ATLAS results for $t\bar{t}H$ Multilepton with 80 fb^{-1}

Kirill Grevtsov on behalf of ATLAS $t\bar{t}H$-ML team

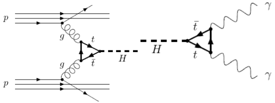
DESY

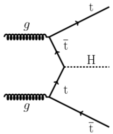
DESY LHC Discussions

October 21, 2019

Top quark Yukawa coupling

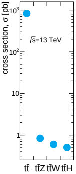
Top quark Yukawa coupling (y_t): Probe Electroweak Symmetry Breaking





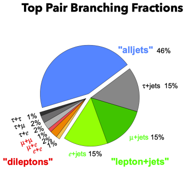
- <u>Indirect</u> measurements:
  - through loops in ggH and its decay to $\gamma\gamma$ (only SM in loop)
- <u>Direct</u> measurements:
  - $t\bar{t}H$ (tH) productions

If new physics scenarios exist, will see significant deviations from SM prediction

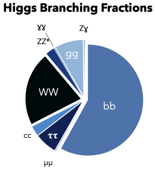


Challenges:

- Low production cross section
- Complex final states:
  - Combination of top quark x Higgs boson decay modes



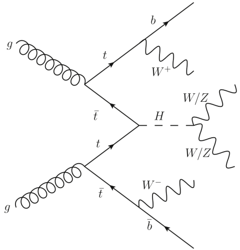
X

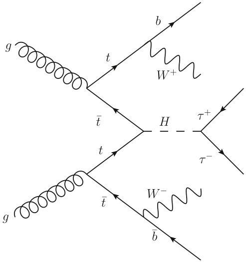


Outline

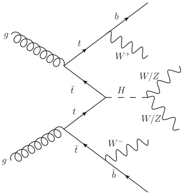
A search for $t\bar{t}H$ in multilepton final states with 80fb^{-1}
ATLAS-CONF-2019-045

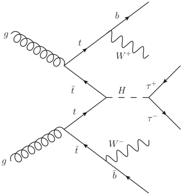
- Analysis strategy
- Object identification
- Event selections and classification
- Background estimates
- Systematic uncertainties
- Results
- Summary

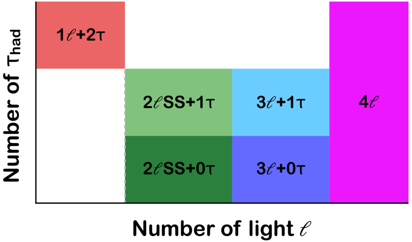




- Analysis targets:
  - Higgs decays to WW, ZZ and $\tau\tau$ (veto $H \to ZZ^* \to 4\ell$)
  - Semi- and dileptonic $t\bar{t}$ decays
- Six orthogonal analysis channels:
  - categorised by the number and flavour of lepton candidates



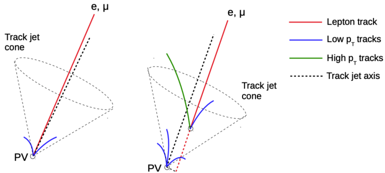




- Backgrounds:
  - Reducible: $t\bar{t}$, V+jets
    - Non-prompt light leptons
    - charge misID
    - electrons from photon conversion
  - Irreducible: $t\bar{t}V$, VV, rare top, etc
    V - vector boson, Z or W
- Signal extraction:
  - A maximum-likelihood fit over 25 event categories

Object identification

- High multiplicity final state:
  - Leptons
    - Calorimeter and track based identification
    - Specific lepton BDT isolation suppressing ℓ from semi-leptonic *b*-decays

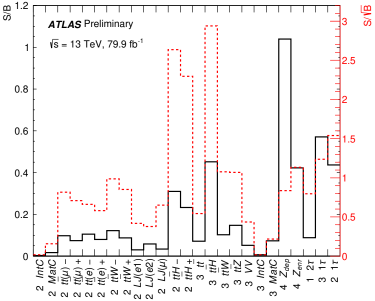


    - BDT to reject charge miss assignment
    - Three categories of selected electron candidates: very tight and two conversion types material conversion (rec.displaced CV) and internal conversion candidates those further suppressed with track invariant masses and conversion radius
  - Jets
    - anti-k$_t$ jets $R = 0.4$
    - Use jet vertex tagger to remove PU jets
    - *b*-tagged jets @70% WP
  - Select τ_{had}
    - 1- and 3-prong
    - multivariate analysis discriminants to reject jets and electrons faking a tau
  - Remove overlapping objects

Event selections and classification

- Six orthogonal channels
  - with further split to gain significance
  - In $2\ell SS0\tau_{had}$ and 3ℓ channels, control regions are defined for the bkg. determination
  - Total 25 categories

- Event level BDTs are employed to further separate signal from bkg[1]
  - $\mathbf{2\ell SS0\tau_{had}}$: a combination of 2 BTDs (vs $t\bar{t}W$, vs. fakes/$t\bar{t}$) in 2D space
  - $\mathbf{3\ell}$: a multi-dimensional BDT (vs. $t\bar{t}W$, $t\bar{t}Z$, VV, fakes/$t\bar{t}$)
    $\rightarrow$ cross-check analysis using **cut-and-count categorisation** were developed for these two most sensitive channels
  - 4ℓ: BDT (vs. $t\bar{t}Z$ and ZZ)
  - $1\ell 2\tau_{had}$: BDT (vs. $t\bar{t}$)

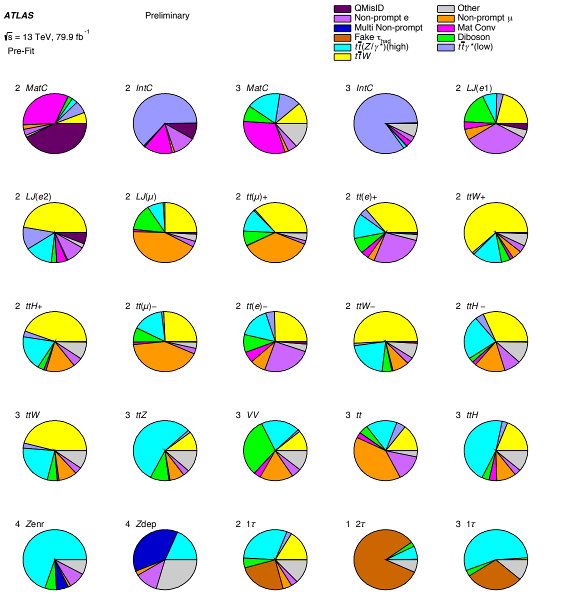


- In total 877 events are selected in data
- There are 170 SM $t\bar{t}H$ events **expected** after selection
  - 0.42% of all produced $t\bar{t}H$ events

[1] except $3\ell 1\tau_{had}$ and $2\ell SS1\tau_{had}$

Background composition

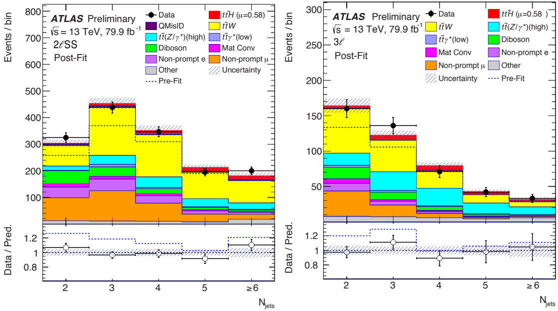
- Dominant backgrounds in channel $\rightarrow$
  - non-prompt: e, μ $(2\ell, 3\ell, 4\ell)$
  - $t\bar{t}W$ $(2\ell, 3\ell)$
  - $t\bar{t}Z$ $(3\ell, 4\ell)$
  - τ_{had} (τ_{had} channels)
- The QmisID and fake τ_{had} are estimated using data-driven techniques
- All the other backgrounds are estimated using the simulation
- Yields of $t\bar{t}W$ and non-prompt are adjusted via normalization factor (NF) in the final Fit



Irreducible background

$t\bar{t}W$ modelling

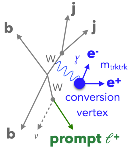
- Modelling of $t\bar{t}W$ is very challenging
  - To minimise dependence of μ_{ttH} on $t\bar{t}W$ prediction, introduce 3 normalization factors: $\hat{\lambda}^{2\ell\mathrm{LJ}}_{ttW} = 1.56^{+0.30}_{-0.28}$, $\hat{\lambda}^{2\ell\mathrm{HJ}}_{ttW} = 1.26^{+0.19}_{-0.18}$, and $\hat{\lambda}^{3\ell}_{ttW} = 1.68^{+0.30}_{-0.28}$.

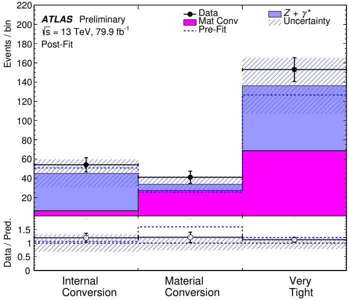


  - Additional uncertainties associated with modelling of b-jet multiplicity and W -boson charge asymmetry introduced to account for observed discrepancy between data and pre-fit bkg.

- The inclusive ttZ/γ^* with $(m(e^+e^-)>1\text{GeV})$ is computed at NLO
- ttZ/γ^* with $m(e^+e^-)<1\text{GeV}$ modelled via PS and is corrected by a scale factor estimated in 2ℓ IntC and 3ℓ IntC:

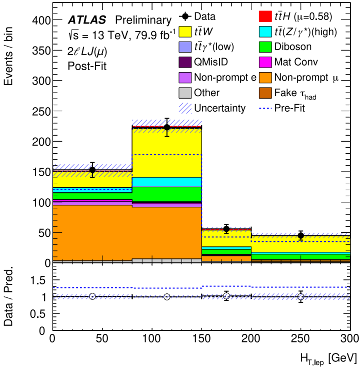
  $\hat{\lambda}_e^{IntC} = 0.83 \pm 0.32$
- Normalization factor is validated in $Z \rightarrow \mu^+\mu^-\gamma^*(\rightarrow e^+e^-)$ region
  - A 25% uncertainty is assigned as extrapolation uncertainty for "internal conversion" CR to other region





Non prompt lepton estimates

- Non-prompt leptons labeled according origin in simulated $t\bar{t}$ sample:
  - heavy (HF) or light (LF) flavour hadron decays
  - material conversion (MatC)
- Event categories introduced in 2ℓSS, 3ℓ are used to estimate fakes:
  - 2ℓ MatC and 3ℓ MatC control regions are enriched in mat conv.
  - 7 CRs in 2ℓ and 1 CR in 3ℓ are enriched in HF
  - Kinematic distribution are used to optimize the sensitivity for HF
    - HT^{lep} (scalar sum of lep) are introduced provide separation against $t\bar{t}W$
  - 3 NF for HF e, μ and mat.conv. are estimated from likelihood fit



- Normalization factor for HF μ, has the largest impact on $t\bar{t}H$

  $\hat{\lambda}_e^{MatC} = 1.61 \pm 0.48$, $\hat{\lambda}_e^{had} = 1.12 \pm 0.38$, and $\hat{\lambda}_\mu^{had} = 1.20 \pm 0.18$
- Estimation of non-prompt is validated in background dominated BDT score distribution

Systematic uncertainties

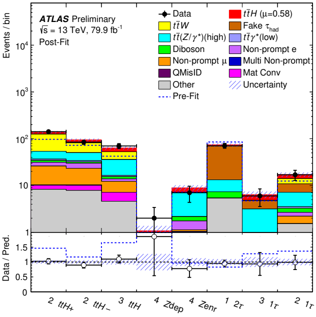
- The largest systematic uncertainties come from $t\bar{t}W$ and $t\bar{t}\ell\ell$ modelling

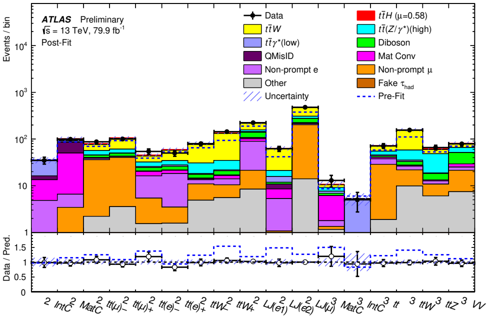
Uncertainty source	$\Delta\hat{\mu}$	
Jet energy scale and resolution	$+0.13$	-0.13
$t\bar{t}(Z/\gamma^*)$ (high mass) modelling	$+0.09$	-0.09
$t\bar{t}W$ modelling (radiation, generator, PDF)	$+0.08$	-0.08
Fake τ_{had} background estimate	$+0.07$	-0.07
$t\bar{t}W$ modelling (extrapolation)	$+0.05$	-0.05
$t\bar{t}H$ cross section	$+0.05$	-0.05
Simulation sample size	$+0.05$	-0.05
$t\bar{t}H$ modelling	$+0.04$	-0.04
Other background modelling	$+0.04$	-0.04
Jet flavour tagging and τ_{had} identification	$+0.04$	-0.04
Other experimental uncertainties	$+0.03$	-0.03
Luminosity	$+0.03$	-0.03
Diboson modelling	$+0.01$	-0.01
$t\bar{t}\gamma^*$ (low mass) modelling	$+0.01$	-0.01
Charge misassignment	$+0.01$	-0.01
Template fit (non-prompt leptons)	$+0.01$	-0.01
Total systematic uncertainty	$+0.25$	-0.22
Intrinsic statistical uncertainty	$+0.23$	-0.22
$t\bar{t}W$ normalisation factors	$+0.10$	-0.10
Non-prompt leptons normalisation factors (HF, material conversions)	$+0.05$	-0.05
Total statistical uncertainty	$+0.26$	-0.25
Total uncertainty	$+0.36$	-0.33

- Fakes impact is reducing its size with more statistics
- Additional uncertainties to cover data/MC disagreements as a function of NBjets and Lepton charge for $t\bar{t}W$
- Jet energy scale and resolution still plays a major role

Fit results

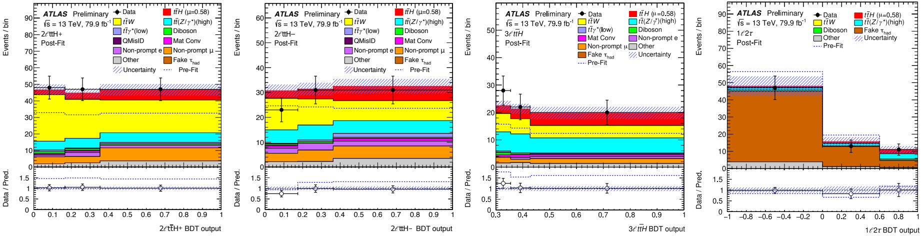
- Simultaneous fit in 25 categories
  - 17 (2ℓSS and 3ℓ) used to determine/constrain different backgrounds
  - 8 signal regions to measure $\sigma_{t\bar{t}H}$





Fit results

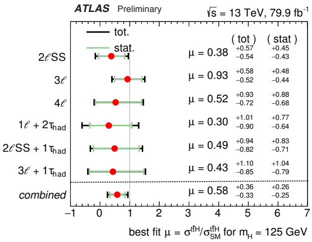
- Simultaneous fit in 25 categories
  - 17 (2ℓSS and 3ℓ) used to determine/constrain different backgrounds
  - 8 signal regions to measure $\sigma_{t\bar{t}H}$
    - a BDT discriminant is used in 2ℓSS, 3ℓ and $1\ell 2\tau_{had}$

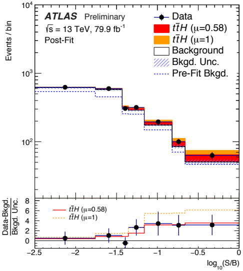


    - the total event yield is used in the remaining four signal regions

Fit results

- Significance with respect to background-only hypothesis = 1.8 (3.1)σ obs (exp)
- The best-fit value of μ is $\hat{\mu_{ttH}} = 0.58^{+0.36}_{-0.33}$
- Data prefers higher $t\bar{t}W$ cross section w.r.t to theoretical predictions:
  - $\hat{\lambda}^{2\ell\mathrm{LJ}}_{ttW} = 1.56^{+0.30}_{-0.28}$, $\hat{\lambda}^{2\ell\mathrm{HJ}}_{ttW} = 1.26^{+0.19}_{-0.18}$, and $\hat{\lambda}^{3\ell}_{ttW} = 1.68^{+0.30}_{-0.28}$





- Consistent results for $\mu_{t\bar{t}H}$ and $t\bar{t}W$ NFs are obtained using cross check analysis based on cut-and-count categorisation
- Alternative scenario with only one $t\bar{t}W$ normalisation:
  - $\hat{\mu_{ttH}} = 0.70^{+0.36}_{-0.33}$ and $\hat{\lambda}_{ttW} = 1.39^{+0.17}_{-0.16}$
  - Compatibility with the main result is 0.59 standard deviation

Summary

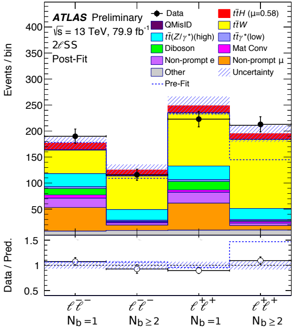
- Search for $t\bar{t}H$ in multilepton final states has been performed with 80fb^{-1}
- For a Higgs at 125 GeV, the measured signal strength is $0.58^{+0.36}_{-0.33}$, which correspond to $1.8\sigma(3.1\sigma)$ obs.(exp.) significance w.r.t bkg.-only hypothesis
- The measured cross section is $\hat{\sigma}(ttH) = 294^{+182}_{-162}$ fb and consistent with SM prediction
- $t\bar{t}W$ normalization is found to be in excess of the theoretical prediction, an improved description of $t\bar{t}W$ is needed to reach greater precision in the feature

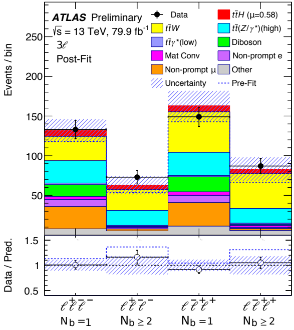
Backup

Irreducible background

$t\bar{t}W$ modelling

- Additional uncertainties associated with modelling of b-jet multiplicity and W -boson charge asymmetry introduced to account for observed discrepancy between data and pre-fit bkg.
- Extrapolation uncertainty:
  - Affects only the shape of the b-jet multiplicity and total charge distributions
  - For b-jet multiplicity, $\pm 25(\mp 35)\%$ for $1(\geq 2)$ b-jets
  - For total charge distribution $\pm 20(\mp 35)\%$ for positive (negative) total charge



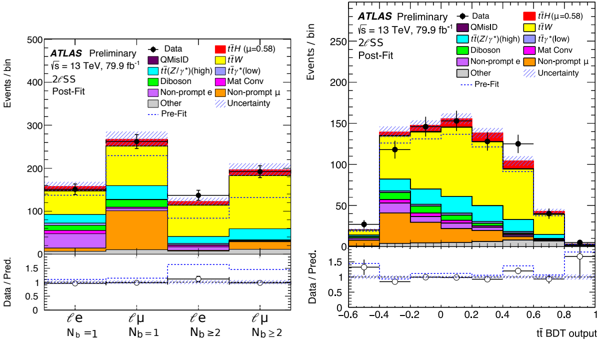


Back to modelling

Non-prompt leptons

Validation

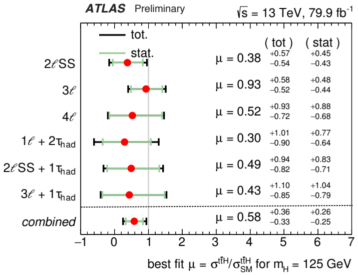
- Comparison between data and signal-plus-background prediction in the $2\ell SS0\tau_{had}$ channel after event selection for (a) the event yield, split in four separate categories depending on the flavour of the sub-leading lepton and the b-jet multiplicity, and (b) the score of the BDT trained to discriminate $t\bar{t}H$ signal from $t\bar{t}$ background

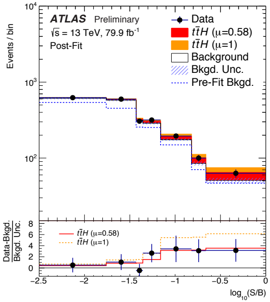


Back to non-prompt estimation

Fit results

- A search for $t\bar{t}H$ in multilepton final states with 80fb^{-1}
  - The significance of the observed (expected) excess above the background-only expectation ($\mu = 0$) is 1.8 (3.1) σ
  - The best-fit value of μ is
    - $\hat{\mu} = 0.58^{+0.26}_{-0.25}\ (\text{stat.})^{+0.19}_{-0.15}\ (\text{exp.})^{+0.13}_{-0.11}\ (\text{bkg. th.})^{+0.08}_{-0.07}\ (\text{sig. th.}) = 0.58^{+0.36}_{-0.33}$
  - Measured $t\bar{t}H$ cross section
    - $\hat{\sigma}(t\bar{t}H) = 294^{+132}_{-127}\ (\text{stat.})^{+94}_{-74}\ (\text{exp.})^{+73}_{-56}\ (\text{bkg. th.})^{+41}_{-39}\ (\text{sig. th.})\ \text{fb} = 294^{+182}_{-162}\ \text{fb}$





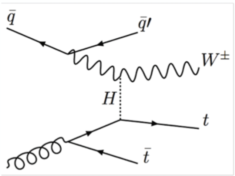
QMisID and τ_{had} fakes

- Charge mis-assignment
  - The electron charge mis-assignment rate vs p_T and $|\eta|$ is measured in data using OS/SS from $Z \to e^{\pm}e^{\mp}$ events for the three types of electrons
  - The total uncertainty for very tight electrons is 30%, with the dominant uncertainty at low p_T originating from the closure tests and at high p_T from limited statistics from $Z \to e^{\pm}e^{\mp}$ events
- Fake τ_{had} estimation
  - In $2\ell SS{+}\tau_{had}$, $3\ell\tau_{had}$
    - A fake τ_{had} factor scale factor as a function of $p_T^{\tau_{had}}$, is determined in $2\ell OS\tau_{had}$ control region
    - Total systematics uncertainty depends on $p_T^{\tau_{had}}$ and is on average ~13 (60)% for one-prong (three-prong) τ_{had}
  - $1\ell{+}2\tau_{had}$:
    - Fake τ_{had} is estimated from control region identical to the signal region but with same charge τ_{had}
    - Fakes(OS) = Data(SS)-Truth(SS) → equal probability for a jet to be reconstructed as a positively or negatively charge τ_{had}
    - An 22% systematics uncertainty from non-closure in CR is taken

Back to bkg composition

$t\bar{t}W$ higher order QCD and EW corrections

- A normalisation factor of **1.2** applied on top of the YR4 cross section for tt̄W
- Origin of the correction factor:
  - Factor **1.11** to account for missing QCD corrections in higher order XS
    - ttW+0j@NLO ➜ tt̄W+0,**1j**@NLO
    - estimated using dedicated samples generated with Sherpa 2.2.1 using the MEPS@NLO prescription, and cross-checked with the NLO generator MadGraph5_aMC@NLO 2.2.1 using the FxFx prescription
  - Factor **1.09** to account for missing EW corrections
    - [1711.02116] shows "subleading" NLO EWK corrections, not included in YR4 XS, can be large
    - primarily because of the large **NLO3 term** driven by the tt̄W+1-jet diagrams with a Higgs boson exchanged in the t-channel



13 TeV

$\delta[\%]$	$\mu = H_T/2$
LO_2	-
LO_3	0.9
NLO_1	50.0 (25.7)
NLO_2	−4.2 (−4.6)
NLO_3	12.2 (9.1)
NLO_4	0.04 (−0.02)