Measurement of $ttH$ production in multileptonic final states at the ATLAS experiment using $36 \, fb^{-1}$ data

Rohin Narayan
Southern Methodist University, Dallas

Motivation

- Fermion masses are generated through Yukawa interaction
- Heaviest SM particle (top) expected to have largest Yukawa coupling ($y_t$) to the Higgs field.

- Indirect constraints to top Yukawa coupling possible through ggH production as well as $H\to\gamma\gamma$ decay processes.
- ATLAS + CMS Run-1 combination of ratio of measured coupling of SM expectation ($\kappa_t$)
  - Assuming no new particles in the loop
  $$\kappa_t \equiv \frac{y_t}{y_t(SM)} = 0.87 \pm 0.15$$

- Direct measurement of top-Yukawa coupling need to disentangle any new physics effect.
- $ttH/tH$ production cross-section measurement is the only direct way to measure $y_t$
Experimental challenges

- Standard Model production cross section: ~507 fb: About 1% of total Higgs cross-section.
- Many final states
- Tiny signal and large backgrounds

**Signal**

- ttH(bb)
- ttH-multileptons
- ttH(γγ)

**Background**

- tt+(HF) jets (irreducible)
- ttV
- tttγ
Analysis strategy

- Targets $H \rightarrow ZZ^*$, $H \rightarrow WW^*$, $H \rightarrow \tau^+\tau^-$
- Events categorized based on number of light leptons and hadronic taus.

- Common jet selection $N_{\text{jet}} \geq 2$ and $N_{\text{bjet}} \geq 1$
- Optimized lepton selection in each category
  - A BDT algorithm to suppress fakes from B-hadron decay

- Light leptons channels dominated by $H \rightarrow WW^*$
- Tau channels dominated by $H \rightarrow \tau^+\tau^-$
Background Estimation

- **Prompt-leptons** and hadronic taus
  - Estimated with MonteCarlo
- Electron charge mis-identification for 2lSS and 2lSS+1τ-had
  - Datadriven estimation: Charge-misid rates in $Z \rightarrow e^+e^-/\mu^+\mu^-$
- **Light-lepton and non-prompt lepton fakes**
  - Source: Semileptonic b-decay, photon conversions
- **Hadronic-tau fakes**
  - Source: light flavor jets and mis-reconstructed electrons.

![Diagram of tt-bar system with leptons, taus, and photons]
Background: Prompt leptons

- Largest prompt backgrounds: $t\bar{t}W$, $t\bar{t}Z$ and $VV$
- Modeled using NLO MC generators (MG5aMC@NLO, Sherpa)
- Systematics: Cross-Section uncertainties, scale-variations, alternate generator comparisons
- Validated in dedicated validation regions
Background: Non prompt leptons

- A dedicated BDT algorithm to suppress Heavy flavor lepton fakes
- Based on lepton track associated to track-jets
- Signal and Background classification and training in MonteCarlo
- Validation in data using Z tag & Probe

<table>
<thead>
<tr>
<th>Variable</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N_{\text{track}}$ in track jet</td>
</tr>
<tr>
<td>IP2 $\log(p_{b}/p_{\text{light}})$</td>
</tr>
<tr>
<td>IP3 $\log(p_{b}/p_{\text{light}})$</td>
</tr>
<tr>
<td>$N_{\text{TrkAtVtx}}$ SV + JF</td>
</tr>
<tr>
<td>$p_{T}^{\text{lepton}}/p_{T}^{\text{track jet}}$</td>
</tr>
<tr>
<td>$\Delta R(\text{lepton, track jet})$</td>
</tr>
<tr>
<td>$p_{T}^{\text{VarCone30}}/p_{T}$</td>
</tr>
<tr>
<td>$E_{T}^{\text{TopoCone30}}/p_{T}$</td>
</tr>
</tbody>
</table>

\[L_{0} = \sqrt{d_{x}^2 + d_{y}^2}\]
Background: Light lepton fakes

**2l/3l**

- Fully data driven method using matrix method.
- Real and fake efficiencies are measured in data in a loose control region and extrapolated to the signal region.
- Method validated in various validation regions

**4l**

- Semi-data driven method in 4l-channel.
- MC split into prompt, heavy and light
- Scale factor determined for each of the components in 3l low-njet CR and extrapolated to 4l region
Background: Tau Fakes

- Relevant for all SR’s which has a hadronic tau.
  - Light lepton fakes are negligible except in $2\ell\text{SS}+1\tau$-had

$1\ell + 2\tau$-had

- Contains two OS $\tau$.
- Fakes are estimated in a SS CR and extrapolated to SR (MC based)

$2\ell\text{OS}+1\tau$-had

- Fake factor method: Measured in a control region inverting part of $\tau$ identification.
- Good data/MC agreement.

- Quite similar composition of fakes in $2\ell\text{OS}+1\tau$-had, $2\ell\text{SS}+1\tau$-had, $3\ell+1\tau$-had
- Data-driven estimates from $2\ell\text{OS}+1\tau$-had used as a correction factor for $\tau$-had fakes in those channels.
- Light-lepton fakes in $2\ell\text{SS}+1\tau$-had estimated using fake factor method.
Fit model

- Except for $3l+1\tau_{\text{had}}$ all sub channels use MVA techniques to further separate signal from background
  - $2\ellSS0\tau$: 2 BDT’s: against $t\bar{t}V$ and $t\bar{t}b\bar{t}$
  - $3\ell\tau$: 5 dimensional multinomial BDT to build categories enriched in $ttH$, $ttW$, $ttZ$, $t\bar{t}b\bar{t}$, diboson
  - $2\ellss+1\tau$: 1 BDT against $t\bar{t}b\bar{t}$
  - Cut and count cross check analysis in high sensitivity sub-channels.

### Table: BDT trained against

<table>
<thead>
<tr>
<th>BDT trained against</th>
<th>$2\ellSS$</th>
<th>$3\ell$</th>
<th>$4\ell$</th>
<th>$1\ell+2\tau_{\text{had}}$</th>
<th>$2\ellSS+1\tau_{\text{had}}$</th>
<th>$2\ellOS+1\tau_{\text{had}}$</th>
<th>$3\ell+1\tau_{\text{had}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Discriminant</td>
<td>Fakes and $t\bar{t}V$</td>
<td>$t\bar{t}$, $t\bar{t}W$, $t\bar{t}Z$, $VV$</td>
<td>$t\bar{t}Z$ / -</td>
<td>$t\bar{t}$</td>
<td>all</td>
<td>$t\bar{t}$</td>
<td>-</td>
</tr>
<tr>
<td>Number of bins</td>
<td>6</td>
<td>5</td>
<td>1 / 1</td>
<td>2</td>
<td>2</td>
<td>10</td>
<td>1</td>
</tr>
<tr>
<td>Control regions</td>
<td>-</td>
<td>4</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

- Simultaneous fit in 12 regions (CR+SR)
- Single bin used in 3l CR’s as well as low stat SR’s
  - $3l+1\tau$
  - $4l$ (z-enriched, z-depleted)
- BDT shape information used in 5 SR’s
Post-Fit: Signal regions

2lSS

2lSS + 1τ

2lSS + 1τ

4l-Zdep

3l

3l + 1τ

1l + 2τ

4l-Zenr
ttHML: Results

- Measured $\mu = 1.6^{+0.5}_{-0.4}$ with a significance with respect to background only hypothesis $4.1\sigma$ (expct. $2.8\sigma$)

- Cross section extrapolated to inclusive phase space

$$\sigma(t\bar{t}H) = 790^{+150}_{-150}(\text{stat})^{+170}_{-150}(\text{syst}) \text{ fb}$$

- Leading systematic uncertainty from ttH modeling followed by jet-energy scale and non-prompt light lepton estimates.
Conclusion

- Presented results of \( \text{ttH-multileptons} \) analysis measured at the ATLAS experiment using \( 36 fb^{-1} \) of data collected in 2015-16.

- Leading systematics that affect the analysis are ttH, ttV-modeling (theory), jet-energy scale, light-lepton fake estimation (experimental).

- In combination with other ttH-channels ATLAS observed ttH-production in pp collisions at the LHC.

- The analysis is being updated to include newer data collected in latest data taking periods.
Backup
Object definitions and Selections

Leptons

- $p_T > 10$ GeV; central $\eta$

<table>
<thead>
<tr>
<th>Isolation</th>
<th>$L$</th>
<th>$L^+$</th>
<th>$L^*$</th>
<th>$T$</th>
<th>$T^*$</th>
<th>$L$</th>
<th>$L^+$</th>
<th>$L^*$</th>
<th>$T$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-prompt lepton BDT</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Identification</td>
<td>Loose</td>
<td>Tight</td>
<td>Loose</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Charge mis-assignment veto</td>
<td>No</td>
<td>Yes</td>
<td>N/A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Traversal impact parameter significance</td>
<td>$</td>
<td>d_0</td>
<td>/\sigma_{d_0}$</td>
<td>$&lt; 5$</td>
<td>$&lt; 3$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Longitudinal impact parameter</td>
<td>$</td>
<td>z_0 \sin \theta</td>
<td>$</td>
<td>$&lt; 0.5$ mm</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Trigger

- A combination of single and di-leptonic trigger

light & b-jets

- $p_T > 25$ GeV, $|\eta| < 2.5$
- b-tag WP: MV2c10 70%

Hadronic Taus

Medium/Tight ID to reject jets

- $p_T > 25$ GeV
- BDT to reject el faking taus
- b-jet veto
- tau vertex is PV

Overlap Removal

<table>
<thead>
<tr>
<th>Keep</th>
<th>Remove</th>
<th>Cone size ($\Delta R$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>electron</td>
<td>electron (low $p_T$)</td>
<td>0.1</td>
</tr>
<tr>
<td>muon</td>
<td>electron</td>
<td>0.1</td>
</tr>
<tr>
<td>electron</td>
<td>jet</td>
<td>0.3</td>
</tr>
<tr>
<td>jet</td>
<td>muon</td>
<td>min(0.4, 0.04 + 10[GeV]/$p_T$ (muon))</td>
</tr>
<tr>
<td>electron</td>
<td>tau</td>
<td>0.2</td>
</tr>
<tr>
<td>muon</td>
<td>tau</td>
<td>0.2</td>
</tr>
<tr>
<td>tau</td>
<td>jet</td>
<td>0.3</td>
</tr>
</tbody>
</table>
\( \bar{t}tH \) ATLAS combination results

<table>
<thead>
<tr>
<th></th>
<th>Run-1</th>
<th>Run-2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Year</td>
<td>2011</td>
<td>2012</td>
</tr>
<tr>
<td>Collision energy</td>
<td>7 TeV</td>
<td>8 TeV</td>
</tr>
<tr>
<td>( \bar{t}tH )-ML</td>
<td>24.8 fb(^{-1})</td>
<td>36 fb(^{-1})</td>
</tr>
<tr>
<td>( \bar{t}tH(\bar{b}b) )</td>
<td>24.8 fb(^{-1})</td>
<td>36 fb(^{-1})</td>
</tr>
<tr>
<td>( \bar{t}tH(\gamma\gamma) )</td>
<td>24.8 fb(^{-1})</td>
<td>79.8 fb(^{-1})</td>
</tr>
<tr>
<td>( \bar{t}tH(Z, Z^*) )</td>
<td>24.8 fb(^{-1})</td>
<td>79.8 fb(^{-1})</td>
</tr>
</tbody>
</table>

- \( \mu_{\bar{t}tH} = 1.32 \pm 0.18 \text{(stat)} \pm ^{+0.21}_{-0.19} \text{(syst)} \)
- Observed significance 6.3 \( \sigma \) (expected 5.1 \( \sigma \))
Electron reconstruction

- First layer (strips): $\Delta \eta \times \Delta \varphi = 0.0031 \times 0.098$
- Second layer: $\Delta \eta \times \Delta \varphi = 0.025 \times 0.0245$
- Third layer: $\Delta \eta \times \Delta \varphi = 0.05 \times 0.0245$

Beam axis, beam spot, d$_0$, pixels, SCT, insertable B-layer, TRT (73 layers), electromagnetic calorimeter, hadronic calorimeter.
Matrix method fakes

- $TT_i$: event with both leptons passing tight selection (Tot. events: $N^{TT}$).
- $TT'_i$: event with leading lepton passing tight selection and subleading lepton failing tight selection (Tot. events: $N^{TT}$).
- $TT_i$: event with leading lepton failing tight selection and subleading lepton passing tight selection (Tot. events: $N^{TT}$).
- $TT'_i$: event with both leptons failing tight selection (Tot. events: $N^{TT}$).

- $rr_i$: event with both leptons being real (Tot. events: $N^{rr}$).
- $rf_i$: event with leading lepton being real and subleading lepton being fake (Tot. events: $N^{rf}$).
- $fr_i$: event with leading lepton being fake and subleading lepton being real (Tot. events: $N^{fr}$).
- $ff_i$: event with both leptons being fake (Tot. events: $N^{ff}$).

\[
\begin{pmatrix}
N^{TT} \\
N^{TT'} \\
N^{TT} \\
N^{TT'}
\end{pmatrix}
= 
\begin{pmatrix}
\varepsilon_{r,1} \varepsilon_{r,2} & \varepsilon_{r,1} \varepsilon_{f,2} & \varepsilon_{f,1} \varepsilon_{r,2} & \varepsilon_{f,1} \varepsilon_{f,2} \\
\varepsilon_{r,1} \varepsilon_{r,2} & \varepsilon_{r,1} \varepsilon_{f,2} & \varepsilon_{f,1} \varepsilon_{r,2} & \varepsilon_{f,1} \varepsilon_{f,2} \\
\varepsilon_{r,1} \varepsilon_{r,2} & \varepsilon_{r,1} \varepsilon_{f,2} & \varepsilon_{f,1} \varepsilon_{r,2} & \varepsilon_{f,1} \varepsilon_{f,2} \\
\varepsilon_{r,1} \varepsilon_{r,2} & \varepsilon_{r,1} \varepsilon_{f,2} & \varepsilon_{f,1} \varepsilon_{r,2} & \varepsilon_{f,1} \varepsilon_{f,2}
\end{pmatrix}
\begin{pmatrix}
N^{rr} \\
N^{rf} \\
N^{fr} \\
N^{ff}
\end{pmatrix}
\]
Fake fraction

- Various sources of light-lepton fakes in 2l and 3l channel