PROCEEDINGS OF SCIENCE

Search for the Standard Model Higgs boson decaying into a WW pair

PoS(LHCPP2013)004

M. Biglietti*
INFN Roma Tre
E-mail: michela.biglietti@cern.ch

P. Govoni
Universita' degli Studi Milano Bicocca
E-mail: pietro.govoni@gmail.com

Results of the Standard Model Higgs boson searches at CMS and ATLAS are presented, in the final state characterised by the presence of two W bosons in the detector. The analyses are performed on the full dataset collected by the LHC at 7 TeV and 8 TeV of centre-of-mass energy and span different Higgs production mechanisms as well as different decay modes of the W bosons. A resonance at low mass is observed with a significance of around 4 σ by both the experiments, compatible with the Standard Model expectations for the Higgs boson behaviour. First measurements of its properties are reported.

VI Italian workshop on p-p physics at the LHC,
8-10 May 2013
Acquario di Genova, Ponte Spinola, Area Porto Antico, Genova, Italy

*Speaker.

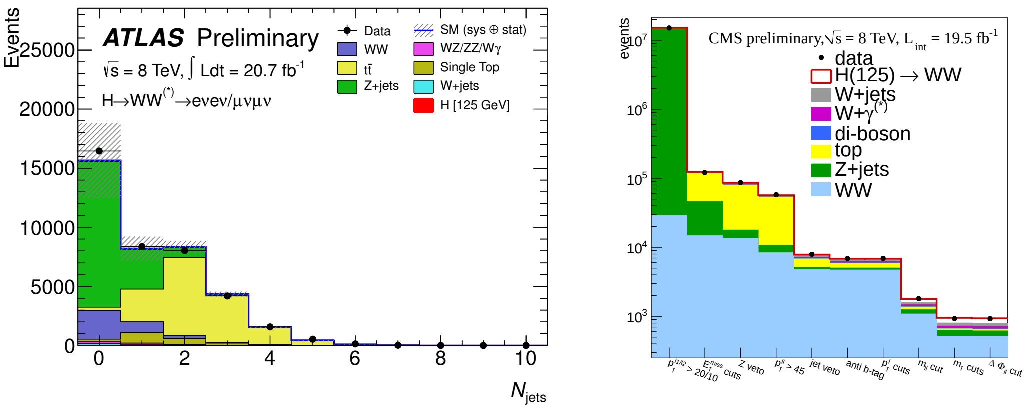

Figure 1: Left: expected rate in jets multiplicity for the fully leptonic ATLAS analysis, after the preselections. Right: events rate after selections, in the CMS analysis, for the 0 jet bin, at 8 TeV centre-of-mass energy. From [1, 2].

Thanks to the large branching ratio, the Standard Model Higgs boson decay into a pair of W bosons has been extensively studied at the LHC both by the CMS and ATLAS collaborations. In fact, for hypothetical masses below 300 GeV, the final state with both Ws decaying into electrons or muons, accompanied by the corresponding neutrinos, is one of the most intense final states in the range of interest. On the other hand, above 300 GeV the favourite Higgs boson decay with at least one detectable lepton in the final state is the semi-leptonic WW one, where one of the W bosons generates a lepton-neutrino pair, while the second one decays into quarks that originate jets in the detector.
Besides these two high intensity cases, the final state with two Ws can also be detected by the LHC detectors when produced via the sub-leading mechanisms of the vector boson fusion and of the Higgs strahlung. The measurement of these final states would confirm the Standard-Model nature of the Higgs production couplings by directly accessing various production diagrams.

1. The fully leptonic final state

Both ATLAS and CMS studied in detail the fully leptonic final state [1, 2], by selecting events containing two high-p_T, isolated electrons or muons of opposite sign accompanied by the presence of genuine missing energy. Since the presence of two neutrinos prevents an accurate reconstruction of the Higgs boson invariant mass, the analyses strongly rely on data-driven techniques to estimate the background contaminations in the signal regions. The analyses have been performed separately depending on the number of jets above a given threshold (tipically 25-30 GeV) present in the events. While in the bin with zero jets the main background is due to the non-resonant WW production, in the bin with one jet events with top quarks are important as well, and become dominant in the bins with more than one jet. Events containing b-jets are therefore rejected to minimise this contribution. Figure 1 shows on the lefthand side the expected rate in jets multiplicity, for the ATLAS analysis, after the preselections, while reports on the righthand side the evolution of the event rates after the various selections put in place by CMS, at 8 TeV, for the 0-jet bin. The WW

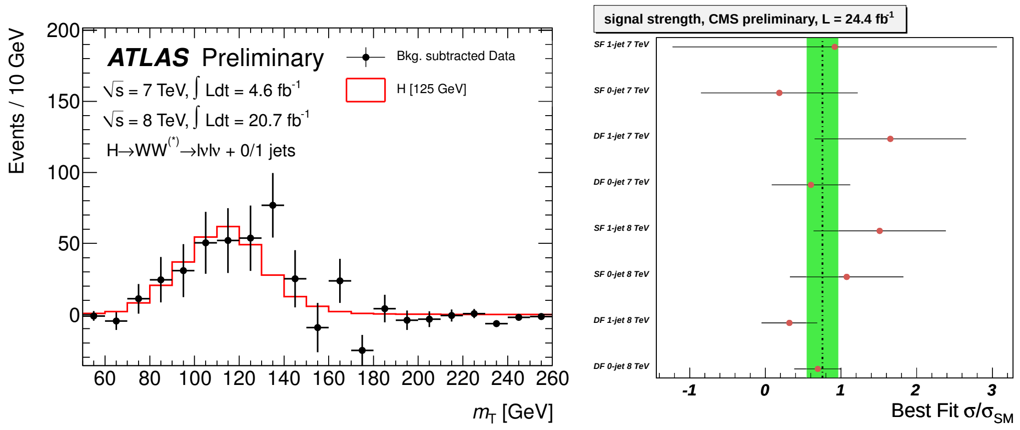

Figure 2: Left: the observed excess after subtracting the expected background, for the simulation (red line) and data (black dots) found by ATLAS. Right: the best fit value of the signal strengh (μ) for the CMS analysis, for the considered sub-channels. From [1, 2].

background is estimated through a Monte Carlo simulation, normalised in signal free regions, while the top background is evaluated by measuring the efficiency of the b-jet veto in data, and using this numbers to extrapolate the contamination from a pure top-enriched region.

Sub-leading backgrounds are due to the production of a Z boson, of a W boson with jets where one jet is wrongly reconstructed as a lepton, the production of diboson pairs other than WW, such WZ, ZZ, $W\gamma^{(*)}$. To cope with the drell-yan contamination (also due to the decay into τ pairs), specific missing-energy variables are used, while the determination of the W+jets background is performed by measuring in side-bands the probability for a jet to be detected as a lepton, and extrapolating the contamination from a W+jets dominated region.

The two-jets case is addressed after vector-boson-fusion selections, which require the presence of two jets with large separation in pseudo-rapidity and large invariant mass.

The signal extraction is performed by the ATLAS experiment with a fit on the spectrum of the transverse mass of the leptons plus missing trasverse energy system, in two regions of the leptons pair invariant mass, while the CMS analysis features a two-dimensional fit in the plane of the two variables. Figure 2 shows, on the lefthand side, the observed excess after subtracting the expected background, for the simulation (red line) and data (black dots) found by ATLAS.

The presence of the low mass resonance, observed also in the $\gamma\gamma$ and ZZ final states, has been measured, with an observed (expected) significance of 4.0 σ (5.1 σ) by CMS and 3.8 σ (3.5 σ) by ATLAS, over the full dataset acquired at 7 and 8 TeV centre-of-mass energy.

The intensity of the signal with respect to the Standard Model expectations has been measured to be:

$$\begin{aligned} \mu_{\text{ATLAS}} &= 1.01 \pm 0.22\,(\text{stat}) \pm 0.19\,(\text{theo syst}) \pm 0.10\,(\text{exp syst}) \pm 0.04\,(\text{lumi})\,, \\ \mu_{\text{CMS}} &= 0.76 \pm 0.13\,(\text{stat}) \pm 0.16\,(\text{syst})\,. \end{aligned}$$

In both the experiments, consistent results are observed in the various sub-channels in which the studies have been subdivided, as shown in figure 2, on the righthand side, for CMS.

The difference in expected significance between the two experiments can be explained by the different signal extraction techniques, by the different threshold applied to the two leptons (20, 10 GeV in case of CMS, 25, 15 GeV in case of ATLAS) and by the different systematics related to the background extraction techniques.

2. The spin characterisation of the Higgs boson

The spin of the new particle will manifest itself in the angular and momentum distributions of the decay products. The decay $H \to \mu\nu e\nu$ provides a relatively large signal yield, which allows for a shape analysis of kinematic distributions despite the poor invariant mass resolution. Other final states are not expected to add much in terms of spin sensitivity due to the presence of large backgrounds which cannot be removed without greatly reducing the acceptance of the spin model considered. Two hypotheses are tested against the observed excess of events in the search, namely a SM Higgs boson with $J^P = 0^+$ and a generic $J^P = 2^+_m$ graviton-like tensor with minimal couplings [3].

ATLAS used [4] two of the most sensitive variables for measuring spin, namely the dilepton invariant mass and the azimuthal separation of the two leptons. In addition to these discriminating variables, 0^+ and 2^+_m hypotheses present different missing transverse energy spectra and different distributions of the dilepton transverse momentum. This motivated a looser event selection compared to the one applied in the rate measurement. The spin hypotheses are tested using a two-dimensional kinematic shape fit. The discriminants used in the fit are outputs of two different boosted decision trees (BDT), trained separately against all backgrounds to identify 0^+ and 2^+_m events, respectively. Two different production processes, $q\bar{q} \to H$ and $gg \to H$, are included for the 2^+_m model and the analysis is performed for different values of the $q\bar{q} \to H$ fraction, f_{qq}.

CMS [1] constructs a signal plus background model for each hypothesis, based on two-dimensional templates in the transverse mass and di-lepton invariant mass, as in the SM Higgs rate analysis described in Section 1. The 2^+_m template includes only the $gg \to H$ production process.

The statistical analysis of the data employs a binned likelihood $\mathscr{L}(\varepsilon, \theta)$ constructed with one parameter of interest ε, which represents the fraction of spin 0^+ signal events in the total signal expectation, such that $\varepsilon = 0$ represents the spin 2^+_m hypothesis. Systematic uncertainties are represented through the nuisance parameters θ. The compatibility between data and the two hypotheses is then estimated using the following test statistic:

$$q = \log \frac{\mathscr{L}(H_{0^+})}{\mathscr{L}(H_{2^+})} = \log \frac{\mathscr{L}(\varepsilon = 1, \hat{\hat{\mu}}_{\varepsilon=1}, \hat{\hat{\theta}}_{\varepsilon=1})}{\mathscr{L}(\varepsilon = 0, \hat{\hat{\mu}}_{\varepsilon=0}, \hat{\hat{\theta}}_{\varepsilon=0})} \,. \tag{2.1}$$

In both the numerator and denominator, the likelihood is maximised over all nuisance parameters to obtain the maximum likelihood estimators $\hat{\hat{\mu}}, \hat{\hat{\theta}}$. Results show that for small f_{qq} the shape differences between kinematic distributions for spin 0^+ and spin 2^+_m particles become smaller and the sensitivity is expected to decrease. A distribution of q, derived from pseudo-experiments, and a summary of the the expected and observed median test statistics for different f_{qq} hypotheses, are shown in Figure 3. For all values of f_{qq}, the spin 0^+ hypothesis is preferred over the spin 2^+_m hypothesis to a confidence level higher than 95% confidence level.

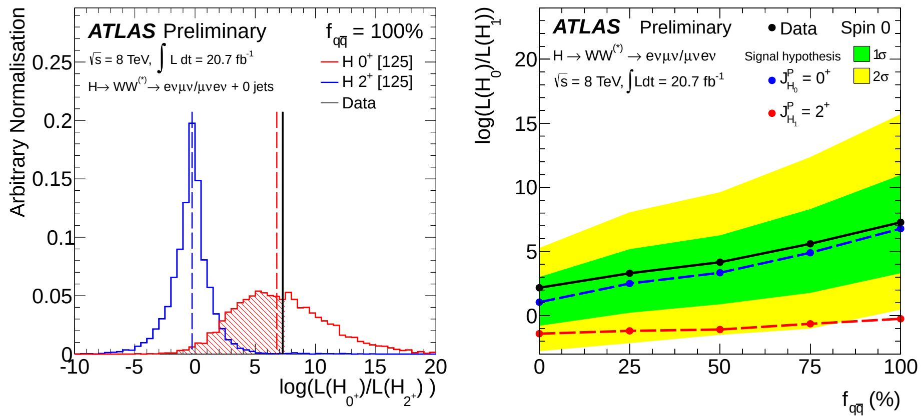

Figure 3: Left: Test statistics distributions for 0^+ (red) and spin 2^+_m (blue) as well as the observed value (solid line) for the $f_{qq} = 1$ working point. Right : The median test statistic for spin 0^+ (blue dashed line) and spin 2^+_m (red dashed line) as well as the observed value (solid black line), for various assumptions of the fraction of $q\bar{q}$ production, f_{qq}. The one- and two-sigma uncertainty bands for spin 0^+ are also shown by the green and yellow regions, respectively. From [4].

3. The low statistics final states

The Vector Boson Fusion (VBF) and Higgs-Strahlung (WH/ZH) processes are two of the four highest cross section production mechanisms of Higgs bosons at the Large Hadron Collider (LHC), and they are important channels for the completeness of the Higgs boson physics program.

For the VBF production, ATLAS performed the analysis on the full dataset acquired at 7 and 8 TeV centre-of-mass energy; the sensitivity to this production mode is obtained by adding candidates with more than one selected jet to the fully leptonic selection. The CMS analysis is performed on about 17 fb^{-1} of integrated luminosity. The VBF-specific selections use the kinematics of the two highest transverse momentum jets in the event. Their rapidity gap and their invariant mass is required to be large. Activity in the rapidity gap between the tag jets is restricted, to reduce the top and gluon-gluon fusion (ggF) contribution to this mode: events with additional jets with trasverse momentum greater then 20-30 GeV inside the rapidity gap are vetoed. The leptons are required to be within the rapidity gap. After the selection the background composition is dominated by top events.

For the ATLAS result, statistical tests of a VBF signal are performed by considering the ggF signal as part of the background. The test defines $\mu_{\rm VBF}$, the signal strength parameter associated with the VBF process, as the parameter of interest. The ggF signal strength $\mu_{\rm ggF}$ is profiled, and is constrained mainly by the $Njet$ <2 signal regions. The observed (expected) significance for the VBF analysis, for the Higgs boson mass of 125 GeV, is 2.5 (1.6) standard deviations. The best-fit measured signal strength at m_H= 125 GeV is

$$\mu_{\rm VBF} = 1.66 \pm 0.67\,(\text{stat.}) \pm 0.42\,(\text{syst.}) = 1.66 \pm 0.79$$

Similarly, $\mu_{\rm ggF}$ has been measured by considering the VBF signal as part of the background.

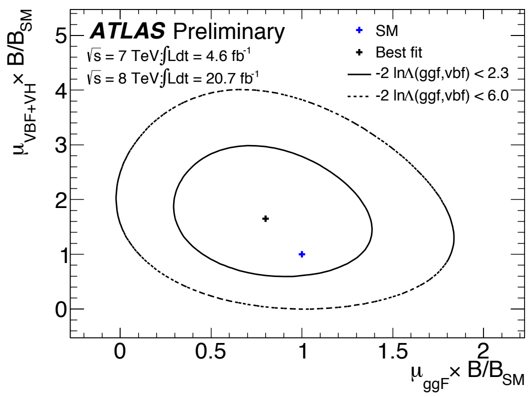

Figure 4: Likelihood contours for separate ggF and VBF signal strength parameters (a) and the likelihood curves for the ratio of the ggF/VBF strength parameters (b). $H \to WW^{(*)} \to \ell\nu\ell\nu$ analysis uses the combined 7 and 8 TeV data. From [2]

The best-fit signal strength at m_H= 125 GeV is

$$\mu_{\rm ggF}{=}0.82\pm0.24\,(\text{stat.})\pm0.28\,(\text{syst.}){=}0.82\pm0.36$$

A two-dimensional likelihood scan of the signal strength for the ggF and VBF production modes is shown in Figure 4. Since the signal strengths in the VBF, WH, and ZH production modes scale with the VH coupling, they are grouped together. The results are consistent with the expected SM values of unity.

The WH production has been studied by ATLAS and CMS in the three-lepton final state. In particular CMS studied this channel using the full dataset acquired at 7 and 8 TeV centre-of-mass energy [7]. This search selects triboson candidates in which all bosons decay leptonically, yielding an experimental signature of three isolated, high transverse momentum leptons (electrons or muons) and large missing transverse energy due to the undetected neutrinos. To further improve the sensitivity, events are split into two categories: events that contain an opposite-sign same-flavor lepton pair (OSSF) and events with same-sign same-flavor leptons (SSSF). While one fourth of the events are selected in the second category, the expected background is rather small since physics processes leading to this final state have small cross section. The main backgrounds are the the W(Z/γ^*) and top events. To reduce the top contribution, events are required to contain high p_T jets, that are not tagged as originating from b quarks. The contamination due to events where one identified lepton is originating from a jet is estimated from a control sample, where one of the two leptons is required to be loosely identified, which is assumed to be dominated by events containing a real prompt lepton only. The extrapolation to the signal region is performed through the probability of loose leptons to be identified also by the full analysis, measured in a multi-jet sample. The the W(Z/γ^*) background is reduced by requiring that all the OSSF lepton pairs have a dilepton mass away from the mass of the Z boson. The normalization of the remaining events is estimated using a three-lepton control region with one of the same-flavour opposite-charge lepton pairs mass compatible with a Z boson mass.

The results are extracted using a shape-based analysis that uses the smallest distance between opposite-charge leptons, $\Delta R_{l^+l^-}$. No significant excess of events is observed with respect to the

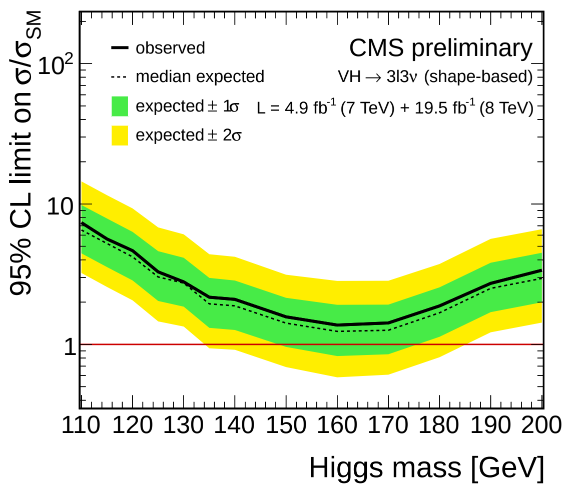

Figure 5: Upper limits at 95% CL, in the SM Higgs scenario, for the WH→3l analysis. From [7].

background prediction, and 95% CL upper limits are determined for the Higgs boson cross section with respect to the SM Higgs boson expectation, σ/σ_{SM}, using the modified frequentist construction CLs. The expected and observed upper limits are shown in Figure 5.

4. The semi-leptonic final state

The semi-leptonic decay of the WW system is exploited by both experiments to exclude the Standard Model Higgs boson at high masses (between 170 GeV and 600 GeV), based on the requirement of an isolated and well identified lepton (electron or muon) with large transverse momentum, together with the presence of two jets and missing energy in the events. In this regime both W bosons are expected to be on-shell, therefore the invariant mass of the decay system can be reconstructed by imposing the W mass constaint to the system composed by the lepton and the missing energy. ATLAS searched [5] for a high mass resonance (above 300 GeV) above the dominant W+jets background considering the gluon fusion and vector boson fusion productions separately, over the first 4.7 fb^{-1} collected. Only events with the dijet invariant mass compatible to the mass of the W boson were selected. The limit extraction was based on a fit to the invariant mass spectrum, with a parametric functional assumption for the background and the signal described by the simulation. The best sensitivity was reached for a 400 GeV Higgs hypothesis, at a value of 1.6 times the Standard Model Higgs production cross-section.
The CMS analysis has been performed over about 17 fb^{-1} of collected data [6], both during the 7 TeV and 8 TeV data taking periods. Besides the selections applied in the ATLAS analysis, the signal in enhanced with respect to the background by exploting a multi-variate discriminant built on the angular distributions of the Higgs decay products. Furthermore, the W+jets background has been evaluated by means of an extrapolation from the regions away from the W mass resonance in the dijet mass. Figure 6 shows the obtained exclusion limit as a function of the Higgs mass hypothesis. With an expected exclusion range of 220-560 GeV, the analysis is able and actually excludes at 95% confidence level the two intervals 225-485 GeV and 550-600 GeV.

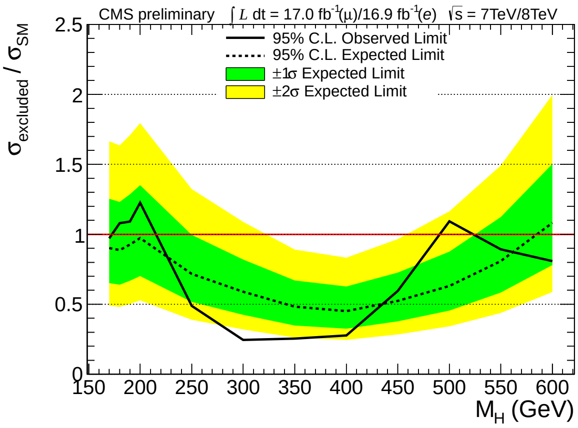

Figure 6: Obtained exclusion limit as a function of the Higgs mass hypothesis, for the CMS analysis, in the semi-leptonic WW final state. From [6].

5. Summary and Outlook

The LHC analyses of the Higgs boson decaying into two W bosons contribute significantly to the Higgs discovery and property measurements. The di-leptons final state has been analized with the full available statistics by both ATLAS and CMS. An observation of a new boson compatible with SM Higgs with a mass around 125 GeV has been reported with a significance of about 4σ. The couplings of the Higgs with the electroweak bosons W/Z have been probed through the analysis of the VBF channel standalone. The di-lepton channel is also sensitive to the spin of the new boson; the analyses indicated a preference towards the spin-0 hypothesis with respect to the spin-2 one, in agreement with the SM expectations. In the high mass regime, the semi-leptonic decay of the WW system is exploited to exclude the mass intervals 225-485 GeV and 550-600 GeV. The available statistics in the WW decay channel is not yet sufficient to conclude on the observation of the associated production of the Higgs boson with vector bosons.
The analyses will be updated and completed using the full dataset available at 7 and 8 TeV. In particular, improvements are expected in the selection and in the knowledge of the systematic uncertainties on the background components. Further improvements on the sensitivity can be obtained also by exploiting multivariate techinques.

After the Long Shutdown, from 2015 onwards, LHC will collect more than 100 fb^{-1} at $\sqrt{s}$=13-14 TeV. The new conditions will require improvements in many areas: for example the trigger must be prepared to run with with increased rates keeping reasonable acceptance for interesting physic channels and key quantities, like the missing energy and the jet reconstruction, need to be kept under control under the higher pileup. Besides this challenging environment, the new data will permit to probe the Standard Model predictions with the vector boson fusion and associated production channels, to test the VV scattering, and to search BSM for addtional resonances decaying into two Ws in the high mass regime. In fact, sub-leading production channels will be probed with enough statistics to become significant by themselves in the Standard Model framework. The quest for a significant discrimination between the picture where the already-discovered resonance fully accounts for the WW scattering spectrum regularisation, and the one where part of it has to be accounted for by another mechanism, is challenging. Nevertheless, only preliminary studies exist on the topic and a full assessment of ATLAS and CMS potential on this has still to be performed in light of the actual performances obtained during the data taking. Eventually, the higher centre-of-mass energy will widen the available phase space to search for new resonances in high mass tails. In this context, it's worth remembering that the presence of the two neutrinos in the fully leptonic final state determines a degradation in the capability of reconstructing the final state invariant mass, rather than preventing it.

References

[1] CMS Collaboration, CMS-PAS-HIG-13-003.

[2] ATLAS Collaboration, ATLAS-CONF-2013-30.

[3] Y. Gao, A. V. Gritsan, Z. Guo, K. Melnikov, M. Schulze and N. V. Tran, Phys. Rev. D **81** (2010) 075022 [arXiv:1001.3396 [hep-ph]].

[4] ATLAS Collaboration, ATLAS-CONF-2013-031.

[5] ATLAS Collaboration, ATLAS-CONF-2012-018

[6] S. Chatrchyan *et al.* [CMS Collaboration], [arXiv:1304.0213 [hep-ex]].

[7] CMS Collaboration, CMS-PAS-HIG-13-009.