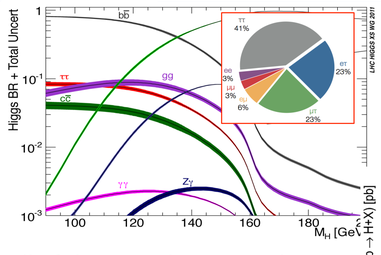
Evidence for the SM Higgs boson in the ττ decay channel using the CMS detector

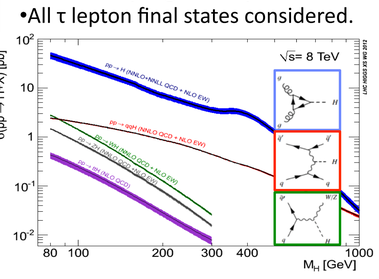
R. Venditti on behalf of the CMS Collaboration
University and INFN Bari

A search for the Standard Model Higgs boson decaying to tau pair has been performed in pp collision data recorded by CMS detector at LHC at centre-of-mass energy 8 TeV (7 TeV) corresponding to an integrated luminosity 20/fb (5/fb). The production modes considered are gluon-gluon fusion, VBF and associated production with a vector boson. The analysis strategy and the resulting evidence for the Higgs boson in tau pair channel are reported.

H→ττ: the theoretical predictions

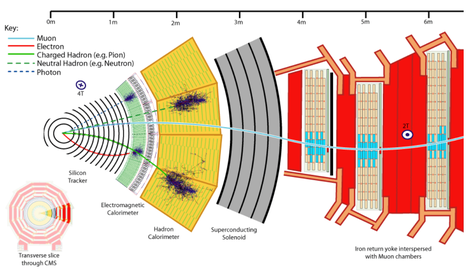


Sizeble SM BR(H→ττ) in the light m_H region (6.3% at m_H=125 GeV/c^2).
•Leading role of H→ττ in probing Yukawa coupling with leptons.
•All τ lepton final states considered.



All of the Higgs boson production mechanism have been exploited.
•Gluon-gluon fusion
•VBF
•Associated production with W/Z

The CMS Experiment



Information from all the subdetectors→Global event description.
Charged and neutral hadrons, photons, muons, electrons are first identified and then used to reconstruct the missing energy (E^{miss}), jets, τ and lepton isolation.

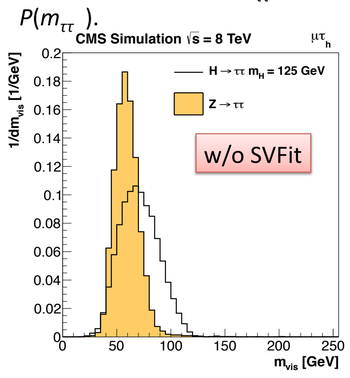
Whole data collected in 2011 and 2012 corresponding to an integrated luminosity of **4.9/fb at 7 TeV and 19.7/fb at 8 TeV** used in the search.

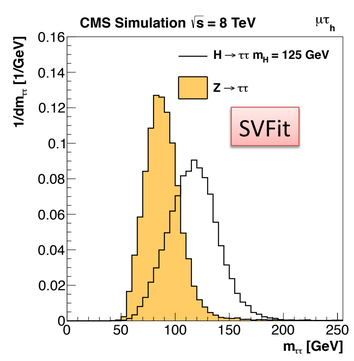
τ-pair invariant mass reconstruction

Neutrinos from τ decay take away a large amount of undetected energy. A kinematic maximum likelihood fit is performed to estimate the mass of ττ system on event-by-event basis

$$P(m^i_{\tau\tau}) = \int \delta\left(m^i_{\tau\tau} - m_{\tau\tau}(\vec{y},\vec{a_1},\vec{a_2})\right) f(\vec{z},\vec{y},\vec{a_1},\vec{a_2})\, d\vec{a_1}\, d\vec{a_2},$$

•a_1, a_2 are measured τ kinematic variables.
•z is the measured missing energy.
•$f(z,y,a_1,a_2) = \mathcal{L}(a_1)*\mathcal{L}(a_2)*\mathcal{L}(z)$.
The best estimate of $m_{\tau\tau}$ is taken to be the value of $m_{\tau\tau}$ that maximizes $P(m_{\tau\tau})$.





✓**Reconstructed mass of the ττ system discriminates best between *H* → ττ signal and the *Z* → ττ background.**
✓Likelihood mass reconstruction leads to a **40% improved sensitivity** when compared to the use of the invariant mass calculated only from the visible decay products of the di-τ pair (visible mass).
✓Used in all no WH channles, where the visible mass is used.

Event Categories

Event sample is split into **mutually exclusive categories**, defined to maximize the sensitivity of the analysis to the presence of a SM Higgs boson. In each channel events are classified according to the number of jets.
•**VBF Tag**: large invariant mass m_{jj}, large $|\Delta\eta_{jj}|$. Central-jet veto. High sensitivity.
•**0 jet**:low sensitivity. Mainly used to constrain the Z →ττ background for the more sensitive categories

		0-jet	1-jet		2-jet	
				$p_T^{\tau\tau}$ > 100 GeV	m_{jj} > 500 GeV $\|\Delta\eta_{jj}\|$ > 3.5	$p_T^{\tau\tau}$ > 100 GeV m_{jj} > 700 GeV $\|\Delta\eta_{jj}\|$ > 4.0
$\mu\tau_h$	$p_T^{\tau h}$ > 45 GeV	high-$p_T^{\tau h}$	high-$p_T^{\tau h}$	high-$p_T^{\tau h}$ boosted	loose VBF tag	tight VBF tag (2012 only)
	baseline	low-$p_T^{\tau h}$	low-$p_T^{\tau h}$			
$e\tau_h$	$p_T^{\tau h}$ > 45 GeV	high-$p_T^{\tau h}$	~~high-$p_T^{\tau h}$~~	high-$p_T^{\tau h}$ boosted	loose VBF tag	tight VBF tag (2012 only)
	baseline	low-$p_T^{\tau h}$	low-$p_T^{\tau h}$			
			E_T^{miss} > 30 GeV			
eμ	p_T^{μ} > 35 GeV	high-p_T^{μ}	high-p_T^{μ}		loose VBF tag	tight VBF tag (2012 only)
	baseline	low-p_T^{μ}	low-p_T^{μ}			
ee, μμ	p_T^{l} > 35 GeV	high-p_T^{l}	high-p_T^{l}		2-jet	
	baseline	low-p_T^{l}	low-p_T^{l}			
$\tau_h\tau_h$ (8 TeV only)			boosted	highly boosted	VBF tag	
	baseline		$p_T^{\tau\tau}$ > 100 GeV	$p_T^{\tau\tau}$ > 170 GeV	$p_T^{\tau\tau}$ > 100 GeV m_{jj} > 500 GeV $\|\Delta\eta_{jj}\|$ > 3.5	

Event Selection

Events are selected and classified in the various channels according to the number of selected electrons, muons, and τ_h. The resulting event samples are independent.
✓**Kinematics selections** according to detector acceptance and trigger thresholds.

Channel	Lepton selection		
$\mu\tau_h$	$p_T^{\mu} > 17$–20	$\|\eta^{\mu}\| < 2.1$	$R^{\mu} < 0.1$
	$p_T^{\tau_h} > 30$	$\|\eta^{\tau_h}\| < 2.4$	$I^{\tau_h} < 1.5$
$e\tau_h$	$p_T^{e} > 20$–24	$\|\eta^{e}\| < 2.1$	$R^{e} < 0.1$
	$p_T^{\tau_h} > 30$	$\|\eta^{\tau_h}\| < 2.4$	$I^{\tau_h} < 1.5$
$\tau_h\tau_h$ (2012 only)	$p_T^{\tau_h} > 45$	$\|\eta^{\tau_h}\| < 2.1$	$I^{\tau_h} < 1$
$e\mu$	$p_T^{\ell_1} > 20$	$\|\eta^{\mu}\| < 2.1$	$R^{\ell} < 0.1$–0.15
	$p_T^{\ell_2} > 10$	$\|\eta^{e}\| < 2.3$	
$\mu\mu$	$p_T^{\mu_1} > 20$	$\|\eta^{\mu_1}\| < 2.1$	$R^{\mu} < 0.1$
	$p_T^{\mu_2} > 10$	$\|\eta^{\mu_2}\| < 2.4$	
ee	$p_T^{e_1} > 20$	$\|\eta^{e}\| < 2.3$	$R^{e} < 0.1$–0.15
	$p_T^{e_2} > 10$		
$\mu + \mu\tau_h$	$p_T^{\mu_1} > 20$	$\|\eta^{\mu}\| < 2.4$	$R^{\mu} < 0.1$–0.2
	$p_T^{\mu_2} > 10$		
	$p_T^{\tau_h} > 20$	$\|\eta^{\tau_h}\| < 2.3$	$I^{\tau_h} < 2$
$e + \mu\tau_h$ / $\mu + e\tau_h$	$p_T^{\ell_1} > 20$	$\|\eta^{e}\| < 2.5$	$R^{\ell} < 0.1$–0.2
	$p_T^{\ell_2} > 10$	$\|\eta^{\mu}\| < 2.4$	
	$p_T^{\tau_h} > 20$	$\|\eta^{\tau_h}\| < 2.3$	$I^{\tau_h} < 2$
$\mu + \tau_h\tau_h$	$p_T^{\mu} > 24$	$\|\eta^{\mu}\| < 2.1$	$R^{\mu} < 0.1$
	$p_T^{\tau_{h,1}} > 25$	$\|\eta^{\tau_h}\| < 2.3$	$I^{\tau_h} < 2$–3
	$p_T^{\tau_{h,2}} > 20$		
$e + \tau_h\tau_h$	$p_T^{e} > 24$	$\|\eta^{e}\| < 2.1$	$R^{e} < 0.1$–0.15
	$p_T^{\tau_{h,1}} > 25$	$\|\eta^{\tau_h}\| < 2.3$	$I^{\tau_h} < 2$
	$p_T^{\tau_{h,2}} > 20$		

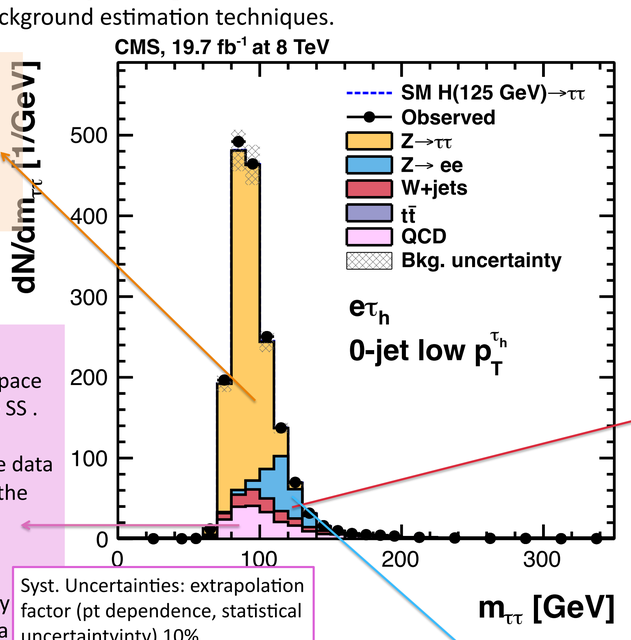
✓**Topological selections** optimized in each channel.
OS charge for leptons assigned to the Higgs boson.
W+jets suppression:Mt(ℓ,E^{miss})<30 GeV ($\ell\tau$), BDT(E^{miss}, ττ kin) ($\ell\tau\tau$)
ttbar suppresion: b-jet veto, BDT (eμ)
Z→$\ell\ell$: light lepton **SS ($\ell\ell\tau$)**

Background Estimation

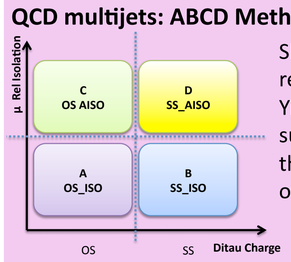
The estimation of the shape and yield of the major backgrounds in each channel is based on the observed data. The experimental systematic uncertainties are thus related to the background estimation techniques.

Z→ττ: Embedding
Main irreducible background. Estimated from data events with loose Z→μμ selections applied. Muons are replaced by the τ leptons taken in simulated Z →ττ events.

Syst. Uncertainties:
Tau reconstruction and trigger efficiency (8%)



QCD multijets: ABCD Method

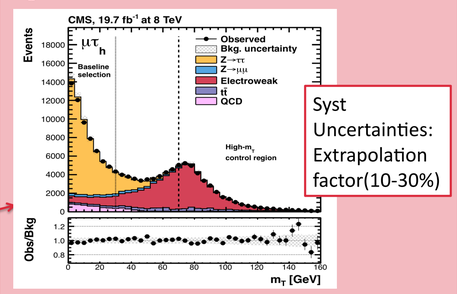


Shape from phase space region where $\ell\tau$ are SS . Yield obtained by subtracting from the data the contribution of the other processes.

Extrapolation into the signal sample is derived by rescaling the yield obtained in the SS region by a factor of 1.06 measured in pure QCD multijet sample obtained by inverting the lepton isolation.

Syst. Uncertainties: extrapolation factor (pt dependence, statistical uncertaintyinty) 10%

W+jets: jet misidentified as τ.
Mt Sideband ($\ell\tau$)
Shape from high-mT control region. The extrapolation factor to the low- *mT* signal region is obtained from the simulation.



Syst Uncertainties:
Extrapolation factor(10-30%)

Fake Rate ($\ell\tau\tau$)
Measure in data jet→τ(ℓ) mis-identification probability in background enriched region *f*. Inverted identification criteria on τ(ℓ) and extrapolate contribution into the signal region by appyling *f/1-f*.

ttbar: Simulation
Yield adjusted to the one observed using a tt-enriched control sample, extracted by requiring b-jets.

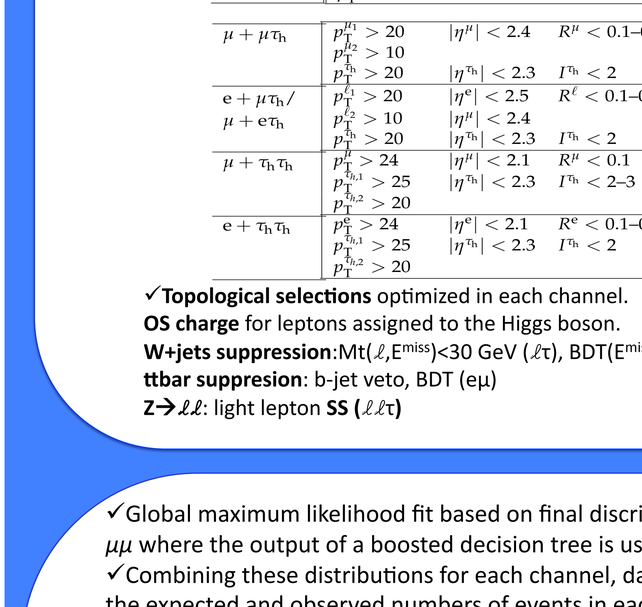
Syst. Uncertainties:
B-tagging eff,
E^{miss} scale,Jet energy scale

Z→$\ell\ell$: Simulation
Estimated from simulation, after rescaling the simulated Drell–Yan yield to the one derived from Z→μμ data. Large contribution in eτ due to the high e→τ fake rate

Syst Uncertainties:
ℓ→τ misidentification rate (20-80%)

Final Results

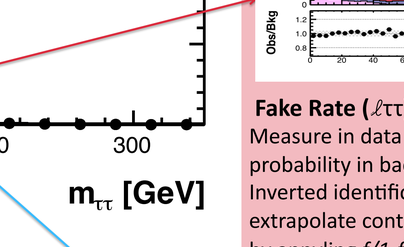
✓Global maximum likelihood fit based on final discriminating variables: $m_{\tau\tau}$ or m^{vis} in all channels except for ee and μμ where the output of a boosted decision tree is used.
✓Combining these distributions for each channel, data-taking period and category in a binned likelihood involving the expected and observed numbers of events in each bin.
✓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 τ pair, multiplied by a signal strength modifier $\mu=\sigma/\sigma(SM)$ treated as free parameter in the fit.
✓Systematic uncertainties represented by nuisance parameters and varied in the fit according to their pdf.



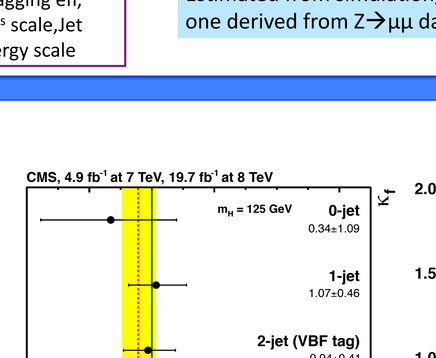
Clear excess of signal events in the observed and expected $m_{\tau\tau}$ distributions for the most sensitive channels.

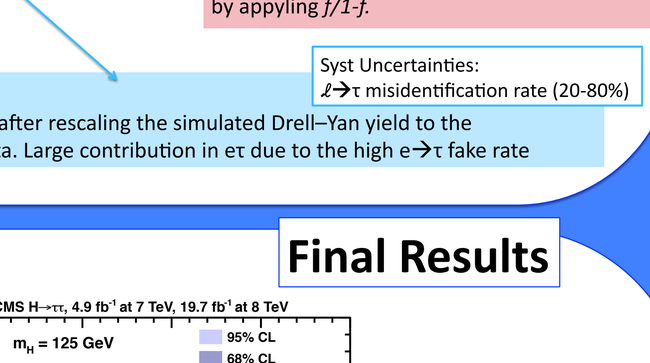
Local p-values and significances as a function of the Higgs boson mass hypothesis. The expected significance at m_H = 125 GeV is 3.6 σ. Observed significance equals 3.4 σ for m_H = 125 GeV.

Best-fit signal strength values per category for m_H = 125 GeV. The combined value corresponds to μ=**0.78 ± 0.27**. The dashed line corresponds to the best-fit μ value.



Likelihood scan as a function of the ratio between measured and the SM value for the coupling of the Higgs boson to vector bosons and fermions for m_H=125. **Observation is consistent with the SM expectation of $k_V = k_f = 1$.**





Combined observed and expected 95% CL upper limit on the signal strength parameter in the background-only hypothesis

The CMS Collaboration, Evidence for the 125 GeV Higgs boson decaying to a pair of τ leptons , CMS-PAS-HIG-13-004, http://arxiv.org/abs/1401.5041

rosamaria.venditti@ba.infn.it