Revisiting combinatorial problem in the dileptonic $t\bar{t}$ production with neural networks

Cosmos (Zhongtian) Dong
University of Kansas

Based on work done in collaboration with
H. Alhazmi, L. Huang, J. Kim, K.C. Kong, D. Shih
Dileptonic ttbar-like events topology

- Imperfect quark charge measurement and missing kinematic information
- Two-fold ambiguity in reconstructing top quark momentum
- We would like to develop an algorithm that finds the correct pairing
- We want to compare existing methods with machine learning approaches
- Event generation using MG5+ Gaussian smearing (following detector resolution given by ATLAS Collaboration)
Method 1: endpoints

- Check if mass variables given by specific pairing violates endpoints
- Event is considered unresolved if both pairings does not violate the endpoints

\[ \max\{m_{b\ell^+}, m_{b\ell^-}\} \leq m_{b\ell}^{\text{max}} = \sqrt{\left(\frac{m_{\text{top}}^2 - m_W^2}{m_W^2}\right)\left(\frac{m_W^2 - m_{\nu}^2}{m_W^2}\right)} \]

\[ M_{T2}(\tilde{m}) \equiv \min_{\vec{q}_{1T}, \vec{q}_{2T}} \{ \max [M_{TP_1}(\vec{q}_{1T}, \tilde{m}), M_{TP_2}(\vec{q}_{2T}, \tilde{m})] \} \]

\[ \vec{q}_{1T} + \vec{q}_{2T} = \vec{p}_T, \]

9906349, Lester, Summers

- We can use other similarly defined mass variable with different constraints such as \( M_{2CG}, M_{2CW}, M_{2CT} \)

Parton level: efficiency 77%  purity 96%
Smeared level: efficiency 74%  purity 87%

Efficiency is defined by percentage of resolved events of all events. Purity is the percentage of correctly resolved events of the resolved events.
Method 2 : Topness

- Chi square minimization over MET constraint
- Smaller Chi square values are considered more likely to be the correct pairing

1212.4495 Graesser, Shelton
1807.11498 Kim, Kong, Matchev, Park

\[
\chi^2_{ij} \equiv \min_{p_T = p_{\nu T} + p_{\bar{\nu} T}} \left[ \frac{\left( m_{b_j \nu}^2 - m_t^2 \right)^2}{\sigma_l^4} + \frac{\left( m_{i \nu}^2 - m_W^2 \right)^2}{\sigma_W^4} \right] \\
+ \left[ \frac{\left( m_{b_j \bar{\nu}}^2 - m_t^2 \right)^2}{\sigma_l^4} + \frac{\left( m_{i \bar{\nu}}^2 - m_W^2 \right)^2}{\sigma_W^4} \right]
\]

Parton level: efficiency 100% purity 87%
Smeared level: efficiency 100% purity 81%
Method 3: KLfitter

- Uses transfer functions and probability densities to estimate likelihood of measured objects given model signature
- Depend on angular correlation and invariant mass

\[ \mathcal{L} = \prod_{x,y} \mathcal{G} \left( E_i^{\text{miss}} | p_{x}^{\nu_1}, p_{y}^{\nu_2}, \sigma_i^{\text{miss}} (m_t, m_W, \eta_{\nu_1}, \eta_{\nu_2}) \right) \cdot \prod_{i=1}^{2} \mathcal{G} \left( \eta_{\nu_i} | m_t \right) \cdot \left( m_{\ell_1,q_1} + m_{\ell_2,q_2} \right)^\alpha \cdot \prod_{i=1}^{2} W_{\text{jet}} \left( E_{\text{jet},i}^{\text{meas}} | E_{\text{jet},i} \right) \cdot \prod_{i=1}^{2} W_{\ell} \left( E_{\ell,i}^{\text{meas}} | E_{\ell,i} \right) \]

Parton level: efficiency 100% purity 85%
Neural network approach

- Uses four-momentum information as inputs
- Formulated as a binary classification problem
- Three hidden layer fully connected neural network
- Using Keras library and trained with Adam optimizer

Parton level: efficiency 100% purity 90%
Smeared level: efficiency 100% purity 85%
Parton level: efficiency 76%  purity 99% (marked by X)

Smeared level: efficiency 54%  purity 99% (marked by X)
Attention mechanism

• Recent studies show success in multi-jets combinatorial assignment in the top quark pair production
  2010.09206 Fenton, Shmakov, Ho, Hsu, Whiteson, Baldi
  2012.03542 Lee, Park, Watson, Yang

• Permutation invariant architecture

• Uses dot product between every object in the sequence

Parton level: efficiency 100% purity 90%

Smeared level: efficiency 100% purity 85%

1706.03762 Vaswani, Shazeer, Parmar, Uszkoreit, Jones, Gomez, Kaiser, Polosukhin
Summary

• Combinatorial ambiguity is everywhere

• Existing methods (Endpoint, Topness, KLfitter etc.) resolve two-fold ambiguity using kinematic features

• Machine learning based methods outperform existing kinematic methods

• The methods can be used for precision measurement i.e. CP phase measurement in the tth production

• The methods can be generalized for arbitrary mass spectrum. (work in progress)
Method 1: endpoints

\[
M_{2CC} = \min \left\{ \max \left[ M_{P_1}(\bar{q}_1, \bar{m}), \ M_{P_2}(\bar{q}_2, \bar{m}) \right] \right\}
\]
\[
\bar{q}_1 T + \bar{q}_2 T = \bar{p}_T
\]
\[
M_{P_1} = M_{P_2}
\]
\[
M_{R_1}^{P_1} = M_{R_2}^{P_2}
\]

\[
M_{2CW}^{(bc)} = \min \left\{ \max \left[ M_{P_1}(\bar{q}_1, \bar{m}), \ M_{P_2}(\bar{q}_2, \bar{m}) \right] \right\}
\]
\[
\bar{q}_1 T + \bar{q}_2 T = \bar{p}_T
\]
\[
M_{P_1} = M_{P_2}
\]
\[
M_{R_1}^{P_1} = M_{R_2}^{P_2} = m_W^2
\]
\[
M_{2CT}^{(ct)} = \min \left\{ \max \left[ M_{P_1}(\bar{q}_1, \bar{m}), \ M_{P_2}(\bar{q}_2, \bar{m}) \right] \right\}
\]
\[
\bar{q}_1 T + \bar{q}_2 T = \bar{p}_T
\]
\[
M_{P_1} = M_{P_2}
\]
\[
M_{R_1}^{P_1} = M_{R_2}^{P_2} = m_t^2
\]
<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Parton purity</th>
<th>Parton efficiency</th>
<th>Smeared purity</th>
<th>Smeared efficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Endpoints method</td>
<td>0.957</td>
<td>0.769</td>
<td>0.874</td>
<td>0.742</td>
</tr>
<tr>
<td>Topness method</td>
<td>0.869</td>
<td>1</td>
<td>0.814</td>
<td>1</td>
</tr>
<tr>
<td>KLfitter</td>
<td>0.85</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neural Network</td>
<td>0.895</td>
<td>1</td>
<td>0.845</td>
<td>1</td>
</tr>
<tr>
<td>Neural Network</td>
<td>0.990</td>
<td>0.762</td>
<td>0.990</td>
<td>0.543</td>
</tr>
<tr>
<td>Attention Network</td>
<td>0.898</td>
<td>1</td>
<td>0.844</td>
<td>1</td>
</tr>
<tr>
<td>Boosted tree</td>
<td>0.861</td>
<td>1</td>
<td>0.824</td>
<td>1</td>
</tr>
<tr>
<td>BDT with masses input</td>
<td>0.894</td>
<td>1</td>
<td>0.840</td>
<td>1</td>
</tr>
</tbody>
</table>
Data generation

- 1 million simulated events at parton level
- 14 TeV center of mass energy

- Use Gaussian smearing to simulate detector effects

\[
\frac{\sigma_{jets}}{E} = \sqrt{\left(\frac{5.3}{E}\right)^2 + \left(\frac{0.74}{\sqrt{E}}\right)^2 + 0.05^2}
\]

\[
\frac{\sigma_e}{E} = \sqrt{\left(\frac{0.3}{E}\right)^2 + \left(\frac{0.1}{\sqrt{E}}\right)^2 + 0.01^2}
\]
Matrix element method

• Uses true neutrino momentum
• 93% find the correct pairing