CMS Physics Analysis Summary

Contact: cms-pag-conveners-higgs@cern.ch 2011/07/22

Search for the standard model Higgs Boson in the decay channel $H \to ZZ \to \ell^-\ell^+q\bar{q}$ at CMS

The CMS Collaboration

Abstract

A search for the standard model Higgs boson decaying to two Z bosons with subsequent decay to a final state with two leptons and two quark-jets, $H \to ZZ \to \ell^-\ell^+q\bar{q}$, is presented. Data corresponding to an integrated luminosity of 1 fb^{-1} of LHC proton-proton collisions at the center-of-mass energy of 7 TeV were collected and analyzed by the CMS experiment. The selection to discriminate between signal and background events is based on kinematic and topological quantities, which include the angular spin correlations of the decay products. The events are classified according to probability of the jets to originate from quarks of light or heavy flavor or from gluons. No evidence for a Higgs boson is found and upper limits on the Higgs boson production cross section are set in the range of masses between 226 GeV/c^2 and 600 GeV/c^2.

1 Introduction

The search for the standard model (SM) Higgs boson in a wide range of masses, from 120 GeV/c^2 to 600 GeV/c^2, is of high priority for experiments at the Large Hadron Collider (LHC) [1]. Direct and indirect measurements [2] suggest that the SM Higgs boson mass should be in the range between 100 GeV/c^2 and 200 GeV/c^2 with part of this range already excluded by direct limits from LEP [3] and the Tevatron [4]. However, these constraints rely on detailed theoretical assumptions and therefore the experimental goal is to cover all accessible masses.

The primary production mechanism of the Higgs boson at LHC is gluon fusion [5–14] with a small but measureable contribution from weak boson fusion (WBF) [15–20]. The decay of a Higgs boson to two light fermions is highly suppressed [21–24]. The Higgs boson decay channels with the highest discovery potential at the LHC have two gauge bosons in the final state. For a Higgs boson mass $m_{\mathrm{H}} < 2m_{\mathrm{W}}$ those final states contain two photons or two weak bosons, of which one is a virtual particle: ZZ^* or WW^*. For $m_{\mathrm{H}} \geq 2m_{\mathrm{W}}$, the most promising final states are those with two real weak bosons: $\mathrm{W}^+\mathrm{W}^-$ for $2m_{\mathrm{W}} \leq m_{\mathrm{H}} < 2m_{\mathrm{Z}}$, and ZZ for $m_{\mathrm{H}} \geq 2m_{\mathrm{Z}}$.

In this note we present an optimized search for a SM Higgs boson decay to two Z bosons with a subsequent decay to two leptons and two quark jets, $\mathrm{H} \to \mathrm{ZZ} \to \ell^-\ell^+\mathrm{q\bar{q}}$. The branching fraction of this decay channel is about 20 times higher than that of $\mathrm{H} \to \mathrm{ZZ} \to \ell^-\ell^+\ell^-\ell^+$. This may lead to better sensitivity to SM Higgs boson production at higher masses, where background can be effectively suppressed kinematically.

Previous searches of a resonance decay to $\mathrm{ZZ} \to \ell^-\ell^+\mathrm{q\bar{q}}$ include a recent ATLAS search for the Higgs boson with data from an integrated luminosity of 35 pb^{-1} [25] and a similar search for a graviton decay by CDF [26], both with sensitivity far from the expectation for a SM Higgs boson.

2 Event Reconstruction

We fully reconstruct the decay chain $\mathrm{H} \to \mathrm{ZZ} \to \ell^-\ell^+\mathrm{q\bar{q}}$, where the charged leptons $\ell^\pm$ are either muons or electrons and quarks are reconstructed as jets in the CMS detector [27]. This analysis uses data from proton-proton collisions, produced at a center-of-mass energy of 7 TeV, corresponding to an integrated luminosity of (1.00 ± 0.06) fb^{-1}. The data were collected by the CMS detector at the LHC during 2010 and 2011.

Although the main sources of background are estimated from data, Monte Carlo (MC) simulations are used to develop and validate these methods. Event generators used for background samples are either MADGRAPH 4.4.12 [28], ALPGEN 2.13 [29], or PYTHIA 6.4.22 [30]. Similarly, signal events are generated using POWHEG [31–33] and a dedicated generator from Ref. [34]. For both signal and background MC, events are simulated using a GEANT4-based model [35] of the CMS detector and processed using the same reconstruction algorithms as for data.

A detailed description of the CMS detector can be found in Ref. [27]. The central feature of the CMS detector is a 3.8 T superconducting solenoid of 6 m internal diameter. Within the field volume are the silicon pixel and strip tracker, the crystal electromagnetic calorimeter (ECAL), and the brass-scintillator hadron calorimeter (HCAL).

The CMS tracker [36] consists of 1440 silicon pixel and 15 148 silicon strip detector modules. The muon system includes barrel drift tubes covering the pseudorapidity range $|\eta| < 1.2$, end-

cap cathode strip chambers ($0.9 < |\eta| < 2.5$), and resistive plate chambers ($|\eta| < 1.6$) [37]. The ECAL consists of nearly 76 000 lead tungstate crystals, which provide coverage in pseudorapidity $|\eta| < 1.479$ in the barrel region and $1.479 < |\eta| < 3.0$ in the two endcap regions. The ECAL has an energy resolution of better than 0.5 % for unconverted photons with transverse energies above 100 GeV. The HCAL [38] consists of a set of sampling calorimeters which utilize alternating layers of brass as absorber and plastic scintillator as active material.

Muons are measured with the all-silicon tracker and the muon system. Electrons are detected in the ECAL as energy clusters and as tracks in the tracker. Both muons and electrons are required to have a momentum transverse to the pp beam direction, p_{T}, greater than 20 GeV/c and 40 GeV/c, for the lower and higher momentum lepton, respectively. They are measured in the pseudorapidity range $|\eta| < 2.4$ for muons, and $|\eta| < 2.5$ for electrons, though for electrons the transition range between the barrel and endcap, $1.44 < |\eta| < 1.57$, is excluded. Both the p_{T} and η requirements are consistent with those in the online trigger selection requiring either one or two charged leptons, either electrons or muons. The details of electron and muon identification criteria are described elsewhere [39]. Muons are required to be isolated from hadronic activity in the detector by restricting the sum of transverse momentum or energy in the tracker, ECAL, and HCAL, within a surrounding cone of $\Delta R \equiv \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} < 0.3$, to be less than 15% of the measured p_{T} of the muon. Electron isolation requirements are similar but vary depending on the shape of the electron shower.

Jets are reconstructed with the Particle Flow (PF) [40] algorithm, which is a full event reconstruction technique with the aim of reconstructing all particles produced in a given collision event through the combination of information from all sub-detectors. Reconstructed particle candidates are clustered to form PF jets with the *anti-k_T* algorithm [41] with radius parameter set to $R = 0.5$. The HCAL, the ECAL, and the tracker data combined in the PF algorithm measure jets with a resolution $\Delta p_{\mathrm{T}}/p_{\mathrm{T}} \approx 4/p_{\mathrm{T}} \oplus 0.2/\sqrt{p_{\mathrm{T}}}$, with p_{T} in GeV/c. Jets that overlap with isolated leptons within $\Delta R = 0.5$ are removed from consideration.

Jets are required to be inside the tracker acceptance ($|\eta| < 2.5$) thus allowing high reconstruction efficiency and precise energy measurements using PF techniques. Jet-energy corrections are applied to account for the non-linear response of the calorimeters to the particle energies and other instrumental effects. These corrections are based on data control samples such as dijet and γ+jet samples. Pile-up (i.e. overlapping minimum-bias events coming from different proton-proton collisions), and the underlying event have an effect on jet reconstruction by contributing additional energy to the measured jet energy. The median energy density due to pile-up is evaluated in each event and the corresponding energy is subtracted from each jet [42]. A jet requirement, primarily based on the energy balance between charged and neutral hadrons in a jet, is applied to remove misidentified jets. A preselection requirement $p_{\mathrm{T}} > 30$ GeV/c is applied to all jets.

Background suppression is primarily based on the dilepton and dijet invariant masses, m_{jj} and $m_{\ell\ell}$. The requirement 75 GeV/c^2 $< m_{jj} <$ 105 GeV/c^2 is applied in order to reduce the Z+jets background, and 70 GeV/c^2 $< m_{\ell\ell} <$ 110 GeV/c^2 to reduce backgrounds without a real Z, such as $\mathrm{t\bar{t}}$. The latter requirement is looser because it is not aimed at discriminating the dominant Z+jets background. The statistical analysis is based on the invariant mass of the Higgs boson candidate, m_{ZZ}, which is calculated after a fit of the final state four momenta, applying the constraint that the dijet invariant mass is consistent with the mass of the Z boson. The experimental resolutions are taken into account in this fit.

Since the Higgs boson carries no spin, the angular distribution of its decay products is independent of the production mechanism. The five angles that fully describe the kinematics of the

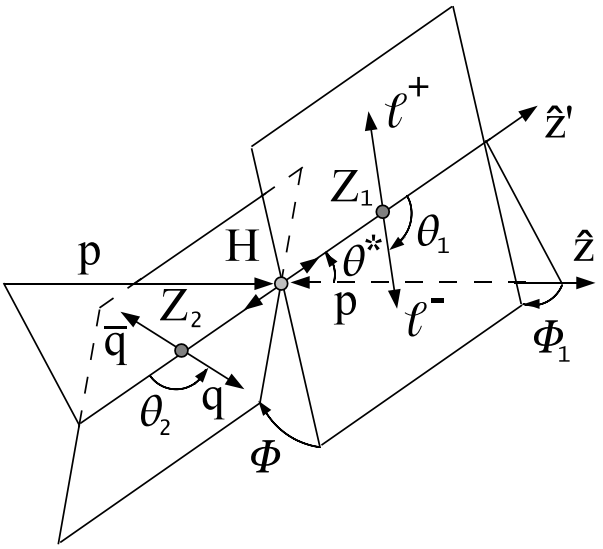


Figure 1: Diagram describing the production and decay angles in the process gg $\to$ H $\to$ ZZ $\to$ $\ell^-\ell^+$q$\bar{\text{q}}$. These angles are defined in the parent particle rest frames (H or Z).

gg $\to$ H $\to$ ZZ $\to$ $\ell^-\ell^+$q$\bar{\text{q}}$ process are discussed in Ref. [34] and shown in Fig. 1. Further kinematic selection exploits these five angular observables, which are only weakly correlated with the invariant masses of the Higgs and the two Z bosons, and with the longitudinal and transverse momenta of the Higgs boson candidate. They contain most of the discriminating power between signal and background. We construct an angular likelihood discriminant (LD) based on the probability ratio of the signal and background hypotheses, as described in Ref. [34]. The signal probability distribution is a correlated five-dimensional parameterization multiplied by empirically determined polynomial acceptance-functions describing non-uniform reconstruction efficiencies in the detector. The background probability is an empirical parameterization taken as a product of independent distributions for each observable. All probabilities are parameterized as functions of m_{ZZ}. Thresholds of the angular likelihood discriminant are selected to optimize the expected sensitivity to the production of a SM Higgs boson and depend on m_{ZZ}.

A powerful handle in the signal to background discrimination is offered by the parton flavor of the jets. Jets in signal events are produced in hadronic decays of a Z boson, and therefore originate from the hadronization of quarks. The flavor of quarks in Z decays is almost equally distributed among the five types d, u, s, c, b, with some preference given to the down-type quarks because of preferential electroweak couplings of the Z. The dominant background is represented by a leptonically-decaying Z boson produced in association with hard jets, a process in which gluon radiation is expected to play a major role. After gluons, the u and d quarks from the protons dominate the jet production associated with the Z. Therefore, the main features that discriminate signal from background are the relatively large contribution of heavy flavor quarks (b and c) and the absence of gluons. We take advantage of both features in the analysis by performing tagging of the b-flavor and introducing a likelihood discriminant which separates gluon and light-quark jets on a statistical basis, as described below.

To identify jets originating from the hadronization of bottom quarks, we use the Track Counting High Efficiency (TCHE) b-tagging algorithm [43, 44]. The data are split into three b-tag categories: a 2 b-tag category is required to have one jet identified with medium ($\sim$70% efficiency) and the other jet with loose ($\sim$80% efficiency) requirements, events not selected in the 2

Table 1: Summary of optimized kinematic and topological selection requirements.

	preselection		
$p_T(\ell^\pm)$	lowest $p_T > 20\,\text{GeV}/c$, highest $p_T > 40\,\text{GeV}/c$		
p_T(jets)	> 30 GeV/c		
$\lvert\eta\rvert(\ell^\pm)$	$(e^\pm) < 2.5$, $(\mu^\pm) < 2.4$		
$\lvert\eta\rvert$(jets)	< 2.5		
	0 b-tag	1 b-tag	2 b-tag
b-tag	none	one loose	medium & loose
angular LD	$> 0.55 + 0.00025\,m_{ZZ}$	$> 0.302 + 0.000656\,m_{ZZ}$	> 0.5
quark-gluon LD	> 0.10	–	–
$2\ln\lambda(\not{E}_T)$	–	–	< 10
m_{jj}	$\in$ [75, 105] GeV/c^2		
$m_{\ell\ell}$	$\in$ [70, 110] GeV/c^2		
m_{ZZ}	analyzed within [183, 800] GeV/c^2		

b-tag region are categorized as 1 b-tag if they have one jet satisfying the loose-tag requirements, the 0 b-tag category contains all the remaining events. The composition of the expected signal and background varies significantly among the three categories.

The 0 b-tag category is dominated by the Z+jets background, and from these events we further select a "gluon-tagged" category, which is then excluded in further analysis. Therefore, we apply an additional requirement that the two jets are not consistent with being gluons, based on three measured quantities. These are the number of charged hadronic particle tracks, the number of photon and neutral hadronic PF candidates, and transverse momentum distribution $\sqrt{\sum p_T^2/(\sum p_T)^2}$, where the sum is extended over all PF candidates in the jet. Gluons have a more intense coupling to the strong field than quarks; therefore, their hadronization favors the production of a larger number of stable particles. This translates into the observation of wider high-multiplicity jets when compared to those generated by final state quarks.

In order to suppress the substantial $t\bar{t}$ background in the 2 b-tag category, a likelihood-ratio-discriminant, λ, is constructed using resolution functions of the PF candidates. This provides a measure of the missing transverse energy significance based on the ratio of the likelihoods of the hypothesis that the event presents a real missing transverse energy ($\not{E}_T$), equal to the value measured with the PF algorithm, and the null hypothesis $\not{E}_T = 0$. We apply a loose requirement, $2\ln\lambda(\not{E}_T) < 10$, in the 2 b-tag category only.

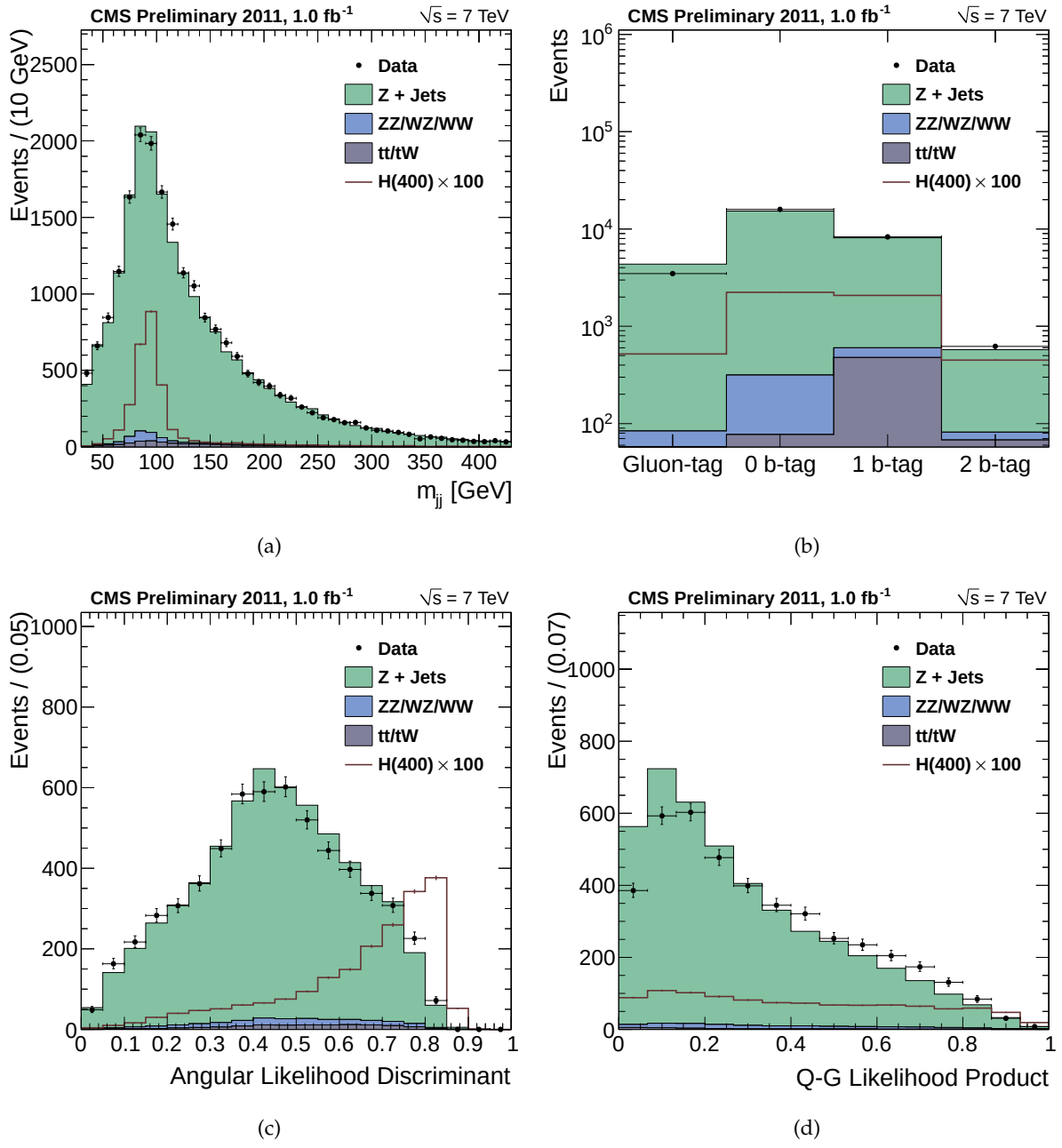
The main selection requirements described above are summarized in Table 1. Good agreement is observed between data and simulated predictions of background distributions after the pre-selection requirements, where the additional contribution of a Higgs boson signal would be indistinguishable above the overwhelming background. Examples of these distributions are given in Fig. 2, where we show the dijet invariant mass m_{jj}, flavor tagging category, including the gluon-tagged category, angular likelihood discriminant, and quark-gluon likelihood discriminant.

When an event contains more that one candidate passing the selection requirements, we retain the one with jets in the highest b-tag category for the analysis. Further ambiguity between multiple candidates is resolved selecting the candidate with (m_{jj}, $m_{\ell\ell}$) values closest to (m_Z,m_Z), m_Z being the nominal Z boson mass. The m_{ZZ} invariant mass distribution of expected background based on simulation as well as the selected data are displayed for the three b-tag categories in

Fig. 3. The main backgrounds include inclusive Z production with either light-flavor or heavy-flavor jets, top quark production, and diboson production such as WZ and ZZ. The expected and observed yields are listed in Table 2.

Table 2: Observed and expected yields with 1 fb^{-1} of data after sequential preselection and all selection requirements. The yields are quoted in the range $183\,\mathrm{GeV}/c^2 < m_{ZZ} < 800\,\mathrm{GeV}/c^2$, while signal extraction requires further analysis of the m_{ZZ} spectrum. The expected background is quoted from the sideband procedure (data) and from simulation (MC). The errors on the expected background from simulation include only statistical uncertainties. The expected Higgs signal yield is combined for the two lepton flavors.

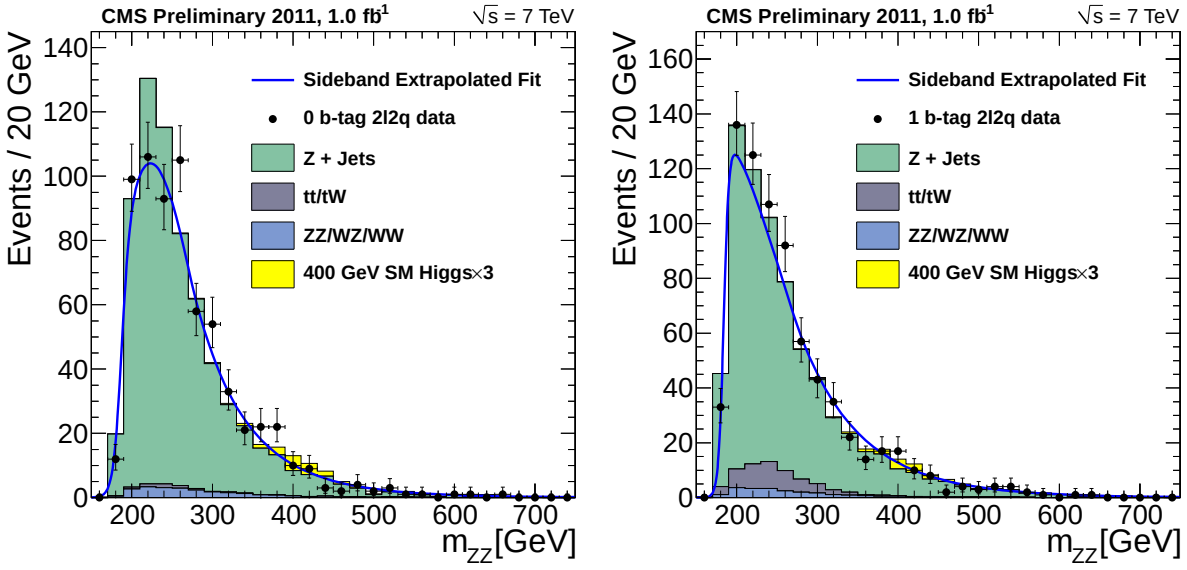
		preselection		
p_T and η		31368		
$m_{\ell\ell}$		24641		
m_{jj}		5451		
		selection		
		0 b-tag	1 b-tag	2 b-tag
$\mu^-\mu^+jj$				
observed yield		359	396	25
exp. background (data)		345.7 ± 17.8	376.4 ± 19.3	24.3 ± 3.7
exp. background (MC)		351.5 ± 5.6	371.2 ± 5.9	23.7 ± 1.6
e^-e^+jj				
observed yield		307	352	30
exp. background (data)		286.4 ± 16.2	334.7 ± 18.2	20.3 ± 3.1
exp. background (MC)		304.3 ± 5.4	332.6 ± 5.8	23.6 ± 1.7
signal expectation (MC)				
Higgs	200 GeV/c^2	2.57 ± 0.37	3.43 ± 0.49	0.70 ± 0.19
	250 GeV/c^2	5.34 ± 0.75	4.75 ± 0.67	1.46 ± 0.38
	300 GeV/c^2	5.84 ± 0.83	4.97 ± 0.69	1.75 ± 0.46
	350 GeV/c^2	6.45 ± 0.94	5.61 ± 0.80	2.15 ± 0.56
	400 GeV/c^2	5.33 ± 0.76	4.83 ± 0.67	1.96 ± 0.50
	450 GeV/c^2	3.54 ± 0.53	3.38 ± 0.48	1.43 ± 0.35
	500 GeV/c^2	2.21 ± 0.34	2.21 ± 0.32	0.97 ± 0.24
	550 GeV/c^2	1.36 ± 0.22	1.41 ± 0.22	0.63 ± 0.16
	600 GeV/c^2	0.83 ± 0.21	0.86 ± 0.21	0.39 ± 0.15



(a) (b)

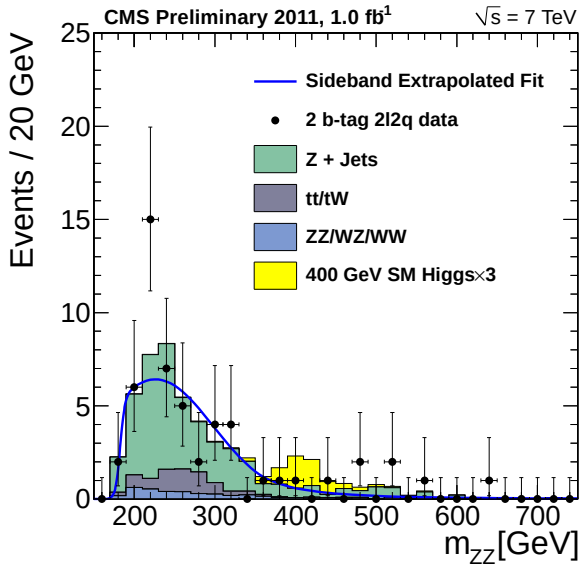
(c) (d)

Figure 2: Distribution of the dijet invariant mass m_{jj} (a), flavor tagging category (b), including the gluon-tagged category, angular likelihood discriminant (c), and quark-gluon likelihood discriminant (d). Points with error bars show distributions of data after preselection requirements, solid histograms depict the background expectation from simulated events with the different components illustrated. Open histograms indicate the expected distribution for a Higgs boson with mass 400 GeV/c^2, multiplied by a factor of 100 for illustration.



(a)

(b)



(c)

Figure 3: The m_{ZZ} invariant mass distribution after final selection in three categories: 0 b-tag (a), 1 b-tag (b), and 2 b-tag (c). Points with error bars show distributions of data, solid histograms depict the background expectation from simulated events with the different components illustrated. Also shown is a hypothetical signal with the mass of 400 GeV/c^2 and cross section 3.0 times that of the SM Higgs, which is smaller than but roughly corresponds to observed exclusion. Solid curved line shows prediction of background from sideband extrapolation procedure.

3 Event Analysis

The primary discrimination power between signal and background is provided by the invariant mass distribution of the Higgs boson, m_{ZZ}. Data containing a Higgs boson signal would have a distinct resonance peak in addition to the continuum background distribution. Therefore, it is critical to have proper estimations of background. The estimates from simulation shown in Fig. 3 provide a good illustration of the expected background, but require further validation of both theoretical predictions, such as production cross section, and detector effects, such as b-tagging efficiency.

We estimate background from the m_{jj} sidebands, defined as $60\,\text{GeV}/c^2 < m_{jj} < 75\,\text{GeV}/c^2$ or $105\,\text{GeV}/c^2 < m_{jj} < 130\,\text{GeV}/c^2$. The composition and distribution of the dominant backgrounds in the sidebands is in agreement with that in the signal region, $75\,\text{GeV}/c^2 < m_{jj} < 105\,\text{GeV}/c^2$. In order to minimize systematic uncertainties from theoretical calculations and detector effects, the expected number of background events, $N_{\text{bkg}}(m_{ZZ})$, is obtained from the number of events in the sidebands, $N_{\text{sb}}(m_{ZZ})$, as follows:

$$N_{\text{bkg}}(m_{ZZ}) = N_{\text{sb}}(m_{ZZ}) \times \frac{N^{\text{sim}}_{\text{bkg}}(m_{ZZ})}{N^{\text{sim}}_{\text{sb}}(m_{ZZ})} = N_{\text{sb}}(m_{ZZ}) \times \alpha(m_{ZZ}), \tag{1}$$

where $\alpha(m_{ZZ})$ is the ratio of the expected number of background events in the signal and sideband regions obtained from simulation. The advantage of this approach is that most of the systematic uncertainties on background cancel in the ratios, such as theoretical cross-section prediction and b-tagging efficiency, while the remaining factor, $\alpha(m_{ZZ})$, reflects small kinematic differences between the signal region and sidebands. It also provides an automatic normalization of the background and adjusts to the shape of the m_{ZZ} mass spectrum should there be any discrepancy between simulation and data.

The above procedure is applied independently in each b-tag category since background composition varies between categories. In all cases, the dominant backgrounds include Z+jets with either light or heavy flavor jets and top background, both of which populate the m_{jj} signal region and the m_{jj} sidebands. The diboson background, accounted for by $\alpha(m_{ZZ})$, amounts to less than 5% in the 0 and 1 b-tag categories and about 10% in the 2 b-tag category.

Additional information about the $t\bar{t}$ background is provided either by the $m_{\ell\ell}$ sidebands or by the mixed-flavor $e^{\pm}\mu^{\mp}jj$ sample, with otherwise identical selections. We illustrate the latter in Fig. 4, its advantage is that the prediction of the top background is nearly identical to the same-flavor final state, while the dominant Z background is absent. We also use a γ+jets control sample, which has advantage of higher statistics in data, to validate the Z+jets production. The energy of the photon is changed to match the mass of the Z in order to reproduce kinematics. Overall good agreement is found with $e^{\pm}\mu^{\mp}jj$ and γ+jets control samples. Both control samples are used for cross-checks of the dominant background contributions.

The m_{ZZ} distributions of selected events, such as those in Fig. 3, are split into six categories based on the b-tag type and the lepton flavor. These events are examined for 60 hypothetical Higgs masses in the range between 226 GeV/c^2 and 600 GeV/c^2, where the mass steps were optimized to account for Higgs resonance mass width and resolution [45]. For each mass hypothesis, we perform a simultaneous likelihood fit of the six m_{ZZ} distributions using the statistical approaches discussed in Ref. [45]. We have also studied an alternative classical cut-based analysis that counts events in regions of the m_{ZZ} distribution and found consistent but systematically better limits with the likelihood fit approach. As the prime method for reporting

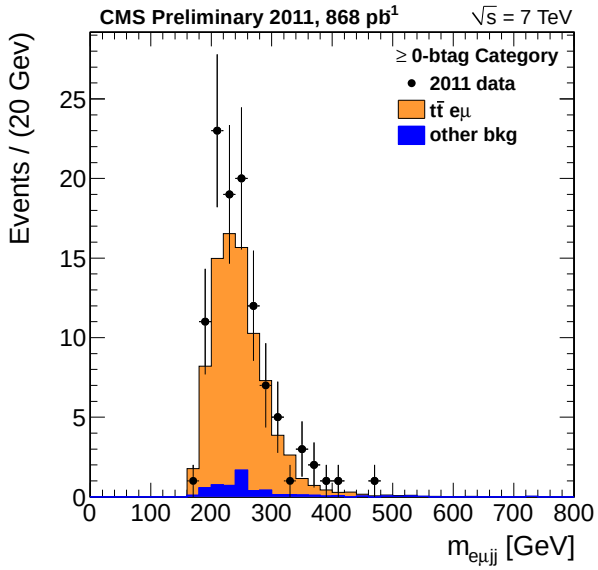


Figure 4: Comparison of the eμjj final state in data (points with error bars) and in the simulation (histogram).

limits we use the CL_s modified frequentist technique [46]. As a complementary method to the frequentist paradigm, we use the Bayesian approach. In this method, the Bayes theorem [47] is invoked to assign a degree of belief to the Higgs hypothesis by calculating the posterior "probability density function" on the signal strength. All results quoted in this paper are validated by using two independent sets of software tools, RooStats package [48] and L&S [49].

The background distribution is parameterized with an empirical probability function with normalization obtained from the m_{jj} sideband distribution rescaled to the m_{jj} signal region. The dominant normalization uncertainty in the sideband procedure is expected to be due to statistical fluctuations in the number of events in the sidebands. The other uncertainty comes from modeling of the invariant mass m_{ZZ} distributions, which we vary according to uncertainties estimated from data. We adopt a conservative approach when errors on parameters of the invariant mass m_{ZZ} probability distribution are treated as uncorrelated resulting in overcoverage of uncertainties. Further improvements are expected with a less conservative approach. Signal distributions are described with a relativistic Breit–Wigner parameterization (with the width Γ according to the SM expectation for each mass hypothesis) convoluted with a Crystal-Ball function [50] reflecting both mismeasurement and misreconstruction of the Higgs boson decay products. The signal reconstruction efficiency and the resolution function are parameterized as a function of the hypothetical Higgs boson mass and are extrapolated to all 60 mass points. The main uncertainties in the signal parameterization are due to the resolution function, which is predominantly affected by the error on the jet energy scale.

The main systematic uncertainties on signal normalization are summarized in Table 3. We consider effects from lepton energy scale, resolution, selection, and trigger; jet resolution and efficiency; pile-up; heavy-quark flavor tagging and quark-gluon discrimination; $\not\!\!E_T$; Higgs production mechanism; LHC luminosity; Higgs cross section and branching fractions. Lepton efficiencies are evaluated from data with a tag-and-probe [39] approach when one lepton from an inclusive sample of Z decays serves as a tag and the efficiency for the reconstruction of the other lepton is calculated. Effects of jet reconstruction are evaluated by the variation of the

Table 3: Summary of systematic uncertainties on signal normalization. Most sources give multiplicative errors on the cross-section measurement, except for the expected Higgs boson production cross section, which is relevant for the measurement of the ratio to the SM expectation.

source	0 b-tag	1 b-tag	2 b-tag
muon reco	2.7%		
electron reco	4.5%		
jet reco	1–5%		
pileup	2%		
b-tagging	3%	1%	20%
gluon-tagging	4.6%	–	–
$\not{E}_{\mathrm{T}}$	–	–	3%
acceptance (PDF)	3%		
acceptance (HQT)	2%	5%	3%
acceptance (WBF)	1-2%		
luminosity	6%		
Higgs cross section	13–18%		

jet energy and resolution within calibration uncertainties using the data control samples such as dijet and γ+jet samples. Effects of pile-up are taken as a difference between reconstruction efficiency with pileup below and above the average expected value, otherwise distributed according to observed values in data. The requirement on the $\not{E}_{\mathrm{T}}$ significance translates into about 3% inefficiency and the resulting uncertainty does not surpass this value. Uncertainty on the b-tagging has been evaluated with a sample of jet events enriched in heavy flavor by requiring a muon to be associated with a jet. This uncertainty includes effects of event pile-up, dependence on fraction of gluon splitting into $b\bar{b}$ in the jet sample, and other variables. Uncertainty on quark-gluon LD selection efficiency was evaluated using the γ+jets sample, which predominantly contains quark jets.

Uncertainties in the production mechanism affect signal acceptance in the detector because both the longitudinal momentum of the Higgs boson, due to PDFs, and the transverse momentum of the Higgs boson, due to QCD initial-state radiation effects, are model-dependent. We follow the PDF4LHC [51–55] recommendation to estimate the uncertainty due to PDF knowledge and to calculate the uncertainty on signal acceptance. We rescale the transverse momentum distribution of the Higgs boson using HQT [56] as a reference and take the full change in efficiency as a systematic uncertainty. Uncertainties on the Higgs boson production cross section are taken from Ref. [57] which includes uncertainties from QCD renormalization and factorization scales, PDFs, and α_s. These uncertainties are separated between the gluon-fusion (gg) and weak boson fusion (WBF) production mechanisms, but uncertainties on the gluon-fusion process dominate in the total production cross section. We also account for a small uncertainty due to a difference in signal acceptance with the gg and WBF production mechanisms while the analysis targets the dominant gluon-fusion production.

In the current study the exclusion limit is compared to the cross section for on-shell Higgs production and decay in the zero-width approximation, and acceptance estimates are obtained with Monte Carlo simulations that are based on ad-hoc Breit-Wigner distributions for describing the Higgs-boson propagation. Recent analyses show that the use of a QFT-consistent Higgs propagator, allowing also for the off-shellness of the Higgs-boson, dynamical QCD scales and interference effects between Higgs signal and backgrounds will result, at Higgs masses above 300 GeV, in a sizable effect on conventionally defined but theoretically consistent parameters

(mass and width) that describe the propagation of an unstable Higgs-boson [57–59]. These effects are estimated to amount to an additional uncertainty on the theoretical cross section of 10–30% for 400–600 GeV/c^2 Higgs masses, but in this work are not included in the limit calculations.

We also consider the Higgs production and decay within the SM4 model with four generations [59–61]. The main difference from the SM Higgs production is 8.3–4.8 times higher cross section for a 200–600 GeV/c^2 SM4 Higgs. We assume the main uncertainties on the SM4 Higgs production cross section to be the same as for the gluon-fusion mechanism in SM but with an additional 10% uncertainty due to the electroweak radiative corrections. This additional uncertainty is added linearly to the uncertainties from QCD renormalization and factorization scales, PDFs, and α_s.

4 Results

Based on the expected normalization and shape of the m_{ZZ} distribution, for signal and background, and the corresponding systematic uncertainties, we generate a large number of random pseudo-experiments. For each of them, the expected background distribution is generated and a likelihood fit is performed. Observed and expected exclusion limits on the product of the Higgs boson production cross section and the branching fraction of $H \to ZZ$ are presented in Fig. 5 using the CL_s technique. Results with the Bayesian approach are consistent with those of CL_s. For comparison, expectation of the production cross section and the branching fraction are shown in the SM and in the SM4 model.

We further incorporate uncertainties on the Higgs production cross section and present a limit on the ratio of the SM Higgs boson production cross section to the SM expectation in Fig. 6. A similar limit on the ratio to the Higgs boson production cross section in the SM4 model is shown in Fig. 7. A range of SM4 Higgs mass hypotheses are excluded between 226 and 445 GeV/c^2 at 95% CL, except for two windows between 261 and 270 GeV/c^2 and between 370 and 381 GeV/c^2. The exclusion limits in Fig. 6 are approaching those of the SM expectation for the Higgs boson production. In the absence of the Higgs boson, these limits are expected to reach the SM expectation with the increased LHC luminosity. At the same time, these presented results along with comparable results from the CMS collaboration [45] obtained with the other channels of the H $\to$ ZZ decay, such as H $\to$ ZZ $\to$ $\ell^-\ell^+\ell^-\ell^+$ and $\ell^-\ell^+\nu\bar{\nu}$, and with the H $\to$ WW decay provide exclusion of the SM Higgs boson at 95% CL in a wide range within the mass window in this search.

5 Summary and Discussion

A search for the standard model Higgs boson decaying into two Z bosons which subsequently decay to two leptons and two quark jets, H $\to$ ZZ $\to$ $\ell^-\ell^+$q$\bar{\text{q}}$, has been presented. Data corresponding to an integrated luminosity of (1.00 ± 0.06) fb^{-1} of proton-proton collisions at center-of-mass energy of 7 TeV have been collected and analyzed by the CMS experiment at LHC. No evidence for a SM-like Higgs boson has been found and upper limits on the production cross section for the SM Higgs boson have been set in the range of masses between 226 GeV/c^2 and 600 GeV/c^2. In this analysis we have excluded a large range of Higgs mass hypotheses in the model with the fourth generation SM-like couplings of the Higgs and, when results are combined with other Higgs decay channels from CMS, excluded [45] a wide range of masses of the SM Higgs within the mass window in this search.

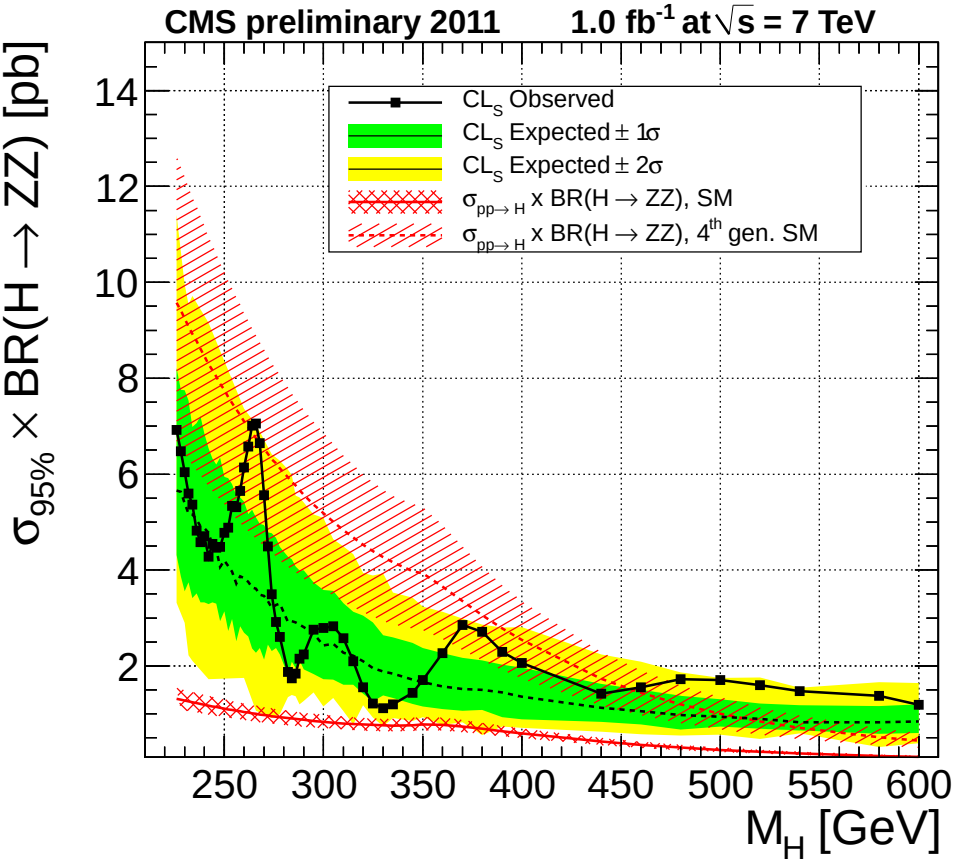


Figure 5: Observed (dashed) and expected (solid) 95% CL upper limit on the product of the Higgs boson production cross section and the branching fraction of $H \to ZZ$ using 1 fb^{-1} of data obtained with the CL$_s$ technique. The 68% and 95% ranges of expectation are also shown with green and yellow bands. The expected product of the SM Higgs production cross section and the branching fraction is shown as a red solid curve with a band indicating theoretical uncertainties at 68%. The same expectation in the SM4 model are shown with a red dashed curve with a band indicating theoretical uncertainties.

References

[1] L. Evans and P. Bryant (eds.), "LHC machine", *JINST* **3** (2008) S08001. doi:10.1088/1748-0221/3/08/S08001.

[2] ALEPH, CDF, D0, DELPHI, L3, OPAL, SLD Collaborations, the LEP Electroweak Working Group, the Tevatron Electroweak Working Group, and the SLD electroweak and heavy flavour groups Collaboration, "Precision electroweak measurements and constraints on the Standard Model", *CERN-PH-EP-2010-09, FERMILAB-TM-2480-PPD, SLAC-PUB-14301* (2010).

[3] ALEPH, DELPHI, L3, OPAL Collaboration, "Search for the Standard Model Higgs boson at LEP", *Phys. Lett.* **B565** (2003) 61. doi:doi:10.1016/S0370-2693(03)00614-2.

[4] CDF, D0, and the TEVNPHWG Working Group Collaboration, "Combined CDF and D0 Upper Limits on Standard Model Higgs Boson Production with up to 8.2 fb^{-1} of Data", arXiv:1103.3233.

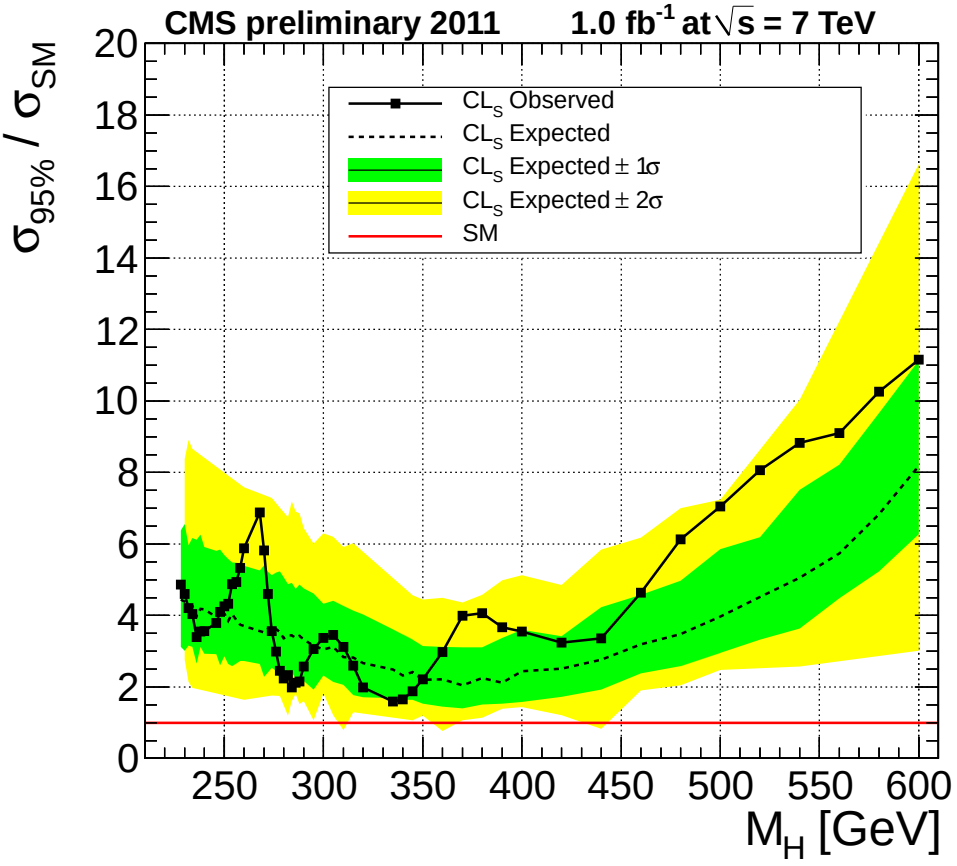


Figure 6: Observed (dashed) and expected (solid) 95% CL upper limit on the ratio of the Higgs boson production cross section to the SM expectation using 1 fb^{-1} of data obtained with the CL_s technique. The 68% and 95% ranges of expectation are also shown with green and yellow bands. The solid line at 1 indicates SM expectation.

[5] S. Dawson, "Radiative corrections to Higgs boson production", *Nucl. Phys.* **B359** (1991) 283–300. doi:10.1016/0550-3213(91)90061-2.

[6] M. Spira, A. Djouadi, D. Graudenz et al., "Higgs boson production at the LHC", *Nucl. Phys.* **B453** (1995) 17–82, arXiv:hep-ph/9504378. doi:10.1016/0550-3213(95)00379-7.

[7] R. V. Harlander and W. B. Kilgore, "Next-to-next-to-leading order Higgs production at hadron colliders", *Phys. Rev. Lett.* **88** (2002) 201801, arXiv:hep-ph/0201206. doi:10.1103/PhysRevLett.88.201801.

[8] C. Anastasiou and K. Melnikov, "Higgs boson production at hadron colliders in NNLO QCD", *Nucl. Phys.* **B646** (2002) 220–256, arXiv:hep-ph/0207004. doi:10.1016/S0550-3213(02)00837-4.

[9] 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.* **B665** (2003) 325–366, arXiv:hep-ph/0302135. doi:10.1016/S0550-3213(03)00457-7.

[10] S. Catani, D. de Florian, M. Grazzini et al., "Soft-gluon resummation for Higgs boson production at hadron colliders", *JHEP* **07** (2003) 028, arXiv:hep-ph/0306211.

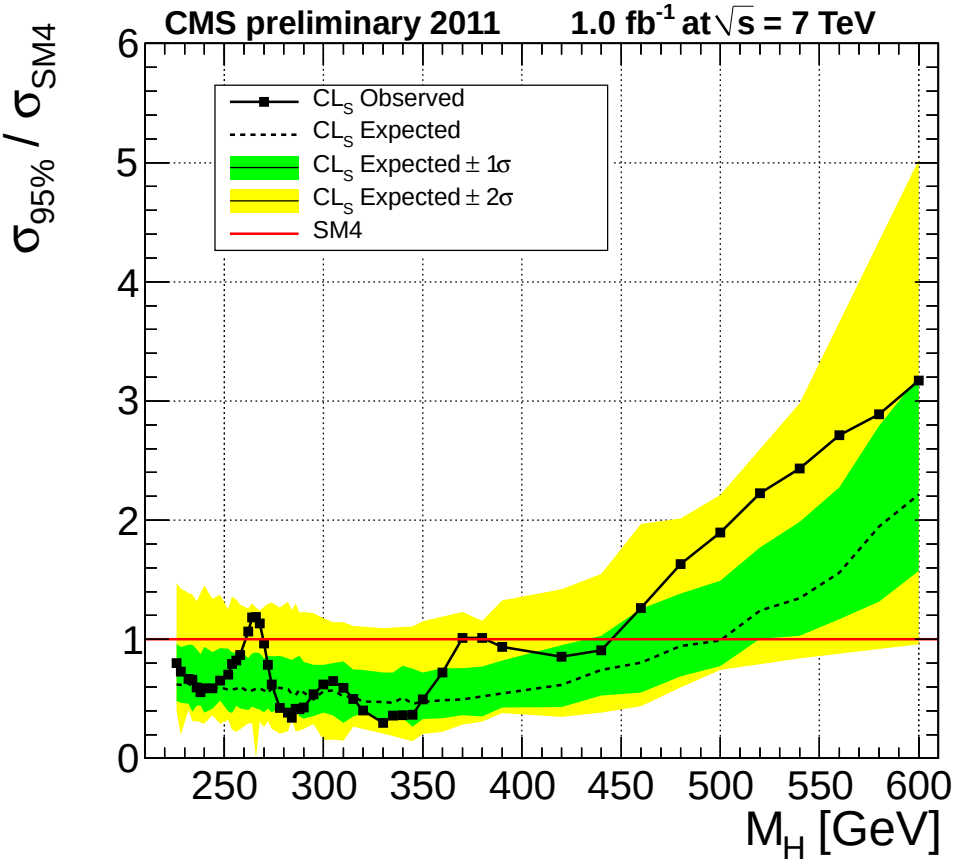


Figure 7: Observed (dashed) and expected (solid) 95% CL upper limit on the ratio of the Higgs boson production cross section to the expectation with the SM4 model using 1 fb^{-1} of data obtained with the CL_s technique. The 68% and 95% ranges of expectation are also shown with green and yellow bands. The solid line at 1 indicates SM4 expectation.

[11] U. Aglietti, R. Bonciani, G. Degrassi et al., "Two-loop light fermion contribution to Higgs production and decays", *Phys. Lett.* **B595** (2004) 432–441, arXiv:hep-ph/0404071. doi:10.1016/j.physletb.2004.06.063.

[12] S. Actis, G. Passarino, C. Sturm et al., "NLO Electroweak Corrections to Higgs Boson Production at Hadron Colliders", *Phys. Lett.* **B670** (2008) 12–17, arXiv:0809.1301. doi:10.1016/j.physletb.2008.10.018.

[13] C. Anastasiou, R. Boughezal, and F. Petriello, "Mixed QCD-electroweak corrections to Higgs boson production in gluon fusion", *JHEP* **04** (2009) 003, arXiv:0811.3458. doi:10.1088/1126-6708/2009/04/003.

[14] D. de Florian and M. Grazzini, "Higgs production through gluon fusion: updated cross sections at the Tevatron and the LHC", *Phys. Lett.* **B674** (2009) 291–294, arXiv:0901.2427. doi:10.1016/j.physletb.2009.03.033.

[15] 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.* **99** (2007) 161803, arXiv:0707.0381. doi:10.1103/PhysRevLett.99.161803.

[16] M. Ciccolini, A. Denner, and S. e. Dittmaier, "Electroweak and QCD corrections to Higgs production via vector-boson fusion at the LHC", *Phys. Rev.* **D77** (2008) 013002, arXiv:0710.4749. doi:10.1103/PhysRevD.77.013002.

[17] T. Figy, C. Oleari, and D. Zeppenfeld, "Next-to-leading order jet distributions for Higgs boson production via weak-boson fusion", *Phys. Rev.* **D68** (2003) 073005, arXiv:hep-ph/0306109. doi:10.1103/PhysRevD.68.073005.

[18] K. Arnold et al., "VBFNLO: A parton level Monte Carlo for processes with electroweak bosons", *Comput. Phys. Commun.* **180** (2009) 1661–1670, arXiv:0811.4559. doi:10.1016/j.cpc.2009.03.006.

[19] P. Bolzoni, F. Maltoni, S.-O. Moch et al., "Higgs production via vector-boson fusion at NNLO in QCD", *Phys. Rev. Lett.* **105** (2010) 011801, arXiv:1003.4451. doi:10.1103/PhysRevLett.105.011801.

[20] T. Figy, S. Palmer, and G. Weiglein, "Higgs production via weak boson fusion in the standard model and the MSSM", arXiv:1012.4789.

[21] A. Djouadi, J. Kalinowski, and M. Spira, "HDECAY: A program for Higgs boson decays in the standard model and its supersymmetric extension", *Comput. Phys. Commun.* **108** (1998) 56–74, arXiv:hep-ph/9704448. doi:10.1016/S0010-4655(97)00123-9.

[22] A. Djouadi, J. Kalinowski, M. Muhlleitner 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. arXiv:1003.1643.

[23] A. Bredenstein, A. Denner, S. Dittmaier et al., "Precise predictions for the Higgs-boson decay H $\to$ WW/ZZ $\to$ 4 leptons", *Phys. Rev.* **D74** (2006) 013004, arXiv:hep-ph/0604011. doi:10.1103/PhysRevD.74.013004.

[24] A. Bredenstein, A. Denner, S. Dittmaier et al., "Radiative corrections to the semileptonic and hadronic Higgs-boson decays H $\to$W W / Z Z$\to$ 4 fermions", *JHEP* **0702** (2007) 080, arXiv:hep-ph/0611234.

[25] ATLAS Collaboration, "Search for a Standard Model Higgs Boson in the Mass Range 200-600 GeV in the Channels H to ZZ to llnunu and H to ZZ to llqq with the ATLAS Detector", *ATLAS-CONF-2011-026* (2011).

[26] CDF Collaboration, "Search for New Heavy Particles Decaying to Z0 Z0 to llll, lljj in p pbar Collisions at Sqrt(s) = 1.96 TeV", arXiv:1102.4566.

[27] CMS Collaboration, "The CMS experiment at the CERN LHC", *JINST* **0803** (2008) S08004.

[28] J. Alwall, P. Demin, S. de Visscher et al., "MadGraph/MadEvent v4: the new web generation", *JHEP* **09** (2007) 028, arXiv:0706.2334. doi:10.1088/1126-6708/2007/09/028.

[29] M. Mangano, M. Moretti, F. Piccinini et al., "ALPGEN, a generator for hard multiparton processes in hadronic collisions", *JHEP* **0307** (2003) 001, arXiv:hep-ph/0206293. doi:10.1088/1126-6708/2003/07/001.

[30] T. Sjöstrand, S. Mrennan and P. Z. Skands, "PYTHIA 6.4 Physics and Manual", *JHEP* **05** (2006) 026, arXiv:hep-ph/0603175.

[31] P. Nason, "A New Method for Combining NLO QCD with Shower Monte Carlo Algorithms", *JHEP* **11** (2004) 040. doi:10.1088/1126-6708/2004/11/040. arXiv:hep-ph/0409146.

[32] S. Frixione, P. Nason, and C. Oleari, "Matching NLO QCD Computations with Parton Shower Simulations: the POWHEG method", *JHEP* **11** (2007) 070. doi:10.1088/1126-6708/2007/11/070. arXiv:0709.2092.

[33] S. Alioli, P. Nason, C. Oleari et al., "NLO Vector-Boson Production Matched with Shower in POWHEG", *JHEP* **07** (2008) 06. doi:10.1088/1126-6708/2008/07/060. arXiv:0805.4802.

[34] Y. Gao et al., "Spin determination of single-produced resonances at hadron colliders", *Phys. Rev.* **D81** (2010) 075022, arXiv:1001.3396. doi:10.1103/PhysRevD.81.075022.

[35] GEANT4 Collaboration, "GEANT4: a simulation toolkit", *Nucl. Instrum. Meth.* **A506** (2003) 250. doi:10.1016/S0168-9002(03)01368-8.

[36] CMS Collaboration, "The CMS tracker system project: technical design report", *CERN/LHCC* **98-006** (1998). CMS TDR 5.

[37] CMS Collaboration, "CMS MUON Technical Design Report", *CERN/LHCC* **32** (1997).

[38] CMS Collaboration, "Performance of CMS Hadron Calorimeter Timing and Synchronization using Test Beam, Cosmic Ray, and LHC Beam Data", *JINST* **5** (2010) T03013, arXiv:0911.4877. doi:10.1088/1748-0221/5/03/T03013.

[39] CMS Collaboration, "Measurements of inclusive W and Z cross sections in pp collisions at $\sqrt{s} = 7$ TeV", *JHEP* **1** (2011) 001, arXiv:1012.2466. doi:10.1007/JHEP01(2011)080.

[40] CMS Collaboration, "Particle Flow Event Reconstruction in CMS and Performance for Jets, Taus, and Emiss", *CMS-PAS-PFT-09-001* (2009).

[41] M. Cacciari, G. P. Salam, and G. Soyez, "The anti-k_t jet clustering algorithm", *JHEP* **04** (2008) 063, arXiv:0802.1189. doi:10.1088/1126-6708/2008/04/063.

[42] M. Cacciari, G. P. Salam, and G. Soyez, "The Catchment Area of Jets", *JHEP* **04** (2008) 005, arXiv:0802.1188. doi:10.1088/1126-6708/2008/04/005.

[43] CMS Collaboration, "Commissioning of b-jet identification with pp collisions at sqrt(s) = 7 TeV", *CMS-BTV-10-001* (2010).

[44] CMS Collaboration, "Performance of b-jet identification in CMS", *CMS-BTV-11-001* (2011).

[45] CMS Collaboration, "Standard model Higgs boson search in pp-collisions at $\sqrt{s} = 7$ TeV and integrated luminosity up to 1.1 fb^{-1}", *CMS-HIG-11-011* (2011).

[46] A. L. Read, "Presentation of search results: the CLs technique", *J. Phys. G: Nucl. Part. Phys.* **28** (2002).

[47] A. O'Hagan and J.J. Forster, "Bayesian Inference", *Kendall's Advanced Theory of Statistics, Arnold, London* **2B** (2004).

[48] L. Moneta, K. Belasco, K. Cranmer et al., "The RooStats Project", *PoS* **ACAT2010** (2010) 057. [arxiv:1009.1003].

[49] Chen, M. and Korytov, A., "Limits and Signifcance". https://mschen.web.cern.ch/mschen/LandS/.

[50] J. Gaiser, "Charmonium Spectroscopy from Radiative Decays of the J/ψ and ψ'". PhD thesis, Stanford University, 1982. Appendix F.

[51] M. Botje, J. Butterworth, A. Cooper-Sarkar et al., "The PDF4LHC Working Group Interim Recommendations", arXiv:1101.0538.

[52] S. Alekhin, S. Alioli, R. D. Ball et al., "The PDF4LHC Working Group Interim Report", arXiv:1101.0536. * Temporary entry *.

[53] H.-L. Lai, M. Guzzi, J. Huston et al., "New parton distributions for collider physics", *Phys.Rev.* **D82** (2010) 074024, arXiv:1007.2241. doi:10.1103/PhysRevD.82.074024.

[54] A. Martin, W. Stirling, R. Thorne et al., "Parton distributions for the LHC", *Eur.Phys.J.* **C63** (2009) 189–285, arXiv:0901.0002. doi:10.1140/epjc/s10052-009-1072-5.

[55] NNPDF Collaboration Collaboration, "Impact of Heavy Quark Masses on Parton Distributions and LHC Phenomenology", arXiv:1101.1300.

[56] G. Bozzi, S. Catani, D. de Florian et al., "Transverse-momentum resummation and the spectrum of the Higgs boson at the LHC", *Nucl. Phys.* **B737** (2006) 73–120, arXiv:hep-ph/0508068. doi:10.1016/j.nuclphysb.2005.12.022.

[57] LHC Higgs Cross Section Working Group, S. Dittmaier, C. Mariotti et al., "Handbook of LHC Higgs Cross Sections: 1. Inclusive Observables", *CERN-2011-002* (CERN, Geneva, 2011) arXiv:1101.0593.

[58] G. Passarino, C. Sturm, and S. Uccirati, "Higgs Pseudo-Observables, Second Riemann Sheet and All That", *Nucl. Phys.* **B834** (2010) 77–115, arXiv:1001.3360. doi:10.1016/j.nuclphysb.2010.03.013.

[59] C. Anastasiou, S. Buehler, F. Herzog et al., "Total cross-section for Higgs boson hadroproduction with anomalous Standard Model interactions", arXiv:1107.0683.

[60] N. B. Schmidt, S. A. Cetin, S. Istin et al., "The Fourth Standart Model Family and the Competition in Standart Model Higgs Boson Search at Tevatron and LHC", *Eur. Phys. J.* **C66** (2010) 119–126, arXiv:0908.2653. doi:10.1140/epjc/s10052-010-1238-1.

[61] Q. Li, M. Spira, J. Gao et al., "Higgs Boson Production via Gluon Fusion in the Standard Model with four Generations", *Phys. Rev.* **D83** (2011) 094018, arXiv:1011.4484. doi:10.1103/PhysRevD.83.094018.