Measurement of $t\bar{t}H$ Production in $H \rightarrow b\bar{b}$ Decay Channel at CMS

ABHISEK DATTA
CORNELL UNIVERSITY
ON BEHALF OF THE CMS COLLABORATION
JULY 30, 2019
Introduction to $t\bar{t}H$

Associated Higgs production with $t\bar{t}$ ($t\bar{t}H$) is the **best direct probe** of the Top-Higgs coupling

- SM Yukawa coupling ($\propto m_t$)
- New physics beyond the SM

Many possible decay channels:
- $H \rightarrow b\bar{b}$
- $H \rightarrow WW^*, \tau\tau, ZZ^*(\text{multilepton})$
- $H \rightarrow \gamma\gamma$
- $H \rightarrow ZZ^* \rightarrow 4l$

$t\bar{t}H$ observation by the combination of all channels by both CMS and ATLAS in 2018

Run 1 + 2016 Data

- $5.1 \text{ fb}^{-1} (7 \text{ TeV}) + 19.7 \text{ fb}^{-1} (8 \text{ TeV}) + 35.9 \text{ fb}^{-1} (13 \text{ TeV})$

$\sigma_{t\bar{t}H} = 0.5 \text{ pb}$ at 13 TeV

$t\bar{t}H(b\bar{b})$ Channel

$t\bar{t}H(b\bar{b})$ is one of the most sensitive channels for $t\bar{t}H$ measurement

- Large (58%) $H \rightarrow b\bar{b}$ branching ratio

But challenging final state:

- Many jets: no unambiguous event reconstruction
- Large irreducible background from $t\bar{t} + b\bar{b}$ production
- Large QCD background in Fully Hadronic channel

This talk:

- Results from CMS with 41.5 $fb^{-1}$ of 2017 data [CMS-PAS-HIG-18-030]
- This is an update of the 2016 analyses
- Combination of 2017 results with 2016 also shown

Improved b-tagging and modelling of parton-shower uncertainties

Different final states depending on $t\bar{t}$ decay:

- Fully Hadronic (FH)
- Single Lepton (SL)
- Dilepton (DL)

Previous CMS Results with 2016 data ($35.9\ fb^{-1}$):

![Graph showing CMS results with 2016 data](JHEP 1903 (2019) 026)

![Graph showing CMS results with 2016 data](JHEP 06 (2018) 101)
Final States for $t\bar{t}H(b\bar{b})$

Different final states depending on $t\bar{t}$ decay:
- Fully Hadronic (FH)
- Single Lepton (SL)
- Dilepton (DL)

Leptons = $e$ or $\mu$ (no explicit $\tau$ reconstruction or veto)

Dedicated selections and analysis methods optimized per channel

Multiple jets (including $b$-tagged jets) in final state:
- Huge combinatorics in event reconstruction
- $t\bar{t} + b\bar{b}$ (10-20 $\times$ signal) irreducible background

<table>
<thead>
<tr>
<th></th>
<th>FH channel</th>
<th>SL channel</th>
<th>DL channel</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of leptons</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>$p_T$ of leptons ($e/\mu$) [GeV]</td>
<td>—</td>
<td>$&gt; 30/29$</td>
<td>$&gt; 25/25$ GeV</td>
</tr>
<tr>
<td>$p_T$ of additional leptons [GeV]</td>
<td>&lt; 15</td>
<td>&lt; 15</td>
<td>&lt; 15</td>
</tr>
<tr>
<td>$</td>
<td>\eta</td>
<td>$ of leptons</td>
<td>&lt; 2.4</td>
</tr>
<tr>
<td>Number of jets</td>
<td>$\geq 6$</td>
<td>$\geq 4$</td>
<td>$\geq 2$</td>
</tr>
<tr>
<td>$p_T$ of jets [GeV]</td>
<td>$&gt; 40$</td>
<td>$&gt; 30$</td>
<td>$&gt; 30, 30, 20$</td>
</tr>
<tr>
<td>$</td>
<td>\eta</td>
<td>$ of jets</td>
<td>&lt; 2.4</td>
</tr>
<tr>
<td>Number of $b$-tagged jets</td>
<td>$\geq 2$</td>
<td>$\geq 2$</td>
<td>$\geq 1$</td>
</tr>
<tr>
<td>$\not{p_T}$ miss</td>
<td>—</td>
<td>$&gt; 20$ GeV</td>
<td>$&gt; 40$ GeV</td>
</tr>
</tbody>
</table>

CMS-PAS-HIG-18-030
$t\bar{t} + b\bar{b}$ Background

Uncertainties in theoretical predictions $\geq 30\%$

Difficulties:

- Complex multi-parton final state
- Multiple very different energy scales

Different Monte Carlo generators tested:
distributions similar within uncertainties

$t\bar{t} + b\bar{b}$ irreducible background in all channels:
- In general, large uncertainties in both theoretical predictions and experimental measurements

$t\bar{t} + b\bar{b}$ experimental measurements also difficult:

- Performed in different $t\bar{t}$ decay channels
- Currently measured precision around 30\%
- Measured cross-section around 30 – 50\% larger than prediction

Data cannot yet distinguish between most models

Therefore, large $t\bar{t} + b\bar{b}$ cross-section uncertainties need to be accounted for in $t\bar{t}H(b\bar{b})$ measurement
Analysis Strategy

Challenges:
- Large combinatorial final state
- Large backgrounds with large uncertainties
- Relatively poor b-jet energy resolution

Direct invariant mass reconstruction of the Higgs not possible

- Event categorization to form signal and control regions to constrain background
- Multivariate Discriminants to separate signal from the large backgrounds
- Final Result: combined maximum-likelihood fit in all the different categories from all three channels

Backgrounds:

\( tt + b \) jets (all channels):
- Modeled from \( tt \) simulation at NLO
- Parton-shower uncertainties important (as shape variations)
- Treat \( tt + b\bar{b}, tt + b \) and \( tt + 2b \) as separate processes based on B-hadron content (rest: \( tt + c\bar{c}, tt + l\bar{f} \)), with separate uncertainties
- Additional 50% normalization uncertainty to all 4 \( tt + HF \) processes

QCD Multijet (Fully Hadronic channel):
- Dedicated background rejection
- QCD Estimation from Data

Other minor backgrounds (all channels):
- Predictions from simulation
Event Categorization:

- Based on Jet Multiplicity:
  - \((4j, \geq 3b)\), \((5j, \geq 3b)\) and \((\geq 6j, \geq 3b)\)

- Multi-classification using Artificial Neural Networks (ANN) per jet category:
  - \(t \bar{t}H\) (signal region)
  - The 5 \(t \bar{t} + \text{jets}\) background processes (control regions)

Input variables used for the ANN:
- Lepton and Jet kinematic variables
- Invariant masses
- Angular differences
- B-tagging information
- Matrix Element Method discriminant

In total, **18 jet-process categories** with different signal and background composition
Single Lepton Channel

Event Categorization:

- Based on Jet Multiplicity:
  - $(4j, \geq 3b), (5j, \geq 3b)$ and $(\geq 6j, \geq 3b)$

- Multi-classification using Artificial Neural Networks (ANN) per jet category:
  - $t\bar{t}H$ (signal region)
  - The 5 $t\bar{t} + \text{jets}$ background processes (control regions)

Input variables used for the ANN:
- Lepton and Jet kinematic variables
- Invariant masses
- Angular differences
- B-tagging information
- Matrix Element Method discriminant

In total, **18 jet-process categories** with different signal and background composition
Single Lepton Channel

Final Discriminant:
In each of the 18 process categories, the corresponding ANN output

Most sensitive signal category:
- $\geq 6j, \geq 3b$
- $t\bar{t}H$ output node
The discriminant here is the ANN output value of the $t\bar{t}H$ node

The $t\bar{t}+jets$ nodes (with higher background content) : constrain background uncertainties

Separation visible between $t\bar{t}H$ signal and $t\bar{t}+jets$ background processes
Dilepton Channel

Event Categorization:

- Based on jet and b-tagged jet multiplicities
  - (3j, 2b) \( \geq 4j, 2b \)
  - (3j, 3b) \( \geq 4j, 3b \)
  - \( \geq 4j, \geq 4b \)

5 jet-btag categories with different signal and background composition

Different categories with different \( tt\bar{t} + \text{jets} \) composition: helps to constrain the backgrounds

\( tt\bar{t} + b\bar{b} \) the dominant background in the most signal sensitive category

---

**CMS Simulation Preliminary**

<table>
<thead>
<tr>
<th>Category</th>
<th>S/B</th>
<th>S/(t\bar{t})</th>
</tr>
</thead>
<tbody>
<tr>
<td>(3 jets, 2 b tags)</td>
<td>0.007, 0.14</td>
<td>0.14</td>
</tr>
<tr>
<td>(3 jets, 3 b tags)</td>
<td>0.0083, 0.28</td>
<td>0.28</td>
</tr>
<tr>
<td>(\geq 4 jets, 2 b tags)</td>
<td>0.0028, 0.53</td>
<td>0.53</td>
</tr>
<tr>
<td>(\geq 4 jets, \geq 4 b tags)</td>
<td>0.0173, 0.88</td>
<td>0.88</td>
</tr>
<tr>
<td>(\geq 4 jets, \geq 4 b tags)</td>
<td>0.0624, 0.89</td>
<td>0.89</td>
</tr>
</tbody>
</table>

Pre-fit expectation

CMS-PAS-HIG-18-030

5 jet-btag categories with different signal and background composition

6% S/B
**Final Discriminant:**

Boosted Decision Trees (BDT) in each jet-btag category

BDTs trained to separate $t\bar{t}H$ from inclusive $t\bar{t}+jets$ background

**Input variables** used for the BDT:
- Lepton and Jet kinematic variables
- Invariant masses
- Angular differences
- B-tagging information
- Matrix Element Method discriminant

Separation visible between $t\bar{t}H$ signal and $t\bar{t}+jets$ background processes
**Event Categorization:**

- Based on jet and b-tagged jet multiplicities
  - (7 jets, 3 b tags), (8 jets, 3 b tags), (≥ 9 jets, 3 b tags)
  - (7 jets, ≥ 4 b tags), (8 jets, ≥ 4 b tags), (≥ 9 jets, ≥ 4 b tags)

**CMS Preliminary**

- FH (7 jets, 3 b tags)
  - S/B = 0.0037, S/√B = 0.76
- FH (8 jets, 3 b tags)
  - S/B = 0.0046, S/√B = 0.86
- FH (≥ 9 jets, 3 b tags)
  - S/B = 0.0055, S/√B = 0.88
- FH (7 jets, ≥ 4 b tags)
  - S/B = 0.0131, S/√B = 0.76
- FH (8 jets, ≥ 4 b tags)
  - S/B = 0.0150, S/√B = 0.98
- FH (≥ 9 jets, ≥ 4 b tags)
  - S/B = 0.0158, S/√B = 1.02

**Pre-fit expectation**

- tH
- Multijet
- t+t+lf
- t+t+c+c
- t+t+b
- t+t+2b
- t+t+b+b
- Other Bkg

QCD is the dominant background

6 jet-btag categories with different signal and background composition

These compositions are after QCD background rejection

1.6% S/B

QCD is the dominant background
Fully Hadronic Channel

QCD Background Rejection:

- $t\bar{t}H$ contains W-boson: Require $m_{qq}$ (for non b-tagged jet pair with closest $m_{qq}$ to W-mass) to be inside the W-mass window
- Jets in QCD separated further: Cut on $\Delta\eta_{jets}$ (average $\Delta\eta$ between furthest separated jets)
- QCD has more gluon jets than $t\bar{t}H$: Cut on quark - gluon likelihood ratio (QGLR) where the QGL separates light-flavor gluon jets from quark jets

QCD Background Estimate:

- Discriminant shape determined from Control Region
- Normalization freely-floating in final fit
- Corrections applied for kinematic differences in b-tagged jets

---

**Signal region**

<table>
<thead>
<tr>
<th>QGLR &gt; 0.5</th>
<th>QGLR &lt; 0.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>CR</td>
<td>Validation CR</td>
</tr>
<tr>
<td>(to extract distribution)</td>
<td>(to validate distribution)</td>
</tr>
<tr>
<td>CR</td>
<td>SR</td>
</tr>
<tr>
<td>(final analysis)</td>
<td>(comparison with data)</td>
</tr>
</tbody>
</table>

---

CMS Preliminary

FH SR (8 jets, 2+ b tags)

Pre-fit expectation

- Data
- $t\bar{t}$
- $W$ jets
- $Z$ jets
- $V+jets$
- $t+V$
- Diboson
- Uncertainty

$N_{b\,\text{tag}} = 2$
$N_{b\,\text{tag loose}} \geq 3$
$N_{b\,\text{tag}} \geq 3$

CMS Preliminary

FH (>6 jets, >2 b tags)

Data
- $t\bar{t}$
- $W$ jets
- $Z$ jets
- $V+jets$
- $t+V$
- Diboson
- Uncertainty

$QGLR (3b)$

---

CMS-PAS-HIG-18-030
Final Discriminant:
Matrix Element Method (MEM)
discriminant in each jet-btag category

\[ \text{Discriminant} = \frac{f(t\bar{t}H)}{f(t\bar{t}H) + \kappa f(t\bar{t} + b\bar{b})} \]

Uses likelihood \((f)\) of event kinematics:
- Based on LO matrix elements
- Detector effects via transfer functions

Constructed to discriminate \(t\bar{t}H\) signal
from \(t\bar{t} + b\bar{b}\) background but works well for \(t\bar{t} + l\bar{f}\) and QCD too
Results from 2017 Data

**Final result** obtained by a maximum likelihood fit of signal and background expectations to data simultaneously in all categories to extract $t\bar{t}H$ signal strength:

- 6 **Fully Hadronic** categories
- 18 **Single Lepton** categories
- 5 **Dilepton** categories

---

Best Fit Signal Strength: $\hat{\mu} = \frac{\hat{\sigma}}{\sigma_{SM}} = 1.49^{+0.44}_{-0.40}$

Observed Significance: $3.7\sigma$ ($2.6\sigma$ expected)
Systematic Uncertainties

Major sources of systematic uncertainties:

- Experimental Sources:
  - B-tagging
  - Jet Energy Scale and Resolution

- Theoretical Sources:
  - $t\bar{t}H$ cross-section
  - $t\bar{t} +$ HF modeling

- QCD background prediction

- Size of simulated samples

Largest Impacts:

- $t\bar{t}H$ cross-section: artefact since measuring $\hat{\sigma}/\sigma_{SM}$
- $t\bar{t} + b\bar{b}$ modeling: 50% assigned prior to fit
Combined Results from 2016 and 2017 Data

2017 results combined with 35.9 fb$^{-1}$ of 2016 data:

- $t\bar{t}H$ and $t\bar{t}$ cross-section uncertainties treated as correlated
- Other modeling uncertainties (including the 50% normalization uncertainty on $t\bar{t} + HF$ processes) uncorrelated
- Most experimental uncertainties also uncorrelated

Best fit Signal Strength: $\hat{\mu} = \frac{\hat{\sigma}}{\sigma_{SM}} = 1.15^{+0.32}_{-0.29}$

Observed Significance: $3.9\sigma$ ($3.5\sigma$ expected)

Evidence of $t\bar{t}H$ production in the $H \rightarrow b\bar{b}$ decay channel at CMS
Combined Results from 2016 and 2017 Data

2017 results combined with $35.9 fb^{-1}$ of 2016 data:

- $t\bar{t}H$ and $tt$ cross-section uncertainties treated as correlated
- Other modeling uncertainties (including the 50% normalization uncertainty on $t\bar{t} + HF$ processes) uncorrelated
- Most experimental uncertainties also uncorrelated

Best fit Signal Strength: $\hat{\mu} = \frac{\hat{\sigma}}{\sigma_{SM}} = 1.15^{+0.32}_{-0.29}$

Observed Significance: $3.9\sigma$ ($3.5\sigma$ expected)

Evidence of $t\bar{t}H$ production in the $H \rightarrow b\bar{b}$ decay channel at CMS

### Post-fit uncertainties (2016+2017)

<table>
<thead>
<tr>
<th>Uncertainty source</th>
<th>$\Delta \hat{\mu}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total experimental</td>
<td>$+0.15/-0.13$</td>
</tr>
<tr>
<td>$b$ tagging</td>
<td>$+0.08/-0.07$</td>
</tr>
<tr>
<td>Jet energy scale and resolution</td>
<td>$+0.05/-0.04$</td>
</tr>
<tr>
<td>Total theory</td>
<td>$+0.23/-0.19$</td>
</tr>
<tr>
<td>Signal</td>
<td>$+0.15/-0.06$</td>
</tr>
<tr>
<td>$t\bar{t}$+hf modelling</td>
<td>$+0.14/-0.15$</td>
</tr>
<tr>
<td>QCD background prediction</td>
<td>$+0.10/-0.08$</td>
</tr>
<tr>
<td>Size of simulated samples</td>
<td>$+0.10/-0.10$</td>
</tr>
<tr>
<td>Total systematic</td>
<td>$+0.28/-0.25$</td>
</tr>
<tr>
<td>Statistical</td>
<td>$+0.15/-0.15$</td>
</tr>
<tr>
<td>Total</td>
<td>$+0.32/-0.29$</td>
</tr>
</tbody>
</table>

Largest source of experimental uncertainty: Size of simulated samples

Largest source of theoretical uncertainty: $b$ tagging

Sensitivity limited mostly by systematic uncertainties

CMS-PAS-HIG-18-030
Summary

$t\bar{t}H$ measurement in the $H \to b\bar{b}$ decay channel at CMS reported for 41.5 fb$^{-1}$ of 2017 data

• Results combined from 3 different $t\bar{t}$ decay channels: Single Lepton, Dilepton and Fully Hadronic

• Event categorization and multivariate discriminants necessary to constrain backgrounds and separate signal from large backgrounds respectively

• Signal strength compatible with SM prediction
  
  o 2017 only : $\hat{\mu} = 1.49^{+0.44}_{-0.40} \ (3.7\sigma)$
  
  o 2016 + 2017 : $\hat{\mu} = 1.15^{+0.32}_{-0.29} \ (3.9\sigma)$

• Sensitivity currently limited mostly by systematic uncertainties
CMS Detector

- Silicon Tracker:
  - Pixel Detector
  - Strip Detector

- Electromagnetic Calorimeter

- Hadronic Calorimeter

- Muon Detectors:
  - Drift Tube
  - Cathode Strip Chamber
  - Resistive Plate Chamber

http://cds.cern.ch/record/2120661
Separating $t\bar{t} + X$ Processes

$t\bar{t} + jets$ events in simulation are classified based on the flavor of the additional jets that do not arise from a $t$-quark decay:

- Jets at particle level with $p_T > 20 \text{ GeV}$, $|\eta| < 2.4$
- HF hadrons associated to jets via ghost-hadron matching

5 $t\bar{t} + X$ processes are defined:

- $t\bar{t} + b\bar{b} : \geq 2$ additional jets containing $\geq 1$ b-hadrons each
- $t\bar{t} + b : 1$ additional jet containing 1 b-hadron
- $t\bar{t} + 2b : 1$ additional jet containing $\geq 2$ b-hadrons
- $t\bar{t} + c\bar{c} : \geq 2$ additional jets containing c-hadrons
- $t\bar{t} + lf :$ otherwise
Yields from All $t\bar{t}$ Decay Channels

<table>
<thead>
<tr>
<th>Process</th>
<th>FH channel</th>
<th>SL channel</th>
<th>DL channel</th>
</tr>
</thead>
<tbody>
<tr>
<td>QCD</td>
<td>2938305 ± 301286</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>$t\bar{t}$+lf</td>
<td>357488 ± 52694</td>
<td>718341 ± 83944</td>
<td>275407 ± 19610</td>
</tr>
<tr>
<td>$t\bar{t}$+c$\bar{c}$</td>
<td>93674 ± 11860</td>
<td>96581 ± 13795</td>
<td>24721 ± 3145</td>
</tr>
<tr>
<td>$t\bar{t}$+b</td>
<td>23737 ± 2892</td>
<td>27222 ± 3749</td>
<td>5613 ± 824</td>
</tr>
<tr>
<td>$t\bar{t}$+2b</td>
<td>14039 ± 2183</td>
<td>10537 ± 2206</td>
<td>1697 ± 351</td>
</tr>
<tr>
<td>$t\bar{t}$+b$\bar{b}$</td>
<td>19730 ± 2413</td>
<td>12770 ± 2050</td>
<td>1813 ± 262</td>
</tr>
<tr>
<td>Single t</td>
<td>24117 ± 1847</td>
<td>38170 ± 3720</td>
<td>14044 ± 1133</td>
</tr>
<tr>
<td>V+jets</td>
<td>31154 ± 2319</td>
<td>14491 ± 1754</td>
<td>2199 ± 264</td>
</tr>
<tr>
<td>$t\bar{t}$+V</td>
<td>2924 ± 228</td>
<td>2963 ± 286</td>
<td>1028 ± 99</td>
</tr>
<tr>
<td>Diboson</td>
<td>354 ± 40</td>
<td>503 ± 61</td>
<td>420 ± 55</td>
</tr>
<tr>
<td>Total bkg.</td>
<td>3505523 ± 339615</td>
<td>921576 ± 97714</td>
<td>326942 ± 23458</td>
</tr>
<tr>
<td>$t\bar{t}$H</td>
<td>2556 ± 164</td>
<td>1747 ± 167</td>
<td>363 ± 22</td>
</tr>
<tr>
<td>Data</td>
<td>3508079</td>
<td>923936</td>
<td>331055</td>
</tr>
</tbody>
</table>

- Yields after baseline selection
- Prior to fit to data
- QCD prediction from simulation
- Uncertainties (statistical + systematic) excluding 50% $t\bar{t}$ + HF normalization uncertainty

CMS-PAS-HIG-18-030
**FH : QCD Rejection and Estimation**

**QCD rejection : angular distance cut**

\[
\Delta \eta_{jets} = \frac{1}{N_{jets}} \sum_{jets} \Delta \eta \text{ (jet, jet furthest apart)}
\]

\[
\Delta \eta_{jets} \leq 2.52
\]

**QCD rejection : QGLR cut**

\[
\text{QGLR} = \frac{L_{qqq}}{L_{qqq} + L_{ggg}}
\]

\[
\text{QGLR} > 0.5
\]

**QCD estimation : from Control Region**

\[
\text{MEM}_{QCD} = \text{MEM}_{Data} - \text{MEM}_{\text{non-QCD background (MC)}}
\]

Estimated shape validated in dedicated validation region

---

[Graph showing data and expected distributions with labels and axes]
Input Variables to ANNs and BDTs

<table>
<thead>
<tr>
<th>Variable</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>MEM</td>
<td>max. element method discriminant</td>
</tr>
<tr>
<td>BLR</td>
<td>likelihood ratio discriminating between events with 4 b-tagged jets and 2 b-tagged jets</td>
</tr>
<tr>
<td>BLR^BLR</td>
<td>ln(BLR/(1 − BLR))</td>
</tr>
<tr>
<td>pT(jet 1)</td>
<td>pT of the 1 jet, ranked in jet pt</td>
</tr>
<tr>
<td>pT(jet 3)</td>
<td>pT of the 3 jet, ranked in jet pt</td>
</tr>
<tr>
<td>∑pT</td>
<td>scalar sum of pT of b-tagged jets</td>
</tr>
<tr>
<td>d(jet)</td>
<td>b-tagging discriminant value of 4 jet, ranked in jet pt</td>
</tr>
<tr>
<td>R2</td>
<td>2. highest b-tagging discriminant value of all jets</td>
</tr>
<tr>
<td>R2̂</td>
<td>average b-tagging discriminant value of all jets</td>
</tr>
<tr>
<td>R2̄</td>
<td>average b-tagging discriminant value of all jets</td>
</tr>
<tr>
<td>R̂</td>
<td>minimal b-tagging discriminant value of all b-tagged jets</td>
</tr>
<tr>
<td>∆(d̂−d̄)^2</td>
<td>squared difference between the b-tagging discriminant value of a b-tagged jet and the average b-tagging discriminant values of all b-tagged jets, averaged over all b-tagged jets</td>
</tr>
<tr>
<td>m^m_{vis}</td>
<td>mass of the jets closest to 125 GeV</td>
</tr>
<tr>
<td>m^m_{vis}</td>
<td>mass of pair of b-tagged jets closest to 125 GeV</td>
</tr>
<tr>
<td>r^m_{vis}</td>
<td>ratio H1/H0 of 0th and first Fox–Wolfram moment computed with all b-tagged jets</td>
</tr>
<tr>
<td>0^0</td>
<td>twist angle between pair of jets with maximum mass</td>
</tr>
<tr>
<td>0^0</td>
<td>twist angle between pair of b-tagged jets with maximum mass</td>
</tr>
<tr>
<td>N^0_{bkg}</td>
<td>multiplicity of dijet pairs with invariant mass between 100 GeV and 140 GeV</td>
</tr>
</tbody>
</table>

Variable Definition

<table>
<thead>
<tr>
<th>Variable</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>m^m_{vis}</td>
<td>mass of pair of jet and b-tagged jet closest in AR</td>
</tr>
<tr>
<td>m^m_{vis}</td>
<td>average mass of all pairs of b-tagged jets</td>
</tr>
<tr>
<td>m^m_{vis}</td>
<td>mass of pair of b-tagged jets with largest mass</td>
</tr>
<tr>
<td>m^m_{vis}</td>
<td>sum yj of pair of closest b-tagged jets</td>
</tr>
<tr>
<td>m^m_{vis}</td>
<td>sum yj of pair of closest b-tagged jets</td>
</tr>
<tr>
<td>m^m_{vis}</td>
<td>largest ∆ between any two jets</td>
</tr>
<tr>
<td>m^m_{vis}</td>
<td>average ∆ between b-tagged jets</td>
</tr>
<tr>
<td>m^m_{vis}</td>
<td>average ∆ between two jets</td>
</tr>
<tr>
<td>m^m_{vis}</td>
<td>average ∆ between a jet and a b-tagged jets</td>
</tr>
<tr>
<td>m^m_{vis}</td>
<td>minimal ∆ between any two jets</td>
</tr>
<tr>
<td>m^m_{vis}</td>
<td>minimal ∆ between any two b-tagged jets</td>
</tr>
<tr>
<td>m^m_{vis}</td>
<td>minimal ∆ between lepton and b-tagged jet</td>
</tr>
<tr>
<td>m^m_{vis}</td>
<td>largest ∆ between any two jets</td>
</tr>
<tr>
<td>m^m_{vis}</td>
<td>sum 1/2(λ_i + λ_j), with λ_i the eigenvalues of the momentum tensor computed with jets</td>
</tr>
<tr>
<td>m^m_{vis}</td>
<td>sum 1/2(λ_i + λ_j), with λ_i the eigenvalues of the momentum tensor computed with b-tagged jets</td>
</tr>
<tr>
<td>m^m_{vis}</td>
<td>sum 1/2(λ_i + λ_j), with λ_i the eigenvalues of the momentum tensor computed with b-tagged jets</td>
</tr>
<tr>
<td>m^m_{vis}</td>
<td>scalar of the jet and lepton yj divided by the sum of the energies of all jets and leptons</td>
</tr>
<tr>
<td>m^m_{vis}</td>
<td>scalar sum of the b-tagged jet yj divided by the sum of the energies of all b-tagged jets</td>
</tr>
<tr>
<td>m^m_{vis}</td>
<td>0th Fox–Wolfram moment computed with all jets</td>
</tr>
<tr>
<td>m^m_{vis}</td>
<td>ratio H1/H0 of 0th and first Fox–Wolfram moment computed with all jets</td>
</tr>
</tbody>
</table>

CMS-PAS-HIG-18-030
## ANN and BDT Hyperparameters

<table>
<thead>
<tr>
<th></th>
<th>(4 jets, ≥ 3 b-tags)</th>
<th>(5 jets, ≥ 3 b-tags)</th>
<th>(≥ 6 jets, ≥ 3 b-tags)</th>
</tr>
</thead>
<tbody>
<tr>
<td>hidden layers</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>nodes per hidden layer</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>loss function</td>
<td>cross-entropy</td>
<td>cross-entropy</td>
<td>cross-entropy</td>
</tr>
<tr>
<td>dropout percentage</td>
<td>0.5</td>
<td>0.5</td>
<td>0.5</td>
</tr>
<tr>
<td>L2 regularisation</td>
<td>$10^{-5}$</td>
<td>$10^{-5}$</td>
<td>$10^{-5}$</td>
</tr>
<tr>
<td>batchsize</td>
<td>5000</td>
<td>5000</td>
<td>5000</td>
</tr>
<tr>
<td>optimiser</td>
<td>Adam($10^{-4}$)</td>
<td>Adam($10^{-4}$)</td>
<td>Adam($10^{-4}$)</td>
</tr>
<tr>
<td>activation function</td>
<td>ELU</td>
<td>ELU</td>
<td>ELU</td>
</tr>
<tr>
<td>last activation</td>
<td>softmax</td>
<td>softmax</td>
<td>softmax</td>
</tr>
<tr>
<td>early stopping percentage</td>
<td>2%</td>
<td>2%</td>
<td>2%</td>
</tr>
<tr>
<td>early stopping min epochs</td>
<td>50</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td>number of input variables</td>
<td>14</td>
<td>15</td>
<td>11</td>
</tr>
<tr>
<td>number of network parameters</td>
<td>22306</td>
<td>22406</td>
<td>22106</td>
</tr>
</tbody>
</table>

### SL Channel:
#### ANN Hyperparameters

<table>
<thead>
<tr>
<th></th>
<th>(3 jets, ≥ 2 b-tags)</th>
<th>(3 jets, ≥ 3 b-tags)</th>
<th>(≥ 4 jets, ≥ 2 b-tags)</th>
<th>(≥ 4 jets, ≥ 3 b-tags)</th>
<th>(≥ 4 jets, ≥ 4 b-tags)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N_{\text{trees}}$</td>
<td>803</td>
<td>1137</td>
<td>1042</td>
<td>1451</td>
<td>1313</td>
</tr>
<tr>
<td>Shrinkage</td>
<td>0.024</td>
<td>0.035</td>
<td>0.035</td>
<td>0.050</td>
<td>0.042</td>
</tr>
<tr>
<td>Bagging Fraction</td>
<td>0.71</td>
<td>0.29</td>
<td>0.22</td>
<td>0.34</td>
<td>0.33</td>
</tr>
<tr>
<td>$N_{\text{cuts}}$</td>
<td>20</td>
<td>62</td>
<td>55</td>
<td>30</td>
<td>51</td>
</tr>
<tr>
<td>Depth</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>ROC AUC</td>
<td>0.78</td>
<td>0.78</td>
<td>0.78</td>
<td>0.78</td>
<td>0.81</td>
</tr>
</tbody>
</table>

### DL Channel:
#### BDT Hyperparameters

---

CMS-PAS-HIG-18-030
MEM Discriminant

Discriminant (per event) based on process matrix elements combined with reconstruction-level information

\[
w(\vec{y} | \mathcal{H}) = \sum_{i=1}^{N_C} \int \frac{dx_a dx_b}{2x_a x_b s} \int \prod_{k=1}^{8} \left( \frac{d^3 p_k}{(2\pi)^3 2E_k} \right) (2\pi)^4 \delta(E, z) \left( p_a + p_b - \sum_{k=1}^{8} p_k \right) \mathcal{R}(x, y) \left( \vec{p}_T, \sum_{k=1}^{8} p_k \right)
\]

\[
x g(x_a, \mu_F) g(x_b, \mu_F) | \mathcal{M}(p_a, p_b, p_1, \ldots, p_8) |^2 W(\vec{y}, \vec{p})
\]
Parton density functions
LO scattering amplitude (Open Loops)
Detector transfer function

Signal/background PDF:
\[
P_{S/B} = \frac{w(\vec{y}|t\bar{t}H)}{w(\vec{y}|t\bar{t}H) + k_{S/B} w(\vec{y}|t\bar{t} + b\bar{b})}
\]
\((t\bar{t} + b\bar{b} \text{ taken as the background hypothesis})\)
Single Lepton Channel – 4 jets, ≥ 3 btags

CMS Simulation Preliminary

SL (4 jets, ≥ 3 b tags) Pre-fit expectation

Pre-fit Distributions

CMS-PAS-HIG-18-030

Post-fit Distributions
Single Lepton Channel – 5 jets, ≥ 3 btags

**CMS Simulation Preliminary**

SL (5 jets, ≥ 3 b tags) Pre-fit expectation

- **t\(\bar{t}\)+lf node**
  - S/B = 0.0035, S/B = 0.42

- **t\(\bar{t}\)+cE node**
  - S/B = 0.0053, S/B = 0.31

- **t\(\bar{t}\)+b node**
  - S/B = 0.0070, S/B = 1.33

- **t\(\bar{t}\)+2b node**
  - S/B = 0.0144, S/B = 0.40

- **t\(\bar{t}\)+bb node**
  - S/B = 0.0026, S/B = 1.35

- **t\(\bar{t}\)H node**
  - S/B = 0.0027, S/B = 1.00

- **Other Bkg**

**CMS-PAS-HIG-18-030**

Pre-fit Distributions

Post-fit Distributions
Single Lepton Channel – ≥ 6 jets, ≥ 3 btags

CMS Simulation Preliminary

SL (6 jets, ≥ 3 b tags) Pre-fit expectation

Pre-fit Distributions

Post-fit Distributions

CMS-PAS-HIG-18-030
Dilepton Channel

CMS Simulation Preliminary

Pre-fit expectation

Pre-fit Distributions

Post-fit Distributions

CMS-PAS-HIG-18-030
Fully Hadronic Channel

**Pre-fit Distributions**

- **CMS Preliminary**
  - FH (7 jets, 3 b tags)
    - $S/B = 0.0237$, $S/B = 0.76$
    - FH (8 jets, 3 b tags)
      - $S/B = 0.0246$, $S/B = 0.76$
    - FH ($> 9$ jets, 3 b tags)
      - $S/B = 0.0235$, $S/B = 0.88$

**Post-fit Distributions**

- Pre-fit expectation
  - $tH$
  - Multijet
  - $t+lf$
  - $t+lf+b$
  - $t+lf+2b$
  - $t+lf+b+5$
  - Other Bkg

**CMS-PAS-HIG-18-030**
# Results

<table>
<thead>
<tr>
<th>Events / Bin</th>
<th>Preliminary CMS (2016) + 41.5 fb⁻¹ (13 TeV)</th>
<th>Preliminary CMS (2017) (13 TeV)</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Event Category</th>
<th>Data / Bkg.</th>
<th>Background</th>
<th>Signal (μ = 1.15)</th>
<th>SM (μ = 1)</th>
</tr>
</thead>
</table>

| FH 3 b-tags | 1.36 ± 0.57 ± 0.36 | +1.68 ± 3.15 | 0.3 σ | (0.2 σ) |
| FH 4 b-tags | −1.54 ± 1.41 ± 1.45 | +0.91 ± 1.08 | −0.90 ± 1.13 | | (0.7 σ) |
| FH combined | −1.69 ± 1.43 ± 1.47 | +0.83 ± 1.16 | −0.83 ± 1.22 | | (0.7 σ) |
| SL 4 jets   | 1.73 ± 2.25 ± 2.21 | +0.88 ± 2.07 | −0.87 ± 2.04 | | 0.8 σ | (0.5 σ) |
| SL 5 jets   | 0.73 ± 0.98 ± 0.97 | +0.47 ± 0.86 | −0.46 ± 0.86 | | 0.8 σ | (1.0 σ) |
| SL ≥ 6 jets | 2.05 ± 0.76 ± 0.69 | +0.31 ± 0.69 | −0.31 ± 0.62 | | 3.0 σ | (1.6 σ) |
| SL combined | 1.84 ± 0.62 ± 0.56 | +0.26 ± 0.56 | −0.26 ± 0.50 | | 3.3 σ | (1.9 σ) |
| DL 3 jets   | −2.35 ± 2.40 ± 2.65 | +2.13 ± 3.85 | −2.06 ± 1.66 | | — | (0.2 σ) |
| DL ≥ 4 jets | 1.57 ± 1.02 ± 0.98 | +0.55 ± 0.86 | −0.53 ± 0.82 | | 1.6 σ | (1.0 σ) |
| DL combined | 1.62 ± 0.90 ± 0.85 | +0.50 ± 0.76 | −0.48 ± 0.70 | | 1.9 σ | (1.2 σ) |
| FH+SL+DL combined | 1.49 ± 0.44 ± 0.40 | +0.21 ± 0.39 | −0.20 ± 0.35 | | 3.7 σ | (2.6 σ) |
| FH+SL+DL combined 2016+2017 | 1.15 ± 0.32 ± 0.29 | +0.15 ± 0.28 | −0.15 ± 0.25 | | 3.9 σ | (3.5 σ) |

Bins of the final discriminants as used in the fit, reordered by the pre-fit expected signal-to-background ratio (S/B). Each of the shown bins includes multiple bins of the final discriminants with similar S/B.
### Systematic Uncertainties

<table>
<thead>
<tr>
<th>Source</th>
<th>Type</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Integrated luminosity</td>
<td>rate</td>
<td>Signal and all backgrounds</td>
</tr>
<tr>
<td>Lepton identification/isolation</td>
<td>shape</td>
<td>Signal and all backgrounds</td>
</tr>
<tr>
<td>Trigger efficiency</td>
<td>shape</td>
<td>Signal and all backgrounds</td>
</tr>
<tr>
<td>Trigger prefiring correction</td>
<td>rate</td>
<td>Signal and all backgrounds</td>
</tr>
<tr>
<td>Pileup</td>
<td>shape</td>
<td>Signal and all backgrounds</td>
</tr>
<tr>
<td>Jet energy scale</td>
<td>shape</td>
<td>Signal and all backgrounds</td>
</tr>
<tr>
<td>Jet energy resolution</td>
<td>shape</td>
<td>Signal and all backgrounds</td>
</tr>
<tr>
<td>b tag b fraction</td>
<td>shape</td>
<td>Signal and all backgrounds</td>
</tr>
<tr>
<td>b tag b stats (linear)</td>
<td>shape</td>
<td>Signal and all backgrounds</td>
</tr>
<tr>
<td>b tag b stats (quadratic)</td>
<td>shape</td>
<td>Signal and all backgrounds</td>
</tr>
<tr>
<td>b tag l fraction</td>
<td>shape</td>
<td>Signal and all backgrounds</td>
</tr>
<tr>
<td>b tag l stats (linear)</td>
<td>shape</td>
<td>Signal and all backgrounds</td>
</tr>
<tr>
<td>b tag l stats (quadratic)</td>
<td>shape</td>
<td>Signal and all backgrounds</td>
</tr>
<tr>
<td>b tag charm (linear)</td>
<td>shape</td>
<td>Signal and all backgrounds</td>
</tr>
<tr>
<td>b tag charm (quadratic)</td>
<td>shape</td>
<td>Signal and all backgrounds</td>
</tr>
<tr>
<td>QGL reweighting</td>
<td>shape</td>
<td>Signal and all backgrounds</td>
</tr>
<tr>
<td>TT_{boson} correction</td>
<td>shape</td>
<td>QCD multiplet estimate</td>
</tr>
<tr>
<td>TF^{loose} correction</td>
<td>shape</td>
<td>QCD multiplet estimate</td>
</tr>
<tr>
<td>TF^{tight} reweighting</td>
<td>shape</td>
<td>QCD multiplet estimate</td>
</tr>
<tr>
<td>Multijet normalisation</td>
<td>rate</td>
<td>QCD multiplet estimate</td>
</tr>
<tr>
<td>Remor. /fact. scales (fD)</td>
<td>rate</td>
<td>Scale uncertainty of NLO f_{tH} prediction</td>
</tr>
<tr>
<td>Remor. /fact. scales (f)</td>
<td>rate</td>
<td>Scale uncertainty of NNLO f_{tH} prediction</td>
</tr>
<tr>
<td>f_{tH} cross sections</td>
<td>rate</td>
<td>Additional 50% rate uncertainty of f_{tH} predictions</td>
</tr>
<tr>
<td>Remor. /fact. scales (f)</td>
<td>rate</td>
<td>Scale uncertainty of NLO single t prediction</td>
</tr>
<tr>
<td>Remor. /fact. scales (V)</td>
<td>rate</td>
<td>Scale uncertainty of NNLO W and Z prediction</td>
</tr>
<tr>
<td>Remor. /fact. scales (VV)</td>
<td>rate</td>
<td>Scale uncertainty of NLO diboson prediction</td>
</tr>
<tr>
<td>PDF (gg)</td>
<td>rate</td>
<td>PDF uncertainty for gg initiated processes except f_{tH}</td>
</tr>
<tr>
<td>PDF (g f_{tH})</td>
<td>rate</td>
<td>PDF uncertainty for f_{tH}</td>
</tr>
<tr>
<td>PDF (gg f_{tH})</td>
<td>rate</td>
<td>PDF uncertainty of gg initiated processes (f_{tW,W,Z})</td>
</tr>
<tr>
<td>PDF (gg)</td>
<td>rate</td>
<td>PDF uncertainty of gg initiated processes (single t)</td>
</tr>
<tr>
<td>PDF shape variations (f_{tH}, f)</td>
<td>shape</td>
<td>Based on the NNPDF replicas, same for f_{tH} and additional jet flavours</td>
</tr>
<tr>
<td>$\mu_R$ scale (f)</td>
<td>shape</td>
<td>Renormalisation scale uncertainty of the f_{tH} ME generator (Powheg), same for additional jet flavours</td>
</tr>
<tr>
<td>$\mu_F$ scale (f)</td>
<td>shape</td>
<td>Factorisation scale uncertainty of the f_{tH} ME generator (Powheg), same for additional jet flavours</td>
</tr>
<tr>
<td>PS scale: ISR (f)</td>
<td>shape</td>
<td>Initial state radiation uncertainty of the PS (for f_{tH} events), same for additional jet flavours</td>
</tr>
<tr>
<td>PS scale: PSR (f)</td>
<td>shape</td>
<td>Final state radiation uncertainty of the PS (for f_{tH} events), same for additional jet flavours</td>
</tr>
<tr>
<td>ME-PS matching (f)</td>
<td>rate</td>
<td>NLO ME to PS matching, kfactor (? ) (for f_{tH} events), independent for additional jet flavours</td>
</tr>
<tr>
<td>Underlying event (f)</td>
<td>rate</td>
<td>Underlying event (for f_{tH} events), independent for additional jet flavours</td>
</tr>
<tr>
<td>Bin-by-bin event count</td>
<td>shape</td>
<td>Statistical uncertainty of the signal and background prediction due to the limited sample size</td>
</tr>
</tbody>
</table>

### Impacts for 2016 + 2017

**CMS Preliminary**

![CMS-PAS-HIG-18-030](image-url)

- $\hat{\mu} = 1.15^{+0.32}_{-0.29}$

**Impacts for 2016 + 2017**

- $f_{tH}$ (renor./fact. scales)
- $f_{tH}$ cross section (50%) (2017)
- $f_{tH}$ cross section (50%) (2016)
- multijet (norm) 9 Jets, 4 b tags (2016)
- $f_{tH}$ cross section (50%) (2016)
- multijet (norm) 8 Jets, 4 b tags (2016)
- multijet (norm) 7 Jets, 3 b tags (2016)
- multijet (norm) 7 Jets, 3 b tags (2017)
- multijet (norm) 4 Jets, 4 b tags (2016)
- multijet (norm) 4 Jets, 3 b tags (2016)
- multijet (norm) 3 Jets, 3 b tags (2016)
- multijet (norm) 2 Jets, 3 b tags (2016)
- multijet (norm) 4 b tags (2016)
- $f_{tH}$ cross section (50%) (2016)
- $f_{tH}$ cross section (50%) (2017)
- b tag charm (linear) (2016)
- b tag charm (linear) (2017)
- $f_{tH}$ cross section (50%) (2017)
- $f_{tH}$ cross section (50%) (2016)
- b tag charm (linear) (2017)

**CMS-PAS-HIG-18-030**

**July 30, 2019**

**DPF 2019**

**34/19**