Search for the standard model Higgs boson decaying to tau pairs produced in association with a W or Z boson with the CMS experiment in $pp$ collisions at $\sqrt{s} = 7$ and 8 TeV

The CMS Collaboration

Abstract

A search for the standard model Higgs boson produced in association with a W or Z boson is presented using data collected in 2011 and 2012 with the CMS detector at the LHC. The topologies studied here have three- or four-leptons final states. The Higgs boson decay channel analysed is the ditau channel where the tau can decay via an electron, a muon or hadronically. The W and Z boson decays used to enhance the signal events are into electron or muon and the dielectrons or dimuons, respectively. The analysis uses $pp$ collision data samples corresponding to integrated luminosities of 5.0 $fb^{-1}$ collected at $\sqrt{s} = 7$ TeV and 19.5 $fb^{-1}$ of 8 TeV. Upper limits at 95% C.L. between 2.9 and 4.6 times the standard model prediction are established on the product of Higgs boson cross section and tau pair decay branching fraction for Higgs masses between 110 and 145 GeV/$c^2$. 
1 Introduction

Spontaneous electroweak symmetry breaking is introduced in the standard model (SM) [1–3] to give mass to the vector bosons (W± and Z) that mediate weak interactions, while keeping the photon, which mediates electromagnetic interactions, massless. This mechanism results in a single scalar in the SM [4–9], the Higgs boson. While the mass of the Higgs boson is a free parameter in the SM, its couplings to the massive vector bosons, Yukawa couplings to fermions, decay branching fractions, and production cross sections in proton-proton collisions are fully defined and well understood theoretically [10]. Gluon fusion (GF), weak vector boson fusion (VBF), associated production (AP) with weak bosons, and associated production with a tt pair (ttH) are the four most important Higgs boson production mechanisms at the Large Hadron Collider (LHC). Although the cross section for AP is an order of magnitude lower than that of the GF mechanism, the presence of isolated high momentum leptons originating from W and Z decays suppresses the backgrounds dramatically, making these channels viable for searches for the Higgs boson.

A resonance consistent with a SM Higgs boson with a mass of about 125 GeV/c² has been observed with a significance of 5.0σ (5.9σ) at the CMS [11] (ATLAS [12]) experiment. At both experiments, the observed excess is driven by the $H \rightarrow \gamma\gamma$, $H \rightarrow ZZ$ and $H \rightarrow WW$ decay modes. While the data are insufficient to exclude the presence of a SM Higgs boson decaying to tau pairs, no excess has been observed in any $H \rightarrow \tau\tau$ searches [11, 13]. It is therefore critical [14] to measure this new resonance in its decays to tau pairs, to determine if it is consistent with the SM Higgs boson or not.

In this document we present searches for the Higgs boson decaying to a tau pair where the Higgs boson is produced in association with a W or Z boson that decays leptonically. While the decays to tau pairs are the dominant Higgs boson signal contribution, the final states used in this paper can additionally be produced by the decay of the Higgs boson into a pair of W bosons that both decay to leptons. The searches use a data sample of proton-proton collisions from an integrated luminosity of 24.5 fb⁻¹ recorded by the Compact Muon Solenoid (CMS) [15] experiment at the LHC. The data are separated into two periods: 5.0 fb⁻¹ [16] was collected in 2011 at $\sqrt{s} = 7$ TeV, and 19.5 fb⁻¹ was collected in 2012 at $\sqrt{s} = 8$ TeV.

Throughout this document, the expression “light lepton,” or symbol $\ell$, will refer to an electron or muon, the symbol $\tau_l$ to a hadronically-decaying tau, and the symbol $L$ to an $e$, $\mu$, or $\tau_l$. Three- and four-lepton events are used to search for Higgs bosons produced in the AP mechanism with the leptonic decay of the W boson (WH) or Z boson (ZH), respectively. The search for WH production is performed in three-lepton final states. Two final states with two light leptons and a hadronic tau decay, $e\mu\tau_l$ and $\mu\mu\tau_l(\ell\ell\tau_l)$, and two final states with one light lepton and two hadronic tau decays, $\ell\tau_l\tau_l$ and $\mu\tau_l\tau_l(\ell\ell\tau_l\tau_l)$, are considered. The search for ZH production is performed in four-lepton ($\ell\ell\ell\ell\ell$) events with a pair of electrons or muons consistent with the decay of a Z boson, and a Higgs boson candidate with one of the following final states: $e\mu$, $e\tau_l$, $\mu\tau_l$, or $\tau_l\tau_l$. Even though the selections of the analyzed decay modes do not include the $H \rightarrow WW$ decays [17], the selected final states are sensitive to this channel.

2 The CMS detector, event reconstruction, and simulation

A more detailed description of the CMS detector can be found in Ref. [15]. The central feature of the CMS apparatus is a superconducting solenoid of 6 m internal diameter, providing a field of 3.8 T. Within the solenoid are the silicon pixel and strip tracker, which cover a pseudorapidity region of $|\eta| < 2.5$. Here, the pseudorapidity is defined as $\eta = -\ln (\tan(\theta/2))$, where $\theta$ is the
polar angle with respect to the direction of the counterclockwise beam. The lead-tungstate crystal electromagnetic calorimeter (ECAL) and the brass/scintillator hadron calorimeter (HCAL) surround the tracking volume and cover $|\eta| < 3$. In addition to the barrel and endcap detectors, CMS has forward calorimetry that extends the coverage to $|\eta| < 5$. The ECAL consists of 75,848 lead-tungstate crystals that provide coverage in pseudorapidity $|\eta| < 1.479$ in the barrel region and $1.479 < |\eta| < 3.0$ in the two endcap (EE) regions. A preshower detector consisting of two planes of silicon sensors interleaved with a total of $3X_0$ of lead is located in front of the EE. Muons are measured in gas-ionization detectors embedded in the steel return yoke, with a coverage of $|\eta| < 2.4$.

The identification of electrons, muons, and hadronically-decaying taus relies crucially on the association of tracks in the tracker with energy depositions in the ECAL for electrons, in the HCAL for charged hadrons, and track segments in the muon system for muons. Photons are identified as ECAL energy depositions without an associated track. All particles are reconstructed using the particle flow (PF) algorithm [18], which focuses on using an optimized combination of subdetector information to reconstruct each individual particle with the highest fidelity. The energy resolution resulting from this reconstruction is between 1 and 3% for the momentum range relevant for this analysis for electrons, muons, photons, and $\tau_h$ candidates, depending on the exact kinematics of the particular particle.

The particles reconstructed by the PF algorithm are used to reconstruct composite objects such as jets, hadronically-decaying taus, and missing transverse energy ($E_T^{\text{miss}}$), defined as the magnitude of the vector sum of the transverse momenta ($p_T$) of all PF objects. The jets are identified using the anti-$k_T$ jet algorithm [19] with a distance parameter of $R = 0.5$. In the 2011 (2012) dataset, an average of 10 (20) interactions (pile-up) occur in each proton bunch crossing. To correct for the contribution to the jet energy due to pile-up, the transverse energy density per unit area ($\rho$) of the pile-up is computed [20, 21] for each event. The energy due to pile-up is estimated as the product of $\rho$ and the area of the jet, and is subtracted from the jet transverse energy ($E_T$) [22]. Jet energy corrections are applied as a function of the jet $E_T$ and $\eta$ [23]. Hadronically-decaying taus are reconstructed using the “hadron-plus-strips” (HPS) algorithm [24], which reconstructs candidates with one or three charged pions and up to two neutral pions. The $E_T^{\text{miss}}$ is further corrected for pile-up effects using a multivariate approach [25].

The PYTHIA (version 6.424) [26] Monte Carlo (MC) event generator is used to generate the simulated Higgs boson samples used in this analysis. The $Z/\gamma^*$, $W$+jets, $ZZ$, $WZ$, and $Z\gamma$ diboson background samples are generated using MADGRAPH 5.1.3 [27] with PYTHIA for hadronization. The generators use the CTEQ6L[28] set of parton distribution functions. While the next-to-leading order (NLO) calculations are used for background cross sections, the cross sections used for the Higgs boson signal samples are computed at next-to-NLO order [10]. For all processes, the detector response is simulated using a detailed description of the CMS detector, based on the GEANT4 package [29], including pile-up distributions as observed in data.

### 3 Event Selection

The dominant reducible backgrounds are $W$, $Z$, and $t\bar{t}$ events with at least one genuine lepton and at least one quark or gluon jet misidentified as a lepton or $\tau_h$. The irreducible backgrounds are $WZ$ and $ZZ$ diboson events which have three or four genuine isolated leptons in the final state.
3.1 Object Identification

All channels presented in this document use a strategy identical to the inclusive CMS $H \rightarrow \tau\tau$ search for identifying $e$, $\mu$, $\tau_h$, and jet candidates. The requirements are described in detail in Ref. [13]. Briefly, all lepton candidates are required to be associated to the reconstructed primary vertex in the event, which is chosen to be the vertex with the highest sum of associated track squared transverse momenta. To reject events where a quark or gluon jet is incorrectly identified as a muon or electron candidate, the sum of the transverse momenta of the particle flow charged hadron, photon, and neutral hadron candidates is computed in a $\Delta R = \sqrt{\Delta \eta^2 + \Delta \phi^2}$ isolation cone of 0.4 about the lepton candidate. This isolation sum ($I$) is corrected for contamination from multiple proton-proton interactions using the $\Delta \beta$ method [30].

The isolation sum is required to be less than 10-30% of the total transverse momentum of the electron or muon, depending on the subchannel. The $\tau_h$ candidates are reconstructed using the HPS algorithm decay mode finding technique. The $\tau_h$ candidates are required to be isolated by using a combination of information from the tracker and the calorimeters (“Combined Isolation” [24]) or using a multivariate analysis (MVA) of reconstructed tracks and energy deposits (“MVA Isolation” [13]). Additionally, $\tau_h$ candidates are required to pass criteria [13] which prevent electrons and muons from being misidentified as $\tau_h$ candidates. Jets are required to satisfy a multivariate discriminant which suppresses jets coming from pile-up interactions which are not associated to the primary event vertex.

3.2 $\ell\ell\tau_h$ Channels

Events are selected online using the double muon or $\mu + e$ trigger in the $\mu\mu\tau_h$ and $e\mu\tau_h$ channels, respectively. The leading (subleading) light lepton candidate is required to have $p_T$ larger than 20 GeV/c (10 GeV/c). Muons, electrons, and $\tau_h$ candidates are required to satisfy $|\eta| < 2.4, 2.5, \text{and } 2.3$, respectively. The $\tau_h$ candidate is required to have $p_T > 20$ GeV/c and pass the tight muon rejection and MVA electron rejection [13]. To further reduce the rate of background events where an electron or muon is misidentified as a $\tau_h$ candidate, $\tau_h$ candidates within 0.4 of a reconstructed muon or electron candidate are removed. The probability for a quark or gluon jet to pass the $\tau_h$ identification is 10 to 100 times greater than the probability for a jet to pass the electron or muon identification. To remove the large background from $Z \rightarrow \ell^+\ell^-$ events with a quark or gluon jet misidentified as a $\tau_h$ candidate, the two light leptons ($e\mu, \mu\mu$) are required to have the same charge.

To suppress ZZ background events, events with additional identified electrons and muons are removed. Events containing jets with $p_T > 20$ GeV/c and $|\eta| < 2.4$ that are identified as coming from b-quarks [31] are discarded to suppress $t\bar{t}$ background events. The lepton candidates in reducible background events have a softer $p_T$ spectrum than those coming from AP Higgs boson production. These backgrounds are reduced by requiring that the scalar sum of the $E_T$ of the lepton candidates ($L_T$) be greater than 80 GeV/c.

3.3 $\ell\tau_h\tau_h$ Channels

Events are selected online using a single isolated $\mu$ trigger or an $e + \tau_h$ cross trigger in the $\mu\tau_h\tau_h$ or $e\tau_h\tau_h$ channel, respectively. The $e + \tau_h$ cross trigger is the same trigger used in the inclusive $e + \tau_h$ analysis. In both the $\mu\tau_h\tau_h$ and $e\tau_h\tau_h$ channels, the light lepton candidate is required to satisfy the identification and isolation criteria, have $p_T > 24$ GeV/c, and to be within a pseudorapidity range of $|\eta| < 2.1$. In the $e\tau_h\tau_h$ channel, the reconstructed $e$ candidate is vetoed if it is within the barrel-endcap transition region of the ECAL, $1.442 < |\eta| < 1.56$, where the increased material decreases the purity performance of the $e$ reconstruction.
The leading and sub-leading $\tau_h$ candidates form the Higgs boson candidate and are required to be of opposite charge, and to have $p_T > 45 \text{ GeV}/c$ and $p_T > 30 \text{ GeV}/c$, respectively. Both $\tau_h$ candidates are required to be within the pseudorapidity range of $|\eta| < 2.3$. The objects which can be mis-identified as $\tau_h$ candidates, real electrons, real muons, and jets, contribute at different rates to the leading and sub-leading $\tau_h$ candidates. Accordingly, the identification selections use different working points for the two $\tau_h$ candidates. The leading $\tau_h$ is required to satisfy the isolation, muon rejection, and electron rejection requirements using the tight, tight, and loose working points [13], respectively. The sub-leading $\tau_h$ is required to satisfy the corresponding requirements using the medium, tight, and medium working points, respectively.

To reduce the contribution from $t\bar{t}$ background events, events containing jets within $|\eta| < 2.4$ and $p_T > 20 \text{ GeV}/c$ that are identified as coming from b-quarks [31] are vetoed. The final state is characterized by a certain amount of missing energy in the transverse plane coming from the undetected neutrinos. Thus only events with $E_T^{\text{miss}}$ larger than $20 \text{ GeV}$ are preserved for further analysis. To remove events where two jets are mis-identified as $\tau_h$ candidates, e.g. $Z \to \ell \ell + 2\text{jet}$, events with an additional reconstructed $\mu$ or $e$ candidate passing the identification criteria are rejected. To remove $Z \to \tau^+\tau^- \to \ell^+\ell^+\tau^+_h\tau^-_h$ events, where an additional jet is mis-identified as a $\tau^\pm$ candidate, the opposite charge $\ell^\pm\tau^\mp_h$ system is required to have an invariant mass greater than $80 \text{ GeV}/c^2$ and $p_T < 50 \text{ GeV}/c$.

To further remove $Z \to ee$ contamination in the $e\tau_h\tau_h$ channel, an additional second electron veto is applied. If there are two opposite sign electrons in the event whose invariant mass falls within the window of $|M(e_1,e_2) - M_Z| < 25 \text{ GeV}/c^2$ around the nominal Z mass the event is discarded. Furthermore, the event is rejected if the invariant mass of the selected electron candidate and the hadronic tau with opposite sign falls within the window of $|M(e,\tau) - M_Z| < 6 \text{ GeV}/c^2$. This requirement recovers events where an electron has been misidentified as a hadronic $\tau$ decay. Finally, events where the selected electron candidate and the opposite sign hadronic tau are separated in $\Delta R$ by less than 0.01 are also discarded.

In addition, to further remove events from $Z \to \tau\tau$ in which one tau decays leptonically and the other decays hadronically, the transverse mass [13] ($m_T$) of the light lepton and $E_T^{\text{miss}}$ system is required to be greater than $20 \text{ GeV}/c^2$ and $50 \text{ GeV}/c^2$, in the $\mu\tau_h\tau_h$ and $e\tau_h\tau_h$ channels, respectively. The $m_T$ requirement further ensures that the selections used in the $\ell\tau_h\tau_h$ channels are exclusive to those used in the GF and VBF CMS $H \to \tau\tau$ search [13].

### 3.4 $\ell\ell\ell\ell$ Channels

The search for ZH production is performed in four-lepton ($\ell\ell\ell\ell$) events with a reconstructed pair of electrons or muons consistent with the decay of a Z boson, and a tau pair (Higgs boson) candidate. The dominant backgrounds in the $\ell\ell\ell\ell$ channels are reducible $Z + 2\text{jet}$ events, where the jets are mis-identified as the Higgs boson candidate, and irreducible ZZ production. Four final states are considered for the Higgs boson candidate: $e\mu$, $e\tau_h$, $\mu\tau_h$, and $\tau_h\tau_h$. The events are selected online using either a double muon or a double electron trigger. All $e$, $\mu$, and $\tau$ candidates are required to have $|\eta| < 2.5$, 2.4, and 2.3, respectively. All four lepton candidates are required to be at least 0.1 in $\Delta R$ away from all other lepton candidates.

The leading and sub-leading leptons associated to the Z boson are required to have $p_T$ larger than 20 GeV/c and 10 GeV/c, respectively, and be of opposite charge, and the total energy deposit in the isolation region must be less than 25% of the lepton $p_T$. To reject background events with leptonic decays of $t\bar{t}$, the invariant mass of the Z candidate is required to be between 60 and 120 GeV/$c^2$. Events with additional identified electrons and muons are removed. Events containing jets with $p_T > 20 \text{ GeV}/c$ and $|\eta| < 2.4$ that are identified as coming from
b-quarks [31] are discarded to suppress \( t \bar{t} \) background events.

The leptons associated to the Higgs boson candidate are required to have opposite charge, \( p_T > 10 \text{ GeV}/c \) if an electron or muon, and \( p_T > 20 \text{ GeV}/c \) if a \( \tau_h \) candidate. The object identification requirements depend on the subchannel, depending on the contributions from the reducible and irreducible background processes. To reduce the contribution from Z+jets backgrounds, muons are required to have the relative isolation less than 0.15 (0.25) in the \( \ell \ell \mu \tau_h \) (\( \ell \ell e \mu \)) channel, electrons are required to have 0.10 (0.25) in the \( \ell \ell e \tau_h \) channel, and \( \tau_h \) candidates are required to pass the tight (medium) combined isolation in the \( \ell \ell \mu \tau_h \), \( \ell \ell e \tau_h \), channels. To reduce the contribution from \( ZZ \) events where an electron or muon is misidentified as a \( \tau_h \), \( \tau_h \) candidates are required to pass the tight, medium, and loose muon rejection working points, and the loose, MVA, and medium electron rejection working points in the \( \ell \ell e \tau_h \), \( \ell \ell \mu \tau_h \), and \( \ell \ell \tau_h \tau_h \) subchannels, respectively.

4 Background Estimation

The irreducible diboson backgrounds are \( WZ \) and \( ZZ \) events in the \( \ell \ell \tau_h \) and \( \ell \tau_h \tau_h \) channels, and \( ZZ \) events in the \( \ell \ell LL \) channels, and are estimated using simulation. The reducible backgrounds have at least one fake lepton in the final state due to a misidentified quark or gluon jet which passes the lepton identification and are estimated solely using data.

The misidentification probabilities as a function of candidate \( p_T \), \( f(p_T) \), for the fake lepton candidates (e, \( \mu \), or \( \tau_h \)) to pass the final identification and isolation criteria are measured in independent, highly pure control samples of multijet, \( W \to \mu \nu + \text{jet} \), and \( Z \to \mu \mu + \text{jet} \) events. The control regions are chosen such that signal events are excluded due to different final state topology, inverted isolation requirements, or both. To minimize possible biases, the same trigger, kinematic, and quality criteria used in the final analysis are applied to the control samples. Sidebands are defined for each channel, where the final identification or isolation criterion is not satisfied for one or more of the final-state lepton candidates. The sidebands are enriched in background events. The number of reducible background events due to a lepton candidate being misidentified in the final selection is estimated by weighting the events according to the observed fake lepton candidates in the sideband and their probability \( f(p_T)/ (1 - f(p_T)) \) to pass the final identification and isolation criteria.

Which lepton candidates are used as “misidentified” objects depends on the charge assignments of the leptons. In the \( \ell \ell \tau_h \) channels, at least one of the two (same-charge) light leptons is misidentified in all reducible backgrounds, and the total fake leptons estimation comes from the sum of estimates using each light lepton in turn. The background from \( \ell^\pm \ell^\mp \) events with a fake \( \tau_h \) and where the charge of one of the light leptons is reconstructed incorrectly is negligible. In the \( \ell \ell LL \) channels, at least one of the Higgs boson candidate leptons is misidentified in all reducible backgrounds, so the total fake estimate is the sum of the estimates from both. The sum of the two estimates from two leptons counts backgrounds where both leptons are fake candidates (\( Z + 2 \text{ jets}, \) multijet, etc.) twice. This double-counting is removed by inverting the identification of both leptons and applying both misidentification weights simultaneously, giving an independent estimate of events with two fake leptons. The estimate of the double-counted background is subtracted from the total sum. In the \( \ell \tau_h \tau_h \) channels, the \( \tau_h \) candidate which has the same charge as the light lepton candidate is misidentified in the background contribution given by the fake tau leptons. Therefore the fake estimate in the \( \ell \tau_h \tau_h \) channel is computed using only the same-charge \( \tau_h \) candidate, and no correction for double counting is required.
Table 1: Observed events and expected yields from the different background processes for the three and four-lepton channels. The uncertainties represent the combined statistical and systematic uncertainties.

<table>
<thead>
<tr>
<th>Process</th>
<th>$\ell\ell\tau_h$</th>
<th>$\ell\tau_h\tau_h$</th>
<th>$\ell\ell LL$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reducible backgrounds</td>
<td>26.3 $\pm$ 4.7</td>
<td>20.8 $\pm$ 4.2</td>
<td>25.2 $\pm$ 10.0</td>
</tr>
<tr>
<td>WZ</td>
<td>35.3 $\pm$ 3.9</td>
<td>6.3 $\pm$ 0.9</td>
<td></td>
</tr>
<tr>
<td>ZZ</td>
<td>2.5 $\pm$ 0.3</td>
<td>0.39 $\pm$ 0.08</td>
<td>27.2 $\pm$ 3.8</td>
</tr>
<tr>
<td>Total bkg.</td>
<td>64.1 $\pm$ 6.2</td>
<td>27.5 $\pm$ 4.3</td>
<td>52 $\pm$ 11</td>
</tr>
<tr>
<td>$VH \to V\tau\tau (m_H = 125 \text{GeV}/c^2)$</td>
<td>3.6 $\pm$ 0.4</td>
<td>1.2 $\pm$ 0.2</td>
<td>2.1 $\pm$ 0.2</td>
</tr>
<tr>
<td>$VH \to VW\tau (m_H = 125 \text{GeV}/c^2)$</td>
<td>0.50 $\pm$ 0.05</td>
<td>0</td>
<td>1.13 $\pm$ 0.09</td>
</tr>
<tr>
<td>Observed</td>
<td>65</td>
<td>36</td>
<td>66</td>
</tr>
</tbody>
</table>

In the $\ell\ell\tau_h$ and $\ell\tau_h\tau_h$ channels the irreducible WZ background is estimated using MC simulation. The ZZ background is largely reduced by the veto of events containing an additional reconstructed $e$, $\mu$, or $\tau_h$ candidate, and is estimated using MC simulation. In the $\ell\ell LL$ channels, WZ events have at least one fake lepton and are estimated using the misidentification probabilities described above. The dominant background comes from the ZZ events, which are estimated using MC simulation.

5 Results

The numbers of events in data and expected yields from all backgrounds are enumerated in Table 1. The observed data are compatible with the background expectation within the errors.

The primary observable used in this analysis is the visible invariant mass of the tau pair associated to the Higgs boson candidate. The visible mass is the invariant mass of all visible tau pair decay products as $e$, $\mu$, $\pi^\pm$ and $\pi^0 \to \gamma\gamma$. In the $\ell\ell\tau_h$ channels, it is not possible to definitively assign the same-charge electrons or muons to either the $W$ or the Higgs boson candidate. However, as the signal is expected to be dominated by $H \to \tau\tau$ decays, the final-state light leptons produced in the decays of the $\tau$ leptons have a softer $p_T$ spectrum than light leptons from $W \to \ell\nu$ decays, as they are associated with two neutrinos instead of one. Accordingly, we assign the subleading light lepton to the Higgs boson candidate rather than to the $W$ boson. The visible mass spectra and the expected Higgs boson contribution are shown in Fig. 1. The yields of each process are determined using the normalization of the signal and background given by the best fit to data using the maximum likelihood method [32] which is also used in the limit setting procedure.

6 Systematic Uncertainties

The efficiencies for the Higgs boson signal and some of the background samples are estimated using MC simulation. Where possible, these efficiencies, e.g., muon reconstruction efficiency, are measured in control regions in data, and residual differences between the efficiencies in the MC simulation and data are corrected by scaling the simulation to match the efficiency measured in data. The uncertainties on the simulation-to-data correction factors are taken as systematic uncertainties, and are propagated to the final results.

The trigger, identification, and isolation efficiencies for electrons and muons are measured with
Figure 1: Observed and expected Higgs boson candidate mass spectra in the ℓℓτ_h (left), ℓℓLL (middle), and ℓℓℓLL (right) channels. The expected contribution from the associated production of a SM Higgs boson with mass m_H = 125 GeV/c^2 is shown by the dashed line.

data using the “tag-and-probe” technique [33] in Z → ℓℓ events. Residual differences between the lepton efficiencies in the MC simulation and data are corrected by scaling the simulation to match the efficiency measured in data. The uncertainty on muon and electron candidate efficiencies depends on the number of muons or electrons in the final state and varies between 1 and 3%. The τ_h identification efficiency is measured with an uncertainty of 6% (per τ_h candidate) using the tag-and-probe technique in Z→ ττ → μτ_h events [24]. This will lead to a 12% systematic uncertainty when two τ_h are reconstructed in the final state. There is a 2% scale uncertainty on the τ_h energy reconstruction, which propagates into a 3% uncertainty on the signal yield in the ℓτ_hτ_h channels, and is negligible in the ℓℓτ_h and 4ℓ channels. The uncertainty on the trigger efficiencies are approximately 1%, with the exception of the ℓτ_hτ_h e + τ_h cross trigger, which has an uncertainty of 3.5%. The uncertainty on the rate of light-quark and gluon jets to be misidentified as b-tagged jets and fail the b-veto is 15%, which propagates into a 1% uncertainty on the yields of the simulated samples. Uncertainties on the jet energy scale (JES) have been evaluated in Z + jet and γ + jet events [23] and vary between 2 and 5%. The effect on the signal yields due to JES is approximately 1%.

In the ℓτ_hτ_h channels, there is a systematic uncertainty on the efficiency of the m_T selection, which depends on the reconstruction of the E_T^{miss}. This uncertainty on the E_T^{miss} reconstruction is driven by the electron, muon, tau, and jet energy scales, and the unclustered energy scale. The unclustered energy, which is defined as the energy not associated with the reconstructed leptons and jets with p_T > 10 GeV/c, has a scale uncertainty is 10%. The uncertainty on the unclustered energy scale results in a 3.7% uncertainty on the ℓτ_hτ_h signal acceptance. Other contributions from the e-veto and μ-veto requirements are found to be 3.8% and 0.7%, respectively.

There is a 2.2% uncertainty [16] on the total integrated luminosity of the collected 7 TeV data sample, and 4.4% on the 8 TeV data sample. Two theoretical systematic uncertainties on the overall signal yield are considered. The uncertainty on the QCD factorization and renormalization scale affect the expected signal cross section. The effect of variations in the parton distribution functions, the value of α_s, and higher-order corrections are propagated to the efficiency of the signal selection using the PDF4LHC prescription [34–38], and affect the expected signal yields between 4 and 15%.

For the fake leptons background estimation, two systematic uncertainties are considered. The dominant effect is statistical and is due to limited numbers of events in the weighted control re-
Table 2: Systematic uncertainties estimated for $\ell\ell\tau_h$, $\ell\tau_h\tau_h$ and $\ell\ell\ell$ channels.

<table>
<thead>
<tr>
<th>Source</th>
<th>Uncertainty</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu$ and e reconstruction efficiencies</td>
<td>from 1 to 3%</td>
</tr>
<tr>
<td>$\tau_h$ identification</td>
<td>from 6 to 12%</td>
</tr>
<tr>
<td>$\tau_h$ energy scale ($\ell\tau_h\tau_h$)</td>
<td>3%</td>
</tr>
<tr>
<td>trigger efficiency</td>
<td>from 1 to 3.5%</td>
</tr>
<tr>
<td>$b$-tagged jet veto</td>
<td>1%</td>
</tr>
<tr>
<td>jet energy scale</td>
<td>1%</td>
</tr>
<tr>
<td>unclustered energy scale ($\ell\tau_h\tau_h$)</td>
<td>3.7%</td>
</tr>
<tr>
<td>e veto ($\ell\tau_h\tau_h$)</td>
<td>3.8%</td>
</tr>
<tr>
<td>$\mu$ veto ($\ell\tau_h\tau_h$)</td>
<td>0.7%</td>
</tr>
<tr>
<td>integrated luminosity 2011(2012)</td>
<td>2.2% (4.4%)</td>
</tr>
<tr>
<td>PDF (PDF4LHC prescription)</td>
<td>from 4 to 15%</td>
</tr>
<tr>
<td>non-prompt background estimation</td>
<td>from 15 to 35%</td>
</tr>
<tr>
<td>WZ cross-section</td>
<td>12%</td>
</tr>
<tr>
<td>ZZ cross-section</td>
<td>10%</td>
</tr>
</tbody>
</table>

gion. The sub-dominant effect is due to an imprecise knowledge of the lepton misidentification probability $f(p_T)$. The statistical uncertainty depends on the subchannel and varies between 5 and 30%. The effect of this uncertainty on the Higgs boson candidate visible mass spectra is taken into account by adding an independent uncertainty [39] for each bin in the spectra. The uncertainty due to the misidentification rate varies between 15-35% and is dominated by the difference between the misidentification rates measured in $Z$, $W$, and multijet events in the $\ell\ell\tau_h$ and $\ell\tau_h\tau_h$ channels and the uncertainty on the measured fake rates in the $\ell\ell\ell$ channels.

For the $\ell\ell\tau_h$ and $\ell\tau_h\tau_h$ channels, the associated uncertainty on the WZ diboson backgrounds is 12% and is taken from the 2011 CMS cross section measurements. For the $\ell\ell\ell$ channels, the theoretical uncertainty of 10% on the ZZ production cross section [10] dominates the uncertainty on the estimate of the ZZ background. A summary of the systematic uncertainties can be seen in Table 2.

7 Limits on Higgs boson production

The data show no evidence for the presence of a Higgs boson signal, and we set 95% CL upper bounds on the Higgs boson cross section. To obtain exclusion limits we use the asymptotic CL$_s$ method [32], based on a binned likelihood of the Higgs boson candidate visible invariant mass spectrum. Only in the 7 TeV data of the $\ell\tau_h\tau_h$ channel a pure counting experiment is being performed due to the very low number of observed events in this category. Systematic uncertainties are represented in the limit computation by nuisance parameters using a log-normal constraint. The statistical limitation concerning the MC simulated signal samples and some of the backgrounds is also considered. Correlated uncertainties among channels are represented by common nuisance parameters.

The observed and median expected 95% CL upper limits on SM Higgs boson production set by the $\ell\ell\tau_h$ and $\ell\tau_h\tau_h$, $\ell\ell\ell$, and combined analyses, are shown in Figs. 2, 3, and 4, respectively. The observed upper limit of the combined analyses is also compared to the median observed limit expected in the presence of a SM Higgs boson, using a Monte Carlo “signal injection” technique. Many pseudo-datasets are prepared, corresponding to the expected spectrum from the estimated backgrounds and the expected contribution of a SM Higgs boson, and the observed limit is computed for each pseudo-dataset. The observed limits set by this analysis are
Figure 2: The expected and observed 95% CL upper limits on SM Higgs boson production set by the $\ell\ell\tau\nu$ channels (left) and $\ell\tau\nu$ channels (right).

Table 3: Expected and observed 95% CL upper limits on SM Higgs boson production. The limit is expressed as $\sigma/\sigma_{SM}$, the ratio of the excluded production rate times branching fraction to that predicted by the SM.

<table>
<thead>
<tr>
<th>$m_{H}$ [GeV/c$^2$]</th>
<th>$-2\sigma$</th>
<th>$-1\sigma$</th>
<th>Median</th>
<th>$+1\sigma$</th>
<th>$+2\sigma$</th>
<th>Obs. Limit</th>
</tr>
</thead>
<tbody>
<tr>
<td>115</td>
<td>1.50</td>
<td>1.99</td>
<td>2.76</td>
<td>3.83</td>
<td>5.09</td>
<td>2.91</td>
</tr>
<tr>
<td>120</td>
<td>1.56</td>
<td>2.07</td>
<td>2.87</td>
<td>3.98</td>
<td>5.29</td>
<td>3.38</td>
</tr>
<tr>
<td>125</td>
<td>1.68</td>
<td>2.24</td>
<td>3.10</td>
<td>4.31</td>
<td>5.72</td>
<td>3.87</td>
</tr>
<tr>
<td>130</td>
<td>1.74</td>
<td>2.31</td>
<td>3.20</td>
<td>4.45</td>
<td>5.91</td>
<td>4.29</td>
</tr>
<tr>
<td>135</td>
<td>1.79</td>
<td>2.38</td>
<td>3.30</td>
<td>4.58</td>
<td>6.08</td>
<td>4.46</td>
</tr>
<tr>
<td>140</td>
<td>1.81</td>
<td>2.40</td>
<td>3.33</td>
<td>4.62</td>
<td>6.14</td>
<td>4.60</td>
</tr>
<tr>
<td>145</td>
<td>1.60</td>
<td>2.13</td>
<td>2.95</td>
<td>4.09</td>
<td>5.44</td>
<td>4.11</td>
</tr>
</tbody>
</table>

found to be statistically compatible with both the absence and presence of a SM Higgs boson. The combined limits are enumerated in Table 3. Both the $H\rightarrow \tau\tau$ and $H\rightarrow WW$ decay modes are considered in the overall limit on SM Higgs boson production. The ratio of the cross section times branching ratio of the two decay modes is considered to be that defined by the SM.

8 Summary

A search for the associated production of standard model Higgs bosons decaying to tau pairs at the CMS experiment is described. The search is conducted using events with three or four isolated leptons in 5.0 fb$^{-1}$ and 19.5 fb$^{-1}$ for 7 and 8 TeV CMS data, respectively. The expected contribution from other SM Higgs boson decay modes is negligible, with the exception of the $H\rightarrow WW$ decay mode. This contribution to the expected signal yield is taken into account for the three- and four-lepton channels. The data are compatible with both the background-only prediction and the presence of a SM Higgs boson. Upper limits of 2.9 to 4.6 times the predicted SM value are set at 95% CL for the product of the SM Higgs boson production cross section and
Figure 3: The expected and observed 95% CL upper limits on SM Higgs boson production set by the $\ell\ell LL$ channels.

Figure 4: The expected and observed 95% CL upper limits on SM Higgs boson production set by the combination of the analyses presented in this document are shown at left. At right, the observed limit is compared to the distribution of observed limits computed using many signal-injected pseudo-datasets.
decay branching fraction in the mass range $110 < m_H < 145 \text{ GeV}/c^2$.

References


References


