THE MATRIX ELEMENT METHOD USED IN THE SEARCH FOR THE ASSOCIATED PRODUCTION OF THE HIGGS BOSON WITH TOP QUARKS AND DECAYING INTO TAU LEPTONS

ALEXANDRE ZABI AND ARTUR LOBANOV
LLR ECOLE POLYTECHNIQUE CNRS-IN2P3 (FR)
On behalf of the CMS Collaboration

The ttH final state

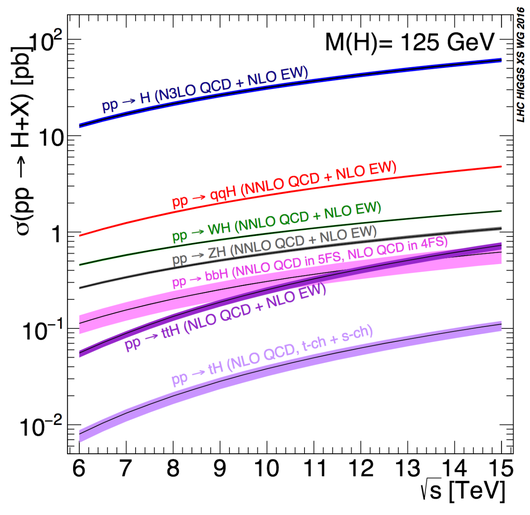
Motivation & Challenges

Motivations:
- **The ttH production mode is a direct probe to coupling of Higgs to top**
- **Increase LHC center-of-mass energy from 8 to 13 TeV, → x4 cross-section 0.51pb @ 13 TeV**
- **The ttH production has recently been observed** Phys. Rev. Lett. 120, 231801 (2018)

Challenges:
The study of the ttH production mode is challenging
→ *Looking for 2 tops + 1 Higgs* (w/ complex final states)

- ***Top decays:***
  t→blv : 1 b-jet + 1 lepton + MET
  t→bqq : 1 b-jet + 2 light jets

- ***Higgs decay:***
  H→bb : 2 b-jets
  H→$\tau\tau$: 1-2 τ_h (+ lepton + MET)
  H→WW/ZZ : leptons (+ jets + MET)
  H→$\gamma\gamma$: 2 photons
  H→ZZ->4l: 4 leptons



Final states with tau leptons: most sensitive final state to ttH H→$\tau\tau$ + 2 SS leptons. Signature with 2 b-jets + 2 light jets + 2 leptons + 1τ_h → almost every SM particle. Signal extraction makes extensive use of MVA discriminants

The Matrix Element Method

A powerful signal extraction method

The Matrix element method (MEM): The MEM is an alternative signal extraction method to Boosted Decision Trees, Neural Networks etc. It combines:
→ Variables from reconstructed objects (y) (such as leptons, taus, jets, MET)
→ Theoretical description in terms of particles in phase-space point (x) (such as quarks, neutrinos, tau before decay etc).
This combination is used to define event weights, which relates to how compatible the event is with a given physics process (signal and background weights are computed:

Parton density functions

Hard-scattering matrix element: computed at LO with MadGraph, OpenLoops…

$$w_\Omega(\mathbf{y}) \propto \sum_p \int d\mathbf{x} dx_a dx_b \frac{f(x_a,Q) f(x_b,Q)}{x_a x_b s} \delta^2(x_a P_a + x_b P_b - \sum p_k) |\mathcal{M}_\Omega(\mathbf{x})|^2 W(\mathbf{y}||\mathbf{x})$$

Numerical integration: performed with VEGAS, Markov chain integrator…

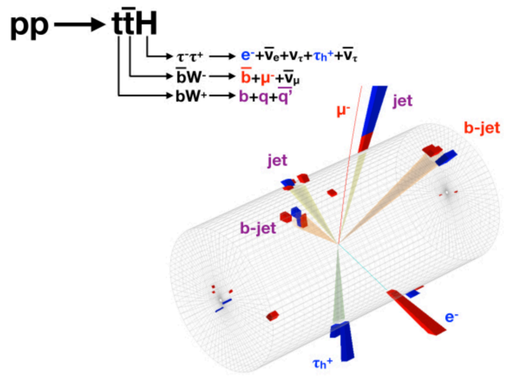
Momentum conservation (+ on-shell conditions): used to reduce dimensionality

Transfer functions: probability to measure y given a parton configuration x

The MEM: The event weight is computed for each hypothesis Ω (Ω=ttH, ttV…) using observables y as inputs and integrating over unmeasured (or poorly measured) quantities x. The final discriminant corresponds to:

$$\mathcal{L}(\mathbf{y}) = \frac{w_S(\mathbf{y})}{w_S(\mathbf{y}) + w_B(\mathbf{y})}$$

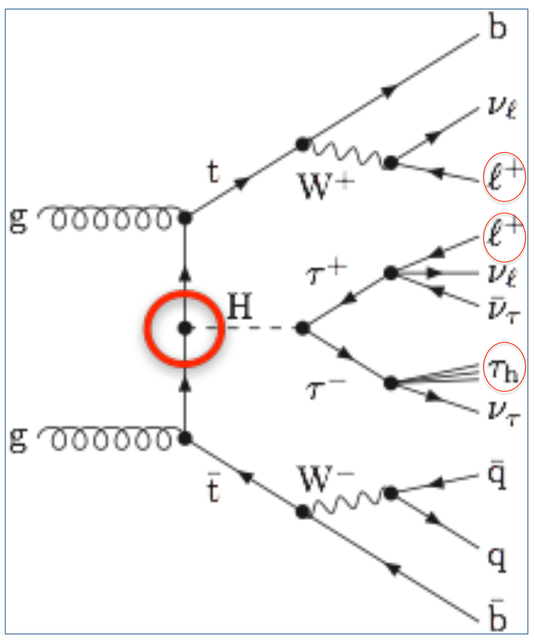
The use of the MEM is very CPU intensive and methods to implement a GPU-based derivation is actively ongoing @ LLR.



The ttH multilepton Analysis

Focus on the two same sign leptons and tau category

The categories with ttH → tau + X are part of the ttH multilepton analysis (CMS PAS HIG-17-018) including final states with more than or one hadronic tau lepton. Using here a dedicated tau identification with a smaller isolation cone size (ΔR=0.3) optimized for high multiplicity events

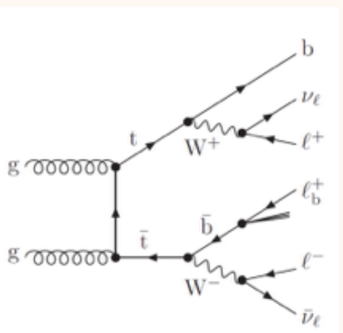


A good case for MEM:
2lSS+1τ_h + ≥ 3 jets (≥ 1 b-tag medium or ≥ 2 b-tag loose)

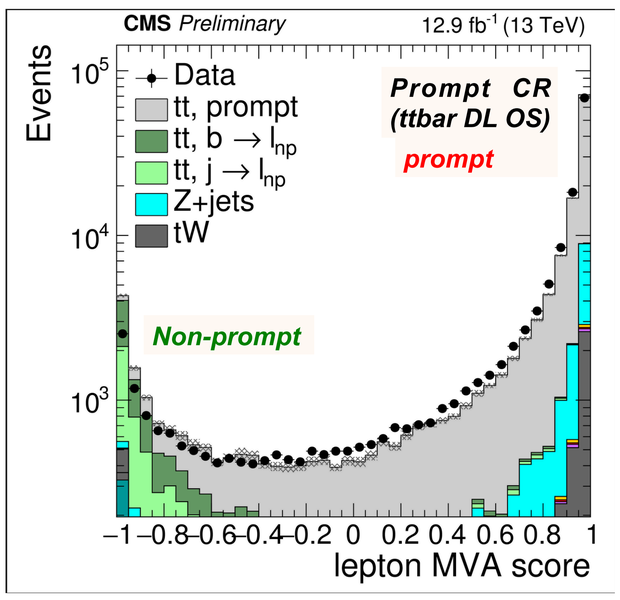
Main sources of background:
- ***Irreducible: ttV, di-boson (evaluated from MC)***
- ***Reducible: non-prompt leptons and charge mis-ID (from data)***

In the analysis presented, the leptons are selected with a MVA, trained to discriminate prompt leptons (from W, Z or τ decays) from non-prompt leptons (from b-jets and mis-identified hadrons)

Inputs to the MVA used:
- ***Isolation***
- ***vertex***
- ***Lepton Identification***
- ***jet variables (…)***



The performance of the MVA was validated in control regions directly using the CMS data.

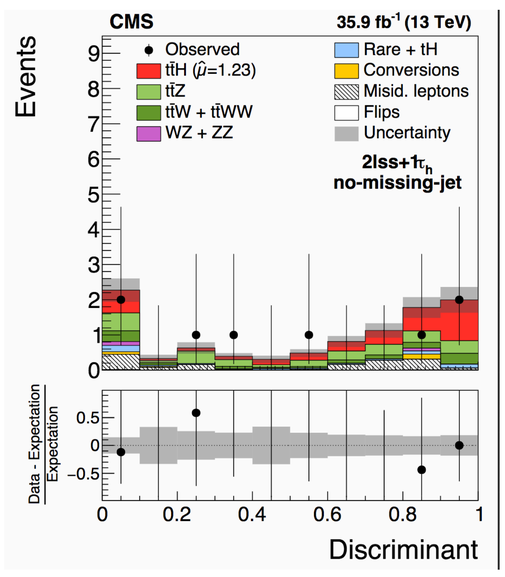


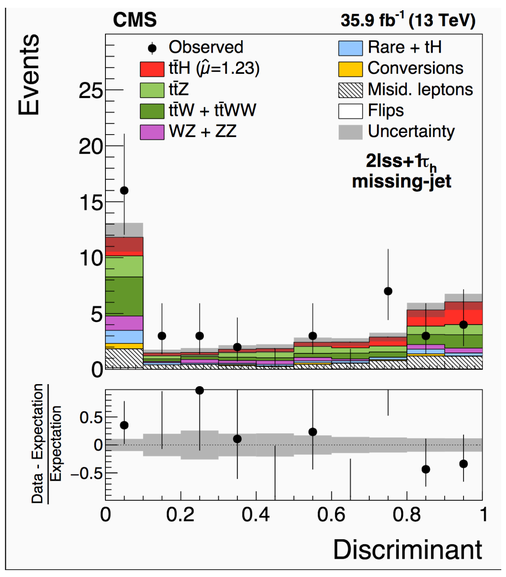
The residual background coming from non-prompt leptons was evaluated using tight-to-loose fake rate method

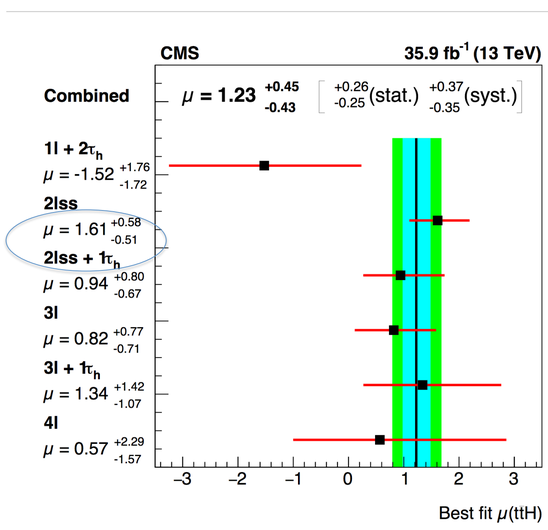
Analysis Results

Based on 2016 CMS data

The impact of the use of the MEM in the 2lSS + 1τ_h
- ***2lSS+1τ_h channel: MEM optimized to discriminate ttH H→$\tau\tau$ signal from ttZ → $\tau\tau$ + ttZ Z→ ll + tt w/ non-prompt lepton***
- ***The events are split between two subcategories based on presence of jets compatible with the W→ qq decay: MEM integration performed on direction of missing jet if needed***

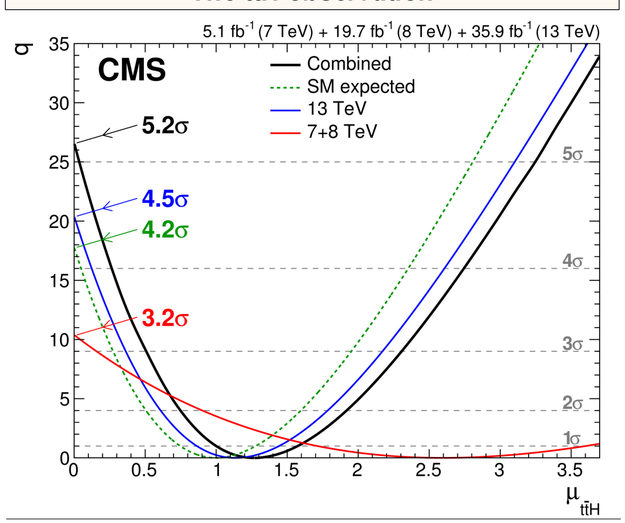






An excess has been observed with respect to the background-only hypothesis in the multilepton final states: 3.2σ observed significance (2.8σ expected). The combined results obtained are compatible with the Standard Model expectation.

The ttH observation



The 2lSS+1τ_h category used as input to the CMS ttH combination→ First observation of the ttH production mode!

CONCLUSION:
- ***In the search for the ttH production mode, the channels involving the presence of τ complement the coverage of the accessible phase space in CMS***
- ***Dedicated signal extraction methods optimized for final states with τ has been used: the MEM has proven to be a powerful tool.***
- ***Final combination with LHC Run I (8 TeV) has allowed to observe the ttH production mode with a significance of 5.2σ (4.2σ expected)***
  Phys. Rev. Lett. 120, 231801 (2018)