IL NUOVO CIMENTO Vol. 37 C, N. 1 Gennaio-Febbraio 2014
DOI 10.1393/ncc/i2014-11696-3

Colloquia: IFAE 2013

# CMS results of search for Higgs decays in two fermions

M. De Gruttola on behalf of the CMS Collaboration
*University of Florida, Institute for High Energy Physics and Astrophysics*
*Department of Physics - Gainesville, Florida, USA*

ricevuto l'1 Ottobre 2013

**Summary.** — Results are presented on the study of the Higgs-like particle at a mass of 125 GeV decaying into final states consisting of either two taus, or a $b$ anti-$b$ quark pair. These are based on the full statistics of about $25\,\text{fb}^{-1}$, collected in 2011 and 2012 at an energy of 7 and 8 TeV, respectively, with the CMS experiment. Leptonic and hadronic decay channels for the tau leptons are included in the search. Different production channels namely inclusive, VBF and associated production have been studied. The $b\bar{b}$ decay channel is studied in associated production with W/Z bosons. The most recent results will be presented which confirm that the new boson couples to fermions with signal strengh close to prediction for a Standard Model-like Higgs boson. The combination of the latest CMS $H \rightarrow \tau^+\tau^-$ and $H \rightarrow b\bar{b}$ results leads to an observation of the coupling of the new boson to fermions with 3.4(3.5) observed (expected) significance.

PACS `14.70.Fm` – $W$ bosons.
PACS `14.70.Hp` – $Z$ bosons.
PACS `14.65.Ha` – Top quarks.

## 1. – Introduction

The CMS and ATLAS collaborations have reported the discovery of a new boson [1,2], with a mass, $m_H$, near 125 GeV and with properties compatible with those of the standard model Higgs boson [3]. To date, significant signals have been observed in channels where the boson decays into gauge boson pairs. The interaction with the fermions and whether the Higgs field serves as source of mass generation in the fermion sector through a Yukawa interaction between the Higgs fields and each fermion has not been firmly established. The recent results of CMS experiment based on the full 2011 and 2012 statistics of collected data is able indeed to assess if there is any indication that the new boson couple to fermions. The content of this paper is divided in four main sections: after a brief introduction to the CMS reconstruction, we proceed in the description of the $H \rightarrow \tau^+\tau^-$ and $H \rightarrow b\bar{b}$ analysis. Finally in the summary section we will state the combination of the results and the implications for Higgs physics sector.

## 2. – The CMS reconstruction

A detailed description of the CMS experiment can be found elsewhere [4].

Particle-flow (PF) algorithm [5] combines the information from all CMS subdetectors to identify and reconstruct the individual particles emerging from all vertices: charged hadrons, neutral hadrons, photons, muons, and electrons. These particles are then used to reconstruct the missing transverse energy ($\not\!\!E_T$), jets, leptons, hadronic $\tau$ decays, $b$-jets and to quantify the isolation of leptons and photons. In addition it allows the identification of the vertex corresponding to the hard-scattering process. Jets are reconstructed from all the PF particles using the anti-$kT$ jet algorithm [6], with a distance parameter of $R = 0.5$. Jets originating from $b$-quark hadronization are identified using the combined secondary-vertex $b$-tagging algorithm [7]. This algorithm combines in an efficient way the information about track impact parameters and secondary vertices within jets in a likelihood discriminant to provide separation of $b$-jets from jets originating from light quarks, gluons and charm quarks. PF $\not\!\!E_T$ is reconstructed from the vectorial sum of the transverse momenta of all particle candidates. Hadronically decaying taus are reconstructed and identified using an algorithm [8] which targets the main decay modes by selecting PF candidates with one charged hadron and up to two neutral pions, or with three charged hadrons.

The SM Higgs boson events are generated with POWHEG [9] interfaced to PYTHIA [10]. Background processes are generated with MADGRAPH [11]. The generated events are processed through a detailed simulation of the CMS detector based on GEANT4 [12] and are reconstructed with the same algorithms as the ones used for data.

## 3. – $H \to \tau^+\tau^-$ search at CMS

Five independent $\tau$-pair final states are studied: $\mu\tau_h$, $e\tau_h$, $e\mu$, $\tau_h\tau_h$, and $\mu\mu$, where $\tau_h$ denotes a reconstructed hadronic $\tau$ decay. In each channel, the signal is separated from the background, and in particular from the irreducible $Z \to \tau^+\tau^-$, using the $\tau$-pair mass $m_{\tau\tau}$ reconstructed from the four-momenta of the visible decay products of the two $\tau$ leptons and from the missing transverse energy $\not\!\!E_T$.

**3**˙1. *Selection*. – The high-level trigger (HLT) requires a combination of electron, muon, and $\tau$ trigger objects. In the $e\tau$ and $\mu\tau$ channels, the offline selection starts with events containing an electron of $p_T > 20\ (24)\,\text{GeV}$ or a muon of $p_T > 17\,(20)\,\text{GeV}$ for the 2011 (2012) dataset, located within $|\eta| < 2.1$, and accompanied by an oppositely charged $\tau$ of $p_T > 20\,\text{GeV}$ within $|\eta| < 2.3$. In the $e\mu$ channel, we select events with an electron satisfying $|\eta| < 2.3$ and an oppositely charged muon within $|\eta| < 2.1$, requiring $p_T > 20\,\text{GeV}$ for the leading lepton and $p_T > 10\,\text{GeV}$ for the sub-leading lepton. To be considered in the offline inclusive event selection, electrons and muons must have a relative isolation better than 10%. In the $\tau_h\tau_h$ channel, both $\tau_h$ are required to have $p_T > 45\,\text{GeV}$ and $|\eta| < 2.1$.

**3**˙2. *Event categoritation*. – The full set of selected events is split into mutually exclusive categories based on the jet multiplicity (considered up to $|\eta| < 4.7$), and on the transverse momentum of the reconstructed $\tau$-decay product.

- VBF: In this category, two jets with $p_T > 30\,\text{GeV}$ are required to tag the vector-boson fusion Higgs-production process. The two jets must have an invariant mass $M_{jj} > 500\,\text{GeV}$ and be separated in pseudorapidity by $\Delta\eta > 3.5$. A rapidity gap is defined by requiring no additional jet with $p_T > 30\,\mathrm{GeV}$ between the two tagging jets.

- 1-jet: Events in this category are required to have at least one jet with $p_T > 30\,\mathrm{GeV}$. This category exploits the production of a high-$p_T$ Higgs boson recoiling against a jet.
- 0-jet: This category contains all events with no jet with $p_T > 30\,\mathrm{GeV}$ and is only used to constrain background normalization, identification efficiencies, and energy scales.

The 0- and 1-jet categories are each further divided into two subsets, using the $p_T$ of the visible $\tau$-decay products, either hadronic or leptonic ($>$ 40 and 50 GeV respectively). We label these subsets low-$p_T$ and high-$p_T$. The SVFit algorithm reconstructs the $\tau$-pair invariant mass $m_{\tau\tau}$ with improved resolution using six parameters: the polar and azimuthal angles of the visible decay product system in the $\tau$ rest frame, the three boost parameters from the $\tau$ rest frame to the laboratory frame, and the invariant mass of the visible decay products.

**3**˙3. *backgrounds and systematics.* – The largest source of background is the Drell-Yan production of $Z \to \tau^+\tau^-$. This contribution is greatly reduced by the 1-jet and VBF selection criteria, and is modelled using embedded event samples, using a loose $Z \to \mu\mu$ selection, in which the reconstructed PF muons are replaced by the PF particles reconstructed from the $\tau$ visible decay products in simulated $Z \to \tau^+\tau^-$ events.

The Drell-Yan production of $Z \to ll$ contributing to the background in the $e\tau_h$ and $\mu\tau_h$ channels is estimated from simulation, after rescaling observed $Z \to \mu\mu$ events.

$W +$ jets production contributes significantly to the $e\tau$ and $\mu\tau$ channels when the $W$ decays leptonically and one jet is misidentified as a $\tau_h$. The background is modelled for these channels using the simulation, and than normalized to the yield observed in data in a high-$m_T$ control region.

The $t\bar{t}$ production process is the main remaining background in the $e\mu$ channel. The predicted yield for all channels is obtained from simulation, with the yield rescaled to the one observed using a $t\bar{t}$-enriched control sample.

QCD multijet events, in which one jet is misidentified as a $\tau$ and another as a lepton, constitute another important source of background in the $e\tau_h$ and $\mu\tau_h$ channels. The background estimation is entirely based on observed data using same charge lepton data.

The main experimental systematic uncertainties affecting the expected signal yield are from the $\tau_h$ identification efficiency (8%), the $\not\!\!E_T$ scale (5%), the integrated luminosity (2.2% in 2011 and 4.2% in 2012), and the jet energy scale (between 2.5 and 5%, depending on the jet $p_T$ and $\eta$). For the theoretical uncertainty the considered scale uncertainty in the VBF production yield is 4%, while the scale uncertainty in the gluon-gluon fusion production yield is 10% in the 1-jet/high-$p_T$ category and 30% in the VBF category.

**3**˙3.1. Results. The signal extraction is done using a profile-likelihood ratio test statistic to search for the presence of a SM Higgs boson signal. The $m_{\tau\tau}$ distributions obtained for each category of the five channels at 7 TeV and 8 TeV are combined in a binned likelihood. Figure 1 left shows the combined observed and expected $m_{\tau\tau}$ distributions, weighting all distributions in each category of each channel by the ratio between the expected signal and background yields for this category in a $m_{\tau\tau}$ interval containing 68% of the signal. An excess compatible with the presence of a 125 SM Higgs boson is

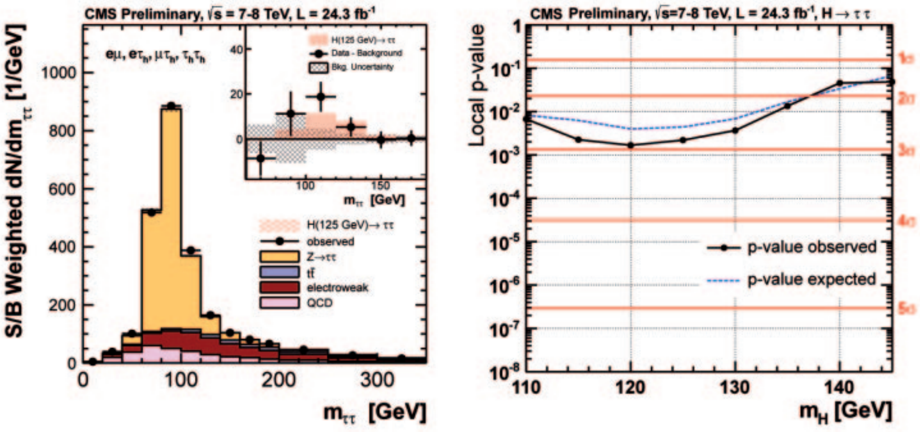

Fig. 1. – Left: combined observed and expected $m_{\tau\tau}$ distributions for the $\mu\tau_h$, $e\tau_h$, $e\mu$ and $\tau_h\tau_h$ channels. Right: observed and expected $p$-value 1-CL$_b$, and the corresponding significance in number of standard deviations.

observed. This excess is quantified in the right plot of fig. 1, which shows the $p$-value 1-CL$_b$ for Higgs-boson mass hypotheses ranging from 110 to 145 GeV. The minimum $p$-value is observed at $m_H = 120\,\text{GeV}$, corresponding to a significance of 2.93 standard deviations. For $m_H = 125.8\,\text{GeV}$, the significance is $2.85\sigma$. The best-fit value for the signal strength combining all channels is $\hat{\mu} = 1.1 \pm 0.4$ at $m_H = 125\,\text{GeV}$.

## 4. – $H \to b\bar{b}$ search with CMS detector

At a mass of 125 GeV the standard model Higgs boson decays predominantly into a $b\bar{b}$ pair. The observation and study of the $H \to b\bar{b}$ decay is therefore essential in determining the nature of the newly discovered boson. Recently, in a search for the standard model Higgs boson where the sensitivity below a mass of 130 GeV is dominated by the channels in which the Higgs boson is produced in association with a weak vector boson and decaying to $b\bar{b}$, the CDF and D0 collaborations have reported evidence for an excess of events at a local significance of 3.0 standard deviations for a mass of 125 GeV [13]. Here a search for the standard model Higgs boson in the VH production mode is presented, where $V$ is either a $W$ or a $Z$ boson and $H \to b\bar{b}$ [14]. The following six channels are included in the search: $W(\mu\nu)H$, $W(e\nu)H$, $W(\tau\nu)H$, $Z(\mu\mu)H$, $Z(ee)H$ and $Z(\nu\nu)H$, all with the Higgs boson decaying to $b\bar{b}$.

4·1. *Selection*. – The event selection is based on the kinematic reconstruction of the vector bosons in their leptonic decay modes and of the Higgs boson decay into two $b$-tagged jets. Backgrounds are substantially reduced by requiring a significant boost of the $p_T$ of the vector boson, $p_T(V)$, or the Higgs boson [15]. In that case the two particles recoil away from each other with a large azimuthal opening angle, $\Delta\phi(V, H)$, between them. For each channel, different regions of $p_T(V)$ boost are considered. The low, intermediate, and high boost regions for the $W(\mu\nu)H$ and $W(e\nu)H$ channels are $100 < p_T(V) < 130\,\text{GeV}$, $130 < p_T(V) < 180\,\text{GeV}$, and $p_T(V) > 180\,\text{GeV}$. For the $W(\tau\nu)H$ a single $p_T(V) > 120\,\text{GeV}$ region is considered. For the $Z(\nu\nu)H$ channel the low, intermediate, and high boost regions are $100 < p_T(V) < 130\,\mathrm{GeV}$, $130 < p_T(V) < 170\,\mathrm{GeV}$ and $p_T(V) > 170\,\mathrm{GeV}$, and for the $Z(\nu\nu)H$ channels, the low and high regions are $50 < p_T(V) < 100\,\mathrm{GeV}$ and $p_T(V) > 100\,\mathrm{GeV}$.

The reconstruction of the $H \to b\bar{b}$ decay is made by requiring the presence of two central ($|\eta| < 2.5$) jets above a minimum $p_T$ threshold, and tagged by the CSV algorithm, requiring that the value of the CSV discriminator be above a certain threshold. After all event selection criteria the dijet invariant mass resolution of the two $b$-jets from the Higgs decay is approximately 10%. The Higgs boson mass resolution is improved by applying regression techniques similar to those used at the CDF experiment [16].

In the final stage of the analysis, to better separate signal from background under different Higgs boson mass hypotheses a boosted decision tree (BDT) algorithm is trained separately at each mass value using simulated samples for signal and all background processes that pass the event selection. The inputs variables used are kinematics (dijet transverse momentum and mass, vector boson transverse momentum, etc.), topological (azimuthal angle between the vector boson and dijet, distance in $\eta - \phi$ between Higgs decay products, number of additional jets, etc.) and $b$-tag of the two jets. The shape of the output distribution of this BDT discriminant is the final discriminant on which a fit is performed to search for events resulting from Higgs boson production.

**4**˙2. *Background control region and systematics.* – Backgrounds arise from production of $W$ and $Z$ bosons in association with jets (from all quark flavors), singly and pair-produced top quarks ($t\bar{t}$), dibosons and QCD multijet processes. Control regions in data are selected to adjust the event yields from simulation for the main background processes and to estimate their contribution in the signal region.

We report also in brief the effects of the most important systematic uncertainties of the analysis. The jet energy scale is varied within one standard deviation as a function of jet $p_T$ and $\eta$. The efficiency of the analysis selection is recomputed to assess the variation in yield. Depending on the process, a 2-3% yield variation is found. The effect of the uncertainty on the jet energy resolution is evaluated by smearing the jet energies according to the measured uncertainty. Depending on the process, a 3-6% variation in yields due to this effect is obtained. Data-to-simulation $b$-tagging scale factors, measured in $t\bar{t}$ events, are applied consistently to jets in signal and background events. These result into yield uncertainties in the 3-15% range, depending on the channel and the specific process. This analysis is performed in the boosted regime, and thus, potential differences in the $p_T$ spectrum of the $V$ and $H$ between data and Monte Carlo generators could introduce systematic effects in the signal acceptance and efficiency estimates. Two calculations are available that estimate the NLO electroweak [17] and NNLO QCD [18] corrections to $VH$ production in the boosted regime. Both the EWK and NNLO QCD corrections are applied to the signal samples. The estimated uncertainties from the NNLO electroweak corrections are 2% for $ZH$ and 2% for $WH$. For the NNLO QCD correction, an uncertainty of 5% for both $ZH$ and $WH$ is estimated. The uncertainty in the background yields that results from the estimates from data is approximately 10%.

The combined effect of the systematic uncertainties results in an increase of about 15% on the expected upper limit on the Higgs boson production cross section and in a reduction of 15% on the expected significance of an observation when the Higgs boson is present in the data at the predicted standard model rate.

**4**˙3. *Results.* – Results are obtained from combined signal and background fits to the shape of the output distributions of the BDT discriminants trained separately for each

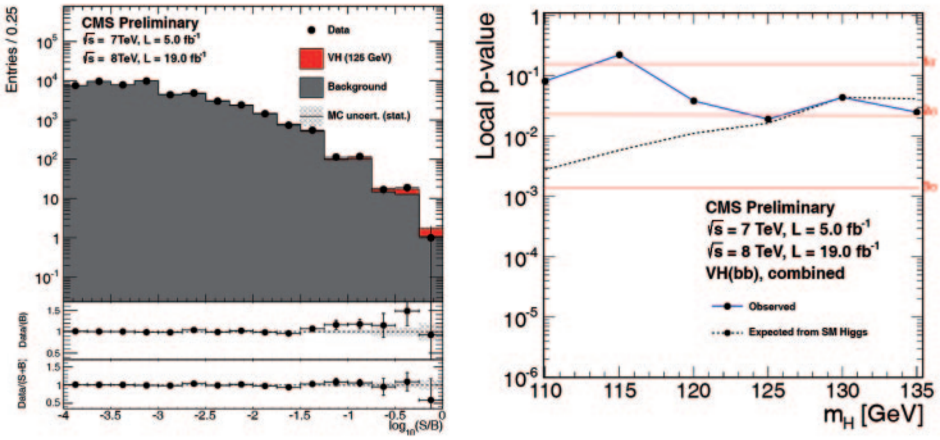

Fig. 2. – Left: combination of all BDT discriminants of the VHbb analysis into a single distribution where all events, for all channels, are sorted in bins of similar expected signal-to-background ratio, as given by the value of the output of their corresponding BDT discriminant. Right: observed and expected $p$-value 1-CL$_b$, and the corresponding significance in number of standard deviations for CMS VHbb analysis.

channel and for each Higgs boson mass hypothesis in the 110–135 GeV range examined. Figure 2 left combines all these discriminants into a single distribution where all events, for all channels, are sorted in bins of similar expected signal-to-background ratio, as given by the value of the output of their corresponding BDT discriminant (trained with a Higgs boson mass of 125 GeV). The observed excess of events in the bins with the largest signal-to-background ratio is consistent with what is expected from the production of a standard model Higgs boson. Probabilities ($p$-values) that the observed excess is due to background fluctuations alone, as a function of the Higgs boson mass hypothesis, are shown in fig. 2 right. For a mass of 125 GeV the excess of observed events is 2.1 standard deviations, and is consistent with the standard model prediction for Higgs boson production. The observed signal strengths for the individual modes are consistent with each other and the value for the signal strength for all modes combined is $1.0 \pm 0.5$.

## 5. – Summary

The most recent results on Higgs decaying to $\tau$ leptons and $b$-quarks with the CMS experiment at LHC have been reported. The measured local significance is 2.9 and $2.1\sigma$ for Higgs decaying to $\tau$ and $b$-jets respectively for $m_H = 125\,\text{GeV}$, with the excess spanning on a wide region from 115 to 135 GeV, given the mass resolution in the two channels. The measured signal strength of $1.1 \pm 0.4$ and $1.0 \pm 0.5$ respectively.

These results confirm that the new boson found in 2012 by CMS and ATLAS collaboration couples to fermions with signal strength close to prediction for a Standard Model-like Higgs boson. The combination of the latest CMS $H \rightarrow \tau^+\tau^-$ and $H \rightarrow b\bar{b}$ results report an excess of coupling to fermions of 3.4 observed significance (3.5 expected), being the first evidence at LHC of Higgs coupling to fermions.

## REFERENCES

[1] ATLAS Collaboration, *Phys. Lett. B*, **716** (2012) 1.
[2] CMS Collaboration, *Phys. Lett. B*, **716** (2012) 30.
[3] Higgs P. W., *Phys. Lett.*, **12** (1964) 132.
[4] CMS Collaboration, *The CMS experiment at the CERN LHC* JINST 0803:S08004 2008.
[5] CMS Collaboration, *Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and MET*, CMS Physics Analysis Summary CMS-PAS-PFT-09-001, (2009).
[6] Cacciari M., Salam G. P. and Soyez G., *JHEP*, **04** (2008) 063.
[7] CMS Collaboration, *Identification of b-quark jets with the CMS experiment*, CMS Physics Analysis Summary CMS-PAS-BTV-12-001 (2012).
[8] CMS Collaboration, *JINST*, **7** (2012) P01001.
[9] Nason P. and Oleari C., *JHEP*, **02** (2010) 037.
[10] Sjostrand T., Mrenna S. and Skands P., *JHEP*, **05** (2006) 026.
[11] Alwall J. *et al.*, *JHEP*, **09** (2007) 028.
[12] GEANT4 Collaboration, *Nucl. Instrum. Methods A*, **506** (2003) 250.
[13] CDF Collaboration, D0 Collaboration, *Higgs Boson Studies at the Tevatron*, arXiv:1303.6346.
[14] CMS Collaboration, *Search for the standard model Higgs boson produced in association with W or Z bosons, and decaying to bottom quarks*, CMS-PAS-HIG-13-012, 2013.
[15] Butterworth J. M. *et al.*, *Phys. Rev. Lett.*, **100** (2008) 242001.
[16] Aaltonen T. *et al.*, *Improved b-jet Energy Correction for $H \to b\bar{b}$ Searches at CDF*, arXiv:1107.3026.
[17] Ciccolini M. *et al.*, *Phys. Rev. Lett.*, **99** (2007) 161803.
[18] Ferrera G., Grazzini M. and Tramontano F., *Associated WH production at hadron colliders: a fully exclusive QCD calculation at NNLO*, arXiv:1107.1164.