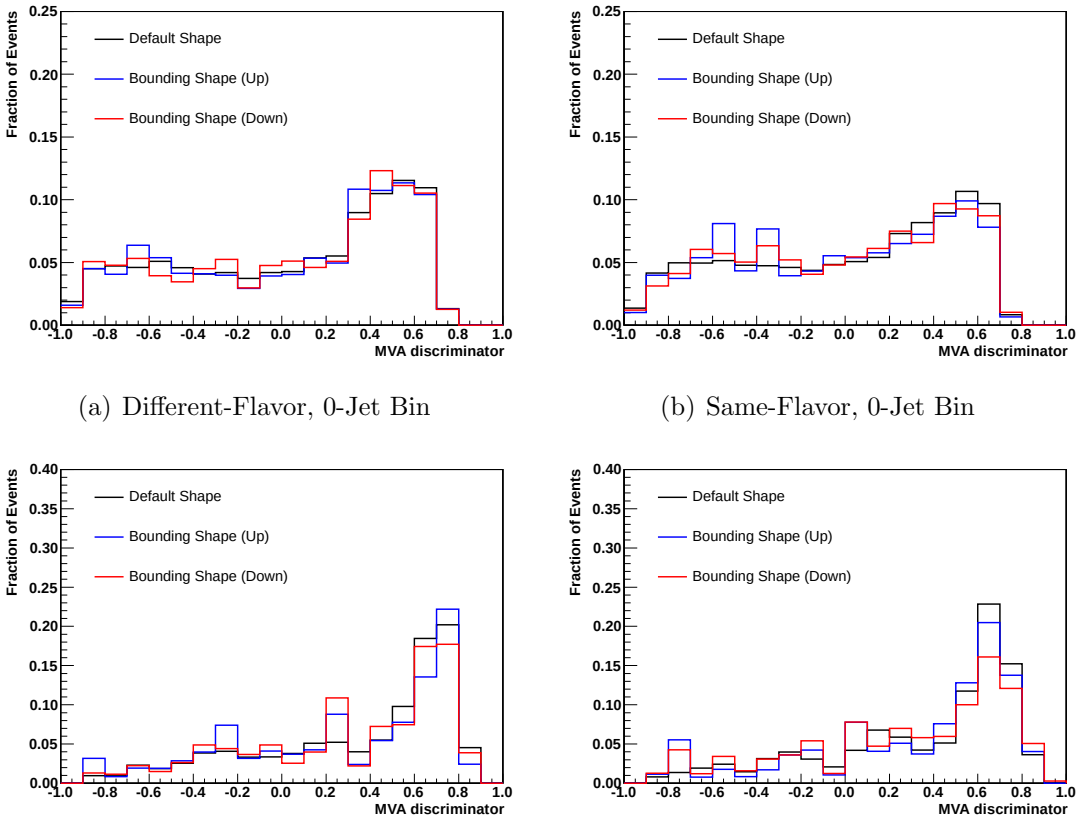
(c) Different-Flavor, 1-Jet Bin

(d) Same-Flavor, 1-Jet Bin

Figure 6-18: The default BDT discriminator distribution shape obtained from the Madgraph generator and alternative shapes obtained from the MC@NLO generator for the W^+W^- background in the $m_{\text{Higgs}} = 130$ GeV/c^2 analysis.

6.5 WZ, ZZ, and $W\gamma^*$ Background

Standard model production of WZ and ZZ decaying to leptons contribute to the signal region as background, either because one or two of the leptons fall outside of the acceptance region, fail the minimum lepton p_T requirement, or are left unidentified. The ZZ process also contributes in the case where one Z decays to a pair of neutrinos



(a) Different-Flavor, 0-Jet Bin

(b) Same-Flavor, 0-Jet Bin

(c) Different-Flavor, 1-Jet Bin

(d) Same-Flavor, 1-Jet Bin

Figure 6-19: The shape of the BDT discriminator distribution for the W^+W^- background in the $m_{\text{Higgs}} = 130$ GeV/c^2 analysis obtained from the default Madgraph prediction is compared to the alternative shapes intended to model the missing higher order corrections beyond the NLO order. These alternative shapes are obtained from the MC@NLO Monte Carlo generator.

and the other Z decays to two leptons. These background contributions and the shape of the BDT discriminator are estimated using Monte Carlo simulation. The $W\gamma^*$ process where the Z/γ^* boson is far off-shell is particularly important because a significant fraction of such events mimic the Higgs signal in a number of kinematic features [91]. The events that contribute to the Higgs signal region are produced in the unusual region of phase space, where the mass of the γ^* approaches the kinematic threshold of twice the lepton mass. Therefore, one must pay particular attention to the input parameters of the Monte Carlo generator.

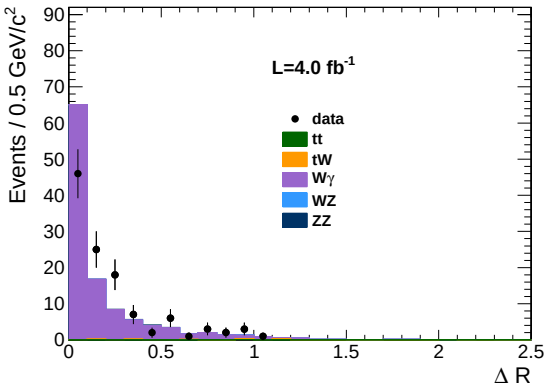
The Madgraph generator [85] was used to produce the $W\gamma^*$ process. The p_T of the two leading leptons is required to be larger than 5 GeV/c at generator level, and no p_T cut is made on the third lepton. We generate Monte Carlo events covering the phase space all the way to the kinematic threshold by allowing the mass of the γ^* to be as small as two times the electron mass and two times the muon mass for the decays to electrons and muons respectively. Finally, the electron mass and muon mass are explicitly retained for the Monte Carlo generation, because we observed that the massless approximation, which fails near the kinematic threshold, results in a significantly incorrect γ^* mass spectrum.

The normalization of the $W\gamma^*$ process is measured in a control region defined by selecting events with three muons or one electron and two muons with low γ^* mass and high missing transverse energy. This data sample is dominated by the $W\gamma^*$ process. The events where the γ^* decays to two electrons suffer from too much fake electron background and are unsuitable for this measurement. Events from this process are likely to have two muons with very close trajectory, and thus we modify the definition of the particle based isolation observable to exclude any muon passing all requirements but the isolation requirement for selecting events in the $W\gamma^*$ control region. This modification of the isolation requirement increases the efficiency for selecting $W\gamma^*$ events in this control region. The p_T of the leading, trailing, and third lepton is required to be larger than 20 GeV/c, 10 GeV/c, and 3 GeV/c, respectively. In the e$\mu^+\mu^-$ case, the electron p_T is further required to be larger than 10 GeV/c. The $\mu^+\mu^-$ pair with the lowest dimuon mass is selected as the pair decaying from the γ^* in the $\mu\mu\mu$ case.

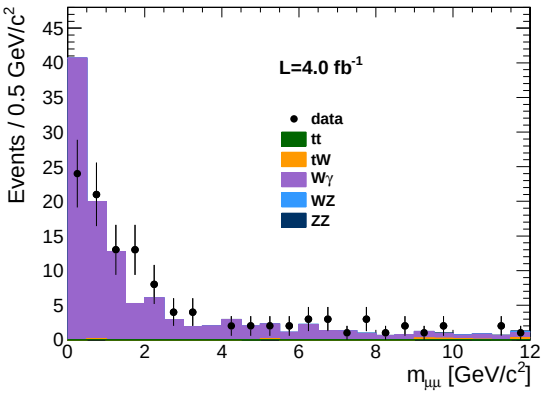
Contributions from non-$W\gamma^*$ backgrounds in the control region are suppressed using a number of different methods. We suppress top backgrounds by requiring that the event has less than two counted jets and that there are no b-tagged jets with p_T greater than 10 GeV/c. We suppress QCD background by requiring that the minimum missing transverse energy is greater than 20 GeV/c, and that the transverse mass formed by the unpaired lepton and the missing transverse energy is greater than

20 GeV/c. Finally, we suppress the WZ background by requiring that the mass of the dimuon pair selected to be from the γ^* decay is less than 12 GeV/c^2.

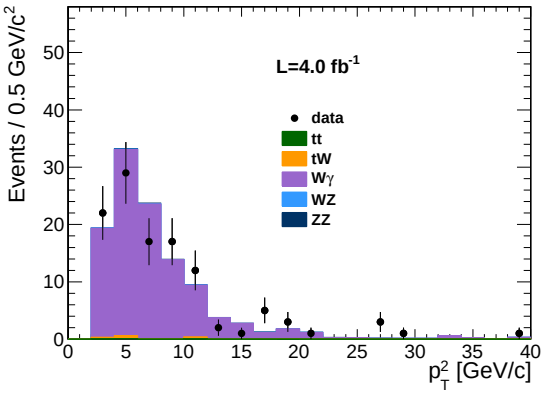
We compare distributions of the ΔR between the two muons from the γ^* decay, the dimuon mass, and the p_T of the softest muon for data and Monte Carlo prediction in Figure 6-20. The $W\gamma^*$ control region measurement is performed using 4 fb^{-1} of integrated luminosity. The small differences between the data shape and the shape predicted by the Monte Carlo simulation is due to the fact that the Monte Carlo simulation mismodels the lepton selection efficiency for leptons with p_T below 10 GeV/c.



(a) ΔR



(b) Dimuon mass



(c) p_T of the third lepton

Figure 6-20: Comparison of data and Monte Carlo prediction for the ΔR between the two muons from the γ^* decay , the dimuon mass, and the p_T of the third lepton in the $W\gamma^*$ control region. The Monte-Carlo-to-data scale factor for the $W\gamma^*$ process of 1.6 has been applied.

To estimate the contribution of fake lepton backgrounds from QCD multijet production and $t\bar{t}$ production to the $W\gamma^*$ control region, we invert the charge requirement for the muon pair selected to result from the γ^* decay and require that these two muons have the same charge. The fake lepton backgrounds are expected to contribute approximately equally to the same charge case as the opposite charge case. We observe no such events in the e$\mu^+\mu^-$ case and four such events in the $\mu\mu\mu$ case, and these yields are used as the prediction for the fake lepton background contribution to the $W\gamma^*$ control region. Table 6.21 summarizes the data yields, non-$W\gamma^*$ background predictions, and the Monte-Carlo-to-data scale factors measured in the $W\gamma^*$ control region. We derive a Monte-Carlo-to-data scale factor of 1.6 ± 0.2 for this process. A systematic uncertainty of 30% is estimated for this normalization accounting for uncertainties in the extrapolation from the control region to the signal region.

Final State	$e\mu\mu$	$\mu\mu\mu$	Combined
Data Yields	26	88	114
Background Prediction from Same Charge Control Region	0	4	4
$W\gamma^*$ MC Prediction	19 ± 1	49 ± 2	68 ± 2
Data to MC Scale Factor	1.4 ± 0.3	1.7 ± 0.2	1.6 ± 0.2

Table 6.21: Data yields, background predictions and the measured Monte Carlo to data scale factors, shown separately for the e$\mu^+\mu^-$ and $\mu\mu\mu$ final states. The measurement is performed in a data sample corresponding to an integrated luminosity of 4 fb^{-1}

.

6.6 Drell-Yan $Z/\gamma^* \rightarrow \tau^+\tau^-$ Background

Neutrinos from leptonic decays of the tau give typical missing transverse energy on the order of 15 GeV/c for $Z/\gamma^* \rightarrow \tau^+\tau^-$ events. Due to the large number of interactions per bunch crossing, the missing transverse energy resolution becomes comparable to the natural missing transverse energy produced by the neutrinos. The missing transverse energy resolution is not well modeled by the Monte Carlo simulation.

Therefore, we use a data driven method to predict this background contribution. Data events from the $Z \to \mu^+\mu^-$ process are selected, requiring two selected muons with a dimuon mass between 70 GeV/c^2 and 110 GeV/c^2. Both muons are subsequently replaced by the decay products of a leptonically decaying tau at the reconstructed particle level. This method is typically referred to as tau-embedding. The missing transverse energy is recomputed taking into account the neutrinos produced by the tau decay. This event sample is normalized to the expected inclusive $Z/\gamma^* \to \tau^+\tau^-$ yield estimated from Monte Carlo simulation corrected by lepton selection efficiency scale factors. The shape of the BDT discriminator for this background process is also predicted using this procedure. After all selection requirements, this background contributes only a fraction of a percent of the total background. The systematic uncertainty for this background prediction is estimated by a closure test of the method on $Z/\gamma^* \to \tau^+\tau^-$ Monte Carlo simulation, where we observe a bias of 10% in the predicted yield for events with one electron and one muon after requiring that the minimum projected missing transverse energy is greater than 20 GeV/c.

6.7 Establishing Control of the W^+W^- signature

Using all the methods described in the previous subsections to estimate the various background processes allows us to predict the total yields and distributions of kinematic observables in the W^+W^- dominated control region obtained after imposing all W^+W^--preselection requirements. The observed yield in data is compared with the predicted background yields in Table 6.22, separately for events with zero and one counted jet. The same comparison is made, separating the sample into events with two same-flavor leptons and events with two different-flavor leptons in Tables 6.23 and 6.24, respectively. In all cases, we predict the observed data to within the total uncertainty.

Once we have verified that the yields are well predicted by the background estimates, we compare the distributions of some important kinematic observables. The

	data	All Bkg	$qq \to W^+W^-$	$gg \to W^+W^-$	$t\bar{t} + tW$	$W +$ jets
0-jet	1319	1289 ± 83	949 ± 73	58 ± 4	124 ± 28	64 ± 23
1-jet	904	915 ± 61	408 ± 53	25 ± 3	337 ± 21	43 ± 16

	WZ/ZZ	$Z/\gamma^* \to \ell\ell$	$W + \gamma$	$Z/\gamma^* \to \tau\tau$
0-jet	34 ± 2	21 ± 5	38 ± 11	1.8 ± 0.4
1-jet	30 ± 1	46 ± 8	14 ± 4	13 ± 3

Table 6.22: The number of observed data events and the number of expected background events are shown for an integrated luminosity of 4.9 fb^{-1}after applying the W^+W^--preselection requirements, separated into events with zero and one counted jet. The total uncertainty in the background prediction include both statistical and systematic components.

	data	All Bkg	$qq \to W^+W^-$	$gg \to W^+W^-$	$t\bar{t} + tW$	$W +$ jets
0-jet	517	491 ± 31	356 ± 27	22 ± 2	43 ± 10	14.5 ± 5.2
1-jet	329	320 ± 21	129 ± 17	8.8 ± 1.2	110 ± 7	7.6 ± 2.8

	WZ/ZZ	$Z/\gamma^* \to \ell\ell$	$W + \gamma$	$Z/\gamma^* \to \tau\tau$
0-jet	20 ± 1	20 ± 5	15 ± 5	0 ± 0
1-jet	12 ± 1	44 ± 7	9 ± 3	0 ± 0

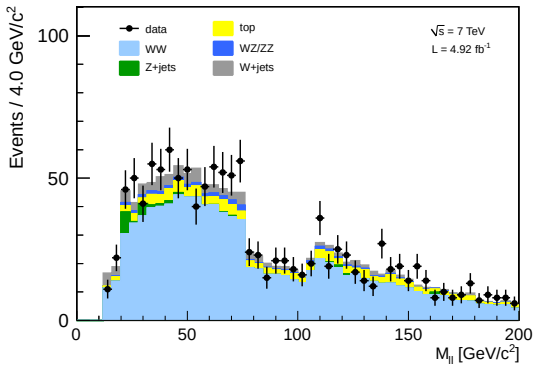
Table 6.23: The number of observed data events and the number of expected background events are shown for an integrated luminosity of 4.9 fb^{-1}after applying the W^+W^--preselection requirements, for events with two same-flavor leptons. The yields for events with zero and one counted jet are reported separately. The total uncertainty in the background prediction include both statistical and systematic components.

	data	All Bkg	$qq \to W^+W^-$	$gg \to W^+W^-$	$t\bar{t} + tW$	$W +$ jets
0-jet	802	798 ± 53	593 ± 46	36 ± 3	81 ± 18	50 ± 18
1-jet	575	595 ± 42	279 ± 36	16 ± 2	227 ± 14	36 ± 13

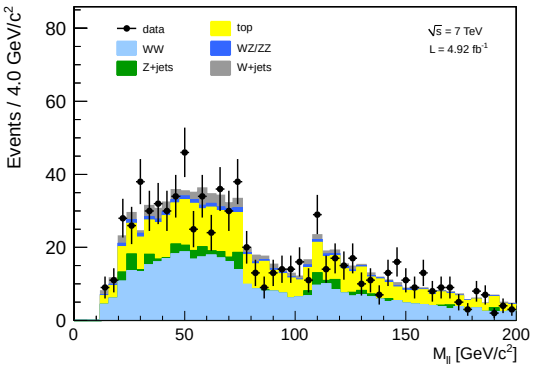
	WZ/ZZ	$Z/\gamma^* \to \ell\ell$	$W + \gamma$	$Z/\gamma^* \to \tau\tau$
0-jet	14 ± 1	0.5 ± 0.7	23 ± 7	1.8 ± 0.4
1-jet	18 ± 1	1.9 ± 0.5	5.3 ± 1.6	13 ± 3

Table 6.24: The number of observed data events and the number of expected background events are shown for an integrated luminosity of 4.9 fb^{-1}after applying the W^+W^--preselection requirements, for events with two different-flavor leptons. The yields for events with zero and one counted jet are reported separately. The total uncertainty in the background prediction include both statistical and systematic components.

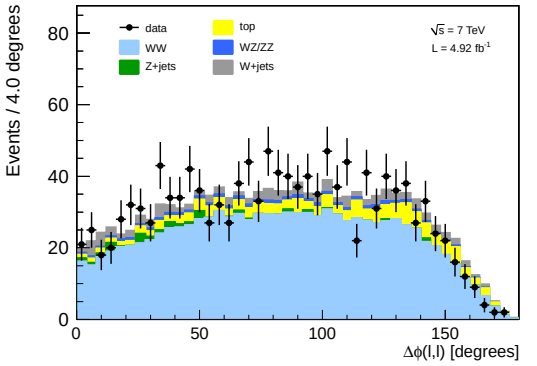
background predictions for the distributions are obtained from the Monte Carlo simulation, where we correct the normalization using the Monte-Carlo-to-data scale factors measured in the previous sections. The dilepton mass and the azimuthal angle in the transverse plane between the two leptons are two of the most discriminating kinematic observables between the Higgs signal and the main W^+W^- background. In Figure 6-21 we compare the observed and predicted distributions of these two observables. The prediction and the observed data are found to be in agreement within uncertainties. The transverse momentum of the trailing lepton, which is most sensitive to the W+jets and $W + \gamma$ background, is also shown in Figure 6-21, where we observe that these backgrounds are well modeled. Finally, in Figure 6-22 we show the dilepton mass distribution for events with two same-flavor leptons and events with two different-flavor leptons separately, because the contribution from the Drell-Yan background differs between the two event samples. We again observe that the backgrounds are well modeled. Overall, these results give us confidence in our background predictions and allow us to proceed ahead with the extraction of the Higgs signal.



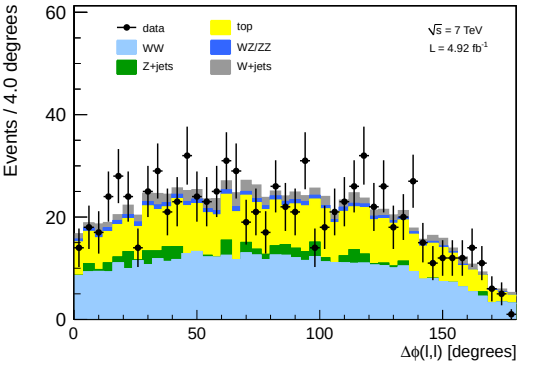
(a) Dilepton Mass, 0-Jet Bin



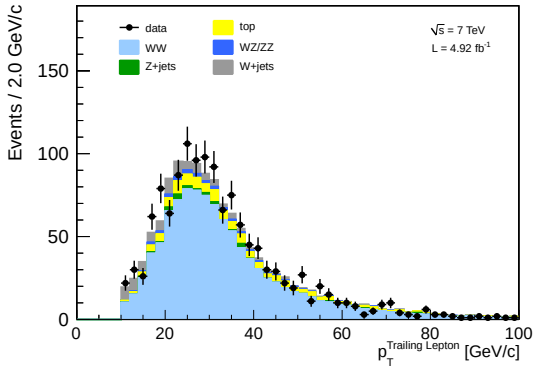
(b) Dilepton Mass, 1-Jet Bin



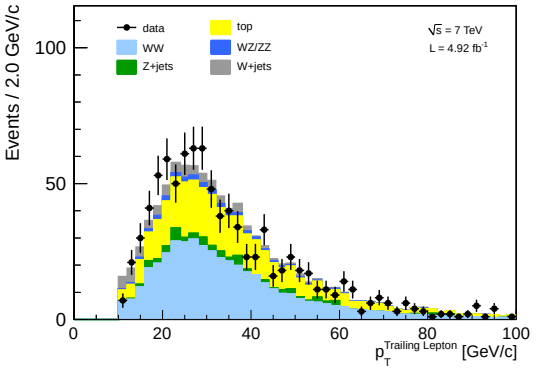
(c) $\Delta\phi_{\ell\ell}$, 0-Jet Bin



(d) $\Delta\phi_{\ell\ell}$, 1-Jet Bin

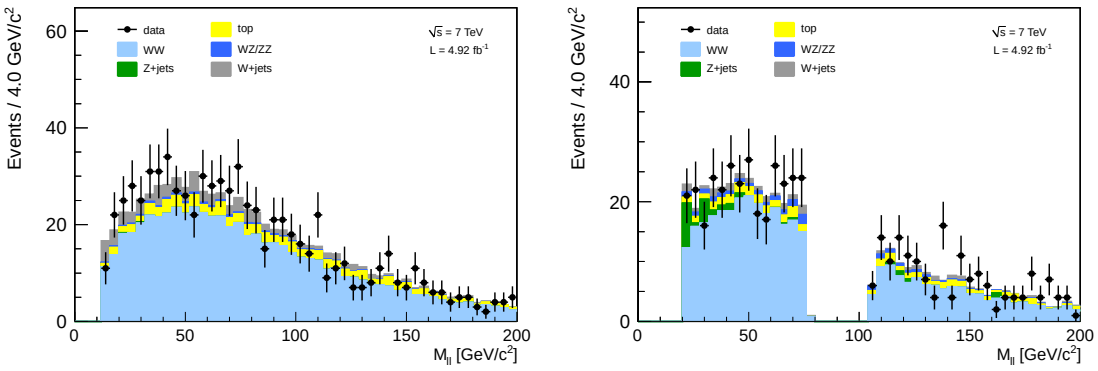


(e) Trailing Lepton p_T, 0-Jet Bin



(f) Trailing Lepton p_T, 1-Jet Bin

Figure 6-21: The background predictions for the dilepton mass distribution, the distribution of the $\Delta\phi$ between the two leptons, and the distribution of the p_T of the trailing lepton is compared with the observed data. This comparison is made for the zero-jet bin and one-jet bin separately. We observe that the background prediction models the observed data very well.



(a) Different-Flavor, 0-Jet Bin

(b) Same-Flavor, 0-Jet Bin

(c) Different-Flavor, 1-Jet Bin

(d) Same-Flavor, 1-Jet Bin

Figure 6-22: We compare the background prediction for the dilepton mass distribution and the observed data separately for the same-flavor and different-flavor final states. The gap in the dilepton mass distribution for the same-flavor cases shows the effect of the Z-veto selection requirement. Again, the background prediction models the observed data very well.

Chapter 7

Systematic Uncertainties

The predictions for the normalization and the distribution of kinematic observables of the signal and the background are affected by a number of systematic effects. These systematic effects may be experimental in nature, such as systematic uncertainties on the energy or momentum measurement of the charged leptons, or uncertainties on the resolution of the missing transverse energy; or they may be theoretical in nature, such as the effect of missing higher order corrections on the signal cross section.

These systematic effects may bias the estimates of the signal and background, and the uncertainty that is introduced by these biases must be propagated to the final result or measurement. In this chapter we enumerate all the systematic effects that have been considered, the constraints placed on them from auxiliary measurements or external information, and the procedures used to propagate these systematic uncertainties to the final result.

In subsequent sections, we will describe first the general procedure to parameterize and propagate the systematic uncertaintes, and then describe the details of the systematic uncertainties resulting from experimental aspects in Section 7.4, and theoretical aspects in Section 7.5.

7.1 Nuisance Parameters

For this analysis, the effect of a particular systematic uncertainty can be viewed as an uncertainty on the normalization of a particular signal or background process. For example, the effect of a mismeasurement of the jet energy scale will induce an uncertainty on the signal yields in the zero-jet, one-jet, and VBF bins. The effect of a systematic uncertainty on the distributions of kinematic observables may also be viewed as an uncertainty on the estimated event yields in each bin of the given observable. To give an example, the systematic uncertainty on the electron efficiency induces a corresponding uncertainty in the estimated yields for each bin of p_T of the electron in the event.

From these specific examples, we see that a generic way to parameterize and propagate systematic uncertainties is to construct some variables that parameterize the underlying quantities that induce each systematic uncertainty, and to make the estimated yields of signal and background processes a function of these variables. We refer to this set of variables as the set of nuisance parameters, and denote them as $\{\theta_k\}$. Thus the predicted yields and distributions are nominally functions of these nuisance parameters θ_k. In general, we do not know what the exact true value of the nuisance parameter is, but we can obtain constraints on these nuisance parameters through auxiliary measurements, or from external information. The external information may be other experimental measurements, or they may be theoretical calculations. We denote the measured value or best estimate of each nuisance parameter θ_k by $\tilde{\theta}_k$.

These measurements or calculations will generally provide some information on the uncertainty of the nuisance parameter estimate. From this information, we construct probability density functions for each nuisance parameter $\rho_k(\theta_k|\tilde{\theta}_k)$, which describes our knowledge about the true value of θ_k given our auxiliary measurement or external information.

In this analysis, we use exclusively the log-normal functional form as the approximation for all nuisance parameter probability density functions. This choice is based

on the fact that all important systematic uncertainties in this analysis are a result of uncertainties on normalizations and cross sections, and the log-normal functional form is well suited to describe such cases [92]. A log-normal random variable X is defined as an exponentiation of a Gaussian random variable: $X = e^{Z(\mu,\sigma)}$, where Z is a Gaussian random variable with mean μ and standard deviation σ. Thus, we may express the dependence of the log-normal random variable X in terms of a standard gaussian random variable Z, with mean of 0 and standard deviation equal to 1, as $X = e^{\mu+\sigma Z}$. This can be rewritten as:

$$X = X_0 \cdot \kappa^Z, \tag{7.1}$$

where $X_0 = e^{\mu}$ is the median of the log-normal distribution, and $\kappa = e^{\sigma}$. The variance of the log-normal distribution is a monotonically increasing function of the parameter κ, and therefore κ is used to describe the size of the systematic uncertainty that is modeled by the log-normal distribution. The probability density function for the log-normal is described by the following functional form:

$$\rho(x|X_0,\kappa) = \frac{1}{\sqrt{2\pi}\ln(\kappa)} \exp\left(-\frac{(\ln(x/X_0))^2}{2(\ln\kappa)^2}\right)\frac{1}{x}. \tag{7.2}$$

The log-normal probability density function is able to address asymmetric uncertainties in the case of non-negative observables like cross sections, in a way that avoids a number of pathologies that affect alternative parametrizations [92]. For relatively small systematic uncertainties, the κ parameter is simply related to the fractional uncertainty $\delta_X = \Delta X/X$ through: $\kappa_X = e^{\delta_X} \sim 1 + \delta_X$.

7.2 Shape Systematic Uncertainties

As described in Section 4.3, we make use of the shape of the BDT discriminator distribution in this analysis to increase the search sensitivity for the Higgs boson. The distributions of the kinematic observables that are used as input to the BDT

may be affected by various systematic uncertainties, such as the lepton momentum scale or the missing transverse energy resolution just to give two examples. Thus, the BDT distribution are also affected by these systematic effects. We describe the procedure used to parameterize the effect of any systematic uncertainty on the shape of the BDT distribution.

As described in the previous section, every systematic uncertainty is described by a nuisance parameter θ. The predicted yields of signal and background processes are functions of θ. The binned prediction for the BDT discriminator distributions can be viewed as N different predicted yields, each of which is a function of θ, where N is the number of bins. We denote f_k^0 as the nominal predicted yield in the k'th bin. For each source of systematic uncertainty we obtain estimates of the size of the effect on the predicted yields by deriving f_k^{up} and f_k^{down}, which represent the upper and lower bounds on the predicted yields induced by the given systematic effect.

To parameterize the systematic effect, we use a linear morphing technique which interpolates from the nominal value to the upper and lower bounds depending on the value of the nuisance parameter. We construct a smooth monotonic function $\chi(\theta)$ whose value ranges between -1 and 1 as the value of the nuisance parameter θ ranges between the one standard deviation upper and lower bounds, and is equal to 0 when θ equals to the best estimate $\tilde{\theta}$. The values of the predicted yields are determined by:

$$f_k(\theta) = \begin{cases} f_k^0 + \chi(\theta) \times (f_k^{\text{up}} - f_k^0) & \theta \geq \tilde{\theta} \\ f_k^0 + \chi(\theta) \times (f_k^0 - f_k^{\text{down}}) & \theta < \tilde{\theta} \end{cases} \tag{7.3}$$

The function χ is nominally chosen to be linear, but slightly modified such that the first and second derivatives of $f_k(\theta)$ are continuous at the best estimate value $\tilde{\theta}$.

7.3 Data Driven Background Estimates

The dominant source of systematic uncertainties for this analysis is from the data driven estimates of the backgrounds. These systematic uncertainties are due to either

the uncertainty introduced by biases intrinsic to the background estimation method, or the statistical uncertainties from the various background dominated control regions. These systematic uncertainties are propagated both for the predicted yields as well as their effects on the predicted distributions of important kinematic observables, in particular their effects on the distribution of the BDT discriminator that is used as a final discriminator observable in the shape analysis. The details of the estimates for these systematic uncertainties were described in Section 6.

7.4 Experimental Systematic Uncertainties

The dominant source of systematic uncertainties for this analysis is from the background estimates. These uncertainties are due to a combination of the statistical uncertainties from the different background dominated control regions, and the uncertainty from potential biases introduced by the background estimation method. The details of these systematic uncertainties have already been described in Chapter 6.

Besides the background estimates, there are a number of experimental systematic uncertainties which, individually, have minor effects but when collectively propagated have a non-negligible effect on the final result. We describe these systematic uncertainties, and the methods used to estimate and propagate them in the subsequent sections.

7.4.1 Luminosity

The measurement of the integrated luminosity collected by CMS is made on the basis of the amount of energy deposition in the HF detector, and cross checked by measurements of the rates of reconstructed clusters in the pixel detector. These rates were calibrated to an absolute luminosity measurement in special collision runs by means of the van der Meer scan technique [93, 94]. The total uncertainty of the integrated luminosity measurement is 2.2%, partly resulting from the uncertainties from the van der Meer scan measurement and partly resulting from the uncertainty

in the estimate of the LHC beam currents.

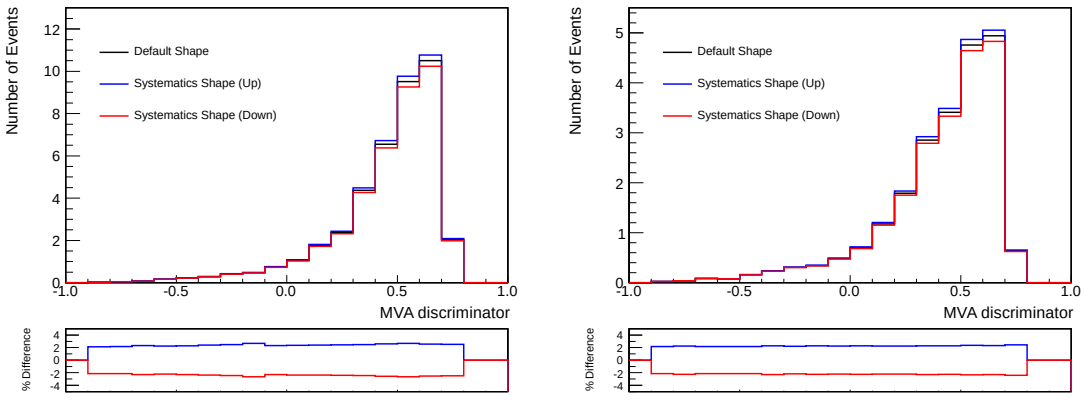
7.4.2 Lepton Selection Efficiency

The electron and muon selection efficiencies were measured in Sections 5.3 and 5.4. The uncertainties in those measurements are dominated by the uncertainty of the background determination in the likelihood fit and range between 5% in the low p_T bins to less than 0.1% in the high p_T bins, where backgrounds are negligible. These uncertainties are added in quadrature with the systematic uncertainty of 2% due to the uncertainties in the signal and background models that enter as input to the likelihood fit method [60].

We systematically shift the efficiency in each bin by the size of its total uncertainty and propagate this effect through to the predicted signal and background yields and the BDT discriminator shape. In the case of the cut-based analysis, this induces an average change in the signal and background yields of roughly 4% for the electron efficiency uncertainty, and 3% for the muon efficiency uncertainty. For the shape analysis, this procedure produces alternative BDT discriminator distributions. The effect of this systematic uncertainty on the shape analysis is parametrized using the linear morphing technique described in Section 7.2 using these alternative shapes to provide the upper and lower bounds. An example of these alternative shapes due to the lepton selection selection efficiency systematic uncertainty for a Higgs boson signal with a mass of 130 GeV/c^2 is shown in Figure 7-1.

7.4.3 Lepton Momentum Scale and Resolution

The energy scale of the electrons and the momentum scale of the muons measured in this analysis have some systematic uncertainty which propagate as a shift in the dilepton mass and transverse mass observables, and affects the efficiency of the p_T cuts. Moreover, there's an uncertainty on the resolution of these energy and momentum measurements that may not be well modeled by the simulation. From studies



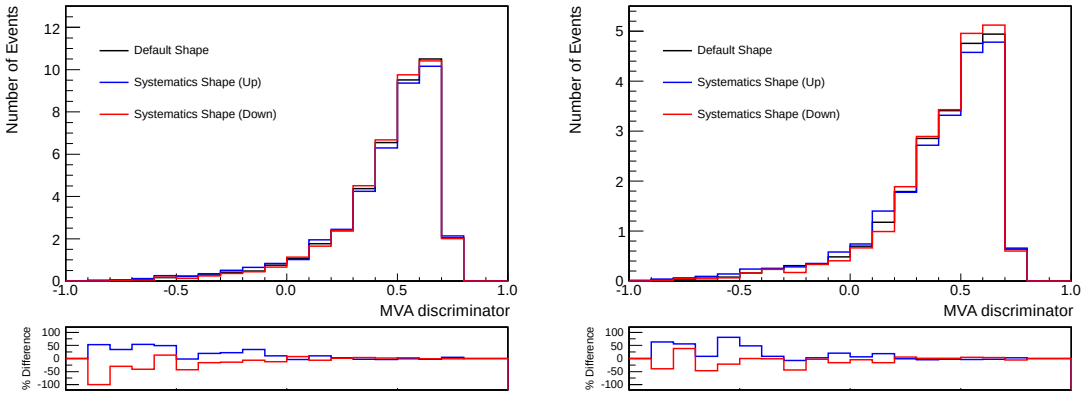
(a) Different-Flavor, 0-Jet (b) Same-Flavor, 0-Jet

Figure 7-1: The default and alternative shapes of the BDT discriminator distribution for the Higgs boson signal in the $m_{\text{Higgs}} = 130$ GeV/c^2 analysis is shown, separately for events with two different-flavor leptons and two same-flavor leptons in the zero-jet bin. The alternative shapes correspond to one standard deviation changes in the nuisance parameter that models the lepton efficiency uncertainty. The fractional differences between the alternative shapes and the default shape is shown at the bottom.

of the lineshape of Z boson decays to two electrons and two muons, we determined the uncertainty on the muon momentum scale to be 1% and the uncertainty on the electron energy scale to be 1% for electrons in the barrel, and 6% for electrons in the endcap. In addition, we measured an uncertainty on the resolution of 1% for muons, and 2% for electrons in the barrel and 6% for electrons in the endcap. The large uncertainty for electrons in the endcap is primarily due to instability and inaccuracy of the time dependant ECAL crystal transparency corrections accounting for changes in the light yield due to radiation damage of the crystal structure.

We propagate these systematic effects to the final result by systematically scaling the p_T of the leptons up and down by the corresponding scale uncertainty, and also smearing it with a Gaussian of width equal to the corresponding resolution uncertainty. For the cut-based analysis, these systematic effects change the signal and background yields by an average of 1.5% for the muon scale and resolution uncertainty, and by an average of 2% for the electron scale and resolution uncertainty. For the shape analysis, we use the linear morphing technique described above using the

alternative shapes obtained through this procedure as the upper and lower bounds. For the resolution, it is conceptually difficult to model a downward bias because it is difficult to perform an unsmearing. Therefore we model this effect instead by mirroring the BDT distribution obtained from the positively biased resolution about the nominal distribution. In Figure 7-2, we show an example of these alternative shapes due to the lepton energy and momentum scale and resolution systematic uncertainty for a Higgs boson signal with a mass of 130 GeV/c^2.



(a) Different-Flavor, 0-Jet

(b) Same-Flavor, 0-Jet

Figure 7-2: The default and alternative shapes of the BDT discriminator distribution for the Higgs boson signal in the $m_{\text{Higgs}} = 130$ GeV/c^2 analysis are shown, separately for events with two different-flavor leptons and two same-flavor leptons in the zero-jet bin. The alternative shapes correspond to one standard deviation changes in the nuisance parameter that models the lepton scale and resolution uncertainty. The fractional differences between the alternative shapes and the default shape is shown at the bottom.

7.4.4 Missing Transverse Energy Resolution

The resolution of the missing transverse energy is not be perfectly simulated by the Monte Carlo simulation. In this analysis, the missing transverse energy is used both as a selection variable and as an input to compute the BDT discriminator. As a result an uncertainty on the missing transverse energy will induce an uncertainty on the selection efficiency and on the shapes of the BDT distribution. For all signal and

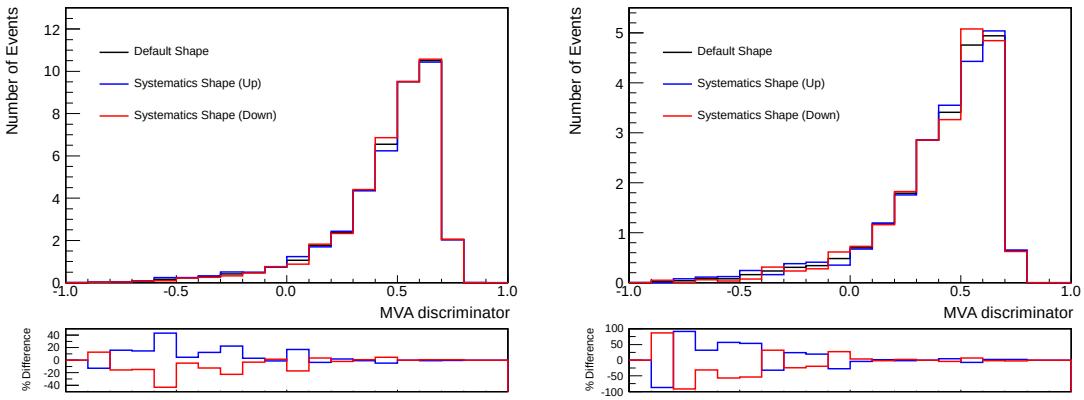
background predictions which use the Monte Carlo simulation, we must propagate the uncertainty due to this possible simulation inaccuracy.

Comparing the observation and the Monte Carlo prediction of the components of the missing transverse energy in the two transverse directions for $Z \rightarrow e^+e^-$ and $Z \rightarrow \mu^+\mu^-$ events, we determine that additional Gaussian smearing is required with standard deviations of 3.2 GeV/c^2, 3.6 GeV/c^2, and 4.3 GeV/c^2 for events with no counted jets, one counted jet, and two or more counted jets, respectively for the particle flow missing transverse energy. For the TrackMET, additional Gaussian smearing is required with widths of 2.1 GeV/c^2, 7.6 GeV/c^2, and 12.4 GeV/c^2 for events with no counted jets, one counted jet, and two or more counted jets, respectively. For all Monte Carlo events, we systematically perform the Gaussian smearing on these observables, and propagate its effect to the final result.

For the cut based analysis, the effect of this systematic uncertainty on the predicted yields depend on the particular process, but is generally about 2%. For the shape analysis, we obtain an alternative BDT discriminator shape through this Gaussian smearing procedure. The difference between this alternative shape and the default shape is mirrored to obtain the upper and lower bounds on the shape of the BDT distribution. The linear morphing technique is used to propagate this effect. The default and alternative shapes corresponding to a one standard deviation change in the nuisance parameter describing the systematic uncertainty due to the missing transverse energy resolution are shown in Figure 7-3 for the Higgs boson signal of mass 130 GeV/c^2.

7.4.5 Jet Energy Scale

The measurement of the energy of jets has associated uncertainties due to poossible systematic effects intrinsic in the jet energy scale calibration procedure. From dijet, photon+jet, and Z+jet balancing studies [74], we have determined that the uncertainties of the particle flow jet energy measurement to be between 3% at low values of pseudorapidity to 5% at very forward regions of the detector. As a conservative

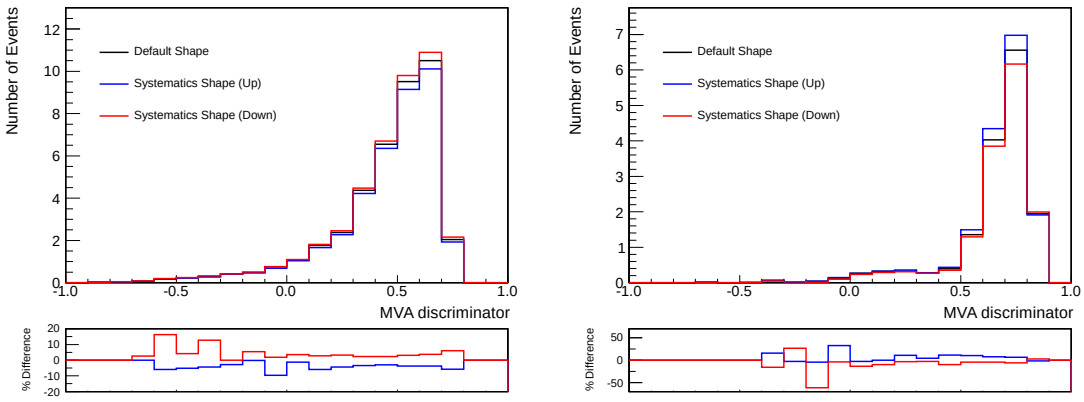


(a) Different-Flavor, 0-Jet (b) Same-Flavor, 0-Jet

Figure 7-3: The default and alternative shapes of the BDT discriminator distribution for the Higgs boson signal in the $m_{\text{Higgs}} = 130$ GeV/c^2 analysis are shown, separately for events with two different-flavor leptons and two same-flavor leptons in the zero-jet bin. The alternative shapes correspond to one standard deviation changes in the nuisance parameter that models the missing transverse energy resolution uncertainty. The fractional differences between the alternative shapes and the default shape is shown at the bottom.

measure, we propagate a 5% systematic uncertainty on the jet energy measurement. We systematically scale the energy of the jets up and down by 5%, and determine its effect on the signal and background predictions.

For the cut-based analysis, we find an average change in the yields of about 2%, 5%, 10%, for the zero-jet bin, the one-jet bin, and the VBF bin, respectively. For the shape analysis, we obtain the alternative BDT discriminator shapes as we scale the jet energies up and down. These shapes are used as the upper and lower bounds in the linear morphing procedure to propagate the effect of this systematic uncertainty. We show the default and alternative shapes corresponding to a one standard deviation change in the nuisance parameter describing the jet energy scale systematic uncertainty in Figure 7-4 for Higgs boson signal events in the zero-jet bin and the one jet bin for a Higgs boson mass 130 GeV/c^2. We also observe that the effects of a change in the jet energy scale for the zero-jet bin and one-jet bin are negatively correlated. This correlation is included when we propagate this systematic uncertainty to the final result.



(a) Different-Flavor, 0-Jet

(b) Different-Flavor, 1-Jet

Figure 7-4: The default and alternative shapes of the BDT discriminator distribution for the Higgs boson signal with two different-flavor leptons in the $m_{\text{Higgs}} = 130$ GeV/c^2 analysis are shown, separately for events with zero counted jets and one counted jet. The alternative shapes correspond to one standard deviation changes in the nuisance parameter that models the jet energy scale uncertainty. The fractional differences between the alternative shapes and the default shape is shown at the bottom.

7.4.6 Pileup

The distribution of the number of pileup interactions per bunch crossing shown in Section 2.1 is obtained from the per bunch crossing instantaneous luminosity measurement using rate measurements in the HF forward calorimeter. This estimate of the distribution of pileup interactions per bunch crossing has some intrinsic systematic uncertainty, both due to the explicit assumptions that are made and due to the uncertainty on the per bunch crossing luminosity measurement. To estimate the effect of this uncertainty on the pileup prediction, we shift the target pileup distribution by its uncertainty, equal to one pileup event, and study its effect on the predicted yields. We observe a change of less than 1%, and neglect this systematic uncertainty.

7.5 Theoretical Systematic Uncertainties

The theoretical systematic uncertainties are factorized into several components, which are approximately independent. We discuss, first, the systematic uncertainty on the

width of the Higgs boson decay at high mass. The second component is the uncertainty on the production cross section due to parton distribution function uncertainties. The third component is the uncertainty on the fraction of events categorized into the different jet bins due to the effect of migrations across jet bins. The last component is the uncertainty on the lepton acceptance and the selection efficiency of all other cuts. The effect of uncertainties on the parton distribution function and the value of α_s, and the effect of higher order corrections are all considered for the latter two components. For the jet categorization, we consider, additionally, the effect of the modeling of soft and collinear radiation as an estimate of the uncertainty due to missing higher order log terms.

7.5.1 Higgs Boson Width

The Monte Carlo generators that we use to predict the signal yield uses the "narrow-width approximation", which models the Higgs boson propagator as a Breit-Wigner distribution. This approximation does not model well the correct shape of the propagator at large Higgs boson masses, above 300 GeV. Following the prescription given in reference [86], we propagate a conservative systematic uncertainty of $150\% \times (m_H/1\text{ TeV})^3$ on the predicted signal yields. The cubic dependence is motivated by the naive assumption that the effect grows proportional to the natural width of the Higgs boson. This prescription results in a systematic uncertainty of 4% at a Higgs boson mass of 300 GeV and a systematic uncertainty of 32% at a mass of 600 GeV.

7.5.2 Parton Distribution Function Systematic Uncertainties

The parton distribution functions (PDF) parameterize the probability for a particular parton with a given longitudinal momentum fraction x and a given momentum transfer Q to be found in a colliding proton. The PDF's are obtained by performing fits to the deep inelastic scattering data from HERA, and various measurements of

the Drell-Yan and jet production processes at the Tevatron. There are a number of different PDF fits addressing various details in slightly different ways [95–100]. For each set of PDF fits, we obtain a nominal PDF as well as a set of systematic error PDF's which represents the effect of each independent systematic uncertainty in the input measurements entering the fit.

The PDF uncertainties are estimated using the procedure outlined in references [101, 102] using the PDF fit results from references [95–100]. The effect of each PDF in the set of systematic error PDF's is propagated to the particular observable, and the differences with respect to the prediction obtained using the nominal PDF are added in quadrature in a particular way. For the observable corresponding to the gluon fusion Higgs boson production cross section, the PDF uncertainty ranges between 7.8% at a Higgs boson mass of 110 GeV/c^2 to 10.5% at a Higgs boson mass of 600 GeV/c^2. The gluon fusion W^+W^- production cross section has an uncertainty of 4% due to PDF's. The vector boson fusion and associated Higgs boson production cross sections have PDF uncertainties of 5% [44], while the W^+W^-, WZ, and ZZ production cross sections have PDF uncertainties of 4% [88].

As a simplification, we approximate all parton distribution function (PDF) uncertainties into a set of two factorized and uncorrelated nuisance parameters. One addresses the uncertainty of the gluon PDF and is used to propagate uncertainties in the production cross section for processes that are predominantly produced via gluon fusion, and the other addresses the uncertainty due to the quark PDF's which are used to propagate the uncertainties in the production cross section for processes with two initial state quarks. All gluon fusion production processes are positively correlated with each other, and all quark initial state production processes are positively correlated with each other.

7.5.3 Systematic Uncertainties on the Jet Bin Fractions

We consider the effect of theoretical systematic uncertainties on the jet counting categorization without imposing any additional selection cuts on the signal [86, 103].

A jet in the event will be counted if it has $p_T > 30$ GeV/c. We define the jet bin fractions f_0, f_1, f_2 as the fraction of events produced with 0, 1, and 2 or more counted jets.

The predictions from any calculation performed to a fixed order in α_s has systematic uncertainties due to the missing higher order terms. This uncertainty is evaluated through the effect of a variation of the renormalization and factorization scales on the predicted jet bin fractions. To estimate the jet bin fractions for a given process, we combine the best calculation of the inclusive production cross section $\sigma_{n_j \geq 0}$, the production cross section for one or more jets $\sigma_{n_j \geq 1}$, and the production cross section for two or more jets $\sigma_{n_j \geq 2}$. The jet bin fractions are computed as:

$$f_0 = (\sigma_{n_j \geq 0} - \sigma_{n_j \geq 1})/\sigma_{n_j \geq 0} \tag{7.4}$$

$$f_1 = (\sigma_{n_j \geq 1} - \sigma_{n_j \geq 2})/\sigma_{n_j \geq 0} \tag{7.5}$$

$$f_2 = \sigma_{n_j \geq 2}/\sigma_{n_j \geq 0}. \tag{7.6}$$

$$\tag{7.7}$$

The systematic uncertainties due to missing higher order corrections in the calculations for $\sigma_{n_j \geq 0}$, $\sigma_{n_j \geq 1}$, and $\sigma_{n_j \geq 2}$, are estimated by varying the renormalization and factorization scales. These uncertainties are expressed as a relative fraction and interpreted as the κ-parameters of the log-normal probability density function used to parameterize them, as discussed in Section 7.1. We relate these uncertainties to the uncertainties on the exclusive cross sections in each jet bin, $\sigma_{n_j=0}$, $\sigma_{n_j=1}$, and $\sigma_{n_j=2}$.

From Equation 7.1, we see that the the log-normal random variable can be expressed in the form $X = X_0 \cdot \kappa^Z$, where Z is a Gaussian random variable. From here on, we make an implicit change of variables and use the value of the Gaussian random variable Z to parameterize the dependence of the inclusive cross sections. This parameter is used as a nuisance parameter and will be denoted by θ. In this way, we express the inclusive cross sections as $\sigma(\theta) = \bar{\sigma} \cdot \kappa^\theta$, where $\bar{\sigma}$ is the median and best estimate of the cross section, and κ is the parameter that describes the uncertainty

on the cross section. More explicitly, we have for each of the inclusive cross sections:

$$\sigma_{n_j\geq 0}(\theta_0) = \bar{\sigma}_{n_j\geq 0}\cdot(\kappa_{\geq 0})^{\theta_0} = \bar{\sigma}_{n_j\geq 0}\cdot\exp(\delta_{\geq 0}\cdot\theta_0) \sim \bar{\sigma}_{n_j\geq 0}\cdot(1+\delta_{\geq 0}\cdot\theta_0), \quad (7.8)$$

$$\sigma_{n_j\geq 1}(\theta_1) = \bar{\sigma}_{n_j\geq 1}\cdot(\kappa_{\geq 1})^{\theta_1} = \bar{\sigma}_{n_j\geq 1}\cdot\exp(\delta_{\geq 1}\cdot\theta_1) \sim \bar{\sigma}_{n_j\geq 1}\cdot(1+\delta_{\geq 1}\cdot\theta_1), \quad (7.9)$$

$$\sigma_{n_j\geq 2}(\theta_2) = \bar{\sigma}_{n_j\geq 2}\cdot(\kappa_{\geq 2})^{\theta_2} = \bar{\sigma}_{n_j\geq 2}\cdot\exp(\delta_{\geq 2}\cdot\theta_2) \sim \bar{\sigma}_{n_j\geq 2}\cdot(1+\delta_{\geq 2}\cdot\theta_2), \quad (7.10)$$

where the subscript i refers to the inclusive cross section with i jets or more, and $\delta_{\geq i}$ is the estimate of the fractional uncertainty for the corresponding inclusive cross section.

Taking the example of the exclusive zero-jet cross section, for typical uncertainties $\delta_{\geq i} \ll 1$ that are relatively small, we have [92]:

$$\begin{aligned}
\sigma_0(\theta_0,\theta_1) &= \sigma_{n_j\geq 0}(\theta_0) - \sigma_{n_j\geq 1}(\theta_1) \\
&= \bar{\sigma}_{n_j\geq 0}\cdot(1+\delta_{\geq 0}\cdot\theta_0) - \bar{\sigma}_{n_j\geq 1}\cdot(1+\delta_{\geq 1}\cdot\theta_1) + O(\theta^2) \\
&= (\bar{\sigma}_{n_j\geq 0} - \bar{\sigma}_{n_j\geq 1}) + \bar{\sigma}_{n_j\geq 0}\cdot\delta_{\geq 0}\cdot\theta_0 - \bar{\sigma}_{n_j\geq 1}\cdot\delta_{\geq 1}\cdot\theta_1 + O(\theta^2) \\
&= \bar{\sigma}_{n_j=0} + \bar{\sigma}_{n_j=0}\cdot\frac{1}{f_0}\cdot\delta_{\geq 0}\cdot\theta_0 - \bar{\sigma}_{n_j=0}\cdot\frac{f_1+f_2}{f_0}\cdot\delta_{\geq 1}\cdot\theta_1 + O(\theta^2) \\
&= \bar{\sigma}_{n_j=0}\cdot\left(1+\frac{1}{f_0}\cdot\delta_{\geq 0}\cdot\theta_0 - \frac{f_1+f_2}{f_0}\cdot\delta_{\geq 1}\cdot\theta_1\right) + O(\theta^2) \\
&= \bar{\sigma}_{n_j=0}\cdot\left(1+\frac{1}{f_0}\cdot\delta_{\geq 0}\cdot\theta_0\right)\cdot\left(1 - \frac{f_1+f_2}{f_0}\cdot\delta_{\geq 1}\cdot\theta_1\right) + O(\theta^2) \\
&= \bar{\sigma}_{n_j=0}\cdot\exp\left(\frac{1}{f_0}\cdot\delta_{\geq 0}\cdot\theta_0\right)\cdot\exp\left(-\frac{f_1+f_2}{f_0}\cdot\delta_{\geq 1}\cdot\theta_1\right) + O(\theta^2) \\
&= \bar{\sigma}_{n_j=0}\cdot\left[(e^{\delta_{\geq 0}})^{\frac{1}{f_0}}\right]^{\theta_0}\cdot\left[(e^{\delta_{\geq 1}})^{-\frac{f_1+f_2}{f_0}}\right]^{\theta_1} + O(\theta^2) \\
&= \bar{\sigma}_{n_j=0}\cdot\left[(\kappa_{\geq 0})^{\frac{1}{f_0}}\right]^{\theta_0}\cdot\left[(\kappa_{\geq 1})^{-\frac{f_1+f_2}{f_0}}\right]^{\theta_1} + O(\theta^2), \quad (7.11)
\end{aligned}$$

demonstrating that $\sigma_{n_j=0}$ can be written as the product of two log-normal random variables with κ parameters equal to $\kappa^0_{\geq 0} = (\kappa_{\geq 0})^{\frac{1}{f_0}}$, and $\kappa^0_{\geq 1} = (\kappa_{\geq 1})^{-\frac{f_1+f_2}{f_0}}$. Carrying out the analogous calculations for the exclusive one-jet and two or more jet cross sections, we obtain the relations between the κ parameters for the exclusive cross sections in terms of the κ parameters for the inclusive cross sections summarized in

Table 7.1. These relations are used to propagate the effect of systematic uncertainties on the inclusive cross section calculations to the prediction of the yields in each jet bin. They will be applied separately for the case of the gluon fusion Higgs boson production process and the W^+W^- production process in the next subsections.

Nuisance Parameter	κ's for $\sigma_{n_j=0}$	κ's for $\sigma_{n_j=1}$	κ's for $\sigma_{n_j=2}$
θ_0	$\kappa^0_{\geq 0} = (\kappa_{\geq 0})^{\frac{1}{f_0}}$	$\kappa^1_{\geq 0} = 1.0$	$\kappa^2_{\geq 0} = 1.0$
θ_1	$\kappa^0_{\geq 1} = (\kappa_{\geq 1})^{-\frac{f_1+f_2}{f_0}}$	$\kappa^1_{\geq 1} = (\kappa_{\geq 1})^{\frac{f_1+f_2}{f_1}}$	$\kappa^2_{\geq 1} = 1.0$
θ_2	$\kappa^0_{\geq 2} = 1.0$	$\kappa^1_{\geq 2} = (\kappa_{\geq 2})^{-\frac{f_2}{f_1}}$	$\kappa^2_{\geq 2} = \kappa_{\geq 2}$

Table 7.1: Table of formulas expressing the κ parameters for the effect of systematic uncertainties on the exclusive cross sections ($\kappa^i_{\geq j}$), in terms of the κ parameters for the inclusive cross sections ($\kappa_{\geq j}$).

Higgs Boson Production

For the Higgs boson produced via the gluon fusion process, the inclusive production cross section is computed at next-to-next-to-leading-order (NNLO) in α_s with the resummation technique used to regularize the singular behavior at low Higgs boson p_T at next-to-next-to-leading log (NNLL) order [44, 49, 104]. These cross sections are computed using the MSTW2008 NNLO parton distribution functions [98] and setting the factorization scale, and renormalization scale to $m_{\text{Higgs}}/2$. The systematic uncertainty on the inclusive gluon fusion Higgs boson production cross section due to missing higher order corrections is reported in the Higgs Cross Section Working Group report [44].

The production cross section for Higgs boson plus one or more jets, and Higgs boson plus two or more jets is computed using version 5.8 of MCFM [105] at next-to-leading-order. The systematic uncertainties $\kappa^{\text{ggH}}_{\geq 1}$, and $\kappa^{\text{ggH}}_{\geq 2}$ due to missing higher order corrections are estimated by performing the following four variations on the factorization scale μ_f and renormalization scale μ_r:

- $\mu_f = m_{\text{Higgs}}$, $\mu_r = m_{\text{Higgs}}$,

- $\mu_f = m_{\text{Higgs}}/4$, $\mu_r = m_{\text{Higgs}}/4$,
- $\mu_f = m_{\text{Higgs}}$, $\mu_r = m_{\text{Higgs}}/2$,
- $\mu_f = m_{\text{Higgs}}/2$, $\mu_r = m_{\text{Higgs}}$,

using MCFM and evaluating the largest positive and largest negative difference from the central value for $\sigma^{\text{ggH}}_{n_j \geq 1}$ and $\sigma^{\text{ggH}}_{n_j \geq 2}$. To interpret the uncertainty parameter consistently with the log-normal parameterization, we symmetrize the uncertainties via the formula:

$$\kappa_{\text{symmetrized}} = \sqrt{e^{\Delta_+} \times e^{\Delta_-}}, \tag{7.12}$$

where $\Delta_{+/-}$ are the relative positive and negative uncertainty respectively. Any asymmetric behavior in these uncertainties are approximately modeled by the log-normal distribution. The results for $\kappa^{\text{ggH}}_{\geq 0}$, $\kappa^{\text{ggH}}_{\geq 1}$, and $\kappa^{\text{ggH}}_{\geq 2}$ are summarized in Table 7.2.

Using the expressions given in Table 7.1, we compute the κ parameters for the exclusive jet bin cross sections from the κ parameters for the inclusive cross sections given in Table 7.2. There are a total of 9 κ parameters, five of which are non-trivial. The values of the non-trivial κ parameters are summarized in Table 7.3 for all Higgs boson masses considered. When one nuisance parameter affects multiple exclusive cross sections, such as the ones parametrized by $\kappa^0_{\geq 1}$ and $\kappa^1_{\geq 1}$, the direction of their correlation is indicated by whether each κ parameter are greater or less than 1. This convention is driven by the use of the log-normal functional form for the propagation of the systematic uncertainties. The effect of the nuisance parameter on the predicted yields is essentially a multiplicate factor. For κ values greater than one, the predicted yield will increase with a corresponding increase of the nuisance parameter, while the predicted yield will decrease if the κ value is less than one. Thus, if both κ values are greater than one or both are less than one, then they are positively correlated. If the

Higgs Boson Mass	$\kappa_{\geq 0}^{\mathrm{ggH}}$	$\kappa_{\geq 1}^{\mathrm{ggH}}$	$\kappa_{\geq 2}^{\mathrm{ggH}}$
115	1.106	1.226	1.149
120	1.104	1.224	1.120
130	1.100	1.230	1.117
140	1.096	1.220	1.129
150	1.095	1.220	1.124
160	1.095	1.221	1.199
170	1.090	1.222	1.175
180	1.089	1.218	1.171
190	1.087	1.217	1.171
200	1.087	1.213	1.197
210	1.085	1.212	1.172
220	1.085	1.210	1.170
230	1.084	1.209	1.197
250	1.083	1.208	1.230
300	1.082	1.208	1.205
350	1.090	1.207	1.209
400	1.075	1.195	1.195
450	1.078	1.194	1.196
500	1.087	1.188	1.174
550	1.089	1.191	1.194
600	1.090	1.187	1.192

Table 7.2: The κ parameters describing the systematic uncertainties due to missing higher order corrections for the inclusive gluon fusion Higgs boson production cross sections are shown. The columns represent the κ-parameters for the fully inclusive case, the case with one or more jets, and the case of two or more jets.

first κ value is greater than 1 and the second is less than 1, then the predicted yields for the first jet bin will increase and the predicted yields for the second jet bin will decrease.

Higgs Boson Mass	$\kappa_{\geq 0}^{\text{ggH,0Jet}}$	$\kappa_{\geq 1}^{\text{ggH,0Jet}}$	$\kappa_{\geq 1}^{\text{ggH,1Jet}}$	$\kappa_{\geq 2}^{\text{ggH,1Jet}}$	$\kappa_{\geq 2}^{\text{ggH,2Jet}}$
115	1.16	0.92	1.28	0.97	1.15
120	1.16	0.92	1.28	0.97	1.12
130	1.15	0.91	1.29	0.98	1.12
140	1.15	0.91	1.28	0.97	1.13
150	1.15	0.90	1.28	0.97	1.12
160	1.15	0.90	1.28	0.96	1.20
170	1.15	0.89	1.28	0.96	1.18
180	1.15	0.89	1.28	0.96	1.17
190	1.15	0.88	1.28	0.96	1.17
200	1.15	0.88	1.27	0.96	1.20
210	1.15	0.88	1.27	0.96	1.17
220	1.15	0.87	1.27	0.96	1.17
230	1.16	0.87	1.27	0.95	1.20
250	1.16	0.86	1.27	0.96	1.17
300	1.17	0.84	1.27	0.95	1.20
350	1.20	0.83	1.27	0.95	1.21
400	1.17	0.82	1.26	0.95	1.20
450	1.19	0.81	1.26	0.95	1.20
500	1.22	0.80	1.25	0.95	1.17
550	1.24	0.78	1.26	0.95	1.19
600	1.25	0.78	1.26	0.94	1.19

Table 7.3: Table of κ parameters modelling the systematic uncertainties for the exclusive jet bin Higgs boson production cross sections due to missing higher order corrections in the inclusive cross section calculations. The subscripts indicate whether the source of each systematic uncertainty comes from the calculation of the fully inclusive cross section ($\sigma_{\geq 0}$), the cross section with one or more jets($\sigma_{\geq 1}$), or the cross section with two or more jets ($\sigma_{\geq 2}$). The superscripts refer to the particular exclusive jet bin that the given systematic uncertainty affects.

There are certain regions of phase space for which the Higgs boson production differential cross sections encounter singular behavior for calculations in fixed orders of α_s. In these regions, typically low Higgs boson p_T and low jet p_T, the calculation no longer gives a physically meaningful result. In the Monte Carlo simulation, this singular behavior is regularized by the effect of parton showers which model soft

and collinear emission at leading-log (LL) order. At low values of jet p_T, the higher order log terms may become large and non-negligible. This effect is particularly important for the exclusive zero-jet cross section where the jet p_T threshold is set at 30 GeV/c, a value of p_T that is approaching the failure region of the fixed order calculation [106, 107].

To estimate the systematic uncertainty due to missing higher order log terms, we compare the exclusive jet bin fractions predicted by two different Monte Carlo generators implementing alternative parton shower and hadronization models. These alternative models for the parton shower and hadronization implement different methods to predict the leading-log term and we use this difference to quantify the size of the effects of higher order log terms. The default prediction is provided by the POWHEG Monte Carlo generator [50], interfaced to Pythia for the parton shower and hadronization [108], and the alternative prediction is provided by the MC@NLO generator [90] interfaced to the HERWIG parton shower and hadronization model [109]. In order to isolate the effect of the parton shower and hadronization model, in both cases we reweight the Higgs boson p_T distribution to the reference prediction given by the NNLO+NNLL resummed calculation, mentioned above.

The κ parameters that describe the systematic uncertainty on the exclusive jet bin cross sections due to missing higher order log terms, as estimated by the difference in the predictions of the two different parton shower models, $\kappa^0_{\rm PS}$, $\kappa^1_{\rm PS}$, $\kappa^2_{\rm PS}$, are summarized in Table 7.4.

W^+W^- Production

As described in Section 6.4, we estimate the W^+W^- background by normalizing the Monte Carlo prediction to the high dilepton-mass control region, where the contamination from the Higgs boson signal is very small for low Higgs boson masses. However, for Higgs boson mass hypotheses of 200 GeV/c^2 and larger, there is significant signal contamination in the W^+W^- control region, rendering the data driven method inapplicable. For such cases, we use the prediction for the W^+W^- background yields

Higgs Boson Mass	κ^0_{PS}	κ^1_{PS}	κ^2_{PS}
115	0.941	1.128	1.212
120	0.940	1.110	1.293
130	0.937	1.113	1.237
140	0.941	1.104	1.168
150	0.942	1.093	1.156
160	0.943	1.084	1.138
170	0.946	1.075	1.108
180	0.947	1.067	1.092
190	0.948	1.068	1.083
200	0.952	1.055	1.059
210	0.948	1.061	1.042
220	0.950	1.061	1.028
230	0.950	1.061	1.024
250	0.955	1.058	0.990
300	0.958	1.061	0.942
350	0.964	1.068	0.889
400	0.966	1.078	0.856
450	0.954	1.092	0.864
500	0.946	1.102	0.868
550	0.931	1.117	0.861
600	0.920	1.121	0.872

Table 7.4: Table of κ parameters that model the systematic uncertainty on the exclusive jet bin cross sections due to missing higher order log terms in the calculation, as a function of the assumed Higgs boson mass.

from the Monte Carlo simulation and theoretical cross sections. In this case the theoretical systematic uncertainties for the cross section and the jet bin fractions must be propagated.

The jet bin migration systematic uncertainty is estimated analogously to the estimate for Higgs boson production. The procedure and results are described in greater detail in reference [86]. We first estimate the uncertainty due to missing higher order corrections on the total inclusive W^+W^- production cross section $\sigma_{n_j\geq 0}^{W^+W^-}$, the inclusive W^+W^- plus one or more jets cross section $\sigma_{n_j\geq 1}^{W^+W^-}$, and the inclusive W^+W^- plus two or more jets cross section $\sigma_{n_j\geq 2}^{W^+W^-}$. This is estimated using MC@NLO [90] varying the renormalization and factorization scales. The uncertainty on $\sigma_{n_j\geq 2}^{W^+W^-}$ is estimated as the difference in the prediction from MC@NLO, which provides a leading log approximation, and the prediction from ALPGEN [110], which provides a leading order calculation. From these calculations we obtain systematic uncertainties on $\sigma_{n_j\geq 0}^{W^+W^-}$, $\sigma_{n_j\geq 1}^{W^+W^-}$, and $\sigma_{n_j\geq 2}^{W^+W^-}$ of 3%, 6%, and 42% respectively. The large systematic uncertainty for the two-jet bin reflects the fact that the calculations are leading order in α_s.

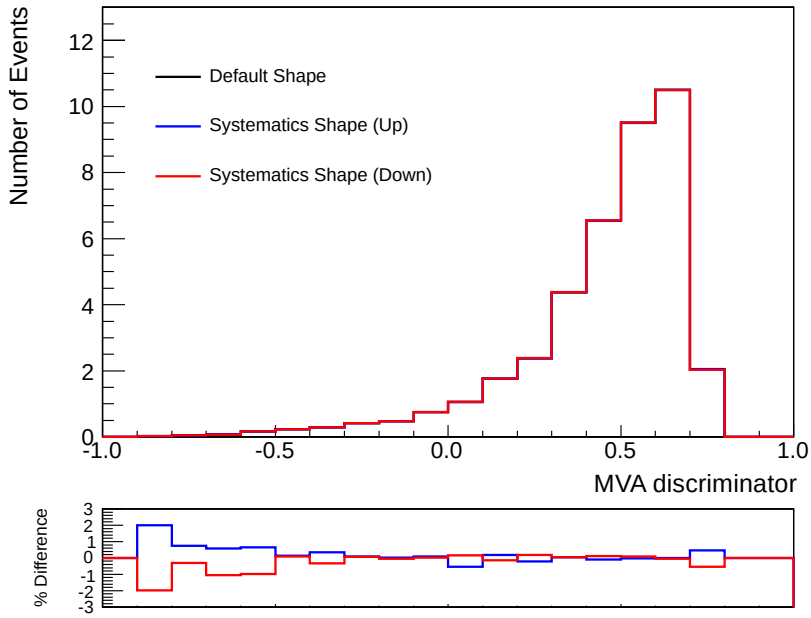
These calculations give the following values of jet bin fractions: $f_0^{W^+W^-} = 0.70$, $f_1^{W^+W^-} = 0.22$, and $f_2^{W^+W^-} = 0.08$. Then, using the expressions from Table 7.1 we compute the relevant κ parameters for the exclusive W^+W^- production cross sections:

- $\kappa_{\geq 0}^{W^+W^-,0-\text{jet}} = 1.042$,
- $\kappa_{\geq 1}^{W^+W^-,0-\text{jet}} = 0.978$,
- $\kappa_{\geq 1}^{W^+W^-,1-\text{jet}} = 1.076$,
- $\kappa_{\geq 2}^{W^+W^-,1-\text{jet}} = 0.914$,
- $\kappa_{\geq 2}^{W^+W^-,2-\text{jet}} = 1.420$.

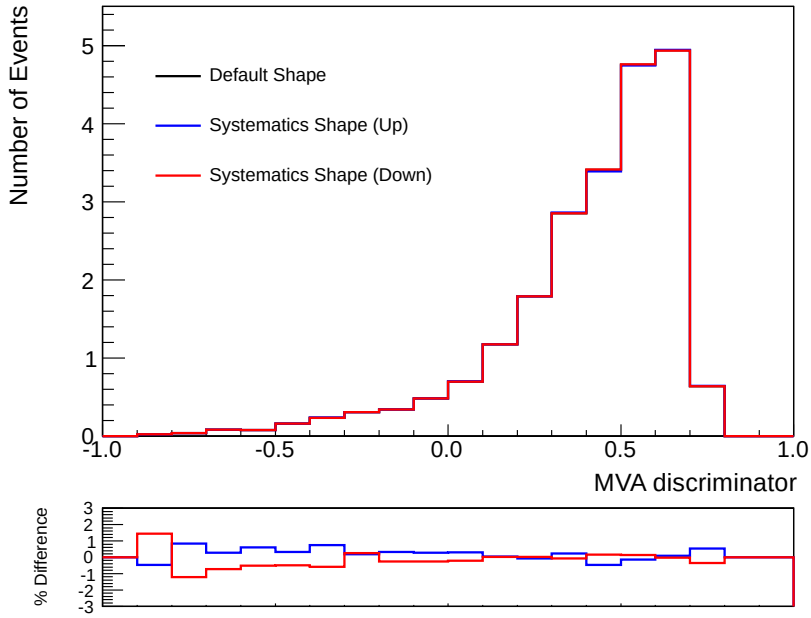
7.5.4 Systematic Uncertainties on the Selection Efficiency and Kinematic Observables

The effect of missing higher order corrections on the selection efficiency of the cut-based analysis for the Higgs boson signal is estimated by comparing the difference in the efficiency computed from the POWHEG NLO Monte Carlo generator by varying the renormalization and factorization scales by factors of 1/2 and 2 from the default scale. We find differences in the efficiency of about 2%, and is propagated for the cut based analysis.

For the BDT shape analysis, we model the effect of missing higher order corrections by reweighting the Higgs boson p_T spectrum to the one obtained from the NNLO+NNLL calculation with the renormalization and factorization scales varied by factors of 2 and 1/2. It is assumed that any changes in kinematic observables are primarily induced through a change of the Higgs boson p_T spectrum. This is a reasonable assumption because all of the inputs to the BDT discriminator are computed from the kinematics of the leptons, and are expected to be not significantly affected by the behavior of the jets. The alternative systematic shapes for the BDT discriminator obtained using this procedure is shown in Figure 7-5 for Higgs boson signal with a hypothesized mass of 130 GeV/c^2 in the zero-jet bin, separately for events with two different-flavor leptons and for events with two same-flavor leptons. The systematic uncertainty is very small due to the fact that the NNLO calculation is fairly precise in predicting the distributions of the kinematic observables. These are propagated using the linear morphing technique described above.



(a) Different-Flavor, 0-Jet



(b) Same-Flavor, 0-Jet

Figure 7-5: The default and alternative shapes of the BDT discriminator distribution for the Higgs boson signal in the $m_{\text{Higgs}} = 130$ GeV/c^2 analysis are shown, separately for events with two different-flavor leptons and two same-flavor leptons in the zero-jet bin. The alternative shapes correspond to one standard deviation changes in the nuisance parameter that models the uncertainty due to missing higher order α_s terms in the NNLO theoretical calculation. The fractional differences between the alternative shapes and the default shape is shown at the bottom.

Chapter 8

Statistical Analysis and Results

Statistical analysis of the observed yields and distributions allow us to transform the observed yields and distributions in the Higgs boson signal region into well defined quantities from which an inference can be made on the existence of the Higgs boson. An observation of an excess of events, alone, is insufficient to make a conclusion on whether the Higgs boson exists or not. The reason is that particle production is generally a quantum mechanical process and therefore it is subject to randomness. Random statistical fluctuations of the background may introduce excesses over the expected background yields simply by chance. Similarly, observing no excess may also be a result of random statistical fluctuations of the signal and background.

The main purpose of performing statistical analysis is to obtain a quantitative measure of the likelihood that a given observation is a result of the hypothesis that the Higgs boson does not exist, referred to as the "background-only hypothesis", and the likelihood that it is the result of the hypothesis that the Higgs boson does exist, refereed to as the "signal+background hypothesis". Inferences on the existing of the Higgs boson may, then, be drawn based on a comparison of these probabilities.

Statistical analyses and the tools used to perform them are well defined mathematical constructions and procedures; they produce quantitative answers to any well defined question that is asked. The key challenge for a physicist is to ask the most appropriate statistical question to answer the particular physical question of interest.

In the context of this analysis, the physical question of interest is to establish the existence or non-existence of the Standard Model Higgs boson. Providing a definitive answer on this question is in fact one of the main goals of the LHC. To this end, the statistical question that we ask is to establish and quantify an observation of Higgs boson production in LHC collision in the W^+W^- decay channel, or in the case of a non-observation, to measure the upper limit to its production cross section.

8.1 Likelihood Construction and Statistical Procedure

To obtain a quantitative measure on the existence or non-existence of hypothesized new particles, a number of statistical procedures have been conventionally used in the past [111–113], each with its own set of advantages and disadvantages. An effort was made by the CMS and ATLAS collaborations to arrive at a common set of statistical procedures in order to facilitate the comparison of results. In this analysis we present results using two of the approaches that were agreed upon. One approach is Bayesian and one is frequentist in the interpretation of its numerical result. These statistical procedures are both described in detail in the 2011 LHC Higgs Combination Group report [92].

The fundamental difference between the Bayesian and the frequentist approaches is in the interpretation of the probability. The frequentist interpretation views the probability as the average frequency of occurrence in the limit of large number of repetitions of the experiment. The Bayesian interpretation views the probability as a degree of belief that the true value of a particular parameter that one measures is a given value. For this analysis, we compute 95% confidence level upper limits for the Higgs boson production cross section. The frequentist approach to the upper limit calculation is to determine the size of the signal cross section for which 5% or less of hypothetical experiments under the signal+background hypothesis would observe no signature indicative of the Higgs boson. The hypothetical results are obtained

by means of throwing pseudoexperiments using the expected signal and background yields. In the Bayesian approach, we determine the size of the signal cross section for which our degree of belief that the true signal cross section is larger is less than 5%.

For both approaches we begin by constructing a likelihood function, which quantifies the probability to observe a particular set of results for a given model. In this search, an independent analysis is performed for each mass hypothesis. For a given mass, there are no free parameters of the theory whose value has not been experimentally measured. Thus, in the statistical procedure, the only parameter that is left unconstrained is the Higgs boson production cross section, the parameter that we intend to measure or to obtain upper bounds on. For ease of interpretation, we transform this parameter into the "signal strength" μ, defined to be the ratio of measured Higgs boson production cross section to the Higgs boson production cross section predicted by the Standard Model.

Thus, the likelihood that we construct is simply a product of Poisson probabilities:

$$\begin{aligned} \mathcal{L}(\text{data}|\mu) &= \text{P}(\text{data}|\mu s + b) \\ &= \prod_i \frac{(\mu_i s_i + b_i)^{n_i}}{n_i!} e^{-(\mu s_i + b_i)} \end{aligned} \tag{8.1}$$

where "data" represents the vector $\{n_i\}$ of the observed number of events in each bin i, and the symbol $\mu s + b$ represents the vector $\{\mu s_i + b_i\}$ of the predicted number of signal and background events for a given signal strength μ in each bin i. It is important to note that the background prediction for each bin i, is in fact the sum of the predicted number of events for each background process j in that bin: $b_i = \sum_j (b_i^j)$. However, we will suppress this sum over background processes in our notation, for brevity.

8.1.1 Systematic Uncertainties

The predicted signal and background yields, s and b, are potentially affected by a number of systematic uncertainties. These effects are parametrized by a set of nuisance parameters $\{\theta_k\}$, which we denote collective by θ. Thus, the predicted yields become functions of the nuisance parameters: $s(\theta)$, and $b(\theta)$.

Sources of systematic uncertainties are assumed to be either fully correlated, or uncorrelated. Partially correlated uncertainties are factorized into uncertainty sources which are either independent or approximated to be independent. Using auxiliary measurements, or external information, we obtain a best estimate of the value of the nuisance parameter $\tilde{\theta}$, and an estimate of its uncertainty. These pieces of information allow us to build the probability density function for the nuisance parameter.

Depending on the statistical approach, we interpret this probability density function in two different ways. In the Bayesian approach, we parametrize our knowledge on θ as $\rho(\theta|\tilde{\theta})$, which describes our degree of belief in what the true value of θ is, given that our best estimate of the nuisance parameter from the auxiliary measurement or the external information is $\tilde{\theta}$. One, subsequently, integrates over the true value of the nuisance parameter given the information encoded in ρ.

On the other hand, for the frequentist approach we use the Bayes' theorem to interpret ρ in a particular way. The Bayes' theorem states:

$$\rho(\theta|\tilde{\theta}) \sim p(\tilde{\theta}|\theta) \cdot \pi(\theta), \tag{8.2}$$

where $\rho(\theta|\tilde{\theta})$ is called the posterior probability, $p(\tilde{\theta}|\theta)$ is the likelihood, and $\pi(\theta)$ is the prior probability. One views the posterior probability as the updated information on the value of the parameter θ that one has after having obtained the result of a particular measurement of θ described by the likelihood. The prior probability encodes all of the information that we have on the value of the parameter θ before making the particular measurement. Using this relation, we can interpret the probability density function for the nuisance parameter $\rho(\theta|\tilde{\theta})$ as the posterior probability resulting from

the particular auxiliary measurement or external information, characterized by the likelihood $p(\tilde{\theta}|\theta)$.

Then the relation given in Equation 8.2 allows us to transform this ρ parametrization into a prediction, $p(\tilde{\theta}|\theta)$, giving us the probability that a measurement of the nuisance parameter will yield the value $\tilde{\theta}$ given that the true value is θ, if we also know the prior π. We encode all our prior knowledge of the nuisance parameters into ρ. Therefore, the pre-measurement prior $\pi(\theta)$ is taken to be a uniform distribution of θ. The uniform distribution is the typical prior distribution that one uses to model the situation where no prior knowledge about θ is assumed. Using these predictive measurement probabilities $p(\tilde{\theta}|\theta)$ and the best estimate $\tilde{\theta}$ from the external measurements as constraints, we construct likelihoods:

$$\mathcal{L}(\text{data}|\mu,\theta) \;=\; \text{P}(\text{data}|\mu s(\theta)+b(\theta))\cdot p(\tilde{\theta}|\theta). \tag{8.3}$$

These likelihoods can then be maximized as a function of the nuisance parameter θ, providing a better measurement of the nuisance parameter using the information in the signal region, and leads to a better search sensitivity.

8.1.2 Bayesian Upper Limits

In the Bayesian approach, we make use of the Bayes' Theorem to construct a posterior probability $L(\mu)$ of the signal strength μ using the observed data and integrating over all values of the nuisance parameters θ characterizing all systematic uncertainties:

$$L(\mu) \;=\; \frac{1}{C}\int_{\theta} p(\text{data}|\mu s+b)\rho(\theta|\tilde{\theta})\pi(\mu)d\theta. \tag{8.4}$$

The probability density function $\rho(\theta|\tilde{\theta})$ parametrizes our knowledge about the value of the nuisance parameters θ given our best estimates $\tilde{\theta}$), and the function $\pi(\mu)$ is the prior probability for the signal strength μ which is taken to be flat for $\mu \geq 0$

and zero otherwise. The constant C is simply a normalization factor for the posterior probability $L(\mu)$.

From this posterior probability we compute the one-sided 95% confidence level upper limit on μ to be the value μ_{95} such that:

$$\int_0^{\mu_{95}} L(\mu)d\mu = 0.95. \tag{8.5}$$

8.1.3 Frequentist Upper Limits

In the frequentist approach, we begin by defining some test statistic q, typically chosen to be a likelihood ratio but can be in principle any quantity that is sensitive to the presence of the signal. Pseudo-data generated from throwing toy Monte Carlo experiments are used to build the probability density function for the test statistic q in the background-only hypothesis and the signal+background hypothesis, and upper limits are derived from these probability density function.

In our case, the test statistic is defined to be a particular modification of the the profile likelihood ratio [114], adapted to models where the signal strength cannot be negative:

$$\tilde{q}_\mu = \begin{cases} -2\ln\frac{\mathcal{L}(\text{data}|\mu,\hat{\theta}_\mu)}{\mathcal{L}(\text{data}|0,\hat{\theta}_0)} & \hat{\mu} < 0 \\ -2\ln\frac{\mathcal{L}(\text{data}|\mu,\hat{\theta}_\mu)}{\mathcal{L}(\text{data}|\hat{\mu},\hat{\theta})} & 0 \leq \hat{\mu} \leq \mu \\ 0 & \hat{\mu} > \mu, \end{cases} \tag{8.6}$$

where $\hat{\theta}_\mu$ represent the values of the nuisance parameters for which the likelihood is maximized with the given observed data and value of the signal strength fixed to μ, and $\hat{\mu}$ and $\hat{\theta}$ are the values of the signal strength and the nuisance parameters, respectively, for which the likelihood is globally maximized with the observed data. To be explicit, the hat symbol in $\hat{\theta}$ is used to indicate that the value is obtained by maximizing the likelihood, while the tilde symbol in $\tilde{\theta}$ as used in Equation 8.2

indicates that the value is the best estimate according to the auxiliary measurement or external information. The behavior of this test statistic for $\hat{\mu} < 0$ protects against unintuitive upper limit results in the case of a downward fluctuation of background. Constraining $\hat{\mu}$ to be less than μ forces the limits that we obtain to be always one-sided, with the consequence of over-coverage in the case of an excess.

First a single fit is performed to determine the best fit values of the signal strength $\hat{\mu}$ and the values of the nuisance parameters $\hat{\theta}$. Then for the numerator, one fit is performed for every fixed value of μ to determine the values of the nuisance parameters $\hat{\theta}_\mu$ that maximizes the likelihood. In the case where the best fit signal strength is negative, the likelihood in the denominator is evaluated for an assumed signal strength of 0. For the case where the best fit signal strength is larger than the signal strength under consideration, the test statistic is set to 0 to ensure that the test statistic does not consider an excess of events in addition to that introduced by the specific signal hypothesis that we are testing as evidence against the signal+background hypothesis.

Then, for the given set of observed data, we compute $\tilde{q}_\mu^{\text{obs}}$ for each hypothesized signal strength μ, and the values of the nuisance parameters that maximize the likelihood $\hat{\theta}_\mu^{\text{obs}}$. Next, we generate toy Monte Carlo pseudo-data, first under the background-only hypothesis using the best fit values of the nuisance parameters assuming no signal, $\hat{\theta}_0^{\text{obs}}$, and then under the signal+background hypothesis with signal strength μ using the best fit values of the nuisance parameters corresponding to that assumed signal strength, $\hat{\theta}_\mu^{\text{obs}}$. In the traditional frequentist approach, the nuisance parameters are fixed when generating pseudoexperiments, but are varied according to its probability density function when they are used to evaluate the likelihood. It has been demonstrated that our procedure, which allows the nuisance parameter to float according to its probability density function in a maximum likelihood fit, yields the equivalent coverage as the traditional frequentist procedure [115].

In each such pseudo-experiment we can evaluate the test statistic $\tilde{q_{\mu'}}$ using the procedure described above. At this point, we explicitly distinguish μ from μ'. The value of the signal strength hypothesis that we test for is denoted by μ and is the signal

strength used to generate the pseudoexperiments. The value of the signal strength that enters the definition of the test statistic is denoted by μ' and is the value used when computing the test statistic from Equation 8.6. From the test statistic $\tilde{q_{\mu'}}$, we can numerically construct its probability density function $f(\tilde{q}_{\mu'}|\mu, \hat{\theta}_{\mu}^{\rm obs})$ under the signal strength equal to μ hypothesis and using the corresponding best fit values of the nuisance parameters $\hat{\theta}_{\mu}^{\rm obs}$.

Although μ and μ' are generally independent quantities, from here on, for the signal hypothesis we will consider only the test statistic for which $\mu' = \mu$ because this choice gives the most sensitive test statistic for a given hypothesized signal strength μ. We also construct the probability density function $f(\tilde{q}_{\mu}|0, \hat{\theta}_{0}^{\rm obs})$ for each test statistic $\tilde{q}_{\mu}$ generated from pseudo-data under the background-only hypothesis and using the corresponding background-only best fit values of the nuisance parameters $\hat{\theta}_{0}^{\rm obs}$.

We define two p-values, p_{μ} and p_b, for the signal+background hypothesis and the background-only hypothesis respectively:

$$p_{\mu}(\mu) = P(\tilde{q}_{\mu} \geq \tilde{q}_{\mu}^{\rm obs}|\text{signal strength } \mu, \hat{\theta}_{\mu}^{\rm obs}) = \int_{\tilde{q}_{\mu}^{\rm obs}}^{\infty} f(\tilde{q}_{\mu}|\mu, \hat{\theta}_{\mu}^{\rm obs}) d\tilde{q}_{\mu}, \quad (8.7)$$

$$1 - p_b(\mu) = P(\tilde{q}_{\mu} \geq \tilde{q}_{\mu}^{\rm obs}|\text{background only}, \hat{\theta}_{0}^{\rm obs}) = \int_{\tilde{q}_{\mu}^{\rm obs}}^{\infty} f(\tilde{q}_{\mu}|0, \hat{\theta}_{0}^{\rm obs}) d\tilde{q}_{\mu}. \quad (8.8)$$

The p-value is the probability of obtaining a test statistic value that is at least as "extreme" as the observed value, where "extreme" is defined by the relative likelihood for obtaining the test statistic value under the given hypothesis. Finally we define $\mathrm{CL}_s(\mu)$ as the ratio of these probabilities:

$$\mathrm{CL_s}(\mu) = \frac{p_{\mu}(\mu)}{1 - p_b(\mu)}. \quad (8.9)$$

Then the confidence level for the exclusion of a signal of signal strength μ is equal to $1 - \mathrm{CL_s}(\mu)$. This estimate of the confidence level is a slight modification to the classical frequentist confidence level $\mathrm{CL_{s+b}}$ computed from the classical likelihood ratio, where an attempt is made to protect the upper limit result against possibly

unintuitive inferences that are drawn in the case of downward fluctuations of the background. Finally to compute the upper limit on the signal strength at the 95% confidence level, μ_{95}, we find the value of μ such that $\mathrm{CL_s}(\mu) = 0.05$.

8.1.4 Asymptotic Approximation of the Frequentist Approach

In the regime of a large number of events, it follows from Wilks theorem [116] that the distribution of the test statistic approaches a particular analytic form [117]. Without the constraint of positive signal strength, the asymptotic analytic form for the probability density function of the profile-likelihood-ratio test statistic is equal to one half times a χ^2 distribution with one degree of freedom. With the positive signal strength constraint, the asymptotic analytic expression becomes:

$$
\begin{aligned}
f(\tilde{q}_\mu|\mu') &= \Phi\left(\frac{\mu'-\mu}{\sigma}\right)\delta(\tilde{q}_\mu) \\
&+ \begin{cases} \frac{1}{2}\frac{1}{\sqrt{2\pi\tilde{q}_\mu}}\exp\left[-\frac{1}{2}\left(\sqrt{\tilde{q}_\mu}-\frac{\mu-\mu'}{\sigma}\right)^2\right] & 0 < \tilde{q}_\mu \leq \mu^2/\sigma^2 \\ \frac{1}{\sqrt{2\pi}(2\mu/\sigma)}\exp\left[-\frac{1}{2}\frac{(\tilde{q}_\mu-(\mu^2-2\mu\mu')/\sigma^2)^2}{(2\mu/\sigma)^2}\right] & \tilde{q}_\mu > \mu^2/\sigma^2, \end{cases}
\end{aligned}
\tag{8.10}
$$

where Φ is the cumulative distribution of the standard Gaussian random variable with a mean of zero and a variance of one, δ is the Dirac delta function, and $\sigma^2 = \mu^2/q_{\mu,\mathrm{asimov}}$. The value $q_{\mu,\mathrm{asimov}}$ is the test statistic evaluated on the Asimov dataset [117], defined as the dataset with the expected signal and background yields and the nominal values of the nuisance parameters. Setting the value of μ' to 0 for the background-only hypothesis, and to μ for the signal hypothesis yields expressions for the probability density functions needed for compute the relevant p-values, and the $\mathrm{CL_s}$ quantity. From those expressions we derive expressions for the median expected limit, and the 1-standard-deviation and 2-standard-deviation expected limit bands. Using these analytic expressions allows us to compute the upper limits very quickly with small computational cost. The 1-standard-deviation and 2-standard-deviation bands calculated using the asymptotic approximation are known to give overly opti-

mistic results for situations with small number of events. Because the signal regions in this analysis largely satisfy the large N approximation the asymptotic approximation is expected to be fairly accurate, and has been verified to be accurate in a number of representative cases.

8.1.5 Expected Limits

The expected limits are computed by generating toy Monte Carlo pseudo-data with the background-only hypothesis and computing the upper limit for each pseudo-experiment using the procedure described in Sections 8.1.2 and 8.1.3. From the limits obtained in each pseudo-experiment we construct a distribution of upper limits. We integrate this distribution to obtain the cumulative probability. The median expected limit is the value of μ_{95} for which the cumulative probability crosses the 50% line. The upper and lower bounds defining the 1-standard-deviation band is obtained by computing the values of μ_{95} for which the cumulative probability of the test statistic distribution reaches 16% and 84%. Similarly the 2-standard-deviation band is defined by the values of μ_{95} for which the cumulative probability of the test statistic distribution reaches 2.5% and 97.5%. In the case of the asymptotic approximation, the median expected limit as well as the 1-standard-deviation and 2-standard-deviation bands are computed analytically without generating pseudo-experiments.

8.2 Signal Region Yields

In this section, we summarize the observations in our selected signal regions, which enter as input to the statistical analysis. For the cut-based analysis, the signal region is separated into five different categories with different background composition and different signal-to-background ratio. We first categorize events into the zero-jet bin, one-jet bin, and VBF bin, depending on the number of counted jets observed in the event. Then for the zero-jet and one-jet bins, we additionally separate events into the same-flavor category and the different-flavor category depending on whether the

event has two leptons of the same flavor or of different flavors, respectively.

The different-flavor zero-jet bin category has the best signal-to-background ratio and the best sensitivity. The expected and observed yields in the signal regions corresponding to each Higgs boson mass hypothesis for this bin is summarized in Table 8.2. The same-flavor zero-jet bin has, in addition, Drell-Yan background. Due to the large uncertainty on this background, the sensitivity of this category is significantly less than the sensitivity of the corresponding different-flavor category. However, it has less W+jets background relative to the different-flavor category due to tighter kinematic cuts. The yields in the zero-jet same-flavor category are summarized in Table 8.2.

The one-jet bins collectively have much larger top backgrounds than the zero-jet bin, and makes top production one of the leading backgrounds. The expected and observed yields are summarized in Table 8.2 for events with two different-flavor leptons and in Table 8.2 for events with two same-flavor leptons. As in the zero-jet bin, the Drell-Yan background is more important in the same-flavor category, while the W+jets background is more important in the different-flavor category.

Finally, the expected and observed yields for the VBF category are summarized in Table 8.2.

8.3 Higgs Boson Production Cross Section Limits

Using the procedures described in Section 8.1 we derive the 95% confidence level upper limits on the Higgs boson production cross section in the $H \rightarrow W^+W^- \rightarrow 2l\ 2\nu$ decay channel. The observed limits, the median expected limit, and the 1-standard-deviation and 2-standard-deviation bands are shown in Figures 8-1 and 8-2 as a function of the Higgs boson mass for the frequentist asymptotic approximation and the Bayesian calculation, respectively. The exact upper limit values are given in numerical form in Tables A.1 and A.2 of the appendix. All observed limits are within one standard deviation of the expectation under the background-only hypothesis.

The upper limit results for the BDT shape analysis are shown in Figures 8-3 and 8-

Mass	Signal	W^+W^-	Top	W+jets	Other	$\sum$Bkg	Data
110	2.8 ± 0.6	41.3 ± 5.0	2.7 ± 0.9	4.8 ± 1.9	5.5 ± 2.1	54.3 ± 5.8	60
115	5.9 ± 1.3	41.3 ± 5.0	2.7 ± 0.9	4.8 ± 1.9	5.5 ± 2.1	54.3 ± 5.8	60
118	8.1 ± 1.8	51.9 ± 6.2	3.5 ± 1.1	5.9 ± 2.3	5.7 ± 2.1	67.0 ± 7.1	74
120	11.9 ± 2.6	61.7 ± 7.4	4.0 ± 1.2	6.2 ± 2.4	5.9 ± 2.1	77.9 ± 8.1	88
122	13.0 ± 2.8	66.3 ± 7.9	4.4 ± 1.3	7.0 ± 2.7	6.0 ± 2.1	83.7 ± 8.7	92
124	16.4 ± 3.5	71.2 ± 8.5	4.7 ± 1.3	7.8 ± 3.0	6.2 ± 2.1	89.9 ± 9.4	99
126	19.6 ± 4.2	75.5 ± 9.0	4.9 ± 1.4	7.8 ± 3.0	6.3 ± 2.1	94.5 ± 9.8	105
128	23.6 ± 5.1	80.0 ± 9.5	5.2 ± 1.4	8.1 ± 3.1	6.4 ± 2.1	99.7 ± 10.3	111
130	32.1 ± 6.8	84.0 ± 10.0	5.9 ± 1.6	8.2 ± 3.1	6.5 ± 2.1	104.5 ± 10.8	114
135	37.8 ± 8.0	87.1 ± 10.3	6.2 ± 1.7	7.4 ± 2.8	4.6 ± 1.5	105.3 ± 10.9	113
140	47.8 ± 10.2	85.2 ± 10.1	6.1 ± 1.6	6.0 ± 2.3	4.4 ± 1.5	101.7 ± 10.5	108
150	51.3 ± 11.2	64.3 ± 7.7	5.4 ± 1.5	3.0 ± 1.2	2.2 ± 0.9	74.9 ± 8.0	81
160	76.2 ± 16.6	45.1 ± 5.4	4.7 ± 1.3	1.7 ± 0.8	1.9 ± 0.9	53.5 ± 5.7	57
170	60.8 ± 13.5	35.3 ± 4.2	4.3 ± 1.3	1.1 ± 0.6	1.5 ± 0.9	42.3 ± 4.5	47
180	43.0 ± 9.5	40.4 ± 4.8	5.2 ± 1.4	0.6 ± 0.4	1.7 ± 0.9	47.8 ± 5.1	56
190	36.7 ± 8.3	63.7 ± 7.5	8.8 ± 2.2	1.3 ± 0.6	1.4 ± 0.3	75.3 ± 7.9	74
200	28.2 ± 6.3	65.1 ± 6.7	10.4 ± 2.6	1.5 ± 0.7	1.6 ± 0.3	78.6 ± 7.2	88
250	16.0 ± 3.9	66.2 ± 6.8	17.0 ± 4.0	2.3 ± 1.0	2.6 ± 0.7	88.1 ± 8.0	94
300	12.0 ± 3.1	53.9 ± 5.6	18.0 ± 4.2	3.0 ± 1.2	2.7 ± 0.7	77.7 ± 7.1	72
350	12.5 ± 3.7	43.6 ± 4.5	17.4 ± 4.1	2.8 ± 1.1	2.6 ± 0.7	66.4 ± 6.2	67
400	10.0 ± 2.9	34.7 ± 3.6	16.0 ± 3.8	2.4 ± 1.0	2.3 ± 0.6	55.4 ± 5.4	54
450	5.7 ± 1.8	20.5 ± 2.2	10.8 ± 2.6	1.6 ± 0.7	1.7 ± 0.5	34.6 ± 3.5	32
500	3.6 ± 1.3	15.9 ± 1.7	7.8 ± 1.9	1.3 ± 0.6	1.6 ± 0.6	26.6 ± 2.7	21
550	2.3 ± 1.0	12.5 ± 1.4	6.4 ± 1.6	1.1 ± 0.5	1.0 ± 0.4	21.1 ± 2.2	18
600	1.4 ± 0.7	10.2 ± 1.1	5.1 ± 1.3	0.8 ± 0.4	0.8 ± 0.3	16.9 ± 1.8	16

Table 8.1: A summary of expected signal and background yields in the selected signal regions of the cut-based analysis for events with two different-flavor leptons and zero counted jets is shown. The rows represent the different signal regions corresponding to each Higgs boson mass hypothesis. The uncertainties include statistical and systematic components.

Mass	Signal	W^+W^-	Drell-Yan	Top	Other	$\sum$Bkg	Data
110	0.8 ± 0.2	22.1 ± 2.7	6.9 ± 4.6	0.6 ± 0.2	2.5 ± 0.8	32.2 ± 5.4	31
115	2.0 ± 0.4	21.8 ± 2.7	7.3 ± 4.8	0.6 ± 0.2	2.5 ± 0.8	32.2 ± 5.5	31
118	2.9 ± 0.6	28.1 ± 3.4	8.6 ± 5.6	0.9 ± 0.3	3.2 ± 0.9	40.7 ± 6.6	40
120	5.0 ± 1.1	35.0 ± 4.2	9.5 ± 5.7	1.2 ± 0.4	3.5 ± 1.0	49.3 ± 7.2	51
122	5.5 ± 1.2	38.9 ± 4.7	10.8 ± 5.9	1.4 ± 0.4	3.6 ± 1.0	54.6 ± 7.6	57
124	7.4 ± 1.6	42.4 ± 5.1	12.4 ± 6.9	1.5 ± 0.4	4.0 ± 1.0	60.3 ± 8.7	65
126	8.9 ± 1.9	46.4 ± 5.6	12.0 ± 8.4	1.6 ± 0.4	4.0 ± 1.0	64.1 ± 10.1	72
128	12.5 ± 2.7	50.4 ± 6.0	11.3 ± 6.1	1.8 ± 0.5	4.3 ± 1.1	67.9 ± 8.7	76
130	16.4 ± 3.5	54.3 ± 6.5	12.5 ± 6.8	2.0 ± 0.5	4.5 ± 1.2	73.4 ± 9.5	86
135	21.2 ± 4.5	59.2 ± 7.0	13.4 ± 7.1	2.4 ± 0.6	5.0 ± 1.3	80.1 ± 10.1	93
140	29.3 ± 6.2	62.8 ± 7.5	13.3 ± 7.0	2.5 ± 0.7	5.1 ± 1.3	83.8 ± 10.3	94
150	39.1 ± 8.6	52.3 ± 6.3	9.1 ± 9.5	3.9 ± 1.1	2.6 ± 0.7	67.9 ± 11.5	83
160	56.7 ± 12.4	36.7 ± 4.4	3.5 ± 3.1	3.1 ± 0.9	1.0 ± 0.4	44.4 ± 5.5	53
170	48.6 ± 10.7	30.4 ± 3.6	3.4 ± 2.9	3.3 ± 0.9	0.9 ± 0.3	38.0 ± 4.8	44
180	38.8 ± 8.6	36.8 ± 4.4	2.5 ± 2.9	4.3 ± 1.2	1.3 ± 0.5	44.8 ± 5.4	44
190	33.1 ± 7.5	55.5 ± 6.6	6.8 ± 4.3	5.9 ± 1.5	1.7 ± 0.5	70.0 ± 8.0	73
200	23.3 ± 5.2	50.7 ± 5.2	4.2 ± 4.3	6.5 ± 1.6	2.1 ± 0.7	63.4 ± 7.0	72
250	8.9 ± 2.2	41.8 ± 4.3	4.0 ± 0.7	8.5 ± 2.0	3.0 ± 0.9	57.2 ± 4.9	63
300	9.1 ± 2.4	38.0 ± 4.0	4.4 ± 2.5	11.4 ± 2.8	5.6 ± 1.4	59.4 ± 5.6	65
350	9.4 ± 2.8	34.3 ± 3.6	3.7 ± 0.5	12.4 ± 3.0	9.9 ± 2.6	60.3 ± 5.4	58
400	8.1 ± 2.4	27.0 ± 2.8	2.8 ± 0.3	10.5 ± 2.5	9.1 ± 2.5	49.5 ± 4.6	44
450	4.4 ± 1.4	15.2 ± 1.6	1.8 ± 0.2	6.3 ± 1.5	4.1 ± 1.2	27.4 ± 2.5	29
500	2.7 ± 1.0	11.0 ± 1.2	1.5 ± 0.2	5.1 ± 1.3	3.4 ± 1.0	21.0 ± 2.0	22
550	1.6 ± 0.7	7.9 ± 0.9	1.2 ± 0.2	3.8 ± 0.9	2.7 ± 0.8	15.6 ± 1.6	19
600	1.0 ± 0.5	6.1 ± 0.7	0.8 ± 0.1	2.9 ± 0.7	1.8 ± 0.6	11.7 ± 1.2	12

Table 8.2: A summary of expected signal and background yields in the selected signal regions of the cut-based analysis for events with two same-flavor leptons and zero counted jets is shown. The rows represent the different signal regions corresponding to each Higgs boson mass hypothesis. The uncertainties include statistical and systematic components.

Mass	Signal	W^+W^-	Top	W+jets	Other	$\sum$Bkg	Data
110	1.3 ± 0.4	13.5 ± 2.2	8.2 ± 0.9	3.3 ± 1.4	2.4 ± 0.5	27.4 ± 2.8	35
115	2.4 ± 0.7	13.5 ± 2.2	8.2 ± 0.9	3.3 ± 1.4	2.4 ± 0.5	27.4 ± 2.8	35
118	3.6 ± 1.0	16.6 ± 2.7	9.8 ± 1.0	4.0 ± 1.6	2.6 ± 0.5	32.9 ± 3.4	41
120	5.6 ± 1.6	19.3 ± 3.1	11.3 ± 1.1	4.4 ± 1.8	2.7 ± 0.5	37.7 ± 3.8	42
122	6.1 ± 1.7	20.8 ± 3.4	12.0 ± 1.1	4.4 ± 1.8	2.8 ± 0.5	40.1 ± 4.0	44
124	7.4 ± 2.0	22.0 ± 3.6	13.3 ± 1.2	4.6 ± 1.8	3.0 ± 0.5	42.8 ± 4.2	46
126	8.8 ± 2.5	23.4 ± 3.8	14.5 ± 1.3	4.6 ± 1.9	3.1 ± 0.5	45.6 ± 4.5	51
128	10.6 ± 3.0	24.9 ± 4.0	15.7 ± 1.4	5.1 ± 2.0	3.4 ± 0.6	49.1 ± 4.8	53
130	13.6 ± 3.9	26.6 ± 4.3	16.3 ± 1.4	5.4 ± 2.2	3.5 ± 0.6	51.8 ± 5.0	60
135	16.3 ± 4.6	27.8 ± 4.5	17.1 ± 1.5	4.6 ± 1.8	3.5 ± 0.6	53.0 ± 5.1	61
140	21.1 ± 5.8	27.6 ± 4.4	16.9 ± 1.5	4.0 ± 1.6	3.3 ± 0.6	51.8 ± 5.0	57
150	26.3 ± 7.3	26.1 ± 4.2	17.2 ± 1.5	1.6 ± 0.8	2.0 ± 0.4	46.9 ± 4.6	47
160	40.0 ± 11.2	21.7 ± 3.5	15.8 ± 1.4	1.4 ± 0.7	1.1 ± 0.1	40.1 ± 3.9	32
170	30.7 ± 8.5	16.9 ± 2.7	12.9 ± 1.2	0.8 ± 0.4	0.6 ± 0.1	31.2 ± 3.0	20
180	23.0 ± 6.3	19.3 ± 3.1	15.8 ± 1.4	0.8 ± 0.4	0.6 ± 0.1	36.4 ± 3.4	25
190	21.9 ± 6.0	31.9 ± 5.1	27.7 ± 2.2	1.0 ± 0.6	1.1 ± 0.1	61.7 ± 5.6	39
200	16.3 ± 4.1	31.5 ± 4.7	33.6 ± 2.6	1.3 ± 0.7	1.3 ± 0.2	67.8 ± 5.4	48
250	10.4 ± 2.7	41.1 ± 6.1	49.6 ± 3.5	2.2 ± 1.0	2.2 ± 0.3	95.1 ± 7.1	77
300	8.7 ± 2.4	38.6 ± 5.8	46.6 ± 3.2	3.1 ± 1.3	2.0 ± 0.4	90.4 ± 6.8	77
350	8.9 ± 2.6	34.8 ± 5.3	39.7 ± 2.8	3.0 ± 1.2	2.0 ± 0.5	79.5 ± 6.1	67
400	7.5 ± 2.2	29.7 ± 4.5	31.3 ± 2.2	3.2 ± 1.3	1.7 ± 0.4	65.9 ± 5.2	57
450	4.6 ± 1.4	19.0 ± 2.9	18.5 ± 1.4	2.1 ± 0.9	1.4 ± 0.4	41.0 ± 3.4	35
500	3.1 ± 1.0	15.5 ± 2.4	14.4 ± 1.1	1.7 ± 0.8	1.2 ± 0.4	32.8 ± 2.8	30
550	2.1 ± 0.8	12.6 ± 2.0	11.3 ± 0.9	1.5 ± 0.7	1.1 ± 0.4	26.6 ± 2.3	25
600	1.4 ± 0.6	10.5 ± 1.6	8.9 ± 0.8	1.3 ± 0.6	1.0 ± 0.4	21.7 ± 2.0	21

Table 8.3: A summary of expected signal and background yields in the selected signal regions of the cut-based analysis for events with two different-flavor leptons and one counted jet is shown. The rows represent the different signal regions corresponding to each Higgs boson mass hypothesis. The uncertainties include statistical and systematic components.

Mass	Signal	W^+W^-	Drell-Yan	Top	Other	$\sum$Bkg	Data
110	0.3 ± 0.1	5.8 ± 1.0	4.9 ± 2.7	2.5 ± 0.4	0.7 ± 0.3	13.9 ± 2.9	11
115	0.7 ± 0.2	5.8 ± 1.0	4.9 ± 2.7	2.5 ± 0.4	0.7 ± 0.3	13.9 ± 2.9	11
118	1.0 ± 0.3	7.2 ± 1.2	6.4 ± 3.1	3.4 ± 0.4	0.8 ± 0.3	17.7 ± 3.4	18
120	1.7 ± 0.5	9.1 ± 1.5	7.9 ± 3.6	4.6 ± 0.6	0.8 ± 0.3	22.5 ± 4.0	25
122	2.2 ± 0.6	10.2 ± 1.7	7.7 ± 3.2	5.1 ± 0.6	1.2 ± 0.4	24.1 ± 3.6	30
124	2.6 ± 0.7	11.0 ± 1.8	7.8 ± 3.3	5.3 ± 0.6	1.2 ± 0.4	25.3 ± 3.9	32
126	3.2 ± 0.9	11.9 ± 2.0	8.2 ± 3.7	5.8 ± 0.6	1.4 ± 0.4	27.3 ± 4.2	35
128	4.2 ± 1.2	13.1 ± 2.2	7.8 ± 3.6	6.3 ± 0.6	1.4 ± 0.4	28.7 ± 4.3	39
130	5.8 ± 1.7	14.2 ± 2.3	8.2 ± 3.7	6.8 ± 0.7	1.8 ± 0.6	31.1 ± 4.5	40
135	7.6 ± 2.1	16.1 ± 2.6	9.8 ± 4.4	7.5 ± 0.7	2.6 ± 0.8	35.9 ± 5.2	44
140	11.2 ± 3.2	17.4 ± 2.8	10.3 ± 4.6	8.2 ± 0.7	2.9 ± 0.9	38.9 ± 5.5	45
150	16.4 ± 4.6	17.6 ± 2.9	7.4 ± 4.2	11.5 ± 1.2	2.4 ± 0.8	38.8 ± 5.3	49
160	26.3 ± 7.4	15.2 ± 2.5	5.8 ± 3.2	10.5 ± 1.1	1.9 ± 0.7	33.4 ± 4.2	48
170	21.2 ± 5.9	13.1 ± 2.1	6.2 ± 3.6	9.6 ± 1.0	1.5 ± 0.7	30.4 ± 4.4	43
180	18.3 ± 5.1	16.1 ± 2.6	6.5 ± 3.6	13.0 ± 1.3	1.5 ± 0.6	37.1 ± 4.7	38
190	15.7 ± 4.3	24.5 ± 4.0	11.9 ± 4.7	21.0 ± 1.8	2.2 ± 0.8	59.6 ± 6.4	51
200	11.2 ± 2.9	21.5 ± 3.2	12.1 ± 4.4	21.5 ± 1.8	2.2 ± 0.8	57.3 ± 5.8	58
250	5.3 ± 1.4	23.4 ± 3.5	8.4 ± 3.4	26.2 ± 2.0	1.9 ± 0.7	60.0 ± 5.3	68
300	5.6 ± 1.5	21.9 ± 3.3	9.6 ± 4.3	26.7 ± 2.0	3.1 ± 0.8	61.3 ± 5.8	73
350	5.9 ± 1.7	20.0 ± 3.1	8.2 ± 2.4	24.9 ± 1.9	3.8 ± 1.0	56.9 ± 4.4	67
400	5.3 ± 1.6	16.7 ± 2.6	7.5 ± 2.3	21.0 ± 1.6	3.7 ± 1.0	48.9 ± 4.0	59
450	3.2 ± 1.0	10.9 ± 1.7	5.7 ± 2.0	12.4 ± 1.1	2.3 ± 0.7	31.4 ± 2.9	38
500	2.1 ± 0.7	8.9 ± 1.4	4.5 ± 1.8	9.7 ± 0.9	2.0 ± 0.7	25.1 ± 2.5	31
550	1.4 ± 0.5	7.2 ± 1.1	3.3 ± 1.5	7.3 ± 0.7	1.8 ± 0.6	19.5 ± 2.1	23
600	0.9 ± 0.4	5.5 ± 0.9	0.6 ± 0.1	5.2 ± 0.5	1.5 ± 0.5	12.8 ± 1.2	18

Table 8.4: A summary of expected signal and background yields in the selected signal regions of the cut-based analysis for events with two same-flavor leptons and one counted jet is shown. The rows represent the different signal regions corresponding to each Higgs boson mass hypothesis. The uncertainties include statistical and systematic components.

Mass	Signal	W^+W^-	Drell-Yan	Top	Other	$\sum$Bkg	Data
110	0.3 ± 0.0	1.7 ± 0.3	2.3 ± 1.8	2.2 ± 1.4	1.4 ± 0.5	7.5 ± 2.3	4
115	0.7 ± 0.1	1.7 ± 0.3	2.3 ± 1.8	2.2 ± 1.4	1.4 ± 0.5	7.5 ± 2.3	4
118	0.9 ± 0.1	1.7 ± 0.3	2.3 ± 1.8	2.2 ± 1.4	1.4 ± 0.5	7.5 ± 2.3	4
120	1.3 ± 0.2	1.7 ± 0.3	2.3 ± 1.7	2.2 ± 1.4	1.4 ± 0.5	7.5 ± 2.3	4
122	1.6 ± 0.2	1.7 ± 0.3	2.3 ± 1.8	2.2 ± 1.4	1.4 ± 0.5	7.5 ± 2.3	4
124	1.8 ± 0.2	1.7 ± 0.3	2.3 ± 1.8	2.2 ± 1.4	1.4 ± 0.5	7.5 ± 2.4	4
126	2.1 ± 0.3	2.0 ± 0.4	3.4 ± 2.3	2.5 ± 1.6	1.5 ± 0.5	9.5 ± 2.9	6
128	2.6 ± 0.3	2.0 ± 0.4	3.4 ± 2.4	2.5 ± 1.6	1.5 ± 0.5	9.5 ± 3.0	6
130	3.0 ± 0.4	2.0 ± 0.4	3.4 ± 2.5	2.5 ± 1.6	1.5 ± 0.5	9.5 ± 3.0	6
135	4.3 ± 0.5	2.2 ± 0.4	3.4 ± 2.5	2.8 ± 1.8	1.8 ± 0.6	10.2 ± 3.1	7
140	5.8 ± 0.7	2.2 ± 0.4	3.4 ± 2.5	2.8 ± 1.8	1.8 ± 0.6	10.2 ± 3.1	7
150	9.1 ± 1.1	2.4 ± 0.5	3.4 ± 2.2	3.4 ± 2.2	1.8 ± 0.6	11.1 ± 3.2	8
160	12.8 ± 1.6	2.4 ± 0.5	3.4 ± 2.3	3.4 ± 2.2	1.8 ± 0.6	11.1 ± 3.3	8
170	15.4 ± 1.9	2.4 ± 0.5	3.4 ± 2.4	3.4 ± 2.2	1.8 ± 0.6	11.1 ± 3.4	8
180	13.8 ± 1.7	2.4 ± 0.5	3.4 ± 2.2	3.4 ± 2.2	1.8 ± 0.6	11.1 ± 3.2	8
190	10.4 ± 1.3	2.4 ± 0.5	3.4 ± 2.4	3.4 ± 2.2	1.8 ± 0.6	11.1 ± 3.3	8
200	8.4 ± 1.0	2.0 ± 0.8	3.4 ± 2.2	3.4 ± 2.2	1.8 ± 0.6	10.7 ± 3.3	8
250	6.0 ± 0.7	3.5 ± 1.4	4.7 ± 3.7	5.8 ± 3.8	1.9 ± 0.6	15.9 ± 5.5	14
300	4.5 ± 0.6	3.5 ± 1.4	4.7 ± 2.6	6.0 ± 3.9	1.9 ± 0.6	16.2 ± 5.0	15
350	3.6 ± 0.5	3.6 ± 1.4	4.7 ± 2.4	6.1 ± 4.0	1.9 ± 0.6	16.3 ± 4.9	15
400	2.6 ± 0.4	3.6 ± 1.4	4.7 ± 2.4	6.2 ± 4.0	1.9 ± 0.6	16.4 ± 4.9	15
450	2.0 ± 0.3	3.6 ± 1.4	4.7 ± 2.4	6.2 ± 4.0	1.9 ± 0.6	16.4 ± 4.9	15
500	1.4 ± 0.3	3.6 ± 1.4	4.7 ± 2.4	6.2 ± 4.0	1.9 ± 0.6	16.4 ± 4.9	15
550	1.1 ± 0.3	3.6 ± 1.4	4.7 ± 2.4	6.2 ± 4.0	1.9 ± 0.6	16.5 ± 5.0	15
600	0.9 ± 0.3	3.7 ± 1.4	4.7 ± 2.4	6.2 ± 4.0	1.9 ± 0.6	16.5 ± 5.0	15

Table 8.5: A summary of expected signal and background yields in the selected signal regions of the cut-based analysis for events passing the VBF selection. The rows represent the different signal regions corresponding to each Higgs boson mass hypothesis. The uncertainties include statistical and systematic components.

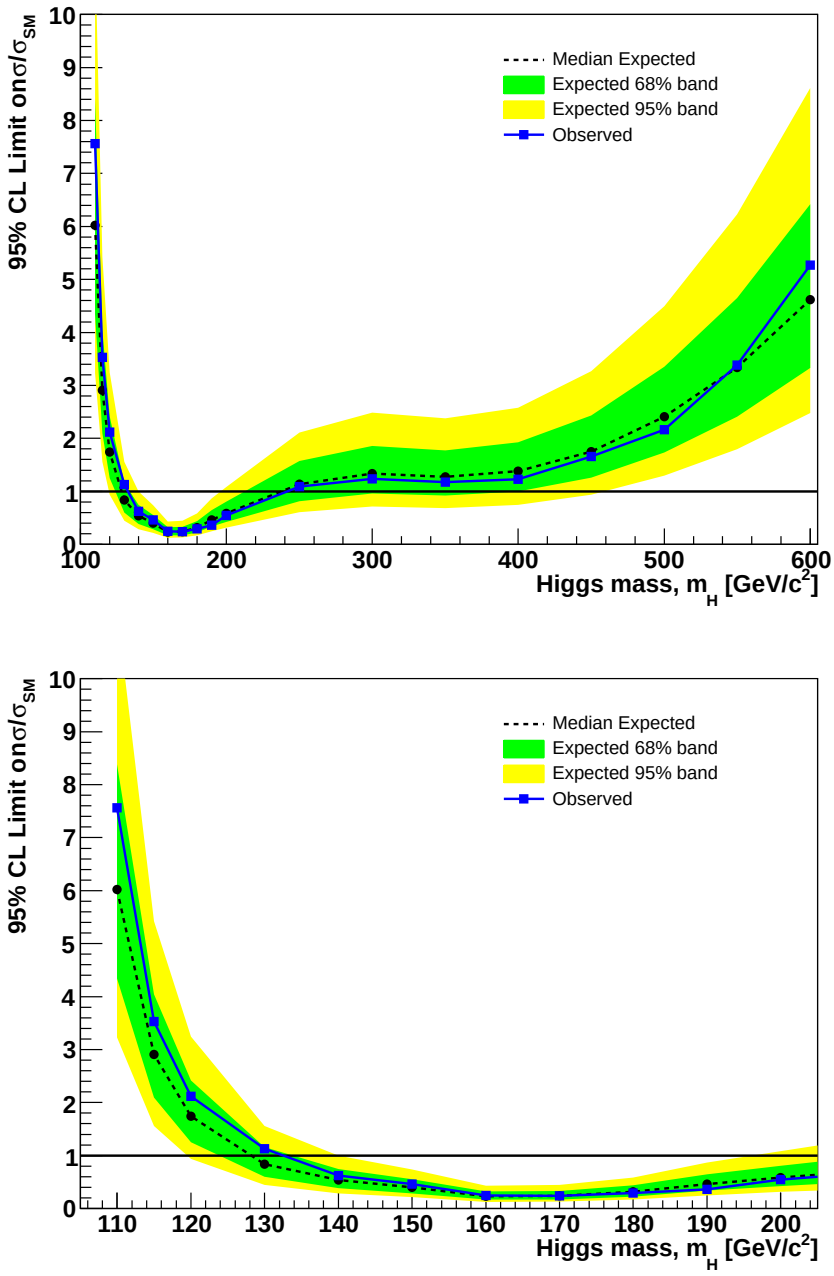


Figure 8-1: Expected and observed upper limits on the Standard Model Higgs boson production cross section are shown as a function of the hypothesized Higgs boson mass for the cut-based analysis using the frequentist asymptotic approximation. In the first plot, we show the upper limits for the full range of Higgs boson masses that were tested, from 110 GeV/c^2 to 600 GeV/c^2. In the second plot, we zoom in to the low mass region between 110 GeV/c^2 and 200 GeV/c^2.

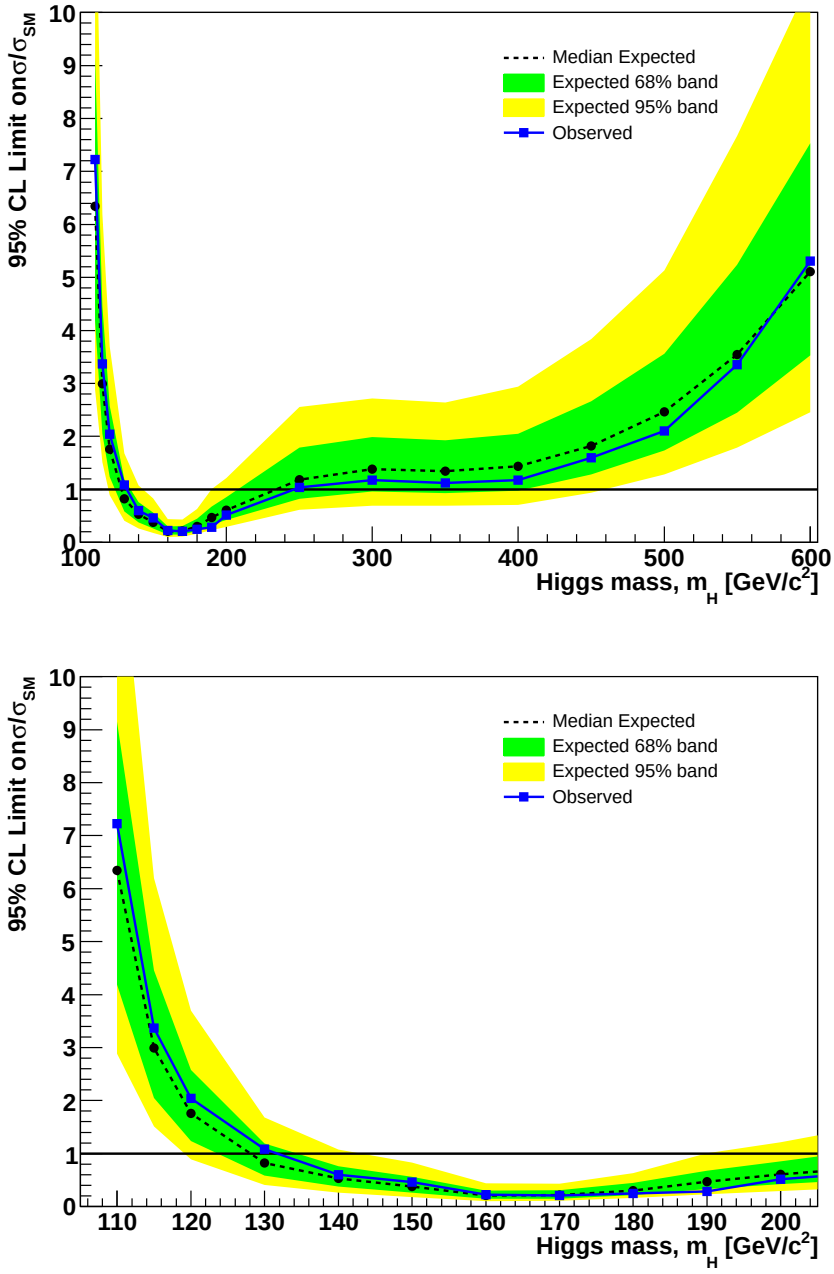


Figure 8-2: Expected and observed upper limits on the Standard Model Higgs boson production cross section are shown as a function of the hypothesized Higgs boson mass for the cut-based analysis using the Bayesian approach. In the first plot, we show the upper limits for the full range of Higgs boson masses that were tested, from 110 GeV/c^2 to 600 GeV/c^2. In the second plot, we zoom in to the low mass region between 110 GeV/c^2 and 200 GeV/c^2.

4 for the frequentist and Bayesian calculations, respectively. They are also given in numerical form in Tables A.3 and A.4 of the appendix. We observe that the expected upper limits using the shape analysis are on average about 20% better, with increasing difference at higher masses. At higher Higgs boson masses, the kinematic differences between the Higgs boson signal and the W^+W^- background become weaker, and therefore additional sensitivity is gained by the use of the BDT, which makes more optimal use of correlations. Side-by-side comparisons between the upper limits obtained using for cut-based analysis and those obtained using the BDT shape analysis are shown in Tables A.5 and A.6 of the appendix for the frequentist and Bayesian calculations, respectively. On average, we expect to exclude the Standard Model predicted Higgs boson production cross section for Higgs boson masses between 126 GeV/c^2 and 260 GeV/c^2, under the background-only hypothesis.

We observe no statistically significant deviations from the background-only expectation in the data. All observed upper limits for the BDT shape analysis fall within a window of two standard deviations around the expected upper limit under the background-only hypothesis. The largest excess observed lies in the low mass region between 115 GeV/c^2 and 130 GeV/c^2, and is lies between one and two standard deviation away from the expectation. While there are conceptual differences between the frequentist and the Bayesian approaches, we observe from Figures 8-3 and 8-4 that the excluded range is, in fact, almost identical. A side-by-side comparison between the upper limits computed using the frequentist asymptotic approximation and those computed using the Bayesian calculation are shown in Table A.7 of the appendix. We, thus, exclude Standard Model Higgs boson production with hypothesized Higgs boson masses between 129 GeV/c^2 and 260 GeV/c^2.

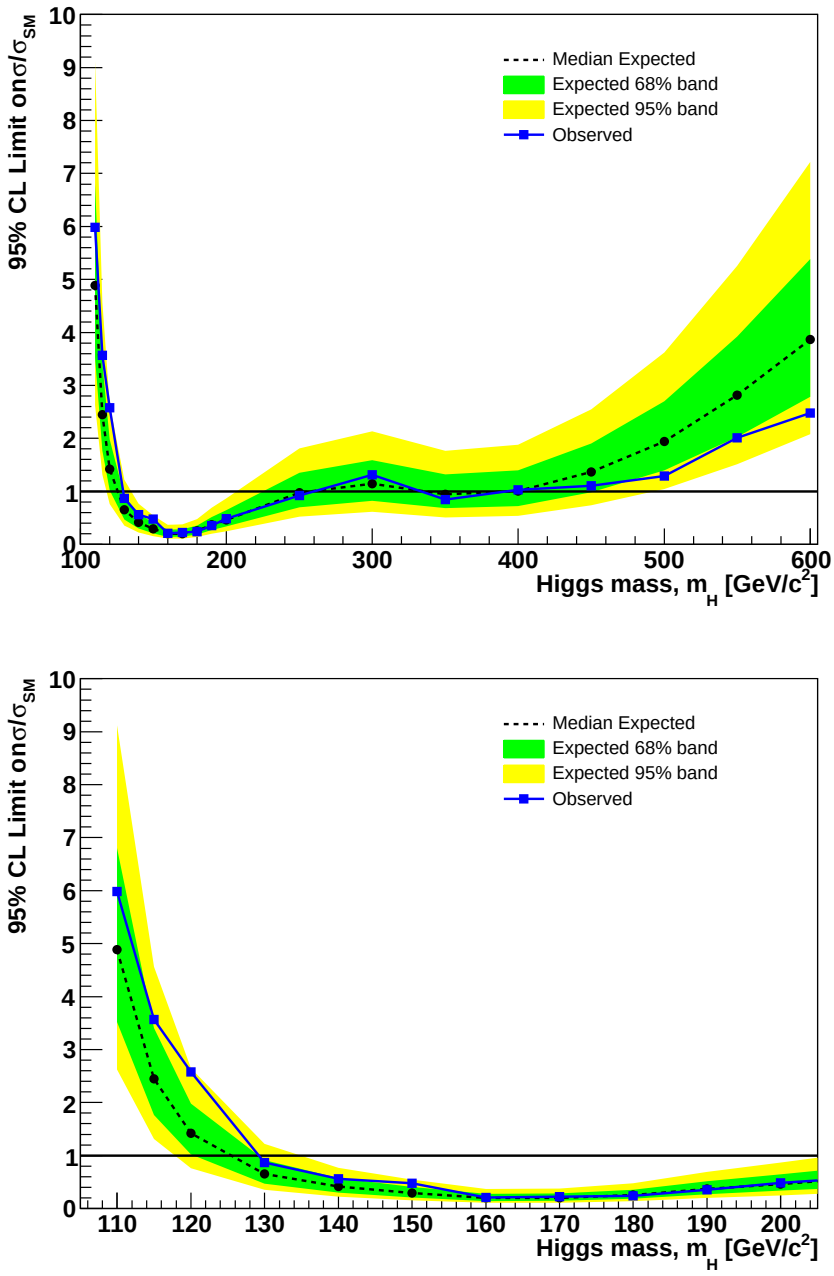


Figure 8-3: Expected and observed upper limits on the Standard Model Higgs boson production cross section are shown as a function of the hypothesized Higgs boson mass for the BDT shape analysis using the frequentist asymptotic approxiamtion. In the first plot, we show the upper limits for the full range of Higgs boson masses that were tested, from 110 GeV/$c^2$to 600 GeV/c^2. In the second plot, we zoom in to the low mass region between 110 GeV/c^2 and 200 GeV/c^2.

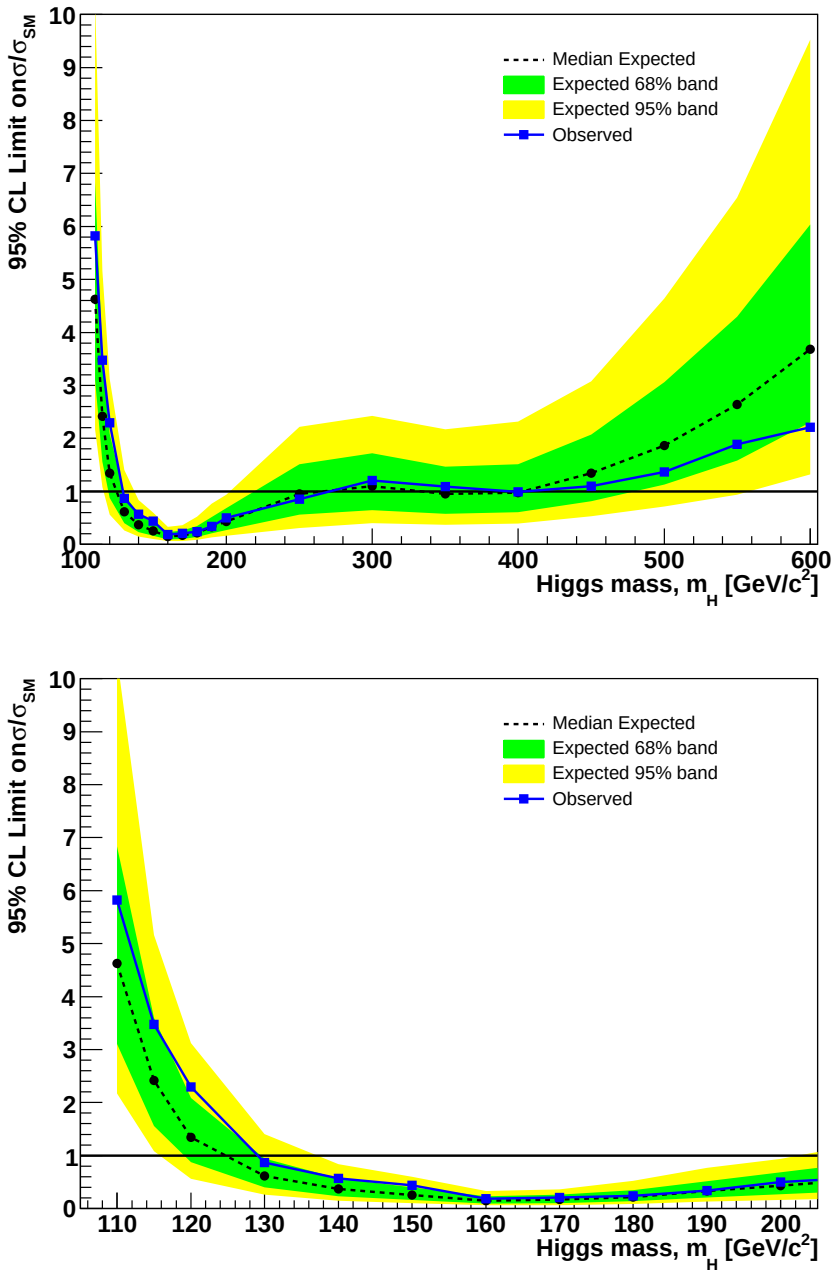


Figure 8-4: Expected and observed upper limits on the Standard Model Higgs boson production cross section are shown as a function of the hypothesized Higgs boson mass for the BDT shape analysis using the Bayesian calculation. In the first plot, we show the upper limits for the full range of Higgs boson masses that were tested, from 110 GeV/$c^2$to 600 GeV/c^2. In the second plot, we zoom in to the low mass region between 110 GeV/c^2 and 200 GeV/c^2.

Chapter 9

Conclusion

We have performed a search for the Higgs boson in the $H \rightarrow W^+W^- \rightarrow 2l\ 2\nu$ decay channel. Methods for suppression of the backgrounds and its precise estimation are the key aspects of this search. To suppress W+jets background, the electron and muon selections are optimized using a multivariate discriminator trained on high purity signal and background control data samples. To suppress the Drell-Yan background, we make use of several complementary estimates of the missing transverse energy. To suppress the top background, we categorize in the number of reconstructed jets and make use of b-tagging techniques. Finally, to suppress W^+W^- background, we make use of features of the kinematic observables. These most important backgrounds are all estimated using data driven techniques and cross checked in various background enhanced control regions.

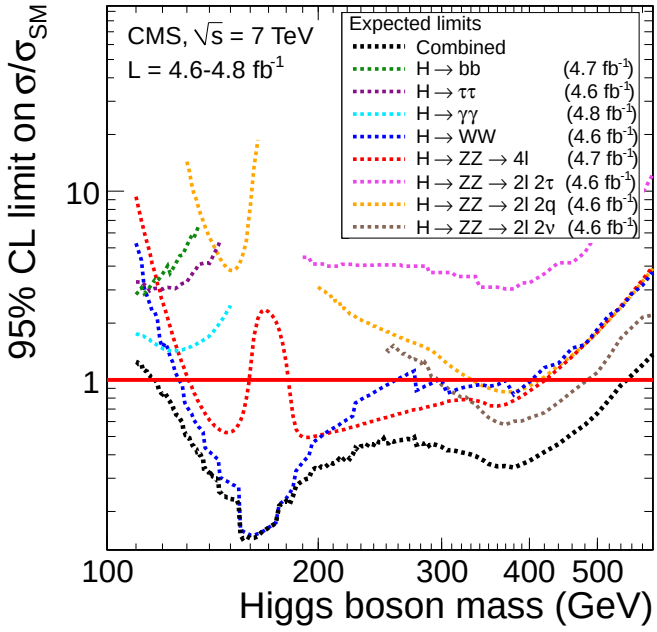
Finally, to extract the signal, we have perform a cut and count analysis as well as a multivariate discriminator shape analysis. Using 4.9 fb^{-1} of integrated luminosity, we expected on average to exclude Standard Model Higgs boson production for masses between 126 GeV/c^2 and 260 GeV/c^2 under the background-only hypothesis. We observed no statistically significant excess in the data, and therefore exclude the Standard Model Higgs boson for masses between 129 GeV/c^2 and 260 GeV/c^2.

To place this result in a global context, we show the expected and observed upper limits on the Higgs boson production cross section obtained through a statistical

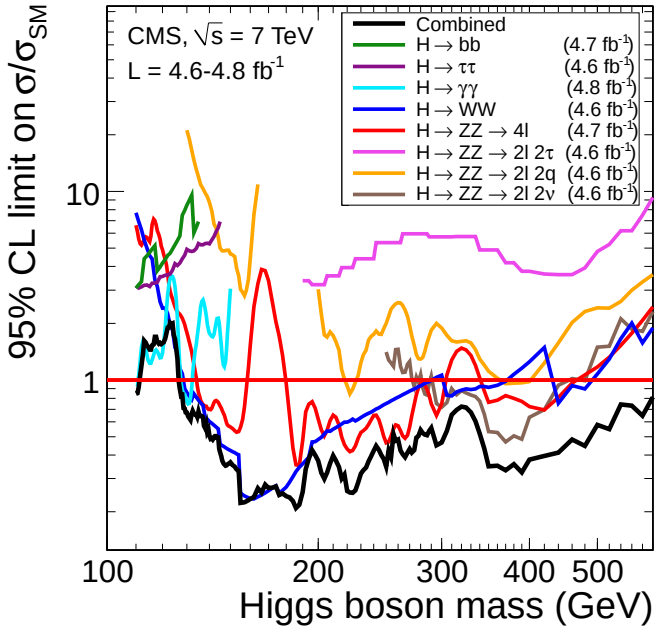
combination of all Higgs boson search channels used in the CMS experiment [118] in Figure 9-1. The particular result for the $H \to W^+W^-$ channel used in this combination is documented in reference [119] and differs from the result discussed in this thesis very slightly. The only difference is in the selection of electrons and muons, for which the result in this thesis is more optimal. From the combined upper limits, we observe that the $H \to W^+W^-$ channel is the most sensitive channel in the mass range between 122 GeV/c^2 and 200 GeV/c^2, and contributes significantly to the combination in the full mass range. From Figure 9-2, showing a zoom-in version of the combination results at low Higgs boson mass, we see that the $H \to W^+W^-$ channel is exceeded in sensitivity by the $H \to \gamma\gamma$ channel below 122 GeV, but maintains good relative sensitivity all the way down to a Higgs boson mass of about 115 GeV/c^2.

The combined CMS upper limits on the Higgs boson production cross section are shown in Figure 9-3. The expected 95% confidence-level exclusion range is between 118 GeV/c^2 and 543 GeV/c^2. The observed exclusion range is between 127 GeV/c^2 and 600 GeV/c^2. The local p-values quantifying the probability that the observation is inconsistent with the background-only hypothesis is shown in Figure 9-4 as a function of the Higgs boson mass. The largest observed excess has a local significance of 3.1 standard deviations. Accounting for the "look-elsewhere effect" [120], this significance is reduced to a global significance of 1.5 standard deviations. More data are required to ascertain the origin of this excess.

In conclusion, for the parameter space favored by the Standard Model electroweak precision fits, the allowed mass range for the Higgs boson has been almost eliminated, with the only remaining region being between 114.4 GeV/c^2 and 127 GeV/c^2. No statistically significant excess has been observed so far, however a small excess has been observed within the allowed mass range. The LHC is expected to deliver another 15 fb^{-1} of integrated luminosity in 2012, increasing the dataset by a factor of four. With this additional data to be collected in 2012, there is a good prospect that a definitive answer can be obtained on the existence or non-existence of the Standard Model Higgs boson in the full mass range.

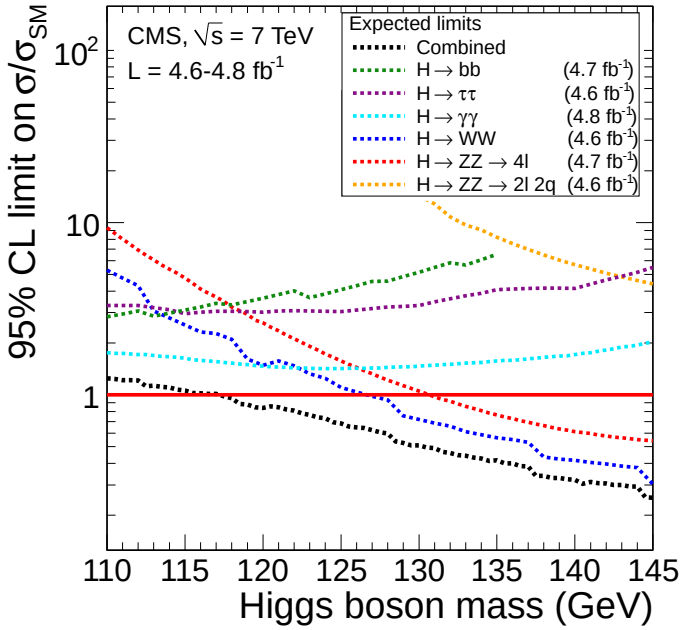


(a) Expected Limits

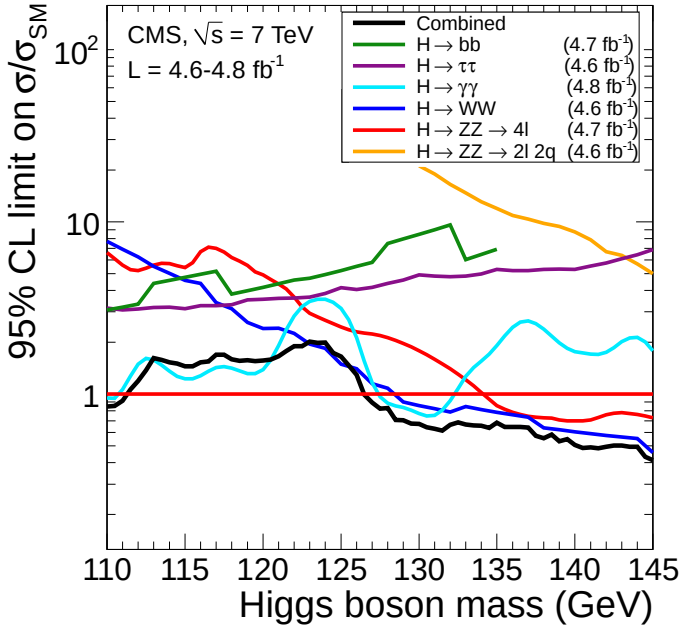


(b) Observed Limits

Figure 9-1: The expected and observed upper limits on the Standard Model Higgs boson production cross section obtained through the statistical combination of all search channels used in CMS is shown as a function of the Higgs boson mass. The $H \rightarrow W^+W^-$ channel, shown in blue, is the most sensitive channel in the mass range between 122 GeV/c^2 and 200 GeV/c^2, and contributes significantly to the combination in the full mass range.



(a) Expected Limits



(b) Observed Limits

Figure 9-2: A zoom-in version of Figure 9-1 shows more clearly the results in the low mass region between 110 GeV/c^2 and 145 GeV/c^2. The $H \rightarrow W^+W^-$ channel is shown in blue and is the most sensitive channel down to a Higgs boson mass of 122 GeV/c^2, where the $H \rightarrow \gamma\gamma$ channel becomes the most sensitive.

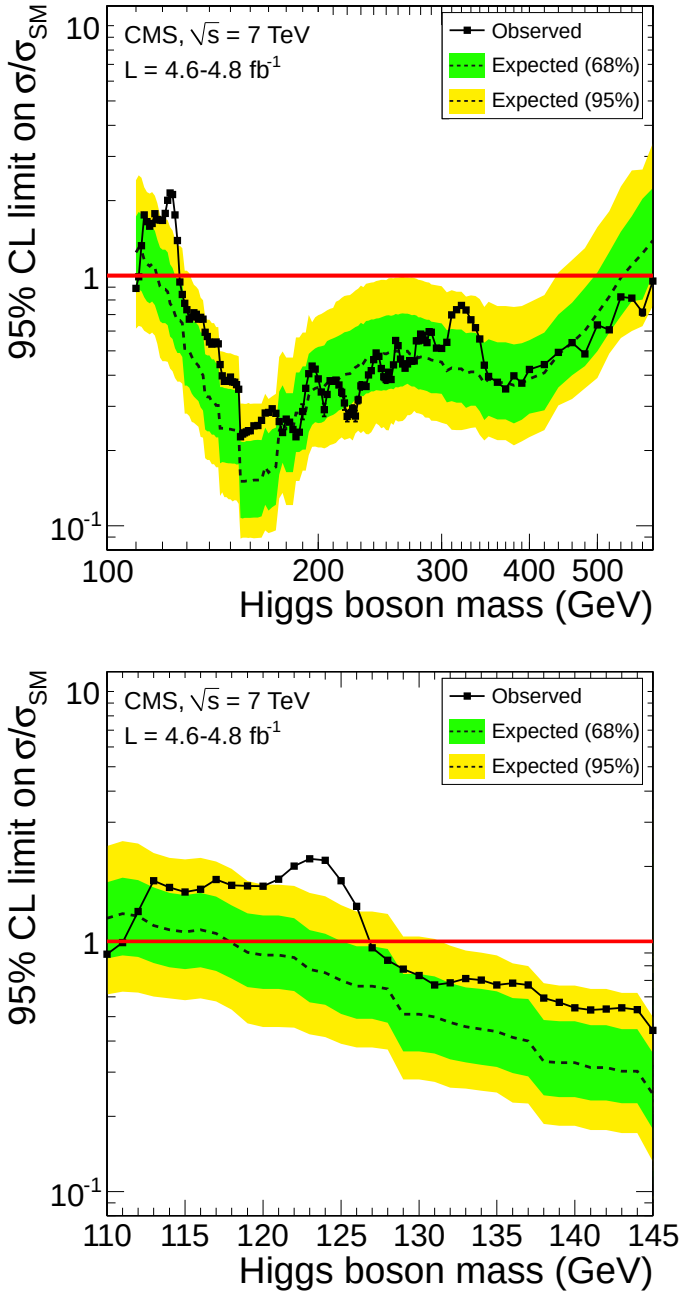


Figure 9-3: The 95% confidence-level upper limits on the Standard Model Higgs boson production cross section obtained through the CMS Higgs combination is shown. The dotted black curve shows the expected upper limit, while the solid curve shows the observed upper limits. The green and yellow regions show the 1-standard-deviation and 2-standard-deviation bands representing the regions that are expected to contain 68% and 95% of all observed excursions from the median. The median expected exclusion mass range is between 118 GeV/c^2 and 543 GeV/c^2The observed exclusion mass range is between 127 GeV/c^2 and 600 GeV/c^2.

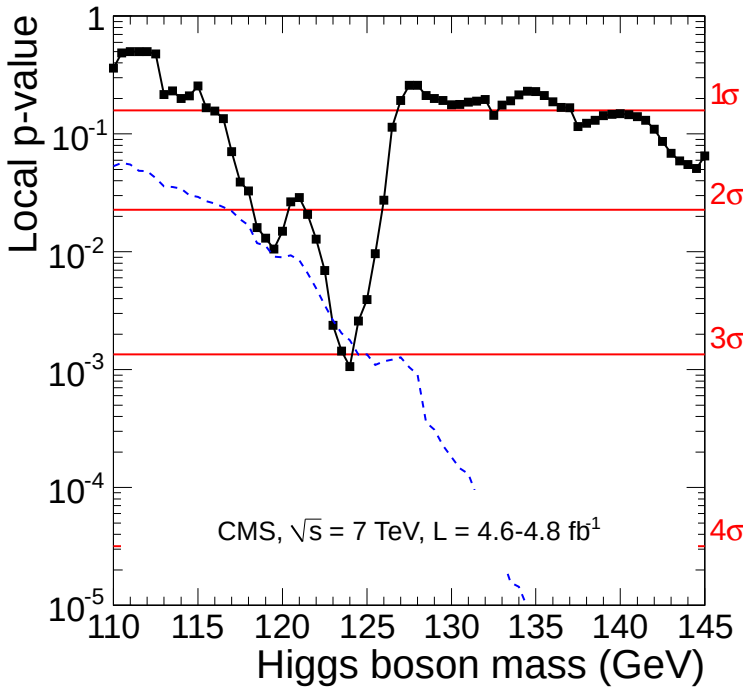


Figure 9-4: The observed local p-value is shown as a function of the Higgs boson mass. The largest observed excess has a local significance of 3.1σ. Accounting for the "look-elsewhere effect", this significance is reduced to a global significance of 1.5σ.

Appendix A

Tables of Upper Limits

Higgs Boson Mass	Observed	Median expected	68% Confidence Level Range	95% Confidence Level Range
110 GeV/c^2	7.57	6.02	[4.34, 8.38]	[3.23, 11.24]
115 GeV/c^2	3.53	2.91	[2.09, 4.04]	[1.56, 5.42]
120 GeV/c^2	2.12	1.74	[1.25, 2.42]	[0.93, 3.24]
130 GeV/c^2	1.13	0.83	[0.60, 1.16]	[0.45, 1.56]
140 GeV/c^2	0.62	0.53	[0.39, 0.74]	[0.29, 1.00]
150 GeV/c^2	0.46	0.40	[0.28, 0.55]	[0.21, 0.74]
160 GeV/c^2	0.25	0.23	[0.17, 0.32]	[0.12, 0.43]
170 GeV/c^2	0.24	0.24	[0.17, 0.33]	[0.13, 0.44]
180 GeV/c^2	0.29	0.31	[0.23, 0.44]	[0.17, 0.58]
190 GeV/c^2	0.36	0.46	[0.33, 0.65]	[0.25, 0.87]
200 GeV/c^2	0.54	0.58	[0.42, 0.81]	[0.31, 1.08]
250 GeV/c^2	1.09	1.13	[0.82, 1.58]	[0.61, 2.11]
300 GeV/c^2	1.23	1.33	[0.96, 1.85]	[0.71, 2.49]
350 GeV/c^2	1.17	1.27	[0.92, 1.77]	[0.68, 2.38]
400 GeV/c^2	1.23	1.38	[1.00, 1.92]	[0.74, 2.58]
450 GeV/c^2	1.66	1.75	[1.26, 2.44]	[0.94, 3.26]
500 GeV/c^2	2.16	2.41	[1.74, 3.35]	[1.29, 4.49]
550 GeV/c^2	3.38	3.34	[2.41, 4.65]	[1.79, 6.23]
600 GeV/c^2	5.27	4.62	[3.33, 6.42]	[2.48, 8.61]

Table A.1: Expected and observed upper limits on the Standard Model Higgs boson production cross section using the frequentist LHC type CL_s approach in the large N asymptotic limit for the cut-based analysis are shown. These upper limits are expressed as ratios to the Higgs boson production cross sections predicted by the Standard Model.

Higgs Boson Mass	Observed	Median expected	68% Confidence Level Range	95% Confidence Level Range
110 GeV/c^2	7.22	6.34	[4.19, 9.15]	[2.88, 13.01]
115 GeV/c^2	3.37	2.99	[2.05, 4.46]	[1.51, 6.19]
120 GeV/c^2	2.04	1.75	[1.24, 2.58]	[0.89, 3.70]
130 GeV/c^2	1.08	0.82	[0.58, 1.19]	[0.41, 1.68]
140 GeV/c^2	0.60	0.53	[0.37, 0.76]	[0.26, 1.08]
150 GeV/c^2	0.46	0.38	[0.26, 0.56]	[0.17, 0.83]
160 GeV/c^2	0.22	0.21	[0.14, 0.30]	[0.10, 0.43]
170 GeV/c^2	0.21	0.21	[0.15, 0.31]	[0.11, 0.43]
180 GeV/c^2	0.25	0.30	[0.21, 0.45]	[0.16, 0.63]
190 GeV/c^2	0.28	0.47	[0.33, 0.67]	[0.23, 1.00]
200 GeV/c^2	0.52	0.60	[0.42, 0.85]	[0.29, 1.21]
250 GeV/c^2	1.04	1.18	[0.82, 1.79]	[0.61, 2.55]
300 GeV/c^2	1.18	1.38	[0.96, 1.99]	[0.69, 2.72]
350 GeV/c^2	1.12	1.34	[0.93, 1.92]	[0.69, 2.64]
400 GeV/c^2	1.17	1.44	[0.98, 2.05]	[0.71, 2.94]
450 GeV/c^2	1.59	1.82	[1.28, 2.66]	[0.94, 3.84]
500 GeV/c^2	2.10	2.46	[1.74, 3.56]	[1.28, 5.13]
550 GeV/c^2	3.35	3.54	[2.45, 5.24]	[1.79, 7.67]
600 GeV/c^2	5.31	5.11	[3.53, 7.53]	[2.46, 10.60]

Table A.2: Expected and observed upper limits on the Standard Model Higgs boson production cross section using the Bayesian approach for the cut-based analysis are shown. These upper limits are expressed as ratios to the Higgs boson production cross sections predicted by the Standard Model.

Higgs Boson Mass	Observed	Median expected	68% Confidence Level Range	95% Confidence Level Range
110 GeV/c^2	5.98	4.89	[3.52, 6.80]	[2.62, 9.12]
115 GeV/c^2	3.56	2.45	[1.76, 3.40]	[1.31, 4.56]
120 GeV/c^2	2.58	1.42	[1.02, 1.98]	[0.76, 2.65]
130 GeV/c^2	0.87	0.66	[0.47, 0.91]	[0.35, 1.22]
140 GeV/c^2	0.56	0.41	[0.30, 0.57]	[0.22, 0.77]
150 GeV/c^2	0.48	0.29	[0.21, 0.41]	[0.16, 0.55]
160 GeV/c^2	0.21	0.20	[0.14, 0.27]	[0.11, 0.37]
170 GeV/c^2	0.23	0.20	[0.15, 0.28]	[0.11, 0.38]
180 GeV/c^2	0.24	0.25	[0.18, 0.35]	[0.14, 0.48]
190 GeV/c^2	0.35	0.37	[0.27, 0.52]	[0.20, 0.69]
200 GeV/c^2	0.49	0.46	[0.33, 0.65]	[0.25, 0.87]
250 GeV/c^2	0.92	0.97	[0.70, 1.35]	[0.52, 1.81]
300 GeV/c^2	1.32	1.14	[0.82, 1.59]	[0.61, 2.13]
350 GeV/c^2	0.85	0.95	[0.68, 1.32]	[0.51, 1.76]
400 GeV/c^2	1.03	1.01	[0.72, 1.40]	[0.54, 1.88]
450 GeV/c^2	1.10	1.37	[0.98, 1.90]	[0.73, 2.55]
500 GeV/c^2	1.29	1.94	[1.40, 2.70]	[1.04, 3.62]
550 GeV/c^2	2.01	2.82	[2.03, 3.92]	[1.51, 5.25]
600 GeV/c^2	2.48	3.87	[2.79, 5.38]	[2.08, 7.22]

Table A.3: Expected and observed upper limits on the Standard Model Higgs boson production cross section using the frequentist LHC-type CL_s approach in the large-N asymptotic limit for the BDT shape analysis are shown. These upper limits are expressed as ratios to the Higgs boson production cross sections predicted by the Standard Model.

Higgs Boson Mass	Observed	Median expected	68% Confidence Level Range	95% Confidence Level Range
110 GeV/c^2	5.83	4.63	[3.10, 6.83]	[2.17, 10.37]
115 GeV/c^2	3.47	2.41	[1.56, 3.60]	[1.08, 5.16]
120 GeV/c^2	2.29	1.35	[0.88, 2.09]	[0.56, 3.13]
130 GeV/c^2	0.87	0.61	[0.40, 0.94]	[0.26, 1.41]
140 GeV/c^2	0.57	0.37	[0.23, 0.56]	[0.15, 0.84]
150 GeV/c^2	0.44	0.25	[0.16, 0.40]	[0.10, 0.60]
160 GeV/c^2	0.18	0.15	[0.09, 0.23]	[0.06, 0.33]
170 GeV/c^2	0.21	0.17	[0.10, 0.26]	[0.07, 0.36]
180 GeV/c^2	0.24	0.21	[0.14, 0.34]	[0.08, 0.52]
190 GeV/c^2	0.34	0.33	[0.21, 0.51]	[0.13, 0.77]
200 GeV/c^2	0.50	0.43	[0.27, 0.68]	[0.16, 0.94]
250 GeV/c^2	0.85	0.95	[0.56, 1.51]	[0.31, 2.22]
300 GeV/c^2	1.20	1.10	[0.65, 1.72]	[0.40, 2.42]
350 GeV/c^2	1.09	0.95	[0.58, 1.47]	[0.37, 2.17]
400 GeV/c^2	0.99	0.98	[0.61, 1.51]	[0.39, 2.31]
450 GeV/c^2	1.10	1.34	[0.81, 2.07]	[0.53, 3.08]
500 GeV/c^2	1.37	1.86	[1.12, 3.06]	[0.71, 4.64]
550 GeV/c^2	1.89	2.64	[1.58, 4.30]	[0.93, 6.55]
600 GeV/c^2	2.21	3.68	[2.30, 6.04]	[1.32, 9.52]

Table A.4: Expected and observed upper limits on the Standard Model Higgs boson production cross section using the Bayesian approach for the BDT shape analysis are shown. These upper limits are expressed as ratios to the Higgs boson production cross sections predicted by the Standard Model.

Higgs Boson Mass	Observed		Median expected	
	Cut-Based	BDT Shape	Cut-Based	BDT Shape
110 GeV/c^2	7.57	5.98	6.02	4.89
115 GeV/c^2	3.53	3.56	2.91	2.45
120 GeV/c^2	2.12	2.58	1.74	1.42
130 GeV/c^2	1.13	0.87	0.83	0.66
140 GeV/c^2	0.62	0.56	0.53	0.41
150 GeV/c^2	0.46	0.48	0.40	0.29
160 GeV/c^2	0.25	0.21	0.23	0.20
170 GeV/c^2	0.24	0.23	0.24	0.20
180 GeV/c^2	0.29	0.24	0.31	0.25
190 GeV/c^2	0.36	0.35	0.46	0.37
200 GeV/c^2	0.54	0.49	0.58	0.46
250 GeV/c^2	1.09	0.92	1.13	0.97
300 GeV/c^2	1.23	1.32	1.33	1.14
350 GeV/c^2	1.17	0.85	1.27	0.95
400 GeV/c^2	1.23	1.03	1.38	1.01
450 GeV/c^2	1.66	1.10	1.75	1.37
500 GeV/c^2	2.16	1.29	2.41	1.94
550 GeV/c^2	3.38	2.01	3.34	2.82
600 GeV/c^2	5.27	2.48	4.62	3.87

Table A.5: A side-by-side comparison is shown between the upper limits obtained using the cut-based analysis and those obtained using the BDT shape analysis. The frequentist asymptotic approximation has been used for the upper limit calculation. These upper limits are expressed as ratios to the Higgs boson production cross sections predicted by the Standard Model. We observe that the BDT shape analysis has on average about 20% better sensitivity.

Higgs Boson Mass	Observed		Median expected	
	Cut-Based	BDT Shape	Cut-Based	BDT Shape
110 GeV/c^2	7.22	5.83	6.34	4.63
115 GeV/c^2	3.37	3.47	2.99	2.41
120 GeV/c^2	2.04	2.29	1.75	1.35
130 GeV/c^2	1.08	0.87	0.82	0.61
140 GeV/c^2	0.60	0.57	0.53	0.37
150 GeV/c^2	0.46	0.44	0.38	0.25
160 GeV/c^2	0.22	0.18	0.21	0.15
170 GeV/c^2	0.21	0.21	0.21	0.17
180 GeV/c^2	0.25	0.24	0.30	0.21
190 GeV/c^2	0.28	0.34	0.47	0.33
200 GeV/c^2	0.52	0.50	0.60	0.43
250 GeV/c^2	1.04	0.85	1.18	0.95
300 GeV/c^2	1.18	1.20	1.38	1.10
350 GeV/c^2	1.12	1.09	1.34	0.95
400 GeV/c^2	1.17	0.99	1.44	0.98
450 GeV/c^2	1.59	1.10	1.82	1.34
500 GeV/c^2	2.10	1.37	2.46	1.86
550 GeV/c^2	3.35	1.89	3.54	2.64
600 GeV/c^2	5.31	2.21	5.11	3.68

Table A.6: A side-by-side comparison is shown between the upper limits obtained using the cut-based analysis and those obtained using the BDT shape analysis. The Bayesian approach has been used for the upper limit calculation. These upper limits are expressed as ratios to the Higgs boson production cross sections predicted by the Standard Model. We observe that the BDT shape analysis has on average about 20% better sensitivity.

Higgs Boson Mass	Observed		Median expected	
	Freqentist	Bayesian	Frequentist	Bayesian
110 GeV/c^2	5.98	5.83	4.89	4.63
115 GeV/c^2	3.56	3.47	2.45	2.41
120 GeV/c^2	2.58	2.29	1.42	1.35
130 GeV/c^2	0.87	0.87	0.66	0.61
140 GeV/c^2	0.56	0.57	0.41	0.37
150 GeV/c^2	0.48	0.44	0.29	0.25
160 GeV/c^2	0.21	0.18	0.20	0.15
170 GeV/c^2	0.23	0.21	0.20	0.17
180 GeV/c^2	0.24	0.24	0.25	0.21
190 GeV/c^2	0.35	0.34	0.37	0.33
200 GeV/c^2	0.49	0.50	0.46	0.43
250 GeV/c^2	0.92	0.85	0.97	0.95
300 GeV/c^2	1.32	1.20	1.14	1.10
350 GeV/c^2	0.85	1.09	0.95	0.95
400 GeV/c^2	1.03	0.99	1.01	0.98
450 GeV/c^2	1.10	1.10	1.37	1.34
500 GeV/c^2	1.29	1.37	1.94	1.86
550 GeV/c^2	2.01	1.89	2.82	2.64
600 GeV/c^2	2.48	2.21	3.87	3.68

Table A.7: A side-by-side comparison is shown between the upper limits obtained using the frequentist asymptotic approximation and those obtained using the Bayesian approach, for the BDT shape analysis. We observe that the two different statistical approaches give upper limits that are within a few percent of each other. Despite conceptual differences between the two approaches, the numerical results that we obtain are very similar.

Bibliography

[1] ALEPH, CDF, D0, DELPHI, L3, OPAL, SLD Collaborations, the LEP Working Group, the Tevatron Electroweak Working Group, and the SLD Electroweak and Heavy flavor Group, "Precision electroweak measurements and constraints on the standard model," 2010.

[2] N. Jarosik, C. Bennett, J. Dunkley, B. Gold, M. Greason, *et al.*, "Seven-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Sky Maps, Systematic Errors, and Basic Results," *Astrophys.J.Suppl.*, vol. 192, p. 14, 2011.

[3] G. Bertone, D. Hooper, and J. Silk, "Particle dark matter: Evidence, candidates and constraints," *Phys.Rept.*, vol. 405, pp. 279–390, 2005.

[4] K. Nakamura *et al.*, "Review of particle physics," *J.Phys.G*, vol. G37, p. 075021, 2010.

[5] M. E. Peskin and D. V. Schroeder, *Introduction to quantum field theory*. Boson, Massachusetts: Addison-Wesley Pub. Co., 1995.

[6] H. Politzer, "Reliable Perturbative Results for Strong Interactions?," *Phys.Rev.Lett.*, vol. 30, pp. 1346–1349, 1973.

[7] D. Gross and F. Wilczek, "Ultraviolet Behavior of Nonabelian Gauge Theories," *Phys.Rev.Lett.*, vol. 30, pp. 1343–1346, 1973.

[8] S. L. Glashow, "Partial symmetries of weak interactions," *Nucl. Phys.*, vol. 22, pp. 579–588, 1961.

[9] S. Weinberg, "A model of leptons," *Phys. Rev. Lett.*, vol. 19, pp. 1264–1266, 1967.

[10] A. Salam, "Weak and electromagnetic interactions," in *Elementary particle physics: relativistic groups and analyticity* (N. Svartholm, ed.), p. 367, Almquvist & Wiskell, 1968. Proceedings of the eighth Nobel symposium.

[11] F. Englert and R. Brout, "Broken symmetry and the mass of gauge vector mesons," *Phys. Rev. Lett.*, vol. 13, pp. 321–323, 1964.

[12] P. W. Higgs, "Broken symmetries, massless particles and gauge fields," *Phys. Rev. Lett.*, vol. 12, pp. 132–133, 1964.

[13] P. W. Higgs, "Broken symmetries and the masses of gauge bosons," *Phys. Rev. Lett.*, vol. 13, p. 508, 1964.

[14] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, "Global conservation laws and massless particles," *Phys. Rev. Lett.*, vol. 13, pp. 585–587, 1964.

[15] P. W. Higgs, "Spontaneous symmetry breakdown without massless bosons," *Phys. Rev.*, vol. 145, pp. 1156–1163, 1966.

[16] T. W. B. Kibble, "Symmetry breaking in non-Abelian gauge theories," *Phys. Rev.*, vol. 155, pp. 1554–1561, 1967.

[17] J. Goldstone, "Field theories with,"

[18] P. W. Anderson, "Plasmons, gauge invariance, and mass," *Phys. Rev.*, vol. 130, pp. 439–442, Apr 1963.

[19] G. Hooft, "Renormalizable lagrangians for massive yang-mills fields," *Nuclear Physics B*, vol. 35, no. 1, pp. 167 – 188, 1971.

[20] F. Hasert *et al.*, "Observation of neutrino-like interactions without muon or electron in the gargamelle neutrino experiment," *Physics Letters B*, vol. 46, no. 1, pp. 138 – 140, 1973.

[21] C. Prescott *et al.*, "Parity non-conservation in inelastic electron scattering," *Physics Letters B*, vol. 77, no. 3, pp. 347–352, 1978.

[22] G. Arnison *et al.*, "Further evidence for charged intermediate vector bosons at the SPS collider," *Physics Letters B*, vol. 129, no. 3-4, pp. 273 – 282, 1983.

[23] G. Arnison *et al.*, "Experimental observation of lepton pairs of invariant mass around 95 gev/c2 at the CERN SPS collider," *Physics Letters B*, vol. 126, no. 5, pp. 398 – 410, 1983.

[24] ALEPH, DELPHI, L3, OPAL Collaborations, and the LEP Working Group for Higgs Boson Searches, "Search for the standard model Higgs boson at LEP," *Phys. Lett. B*, vol. 565, p. 61, 2003.

[25] e. Voss, Rudiger and e. Breskin, Amos, "The CERN Large Hadron Collider, accelerator and experiments," 2009. Contents: Vol. 1. LHC machine, ALICE, and ATLAS – Vol. 2. CMS, LHCb, LHCf, and TOTEM.

[26] CDF and D0 Collaborations, "Combination of Tevatron searches for the standard model Higgs boson in the W^+W^- decay mode," *Phys. Rev. Lett.*, vol. 104, p. 061802, 2010. A more recent, unpublished, limit is given in preprint arXiv:1103.3233.

[27] T. Aaltonen *et al.*, "Search for a Higgs boson decaying to two W bosons at CDF," *Phys. Rev. Lett.*, vol. 102, p. 021802, 2009.

[28] T. Aaltonen and others, "Measurement of the W^+W^- production cross section and search for anomalous $WW\gamma$ and WWZ couplings in $p\bar{p}$ collisions at $\sqrt{s}$=1.96 TeV," *Phys. Rev. Lett.*, vol. 104, p. 201801, 2010.

[29] S. Dawson, "Radiative corrections to Higgs boson production," *Nucl. Phys. B*, vol. 359, pp. 283–300, 1991.

[30] M. Spira *et al.*, "Higgs boson production at the LHC," *Nucl. Phys. B*, vol. 453, pp. 17–82, 1995.

[31] R. V. Harlander and W. B. Kilgore, "Next-to-next-to-leading order Higgs production at hadron colliders," *Phys. Rev. Lett.*, vol. 88, p. 201801, 2002.

[32] C. Anastasiou and K. Melnikov, "Higgs boson production at hadron colliders in NNLO QCD," *Nucl. Phys. B*, vol. 646, pp. 220–256, 2002.

[33] S. Alioli, P. Nason, C. Oleari, and E. Re, "NLO vector-boson production matched with shower in POWHEG," *JHEP*, vol. 0807, p. 060, 2008.

[34] V. Ravindran, J. Smith, and W. L. van Neerven, "NNLO corrections to the total cross section for Higgs boson production in hadron hadron collisions," *Nucl. Phys. B*, vol. 665, pp. 325–366, 2003.

[35] S. Catani, D. de Florian, M. Grazzini, and P. Nason, "Soft-gluon resummation for Higgs boson production at hadron colliders," *JHEP*, vol. 07, p. 028, 2003.

[36] S. Actis *et al.*, "NLO electroweak corrections to Higgs boson production at hadron colliders," *Phys. Lett. B*, vol. 670, pp. 12–17, 2008.

[37] C. Anastasiou, R. Boughezal, and F. Petriello, "Mixed QCD-electroweak corrections to Higgs boson production in gluon fusion," *JHEP*, vol. 04, p. 003, 2009.

[38] D. de Florian and M. Grazzini, "Higgs production through gluon fusion: updated cross sections at the Tevatron and the LHC," *Phys. Lett. B*, vol. 674, pp. 291–294, 2009.

[39] M. Ciccolini, A. Denner, and S. Dittmaier, "Strong and electroweak corrections to the production of Higgs + 2-jets via weak interactions at the LHC," *Phys. Rev. Lett.*, vol. 99, p. 161803, 2007.

[40] M. Ciccolini, A. Denner, and S. Dittmaier, "Electroweak and QCD corrections to Higgs production via vector-boson fusion at the LHC," *Phys. Rev. D*, vol. 77, p. 013002, 2008.

[41] K. Arnold *et al.*, "VBFNLO: A parton level monte carlo for processes with electroweak bosons," *Comput. Phys. Commun.*, vol. 180, pp. 1661–1670, 2009.

[42] O. Brein, A. Djouadi, and R. Harlander, "NNLO QCD corrections to the Higgs-strahlung processes at hadron colliders," *Phys. Lett. B*, vol. 579, pp. 149–156, 2004.

[43] M. L. Ciccolini, S. Dittmaier, and M. Krämer, "Electroweak radiative corrections to associated WH and ZH production at hadron colliders," *Phys. Rev. D*, vol. 68, p. 073003, 2003.

[44] LHC Higgs Cross Section Working Group, S. Dittmaier, C. Mariotti, G. Passarino, R. Tanaka (Eds.), "Handbook of LHC Higgs cross sections: Inclusive observables," CERN Report CERN-2011-002, 2011.

[45] A. Djouadi *et al.*, "An update of the program HDECAY," in *The Les Houches 2009 workshop on TeV colliders: The tools and Monte Carlo working group summary report*, 2010.

[46] A. Denner *et al.*, "Standard model Higgs-boson branching ratios with uncertainties," *Eur. Phys. J. C*, vol. 71, p. 1753, 2011.

[47] A. Bredenstein *et al.*, "Precise predictions for the Higgs-boson decay H $\rightarrow$ WW/ZZ $\rightarrow$ 4 leptons," *Phys. Rev. D*, vol. 374, p. 013004, 2006.

[48] A. Bredenstein, A. Denner, S. Dittmaier, and M. M. Weber, "Radiative corrections to the semileptonic and hadronic Higgs-boson decays H $\rightarrow$W W / Z Z$\rightarrow$ 4 fermions," *JHEP*, vol. 0702, p. 080, 2007.

[49] D. de Florian, G. Ferrera, M. Grazzini, and D. Tommasini, "Transverse-momentum resummation: Higgs boson production at the Tevatron and the LHC," *JHEP*, vol. 2011, p. 64, 2011.

[50] S. Alioli, P. Nason, C. Oleari, and E. Re, "NLO Higgs boson production via gluon fusion matched with shower in POWHEG," *JHEP*, vol. 04, p. 002, 2009.

[51] M. Dittmar and H. K. Dreiner, "How to find a Higgs boson with a mass between 155 GeV and 180 GeV at the CERN LHC," *Phys. Rev. D*, vol. 55, p. 167, 1996.

[52] M. Bajko *et al.*, "Report of the Task Force on the Incident of 19th September 2008 at the LHC. oai:cds.cern.ch:1168025," Tech. Rep. LHC-PROJECT-Report-1168. CERN-LHC-PROJECT-Report-1168, CERN, Geneva, Mar 2009.

[53] S. Chatrchyan *et al.*, "The CMS experiment at the CERN LHC," *JINST*, vol. 3, p. S08004, 2008.

[54] C. Collaboration, J. Puerta-Pelayo, M. Benettoni, and D. Dattola, "Images of cms muons - drift tubes (dt)." CMS Collection., Nov 2008.

[55] C. Eck, J. Knobloch, L. Robertson, I. Bird, K. Bos, N. Brook, D. Dllmann, I. Fisk, D. Foster, B. Gibbard, C. Grandi, F. Grey, J. Harvey, A. Heiss, F. Hemmer, S. Jarp, R. Jones, D. Kelsey, M. Lamanna, H. Marten, P. Mato-Vila, F. Ould-Saada, B. Panzer-Steindel, L. Perini, Y. Schutz, U. Schwickerath, J. Shiers, and T. Wenaus, *LHC computing Grid: Technical Design Report. Version 1.06 (20 Jun 2005)*. Technical Design Report LCG, Geneva: CERN, 2005.

[56] CMS Collaboration, "Electron reconstruction and identification at $\sqrt{s} = 7$ TeV," CMS Physics Analysis Summary CMS-PAS-EGM-10-004, 2010.

[57] R. Fruhwirth, "Application of Kalman filtering to track and vertex fitting," *Nucl.Instrum.Meth.*, vol. A262, pp. 444–450, 1987.

[58] H. Bethe and W. Heitler, "On the Stopping of fast particles and on the creation of positive electrons," *Proc.Roy.Soc.Lond.*, vol. A146, pp. 83–112, 1934.

[59] W. Adam, R. Fruhwirth, A. Strandlie, and T. Todorov, "Reconstruction of electrons with the Gaussian sum filter in the CMS tracker at LHC," *eConf*, vol. C0303241, p. TULT009, 2003.

[60] V. Khachatryan *et al.*, "Measurements of inclusive W and Z cross sections in pp collisions at $\sqrt{s} = 7$ TeV," *JHEP*, vol. 01, p. 080, 2011.

[61] CMS Collaboration, "Measurement of WW production and search for the Higgs boson in pp collisions at $\sqrt{s} = 7$ TeV," *Phys. Lett. B*, vol. 699, p. 25, 2011.

[62] M. Cacciari, G. P. Salam, and G. Soyez, "Fastjet user manual. oai:cds.cern.ch:1402449," Tech. Rep. arXiv:1111.6097. CERN-PH-TH-2011-297, Nov 2011. Comments: 69 pages. FastJet 3 is available from http://fastjet.fr/.

[63] M. Cacciari, G. P. Salam, "Dispelling the N^3 myth for the k_t jet-finder," *Phys. Lett. B*, vol. 641, p. 57, 2006.

[64] M. Cacciari, G. P. Salam, "Pileup subtraction using jet areas," *Phys. Lett. B*, vol. 659, p. 119, 2008.

[65] J. H. Friedman, "Greedy function approximation: A gradient boosting machine," *Annals of Statistics*, vol. 29, pp. 1189–1232, 2000.

[66] A. Hoecker, J. Stelzer, F. Tegenfeldt, H. Voss, K. Voss, A. Christov, S. Henrot-Versill, M. Jachowski, A. Krasznahorkay, Y. Mahalalel, X. Prudent, and P. Speckmayer, "Tmva - toolkit for multivariate data analysis with root: Users guide. oai:cds.cern.ch:1019880," Tech. Rep. CERN-OPEN-2007-007, CERN, Geneva, Mar 2007.

[67] CMS Collaboration, "Search for the higgs boson in the fully leptonic W^+W^- final state," no. CMS-PAS-HIG-11-003, 2011.

[68] CMS Collaboration, "Performance of muon identification in pp collisions at $\sqrt{s}$ = 7 TeV," CMS Physics Analysis Summary CMS-PAS-MUO-10-002, 2010.

[69] CMS Collaboration, "Particle–flow event reconstruction in CMS and performance for jets, taus, and $E_{\mathrm{T}}^{\mathrm{miss}}$," CMS Physics Analysis Summary CMS-PAS-PFT-09-001, 2009.

[70] CMS Collaboration, "Commissioning of the particle-flow reconstruction in minimum-bias and jet events from pp collisions at 7 TeV," *CMS Physics Analysis Summary*, vol. CMS-PAS-PFT-10-002, 2010.

[71] M. Cacciari and G. P. Salam and G. Soyez, "The anti-k_t jet clustering algorithm," *JHEP*, vol. 04, p. 063, 2008.

[72] S. Catani, Y. Dokshitzer, M. Seymour, and B. Webber, "Longitudinally-invariant kt-clustering algorithms for hadron-hadron collisions," *Nuclear Physics B*, vol. 406, no. 1-2, pp. 187 – 224, 1993.

[73] S. D. Ellis and D. E. Soper, "Successive combination jet algorithm for hadron collisions," *Phys. Rev. D*, vol. 48, pp. 3160–3166, Oct 1993.

[74] CMS Collaboration, "Determination of jet energy calibration and transverse momentum resolution in CMS," *JINST*, vol. 6, p. 11002, 2011.

[75] CMS collaboration, "CMS MET Performance in Events Containing Electroweak Bosons from pp Collisions at $\sqrt{s} = 7$ TeV," *CMS PAS JME-2010-005*, 2010.

[76] CMS Collaboration, "Algorithms for b jet identification in CMS," CMS Physics Analysis Summary CMS-PAS-BTV-09-001, 2009.

[77] CMS Collaboration, "Commissioning of b-jet identification with pp collisions at $\sqrt{s} = 7$ TeV," CMS Physics Analysis Summary CMS-PAS-BTV-10-001, 2010.

[78] P. Jenni, M. Nessi, M. Nordberg, and K. Smith, *ATLAS high-level trigger, data-acquisition and controls: Technical Design Report*. Technical Design Report ATLAS, Geneva: CERN, 2003.

[79] "Precision electroweak measurements on the z resonance," *Physics Reports*, vol. 427, no. 5-6, pp. 257 – 454, 2006.

[80] S. Agostinelli *et al.*, "GEANT4: A simulation toolkit," *Nucl. Instrum. Meth.*, vol. A506, p. 250, 2003.

[81] "Search for the standard model higgs boson in the H→WW(*)→lvlv decay mode with 4.7 fb^{-1} of atlas data at sqrt(s) = 7," tech. rep.

[82] S. Frixione, P. Nason, and C. Oleari, "Matching NLO QCD computations with parton shower simulations: the POWHEG method," *JHEP*, vol. 11, p. 070, 2007.

[83] J. M. Campbell and F. Tramontano, "Next-to-leading order corrections to Wt production and decay," *Nucl. Phys.*, vol. B726, pp. 109–130, 2005.

[84] C. D. White, S. Frixione, E. Laenen, and F. Maltoni, "Isolating Wt production at the LHC," *JHEP*, vol. 0911, p. 074, 2009.

[85] J. Alwall, P. Demin, S. de Visscher, R. Frederix, M. Herquet, F. Maltoni, T. Plehn, D. L. Rainwater, and T. Stelzer, "MadGraph/MadEvent v4: the new web generation," *JHEP*, vol. 09, p. 028, 2007.

[86] LHC Higgs Cross Section Working Group, S. Dittmaier, C. Mariotti, G. Passarino, and R. Tanaka (Eds.), "Handbook of LHC Higgs Cross Sections: 2. Differential Distributions," *CERN-2012-002*, CERN, Geneva, 2012.

[87] T. Binoth, M. Ciccolini, N. Kauer, and M. Krämer, "Gluon-induced W-boson pair production at the LHC," *JHEP*, vol. 12, p. 046, 2006.

[88] J. Campbell, R.K. Ellis, C. Williams, "Vector boson pair production at the LHC," *arxiv:hep-ph/1105.0020*, p. 028, 2011.

[89] J. M. Campbell and R. K. Ellis, "MCFM for the Tevatron and the LHC," *Nucl. Phys. Proc. Suppl.*, vol. 205-206, pp. 10–15, 2010.

[90] S. Frixione and B. R. Webber, "Matching NLO QCD computations and parton showers simulations," *JHEP*, vol. 0206, p. 029, 2002.

[91] R. C. Gray, C. Kilic, M. Park, S. Somalwar, and S. Thomas, "Backgrounds to higgs boson searches from asymmetric internal conversion," Tech. Rep. arXiv:1110.1368, Oct 2011. Comments: 16 pages, 5 figures.

[92] "Procedure for the LHC Higgs boson search combination in summer 2011," Tech. Rep. ATL-PHYS-PUB-2011-011, CERN, Geneva, Aug 2011.

[93] CMS Collaboration, "Absolute calibration of the CMS luminosity measurement: Summer 2011 update," CMS Physics Analysis Summary CMS-PAS-EWK-11-001, 2011.

[94] CMS Collaboration, "Absolute calibration of the luminosity measurement at cms: Winter 2012 update," no. CMS-PAS-SMP-12-008, 2012.

[95] H.-L. Lai, M. Guzzi, J. Huston, Z. Zi, P. Nadolsky, J. Pumplin, and C.-P. Yuan, "New parton distributions for collider physics," *Phys. Rev. D*, vol. 82, p. 074024, 2010.

[96] H.-L. Lai, J. Huston, Z. Zi, P. Nadolsky, J. Pumplin, D. Stump, and C.-P. Yuan, "Uncertainty induced by QCD coupling in the CTEQ global analysis of parton distributions," *Phys. Rev. D*, vol. 82, p. 054021, 2010.

[97] H.-L. Lai, M. Guzzi, J. Huston, Z. Li, P. M. Nadolsky, J. Pumplin, and C.-P. Yuan, "New parton distributions for collider physics," *Phys. Rev. D*, vol. 82, p. 074024, 2010.

[98] A. D. Martin *et al.*, "Parton distributions for the LHC," *Eur. Phys. J. C*, vol. 63, pp. 189–285, 2009.

[99] R. D. Ball *et al.*, "Impact of heavy quark masses on parton distributions and LHC phenomenology," *Nucl. Phys. B*, vol. 849, 2011.

[100] R. D. Ball, L. D. Debbio, S. Forte, A. Guffanti, J. I. Latorre, J. Rojo, and M. Ubiali, "A first unbiased global nlo determination of parton distributions and their uncertainties," *Nuclear Physics B*, vol. 838, no. 1-2, pp. 136 – 206, 2010.

[101] M. Botje, J. Butterworth, A. Cooper-Sarkar, A. de Roeck, J. Feltesse, S. Forte, A. Glazov, J. Huston, R. McNulty, T. Sjostrand, and R. Thorne, "The PDF4LHC Working Group Interim Recommendations," Tech. Rep. arXiv:1101.0538, Jan 2011. Comments: 12 pages.

[102] S. Alekhin, S. Alioli, R. D. Ball, V. Bertone, J. Blumlein, M. Botje, J. Butterworth, F. Cerutti, A. Cooper-Sarkar, A. de Roeck, L. Del Debbio, J. Feltesse, S. Forte, A. Glazov, A. Guffanti, C. Gwenlan, J. Huston, P. Jimenez-Delgado, H.-L. Lai, J. I. Latorre, R. McNulty, P. Nadolsky, S.-O. Moch, J. Pumplin, V. Radescu, J. Rojo, T. Sjostrand, W. J. Stirling, D. Stump, R. S. Thorne, M. Ubiali, A. Vicini, G. Watt, and C. P. Yuan, "The PDF4LHC Working Group Interim Report," Tech. Rep. arXiv:1101.0536, Jan 2011. Comments: 35 pages.

[103] I. Stewart and F. Tackmann, "Theory uncertainties for higgs and other searches using jet bins," *Phys. Rev. D*, vol. 85, p. 034011, 2012.

[104] G. Bozzia, S. Catani, D. de Florian, and M. Grazzini, "Transverse-momentum resummation and the spectrum of the Higgs boson at the LHC," *Nucl. Phys. B*, vol. 737, p. 73, 2006.

[105] J. Campbell, R.K. Ellis, G. Zanderighi, "Next-to-leading order Higgs + 2 jet production via gluon fusion," *JHEP*, vol. 0610, p. 028, 2006.

[106] C. Anastasiou, G. Dissertori, and F. Stoeckli, "NNLO QCD predictions for the H→WW→ $l\nu l\nu$ signal at the LHC," *Journal of High Energy Physics*, vol. 2007, no. 09, p. 018, 2007.

[107] C. Anastasiou, G. Dissertori, F. Stoeckli, and B. R. Webber, "QCD radiation effects on the H→WW→ $l\nu l\nu$ signal at the LHC," *Journal of High Energy Physics*, vol. 2008, no. 03, p. 017, 2008.

[108] T. Sjöstrand, S. Mrenna, and P. Skands, "Pythia 6.4 physics and manual," *JHEP*, vol. 05, p. 026, 2006.

[109] G. Corcella, I. G. Knowles, G. Marchesini, S. Moretti, K. Odagiri, P. Richardson, M. H. Seymour, and B. R. Webber, "HERWIG 6.5 release note," Tech. Rep. hep-ph/0210213. CAVENDISH-HEP-2002-17. CERN-TH-2002-270. DAMTP-2002-124. IPPP-2002-58, CERN, Geneva, Oct 2002.

[110] M. L. Mangano, M. Moretti, F. Piccinini, R. Pittau, and A. D. Polosa, "Alpgen, a generator for hard multiparton processes in hadronic collisions," *JHEP*, vol. 07, p. 001, 2003.

[111] A. O'Hagan, J.J. Forster, "Bayesian inference," *Kendalls Advanced Theory of Statistics, Arnold, London*, vol. 2B, 2004.

[112] A. L. Read, "Presentation of search results: the CLs technique," *J. Phys. G: Nucl. Part. Phys.*, vol. 28, p. 2693, 2002.

[113] T. Junk, "Confidence level computation for combining searches with small statistics," *Nucl. Instrum. Meth.*, vol. A434, p. 435, 1999.

[114] R. D. Cousins, J. T. Linnemann, and J. Tucker, "Evaluation of three methods for calculating statistical significance when incorporating a systematic uncertainty into a test of the background-only hypothesis for a poisson process," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 595, no. 2, pp. 480 – 501, 2008.

[115] K. Cranmer, "Statistical challenges for searches for new physics at the LHC," Tech. Rep. physics/0511028, Nov 2005.

[116] S. S. Wilks, "The large-sample distribution of the likelihood ratio for testing composite hypotheses," *The Annals of Mathematical Statistics*, vol. 9, no. 1, pp. pp. 60–62, 1938.

[117] Glen Cowan, Kyle Cranmer, Eilam Gross, and Ofer Vitells, "Asymptotic formulae for likelihood-based tests of new physics," *Eur.Phys.J.*, vol. C71, p. 1554, 2011.

[118] S. Chatrchyan *et al.*, "Combined results of searches for the standard model higgs boson in pp collisions at $\sqrt{s} = 7$ tev," *Physics Letters B*, vol. 710, no. 1, pp. 26 – 48, 2012.

[119] S. Chatrchyan *et al.*, "Search for the standard model higgs boson decaying to w^+w^- in the fully leptonic final state in pp collisions at $\sqrt{s} = 7$ tev," *Physics Letters B*, vol. 710, no. 1, pp. 91 – 113, 2012.

[120] E. Gross and O. Vitells, "Trial factors for the look elsewhere effect in high energy physics," *Eur.Phys.J*, vol. C70, p. 525, 2010.