Search for boosted $t\bar{t}H(H \rightarrow b\bar{b})$ with the ATLAS detector

Emma Winkels

26 March 2018

IOP Joint APP and HEPP annual meeting Bristol, UK

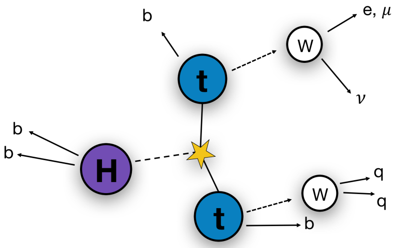
University of Sussex

ATLAS EXPERIMENT

Introduction

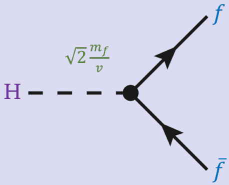
$t\bar{t}H$ production is a rare Higgs boson production mode at the LHC ($\sim 1\%$).

- Recent evidence for $t\bar{t}H$ with ATLAS using $36 fb^{-1}$ of data collected in 2015-2016 at $\sqrt{s} = 13$ TeV. This achievement was a combination of several results:
  - $H \to b\bar{b}$: arXiv 1712.08895
  - $H \to ZZ* \to 4l$: arXiv 1712.02304
  - $H \to \gamma\gamma$: arXiv 1802.04146
  - $H \to$ Multilepton and combination: arXiv 1712.08891
- Will discuss here the $H \to b\bar{b}$ analysis with focus on the boosted channel.
- First time that a boosted channel is included in this analysis.



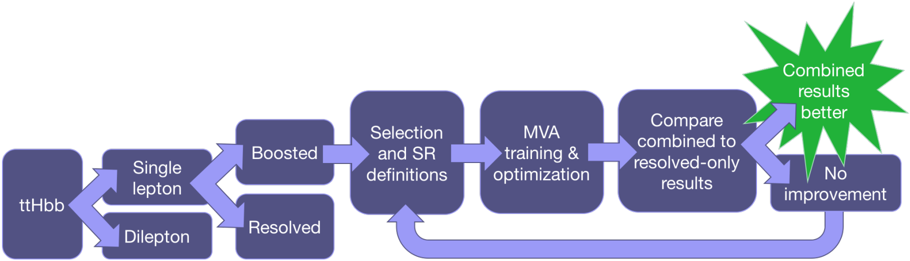
Motivation

All current Higgs measurements consistent with the SM. Yukawa coupling is proportional to fermion mass; heaviest fermion is the top. Any deviations in y_t would give indication for new physics.

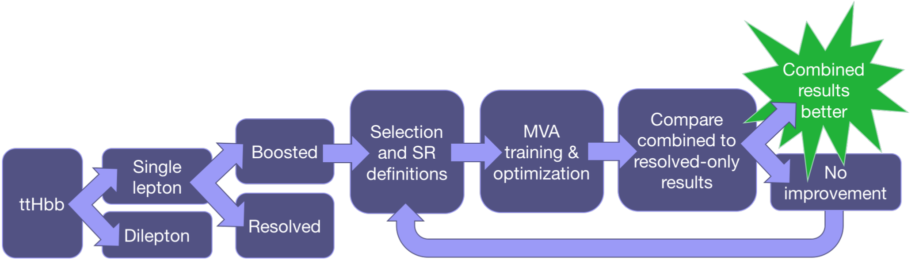


- $t\bar{t}H$ is direct probe of the top Yukawa coupling whereas e.g. ggF only accesses it via loop effects.
- Decay to $b\bar{b}$ has the largest branching ratio in the SM (58%)
- Boosted channel selects high-P_T events which leads to different kinematics and a simplified combinatorial background compared to the resolved channel.

Analysis overview

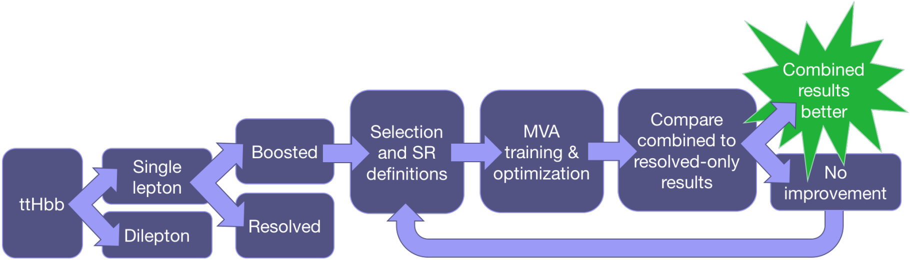


Analysis overview



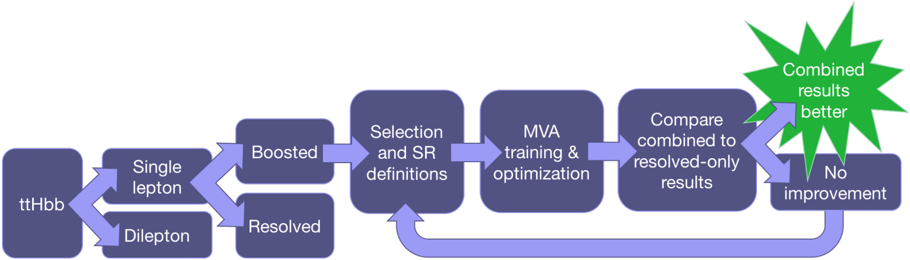
- The channels are split into regions based on the number of jets and b-jets (at different efficiencies) in the events. The boosted region gets priority over the resolved single-lepton regions.

Analysis overview



- The channels are split into regions based on the number of jets and *b*-jets (at different efficiencies) in the events. The boosted region gets priority over the resolved single-lepton regions.
- Multivariate analysis to discriminate between signal and background events

Analysis overview

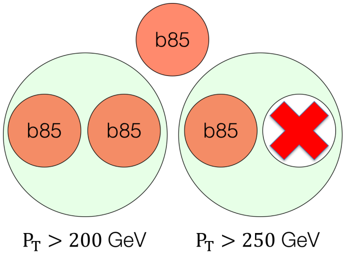


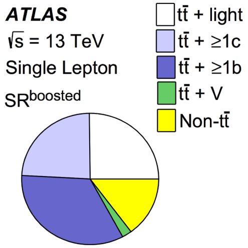
- The channels are split into regions based on the number of jets and *b*-jets (at different efficiencies) in the events. The boosted region gets priority over the resolved single-lepton regions.
- Multivariate analysis to discriminate between signal and background events
- The boosted analysis is combined with resolved in the final result.

Boosted signal region

Decay products of boosted Higgs and top are collimated such that we can capture them in a large jet. In this analysis, we recluster our standard anti-k_T jets with $R = 0.4$ into larger $R = 1.0$ jets. We require events to have:

- **Higgs candidate**: 1 large reclustered jet with $\geq$ 2 b-tagged subjets
- **Top candidate**: 1 large reclustered jet with 1 b-tagged subjet and $\geq$ 1 b-veto
- $\geq$**1 additional** ***b*****-jet** outside of the top and Higgs candidates

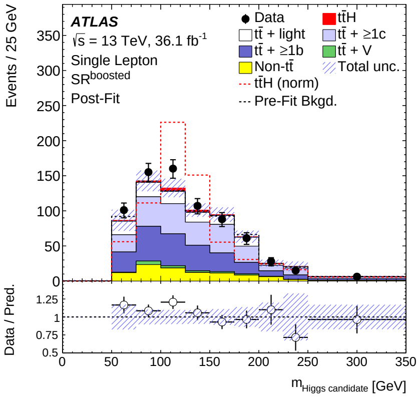




Multivariate Analysis

Boosted Decision Tree (BDT) to separate signal from background events. Eight variables are chosen as input:

- The Higgs candidate mass
- The first splitting scale of the Top candidate
- 4 angular variables between various objects in the event
- 2 pseudo-continuous *b*-tagging variables: the *b*-tagged jets are divided into 4 categories depending on their W.P. percentage. These categories are assigned a score which is used to construct input variables to the BDT.



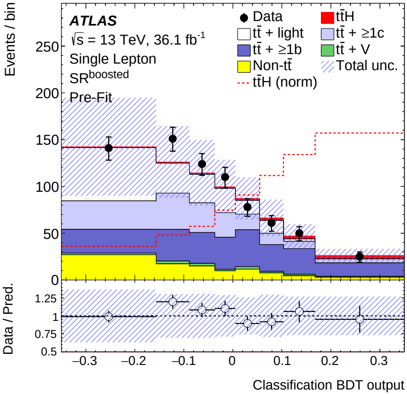
Statistical analysis

- The boosted region is fitted together with 8 resolved signal regions and 10 resolved control regions.
- Classification BDT used in each SR, in CRs we fit the total event yield except for two $t\bar{t}+\geq 1c$ dominated CRs where $H_T^{had} = \Sigma p_T^{jet}$ is used.
- Binned profile likelihood performed simultaneously across all regions.
- Normalizations of $t\bar{t}+\geq 1b$ and $t\bar{t}+\geq 1c$ are left free-floating to be constrained by the fit.
- Extract $\mu_{t\bar{t}H} = \sigma_{obs}^{t\bar{t}H}/\sigma_{SM}^{t\bar{t}H}$ correlated across all regions.

ewinkels@cern.ch

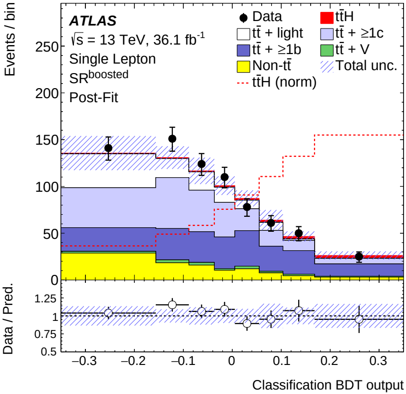
BDT behaviour in fit

Pre-fit



→

Post-fit

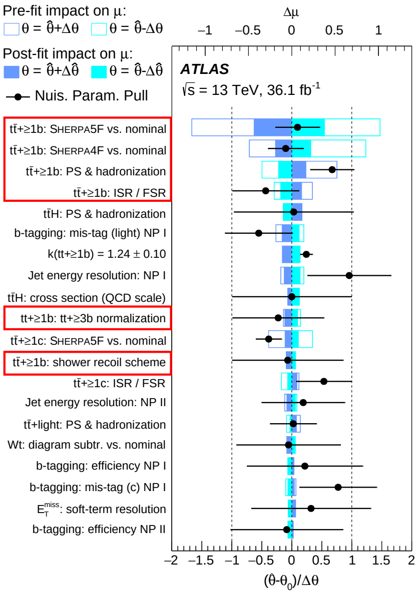


ewinkels@cern.ch

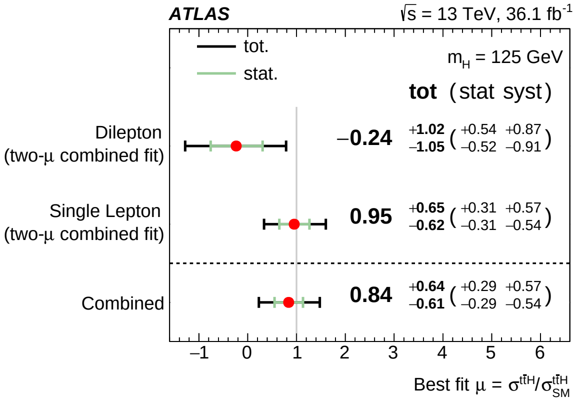
Systematics

The analysis is limited by systematics; main sources are:

- Modeling of $t\bar{t}+\geq 1b$: compare different MC generators to assess uncertainties. This uncertainty also dominates combined result.
- Background modeling MC statistics
- b-tagging uncertainties
- Jet uncertainties

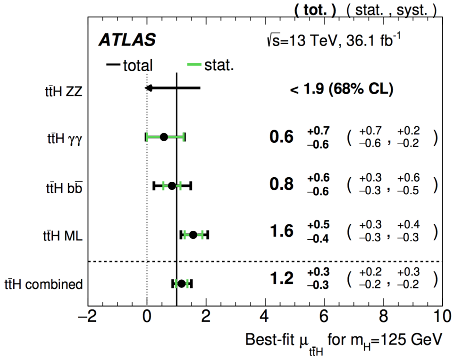


Results $t\bar{t}H(H \to b\bar{b})$



- The $t\bar{t}H(H \to b\bar{b})$ signal is seen with an obs (exp) significance of 1.4σ (1.6σ).
- Best-fit value found of $\mu_{t\bar{t}H} = 0.84 \pm 0.29(\text{stat})\ ^{+0.57}_{-0.54}(\text{syst})$.
- Excluded $\mu > 2.0$ at 95% CL.
- Currently the boosted SR does not add significant sensitivity to the channel.

Results from combination



Channel	Best-fit μ		Significance	
	Observed	Expected	Observed	Expected
Multilepton	$1.6\,^{+0.5}_{-0.4}$	$1.0\,^{+0.4}_{-0.4}$	4.1σ	2.8σ
$H \rightarrow b\bar{b}$	$0.8\,^{+0.6}_{-0.6}$	$1.0\,^{+0.6}_{-0.6}$	1.4σ	1.6σ
$H \rightarrow \gamma\gamma$	$0.6\,^{+0.7}_{-0.6}$	$1.0\,^{+0.8}_{-0.6}$	0.9σ	1.7σ
$H \rightarrow 4\ell$	< 1.9	$1.0\,^{+3.2}_{-1.0}$	—	0.6σ
Combined	$1.2\,^{+0.3}_{-0.3}$	$1.0\,^{+0.3}_{-0.3}$	4.2σ	3.8σ

Combination results in best fit value $\mu_{t\bar{t}H} = 1.17 \pm 0.19(\text{stat})^{+0.27}_{-0.23}(\text{syst})$ with an obs (exp) significance of 4.2σ (3.8σ).

Evidence for $t\bar{t}H$ production!

Conclusions

- Very challenging analysis with one of the main problems being the $t\bar{t}+ \geq 1b$ irreducible background and modeling thereof.
- Measured $\mu_{t\bar{t}H}$ consistent with both SM and background only hypotheses in the $b\bar{b}$ channel. The paper was accepted by PRD.
- The boosted channel is not currently adding significant sensitivity to the resolved analysis. Many studies are on-going to improve the performance for the full Run 2 analysis. More data and better MC statistics will also allow us to design tighter SR.
- Systematics limited which means we have to rethink our approach to dealing with these uncertainties.
- The four-channel combination in ATLAS with 2015+2016 data has found evidence for $t\bar{t}H$ production.