A combination of searches for the invisible decays of the Higgs boson using the CMS detector

The CMS Collaboration

Abstract

Searches for invisible decays of the Higgs boson using Run 1 proton-proton collision data from the CMS experiment at the LHC are combined. The searches cover a wide range of Higgs boson production modes namely gluon fusion, vector boson fusion and vector boson associated production. The data samples used were collected in 2011 and 2012 and correspond to integrated luminosities of up to 4.9 fb$^{-1}$ at 7 TeV and up to 19.7 fb$^{-1}$ at 8 TeV. Data, in all channels, are consistent with the expected standard model backgrounds. Assuming standard model Higgs boson cross sections and corresponding detector acceptances, limits are set on the invisible branching fraction for each production mode separately and for the full combination of all channels. From the combination of all channels, we set an observed (expected) upper limit on the invisible branching fraction for $m_H = 125$ GeV to be 36% (30%) at 95% confidence level.
1 Introduction

Invisible decays of the Higgs boson are predicted in several proposed extensions of the Standard Model (SM). For example, the Higgs boson can decay to neutralinos in supersymmetric models [1], or graviscalars in models with extra dimensions [2, 3]. In general, interactions of the Higgs boson with the unknown dark matter (DM) sector may introduce invisible decay modes, and bounds on these decays can constrain DM models [4–6].

In this note we describe the combination of the results of the CMS analyses searching for an invisibly decaying Higgs boson using the data from Run 1 of the LHC [7–9]. To identify an invisibly decaying Higgs boson it must be produced in association with other particles. The searches used cover the three associated production modes of the Higgs boson with the highest standard model (SM) cross sections. All of the searches use events with a large missing transverse momentum, defined as the negative vector sum of the reconstructed momenta of particles in the transverse plane, in association with one or more high energy, reconstructed objects.

The analysis with the best sensitivity targets the vector boson fusion (VBF) mode, where the Higgs boson is produced in association with two quarks, as shown in Fig. 1 (left). This analysis benefits from a large SM cross section, but also suffers from large backgrounds due to its two jets plus missing transverse energy ($E_T^{\text{miss}}$) final state.

There is also an analysis targeting the gluon fusion (ggH) production mode, as shown in Fig. 1 (center). This production mode has the highest SM cross section, however it normally results in the Higgs boson being created alone, and thus leaving no characteristic signature in the detector if it decays invisibly. Therefore, the only way to detect this production mode is to look for events with initial state radiation and $E_T^{\text{miss}}$. These “monojet” events provide an identifiable topology and their SM cross section is still approximately 10 times that of VBF, however, the signal acceptance after selection to remove background is small.

Finally there are several analyses with categories targeting the vector boson, V, (W or Z) associated production mode, VH, as shown in Fig. 1 (right). This production mode has a smaller SM cross section, but the presence of the V-boson provides a variety of identifiable final states with relatively low backgrounds. We consider the case where the V decays hadronically, referred to as V(had)H-tagged, and ZH production where the Z boson decays to electrons and muons, referred to as Z(ℓℓ)H(inv), or bb, referred to as Z(bb)H(inv).

Figure 1: Example Feynman diagrams for Higgs production in the VBF (left), ggH (center) and VH (right) channels. The Higgs boson is assumed to decay invisibly.
2 Analyses entering the combination

The analyses included for the combination are summarised in Table 1. An outline of their selection and signal extraction methods as well as any modifications made for this combination to the previously published analyses are given in the remainder of this section. The selections used for each analysis are designed to be mutually exclusive such that each category is statistically independent from the others.

Table 1: Summary of the analyses included in the combination. The first column is the name of the analysis. The second and third columns give the integrated luminosity of the 7 and 8 TeV data sets used by each analysis. The fourth column contains the names of the categories in each analysis and the fifth column gives the proportion of signal events expected to come from each Higgs boson production mode.

<table>
<thead>
<tr>
<th>Analysis</th>
<th>Luminosity (fb$^{-1}$)</th>
<th>Category</th>
<th>Expected signal composition</th>
</tr>
</thead>
<tbody>
<tr>
<td>VBF</td>
<td>19.2</td>
<td>2-jet VBF</td>
<td>92% VBF, 8% ggH</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Monojet</td>
<td>70% ggH, 20% VBF, 6% WH, 3% ZH</td>
</tr>
<tr>
<td>Monojet+V(had)H-tagged</td>
<td>19.7</td>
<td>V(had)H boosted</td>
<td>47% WH, 25% ggH, 23% ZH, 5% VBF</td>
</tr>
<tr>
<td></td>
<td></td>
<td>V(had)H resolved</td>
<td>39% ggH, 32% WH, 18% ZH, 11% VBF</td>
</tr>
<tr>
<td>Z(\ell\ell)H(inv)</td>
<td>19.7 (4.9)</td>
<td>$e^+e^-$ - 0-jet</td>
<td>100% ZH</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$e^+e^-$ - 1-jet</td>
<td>100% ZH</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\mu^+\mu^-$ - 0-jet</td>
<td>100% ZH</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\mu^+\mu^-$ - 1-jet</td>
<td>100% ZH</td>
</tr>
<tr>
<td>Z(b\overline{b})H(inv)</td>
<td>18.9</td>
<td>2-b-jet - low $E_T^{miss}$</td>
<td>100% ZH</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2-b-jet - medium $E_T^{miss}$</td>
<td>100% ZH</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2-b-jet - high $E_T^{miss}$</td>
<td>100% ZH</td>
</tr>
</tbody>
</table>

2.1 Vector Boson Fusion

The VBF mode benefits from a large signal to background ratio due to the presence of two reconstructed jets in the final state which typically have a large rapidity gap between them without any hadronic activity. The invariant mass of the dijet system is typically large in VBF events. We therefore require two jets with the VBF topology and large $E_T^{miss}$ which is well separated from any jets. The latter requirement acts to reduce backgrounds from QCD multijet events for which the $E_T^{miss}$ generated is typically due to a mis-measured jet. We veto events with electrons or muons present to remove background from V+jets processes [8].

After selection the main backgrounds are Z+jets, where the Z boson decays to neutrinos, and W+jets, where the W decays leptonically and the lepton is either outside the detector acceptance or not reconstructed. The contributions of these processes are estimated using lepton-enriched control regions defined by inverting the lepton vetoes. Multiplicative scale factors obtained from Monte Carlo (MC) simulation are then used to extrapolate from the control region to the signal region. The QCD multijet background is estimated using a control region with inverted jet-$E_T^{miss}$ separation requirements. Minor backgrounds from top, dibosons and
Drell-Yan are taken directly from MC simulation. These background estimations and the observed number of events in data are used to perform a counting experiment.

### 2.2 Monojet+V(had)H-tagged

The events selected by this analysis have one or more energetic jet and large $E_T^{\text{miss}}$. Events with well identified electrons, photons, muons or taus are rejected. The events are then separated into three categories based on the jet topology, referred to as “boosted”, “resolved” and “monojet”. The monojet category targets ggH production with initial state radiation and the other two categories target V(had)H production. The categorisation is sequential with events first being considered for the boosted category, followed by the resolved category, and finally the monojet category [9].

In order to remove any overlap with the VBF analysis events with two jets are rejected if the jets fall into the VBF signal region. This veto is the only difference between the analysis described here and that previously published in Ref [9]. When comparing the number of signal events expected from MC before and after applying the veto, we find that the veto removes approximately 4% of events from the monojet category, less than 1% of events from the boosted category and no events from the resolved category.

The boosted category targets events with a hadronically decaying vector boson, in which the two resulting jets are merged due to its large transverse momentum ($p_T$). The events in this category are required to contain a “fat jet”, which is obtained using the Cambridge/Aachen algorithm [10] with a distance parameter ($\Delta R = \sqrt{\Delta \phi^2 + \Delta \eta^2}$) of 0.8, that is tagged as likely to contain a hadronic V [11, 12]. Additional jets with $\Delta R < 0.5$ with respect to the fat jet are allowed to accommodate initial state radiation, while events with more than one additional central jet with $\Delta R$ to the fat jet larger than 0.5 are rejected from this category to remove tt background.

The resolved category targets events with a hadronically decaying vector boson whose $p_T$ is not large enough for it’s decay products to be contained in a single fat jet. The events are required to have two jets whose dijet mass is consistent with originating from a V-boson. They are also required to be identified by a multivariate V-tagger [13] as likely to contain a hadronically decaying vector boson. Events with jets tagged as being from b-hadron decays are rejected from this category.

The monojet category targets events with a single jet originating from initial or final state radiation. At least one jet with $p_T > 150$ GeV is required. As in the boosted category, a second jet with $\Delta R < 0.5$ with respect to the leading jet is allowed to accommodate for additional radiation, however, events with three or more central jets are rejected.

The dominant backgrounds in this analysis are Z+jets events where the Z boson decays to neutrinos and leptonically decaying W+jets events where the lepton is not reconstructed. The Z+jets background is constrained using two data control regions, one requiring two muons compatible with a Z boson decay and a second requiring a photon which supplements the limited number of events available in the dimuon region. The W+jets background is controlled using a single muon data control region. Further minor backgrounds from dibosons, top and QCD are taken from MC simulation. A simultaneous fit to the $E_T^{\text{miss}}$ distribution in the signal and background control regions is used to extract limits.
2.3 Z(ℓℓ)H

In this analysis we require the presence of two high $p_T$, isolated, opposite-sign, same-flavour leptons (either electrons or muons) compatible with the Z boson mass, large $E_T^{\text{miss}}$ and either one or no jets. To reduce the contribution from background, events containing b-tagged jets or additional leptons are vetoed and we require the dilepton system and the $E_T^{\text{miss}}$ to be opposite each other and balanced in magnitude in the transverse plane [7].

The selected events are split into four categories depending on the flavour of the leptons and the number of jets present. After the signal region selection the dominant backgrounds come from the diboson processes ZZ and WZ, which are modelled using MC simulation. Remaining smaller backgrounds from t̅t̅, WW and W+jets are estimated together using control regions with opposite-sign, opposite-flavour leptons. The background observed in these control regions is then related back to the signal region using the ratio between same-flavour and opposite-flavour events observed in the sidebands of the Z boson mass peak.

In the 8 TeV dataset a two-dimensional fit to the azimuthal angular separation between the two leptons and the transverse mass ($m_T$) of the dilepton-$E_T^{\text{miss}}$ system is carried out to extract the signal yield. For the 7 TeV dataset a one-dimensional fit to the $m_T$ is used, due to limited numbers of events in the control regions.

2.4 Z(bb)H

The event selection for this analysis requires significant $E_T^{\text{miss}}$, with a direction which is separated from jets and a pair of jets consistent with a Z boson decaying to b̅b. The selected events are separated into three categories according to the amount of $E_T^{\text{miss}}$ in the event, namely the “low” ($100 < E_T^{\text{miss}} < 130$ GeV), “intermediate” ($130 < E_T^{\text{miss}} < 170$) and “high” ($E_T^{\text{miss}} > 170$ GeV) $E_T^{\text{miss}}$ regions. In all categories events with leptons are vetoed, and in the low $E_T^{\text{miss}}$ category events with two or more additional jets are also removed [7].

The remaining backgrounds come from V+jets, t̅t̅, top, diboson and QCD multijet processes. MC simulation normalised to control regions in data is used to model all of these processes, and in all cases the normalisation factor obtained is close to one. The control regions are constructed by combinations of inverting and relaxing the selection requirements to provide a region enriched in each background.

The signal yield after this selection is estimated using a boosted decision tree (BDT) trained using signal and background events from MC simulation. The shape of the background events is taken from simulation and validated by comparing the distributions in simulation with the data in several, signal free control regions. The agreement is found to be well within the systematic uncertainties ascribed to the shape of these distributions.

3 Combination

None of the above analyses sees a significant excess. By assuming SM Higgs boson production cross sections and corresponding detector acceptances, the results of the individual analyses may be combined and interpreted as a limit on the invisible branching fraction of the Higgs boson, $B(H \rightarrow \text{inv})$. The limits are placed at 95% confidence level (CL) using an asymptotic CLs method [14–16], following the standard LHC Higgs combination procedure [17]. Systematic uncertainties are treated as nuisance parameters in a frequentist paradigm, as described in [17].
The statistical procedure takes full account of correlations between the nuisance parameters in each of the analyses. The uncertainties correlated are those associated with the signal uncertainty, due to PDF and renormalization/factorization scale variation uncertainties, the diboson cross section uncertainties, the total integrated luminosity uncertainty and the lepton momentum scale uncertainties and the $E_\text{miss}$ energy scale and resolution uncertainties. The $E_\text{miss}$ uncertainties are uncorrelated between the monojet+$V$(had)H-tagged analysis and the other analyses due to differences in the algorithm used to reconstruct the $E_\text{miss}$ and its calibration. Where simulation is used to model the $E_\text{miss}$, we propagate uncertainties on the unclustered energy, the jet energy corrections and the lepton scales to the $E_\text{miss}$ treating the lepton and jet energy correction uncertainties correlated with the other respective jet energy correction and lepton energy scale nuisances. For the mono-$V$(had) categories, an additional correction to match the recoil is applied, the uncertainties on this correction are propagated to the signal model and validated on $\gamma+$jet events.

The jet energy uncertainties in the $Z$(b$b$)H(inv)analysis are estimated using a different technique from that used in the other analyses, so are not correlated [18]. The jet kinematics in the monojet+$V$(had)H-tagged analysis are also distinct from those in the other analyses, and therefore, the jet energy uncertainties are correlated only between the $Z$(l$l$)H(inv)and VBF analyses.

The 95% CL upper limits on $\mathcal{B}(H \rightarrow \text{inv})$ obtained from the searches targeting each production mode individually and from the combination of all searches are given in Table 2. A scan of the profile likelihood versus the invisible branching fraction is also shown in Fig 2. The profile likelihood and CLs methods are not expected to give exactly the same limits. In addition to the individual analyses and the full combination, for each production mode we have performed a combination of all the categories targeting that mode. The results of these combinations are shown in Fig 3.

<table>
<thead>
<tr>
<th>Channel</th>
<th>Observed (expected) upper limits on $\mathcal{B}(H \rightarrow \text{inv})$ (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>VBF-tagged</td>
<td>57 (40)</td>
</tr>
<tr>
<td>VH-tagged</td>
<td>60 (69)</td>
</tr>
<tr>
<td>ggH-tagged</td>
<td>67 (71)</td>
</tr>
<tr>
<td>Combined</td>
<td>36 (30)</td>
</tr>
</tbody>
</table>

### 4 Summary

The CMS searches for invisible decays of Higgs bosons using data from Run 1 of the LHC have been combined. The searches target the VBF, VH and ggH production modes. The observation in data shows no significant difference compared to the expectation from SM backgrounds in all channels. Using the CLs method and assuming SM production cross sections and corresponding detector acceptances we set upper limits on the invisible branching fraction of the Higgs boson for each channel separately, for the combination of all channels contributing to each Higgs boson production mode and for the combination of all channels. The observed (expected) upper limit set on the Higgs boson invisible branching fraction for $m_H = 125$ GeV is 36 (30)%, at 95% confidence level.
Summary

Figure 2: Log-likelihood versus $B(H \to \text{inv})$. The solid curves represent the observation in data and the dashed curves represent the median expected result for no invisible decays of the Higgs boson.

Figure 3: Expected and observed 95% CL upper limits on production cross section times $B(H \to \text{inv})$ normalised to the SM Higgs boson production cross section obtained from the combination of all channels targeting each Higgs boson production mode.
References


