Search with the CMS detector for the SM Higgs boson decaying to a pair of hadronically decaying τ-leptons, produced in association with a $W^{\pm}$ boson

Anthony Rose
On behalf of the CMS collaboration

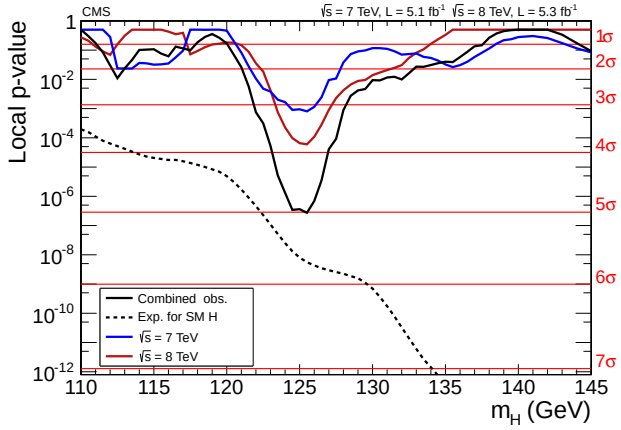
2014 Phenomenology Symposium
May 5, 2014

Outline

1. Motivation
2. Tau identification
3. $W^{\pm}H \to \tau_h\tau_h\ell^{\pm}$ analysis

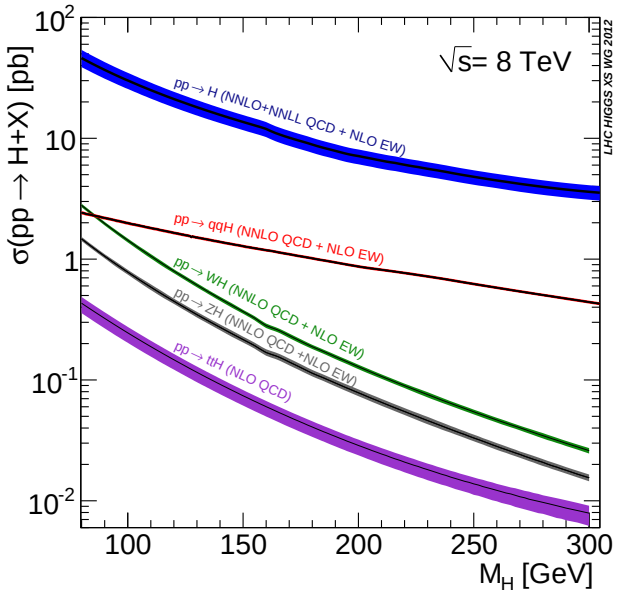
Motivation

Following the July 2012 discovery of a new particle compatible with the SM Higgs there was still a lot of work to be done.



Variety of Higgs production mechanisms to explore:

- → Gluon-gluon fusion
- → Vector boson fusion
- → Associated production
- → $t\bar{t}H$



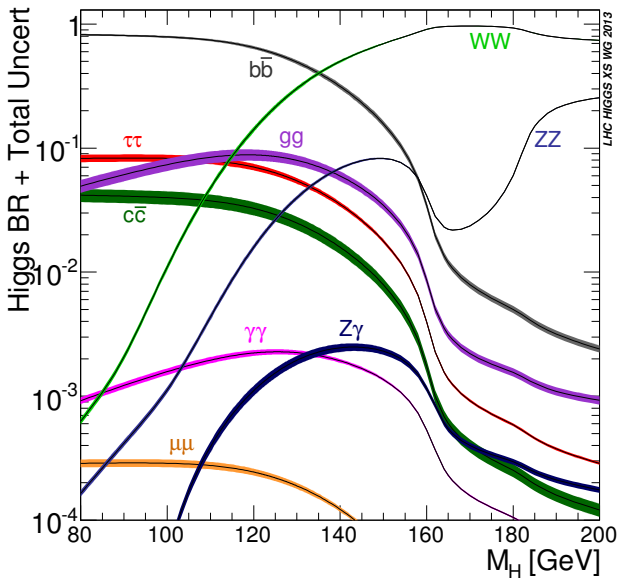
Associated production has a smaller cross section than the dominant processes, but leptonic decay of the $W^{\pm}/Z$ boson gives extra ability to distinguish these events from backgrounds.

$H \to \tau\tau$

Along with a variety of production mechanisms, there are also many Higgs decay modes to be studied.

At $m_H \approx 125$ GeV, $H \to \tau\tau$ is particularly interesting.

- ✓ Relatively high branching ratio
- ✓ Lower background than $b\bar{b}$ and gg
- X τ-leptons more difficult to reconstruct than e, μ
- X Still a large background ($Z \to \tau\tau$), but associated production of vector boson provides extra light lepton

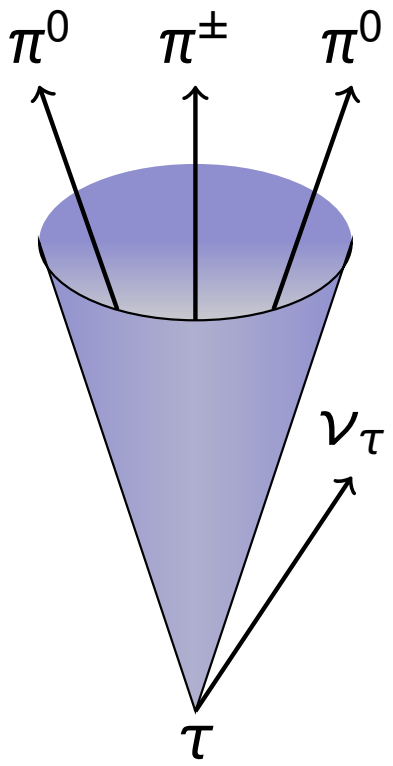


It is crucial to search for the Higgs in every possible channel and measure properties

Tau identification at CMS

τ-leptons

- → Heaviest lepton ($m_\tau = 1.78$ GeV)
- → Short lifetime ($c\tau = 87\mu m$)
- → Decays to light leptons $\sim 35\%$
- → Decays to hadrons $\sim 65\%$
  - 1-prong decays ($\sim 85\%$)
    - $\tau^\pm \to \pi^\pm \nu_\tau$
    - $\tau^\pm \to \pi^\pm \pi^0 \nu_\tau$
  - 3-prong decays ($\sim 15\%$)
    - $\tau^\pm \to \pi^\pm K_S \to \pi^\pm \pi^\pm \pi^\mp \nu_\tau$



Detector signature:

Isolated and collimated jets with low charged track multiplicty

Isolation requirement vital in rejecting jet fakes

Reconstruction performed with *Hadrons Plus Strips* algorithm:

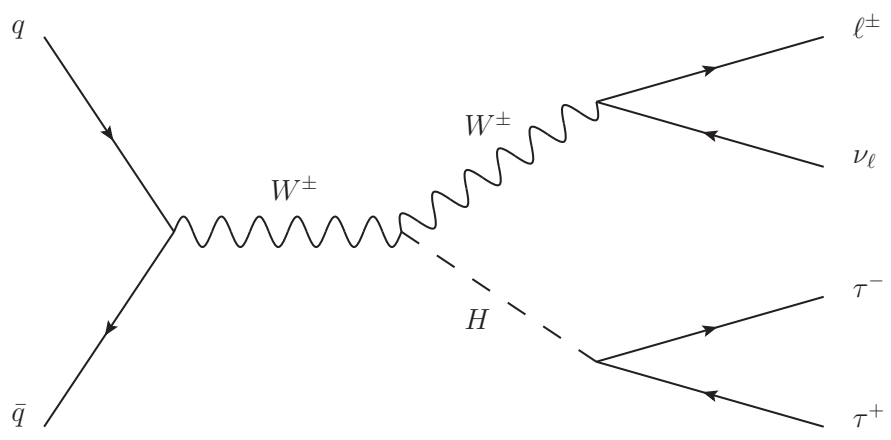
- → Charged hadrons reconstructed with Particle Flow algorithm, combined with γ from π^0 in strips of η.

$W^{\pm}H \rightarrow \tau_h\tau_h\ell^{\pm}$ analysis

Analysis overview

- → Search for production of Standard Model Higgs boson, in association with a $W^{\pm}$ boson
- → Higgs decays to a pair of τ-leptons, which decay hadronically
- → $W^{\pm}$ boson decays to a light lepton (e/μ) - seperate analyses
- → Analysed full CMS Run 1 dataset: 5 fb^{-1} at 7 TeV (2011) and 19.7 fb^{-1} at 8 TeV(2012)

This analysis forms part of wider VH and inclusive $H \rightarrow \tau\tau$ searches at CMS



Final state signature:

- → Two high-p_T τ-leptons
- → One high-p_T light lepton (e/μ)
- → Missing transverse energy from escaping ν

$W^{\pm}H \rightarrow \tau_h\tau_h\ell^{\pm}$ analysis

Analysis strategy and selection

- → Signal selection designed to ensure high trigger efficiency and reject 'fake' signals
- → Reducible backgrounds estimated with a data driven method
- → Irreducible background prediction from events generated via Monte Carlo simulation
- → Statistical interpretation of results using the visible di-tau mass shape

Taus

- p_T > 25(20) GeV, $|\eta| < 2.1$
- Isolated, close to PV
- Tight μ, loose e rejection

Leptons

- p_T >24 GeV, $|\eta| < 2.1$
- Isolated, close to PV
- Tight ID - μ channel
- Loose ID - e channel

Additional requirements

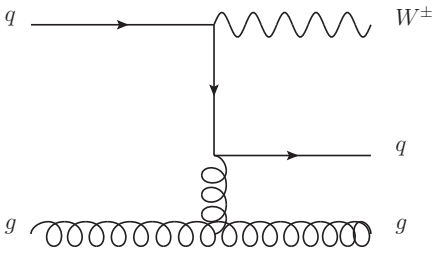
- Single μ trigger / $e + \tau$ trigger
- Objects separated in ΔR[1]
- Opposite sign τ-lepton pair
- Tight b-jet veto
- Veto events with additional ℓ (p_T > 15 GeV)

[1] $\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2}$

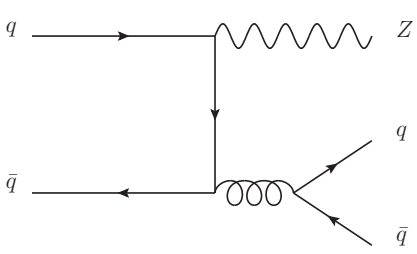
Backgrounds

Wide variety of background processes to consider

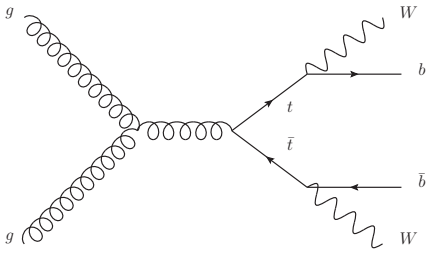
The majority contain jets misidentified as τ-leptons:



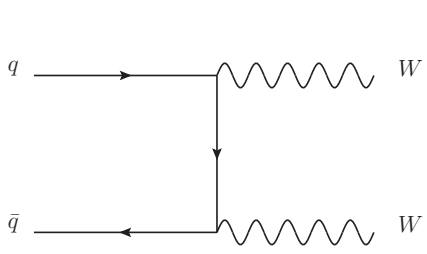
$W+$ jets



$Z+$ jets



$t\bar{t}$



WW

- **Reducible** backgrounds all contain at least one jet misidentified as a τ
- τ same sign as light lepton is always fake

- **Irreducible** backgrounds are:
  - → WZ: $W \to \ell\nu$, $Z \to \tau\tau$ (very hard to distinguish from WH signal)
  - → $ZZ \to \tau\tau\ell\ell$ (one lepton not reconstructed)

Multivariate discriminator against fake taus

After object selection the dominant background is W+jets (jets misidentified as τ-candidates)

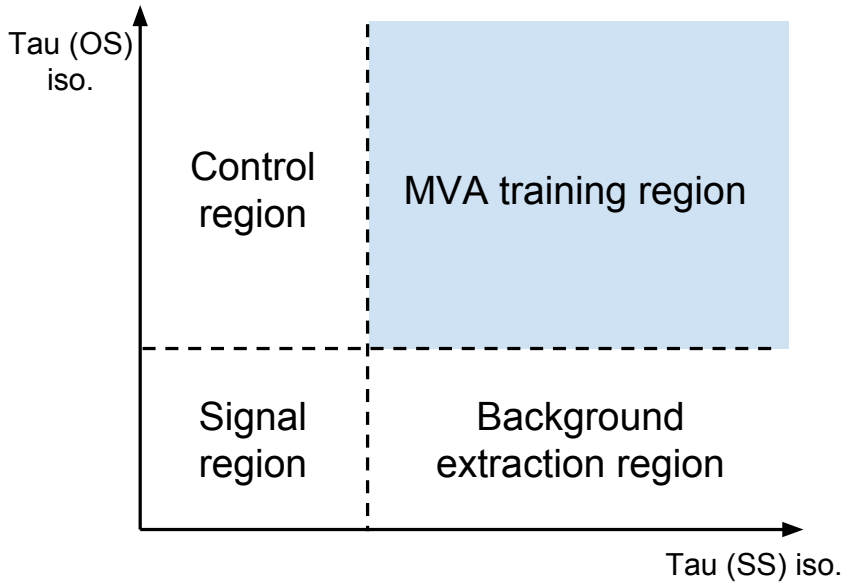
A Boosted Decision Tree is used to provide discrimination against these fakes.

Several kinematic variables are input to the BDT, which is trained in the region with all analysis cuts applied but τ-isolation inverted (enriched with fakes).

This region was found to be dominated by W+jets events in MC.

BDT inputs

- $\tau_1\ p_T$
- $\tau_2\ p_T$
- $\frac{(\tau_1\tau_2)p_T}{\tau_1 p_T + \tau_2 p_T}$
- $\Delta R(\tau_1, \tau_2)$
- E_T^{miss}



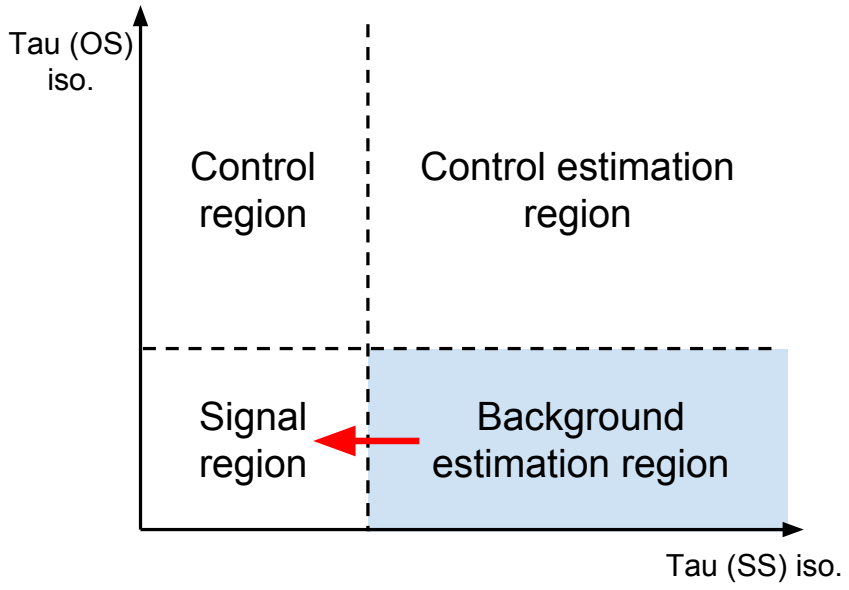
Estimation of reducible backgrounds

Reducible backgrounds were estimated in a data-driven way via a ***fake rate*** method.

The τ-candidate same sign as lepton is always fake in these cases.

We exploit this fact in our approach.

- → Fake Rate = P(isolated)/P(not isolated)
- → Measure this with data separately in W and Z enriched regions
- → Measure as a function of p_T and η
- → Apply using region where the τ same sign to lepton has isolation inverted
- → Procedure tested by estimating the content in the control region, with calculation from control estimation region



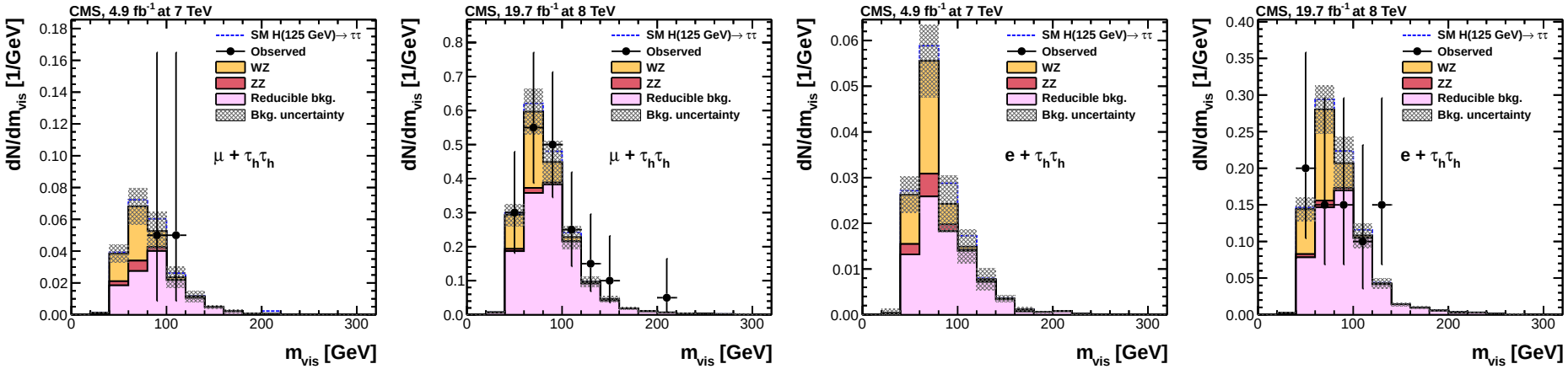
This fake rate used to predict the background content in the signal region

Results

Results consistent with background predictions

Slight downward fluctuation observed in total yield.

Event category	Signal	Background	Data	$\frac{S}{S+B}$
$\mu + \tau_h\tau_h$ 7 TeV	0.35 ± 0.03	4.1 ± 0.4	2	0.098
$\mu + \tau_h\tau_h$ 8 TeV	1.57 ± 0.12	35.2 ± 2.1	38	0.054
$e + \tau_h\tau_h$ 7 TeV	0.23 ± 0.02	2.7 ± 0.2	0	0.101
$e + \tau_h\tau_h$ 8 TeV	0.87 ± 0.08	16.5 ± 1.1	15	0.062



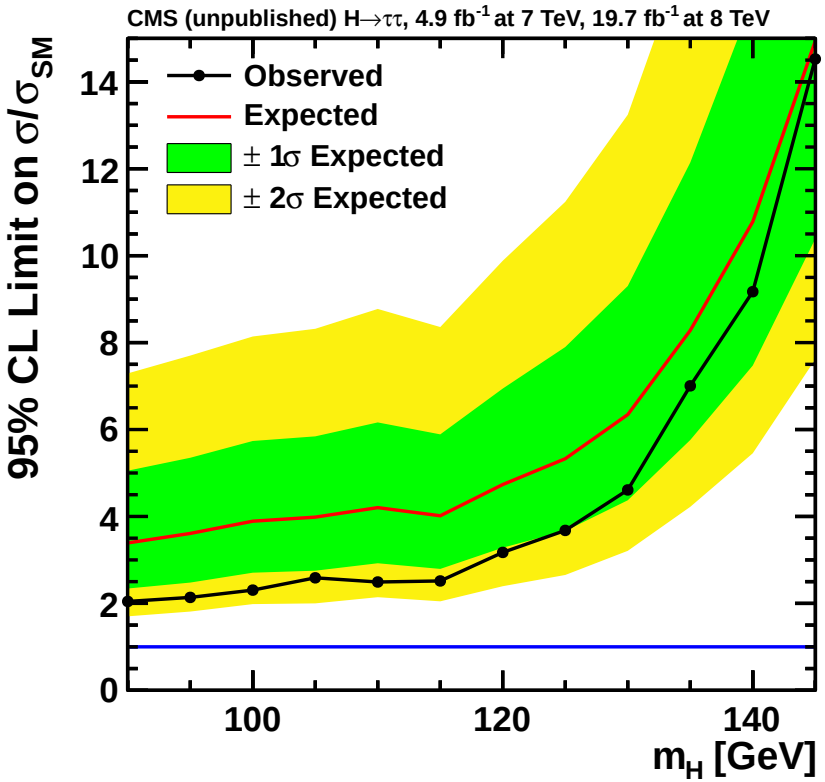
7 TeV: $\mu + \tau\tau$ 8 TeV: $\mu + \tau\tau$ 7 TeV: $e + \tau\tau$ 8 TeV: $e + \tau\tau$

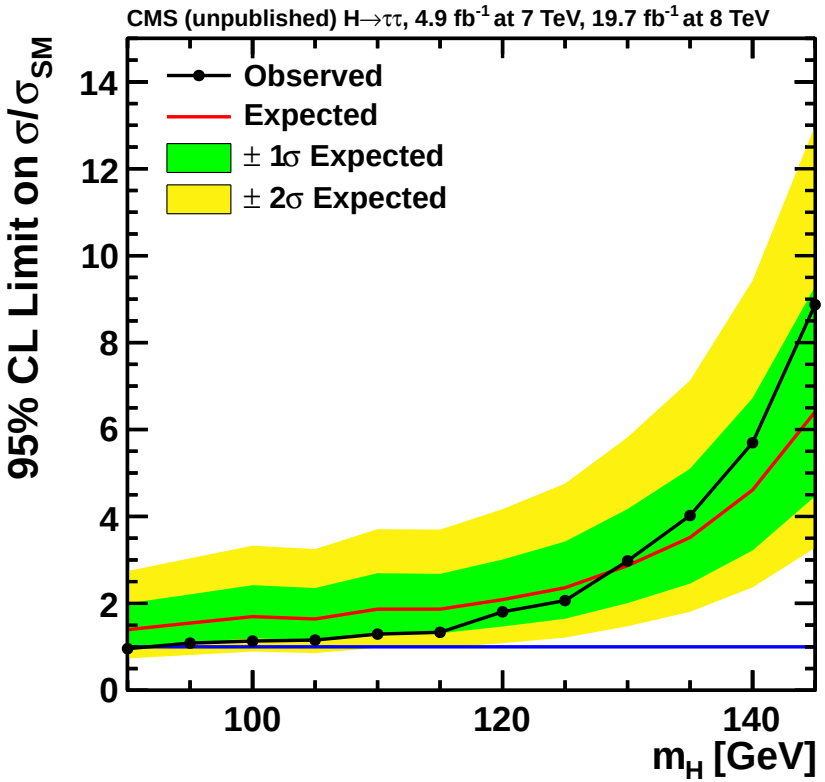
Limits

Final observed limit from this analysis of $\sim 3.5\sigma_{SM}$ at $m_H = 125$ GeV

This was included as part of a wider VH contribution to the inclusive $H \to \tau\tau$ search.
VH also included $WH \to \tau_h\tau_\ell\ell$, $ZH \to \tau_h\tau_h\ell\ell$, $ZH \to \tau_h\tau_\ell\ell\ell$, $ZH \to \tau_\ell\tau_\ell\ell\ell$

Final observed VH combined limit of $\sim 2\sigma_{SM}$ at $m_H = 125$ GeV



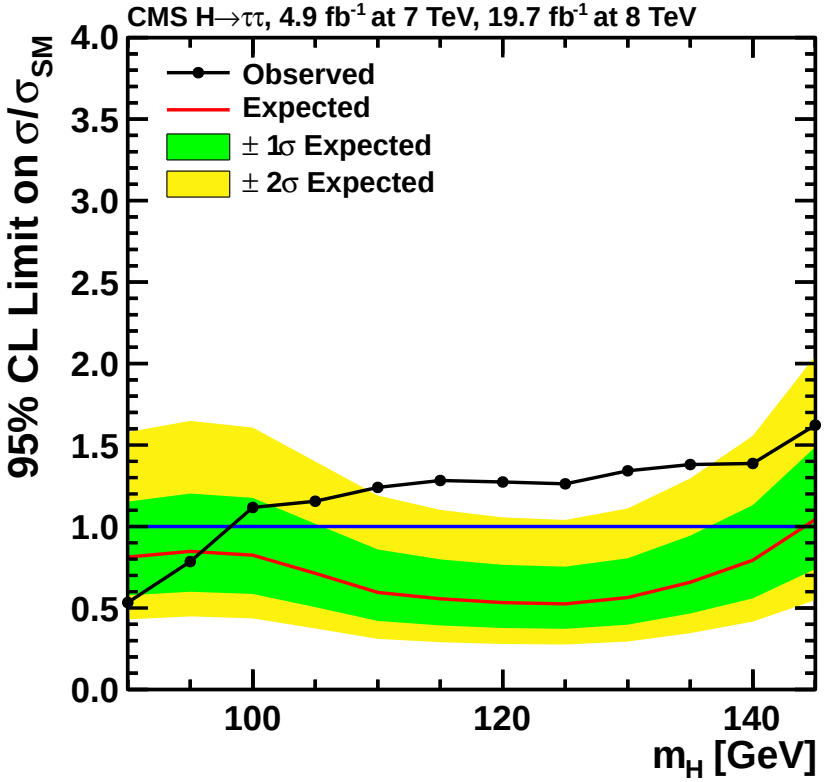


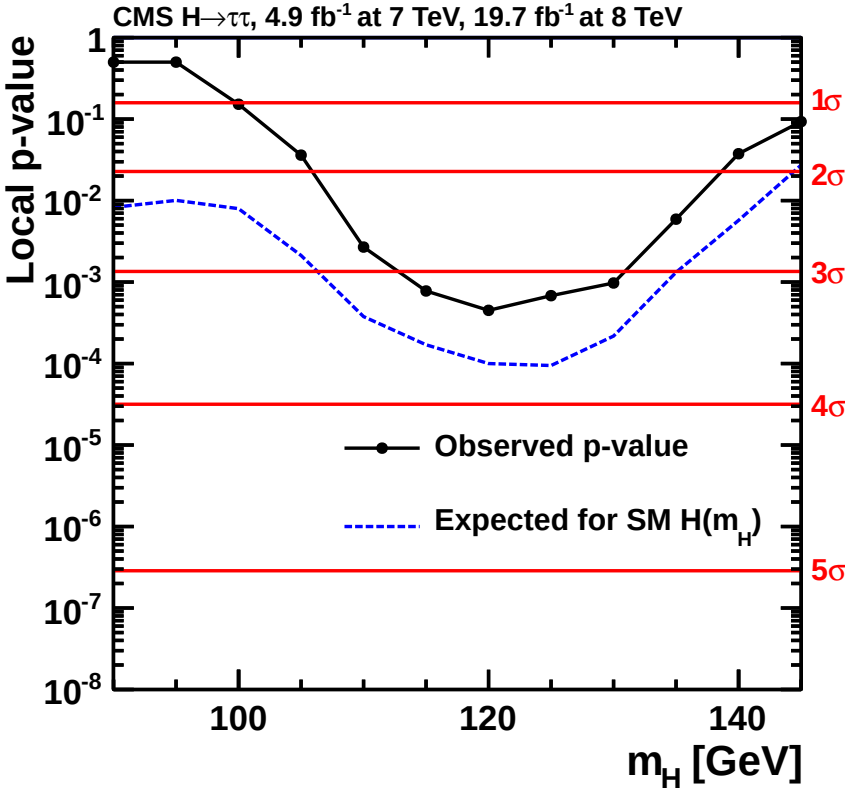
Inclusive $H \to \tau\tau$ limit

The VH analyses were included in the final inclusive $H \to \tau\tau$ search

Reported an excess $> 3\sigma$ at $m_H = 125$ GeV![2]

These observations constitute evidence for the Higgs boson decaying to a pair of τ-leptons

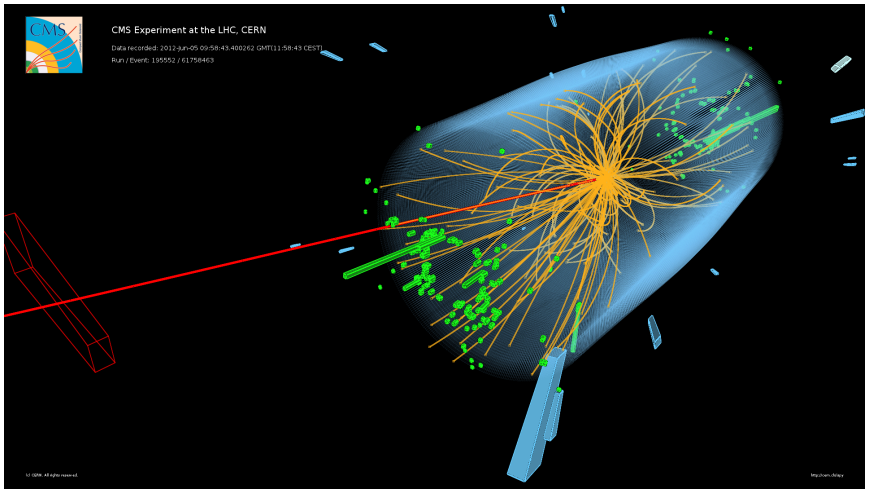




[2] arXiv:1401.5041

Conclusion

- Described the search at CMS for $WH \rightarrow \tau_h \tau_h \ell$
- Currently this channel doesn't have sensitivity by itself, but forms part of a wider VH and inclusive $H \rightarrow \tau\tau$ search
- Further details in arXiv:1401.5041: *Evidence for the 125 GeV Higgs boson decaying to a pair of τ leptons*
- Looking forward to the next iteration of this analysis in 2015, with more data and at a higher centre-of-mass energy!



Candidate $H \rightarrow \tau\tau$ event in $\mu\tau_h$ channel of inclusive analysis

Backup

Lepton selection

Channel	HLT requirement	Lepton selection		
$\mu\tau_h$	$\mu(12\text{–}18)$ & $\tau_h(10\text{–}20)$	$p_T^{\mu} > 17\text{–}20$ $p_T^{\tau_h} > 30$	$\lvert\eta^{\mu}\rvert < 2.1$ $\lvert\eta^{\tau_h}\rvert < 2.4$	$R^{\mu} < 0.1$ $I^{\tau_h} < 1.5$
$e\tau_h$	$e(15\text{–}22)$ & $\tau_h(15\text{–}20)$	$p_T^{e} > 20\text{–}24$ $p_T^{\tau_h} > 30$	$\lvert\eta^{e}\rvert < 2.1$ $\lvert\eta^{\tau_h}\rvert < 2.4$	$R^{e} < 0.1$ $I^{\tau_h} < 1.5$
$\tau_h\tau_h$ (2012 only)	$\tau_h(35)$ & $\tau_h(35)$ $\tau_h(30)$ & $\tau_h(30)$ & jet(30)	$p_T^{\tau_h} > 45$	$\lvert\eta^{\tau_h}\rvert < 2.1$	$I^{\tau_h} < 1$
$e\mu$	$e(17)$ & $\mu(8)$ $e(8)$ & $\mu(17)$	$p_T^{\ell_1} > 20$ $p_T^{\ell_2} > 10$	$\lvert\eta^{\mu}\rvert < 2.1$ $\lvert\eta^{e}\rvert < 2.3$	$R^{\ell} < 0.1\text{–}0.15$
$\mu\mu$	$\mu(17)$ & $\mu(8)$	$p_T^{\mu_1} > 20$ $p_T^{\mu_2} > 10$	$\lvert\eta^{\mu_1}\rvert < 2.1$ $\lvert\eta^{\mu_2}\rvert < 2.4$	$R^{\mu} < 0.1$
ee	$e(17)$ & $e(8)$	$p_T^{e_1} > 20$ $p_T^{e_2} > 10$	$\lvert\eta^{e}\rvert < 2.3$	$R^{e} < 0.1\text{–}0.15$
$\mu + \mu\tau_h$	$\mu(17)$ & $\mu(8)$	$p_T^{\mu_1} > 20$ $p_T^{\mu_2} > 10$ $p_T^{\tau_h} > 20$	$\lvert\eta^{\mu}\rvert < 2.4$ $\lvert\eta^{\tau_h}\rvert < 2.3$	$R^{\mu} < 0.1\text{–}0.2$ $I^{\tau_h} < 2$
$e + \mu\tau_h$/ $\mu + e\tau_h$	$e(17)$ & $\mu(8)$ $e(8)$ & $\mu(17)$	$p_T^{\ell_1} > 20$ $p_T^{\ell_2} > 10$ $p_T^{\tau_h} > 20$	$\lvert\eta^{e}\rvert < 2.5$ $\lvert\eta^{\mu}\rvert < 2.4$ $\lvert\eta^{\tau_h}\rvert < 2.3$	$R^{\ell} < 0.1\text{–}0.2$ $I^{\tau_h} < 2$
$\mu + \tau_h\tau_h$	$\mu(24)$	$p_T^{\mu} > 24$ $p_T^{\tau_{h,1}} > 25$ $p_T^{\tau_{h,2}} > 20$	$\lvert\eta^{\mu}\rvert < 2.1$ $\lvert\eta^{\tau_h}\rvert < 2.3$	$R^{\mu} < 0.1$ $I^{\tau_h} < 2\text{–}3$
$e + \tau_h\tau_h$	$e(20)$ & $\tau_h(20)$ $e(22)$ & $\tau_h(20)$	$p_T^{e} > 24$ $p_T^{\tau_{h,1}} > 25$ $p_T^{\tau_{h,2}} > 20$	$\lvert\eta^{e}\rvert < 2.1$ $\lvert\eta^{\tau_h}\rvert < 2.3$	$R^{e} < 0.1\text{–}0.15$ $I^{\tau_h} < 2$

Lepton selection for the LL' and $\ell + \ell\tau_h$ channels. The HLT requirement is defined by a combination of trigger objects with p_T over a given threshold. The p_T and I^{τ_h} thresholds are given in GeV. The indices 1 and 2 denote, respectively, the leptons with the highest and next-to-highest p_T.

Systematics

Uncertainty	Affected processes	Change in acceptance
Tau energy scale	signal & sim. backgrounds	1–29%
Tau ID (& trigger)	signal & sim. backgrounds	6–19%
e misidentified as τ_h	$Z \to ee$	20–74%
μ misidentified as τ_h	$Z \to \mu\mu$	30%
Jet misidentified as τ_h	$Z + \text{jets}$	20–80%
Electron ID & trigger	signal & sim. backgrounds	2–6%
Muon ID & trigger	signal & sim. backgrounds	2–4%
Electron energy scale	signal & sim. backgrounds	up to 13%
Jet energy scale	signal & sim. backgrounds	up to 20%
E_T^{miss} scale	signal & sim. backgrounds	1–12%
$\varepsilon_{\text{b-tag}}$ b jets	signal & sim. backgrounds	up to 8%
$\varepsilon_{\text{b-tag}}$ light-flavoured jets	signal & sim. backgrounds	1–3%
Norm. Z production	Z	3%
$Z \to \tau\tau$ category	$Z \to \tau\tau$	2–14%
Norm. $W + \text{jets}$	$W + \text{jets}$	10–100%
Norm. $t\bar{t}$	$t\bar{t}$	8–35%
Norm. diboson	diboson	6–45%
Norm. QCD multijet	QCD multijet	6–70%
Shape QCD multijet	QCD multijet	shape only
Norm. reducible background	Reducible bkg.	15–30%
Shape reducible background	Reducible bkg.	shape only
Luminosity 7 TeV (8 TeV)	signal & sim. backgrounds	2.2% (2.6%)
PDF (qq)	signal & sim. backgrounds	4–5%
PDF (gg)	signal & sim. backgrounds	10%
Norm. ZZ/WZ	ZZ/WZ	4–8%
Norm. $t\bar{t} + Z$	$t\bar{t} + Z$	50%
Scale variation	signal	3–41%
Underlying event & parton shower	signal	2–10%
Limited number of events	all	shape only

Systematic uncertainties, affected samples, and change in acceptance resulting from a variation of the nuisance parameter equivalent to one standard deviation. Several systematic uncertainties are treated as (partially) correlated for different decay channels and/or categories.

Event yields

Event category	Signal	Background	Data	$\frac{S}{S+B}$
$\ell\ell + LL'$				
$\mu\mu + \mu\tau_h$ 7 TeV	0.111 ± 0.005	2.4 ± 0.3	2	0.103
$\mu\mu + \mu\tau_h$ 8 TeV	0.427 ± 0.021	10.5 ± 0.6	12	0.092
$ee + \mu\tau_h$ 7 TeV	0.087 ± 0.004	1.5 ± 0.1	2	0.135
$ee + \mu\tau_h$ 8 TeV	0.385 ± 0.018	7.6 ± 0.4	11	0.149
$\mu\mu + e\tau_h$ 7 TeV	0.078 ± 0.004	2.2 ± 0.1	1	0.092
$\mu\mu + e\tau_h$ 8 TeV	0.293 ± 0.014	12.2 ± 0.6	8	0.081
$ee + e\tau_h$ 7 TeV	0.075 ± 0.004	2.2 ± 0.1	4	0.077
$ee + e\tau_h$ 8 TeV	0.279 ± 0.013	10.2 ± 0.5	13	0.063
$\mu\mu + \tau_h\tau_h$ 7 TeV	0.073 ± 0.006	0.8 ± 0.1	0	0.195
$\mu\mu + \tau_h\tau_h$ 8 TeV	0.285 ± 0.022	5.8 ± 0.4	4	0.150
$ee + \tau_h\tau_h$ 7 TeV	0.061 ± 0.004	1.1 ± 0.1	1	0.127
$ee + \tau_h\tau_h$ 8 TeV	0.260 ± 0.020	4.8 ± 0.4	9	0.148
$\mu\mu + e\mu$ 7 TeV	0.051 ± 0.002	1.0 ± 0.1	3	0.100
$\mu\mu + e\mu$ 8 TeV	0.202 ± 0.008	5.1 ± 0.3	9	0.105
$ee + e\mu$ 7 TeV	0.045 ± 0.002	1.0 ± 0.0	1	0.077
$ee + e\mu$ 8 TeV	0.185 ± 0.007	4.0 ± 0.2	4	0.082
$\ell + \tau_h\tau_h$				
$\mu + \tau_h\tau_h$ 7 TeV	0.35 ± 0.03	4.1 ± 0.4	2	0.098
$\mu + \tau_h\tau_h$ 8 TeV	1.57 ± 0.12	35.2 ± 2.1	38	0.054
$e + \tau_h\tau_h$ 7 TeV	0.23 ± 0.02	2.7 ± 0.2	0	0.101
$e + \tau_h\tau_h$ 8 TeV	0.87 ± 0.08	16.5 ± 1.1	15	0.062
$\ell + \ell'\tau_h$				
$\mu + \mu\tau_h$ 7 TeV	0.33 ± 0.02	3.2 ± 0.4	2	0.090
$\mu + \mu\tau_h$ low L_T 8 TeV	0.72 ± 0.03	20.7 ± 2.2	19	0.046
$\mu + \mu\tau_h$ high L_T 8 TeV	0.72 ± 0.02	8.4 ± 1.3	7	0.102
$e + \mu\tau_h/\mu + e\tau_h$ 7 TeV	0.47 ± 0.03	6.2 ± 1.0	6	0.074
$e + \mu\tau_h/\mu + e\tau_h$ low L_T 8 TeV	0.92 ± 0.03	24.6 ± 3.2	30	0.041
$e + \mu\tau_h/\mu + e\tau_h$ high L_T 8 TeV	1.15 ± 0.04	13.9 ± 2.0	11	0.109

Observed and predicted event yields in the $\ell\ell + LL'$ and $\ell + L\tau_h$ channels. The event yields of the predicted background distributions correspond to the result of the global fit. The signal yields are normalized to the SM prediction. Only SM Higgs boson production ($m_H = 125$ GeV) in association with a W or Z boson is considered as a signal process.