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[*][†]

Laboratoire Leprince Ringuet CNRS-IN2P3, Ecole polytechnique, Paris, France
E-mail: Alexandre.Zabi@cern.ch, Artur.Lobanov@cern.ch

Latest results of CMS searches for a Higgs boson produced in association with top quarks in final states with tau leptons are reported. This contribution will specifically focus on technical aspects related to the Matrix Element Method implementation and on its impact on the sensitivity of the analysis. The analysis presented here uses proton-proton collision data collected at center-of-mass energies of 13 TeV during the Run II of the LHC collected by the CMS experiment in 2016. An excess has been observed with respect to the background-only hypothesis in the multilepton final states: 3.2σ observed significance (2.8σ expected).

1. Introduction

The direct measurement of the Higgs coupling to the top quark can be achieved through the observation of the $t\bar{t}H$ production mode. Given the mass of the top quark, its coupling (y_t) to the Higgs boson is expected to approach unity. A precise measurement of y_t represents an opportunity to constrain physics beyond the Standard Model. The combination of the Run I and Run II (2016) datasets, representing a total of 60.7 fb^{-1} of data collected by CMS with a center-of-mass energies of 7, 8 and 13 TeV, allowed to achieve the first observation of the $t\bar{t}H$ signal with a significance of 5.2 standard deviations over the expectation from the background-only hypothesis [1]. In what follows, we focus on the $t\bar{t}H$ multilepton analysis [2] and the use of the Matrix Element Method (MEM) as signal extraction technique.

2. The $t\bar{t}H$ multilepton analysis

The search for the $t\bar{t}H$ production mode is particularly challenging: the cross section for this process is 0.51 pb^{-1} at $\sqrt{s} = 13$ TeV and a large range of possible final states are raising from the decay modes of the top quarks and the Higgs boson. We focus here on the multilepton channel where the Higgs decays into a tau lepton pair, one of the quark top decays into a semilepton-ically and the other one hardonically (noted 2lSS+1τ_h category). The final state contains 2 b quarks, 2 same sign leptons, one hardonically decaying tau lepton and a pair of light quarks as seen on Fig. 1. The main sources of backgrounds considered are the $t\bar{t}V$ (Z or W) and diboson production as well as reducible sources such as backgrounds with non-prompt leptons and charge misidentification.

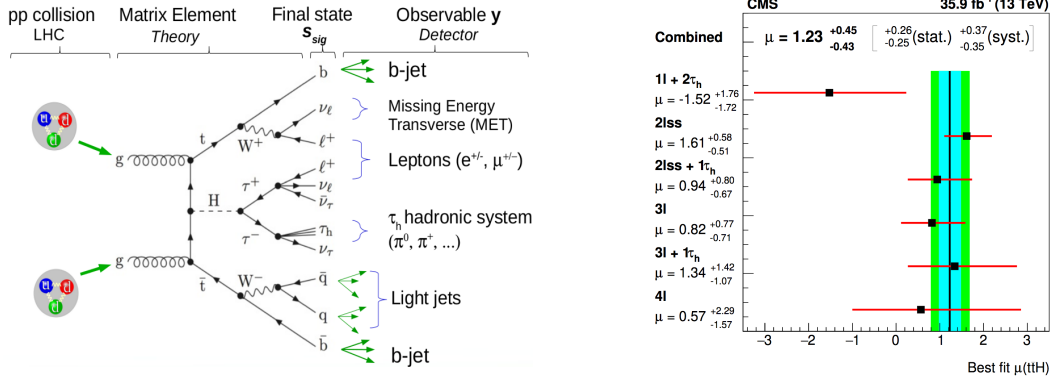[*]Speaker.

[†]on behalf of the CMS Collaboration

Figure 1: (Left) The final state considered in the analysis presented: $b\bar{b}$, $q\bar{q}$ pairs, hardonically decaying τ_h, 2 same sign leptons and 3 neutrinos $\nu's$. (Right) Signal rates μ, in units of the SM tt̄H production rate, measured in each of the categories individually and for the combination of all six categories [1]. The blue (green) band corresponds to the statistical (total) uncertainty on the combined signal rate.

3. The Matrix Element Method used on 2016 LHC Run II data

The Matrix Element Method (MEM) is a signal extraction method alternative to Boosted Decision Trees, Neural Networks, etc. For each event in a given data sample, an associated weight is computed (eq. 3.1) , which can be used to indicate how compatible this event is with a given physics process. Both signal and background weights are computed. The principle is to combine variables from the reconstructed objects (**y**) such as jets, leptons, hadronically decaying taus or missing transverse energy and a theoretical description of the event in terms of particles (quarks, neutrinos etc.) at a phase-space point (**x**).

$$w_i(y) = \frac{1}{\sigma_i} \sum_p \int \mathbf{dx} dx_a dx_b \frac{f(x_a,Q).f(x_b,Q)}{x_a x_b S} \delta^4(x_a P_a + x_b P_b - \Sigma P_k) |M_i(\mathbf{x})|^2 W(\mathbf{y}||\mathbf{x}) \quad (3.1)$$

The event weight $w_i(y)$ integrates all possible contributions of the matrix element $|M_i(\mathbf{x})|^2$ for a physics process i convoluted by the Transfer Function $W(\mathbf{y}||\mathbf{x})$ introducing the detector resolution effects on the measured final state described by **y** given the parton configuration **x**. The term $\frac{f(x_a,Q).f(x_b,Q)}{x_a x_b S}$ corresponds to the parton density functions and $\delta^4(x_a P_a + x_b P_b - \Sigma P_k)$ imposes momentum conservation. All the parton combinations are described by the Σ_p term and σ_i is a normalization factor tuned to achieve the best signal-to-background separation when combining the weights into a likelihood ratio. The computation of these weights are CPU intensive and a GPU-based derivation has been proposed successfully [3]. An excess has been observed with respect to the background-only hypothesis in the multilepton final states: 3.2σ observed significance (2.8σ expected) [2]. The combined results obtained are compatible with the Standard Model expectation. Fig. 1 shows the signal strength in all final states considered in this analysis. The 2lSS+1τ_h category was used as input to the CMS tt̄H combination leading to the first observation of this production mode. The results obtained in the 2016 tt̄H analysis demonstrated a successful implementation of the Matrix Element Method as signal extraction technique.

References

[1] *Observation of tt̄H Production* A. M. Sirunyan et al. (CMS Collaboration) Phys.Rev.Lett.120, 231801

[2] *Evidence for associated production of a Higgs boson with a top quark pair in final states with e, µ, and τ_h at $\sqrt{s}$ = 13 TeV*, A. M. Sirunyan et al. (CMS collaboration), JHEP 08 (2018) 066

[3] *Deployment of a Matrix Element Method code for the ttH channel analysis on GPU's platform*, G. Grasseau et al. (LLR-Ecole Polytechnique) Proceedings of the CHEP 2018 Conference (submitted)