CMS Searches With Boosted Top Quarks

Reyer Band
University of California, Davis
On Behalf of the CMS Collaboration

11th BOOST 2019
25 July 2019
CMS Searches With Boosted Top Quarks... Using Deep Neural Networks?

Reyer Band
University of California, Davis
On Behalf of the CMS Collaboration

11th BOOST 2019
25 July 2019
Boosted Tops At CMS

Many theories of new physics contain heavy particles which couple to top quarks

- A few recent results from CMS:
  - Vector-Like Quark Pair Production
  - A Search For Top Superpartners

For other CMS results, see Christine’s talk from BOOST18
\( \tilde{t}\tilde{t} \) - Signal

- SUSY top super partner \( \tilde{t} \) may decay to a top quark
- Look for events with a single lepton from a \( W \), itself either from a top or chargino
- MET expected from neutralino and leptonic \( W \)
$\tilde{t}\tilde{t}$ - Signal

- SUSY top super partner $\tilde{t}$ may decay to a top quark

- Look for events with a single lepton from a $W$, itself either from a top or chargino

- MET expected from neutralino and leptonic $W$

- Hadronic decay of the top a key signature in two decay modes
  - If $m_{\tilde{t}} \gg m_{\tilde{\chi}} + m_t$ ("uncompressed"), top quarks will be boosted
  - If not boosted, identify the resolved jets.

- Search performed in the **full Run 2** dataset (137 fb$^{-1}$)
Tag tops as either **merged** or **resolved**, boost depends on mass splitting

**Resolved:**
- Must be at least 3 (AK4) jets
- Tag with **DeepResolved** network - trained on jet quantities and trijet variables
Tag tops as either **merged** or **resolved**, boost depends on mass splitting

**Resolved:**
- Must be at least 3 (AK4) jets
- Tag with **DeepResolved** network - trained on jet quantities and trijet variables

**Merged:**
- Heavily boosted top quark identified as one AK8 jet
- Tag with **DeepAK8** network. See **Meenakshi Narain’s talk** for all the details

First search to make use of these neural networks
**ttbar - Background**

- **Lost lepton** from dilepton event
  - Mostly from ttbar->2l
  - Measure transfer factor in MC by inverting lepton veto (right)
  - Reduced by t

- **W lepton** (besides top)
  - W+Jets, WW with 1 lepton
  - Genuine MET from neutrino

- **tt semi-leptonic**
  - High $M_T$ requirement kills this background
  - Predicted from simulation

- **Z->MET**
  - Produced with W or ttbar
  - May have actual hadronic top

\[
N_{SR}^{\text{lost-}l} = N_{CR} \times \frac{M_{SR}^{\text{lost-}l}}{M_{CR}^{ll}}
\]

**CMS Preliminary**

- 137 fb\(^{-1}\) (13 TeV)

- Observed
- Lost lepton
- 1/ from top
- 1/ not from top
- $Z\rightarrow \nu\bar{\nu}$

**Graph**

- Events / 25 GeV

**Document Reference**

SUS-19-009
**t̃t̃ - t_{mod}: Modified topness**

χ²-like variable

- t̃t̃→2l system with missing lepton
  - Missing neutrino and a W
  - System under-constrained
- Expect low S (or t_{mod}) from t̃t̃ and high from signal

\[ t_{mod} = \ln(\min S) \]
$\tilde{t}\tilde{t} - t_{\text{mod}}$: Modified topness

- $\chi^2$-like variable
- $t\bar{t} \rightarrow 2l$ system with missing lepton
  - Missing neutrino and a $W$
  - System under-constrained
- Expect low $S$ (or $t_{\text{mod}}$) from $t\bar{t}$ and high from signal

$$t_{\text{mod}} = \ln(\text{min } S)$$

$$S = \frac{\left(m_W^2 - (p_\nu + p_\ell)^2\right)^2}{a_W^4} + \frac{\left(m_t^2 - (p_b + p_W)^2\right)^2}{a_t^4}$$

- $W$ mass constraint from visible lepton
- Top mass constraint from missing lepton
\( \tilde{t} \tilde{t} - \text{t}_{\text{mod}} \): Modified topness

- **χ^2-like variable**
- \( \tilde{t} \tilde{t} \rightarrow 2l \) system with missing lepton
  - Missing neutrino and a W
  - System under-constrained
- Expect low S (or \( t_{\text{mod}} \)) from \( \tilde{t} \tilde{t} \) and high from signal

\[ t_{\text{mod}} = \ln(\min S) \]

\[ S = \frac{(m_W^2 - (p_\nu + p_\ell)^2)^2}{a_W^4} + \frac{(m_t^2 - (p_b + p_W)^2)^2}{a_t^4} \]

- W mass constraint from visible lepton
- Top mass constraint from missing lepton

**SUS-19-009**
- Results

137 fb⁻¹ (13 TeV)

<table>
<thead>
<tr>
<th>N_j</th>
<th>t_{mod}</th>
<th>M_{fb} [GeV]</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>2–3</td>
<td>&gt; 10</td>
</tr>
<tr>
<td>B</td>
<td>2–3</td>
<td>&gt; 10</td>
</tr>
<tr>
<td>C</td>
<td>≥ 4</td>
<td>≤ 0</td>
</tr>
<tr>
<td>D</td>
<td>≥ 4</td>
<td>≤ 0</td>
</tr>
<tr>
<td>E</td>
<td>≥ 4</td>
<td>0–10</td>
</tr>
<tr>
<td>F</td>
<td>≥ 4</td>
<td>0–10</td>
</tr>
<tr>
<td>G</td>
<td>≥ 4</td>
<td>&gt; 10</td>
</tr>
<tr>
<td>H</td>
<td>≥ 4</td>
<td>&gt; 10</td>
</tr>
</tbody>
</table>

X0: Inclusive
X1: Untagged
X2: Boosted top
X3: Resolved top

| | |
|---|---|---|
| I: N_j ≥ 5, N_{b,med} ≥ 1 |
| J: N_j ≥ 3, N_{b,soft} ≥ 1 |
\( \tilde{t}\tilde{t} - \text{Results} \)

**CMS Preliminary**

137 fb\(^{-1}\) (13 TeV)

\( pp \rightarrow \tilde{t}\tilde{t} \rightarrow t\tilde{\chi}_1 \)

Approx. NNLO+NNLL exclusion

- **Observed \( \pm 1 \sigma_{\text{theory}} \)**
- **Expected \( \pm 1 \sigma_{\text{experiment}} \)**

\( m_{\tilde{\chi}^0} - m_{\tilde{t}} - m_\chi < 25 \text{ GeV} \)

\( \sigma_{\text{95\% CL upper limit on cross section [pb]}} \)

\( m_{\tilde{t}} \) vs. \( m_{\tilde{\chi}^0} \) graph
Results

CMS Preliminary 137 fb⁻¹ (13 TeV)

pp → \tilde{t}\tilde{t}, \tilde{t} → b \tilde{\chi}_1^0 → b W^+ \tilde{\chi}_1^0 or \tilde{t} → t \tilde{\chi}_1^0

Approx. NNLO+NNLL exclusion

- Observed ± 1 \sigma_{theory}
- Expected ± 1 \sigma_{experiment}

m_{\tilde{\chi}_1^0} = (m_t + m_{\tilde{\chi}_1^0})/2

95% CL upper limit on cross section [pb]

SUS-19-009
Vector-like quarks (VLQs) are not constrained by observed Higgs production cross section

- Left- and right-handed components transform equally under SU(2) (‘vector-like’)

- Consider pair production in all-hadronic final states

Two searches performed with different tagging approaches:

- One targeting $T \rightarrow bW$ decay

- One targeting all modes, utilizing a novel neural net based tagger: “Boosted Event Shape Tagger” (BEST)

- Perform these searches in the 2016 dataset (35.9 fb$^{-1}$)
VLQ - BEST Procedure

- Hypothesize different particles as origin of jet: top, H, Z, W
  - Jet constituents should be isotropic in the rest frame of that particle
  - Use particle masses to apply different boosts to jet constituents
  - In each reference frame, calculate event shape variables

- Also use some lab frame inputs, such as soft drop mass and CSV

- Six possible classifications:
  - $t, W, Z, H, b, \text{light [u/d/s/c/g]}$
VLQ - BEST Inputs

- **Longitudinal Asymmetry** - momentum balance along jet axis in rest frame
  - Balanced momentum ($A_L = 0$) more likely in correct frame
  - Calculate value in each frame

\[ A_L = \frac{\sum p_L^{jet}}{\sum p^{jet}} \]

- CMS (13 TeV)
  - 0.5 TeV < Jet $p_T$ < 1.5 TeV
  - Jet $m_{SD} > 10$ GeV
  - jet: Top Jets, Z Jets, W Jets, H Jets, b Jets, Light Jets
**VLQ - BEST Inputs**

- **Longitudinal Asymmetry** - momentum balance along jet axis in rest frame
  - Balanced momentum ($A_L = 0$) more likely in correct frame
  - Calculate value in each frame

```
0.5 TeV < Jet $p_T < 1.5$ TeV
Jet $m_{SD} > 10$ GeV
```

- **Soft Drop Mass** - Well known variable for identifying jets from heavy objects
  - Small cut on $m_{SD} > 10$ GeV to reduce light background
Train fully connected deep neural network on 500K simulated jets

- Require jets with $p_T > 500$ GeV, to ensure decay products are fully merged

- Achieve good discrimination between particle types

- Tagging rates found to be close to simulation
- Neural Net-based analysis signal regions include exactly 4 jets
  - Count multiplicities of objects according to BEST classification
  - Unique set of $(N_t, N_H, N_W, N_Z, N_b, N_j)$, sum of the $N_i = 4$
- 126 independent signal regions

### CMS

<table>
<thead>
<tr>
<th>Events per category</th>
<th>Observed</th>
<th>Z+jets</th>
<th>Multijet (DD)</th>
<th>t(\bar{t})+jets</th>
<th>W+jets</th>
<th>t(\bar{t})V(\bar{t})f(\bar{f})</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data / BG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

35.9 fb\(^{-1}\) (13 TeV)
Neural Net-based analysis signal regions include exactly 4 jets
  - Count multiplicities of objects according to BEST classification
    - Unique set of \((N_t, N_H, N_W, N_Z, N_b, N_j)\), sum of the \(N_i = 4\)

126 independent signal regions

Cut-based analysis counts multiplicities of \(W, b\) tags, requiring > 1 of each tag

- Tag \(W\) jets via cuts on Soft Drop Mass and \(\tau_{21}\)
- Tag \(b\) jets with cuts on CSV discriminant

4 independent signal regions

In both analyses, use \(H_T\) distribution to discriminate signal from background
VLQ - Background Estimate

- Dominant background is QCD
  - Estimate this from observed data
- **Cut-based** analysis uses ABCD method
  - $H_T$, VLQ mass difference sideband
- **NN-based** analysis measures a tagging rate in independent 3 jet sample
  - 3-jet events entirely QCD
  - Apply rates to whole 4-jet sample to get QCD distribution per category
  - Other backgrounds taken from simulation: $W$+jets, $t\bar{t}$

\[
\epsilon_X(p_T) = \frac{\text{Number of jets with BEST class } X}{\text{Number of jets}}
\]

\[
R = \sum_{\text{events}} [r] = \sum_{\text{events}} \left[ \sum_{\text{perms}} \left( \prod_{i=1}^{4} \epsilon_{X_i}(p_T(i)) \right) \right]
\]

35.9 fb$^{-1}$ (13 TeV)
VLQ Results — TT Limits By Branching Fraction

- Scan limits over the branching fractions of the VLQ
  - BEST sensitive to $T\to tH$ and $T\to tZ$

![Scan limits over the branching fractions of the VLQ](image)

- Expected
- Observed

35.9 fb$^{-1}$ (13 TeV)

**CMS**

- Observed T quark mass limit [GeV]
  - NN analysis

- Expected T quark mass limit [GeV]
  - NN analysis

**CMS**

- Expected
- Observed

B2G-18-005
Scan limits over the branching fractions of the VLQ

- Cut-based analysis performs better in T->bW corner
Summary

- Several new searches by CMS using boosted tops have been presented
  - Both make use of new tagging techniques based on deep neural nets
  - More advances in this field being made, look forward to even more results at the next BOOST conference!

- See Meenakshi Narain’s talk from this conference for more information on CMS tagging techniques featuring top quarks

Vector-like quark pair production

- X5/3 → tW(LH)
- X5/3 → tW(RH)
- YY, B(Y → bW) = 100%
- BB doublet
- BB singlet
- BB, B(B → bH) = 100%
- BB, B(B → tW) = 100%
- TT doublet
- TT singlet
- TT, B(T → bH) = 100%
- TT, B(T → tZ) = 100%
- TT, B(T → bW) = 100%
- QQ doublet
- QQ singlet
- QQ, B(Q → qH) = 100%
- QQ, B(Q → qZ) = 100%
- QQ, B(Q → qW) = 100%

- √s = 13 TeV
- √s = 8 TeV

CMS, EPS-HEP 2019
Backup Material
SUS - Pre-selection

- Require exactly one isolated lepton
- Veto extra isolated leptons with $p_T > 5$ GeV
- Veto tracks (potentially from taus) with $p_T > 10$ GeV
- Medium b-tag requirement (special soft $>20$ GeV tag for compressed spectra)
- $M_T > 150$
- MET $> 250$, with a separation from leading jets $\Delta\Phi > 0.8$ (0.5 for compressed)
- At least 2 AK4 jets
SUS - Signal Regions

- Uncompressed spectra:
  - $m_\tilde{t} > m_t + m_\chi$
  - Split categories by:
    - Number of jets
    - $t_{\text{mod}}$
    - $M_{lb}$
    - MET
  - 39 categories

- Compressed spectra:
  - $m_\tilde{t} \sim m_t + m_\chi$ or $\sim m_W + m_b + m_\chi$
  - Softer b jets
  - Leading jet from ISR
  - Veto high $p_T$ lepton
  - Split by MET bin
  - Number of jets: 3 or 5
  - 10 categories

### Uncompressed

<table>
<thead>
<tr>
<th>Label</th>
<th>$N_j$</th>
<th>$t_{\text{mod}}$</th>
<th>$M_{lb}$ [GeV]</th>
<th>Top tagging category</th>
<th>$E_{T,\text{miss}}$ bins [GeV]</th>
</tr>
</thead>
<tbody>
<tr>
<td>A0</td>
<td>2–3</td>
<td>$\geq 10$</td>
<td>$&lt; 175$</td>
<td>–</td>
<td>[600, 750, +∞]</td>
</tr>
<tr>
<td>A1</td>
<td></td>
<td></td>
<td>U</td>
<td>[350, 450, 600]</td>
<td></td>
</tr>
<tr>
<td>A2</td>
<td></td>
<td></td>
<td>M</td>
<td>[250, 600]</td>
<td></td>
</tr>
<tr>
<td>B</td>
<td></td>
<td>$\geq 175$</td>
<td>–</td>
<td>[250, 450, 700, +∞]</td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>&lt; 0</td>
<td>$&lt; 175$</td>
<td>–</td>
<td>[350, 450, 550, 650, 800, +∞]</td>
<td></td>
</tr>
<tr>
<td>D</td>
<td></td>
<td>$\geq 175$</td>
<td>–</td>
<td>[250, 350, 450, 600, +∞]</td>
<td></td>
</tr>
<tr>
<td>E0</td>
<td></td>
<td></td>
<td>U</td>
<td>[450, 600, +∞]</td>
<td></td>
</tr>
<tr>
<td>E1</td>
<td></td>
<td>$0–10$</td>
<td>$&lt; 175$</td>
<td>U</td>
<td>[250, 350, 450]</td>
</tr>
<tr>
<td>E2</td>
<td></td>
<td></td>
<td>M</td>
<td>[250, 350, 450]</td>
<td></td>
</tr>
<tr>
<td>E3</td>
<td></td>
<td></td>
<td>R</td>
<td>[250, 350, 450]</td>
<td></td>
</tr>
<tr>
<td>F</td>
<td>$\geq 4$</td>
<td>$\geq 175$</td>
<td>–</td>
<td>[250, 350, 450, +∞]</td>
<td></td>
</tr>
<tr>
<td>G0</td>
<td></td>
<td></td>
<td>U</td>
<td>[450, 550, 750, +∞]</td>
<td></td>
</tr>
<tr>
<td>G1</td>
<td></td>
<td>$\geq 10$</td>
<td>$&lt; 175$</td>
<td>U</td>
<td>[250, 350, 450]</td>
</tr>
<tr>
<td>G2</td>
<td></td>
<td></td>
<td>M</td>
<td>[250, 350, 450]</td>
<td></td>
</tr>
<tr>
<td>G3</td>
<td></td>
<td></td>
<td>R</td>
<td>[250, 350, 450]</td>
<td></td>
</tr>
<tr>
<td>H</td>
<td></td>
<td></td>
<td>–</td>
<td>[250, 500, +∞]</td>
<td></td>
</tr>
</tbody>
</table>
**SUS - $M_{lb}$**

- Invariant mass of lepton+b quark in a $tt$ system is bounded

\[ M_{lb} \leq M_t \sqrt{1 - \frac{M_W^2}{M_t^2}} \]

- Mass constraint applies to signal as well, in T2tt channel

- Violated by backgrounds from other processes, like W+jets, and other signal channels
SUS - Control Regions

- Control region samples used for the background estimation
- MET distribution used for the dilepton transfer factor
  - Require presence of a second isolated lepton with $p_T > 10$ GeV
  - Add trailing lepton $p_T$ to the MET
- $W$+jets estimated from the 0b control region
  - The “b” jet used for $M_{lb}$ distribution is jet with highest deepCSV value

![Graph 1: 2 Leptons](image1)

![Graph 2: 0b](image2)
## SUS - Systematics

<table>
<thead>
<tr>
<th>Source</th>
<th>Signal</th>
<th>Lost lepton</th>
<th>$1\ell$ background</th>
<th>$Z \rightarrow \nu\bar{\nu}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data statistical uncertainty</td>
<td>—</td>
<td>5–50%</td>
<td>4–30%</td>
<td>—</td>
</tr>
<tr>
<td>Simulation statistical uncertainty</td>
<td>6–36%</td>
<td>3–68%</td>
<td>5–70%</td>
<td>4–41%</td>
</tr>
<tr>
<td>$t\bar{t}$ $E_T^{\text{miss}}$ modeling</td>
<td>—</td>
<td>3–50%</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>QCD scales</td>
<td>1–5%</td>
<td>0–3%</td>
<td>2–5%</td>
<td>1–40%</td>
</tr>
<tr>
<td>Parton distribution</td>
<td>—</td>
<td>0–4%</td>
<td>1–8%</td>
<td>1–12%</td>
</tr>
<tr>
<td>Pileup</td>
<td>1–5%</td>
<td>1–8%</td>
<td>0–5%</td>
<td>0–7%</td>
</tr>
<tr>
<td>Luminosity</td>
<td>2.3–2.5%</td>
<td>—</td>
<td>—</td>
<td>2.3–2.5%</td>
</tr>
<tr>
<td>$W + b$ cross section</td>
<td>—</td>
<td>—</td>
<td>20–40%</td>
<td>—</td>
</tr>
<tr>
<td>$Z \rightarrow \nu\nu$ estimate</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>5–10%</td>
</tr>
<tr>
<td>System recoil (ISR)</td>
<td>1–13%</td>
<td>0–3%</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Jet energy scale</td>
<td>2–24%</td>
<td>1–16%</td>
<td>1–34%</td>
<td>1–28%</td>
</tr>
<tr>
<td>$E_T^{\text{miss}}$ resolution</td>
<td>—</td>
<td>1–10%</td>
<td>1–5%</td>
<td>—</td>
</tr>
<tr>
<td>Trigger</td>
<td>2–3%</td>
<td>1–3%</td>
<td>—</td>
<td>2–3%</td>
</tr>
<tr>
<td>Lepton efficiency</td>
<td>3–4%</td>
<td>2–12%</td>
<td>—</td>
<td>1–2%</td>
</tr>
<tr>
<td>Merged top tagging efficiency</td>
<td>3–6%</td>
<td>—</td>
<td>—</td>
<td>5–10%</td>
</tr>
<tr>
<td>Resolved top tagging efficiency</td>
<td>5–6%</td>
<td>—</td>
<td>—</td>
<td>3–5%</td>
</tr>
<tr>
<td>b tagging efficiency</td>
<td>0–2%</td>
<td>0–1%</td>
<td>1–7%</td>
<td>1–10%</td>
</tr>
<tr>
<td>Soft b tagging efficiency</td>
<td>2–3%</td>
<td>0–1%</td>
<td>0–1%</td>
<td>0–5%</td>
</tr>
</tbody>
</table>
DeepAK8+BEST Performance In Simulation

Simulation Preliminary
Top quark tagging, $\mathbb{B} = 30\%$

- $300 < p_{T,\text{truth}} < 2000$ GeV, $|\eta_{\text{truth}}| < 2.4$
- $105 < m_{SD} < 210$ GeV
- $110 < m_{SD}^{\text{CA15}} < 210$ GeV
- $140 < m_{\text{HOTVR}} < 220$ GeV

CMS

Efficiency vs. $p_{T,\text{truth}}$ [GeV] for different algorithms and kinematic cuts.

Misidentification rate vs. $p_{T,\text{truth}}$ [GeV] for different algorithms and kinematic cuts.
DeepAK8 Top Tagging Validation

CMS Preliminary

35.9 fb⁻¹ (13 TeV)

single-µ sample
\( p_T (\text{AK8 jet}) > 200 \text{ GeV} \)
\( \eta (\text{AK8 jet}) < 2.4 \)

- Data
- \( \bar{t}t \)
- single-\( t \)
- QCD multijet
- W+jets
- ttV
- VV
- Bkg unc (stat)
- Bkg unc (stat+syst)

Events / bin

\( N_{\text{obs}} / N_{\text{exp}} \)

DeepAK8 (t vs. QCD)
BEST Top Tagging Validation

35.9 fb^{-1} (13 TeV)

CMS Preliminary

Events / bin

single-μ sample
p_T (AK8 jet)>200 GeV
\eta(AK8 jet)|<2.4

SM (Herwig)

Data
tt
single-t
QCD multijet
W+jets
ttV
VV
Bkg unc (stat)
Bkg unc (stat+syst)

N_{obs} / N_{exp}

0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1

BEST (t vs. QCD)
### VLQ - Event Selection - Cut Based

- Use tagging multiplicities of W’s and b’s to define signal regions.

- Suppress background by requiring multiple W/b tags and high hadronic activity:
  - $H_T > 1200$ GeV, where $H_T$ is the scalar sum of all jet’s transverse momenta
  - **Signal** requires > 0 of each tag.

- Signal-poor areas used for validating multijet QCD background estimation.

- **Sideband** used to extract a normalization scale factor.

<table>
<thead>
<tr>
<th>2W0b</th>
<th>2W1b</th>
<th>2W2b</th>
</tr>
</thead>
<tbody>
<tr>
<td>1W0b</td>
<td>1W1b</td>
<td>1W2b</td>
</tr>
<tr>
<td>0W0b</td>
<td>0W1b</td>
<td>0W2b</td>
</tr>
</tbody>
</table>

= 0 W tags, any b tags: Validation Regions
> 0 W tags, = 0 b tags: Sideband Regions
> 0 W tags, > 0 b tags: Signal Regions.
Final signal regions for cut-based analysis (post-fit).

- **2W2b**
  - (35.9 fb$^{-1}$ (13 TeV))
  - Observed:
    - Multijet (DD)
    - Tt+jets
    - W+jets
    - TT (0.8 TeV)
    - TT (1.2 TeV)
    - TT (1.6 TeV)
  - BG Statistical Error

- **2W1b**
  - (35.9 fb$^{-1}$ (13 TeV))
  - Observed:
    - Multijet (DD)
    - Tt+jets
    - W+jets
    - TT (0.8 TeV)
    - TT (1.2 TeV)
    - TT (1.6 TeV)
  - BG Statistical Error

- **1W2b**
  - (35.9 fb$^{-1}$ (13 TeV))
  - Observed:
    - Multijet (DD)
    - Tt+jets
    - W+jets
    - TT (0.8 TeV)
    - TT (1.2 TeV)
    - TT (1.6 TeV)
  - BG Statistical Error

- **1W1b**
  - (35.9 fb$^{-1}$ (13 TeV))
  - Observed:
    - Multijet (DD)
    - Tt+jets
    - W+jets
    - TT (0.8 TeV)
    - TT (1.2 TeV)
    - TT (1.6 TeV)
  - BG Statistical Error
VLQ - BEST - Signal Region Summary

» 126 signal regions, summarized here by inclusive BEST tag categories.
VLQ - BEST Architecture

- Feed-forward deep neural network with 3 hidden layers, 40 nodes each.
- Multi classification algorithm with six outputs.
- Node with highest output determines the jet label.
- 59 input features, 500,000 jets used to train the network.
- Use samples of high-pT boosted SM particles from high-mass resonance decays for training:
  - $Z' \rightarrow tt$, RSG-$\rightarrow HH$, etc.

```
59 input features → Dense 40 Nodes → Dense 40 Nodes → Dense 40 Nodes → top H W Z b u/d/s/c+g
```
DeepAK8 (Merged Tagger) Structure

- Trained on 40 million jets
- First 100 reconstructed particles in a jet used as inputs
  - 42 features per particle, such as basic kinematics, tracking info
- Up to 7 secondary vertices fed into parallel CNN
  - 15 features per SV, including displacement and quality
- Both CNNs are fed into one fully connected layer
- Discriminate between decay modes of heavy objects
## VLQ - BEST Variable List

<table>
<thead>
<tr>
<th>NN Input Quantities</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Sphericity ( (t, W, Z, H) )</td>
<td>Jet Soft-Drop Mass</td>
</tr>
<tr>
<td>Isotropy ( (t, W, Z, H) )</td>
<td>Jet ( \eta )</td>
</tr>
<tr>
<td>Aplanarity ( (t, W, Z, H) )</td>
<td>Jet ( \tau_{21} )</td>
</tr>
<tr>
<td>Thrust ( (t, W, Z, H) )</td>
<td>Jet ( \tau_{32} )</td>
</tr>
<tr>
<td>Jet Asymmetry ( A_L ) ( (t, W, Z, H) )</td>
<td>Jet Charge</td>
</tr>
<tr>
<td>Fox-Wolfram ( H_1/H_0 ) ( (t, W, Z, H) )</td>
<td>Maximum Subjet CSV Value</td>
</tr>
<tr>
<td>Fox-Wolfram ( H_2/H_0 ) ( (t, W, Z, H) )</td>
<td>Subjet 1 CSV Value</td>
</tr>
<tr>
<td>Fox-Wolfram ( H_3/H_0 ) ( (t, W, Z, H) )</td>
<td>Subjet 2 CSV Value</td>
</tr>
<tr>
<td>Fox-Wolfram ( H_4/H_0 ) ( (t, W, Z, H) )</td>
<td>( m_{13} ) ( (t,W,Z,H) )</td>
</tr>
<tr>
<td>( m_{12} ) ( (t,W,Z,H) )</td>
<td>( m_{1234} ) ( (t,W,Z,H) )</td>
</tr>
<tr>
<td>( m_{23} ) ( (t,W,Z,H) )</td>
<td></td>
</tr>
</tbody>
</table>
## VLQ - Systematic Uncertainties

<table>
<thead>
<tr>
<th>Source</th>
<th>Uncertainty</th>
<th>Contribution to:</th>
<th>Applies to samples:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diboson cross section</td>
<td>50%</td>
<td>✓</td>
<td>VV only</td>
</tr>
<tr>
<td>Rare top quark process cross sections</td>
<td>50%</td>
<td>✓</td>
<td>t(\bar{t}V), t(\bar{t})t(\bar{t})</td>
</tr>
<tr>
<td>Higgs boson cross section</td>
<td>50%</td>
<td>✓</td>
<td>H only</td>
</tr>
<tr>
<td>W+jets cross section</td>
<td>15%</td>
<td>✓</td>
<td>W+jets only</td>
</tr>
<tr>
<td>Z+jets cross section</td>
<td>15%</td>
<td>✓</td>
<td>Z+jets only</td>
</tr>
<tr>
<td>Integrated luminosity measurement</td>
<td>2.5%</td>
<td>✓</td>
<td>All simulation</td>
</tr>
<tr>
<td>Pileup reweighting</td>
<td>±1σ((p_T, \eta))</td>
<td>✓</td>
<td>All simulation</td>
</tr>
<tr>
<td>Jet energy scale</td>
<td>±1σ((\eta))</td>
<td>✓</td>
<td>All simulation</td>
</tr>
<tr>
<td>Jet energy resolution</td>
<td>±1σ((p_T))</td>
<td>✓</td>
<td>All simulation</td>
</tr>
<tr>
<td>Parton distribution functions</td>
<td>±1σ</td>
<td>✓</td>
<td>t(\bar{t}), VLQ</td>
</tr>
<tr>
<td>Renormalization and factorization scales</td>
<td>±1σ</td>
<td>✓</td>
<td>t(\bar{t}), VLQ</td>
</tr>
<tr>
<td>CSVv2 discriminant reshaping</td>
<td>(\delta(\text{wgt.}, \text{unwgt.}))</td>
<td>✓</td>
<td>All simulation</td>
</tr>
<tr>
<td>BEST classification fractions</td>
<td>±1σ((p_T))</td>
<td>✓</td>
<td>QCD multijet</td>
</tr>
<tr>
<td>BEST classification scale factor</td>
<td>5%</td>
<td>✓</td>
<td>All simulation</td>
</tr>
<tr>
<td>BEST misclassification scale factor</td>
<td>5%</td>
<td>✓</td>
<td>All simulation</td>
</tr>
<tr>
<td>Trigger</td>
<td>2%</td>
<td>✓</td>
<td>All simulation</td>
</tr>
<tr>
<td>W tag scale factor</td>
<td>±1σ</td>
<td>✓</td>
<td>All simulation</td>
</tr>
<tr>
<td>Soft drop jet mass scale</td>
<td>±1σ</td>
<td>✓</td>
<td>All simulation</td>
</tr>
<tr>
<td>Soft drop jet mass resolution</td>
<td>±1σ</td>
<td>✓</td>
<td>All simulation</td>
</tr>
<tr>
<td>b tag scale factor</td>
<td>±1σ</td>
<td>✓</td>
<td>All simulation</td>
</tr>
<tr>
<td>Extrapolation fit</td>
<td>±1σ</td>
<td>✓</td>
<td>Background from data</td>
</tr>
<tr>
<td>Normalization of 1W background prediction</td>
<td>1.9%</td>
<td>✓</td>
<td>Background from data</td>
</tr>
<tr>
<td>Normalization of 2W background prediction</td>
<td>1.1%</td>
<td>✓</td>
<td>Background from data</td>
</tr>
</tbody>
</table>
Tagging Boosted Jets

Heavy object (e.g. W boson) decays to two quarks
If W is sufficiently boosted, resulting jets will overlap, appearing as 1 jet
Can identify decays of heavy objects from typical light jets by mass and composite structure

“Soft Drop” Mass
Remove contribution of soft radiation to jet mass.
Resulting dist. is centered on original particle mass.

Cut-based: Require mass in the range \([65, 105]\) GeV.
One of the most sensitive variables in the NN as well.

N-subjettiness: “\(\tau_N\)"
A measure of consistency with N subjets.
Cut on the ratio of \(\tau_N\) to \(\tau_{N-1}\)

Cut-based: For W jet tagging, \(\tau_{21} < 0.55\)

\[
\tau_N = \frac{1}{d_0} \sum_k p_T, k \min \{\Delta R_{1,k}, \Delta R_{2,k}, \cdots, \Delta R_{N,k}\}
\]
Expected B quark mass limit [GeV]

Cut-based analysis

NN analysis

CMS

35.9 fb⁻¹ (13 TeV)

B2G-18-005
VLQ Results — BB Limits By Branching Fraction

- Scan limits over the branching fractions of the VLQ
  - BEST sensitive to $B \to tW$, $B \to bZ$ and $B \to bH$ covered by other analysis

![Graph showing branching fractions](image-url)

- Observed $B$ quark mass limit [GeV]
- Cut-based analysis
- NN analysis

- 35.9 fb$^{-1}$ (13 TeV)
Summary Of Limits on Production of Single VLQs

Vector-like quark single production

σ(Xtq) x B(X → tW), RH
σ(Xtq) x B(X → tW), LH
σ(Btq) x B(B → tW), RH
σ(Bbq) x B(B → tW), RH
σ(Bbq) x B(B → tW), LH
σ(Bbq) x B(B → bH), LH
σ(Bbq) x B(B → bZ), LH
σ(Btq) x B(B → bZ), LH
σ(Ttq) x B(T → tZ), RH
σ(Tbq) x B(T → tZ), LH
σ(Ttq) x B(T → tH), RH
σ(Tbq) x B(T → tH), LH
σ(Tbq) x B(T → tW), LH
σ(Ybq) x B(Y → bW)
σ(Qq) x B(Q → qZ)
σ(Qq) x B(Q → qW)

95% CL Upper Cross Section Limits (pb)

CMS, EPS-HEP 2019