Search for the Higgs Boson Decaying into Pairs of Tau Leptons Further Decaying into Same-Flavour Leptons

Karlsruhe Institute of Technology

The Role of $H \to \tau\tau$ in Higgs Boson Searches

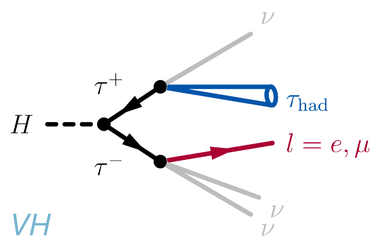
- **July 4th, 2012**: Observation of Higgs-like boson with mass near 125 GeV
- **March 2013**: First indications of coupling to tau leptons
- Tau channel important for measurement of the **Yukawa coupling of the Higgs boson to fermions**
- **Up to four neutrinos in the final state** complicate the mass reconstruction and therefore also the separability from SM background processes
- **Here: Emphasis on same-flavour lepton sub-channel $H \to \tau\tau \to \mu\mu$**

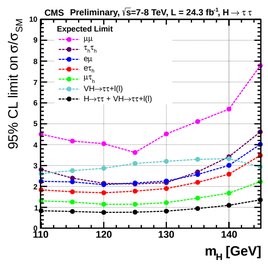
The $H \to \tau\tau$ Analysis in General

- The analysis presented here coresponds to data taken at the CMS experiment at the LHC at center of mass energies of 7 resp. 8 TeV with an integrated luminosity of 24.3 fb^{-1}

Decay Topology and Channels

- Semi-leptonic: $\mu\tau_h$, $e\tau_h$
- Fully leptonic: $e\mu$, $\mu\mu$
- Fully hadronic: $\tau_h\tau_h$
- *Missing: ee (aim to include soon)*
- Associated production: *VH*





- (Expected) sensitivities of individual channels depend strongly on
  - branching ratios
  - background contributions

Background Processes

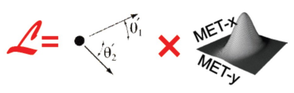
- **Drell-Yan** (most important: $Z \to \tau\tau$)
- Top-pair production
- Diboson production
- QCD
- W + jets

Reference: CMS-PAS-HIG-13-004

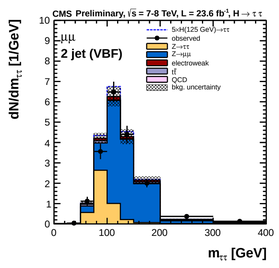
Mass Reconstruction

Reconstructed Mass of the Ditau System

- Reconstructed mass of the ditau system discriminates best between $H \to \tau\tau$ signal and the $Z \to \tau\tau$ background
- Mass hypothesis is given by a maximum likelihood method, where the likelihood function contains two parts:



  - Compatibility of the measured decay kinematics with the phase space information given by matrix element calculations
  - Compatibility of the measured missing transverse energy (MET) with the predicted kinematics of the neutrinos
- Algorithm yields broad distributions for events without genuine MET, e.g. $Z \to \mu\mu$

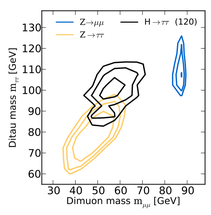


- ➜ Example of the final mass distribution in the VBF category

Mass of the Visible Decay Products

- Neglect contributions from invisible decay products
- ➜ Underestimation of the mass for $\tau\tau$ events
- ➜ Yields narrow mass peak for $Z \to \mu\mu$ events and therefore discriminates strongly between events with prompt leptons and $\tau\tau$ events

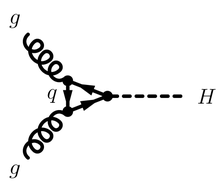
Combined Separation Power



- **Ditau mass** separates better between signal and $Z \to \tau\tau$
- **Dimuon mass** separates better between signal and $Z \to \mu\mu$

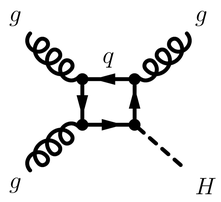
Higgs Production – Event Categories

0 Jets – low and high p_T



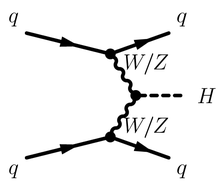
- **Gluon fusion** is the most dominant Higgs production mode at the LHC
- Large background
- ➜ Used only for constraining uncertainties

1 Jet – low and high p_T



- Jet recoils from boosted Higgs
- ➜ Better missing transverse energy (MET) resolution
- ➜ Harder p_T spectrum in Higgs events compared to Drell-Yan

Vector Boson Fusion (VBF)



- Two forward jets and low hadronic activity in barrel region
- ➜ Clear distinction from SM backgrounds

Peculiarities of the $H \to \tau\tau \to \mu\mu$ Channel

Challenges

- Small branching ratio: $BR(\tau\tau \to \mu\mu) \approx 3\%$
- Four neutrinos in final state reduce mass resolution
- **Additional overwhelming irreducible $Z \to \mu\mu$ background**: about 95 % after preselection
- ➜ **Two main backgrounds to account for**:
  - $Z \to \mu\mu$ as the largest irreducible background
  - $Z \to \tau\tau \to \mu\mu$ as the background whose detector signature hardly differs from the one of the Higgs signal

Analysis Strategy

1. **Event categorisation**
   - ➜ Account for different production processes
2. **MVA based selection of signal like events**
   - ➜ Suppress $Z \to \mu\mu$ background
3. **Background estimationon**
   - ➜ Data-driven as far as possible
4. **Statistical inference based on 2D Likelihoods**
   - ➜ Acoount for two main DY backgrounds based on the visible mass and the reconstructed ditau mass

Background Estimation

$Z \to \mu\mu$ — DCA Template Fits

- Data-driven estimation by correcting the MC based on template fits of the distance of closest approach (DCA) of the two muons
  - The DCA variable is only weakly to other BDT input variables
  - Fits are performed in bins of the two mass variables and a BDT discriminator excluding the DCA variable as input variable
  - ➜ Both the shape and the normalisation are corrected to fit the data

$Z \to \tau\tau$ — Embedding

- Shape is taken from embedded data sample
  - Muons in selected $Z \to \mu\mu$ data events are replaced by simulated tau leptons
  - ➜ Underlying event and pile-up remain from data
- Normalisation is corrected by the expectation in the full simulation

QCD — Same-sign Charge Data Sample

- Shape taken from data sample where lepton pairs with same-sign charge are selected
- Normalisation extracted from data samples where the isolation criteria have been inverted

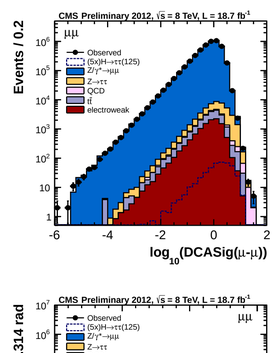
Other Backgrounds

- Shape and normalisation taken from Monte Carlo simulations
- Perfromance controlled in sideband regions

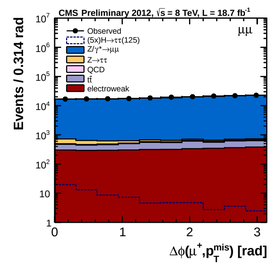
MVA Selection of Signal-like Events

Discriminating Variables as Inputs for BDTs

- Dilepton kinematic variables
- Variables evaluating the origin of the MET
- Two important examples:



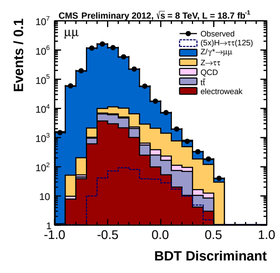
- Significance of the DCA of the leptons
- ➜ Prompt leptons in $Z \to \mu\mu$ originate from the same vertex
- ➜ Leptons from τ decays originate from two different secondary vertices



- Azimuthal angle between one lepton and the MET
- ➜ Flight directions of leptons and genuine MET in $\tau\tau$ events are correlated with each others
- Additional variables describing the two forward jets in the VBF category (mass and distance in the pseudorapidity)

BDT Discriminators and Selection

- **Boosted Decision Trees** (BDTs) are exploited to further suppress the contamination from Drell-Yan processes
- Trainings are performed for all Higgs mass hypotheses (110 to 145 GeV) at once
- Trainings are performed in two categories
  - 0 and 1 jet category inclusively
  - VBF category separately (with additional variables)
- BDT outputs discriminate strongly against the main $Z \to \mu\mu$ background

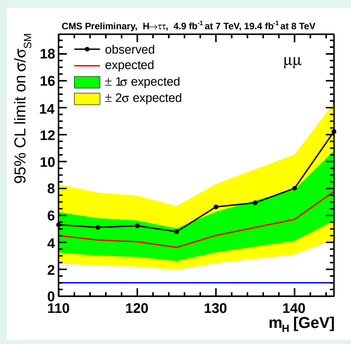


- Signal-like events are selected after cuts on the discriminators
- Cut thresholds are optimised for the significance $\frac{S}{\sqrt{S+B}}$ in the selected sample
- Optimisation in each event category separately
- The $Z \to \mu\mu$ remains the largest background after all selection steps

Conclusion

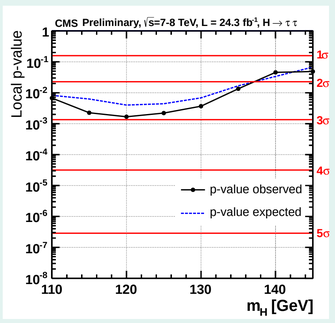
Statistical Inference Based on 2D Likelihoods

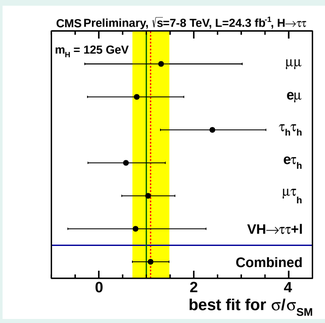
- Limits on the higgs production cross section are calculated based on 2D distributions of the visible and the reconstructed ditau mass



- **Good performance, given the challenges of this channel, is achieved**
  - ➜ MVA Selection
  - ➜ Mass reconstruction
  - ➜ Background estimation

Results for all $H \to \tau\tau$ Channels Combined





- Observe excess over broad mass range
- Max. local significance: 2.94 σ at $m_H = 120$ GeV
- Early measurements indicate compatibility with SM Higgs boson ($m_H \approx 125$ GeV)

KIT – University of the State of Baden-Württemberg and National Laboratory of the Helmholtz Association

Thomas Müller (KIT) on behalf of the CMS Collaboration

EPS HEP 2013, Stockholm

www.kit.edu