EVIDENCE OF THE SM HIGGS BOSON IN THE DECAY CHANNEL INTO TAU LEPTONS

R.A. MANZONI
INFN & Università degli Studi - Milano Bicocca,
piazza della Scienza 3, 20126 Milan, Italy

A search for the standard model Higgs boson decaying into a pair of tau leptons is performed using events recorded by the CMS experiment at the LHC in 2011 and 2012. The dataset corresponds to an integrated luminosity of 4.9 fb^{-1} at a centre-of-mass energy of 7 TeV and 19.7 fb^{-1} at 8 TeV. An excess of events is observed over the expected background contributions, with a local significance larger than 3 standard deviations for m_{H} values between 110 and 130 GeV.

1 Analysis motivations

After the discovery of a new boson with mass close to 125 GeV, announced by ATLAS [1] and CMS [2] Collaborations on July 4th 2012, the focus aimed at discerning the properties of this new particle and whether it behaves like the Higgs boson predicted in the Standard Model.

In the SM, the masses of the fermions are generated by the interaction between the Higgs' and the fermions' fields through the Yukawa couplings. The measurements of these couplings are therefore fundamental to completely identify this boson as the SM Higgs boson.

The di-τ decay channel is the most promising, offering low background contamination compared to the $\text{b}\bar{\text{b}}$ channel and large predicted rate compared to other leptonic channels.

2 Analysis strategy

The analysis [3] is designed to probe the three main Higgs boson production mechanisms at LHC: the gluon fusion, the Vector Boson Fusion (VBF) and the production in association to a W or a Z boson.

2.1 Physics objects

Several physics objects are involved in the analysis: electrons, muons, hadronically decaying taus (τ_h), jets and missing transverse energy ($\not{E}_T$). All of them are built upon the *Particle-Flow* (PF) algorithm [4,5,6], that exploits the information from the all the CMS subdetectors to identify the stable particles generated in the event.

Fig. 1 shows the mass of the visible products of the τ_h, reconstructed using the *Hadron Plus Strip* (HPS) algorithm [7] that identifies the main decay modes of the τ_h, out of charged hadrons and photons.

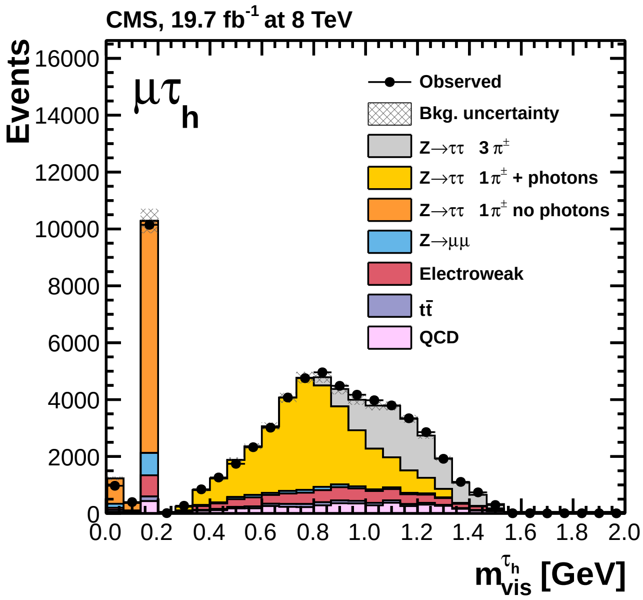

Figure 1 – Observed and predicted distributions for the visible τ_h mass, $m_{\mathrm{vis}}^{\tau_h}$, in the $\mu\tau_h$ channel after the baseline selection. The $\mathrm{Z} \to \tau\tau$ contribution is then split according to the decay mode reconstructed by the HPS algorithm as shown in the legend. The mass distribution of the τ built from one charged hadron and photons peaks near the mass of the intermediate $\rho(770)$ resonance; the mass distribution of the τ built from three charged hadrons peaks around the mass of the intermediate $a_1(1260)$ resonance. The τ_h built from one charged hadron and no photons are reconstructed with the $\pi^{\pm}$ mass, assigned to all charged hadrons by the PF algorithm, and constitute the main contribution to the third bin of this histogram. The "Bkg. uncertainty" band represents the combined statistical and systematic uncertainty in the background yield in each bin. The expected contribution from the SM Higgs signal is negligible.

2.2 Backgrounds

In the $\mu\tau_h$, $\mathrm{e}\tau_h$, $\tau_h\tau_h$ and $\mathrm{e}\mu$ channels, the largest source of irreducible background is the Drell-Yan production of $\mathrm{Z} \to \tau\tau$. Other important sources of background are constituted by W+jets processes, contributing significantly to the $\mu\tau_h$ and $\mathrm{e}\tau_h$ channels when the W boson decays leptonically and a jet is misidentified as a τ_h, QCD processes, contributing mostly to the $\tau_h\tau_h$ channel when two jets are misidentified as τ_h, and $\mathrm{t\bar{t}}$ processes decaying into the fully leptonic, opposite flavor final state, contributing to the $\mathrm{e}\mu$ channel.

The Drell-Yan production of $\mathrm{Z} \to \ell\ell$ is the largest background in the $\ell\ell$ channels whilst the di-bosons processes are a relevant source of contamination for the WH and ZH channels.

All the most important backgrounds are entirely derived from data or controlled in data sidebands.

2.3 Event classification

For the Higgs boson production through gluon fusion and VBF, all the six di-τ final states are considered: $\mu\tau_h$, $\mathrm{e}\tau_h$, $\tau_h\tau_h$, $\mathrm{e}\mu$, $\mu\mu$, and ee.

The events are then classified in mutually exclusive categories, according to the number of reconstructed jets in the final state and, in particular, the contribution of the VBF production process is enhanced by requiring a large rapidity gap between the two jets with the highest transverse momentum. Moreover, the transverse momentum of the τ (for $\mu\tau_h$, $\mathrm{e}\tau_h$ channels) or the μ (for $\mathrm{e}\mu$ channel) and the transverse momentum of the Higgs candidate are used to further split the events in order to enhance the sensitivity.

For the Higgs boson production in association with a vector boson, the presence of one or two additional light leptons, compatible with the presence of a W or a Z boson respectively, is

required.
Four channels are considered for at the WH process: $\mu + \mu\tau_h$, $\mu + \mathrm{e}\tau_h/\mathrm{e} + \mu\tau_h$, $\mu + \tau_h\tau_h$ and $\mathrm{e} + \tau_h\tau_h$, whilst eight channels are aimed at ZH process: $\ell\ell + \mu\tau_h$, $\ell\ell + \mathrm{e}\tau_h$, $\ell\ell + \tau_h\tau_h$, $\ell\ell + \mathrm{e}\mu$, where $\ell\ell$ can be either $\mu\mu$ or ee.

2.4 Signal extraction

The signal extraction is carried out through a global, binned, maximum likelihood fit based on the following final discriminating variables:

- the invariant mass of the di-τ pair, $m_{\tau\tau}$, reconstructed using the SVFIT algorithm [3] from the four momenta of the visible decay products of the Higgs and the $\not\!\!E_T$, is used for all the channels except for $\mu\mu$, ee and WH;
- the mass of the visible decay products of the Higgs, m_{vis} is used for WH channels. The $m_{\tau\tau}$ is not used because the contributions to the $\not\!\!E_T$ are from both the neutrinos coming from the Higgs and the neutrino coming from the W boson;
- the output distribution of a multivariate discriminator, comprising various kinematic and topological variables, is used for $\mu\mu$ and ee channels

The expected number of signal events is the one predicted by the SM for the production of a Higgs boson of mass m_{H} decaying into a pair of τ leptons, multiplied by a signal strength modifier μ treated as free parameter in the fit. The contribution from the $\mathrm{H} \to \mathrm{WW}$ process, decaying into lepton pairs, ee, $\mu\mu$ or eμ is non negligible, especially in the eμ case, and it is treated as background throughout except where explicitly noted.

The systematic uncertainties are represented by nuisance parameters that are varied in the fit according to their probability density function, affecting both the shape and the yields of the different contributions.

3 Results

An excess of events over the background-only hypothesis is observed with a local significance larger than 3 standard deviations for Higgs boson mass hypotheses between $m_{\mathrm{H}} = 115$ and 130 GeV, and equal to 3.2 standard deviations at $m_{\mathrm{H}} = 125$ GeV, to be compared to an expected significance of 3.7 standard deviations (Fig. 2, right).

The best-fit value for μ, combining all channels, is $\hat{\mu} = 0.78 \pm 0.27$ at $m_H = 125$ GeV.

The mass of the Higgs boson has been measured by performing a scan of the negative log likelihood as a function of mass m_{H} (Fig. 3, left), resulting in a best-fit mass $m_H = 122 \pm 7$ GeV.

The results are compatible with the prediction of a SM Higgs boson at 125 GeV and they constitute the evidence for the coupling between the τ lepton and the Higgs boson discovered by the ATLAS and CMS Collaborations.

References

1. ATLAS Collaboration, *Phys. Lett.* B **716**, 1-29 (2012).
2. CMS Collaboration, *Phys. Lett.* B **716**, 30-62 (2012).
3. CMS Collaboration, http://arxiv.org/abs/arXiv:1401.5041, (2013).
4. CMS Collaboration, *CMS-PAS-PFT-09-001*, (2009).
5. CMS Collaboration, *CMS-PAS-PFT-10-002*, (2010).
6. CMS Collaboration, *CMS-PAS-PFT-10-003*, (2010).
7. CMS Collaboration, *JINST* **7**, P01001 (2012).

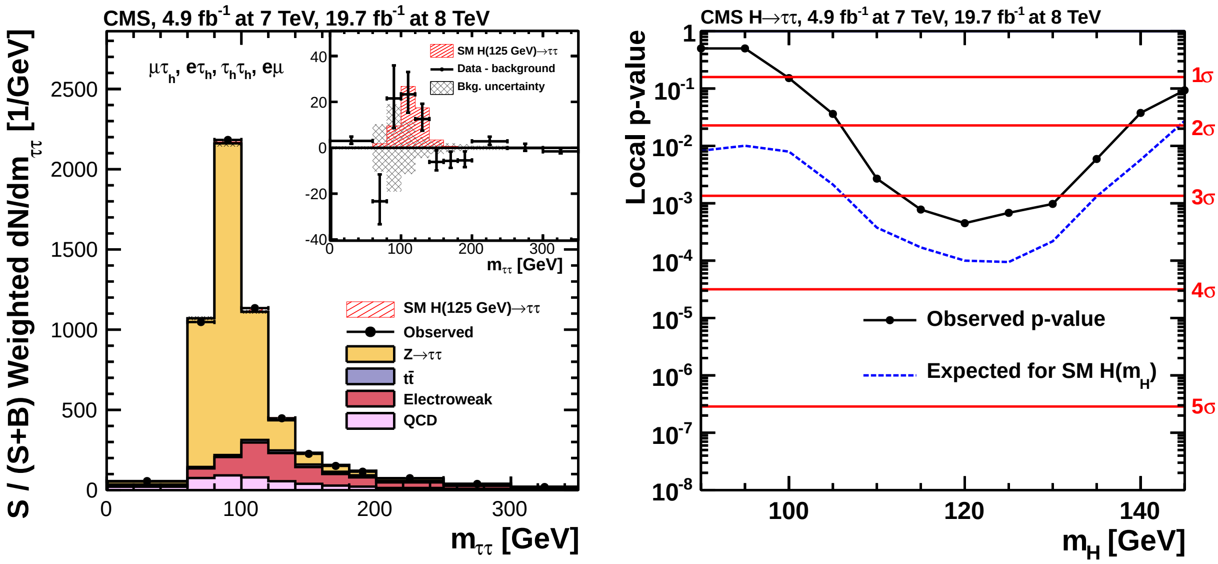

Figure 2 – The plot on the left shows the combined observed and predicted $m_{\tau\tau}$ distributions for the $\mu\tau_h$, eτ_h, $\tau_h\tau_h$, and eμ channels. The signal distribution is normalized to the SM prediction ($\mu = 1$). The distributions obtained in each category of each channel are weighted by the ratio between the expected signal and signal-plus-background yields in the category, obtained in the central $m_{\tau\tau}$ interval containing 68% of the signal events. The inset shows the corresponding difference between the observed data and expected background distributions, together with the signal distribution for a SM Higgs boson at $m_H = 125$ GeV.
The plot on the right shows the local p-value and significance in number of standard deviations as a function of the SM Higgs boson mass hypothesis for the combination of all decay channels. The observation (solid line) is compared to the expectation (dashed line) for a SM Higgs boson with mass m_{H}.

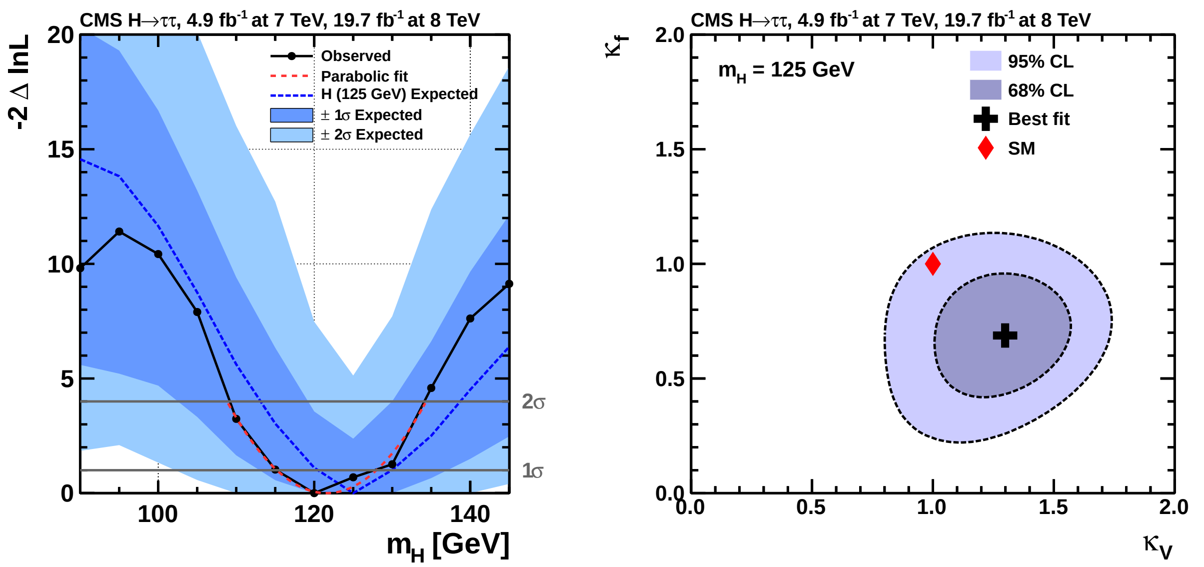

Figure 3 – Scan of the negative log-likelihood difference, $-2\Delta\ln\mathcal{L}$, as a function of m_{H} (left) and as a function of κ_V and κ_f (right). For each point, all nuisance parameters are profiled. The observation (solid line) is compared to the expectation (dashed line) for a SM Higgs boson with mass $m_H = 125$ GeV. For the likelihood scan as a function of κ_V and κ_f, and for this measurement only, the $H \to WW$ contribution is treated as a signal process.