PROCEEDINGS OF SCIENCE

Search for $t\bar{t}H$ production in high-p_{T} regimes with the ATLAS detector

Silvia Biondi*
University "Alma Mater Studiorum" & INFN - Bologna, Italy
E-mail: silvia.biondi@cern.ch

The associated production of the Higgs boson with a pair of top/anti-top quarks ($t\bar{t}H$) is the only process providing the direct access to the measurement of the Yukawa coupling between the Higgs boson and the top quark. The presented results exploit the data collected during 2015 and 2016 by the ATLAS experiment during LHC collisions at a center-of-mass energy of 13 TeV. Multivariate analysis techniques are used in order to discriminate the signal from the very large backgrounds arising mainly from top-quark pair production. In addition, for the first time the analysis uses algorithms specifically designed to cope with the challenging reconstruction of hadronically decaying high-p_{T} bosons and top quarks. The combined best fit value of the signal strength is found to be $0.84^{+0.64}_{-0.61}$, while a value greater than 2.0 is excluded at 95% confidence level while the expected upper limit is $\mu < 1.2$ in the absence of a $t\bar{t}H$ signal.

Sixth Annual Conference on Large Hadron Collider Physics (LHCP2018)
4-9 June 2018
Bologna, Italy

*on behalf of the ATLAS Collaboration

https://pos.sissa.it/

PoS(LHCP2018)005

1. Introduction

In this document an analysis strategy, used for the first time, is presented to improve the high-p_{T} spectrum regimes in an already existing $t\bar{t}H$ production mode analysis. The results obtained with data collected during 2015 and 2016 by the ATLAS experiment, corresponding to an integrated luminosity of $36.1\,\mathrm{fb}^{-1}$, are presented as well.

The process considered in the analysis involves a top quark loop, which allow to directly measure the top-Higgs Yukawa coupling for the first time. Up to now only indirect constraints on this coupling have been possible, from the gluon-gluon fusion production, followed by $H \rightarrow \gamma\gamma$ decay. Moreover, the considered production mode has the highest cross section increase as a function of the energy with respect to the other production modes.

2. The boosted topology

At high energies, explored by LHC during the Run-2 phase, heavy particles are often produced with large transverse momentum (called boosted particles) and for this reason this case is called *boosted topology*. The decay products of such particles (mainly W, Z and H bosons and top quarks) have a large Lorentz boost which results in a collimation of the objects around the momentum direction of the boosted original particle in the detector rest frame.

Different techniques have been tested in order to explore the high-p_{T} region and to mantain a good efficiency in both the reconstruction and identification of the objects. For this reason, both large-R jets (tagging top quark and Higgs boson) and re-clustering techniques [1] have been tested for the boosted $t\bar{t}H$ analysis, trying to detect the most optimal reconstruction technique and to improve the significance of the analysis.

Many studies have been performed leading to the decision to use the re-clustering technique: this is an innovative way to reconstruct the jets iside the events and it has been introduced for the first time in this analysis. The re-clustering algorithms use the sequential recombination explained in detail in Ref. [2].

The clustering technique is the best choice since it allows to use different radius parameter R of the jets, without the limit of a given calibration provided by the performances groups. In fact, the technique introduces a new angular scale $r < R$, such that jets of radius r can be used to build large radius R jets. In this case, the fully calibrated small r jets can make the calibration of the re-clustered large R jets automatic. Consequently, no additional calibrations are required and a broader class of algorithms and jet radius parameters can be used in the analysis.

The application of this technique in the $t\bar{t}H$ analysis allows to reconstruct the Higgs boson candidate mass, as shown in Fig. 1, using anti-k_t small r jets (with $r = 0.4$), re-clustered into large-R jets with $R = 1.0$, $p_{\mathrm{T}} > 200$ GeV, $|\eta| < 2$ and $m > 50$ GeV.

3. Signal region selection and analysis strategy

The analysis described in this document studies the process in Fig. 2, in which the $t\bar{t}$ system decays either semi-leptonically (one top quark decaying leptonically and one decaying hadronically, called *single-lepton* channel) or leptonically (both top quarks decaying leptonically, called *dilepton* channel). In both cases the Higgs boson decays in two b quarks ($H \rightarrow b\bar{b}$).

PoS(LHCP2018)005

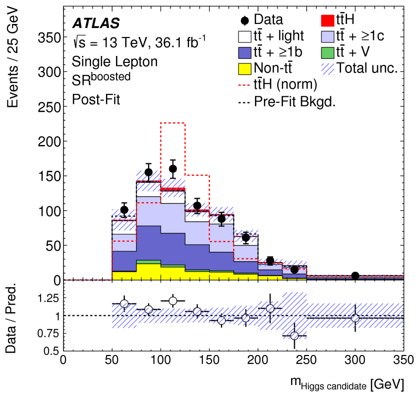

Figure 1: Higgs boson candidate mass distribution, reconstructed with the re-clustering technique, in the boosted $t\bar{t}H$ channel. Anti-k_t small R jets (with $R = 0.4$), re-clustered into large-R jets with $R = 1.0$, $p_T > 200$ GeV, $|\eta| < 2$ and $m > 50$ GeV have been used.

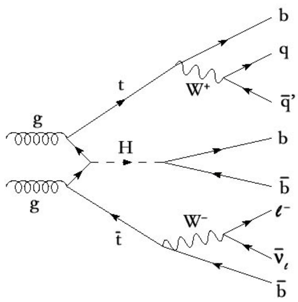

Figure 2: Feynman diagram of the $t\bar{t}H$ process studied in the analysis described in this document.

In the boosted regime of the analysis, which is in the semi-leptonic channel, the event selection requires:

- exactly one lepton (e or μ with $p_T > 27$ GeV);
- one Higgs candidate: one re-clustered jet ($R = 1.0$, $p_T > 200$ GeV) with two associated b-tagged jets;
- one top candidate: one re-clustered jet ($R = 1.0$, $p_T > 250$ GeV) with one associated b-tagged jet and one non-b-tagged jet;
- one b-tagged jet outside the two re-clustered jets ($\Delta R > 1.0$).

The analysis strategy requires the combination of the boosted regime with the two resolved channels (*single-lepton* and *dilepton*) that use the standard object identification and reconstruction algorithms. A BDT (Boosted Decision Tree) discriminant distribution, built using a MultiVariate Analysis technique (MVA [4]) and event kinematics and topology and b-tagging information, is introduced into a combined fit procedure, together with the BDT distributions of the resolved signal regions. A Likelihood function is used to determine the best value of the signal strength ($\mu_{t\bar{t}H} =$

$\frac{\sigma(t\bar{t}H)_{obs}}{\sigma(t\bar{t}H)_{SM}}$). The aim of the analysis is the estimation of the $t\bar{t}H$ signal strength μ and its 95% CL upper limit.

The BDT distribution for the boosted channel, before and after the fit procedure is shown in Fig. 3.

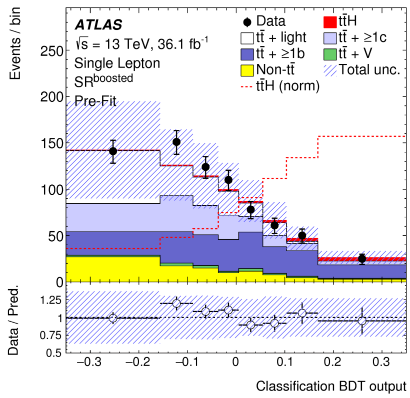

Figure 3: Data and MonteCarlo comparison for the BDT discriminant in the boosted signal region, before (left) and after (right) the combined dilepton and single-lepton fit to the data. [3]

4. Combined results

Combined results have been obtained with the 2015 and 2016 datasets, corresponding to an integrated luminosity of 36.1 fb^{-1} ($\sqrt{s} = 13$ TeV), with a boosted category included in the analysis for the very first time in the ATLAS experiment.

The best fit values of μ and its 95% CL limit are shown in Fig. 4 for each channel and for their combination.

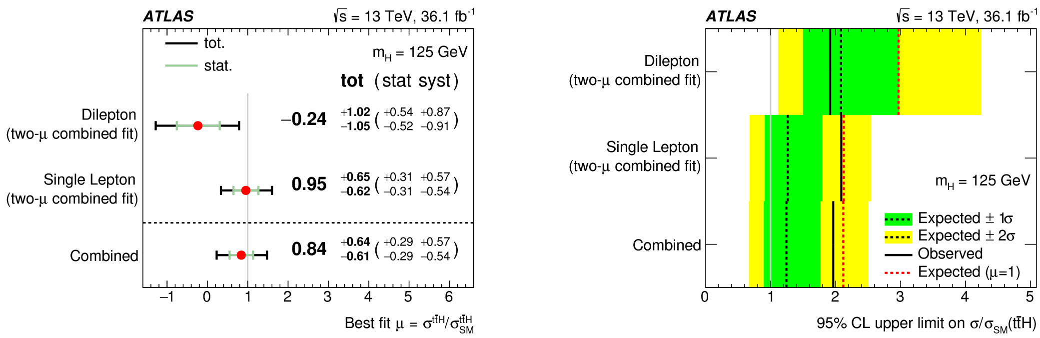

Figure 4: Signal strength measured (left) and 95% CL limits on the signal strength (right) for the single-lepton (resolved and boosted), dilepton and their combination. [3]

At the moment, the analysis is limited in its sensitivity by the systematic uncertainties. The main contribution comes from the main background, which is the $t\bar{t}$ production process in association with Heavy Flavor jets (jets coming from b or c quarks), as shown in the Table 1. Since the theoretical uncertainty related to this background is high, the normalisation factors of the two contributions ($t\bar{t}$+b-jets and $t\bar{t}$+c-jets) have been left freely floating in the fit. In fact, they are estimated from Monte Carlo simulations and constrained by dedicated control regions, in order to take into account both the experimental and theoretical uncertainties for the process. Other important contributions come from the signal modelling and the uncertainties related to the b-tagging algorithm and the Jet Energy Scale and Resolution (JES/JER).

Uncertainty source	$\Delta\mu$	
$t\bar{t}+\geq 1b$ modeling	+0.46	-0.46
Background-model stat. unc.	+0.29	-0.31
b-tagging efficiency and mis-tag rates	+0.16	-0.16
Jet energy scale and resolution	+0.14	-0.14
$t\bar{t}H$ modeling	+0.22	-0.05
$t\bar{t}+\geq 1c$ modeling	+0.09	-0.11
JVT, pileup modeling	+0.03	-0.05
Other background modeling	+0.08	-0.08
$t\bar{t}$+light modeling	+0.06	-0.03
Luminosity	+0.03	-0.02
Light lepton (e, μ) id., isolation, trigger	+0.03	-0.04
Total systematic uncertainty	+0.57	-0.54
$t\bar{t}+\geq 1b$ normalisation	+0.09	-0.10
$t\bar{t}+\geq 1c$ normalisation	+0.02	-0.03
Intrinsic statistical uncertainty	+0.21	-0.20
Total statistical uncertainty	+0.29	-0.29
Total uncertainty	+0.64	-0.61

Table 1: Breakdown of the contributions to the uncertainties in μ, all evaluated after the fit.

An improvement in the understanding of the main background and in the available statistic of the simulation datasets used in the analysis will be crucial for the future efforts and will certainly help in improving the sensitivity of this channel.

References

[1] B. Nachman, P. Nef, A. Schwartzman, M. Swiatlowski and C. Wanotayaroj, *Jets from Jets: Re-clustering as a tool for large radius jet reconstruction and grooming at the LHC*, JHEP 02 (2015) 075.

[2] S. Biondi, *Search for the $t\bar{t}H$ production in high-p_{T} regimes with the ATLAS detector*, ATL-PHYS-PROC-2017-145.

[3] M. Aaboud *et al.* (ATLAS Collaboration), *Search for the standard model Higgs boson produced in association with top quarks and decaying into a $b\bar{b}$ pair in pp collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector*, Phys. Rev. D 97, 072016.

[4] A. Hoeker *et al.*, *TMVA - Toolkit for Multivariate Data Analysis*, CERN-OPEN-2007-007.