Search for a Standard Model-like Higgs boson decaying into $WW \rightarrow \ell \nu q\bar{q'}$ in pp collisions at $\sqrt{s} = 8$ TeV

The CMS Collaboration

Abstract

A search for a Standard Model (SM) Higgs boson decaying into the WW final state is performed with an integrated luminosity of up to $19.3\,\text{fb}^{-1}$ of pp collisions at $\sqrt{s} = 8$ TeV in the high mass regime $600 < m_H < 1000$ GeV. The search is performed in the semi-leptonic channel where the hadronically decaying W boson is highly boosted and its decay products are contained in one jet. Jet substructure techniques are used in identifying the hadronically decaying W. The results are also interpreted in a beyond the Standard Model heavy Higgs scenario, based on an effective theory which predicts the existence of two Higgs-like scalar particles, with the mass of the lighter one being around 125 GeV, that completes the unitarization of the WW scattering.
1 Introduction

The Standard Model (SM) of electroweak interactions [1–3] relies on the existence of the Higgs boson, \( H \), a scalar particle associated with the field responsible for spontaneous electroweak symmetry breaking [4–9]. The observation of a Higgs boson with a mass of 125 GeV [10, 11] is consistent with the theoretical constraint coming from the unitarization of WW scattering at high energies [12–21]. However, there is still a possibility that the newly discovered particle has no connection to the electroweak symmetry breaking mechanism [22, 23]; thus, it is important to continue searching for the SM Higgs boson in the high mass regime to further confirm the properties of the new boson at 125 GeV. In addition, several popular scenarios, such as general two-Higgs-doublet models (for a review see [24]) or models in which the SM Higgs boson mixes with a heavy electroweak singlet [25–38], predict the existence of additional resonances at high mass, with couplings similar to the SM Higgs boson. This study reports the search for a SM Higgs-like boson decay to two W bosons in the mass range from \( m_H > 600 \) to \( m_H < 1000 \) GeV.

The measurement of the di-boson mass spectrum in the high mass range is important regardless of the role the new boson plays in electroweak symmetry breaking, for example to investigate anomalous triple gauge couplings or to search for new resonances in this regime. In addition, by exploring the WW production with the vector-boson fusion signature, it will be possible to investigate the unitarization of the WW scattering process. The semi-leptonic final state investigated in this analysis can be triggered and separated from the QCD production due to the presence of one lepton from the leptonic decay of one W boson, and shows a large branching ratio due to the hadronic decay of the other W boson. Because of the large invariant mass searched for, the W bosons have a large transverse momentum, which causes the hadronic decay products to be collimated to the point of being reconstructed as a single jet in the detector. Jet substructure techniques are therefore employed for identifying single jets that have originated from a highly boosted, hadronically decaying W boson to identify this particular signature. Because of these characteristics, this study is a benchmark for future di-boson scattering measurements at high mass at the LHC.

The analysis has been performed on the pp-collision data sample recorded by the CMS detector at the center-of-mass energy of 8 TeV, corresponding to an integrated luminosity of 19.3 fb\(^{-1}\). The obtained results are interpreted as a search for a SM Higgs boson, as well as in the context of a beyond the SM (BSM) scenario, based on an effective theory where a second scalar particle besides the resonance at 125 GeV completes the unitarization of the WW scattering [39].

2 The CMS detector

The CMS detector, described in detail in [40], is a multipurpose apparatus designed to study high transverse momentum (pT) physics processes in proton-proton collisions. It consists of pixel and silicon-strip trackers up to a pseudorapidity of \( |\eta| < 2.5 \) which, together with a 3.8 Tesla solenoid, provide a track momentum resolution of 1% at 100 GeV; a granular electromagnetic crystal-calorimeter extending up to \( |\eta| < 3 \) with an energy resolution of about 3%/\( \sqrt{E} \) [41]; a hadronic calorimeter extending up to \( |\eta| < 5 \) with an energy resolution of 100%/\( \sqrt{E} \); and an efficient muon system capable of reconstructing and identifying muons up to \( |\eta| < 2.4 \). The detector is nearly hermetic, allowing for measurements of the missing transverse energy (\( \cancel{E}_T \)) in the event. The right-handed coordinate system is used by CMS [40] with the origin at the nominal interaction point, the x-axis pointing to the center of the LHC, the y-axis pointing up perpendicular to the LHC plane, and the z-axis along the counterclockwise-beam direction. The polar angle \( \theta \) is measured from the positive z-axis and the azimuthal angle \( \phi \) is measured
in the x-y plane. The pseudorapidity is defined as $\eta = -\ln(\tan \frac{\theta}{2})$.

## 3 Data and Simulation samples

Data have been collected by the CMS experiment with single lepton triggers, with a typical on-line $p_T$ threshold of 24 GeV for muons and 27 GeV for electrons. The pseudorapidity range for muons (electrons) extends up to $|\eta| < 2.4$ ($2.5$).

The main source of background for this analysis is due to single $W$ boson production in association with jets and it is primarily estimated from data. Monte Carlo (MC) simulations are used to estimate residual contributions of the background, and to estimate efficiency of the Higgs signal. We use the MADGRAPH 5 1.3.30 [42] event generator to simulate the $W$ boson and Drell-Yan production in association with up to four jets, referred to as $W$+jets and $Z$+jets, respectively, which are showered with the PYTHIA [43] generator. In addition, we use the HERWIG [44] generator to produce an alternate sample of $W$+jets events which has a different parton shower model than PYTHIA, in order to estimate systematic uncertainties. Top-quark pair events are simulated using POWHEG 1.0 [45–48]. Single top production is also modeled using POWHEG [45–48]. Electroweak diboson ($WW$ and $WZ$ semi-leptonic decay) processes are simulated using PYTHIA. The POWHEG generator has been used to produce Higgs signal events, and the showering has been performed with PYTHIA. For this analysis, samples with Higgs mass hypotheses ranging from 600 to 1000 GeV have been used, where the Higgs mass is reweighted according to the complex pole reweighting scheme calculation [49–52]. We also correct the signal for the effect of interference with the SM $gg\to WW$ production process using the method proposed in [50] computing the interference effect with the MCFM generator [53]. All simulated samples are corrected for any data-MC difference in the trigger and physics objects identification efficiencies.

In the BSM interpretation, we search for a electroweak singlet scalar where a heavy Higgs boson mixes with the recently discovered candidate with a mass close to 125 GeV. Phenomenologically the couplings of the two gauge eigenstates (SM and singlet) become related by unitarity and the original coupling strength of the light Higgs boson is therefore reduced with respect to the SM case. If we define $C$ ($C'$) as the scale factor of the couplings of the low (high) mass with respect to the SM, one can write $C^2 + C'^2 = 1$ as the unitarity condition to be preserved. Indirectly we can set an upper limit at 95% CL on $C'^2 < 0.446$ using the signal strength fits to the $H(125)$ candidate as obtained in [54]. The heavy Higgs cross cross-section is modified by a factor, $\mu'$, and the modified width is $\Gamma'$; they are defined as:

\[
\mu' = C^2(1 - BR_{\text{new}}) \quad (1)
\]

\[
\Gamma' = \Gamma_{\text{SM}} \times \frac{C'^2}{(1 - BR_{\text{new}})} \quad (2)
\]

where $BR_{\text{new}}$ is the branching ratio of the heavy Higgs to non-SM-like decay modes. The BSM heavy Higgs line shape is evaluated by reweighting the SM POWHEG samples following the procedure outlined in [55]. The model is implemented as a rescaling of the SM at NLO in QCD and LO in EWK. In this reweighting procedure we set as target line shape a relativistic Breit-Wigner with a narrower signal width with respect to the width of the SM Higgs boson. Notice that due to the poor resolution of this channel departures from the Breit-Wigner approximation are not expected to modify significantly the final results. The contribution from the interference term between the BSM Higgs and the background is furthermore assumed to scale according to the modified coupling of the Higgs boson as: $(\mu + I)_{\text{BSM}} = \mu_{\text{SM}}C'^2 + I_{\text{SM}}C'$ where $\mu$ ($I$) is the signal strength (interference) in the BSM and SM cases. This assumption is based on the
hypothesis that the couplings are similar to the SM case and simply re-scaled due to unitarity constraints. Intuitively one may expect that if the new resonance has a very small width its production will tend to interfere less with the background continuum. By using the MCFM generator we have cross-checked further this hypothesis. We assign an extra uncertainty of 100% to the interference term after the results obtained from these generator level studies.

4 Event Reconstruction

The final state signature of this analysis is an isolated electron or muon, missing transverse energy ($E_T$), and a jet containing the entire hadronic decay of a $W$ boson.

4.1 Leptonically decaying $W$ boson, $W_\ell$

Muons are measured with the tracker and the muon system, within $|\eta| < 2.1$. Electrons are detected as tracks in the tracker pointing to energy clusters in the ECAL, within $|\eta| < 2.5$, excluding the transition region between the barrel and endcap, $1.44 < |\eta| < 1.57$. Muons (electrons) are required to have a momentum transverse to the beam direction, $p_T$, greater than 30 GeV (35 GeV). The lepton candidates are required to be compatible with the primary vertex of the event, which is chosen as the vertex with the highest $\sum p_T^2$ of its associated tracks. According to simulation, this requirement provides the correct assignment for the primary vertex in more than 99% of cases in both signal and background events. Charged leptons from $W$ boson decays are required to be isolated from other activity in the event and to be consistent with the primary vertex. The isolation of $e$ or $\mu$ leptons is therefore ensured by applying requirements on the sum of the transverse energies of all reconstructed particle-flow particles [56], charged or neutral, within a cone of $\Delta R < 0.4$ around the lepton direction, after subtracting the average pileup energy estimated using a jet area technique [57] on an event-by-event basis. We define $\Delta R \equiv \sqrt{(\Delta \eta)^2 + (\Delta \phi)^2}$ where $\Delta \eta$ and $\Delta \phi$ are differences in pseudorapidity and in azimuthal angle between two measured particles in radians. To reduce the backgrounds from the Drell-Yan and SM diboson processes, we veto the presence of any other loosely identified lepton ($p_T > 20$ GeV for electrons, 10 GeV for muons) in the event.

The leptonic $W$ candidate is reconstructed from the lepton and the missing transverse energy, $E_T$, in the event. Therefore, an accurate $E_T$ measurement is essential for reconstructing the full event kinematics of the $WW$ system. We use $E_T$ measured in the event from particle-flow reconstruction. The $E_T$ resolution, measured as a function of the sum $E_T (\sum E_T)$ of the particle-flow objects in the event, varies from 4% at $\sum E_T = 60$ GeV to 10% at $\sum E_T = 350$ GeV[58]. We require $E_T > 25$ (30) GeV for each event in the muon (electron) data. In addition, the unmeasurable longitudinal component of the neutrino momentum is reconstructed by requiring the lepton-neutrino pair to have the invariant mass of a $W$ boson. The ambiguity in the second-order equation is resolved by taking the solution that yields the smallest $|p_z|$ value which, when evaluated in simulation, is closest to the truth $|p_z|$ approximately 70% of the time.

4.2 Hadronically decaying $W$ boson, $W_{\text{had}}$

Jets are reconstructed from calorimeter and tracker information using the particle flow algorithm, combining the information from all CMS sub-detectors to reconstruct each particle. In addition, charged-particle tracks not originating at the primary vertex are not considered for jet clustering [59]. A jet quality requirement, primarily based on the energy balance between charged and neutral hadrons, is also applied.

Two different clustering algorithms, the anti-$k_T$ (AK) and Cambridge-Aachen clustering al-
Event Selection

Cambridge-Aachen jets are clustered with a distance parameter of $\Delta R = 0.8$ (CA8) and are used for reconstructing the hadronically decaying $W$ boson, $W_{\text{had}}$, where the entire $W$ boson decays into a single jet. Anti-$k_T$ jets clustered with a distance parameter of $\Delta R = 0.5$ (AK5) are used primarily for b-quark jet tagging. The CA8 jets with larger $R$ parameter are used to increase signal acceptance of $W_{\text{had}}$.

For signal events falling in the detector acceptance, approximately 65% (82%) of the $W_{\text{had}}$ decay products are contained in a cone of $\Delta R < 0.8$ for a signal mass of 600 GeV (1 TeV). Alternatively, approximately 10% (42%) of the $W_{\text{had}}$ decay products are separated by a distance of $\Delta R < 0.5$ for a signal mass of 600 GeV (1 TeV) and would not be reconstructed by the standard CMS AK5 jet finding algorithm; thus, the larger cone CA8 jet affords more signal acceptance in the single jet signature while not losing events when the decay products are separated by $\Delta R < 0.5$. In addition, in the analysis where the jets are resolved, the combinatorial background is much larger.

5 Event Selection

The large invariant mass of the signal translates into a large transverse momentum of the $W$ bosons. Therefore, the $p_T$ of both leptonically decaying $W$ and of the hardest jet in the event are required to be above 200 GeV. We choose the highest $p_T$ CA8 jet in the event to uniquely be the $W_{\text{had}}$ candidate. At the same time, the transverse missing energy is required to be above 50 GeV (70 GeV) for the muon (electron) channel to suppress the contribution to the analysis from QCD. In addition, there are specific topological cuts which require that the $W$ bosons are back-to-back:

- $\Delta R_{l,j} > \pi/2$; the distance between the lepton and the jet should be large
- $\Delta \Phi_{E_{T},j} > 2.0$; the azimuthal distance between the missing energy and the jet should be large
- $\Delta \Phi_{V,j} > 2.0$; the azimuthal distance between the leptonically decaying $W$ boson and the jet should be large

In addition to these kinematic cuts, to reduce the amount of the $t\bar{t}$ background, we veto the presence of any b-tagged AK5 jets in the event by selecting on the CSV discriminant [63], choosing a 70% on signal efficiency as optimal working point. After applying the kinematic selections described above, we make comparisons of the data and MC for various kinematic observables in Fig. 1. The $p_T$ of $W_l$ (left) and CA8 jet $p_T$ (right) for the muon (top) and electron (bottom) channels are presented. From Fig. 1, we find that $W$+jets events are predominant background and the next leading background contribution comes from $t\bar{t}$ events. In Fig. 1, the signal distributions are also shown for a SM signal mass hypothesis of 600 GeV where the gluon fusion (ggH) and vector boson fusion (vbfH) production mechanisms are shown separately.

After these kinematic selections, cuts are applied on the jet substructure to further discriminate jets coming from a hadronically decaying $W$ boson from QCD jets originating from quarks and gluons. The primary observable used to separate jets originating from $W$ boson decays from QCD ones is the jet mass itself with improved separation between signal and background by means of a jet grooming algorithm [64]. In particular, for this analysis the pruning implementation has been chosen [65], as it is specifically designed to remove soft QCD and pileup contributions from the jet. Fig. 2, left, shows the invariant mass of the $W$ boson candidates, after the analysis selections described in Section 5, for a signal SM Higgs boson of 600 GeV and for the simulated background of $W$+jets. Dashed lines correspond to the distributions before the
Figure 1: Leptonic $W p_T$ (left) and CA8 jet $p_T$ (right) for the muon (top) and electron (bottom) channels.
application of the pruning algorithm, while continuous lines after pruning. Jets reconstructed from \( W_{\text{had}} \) show a distinctive peak around the W mass, which is narrowed by the pruning, while the background jets get correctly assigned a smaller mass on average, enhancing the separation between the two samples. In addition to the pruned jet mass, we use additional information about the jet shape to further reduce QCD jets coming from the background. N-subjettiness was introduced in [66] and is a generalized jet shape observable which defines a measure, \( \tau_N \), for a jet to have N subjets where \( \tau_N \) is defined as:

\[
\tau_N = \frac{1}{d_0} \sum_i p_T^{i} \min\{ (\Delta R_{1,i})^\beta, (\Delta R_{2,i})^\beta, \ldots, (\Delta R_{N,i})^\beta \}
\]

\[d_0 = \sum_i p_T^{i}(R_0)^\beta\]  \hspace{1cm} (3)

The sum runs over all particle constituents of the jet and are minimized based on their distance to the N subjets axes. The characteristic jet radius, \( R_0 \), is taken from the original clustering algorithm and in this study is, \( R_0 = 0.8 \), with a \( \beta = 1 \). The value of \( \tau_N \) tends to zero as the jet becomes more consistent with N subjets. Ratios of \( \tau_N \) are found to be especially powerful in separating signal from background and we use \( \tau_2/\tau_1 \) as the additional discriminating observable. In Fig. 2, the \( \tau_2/\tau_1 \) distribution can be seen after selecting a the jet to have a pruned mass from 65-105 GeV. In this figure, one can see that \( \tau_2/\tau_1 \) provides additional discrimination between signal and background. Two signal masses are shown, 600 GeV and 1000 GeV, where the difference between the two is small. We cut on the N-subjettiness variable \( \tau_2/\tau_1 \) with one-pass \( k_T \) axis optimization in order to further reduce background contributions. The cut values on the pruned jet mass and \( \tau_2/\tau_1 \) are given in Sec. 5.

![Figure 2](image-url)

Figure 2: Signal and background distributions from simulation of pruned jet mass (left) and \( \tau_2/\tau_1 \) (right) after analysis level cuts described in Section 5. On the left plot, we also show the ungroomed jet mass as dotted lines to show the effect of pruning.

The pruned jet mass and \( \tau_2/\tau_1 \) cuts are exclusively used to identify the hadronically decaying W boson, \( W_{\text{had}} \). A jet passing this criteria is referred to as a W-tagged jet. The final WW system is defined using the four-vector of \( W_\ell \) and the ungroomed four-vector of \( W_{\text{had}} \).

Since the jet substructure reproduced in Monte Carlo events depends on the details used in the simulation, a sample of \( W_{\text{had}} \) have been isolated to study the effect of jet selections in
data. By applying all analysis requirements, but instead to require at least one b-tagged jet in the event, a sample of W boson decaying hadronically into a single jet can be isolated in a sample of nearly pure $t\bar{t}$ and single-top events. We use this top-enriched control sample to validate the hadronically decaying W boson selection. The distribution of $\tau_2/\tau_1$ in this control sample is shown in left plot of Fig. 3, while the right shows the pruned jet mass distribution after applying a cut on $\tau_2/\tau_1 < 0.5$. From the comparison between data and Monte Carlo, a normalization correction factor for $t\bar{t}$ and single top is evaluated in the signal region. The scale factor is measured to be $0.95 \pm 0.06 (0.92 \pm 0.06)$ in the muon (electron) channel.

A simultaneous fit to the jet mass distributions before and after the jet mass and $\tau_2/\tau_1$ cuts are used to extract a data-to-MC efficiency scale factor for identifying merged W bosons. The scale factor for W-tagging is $0.95 \pm 0.10 (0.86 \pm 0.10)$ in the muon (electron) channel. In addition, the $W_{had}$ peak mass and resolution are extracted from the same fit and are measured to be $83.4 \pm 0.4$ GeV and $7.4 \pm 0.4$ GeV in simulation and $84.5 \pm 0.4$ GeV and $8.7 \pm 0.6$ GeV in data. The larger resolution in data with respect to that of the simulation is in agreement with past CMS measurements of jet energy resolution [67].

![Figure 3](image_url)

Figure 3: The $\tau_2/\tau_1$ distribution in the top-enriched control sample is shown on the left. The pruned jet mass distribution in the top-enriched control sample after applying a cut on $\tau_2/\tau_1$ for the muon channel is shown on the right including the fit to both data and simulation.

### 6 Background and Signal Estimation

The final discriminating variable in the analysis is the shape of the three-body $m_{llJ}$ distribution. The signal normalization and shape is estimated from MC with data-to-MC corrections applied for mass scale, normalization and resolution extracted from a top-enriched control sample. The non-dominant background processes $t\bar{t}$, single top and dibosons (WW/WZ) are estimated with the simulation, taking from MC both the shapes and the normalizations, appropriately corrected to match data by using the scale factors derived in the top-enriched control sample. Background contributions from inclusive QCD processes are negligible due to the large requirement on $E_T$.

We estimate the main W+jets background contribution extracted from data via the sideband method to extract the shape and normalization. As mentioned previously, the signal region
Background and Signal Estimation

(SR) is defined around the W boson mass, requiring the pruned jet mass to fall within \( m_J = [65-105] \) GeV, while an event falls in the lower sideband (LSB) if \( m_J \) has a value in \([40-65]\) GeV, or in the upper sideband region when \( m_J = [105,130] \) GeV.

The W+jets normalization is obtained from a sideband fit of data in the \( m_J \) spectrum, as shown in Fig. 4 (left), where the signal region is not used in the fit. All background contributions except for W+jets are fixed from the simulation, and all the distributions are parametrized with functions determined with dedicated fits on each simulated sample. The uncertainty band includes contributions from the fit parameter errors and normalization uncertainties for the non-dominant background contributions detailed more in Sec. 7.

The W+jets shape in the signal region is determined from the lower sideband, through an extrapolation function \( \alpha_{MC}(m_{\ell \nu j}) \) derived from the W+jets simulation, defined as:

\[
\alpha_{MC}(m_{\ell \nu j}) = \frac{F_{MC,SR}(m_{\ell \nu j})}{F_{MC,LSB}(m_{\ell \nu j})}
\]

where the function \( \alpha_{MC}(m_{\ell \nu j}) \) is determined from MC in order to account for the correlations between the jet mass and the mass of the three-body system and \( F_{MC,SR}(m_{\ell \nu j}) \) and \( F_{MC,LSB}(m_{\ell \nu j}) \) are the probability density functions used to describe the \( m_{\ell \nu j} \) spectrum in simulation for the signal region and low sideband region, respectively. In particular, the HERWIG sample is used, as this generator is known to better describe the jet substructure [64]. The shapes of \( \alpha_{MC}(m_{\ell \nu j}) \) can be seen in Fig. 4 (right). The red and blue lines are the fits of the \( m_{\ell \nu j} \) shape in the signal and in the low sideband region, respectively. Their ratio is given by the black line and the black (green) shaded regions correspond to the \( 1\sigma \) (2\( \sigma \)) bands of the fit parameters. The structure of the uncertainty band is due to the functional form that is chosen. The difference between the HERWIG and PYTHIA determinations is taken as an additional systematic effect due to the choice of the alternate parton shower model, labeled as “Alternate PS” in Fig. 4. The systematic

![Figure 4: Data and MC distribution for the \( m_J \) sideband and signal region is shown on the left. The normalization W+jets contribution is extracted from a fit to the sideband region excluding the signal one. On the right, the extrapolation function \( \alpha \) from sideband to signal region is shown including the statistical uncertainty bands of fit parameters and systematics related to parton shower and fit function models. The original sideband and signal region fits for \( m_{\ell \nu j} \) are also shown on the same figure.](image)
uncertainty associated with the chosen fit function for the sideband and signal regions is also considered and is shown as "Alternate Function" in Fig. 4.

The W+jets shape in the signal region is then extrapolated from the lower sideband through the $\alpha_{MC}(m_{\ell\nu})$ function:

$$F_{\text{data,SR}}(m_{\ell\nu}) = \alpha_{MC}(m_{\ell\nu}) \times F_{\text{data,LSB}}(m_{\ell\nu})$$

(5)

Fig. 5 shows, on top, the $m_{\ell\nu}$ distribution in the LSB for data and MC, for muons on the left and electrons on the right. The simulated distributions are parametrized with functions determined with dedicated fits on each single simulated sample. The non-dominant backgrounds are fixed to MC expectations corrected by data-to-MC scale factors, while the W+jets parameters are left free to float and the uncertainty bands reflect the fit parameter errors. This latter component is then scaled according to Eq. 5 to get the W+jets shape in the signal region. In the bottom row of Fig. 5, the obtained distributions in the signal region are shown, for muons on the left and electrons on the right. The uncertainty band takes into account both the fit uncertainty on $m_{\ell\nu}$ spectrum in the LSB region, the uncertainty on alpha function as shown in Fig. 4 and normalization uncertainties for the non-dominant background contributions detailed more in Sec. 7.

The SM and BSM Higgs hypotheses are modeled using a Crystal Ball function in both gluon fusion and vector boson fusion production modes. The model is able to account for large interference effects between the signal and background which increase with the mass of the hypothesized Higgs resonance. The shapes for these signal models are also shown in Fig. 5.

7 Systematic uncertainties

In this section, we describe the systematic uncertainties contributing to the analysis. The uncertainty from the luminosity is taken to be 4.4% [68]. Uncertainties related to the SM Higgs signal cross-sections are taken from [69, 70] are at most 13% and are dependent on the signal hypothesis mass. Uncertainties on the Parton Distribution Functions (PDFs) and the QCD scale are also taken from [69, 70]. Uncertainties related to the cross-sections of non-dominant background contributions are related to scale factors described in Section. 6. Additional conservative uncertainties of 30% for single top and WW/WZ are included for the boosted topology to coincide with similar uncertainties assigned in [71]. The effect of standard jet energy correction systematic uncertainties [67] are evaluated on the jet mass. The effect on the normalization of the jet mass distribution is approximately 2% including the effect from the fit of the merged W peak in the top-enriched control sample. The normalization uncertainty for $t\bar{t}$ and signal top are taken from the top-enriched control region. We also account for the shape variations of the $t\bar{t}$ contribution varying the matching and renormalization scales and find the effect to be negligible. The pure W-tagging scale factor is taken from a simultaneous fit on the merged W same in the $t\bar{t}$ control region which has systematic uncertainty of 10%. The W+jets shape and normalization uncertainties comes from three contributions: the uncertainties on the fit parameters, differences in parton shower models hypotheses and the difference coming from alternative shape functions. The size of the shape uncertainties can be seen in Fig. 4. Table 1 shows a summary of the systematic uncertainties related to normalization of expected signal and background yields.
Figure 5: Top: distributions in the LSB for data and MC for muons (left) and electrons (right). Bottom: final \( m_{\ell\nu j} \) distributions in the signal region are shown for the muon (left) and electron (right) channels.
Table 1: Summary of systematic uncertainties related to normalization of expected signal and background yields. The symbol † denotes a mass-dependent uncertainty.

<table>
<thead>
<tr>
<th>Syst. uncertainty</th>
<th>4.4%</th>
<th>4.4%</th>
<th>4.4%</th>
<th>4.4%</th>
<th>4.4%</th>
<th>4.4%</th>
</tr>
</thead>
<tbody>
<tr>
<td>lumi</td>
<td></td>
<td></td>
<td></td>
<td>4.4%</td>
<td>4.4%</td>
<td>4.4%</td>
</tr>
<tr>
<td>Higgs QCD scale</td>
<td>6.5% †</td>
<td>1.3% †</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Higgs PDF+α_s</td>
<td>12.1% †</td>
<td>5.9% †</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intf (sig/bkg)</td>
<td>10.0%</td>
<td>50.0%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bkg cross-section</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>30.0%</td>
<td>30.0%</td>
</tr>
<tr>
<td>W+jets norm.</td>
<td></td>
<td>8%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>W-tagging</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>10.0%</td>
<td></td>
</tr>
<tr>
<td>tf norm.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jet mass/energy scale</td>
<td>2%</td>
<td>2%</td>
<td></td>
<td>2%</td>
<td>2%</td>
<td>2%</td>
</tr>
<tr>
<td>W+jets shape</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>see Sec. 6</td>
<td></td>
</tr>
<tr>
<td>b-tagging</td>
<td>2.5%</td>
<td>2.5%</td>
<td></td>
<td></td>
<td>2.5%</td>
<td>2.5%</td>
</tr>
<tr>
<td>Trigger (e &amp; µ)</td>
<td>1%</td>
<td>1%</td>
<td></td>
<td>1%</td>
<td>1%</td>
<td>1%</td>
</tr>
<tr>
<td>Selection Eff. (e &amp; µ)</td>
<td>2%</td>
<td>2%</td>
<td></td>
<td>2%</td>
<td>2%</td>
<td>2%</td>
</tr>
</tbody>
</table>

8 Results

No significant excesses have been observed in the $m_{ℓνj}$ spectrum, which can be seen in Fig. 5. Therefore, we set upper limits on the SM Higgs boson production cross section. We use the modified frequentist asymptotic CLs [72, 73] with profile likelihood as the test statistic in computing the exclusion limits. The limits are computed using an unbinned shape analysis where examples of the shapes can be seen in Fig. 5. Exclusion limits at a 95% confidence level are presented on the production cross section for the Higgs boson compared to the SM expectation in Fig. 6. An expected sensitivity to exclude the SM Higgs boson varies from 1.05 times the SM cross-section at 600 GeV to 4.6 times the SM cross-section at 1000 GeV. We are able to exclude at 1.1 (4.1) times the SM Higgs cross-section for a mass of 600 (1000) GeV.

Regarding the BSM electroweak singlet heavy Higgs, the left side of Fig. 7 shows the 95% exclusion limit on the cross-section times branching ratio into a WW pair, as a function of the mass of the signal hypothesis. Several values of the $C^2$ parameter, introduced in Eq. 2, are considered while the $BR_{\text{new}}$ is kept at zero. The expected theory cross-section is also shown. The right side of Fig. 7 shows the same limit, for a fixed signal mass of 600 GeV as a function of $BR_{\text{new}}$ for several values of the $C^2$. In both cases, expected limits are shown as dashed lines, while observed limits as continuous ones. We find that the typical upper limit on the $σ_{95\%} × BR_{WW}$ ranges from ~60-400 fb when $BR_{\text{new}} = 0$ where $C^2$ ranges from 0.3-1.0. As a function of $BR_{\text{new}}$, the upper limit is fairly consistent due to the competing effects of an increasing width and decreasing cross-section as $BR_{\text{new}}$ increases.

9 Conclusions

A search for a Higgs boson decaying into the WW final state is performed with an integrated luminosity of up to 19.3 fb$^{-1}$ of pp collisions at $\sqrt{s} = 8$ TeV in the high mass regime $600 < m_H < 1000$ GeV. The search is performed where one W boson decays leptonically and the other W decays hadronically. The hadronic W is highly boosted and the decay products are contained in one jet. Requirements on the pruned jet mass and N-subjettiness, $τ_2/τ_1$, are used to separate these jets from merged W bosons from QCD jets. Benchmark searches are performed for high mass SM Higgs bosons and 95% exclusion is found ranging from 1.1 (4.2) times the Standard Model expected cross-section for the 600 (1000) GeV mass hypothesis. The results are
Figure 6: The 95% CL limit on $\sigma/\sigma_{SM}$ for a Higgs boson decaying to $WW \rightarrow l\nu q\bar{q}'$.

Figure 7: On the left, BSM exclusion limits for a signal mass hypothesis of 600 GeV as a function of mass for various values of $C^2$ where $BR_{\text{new}} = 0$. On the right, BSM exclusion limits for a signal mass hypothesis of 600 GeV as a function of $BR_{\text{new}}$ for various values of $C^2$ where $m_H = 600$ GeV.
also interpreted in a beyond the Standard Model scenario with a heavy electroweak singlet mixing with the new Higgs-like boson at 125 GeV. Exclusion limits are presented ranging from approximately 60–400 fb on the cross-section times branching ratio to $WW$ where $C^2$ ranges from 0.3–1.0.

Acknowledgments

We wish to congratulate our colleagues in the CERN accelerator departments for the excellent performance of the LHC machine. We thank the technical and administrative staff at CERN and other CMS institutes, and acknowledge support from: FMSR (Austria); FNRS and FWO (Belgium); CNPq, CAPES, FAPERJ, and FAPESP (Brazil); MES (Bulgaria); CERN; CAS, MoST, and NSFC (China); COLCIENCIAS (Colombia); MSES (Croatia); RPF (Cyprus); Academy of Sciences and NICPB (Estonia); Academy of Finland, MEC, and HIP (Finland); CEA and CNRS/IN2P3 (France); BMBF, DFG, and HGF (Germany); GSRT (Greece); OTKA and NKTH (Hungary); DAE and DST (India); IPM (Iran); SFI (Ireland); INFN (Italy); NRF and WCU (Korea); LAS (Lithuania); CINVESTAV, CONACYT, SEP, and UASLP-FAI (Mexico); MSI (New Zealand); PAEC (Pakistan); SCSR (Poland); FCT (Portugal); JINR (Armenia, Belarus, Georgia, Ukraine, Uzbekistan); MST, MAE and RFBR (Russia); MSTD (Serbia); MICINN and CPAN (Spain); Swiss Funding Agencies (Switzerland); NSC (Taipei); TUBITAK and TAEK (Turkey); STFC (United Kingdom); DOE and NSF (USA). Individuals have received support from the Marie-Curie programme and the European Research Council (European Union); the Leventis Foundation; the A. P. Sloan Foundation; the Alexander von Humboldt Foundation; the Belgian Federal Science Policy Office; the Fonds pour la Formation à la Recherche dans l’Industrie et dans l’Agriculture (FRIA—Belgium); the Agentschap voor Innovatie door Wetenschap en Technologie (IWT-Belgium); and the Council of Science and Industrial Research, India.

References


