FERMILAB-CONF-12-613-CMS-PPD

Observation of a Higgs-like Boson in CMS at the LHC

Pushpalatha C. Bhat

Fermi National Accelerator Laboratory
Batavia, IL 60510, USA.

for the CMS Collaboration

Abstract

After decades of pursuit at lepton and hadron colliders around the world, a Higgs-like boson has been observed at the Large Hadron Collider (LHC). We present results of searches by the CMS collaboration for the production of the standard model Higgs boson in proton-proton collisions, using data samples corresponding to integrated luminosities of 5.1 fb^{-1} at $\sqrt{s}$ = 7 TeV and 5.3 fb^{-1} at $\sqrt{s}$ = 8 TeV. The searches have been performed in five Higgs decay modes, $\gamma\gamma$, ZZ, WW, $\tau^+\tau^-$, and $b\bar{b}$, in the mass range from 110 up to 160 GeV. Multivariate methods have been used in many aspects of the analyses to extract optimal results. An excess of events has been observed above the expected background, with a significance of 5.0 standard deviations. Most of the excess is in the high resolution $\gamma\gamma$ and ZZ decay modes. A fit to the signal in these modes yields a mass for the boson of 125.3±0.4(stat.)±0.5(syst.) GeV. Since the new boson decays into two photons, its spin can be inferred to be different from one.

Keywords:
Higgs boson, Higgs-like boson, CMS experiment, Multivariate methods, Observation, LHC

1. Introduction

The Higgs mechanism [1] is a vital part of the Standard Model (SM) [2] of particle physics that provides for electroweak symmetry breaking (EWSB), endowing the W and Z bosons with mass while the photon remains massless. Inclusion of the Higgs mechanism in the SM via a weak isospin doublet of fundamental scalar Higgs field also implies the existence of a spin zero particle, the Higgs boson. Each of the quarks and leptons would interact with the Higgs field with a characteristic Yukawa coupling, thereby acquiring a mass. Since the discovery of the top quark in 1995, the pursuit of this agent of EWSB has been the top priority in experimental high energy physics.

While the SM does not predict the mass of the Higgs boson (m_{H}), it is expected, from general arguments, to be below 1 TeV, and the precision electroweak measurements imply $m_{\mathrm{H}} < 152$ GeV at 95% C.L. [3]. Direct searches at the CERN electron-positron collider (LEP) yielded $m_{\mathrm{H}} > 114.4$ GeV at 95% C.L.[4]. Significant upgrades to the Fermilab Tevatron collider complex [5] over the past decade provided proton-antiproton collision data in excess of 10 fb^{-1} at $\sqrt{s}$ =1.96 TeV to the CDF and D0 experiments. The two experiments together had excluded the SM Higgs boson mass range of 162–166 GeV at 95% C.L. [6], and recently reported an excess at low masses [7].

One of the primary scientific goals of the Large Hadron Collider (LHC) at CERN has been to discover the Higgs boson, if it exists. The experiments at the LHC have been collecting collision data since 2010, at $\sqrt{s}$= 7 TeV through 2011 and at 8 TeV in 2012. Using the 2011 datasets, the CMS experiment had excluded a range of masses from 127 to 600 GeV while the ATLAS experiment excluded the ranges 114.4–116.6, 119.4–122.1 and 129.2–541 GeV at 95% C.L. [8]. An excess of events in the SM Higgs boson searches was reported in December 2011 by both CMS and ATLAS experiments, around a mass of 125 GeV. The Tevatron experiments have reported [7] an excess of events in a broad mass range of 120–135 GeV, in WH and ZH channels with $\mathrm{H} \rightarrow b\bar{b}$ decay. Here we present the searches for the SM Higgs boson by the CMS experiment, using 5.1 fb^{-1} at $\sqrt{s}$= 7 TeV and 5.3 fb^{-1} at $\sqrt{s}$= 8 TeV that led to the observation of a new Higgs-like boson

Email address: pushpa@fnal.gov (Pushpalatha C. Bhat)

Operated by Fermi Research Alliance, LLC under Contract No. De-AC02-07CH11359 with the United States Department of Energy.

with a mass of ~ 125 GeV, reported on July 4, 2012 at CERN [9].

2. Higgs Production and Decay

In proton-proton collisions at the LHC, the main processes for the production of SM Higgs bosons are: gluon-gluon fusion with the largest cross section (gg → H via heavy quark loops); $qq' \to qq'H$ via vector boson (WW or ZZ) fusion (VBF); associated production $q\bar{q}' \to WH, ZH$; and $gg \to t\bar{t}H$ via $t\bar{t}$ fusion ($t\bar{t}H$). For m_H=125 GeV, the cross sections [10] for the gluon-gluon fusion process at $\sqrt{s}$= 7 and 8 TeV are about 15.3 and 19.5 pb, respectively; those for the VBF and WH+ZH are about an order of magnitude smaller and $t\bar{t}H$ more than two orders of magnitude smaller.

The decay modes in which to search for the SM Higgs boson very much depends on its mass. CMS has carried out searches in the Higgs decay modes of $\gamma\gamma$, ZZ, WW, $\tau\tau$, and $b\bar{b}$. Of these, the first three are the most sensitive decay modes in the low m_H range of 100–160 GeV. The $H \to \gamma\gamma$, $H \to ZZ \to \ell^+\ell^-\ell^+\ell^-$ (or 4ℓ) decays allow reconstruction of the mass of the parent particle, the Higgs boson, with very high resolution. These two channels, in fact, contribute most significantly to the discovery and are described in some detail here. Results with all five channels combined are also reported.

3. The LHC and the CMS Detector

The Large Hadron Collider (LHC) [11] at CERN, is installed in a 27 km circular ring, 100 m underground, straddling the Franco-Swiss borders near Geneva. The LHC accelerates two beams of protons in clockwise and anti-clockwise directions and collides them in the designated experimental collision points. It has been providing excellent operational performance since March 2010, with center of mass energy of collisions of 7 TeV during 2010-11 and 8 TeV in the 2012 run. Each beam, currently, consists of 1368 colliding bunches with bunch intensity $\sim 1.5 \times 10^{11}$ protons. The highest peak luminosity achieved to date is 7.5×10^{33} cm^{-2}sec^{-1}.

The CMS detector is one of the two large general purpose detectors at the LHC, designed and built to have capabilities to detect Higgs bosons in several decay channels. The CMS detector comprises the following major detector systems: an all-silicon pixel and strip tracker, a lead tungstate crystal electromagnetic calorimeter (ECAL) and a brass/scintillator sampling hadron calorimeter (HCAL), all housed within a superconducting solenoid of 6 m internal diameter providing 3.8 T axial field; a muon spectrometer with gas-ionization chambers embedded in the steel flux-return yoke, and a forward Cherenkov calorimeter based on steel absorber and quartz scintillation fibers. The HCAL barrel and endcaps cover the pseudorapidity range $|\eta| < 3.0$, where $\eta = -\ln[\tan(\theta/2)]$, and θ is the polar angle. The hadronic forward (HF) calorimeters extend the calorimeter coverage up to $|\eta|$= 5.0. Muons are measured in the range $|\eta| < 2.4$, with detection planes based on three technologies: drift tubes ($|\eta| < 1.2$), cathode strip chambers ($0.9 < |\eta| < 2.4$), and resistive plate chambers ($|\eta| < 1.6$). The CMS detector is described in detail elsewhere [12].

The two-level CMS trigger system ensures that potentially interesting events are recorded with high efficiency. While the design collision rate is ~ 32 MHz, events are recorded at the rate of around 0.5 kHz, for offline reconstruction and analysis. A high performance and highly distributed computing system is used to execute numerous tasks, including the reconstruction and analysis of the collected data, as well as the generation and detailed detector simulation of Monte Carlo (MC) event samples.

4. Event Reconstruction

The CMS "Particle-Flow" algorithm [13] is used to optimally combine all of the sub-detector information to reconstruct and identify the final state particles or objects such as photons, electrons, muons, charged and neutral hadrons.

Charged particles are tracked within the pseudorapidity range $|\eta| < 2.5$. The tracker has a track-finding efficiency larger than 99% for muons with transverse momentum $p_T > 1$ GeV, a transverse momentum resolution between 2 and 3% for charged tracks of $p_T \approx 100$ GeV in the central region ($|\eta| < 1.5$), and a very good b-vertex-tagging capability.

Jets are reconstructed by clustering the particle-flow objects, using an anti-k_T algorithm [14]. The jet energy is corrected for pile-up (presence of multiple events in each crossing of the bunches) and underlying event effects using the η-dependent transverse momentum density, calculated on an event-by-event basis. Jet energy corrections are between 5% and 10% of the true momentum over the whole p_T spectrum. The jet momentum resolution is $\sigma(p_T) = 85\%/\sqrt{p_T/GeV} \bigoplus 4\%$ for central jets. A b-jet tagging efficiency (using secondary vertex, tracking and kinematic information) of more than 50% is achieved with a rejection factor of ≈ 200 for light-quark jets, as measured in $t\bar{t}$ events in data.

The fine-grained hermetic ECAL (along with Lead/Silicon preshower detectors in the endcaps) pro-

vide excellent resolution for the measurement of electromagnetic showers. Photons and electrons are reconstructed by clustering energy deposits in ECAL and requiring the shapes in η and ϕ (azimuth) to conform to the expected shower shapes. For electrons, ECAL and tracker measurements are combined. The diphoton mass resolution in the central barrel region is estimated to be 1.1 GeV at a mass of 125 GeV. In the 7 TeV dataset, the dielectron mass resolution at the Z boson mass is 1.56 GeV in the barrel and 2.57 GeV in the endcaps, while for the 8 TeV data set, the corresponding values are 1.61 GeV and 3.75 GeV. Muons are reconstructed using tracks in the silicon tracker and the muon systems. The efficiency to reconstruct a muon with $p_T >$ 5 GeV is larger than 95%. The dimuon mass resolution at the Υ mass, dominated by instrumental effects, is measured to be 0.6% in the barrel region.

More specialized identification and energy corrections are applied to the objects specific to a particular analysis, and they are discussed in the appropriate sections that follow.

5. Searches for the Standard Model Higgs Boson

The sensitivity of a search in a decay channel, for a given Higgs mass, depends on the cross section times branching fraction, the signal selection efficiency, signal to background ratio after candidate selection, and the mass resolution. The cross sections and branching ratios along with predicted uncertainties used in the present study are those compiled in Ref. [10]. Samples of simulated events for signal and background processes were used to study signal and background efficiencies as well as a number of other aspects of the analyses. The signal MC samples for most production processes and decay modes were generated using the next-to-leading-order (NLO) matrix-element generator code POWHEG [15] interfaced with PYTHIA 6.4 program [16] for fragmentation and parton showering. In some cases, the PYTHIA code was directly used to generate the signal samples. Signal samples for $\mathrm{H} \to \mathrm{b\bar{b}}$ were produced by interfacing POWHEG with HERWIG++[17]. The MC events for background studies were generated using POWHEG, PYTHIA and MADGRAPH [18]. The MC events were processed through full detector simulation using GEANT4 [19].

Multivariate analysis (MVA) methods [20] have been extensively used in many aspects of the analyses, in all channels, to extract maximal information from data and obtain the best possible physics results. MVA methods are used in electron and photon identification, object energy corrections, vertex association, and signal/background discrimination. In most instances, Boosted Decision Tree (BDT) methods implemented in the TMVA software package [21] are used.

The statistical methodology for combining results from different channels to obtain exclusion limits and p-values, was developed by the CMS/ATLAS Higgs combination group [22]. The exclusion limits are calculated using the modified frequentist criterion $\mathrm{CL_s}$. The presence of a signal is quantified in terms of the background-only p-value, which is the probability for the expected background to fluctuate upwards to give an excess of events as large as, or larger than, the observed number of events. The p-value found at a given m_{H} is called the local p-value and the probability for such an excess anywhere in a specified mass range is called the global p-value. Since there is a finite probability of finding an excess from random fluctuations anywhere in the mass range, the global p-value will be larger than the local p-value – this effect is called the look elsewhere effect (LEE) and has to be considered for the interpretation of the final results. The local and global p-values are expressed in terms of standard deviations (σ) using the one-sided Gaussian tail distribution. The magnitude of the signal is denoted by the signal strength, the observed cross section relative to the SM expectation, $\sigma/\sigma_{\mathrm{SM}}$.

5.1. $\mathrm{H} \to \gamma\gamma$ *Channel*

The analysis in this channel is essentially a bump hunt in the diphoton invariant mass spectrum in the mass range of 110–150 GeV on a large falling background from QCD production of two photons and from Z+X events (X=jets and/or other objects) due to instrumental effects where jet fragments fake one or more of the photons. Because of the excellent electromagnetic resolution for measuring photons, we expect to see a narrow peak from the Higgs boson decay into two photons.

The photon energy is reconstructed by summing appropriate cluster energies in ECAL; energy in the preshower detectors, where available, for $|\eta| > 1.65$ is added. Corrections to the raw energy due to shower containment and losses due to conversions in the material upstream of the calorimeter are derived using a multivariate regression BDT trained with variables characterizing the energy clusters and shower shape. A multivariate BDT discriminant is also used in the selection of photons to suppress background from jets faking photons. The calibration of the energy scale, photon selection efficiency, energy resolution and associated uncertainties are estimated using $\mathrm{Z} \to \mathrm{ee}$ data events, where the track information is ignored.

The event selection requires two high p_T, isolated photons with $|\eta| < 2.5$ (excluding $1.44 < |\eta| < 1.57$,

the barrel-endcap transition region). For the leading and sub-leading photons, $p_T > m_{\gamma\gamma}/3$ and $p_T > m_{\gamma\gamma}/4$ are required where $m_{\gamma\gamma}$ is the invariant mass of the photon-pair. In the VBF events with two tagged forward jets in addition to two photons, the p_T threshold for the leading photon is raised to $m_{\gamma\gamma}/2$, for better background suppression. The two jets are required to have p_T greater than 30 and 20 GeV, $\Delta\eta > 3.5$, and the dijet invariant mass greater than 350 (250) GeV for 7 (8) TeV data sets. In case of the 8 TeV data set, a second "tight" dijet category is created where both jets are required to have $p_T > 30$ GeV and dijet invariant mass, $m_{jj} > 500$ GeV. In addition, the difference in pseudorapidity and the azimuthal angle between the dijet and the diphoton system are required to be less than 2.5 units and 2.6 radians, respectively. Since there are many interactions per bunch crossing, the association of the photon-pair to the correct primary vertex is made using a BDT discriminant trained on the kinematics of tracks and their correlation with the diphoton kinematics.

To enhance the sensitivity of the search, the diphoton candidate events are split into several mutually exclusive categories and fits to the diphoton invariant mass spectra are performed separately in each of these categories. The splitting is based on a BDT multivariate discriminant built using the following inputs: i) quality of the photons as determined by the electromagnetic shower shape and isolation, ii) expected mass resolution, iii) probability of locating the diphoton vertex within 10 mm of its true vertex, and iv) kinematics of photons and the diphoton system. The input variables are constructed so as to contain no information about the diphoton invariant mass. The training of the BDT is performed using signal MC events and prompt diphoton background events. The splitting into five categories using the MVA discriminant is optimized for best sensitivity to find the SM Higgs boson signal. Events in the category with smallest S/B (which corresponds to small discriminant values) are rejected and events in the remaining four categories are kept. The VBF category events are also required to pass the minimal discriminant cut.

A signal plus background fit is performed to the diphoton invariant mass distribution in each of the categories in the range $100 < m_{\gamma\gamma} < 180$ GeV. The background is fit to a polynomial function of degree 3–5 (the degree chosen to minimize bias in the shape). The diphoton invariant mass distribution with each event weighted by the S/(S+B) value of its category is shown in Fig. 1, where S and B are the number of signal and background events, respectively, calculated from fits, integrated over a window of $2\sigma_{eff}$, around 125 GeV. The inset of the figure shows the combined, unweighted invariant mass spectrum in a smaller range around 125 GeV. A clear and significant signal peak can be seen in both plots.

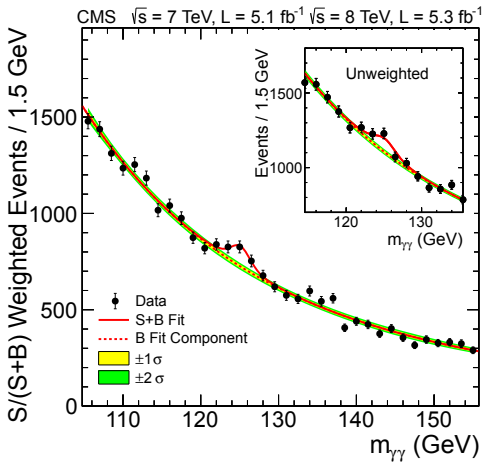

Figure 1: The diphoton invariant mass distribution with each event weighted by the S/(S+B) value of its category. The lines represent the background and background+signal fits, and the bands represent the $\pm 1\sigma$ and $\pm 2\sigma$ uncertainties in the background estimate. The inset shows the unweighted invariant mass distribution in the region of the signal peak.

Table 1 shows the number of signal and background events from fitting of the $m_{\gamma\gamma}$ distributions in each category. The results from the simultaneous fit in all the eleven categories are used to extract confidence level for exclusion or discovery of a SM Higgs boson. The expected 95% C.L. upper limits on the signal strength in the range $110 < m_{\mathrm{H}} < 140$ GeV for the combined data set are below 1.0 (0.76 at $m_{\mathrm{H}} = 125$ GeV). The observed limit indicates the presence of a significant excess at $m_{\mathrm{H}} \approx 125$ GeV.

At m_{H}=125 GeV, the local p-value corresponds to a signal significance of 4.1σ. (The local p-values as a function of m_{H} are shown in Fig. 7.) The best-fit signal strength at 125 GeV is $\sigma/\sigma_{\mathrm{SM}} = 1.6 \pm 0.4$.

5.2. $\mathrm{H} \to \mathrm{ZZ} \to 4\ell$ *Channel*

In this channel, as in the case of $H \to \gamma\gamma$, we expect to see a narrow peak in the invariant mass spectrum of the final state objects, from Higgs decay into two Z bosons, each of which decay into two same flavor, opposite sign, leptons. The search for these decays in the m_{H} range of 110–160 GeV in 4e, 4μ and $2e2\mu$ final states are presented here. The dominant background contributions come from the direct ZZ production via quark-antiquark annihilation or gluon-gluon fusion. Other background

Table 1: Expected number of fitted signal events at m_{H}=125 GeV and background events (per GeV, at $m_{\gamma\gamma}$=125 GeV), in the eleven categories in the $\mathrm{H} \to \gamma\gamma$ channel. See text for more details.

	7 TeV (5.1 fb^{-1})		8 TeV (5.3 fb^{-1})	
Categories	Signal	Background	Signal	Background
Cat. 0	3.2	3.3±0.4	6.1	7.4 ± 0.6
Cat. 1	16.3	37.5±1.3	21.0	54.7 ± 1.5
Cat. 2	21.5	74.8±1.9	30.2	115.2 ± 2.3
Cat. 3	32.8	193.6±3.0	40.0	256.5 ± 3.4
Dijet Tag	2.9	1.7±0.2	-	-
Dijet Tight	-	-	2.6	1.3 ± 0.2
Dijet Loose	-	-	3.0	3.7 ± 0.4

processes include $\mathrm{Z} + \mathrm{b\bar{b}}$ and $\mathrm{t\bar{t}}$ events with four lepton final states as well as Z+jets and WZ+jets where jets might be misidentified as leptons.

The event selection requires two pairs of same flavor, oppositely charged leptons – isolated electrons with $p_T > 7$ GeV and $|\eta| < 2.5$ and isolated muons with $p_T > 5$ GeV and $|\eta| < 2.4$ are selected. A multivariate classifier trained using W+jets events is used to select electrons. The combined efficiencies for electrons and muons are measured using Z boson events. For muons with $p_T < 15$ GeV, the efficiencies are measured using J/ψ decays. Any photon in the vicinity of a selected lepton, that could be from a final state radiation (FSR) is included in the computation of lepton-lepton pair invariant mass. One of the two lepton pairs is required to have a mass in the range of 40–120 GeV and the other pair in the range of 12–120 GeV.

The distribution of four lepton invariant mass of the selected events is shown in Fig. 2. The background from direct ZZ (and $\mathrm{Z}\gamma^*$) production is obtained from MC simulations. Other backgrounds are estimated using data. Control background samples of data events with like-sign leptons, or with unlike-sign leptons which fail one or more lepton selection criteria, are used to measure event rates in the background region. These rates are then folded with the measured probability for a reconstructed lepton to pass the required criteria to estimate background in the signal region. The $m_{4\ell}$ distributions from background contributions are shown by the filled histograms in Fig. 2. A clear peak in data above the expected background can be seen. The expected signal peak for a 125 GeV Higgs is shown by the open (red) histogram. There is also a peak at the Z boson mass from $\mathrm{Z} \to 4\ell$ decays. The expected number of background and signal events and selected data events are shown in Table 2.

Again, to enhance the sensitivity of the analysis, we resort to a multivariate analysis of the data. A multivariate discriminant (MELA, for Matrix Element Likelihood Analysis) is built using five angle variables and invariant masses of the two pairs of leptons. These variables fully describe the kinematics of the Higgs decay mode to 4ℓ and are used to construct a kinematic discriminant, $K_D = P_{sig}/(P_{sig} + P_{bkg})$ [23] ($\mathrm{q\bar{q}} \to \mathrm{ZZ/Z}\gamma$ used for background).

Fig. 3 shows the distribution of K_D vs. $m_{4\ell}$ for data events selected in the analysis superposed on the expected signal density for m_{H} =125 GeV. A clustering of data events can be seen around 125 GeV with high values of K_D, as expected for signal events. The $m_{4\ell}$ distribution of data events with $K_D > 0.5$ is shown in the inset of Fig. 2.

Two dimensional likelihood fits are performed for data in the three final states for the 7 TeV and 8 TeV data sets for different values of m_H in the considered mass range. The expected 95% C.L. upper limit on the signal strength, $\sigma/\sigma_{\mathrm{SM}}$, in the background only hypothesis, has a value of 0.6 at m_{H} =125 GeV. The observed limits indicate a significant excess in the range $120 < m_{\mathrm{H}} < 130$ GeV. The minimum local p-value, at m_{H} =125.6 GeV corresponds to a signal significance of 3.2σ (expected 3.8σ). The best-fit signal strength at 125.6 GeV is $\sigma/\sigma_{\mathrm{SM}} = 0.7^{+0.4}_{-0.3}$.

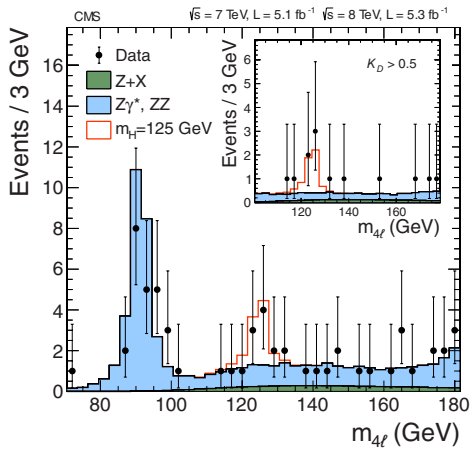

Figure 2: The four-lepton invariant mass distribution from the $\mathrm{H} \to \mathrm{ZZ} \to 4\ell$ channel analysis. The inset shows the invariant mass distribution for events with the MELA discriminant K_D >0.5.

5.3. Low Resolution Channels

The $\mathrm{H} \to \mathrm{WW}$ decay mode, with both W's decaying leptonically has very good sensitivity to find a SM Higgs boson signal down to 120 GeV mass, but has poor mass resolution. For Higgs search in this channel, events with two isolated leptons (electrons and

Table 2: Expected number of background and signal events and number of observed events in H → ZZ → 4ℓ channel. The number of background and data events in the signal region (121.5< $m_{4\ell}$ <130.5 GeV) are also shown.

Channel	$4e$	4μ	$2e2\mu$	4ℓ
Background	4.0±1.0	6.6±0.9	9.7±1.8	20±3
Data	6	6	9	21
Signal ($m_{\rm H}$=125 GeV)	1.36±0.22	2.74±0.32	3.44±0.44	7.54±0.78
Background (signal region)	0.7±0.2	1.3±0.1	1.9±0.3	3.8± 0.5
Data (signal region)	1	3	5	9

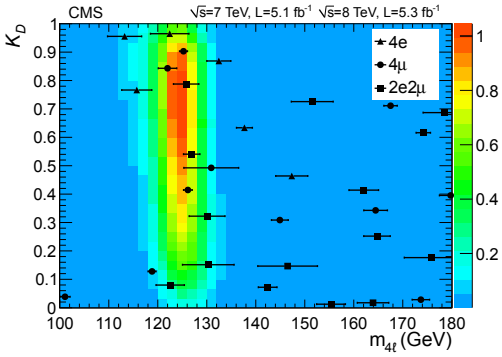

Figure 3: The MELA discriminant vs the four-lepton invariant mass in the H → ZZ → 4ℓ channel. Data events are shown as symbols (defined in the legend) superposed on the signal density expected for 125 GeV Higgs boson events.

muons) of p_T > 20 and 10 GeV, and up to two jets with p_T > 30 GeV and large missing transverse energy (E_T^{miss}) are used. Event selection criteria are further optimized separately for different event categories to suppress different dominating backgrounds.

The dominant background for the most sensitive zero-jet eμ category comes from nonresonant WW production, and the next largest contribution comes from the W+jets process with a fake lepton, followed by top quark events. In ee and $\mu\mu$ event categories, the background mainly comes from diboson and Drell-Yan production. The one-jet and two-jet categories have backgrounds from top, WW and W+jets events.

To suppress dominant non-resonant WW background, either a multivariate selection (applied in 7 TeV data analysis) or cuts on single variables are used. The dilepton invariant mass, $m_{\ell\ell}$, is one the best discriminants against this background. A multivariate selection is also applied to reduce Drell-Yan background in ee and $\mu\mu$ channels. To suppress top-quark background, b-tagging information is used. The backgrounds in data, after the event selection, are estimated using control samples of data enriched in relevant backgrounds.

In the most sensitive zero-jet $e\mu$ category, 124.2±12.4 events are expected from background, 23.9±5.2 events from signal with m_H=125 GeV, and 158 data events are observed. A plot of the $m_{\ell\ell}$ distributions for 8 TeV data, background and signal is shown in Fig. 4. Combining all categories, a broad excess above the expected 95% C.L. limits is seen in the mass range 110–160 GeV. The observed excess is consistent with a SM Higgs boson with a mass around 125 GeV. The expected significance at 125 GeV is 2.4σ, while the observed significance is 1.6σ.

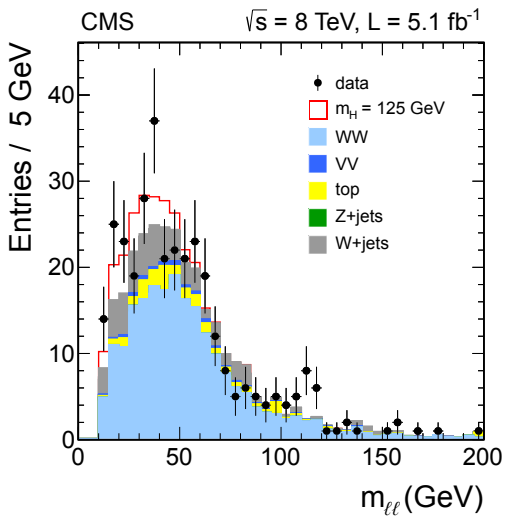

Figure 4: The distribution of dilepton invariant mass for data events ($\sqrt{s}$= 8 TeV) compared with background and signal+background distributions in the H → WW → $\ell\nu\ell\nu$ channel analysis (zero-jet eμ category).

The H → $\tau\tau$ decay mode is analyzed in four exclusive sub-channels: eμ, $\mu\mu$, e$\tau_{\rm h}$, and $\mu\tau_{\rm h}$, where electrons and muons come from leptonic τ decays, and τ_h denotes hadronic decays of τ (reconstructed as three prong or one prong decays). The search is carried out in the mass range of 110–145 GeV, and a signal, if any, is expected to show up as a broad excess in the distribution of the invariant mass of the τ-pair, $m_{\tau\tau}$. All backgrounds – $\bar{\rm t}$t and Drell-Yan in the eμ and $\mu\mu$ channels, and W and Z production with fake $\tau_{\rm h}$ candidates in the hadronic channels, are estimated using data. The expected median

95% C.L. limits on $\sigma/\sigma_{\rm SM}$ is pretty flat (1.3–1.6) in the Higgs mass range 110–130 GeV. The values of the expected and observed limits are 1.3 and 1.1, respectively, for a mass of 125 GeV. The expected significance for a SM Higgs boson of mass 125 GeV is 1.4σ. However, no excess is observed in this channel.

For $m_{\rm H} < 135$ GeV, the H $\to$ b$\bar{\rm b}$ decay has the largest branching fraction of the five decay modes considered, but it is overwhelmed by background from QCD production of b−quarks. Therefore, the strategy in the analysis is to look for associated production of Higgs boson with a W or a Z boson. The following channels are explored: ZH $\to$ ee b$\bar{\rm b}$ or $\mu\mu$b$\bar{\rm b}$, ZH $\to \nu\nu$b$\bar{\rm b}$ and WH $\to$ eνb$\bar{\rm b}$ or $\mu\nu$b$\bar{\rm b}$. Dominant backgrounds arise from W/Z+jets, t$\bar{\rm t}$, single top-quark, and diboson production (with one of the bosons decaying hadronically). Significant background rejection is achieved by requiring a large p_T for the dijet, minimal additional jet activity, and the vector boson to be back-to-back with the dijet in azimuth. Further discrimination is obtained using a BDT trained on kinematic and topological variables. The expected 95% C.L. limit on $\sigma/\sigma_{\rm SM}$ ranges from 1.2 at m_H=110 GeV to 2.8 at 135 GeV. The expected significance at m_H=125 GeV is 1.9σ while the observed significance is 0.7σ.

5.4. Combined Results

The methodology for combination of results from various channels and their interpretation is discussed elsewhere as noted at the beginning of the section. The combination takes into account the statistical and systematic uncertainties in measurements as well as theoretical uncertainties such as those on QCD scale and parton distribution functions. The observed CL$_{\rm s}$ values from the combination of all five channels are shown in Fig. 5 along with the expected median values and the 68% and 95% C.L. bands for background-only hypothesis. These and our previous search results [8] exclude a SM Higgs boson in the mass ranges of 110–121.5 GeV and 128–600 GeV with a confidence level of 95% or higher. However, the range 121.5 $< m_{\rm H} <$ 128 GeV cannot be excluded at 95% C.L. due to the observed excess of events. The observed local p-values as a function of Higgs boson mass hypothesis for $\gamma\gamma$ and ZZ channel results combined, and for all five channels combined, are shown in Figs. 6 and 7, respectively. The combined significance of the $\gamma\gamma$ and ZZ channels is about 5.0σ at $m_{\rm H}$= 125.5 GeV. When the WW channel is added, the significance becomes 5.1σ. The combination of all five channels gives an observed significance of 5.0σ at a mass of 125.5 GeV, while the expected significance is 5.8σ. The global p-value calculated by correcting for LEE in the search range of 115–130 (110–145) GeV corresponds to a significance of 4.6σ (4.5σ). An unconstrained fit to the signal in the $\gamma\gamma$, ZZ$\to$ 4ℓ channels, yields a best-fit mass value for the boson of 125.3±0.4(stat.)±0.5(syst.) GeV. The best-fit value for $\sigma/\sigma_{\rm SM}$ vs. $m_{\rm H}$ with all five channels combined is shown in Fig. 8; the value at 125.5 GeV is 0.87±0.23.

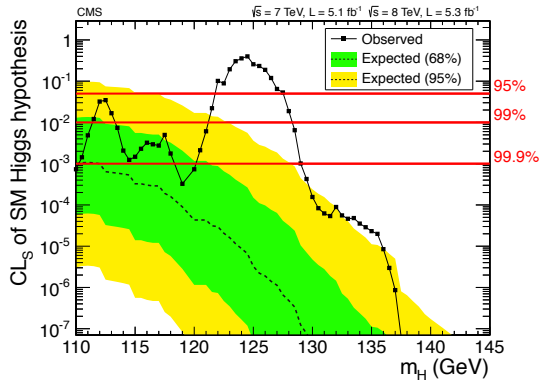

Figure 5: The CL$_{\rm s}$ values for the SM Higgs boson hypothesis for the five channels reported. For background-only hypothesis, the median expected values (dashed line) and the 68% and 95% C.L. bands are shown. The observed values are shown as solid points. The horizontal lines indicate CL$_{\rm s}$ values of 0.05, 0.01 and 0.001, corresponding to 95%, 99% and 99.9% C.L. exclusion limits for mass regions below the observed CL$_{\rm s}$.

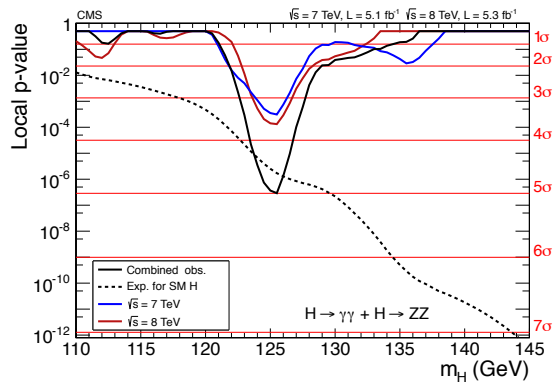

Figure 6: The observed local p-values for H $\to \gamma\gamma$ and H $\to$ ZZ $\to 4\ell$ channels combined. The dashed line shows the expected p-values for the combined dataset.

6. Conclusions

A new Higgs-like boson, with a mass around 125 GeV, has been observed in searches for the standard model Higgs boson in proton-proton collision data collected with the CMS detector at the LHC. The searches in decay modes $\gamma\gamma$, ZZ, WW, $\tau\tau$ and b$\bar{\rm b}$, using data corresponding to 5.1 fb^{-1} at $\sqrt{s}$=7 TeV and 5.3 fb^{-1} at

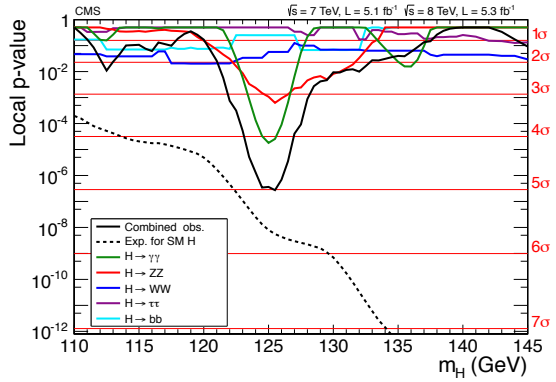

Figure 7: The observed local p-values for the five individual channels analyzed, and for the combination, as a function of the SM Higgs boson mass. The dashed line shows the expected p-values for the combination.

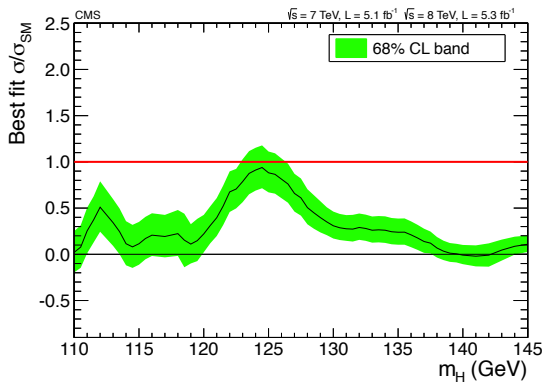

Figure 8: The best-fit production cross section for the observed boson relative to the predicted SM cross section from the analyses of five channels combined.

$\sqrt{s}$=8 TeV, yield a combined signal significance of 5.0σ, while the expected significance is 5.8σ. The high mass resolution $\gamma\gamma$ and ZZ channels show significant signal peaks in the $m_{\gamma\gamma}$ and $m_{4\ell}$ invariant mass distributions, respectively. A fit to these signals yields a mass for the new boson to be 125.3±0.4(stat.)±0.5(syst.) GeV. A broad excess consistent with this mass is also seen in the WW channel. The best-fit value for the production cross section relative to SM, using results from all five channels, is $\sigma/\sigma_{\rm SM}$=0.87±0.23 at $m_{\rm H}$=125.5 GeV. The decay to two photons indicates that the spin of the new boson is different from one. The results, so far, are consistent with the expectations for a standard model Higgs boson. With more data being collected at the LHC this year, it should be possible to begin to elucidate the nature of this newly observed Higgs-like particle – whether it has all the properties expected of the standard model Higgs boson or is a harbinger of new physics beyond the standard model.

Acknowledgements

We would like to congratulate and thank our colleagues in the CERN accelerator departments for the excellent performance of the LHC. We acknowledge the support from all the funding agencies that contributed to the construction and operation of the CMS detector. The author's work is supported by the Fermi Research Alliance, LLC under Contract No. DE-AC02-07CH11359 with the U.S. Department of Energy.

References

[1] F. Englert and R. Brout, Phys. Rev. Lett. **13** (1964) 321; P.W. Higgs, Phys. Lett. **12** (1964) 132; *ibid* Phys. Rev. Lett. **13** (196) 508; G.S. Guralink, C.R. Hagen, and T.W.B. Kibble, Phys. Rev. Lett. bf 13 (1964) 585; P.W. Higgs, Phys. Rev. **145** (1966) 1156; T.W.B. Kibble, Phys. Rev. **155** (1967) 1554.

[2] S.L. Glashow, Nucl. Phys. **22** (1961) 579; S. Weinberg, Phys. Rev. Lett. **19** (1967) 1264; A. Salam, in *Elementary Particle Physics: relativistic groups and analyticity*, N. Svartholm, ed., p. 367, Almquist & Wiskell, (1968).

[3] ALEPH, CDF, D0, DELPHI, L3, OPAL, SLD Collaborations, Working groups report, CERN PH-EP-2010-095 (2010); http://lepewwg.web.cern.ch/LEPEWWG/plots/winter2012/.

[4] ALEPH, DELPHI, L3, OPAL Collaborations, and LEP working group for Higgs boson searches, Phys. Lett. **B565** (2003) 61.

[5] P.C. Bhat and W.J. Spalding, AIP Conf. Proc. **753** (2005) 30.

[6] CDF and D0 Collaborations, Phys. Rev. Lett. **104** (2010) 061802.

[7] CDF and D0 Collaborations, and the Tevatron New Physics and Higgs working group, arXiv:1207.0449; CDF and D0 Collaborations, Phys. Rev. Lett. **109** (2012) 071804.

[8] CMS Collaboration, Phys. Lett. **B710** (2012) 26; ATLAS Collaboration, Phys. Rev. **D86** (2012) 032003.

[9] The CMS Collaboration, Phys. Lett. **716** (2012) 30;The ATLAS Collaboration, Phys. Lett. **716** (2012) 1.

[10] LHC Higgs Cross Section Working Group, "Handbook of LHC Higgs Cross Sections:1. Inclusive Observables", CERN, Geneva (2011), arXiv:1101.0593.

[11] L. Evans and P. Bryant (Editors), "LHC Machine", JINST, **3** (2008) S08001.

[12] CMS Collaboration, "The CMS Experiment at the CERN LHC", JINST, **3** (2008) S08004.

[13] CMS Collaboration, CMS Physics Analysis Summary, CMS-PAS-PFT-10-001 (2010).

[14] M. Cacciari and G.P. Salam, JHEP **04** (2008) 063.

[15] S. Alioli, P. Nason, C. Oleari, et al., JHEP **04** (2009) 002.

[16] T. Sjostrand, S. Mrenna, and P.Z. Skands, JHEP **05** (2006) 026.

[17] S. Gieseke, D. Grellscheid, K. Hamilton, et al., arXiv:hep-ph/0609306 (2006).

[18] J. Alwall, P. Demin, S. de Visscher et al., JHEP **09** (2007) 028.

[19] GEANT4 Collaboration, Nucl. Instrum. Meth. A **A 506** (2003) 250.

[20] P.C. Bhat, Annu. Rev. Nucl. Part. Sci., **61** (2011) 281; H.B. Prosper, PoS ACAT08 (2008) 010.

[21] H. Voss, et al., PoS ACAT07 (2007) 040.

[22] ATLAS and CMS Collaborations, Higgs Combination Group, ATL-PHYS-PUB 2011-11, CMS NOTE 2011/005 (2011).

[23] CMS Collaboration, JHEP **04** (2012) 036.