PROCEEDINGS OF SCIENCE

The search for top-antitop-Higgs associated production at CMS

Fabrizio Margaroli[*][†]
Rome Sapienza University & INFN Roma 1
E-mail: fabrizio.margaroli@roma1.infn.it

The top quark and the Higgs boson are the two most massive elementary – and are they really so? – particles discovered so far. Understanding the relation between the two particles is of fundamental importance in order to investigate the Yukawa structure of the Standard Model (SM) lagrangian, to better constrain the theory using current data, and to probe for new physic appearing indirectly in loops or directly in events with top quark and Higgs in the final state. The CMS collaboration performed several searches for associated top-antitop-Higgs production that probe a multitude of combinations of both the top-antitop system final state, and of Higgs boson decays. The analyses are quickly approaching the sensitivity to the predicted Standard Model cross section for this extremely interesting process.

The European Physical Society Conference on High Energy Physics -EPS-HEP2013
18-24 July 2013
Stockholm, Sweden

---

[*] Speaker.
[†] On behalf of the CMS collaboration.

1. Introduction

The discovery in 2012 of a new heavy boson [1, 2] at the Large Hadron Collider represents an historical milestone in our understanding of nature. To examine whether this particle truly plays the role of the electroweak symmetry breaking (ESB) mechanism agent, it will be crucial to study its coupling to all known particles. A coupling of special interest is the one of the new boson to the top quark. In fact, due to the top quark very large mass [3] the top quark plays a special role in the ESB, pointing to the need for new physics. The Higgs boson has been discovered mainly through its direct coupling with the other known heavy bosons (W/Z) and so far only indirectly with fermions through loops. As a matter of fact, the Yukawa structure of the coupling of the Higgs to fermions is largely unexplored: as of today, we only have mild suggestions of direct Higgs coupling to b quarks [4] and to τ leptons [5]. On the other hand, a multitude of new physics scenarios could be hiding in those particle loops. Studying the direct coupling of Higgs to fermions through the associated production of top quarks together with the new heavy boson in yet unexplored production and decay channels utilizing the CMS experiment at the Large Hadron Collider (LHC) could lead to possible deviations from the standard model predictions in the top-Higgs interaction, as foreseen by natural new physics scenarios. In fact, the SM theory appears unnaturaly fine-tuned. This unnatural fine-tuning would be removed by the existence of exotic partners of the top quark that could be either of fermionic nature, such as those predicted by Composite Higgs and Little Higgs, or of bosonic nature, such as the supersymmetric top partner. In particular, both Composite/little Higgs and SUSY would predict final states that would largely overlap with the ones needed for studying top and Higgs production in the Standard Model, i.e. one or more top quarks, in addition to one or more Higgs bosons [6, 7].

The search for $\mathrm{t\bar{t}H}$ production is experimentally very challenging for a multitude of reasons. First, the theoretical NNLO prediction for this process at 8 TeV collisions amounts to only 130 fb [8]. Approximately one for every 10^9 LHC collisions gives rise to a Higgs boson; however, only about one every 200 events the Higgs boson is produced in association with top quarks. Also, both top quarks and Higgs bosons are extremely short-lived. Each top quark decays approximately 99.9% of the times in a W boson and a b quark. The W boson decays approximately 67% of the times into a quark-antiquark pair, and the rest into a charged-neutral lepton pair. For these searches, it is assumed that the Higgs boson with and decay rates are the ones predicted by the SM. It is thus expected to detect two-to-four particles coming from the Higgs decay, and six particles coming from the top-antitop quark system decay, leading to some of the busiest events under study in high energy physics. In order to reach sensitivity to such a small signal, the search for $\mathrm{t\bar{t}H}$ production mandates a careful choice of only a handful of the allowed signatures. Backgrounds are generally very large as the signature is dominated by the presence of the top-antitop pair, and the cross section for $t\bar{t}+X$ production is three orders of magnitude larger, approximately 230 pb [9].

The search for $\mathrm{t\bar{t}H}$ production at the Tevatron collider has been performed only in the $\mathrm{H} \to \mathrm{b\bar{b}}$ channel [10, 11] to maximize signal yields. Approximately 3000 $\mathrm{t\bar{t}H}$ events would have been produced by the LHC collisions during the 7 and 8 TeV center-of-mass energy runs. A portion of LHC data has been already analyzed looking for top-antitop-Higgs, Higgs to bottom quarks by ATLAS [12] and CMS [13] experiments. The larger available dataset enables the exploration of a broader range of final states.

2. General features

The CMS detector is described elsewhere [14]. Simulated samples for the SM Higgs boson signal and for background processes are used to optimize the event selection and to evaluate the acceptance and systematic uncertainties. The $t\bar{t}H$ signal is modeled with PYTHIA while most backgrounds have been generated with MADGRAPH combined with PYTHIA for the parton shower and hadronization; single top quark production has been generated with Powheg. All events are processed through a detailed simulation of the CMS detector based on GEANT4 and are reconstructed with the same algorithms that are used in data. The simulations include pileup interactions, the multiplicity of which matches the distribution observed with data. All events from data and simulated samples are required to pass the same trigger conditions. The identification of charged leptons, missing transverse energy, and jets proceeds as usual for CMS analyses. The identification of b quarks is very important due to the presence of two b jets in a any $t\bar{t}H$ final state; the b-tagging algorithm combines both secondary vertex information and track impact parameter information in a likelihood discriminant. The data sample used in this analysis was collected with single lepton triggers, dilepton triggers and diphoton triggers. The data correspond to a total integrated luminosity of 19.5 fb^{-1} for 8 TeV collisions. The single and dilepton triggered data from 5.0 fb^{-1} collisions have been used as well. Throughout this paper “leptons" is intended to refer to electrons or muons only – taus are mentioned explicitly whenever used.

3. Higgs to b quarks and taus

The signal is characterized by a large number of particles in the final state. In particular, the top-antitop quark pair always provides at least two jets, with additional jets coming from either top or Higgs decays. The analysis of events with $H \to b\bar{b}$ decay proceeds as follows: first, the sample is split into events with exactly one high-p_T lepton – the semileptonic top-antitop decays – or two high-p_T leptons – the dileptonic top final state. Both samples are further split into multiple subsamples characterized by different jet multiplicities. The Higgs boson is reconstructed within combinatorial ambiguities whenever a sufficient number of objects are identified. The signal is discriminated from the dominant $t\bar{t}$+jets background by means of Boosted Decision Trees (BDT) that exploit the peculiar kinematical and topological characteristics of the top-antitop-Higgs production. Figure 1 shows as illustration the jet multiplicity distribution (left) for lepton+jet events, and a BDT discriminant for a particular jet multiplicity category (right). For the sample where the Higgs decays to taus, BDTs are used again, where the discrimination power comes from combining observables associated to the quality of the reconstructed taus. For all of the above channels, the sensitivity to the signal is estimated by a likelihood fit over the binned BDT distributions. No statistically significant excesses have been found at this stage in the region compatible with a Higgs boson, thus upper limit on the Higgs production cross sections have been set.

4. Higgs to photons

The excellent resolution of the CMS electromagnetic calorimeter facilitates the reconstruction of a narrow mass peak when the Higgs boson decays to two photons. In addition to the two high-p_T photons triggering the recording of the events, the presence of at least two jets is required. To

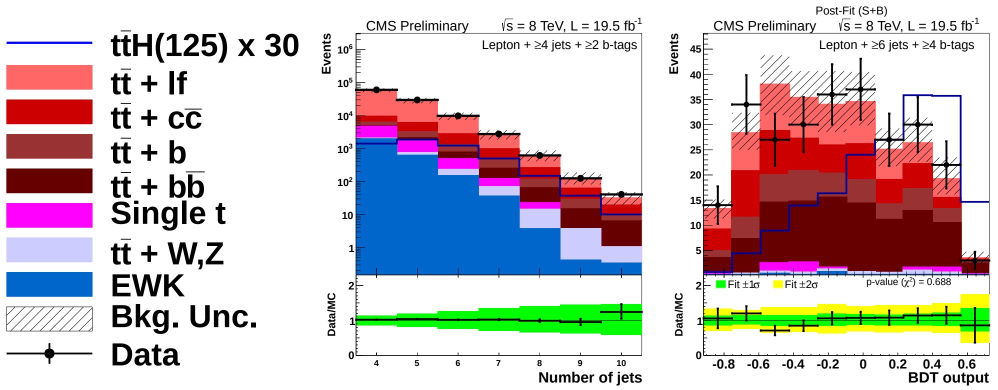
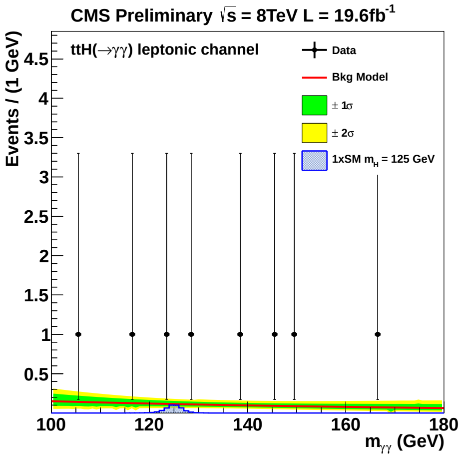


Figure 1: Left plot: jet multiplicity distribution for events with Higgs decaying to b jets and top pairs in the lepton+jets channel. The uncertainty band includes both statistical and systematic uncertainties. The $t\bar{t}H$ signal (m_H = 125 GeV) is normalized to $\sim 30\times$ SM expectation. Right plot shows the final BDT output for events with $\geq$6 jets and $\geq$4 b-tags.

increase sensitivity and maximize acceptance, all $t\bar{t}$ system decays are collected, split into events with at least five high-p_T jets (hadronic channel), and events with at least two high-p_T jets and one high-p_T lepton (leptonic channel). In both instances at least one jet is required to be b-tagged. All jets are required to be sufficiently energetic as to suppress the contamination of the dominant Higgs production process. The contribution of single top plus Higgs production [17] has not been estimated so far in this analysis but its cross section σ_{tHq} is only about 1/10 of $\sigma_{t\bar{t}H}$ and has different kinematics, thus its contribution is expected to be small. Given the complex background composition and the poorly known cross sections of the participating processes, its estimation is entirely data-driven. The sensitivity to the signal is estimated by a likelihood fit over the diphoton invariant mass distribution; the diphoton distributions for the hadronic and leptonic channels are shown in Fig. 2. No statistically significant excess has been found in the region compatible with a Higgs boson, thus upper limit on the Higgs production cross sections have been set.

5. Combination

The combination of all the above channels is accomplished using the same techniques employed in the global CMS Higgs combination [1]. Theoretical systematic uncertainties impacting more than one analysis in the combination are treated as fully correlated, with the exception of the uncertainties connected with $t\bar{t}+b\bar{b}$, $t\bar{t}+b$, and $t\bar{t}+c\bar{c}$ which are not treated as correlated because the uncertainties on those backgrounds are parameterized differently between the 7 TeV and 8 TeV analyses. Experimental uncertainties relevant to the $H\rightarrow b\bar{b}$, $H\rightarrow\tau^+\tau^-$, or $H\rightarrow\gamma\gamma$ analyses using the 8 TeV dataset are treated as fully correlated, but are not considered correlated with the corresponding uncertainties from the 7 TeV analysis. Figure 3, left sides, shows the expected and observed limit from the individual analyses and from their combination, for an assumed Higgs boson mass of 125 GeV. Combining these analyses improves the expected limit by 34% compared to the best individual result for a Higgs mass of 125 GeV. Figure 3, right side, shows the best fit signal strength μ for each channel contributing to the combination. For the combination of all $t\bar{t}H$



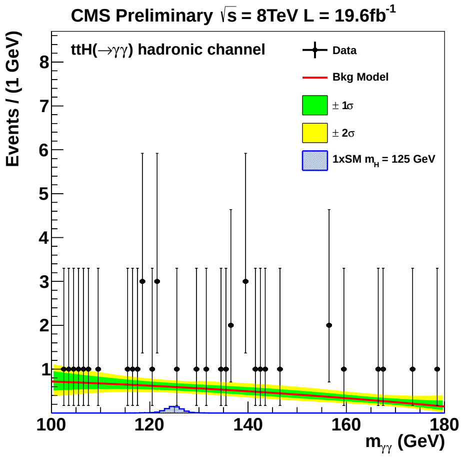


Figure 2: Diphoton distribution in events with hadronic (left) or leptonic (right) $\mathrm{t\bar{t}}$ decays. The red line shows the background model, surrounded by the 68% (green) and 95% (yellow) uncertainty bands.

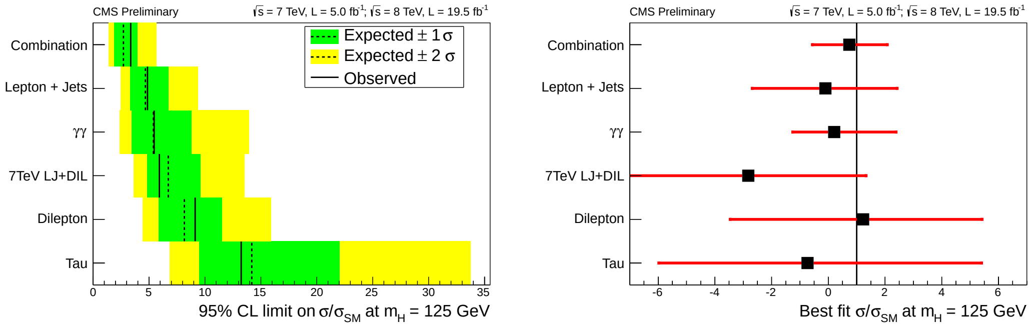


Figure 3: The observed and expected 95% CL upper limits (left) and best-fit values (right) of $\mu = \sigma/\sigma_{SM}$ for the lepton + jets (LJ), dilepton (DIL), $\tau\tau$ and $\gamma\gamma$ channels separately from the 2012 8 TeV dataset, the combination of the lepton + jets and dilepton channels from the 2011 7 TeV dataset, and the combination of all of the channels, for $m_H = 125\,\mathrm{GeV}$.

channels, the median expected limit for $\mathrm{m_H}$ = 125 GeV is $2.7 \times \sigma_{SM}$ while the observed limit is $3.4 \times \sigma_{SM}$, and the best-fit value for μ is 0.7 ± 1.3 (68% CL).

6. Conclusions

Several new searches for the standard model Higgs boson produced in association with a top-quark pair have been performed at the CMS experiment using the 2011 and 2012 data samples, corresponding to integrated luminosities respectively of 5.0 and 19.5 fb^{-1} at $\sqrt{s} = 7$ and 8 TeV collisions. A multitude of combination of the top-antitop system, and Higgs system, decays have been analyzed, and different strategies enacted. In particular, the searches have been optimized for the $\mathrm{H} \to \mathrm{b\bar{b}}$, $\mathrm{H} \to \tau^{+}\tau^{-}$, $\mathrm{H} \to \gamma\gamma$ decay modes. Combining the results from all the analyzed channels, the observed and expected limits on the cross section for Higgs boson production in association with top-quark pairs for a Higgs boson mass of 125 GeV are 3.4 and 2.7 times the standard model expectation, respectively. The best-fit value for the signal strength μ is 0.7 ± 1.3 (68% CL). Extrapolation of current CMS results obtained using 7 and 8TeV collisions to a few hundreds fb^{-1} of luminosity collected at higher center-of-mass energies of the coming 2015 run already outperform the estimations on the precision on the top-Higgs Yukawa coupling in the CMS technical design report [18], and the sensitivity achievable in several years of ILC running [19].

References

[1] S. Chatrchyan *et al.* [CMS Collaboration], Phys. Lett. B **716**, 30 (2012) arXiv:1207.7235

[2] G. Aad *et al.* [ATLAS Collaboration], Phys. Lett. B **716**, 1 (2012) arXiv:1207.7214

[3] Tevatron Electroweak Working Group for the CDF and D0 Collaborations, arXiv:1305.3929

[4] T. Aaltonen *et al.* [CDF and D0 Collaborations], Phys. Rev. Lett. **109**, 071804 (2012) arXiv:1207.6436

[5] S. Chatrchyan *et al.* “Search for the Standard-Model Higgs boson decaying to tau pairs in proton-proton collisions at sqrt(s) = 7 and 8 TeV" CMS-HIG-13-004

[6] A. De Simone, O. Matsedonskyi, R. Rattazzi and A. Wulzer, JHEP **1304**, 004 (2013) [arXiv:1211.5663 [hep-ph]].

[7] K. Howe and P. Saraswat, JHEP **1210**, 065 (2012) [arXiv:1208.1542 [hep-ph]].

[8] S. Dawson, C. Jackson, L. H. Orr, L. Reina and D. Wackeroth, Phys. Rev. D **68**, 034022 (2003) hep-ph/0305087

[9] M. Czakon, P. Fiedler and A. Mitov, Phys. Rev. Lett. **110**, 252004 (2013) arXiv:1303.6254

[10] A. Azatov *et al.* [D0 Collaboration], “Search for the Standard Model Higgs boson in the $t\bar{t}H \to t\bar{t}b\bar{b}$ channel", D0 Conference Note CONF-5739, (2008)

[11] T. Aaltonen *et al.* [CDF Collaboration], Phys. Rev. Lett. **109**, 181802 (2012) arXiv:1208.2662

[12] G. Aad *et al.* [ATLAS Collaboration], “Search for the Standard Model Higgs boson produced in association with top quarks in proton-proton collisions at Ãs = 7 TeV using the ATLAS detector,” ATLAS-CONF-2012-135

[13] S. Chatrchyan *et al.* [CMS Collaboration], JHEP **1305**, 145 (2013) arXiv:1303.0763

[14] S. Chatrchyan *et al.* [CMS Collaboration], JINST **3**, S08004 (2008).

[15] S. Chatrchyan *et al.* [CMS Collaboration], “Search for top-antitop-Higgs production in events where Higgs decays to photons at $\sqrt{s} = 8$ TeV collisions" CMS-HIG-13-015

[16] S. Chatrchyan *et al.* [CMS Collaboration], “Search for Higgs Boson Production in Association with a Top-Quark Pair and Decaying to Bottom Quarks or Tau Leptons" CMS-HIG-13-019

[17] S. Biswas, E. Gabrielli, F. Margaroli and B. Mele, JHEP **07**, 073 (2013) arXiv:1304.1822

[18] CMS physics TDR J. Phys. G: Nucl. Part. Phys. 34 995-1579

[19] R. Yonamine et al. Phys.Rev. D84 (2011) 014033