Multiple Event Interpretations in Jet Physics

Yang-Ting Chien

Los Alamos National Laboratory

July 1, 2014
"LHC After the Higgs", Santa Fe

In collaboration with D. Farhi, D. Krohn, A. Marantan, D. L. Mateos, M. Schwartz
Outline

- LHC after the Higgs
  - Search for BSM physics
  - Precision measurement
  - Jet physics
- Multiple event interpretations
  - Q-jets and Q-events (jet sampling)
  - T-jets (telescoping jets)
- Demonstration: the higgs search in $ZH \rightarrow e^+e^-b\bar{b}$
  - From cut-and-count towards extended maximum likelihood fit
  - Combining and comparing T-jets with Multivariate analysis
- Results

reference:
1407.xxxx (to appear before July 4th),
Telescopin jets: 1304.5240, Jet sampling: 1304.2394, Q-jets: 1201.1914
Higgs discovery on July 4, 2012

- We saw $h \rightarrow \gamma\gamma$

- Can we see $h \rightarrow b\bar{b}$?

- Clean signal but small cross section
  - Then the LHC was shutdown in early 2013
  - Restart in early 2015

- No (Run 1), and yes (Run 2)
- Extraction of the bottom and top Yukawa couplings needs more statistics and better jet physics techniques

Large branching fraction but also large background
One important quest in jet physics

- **Quark-gluon discrimination**
  - Many SM and BSM signals have quark-heavy final states
  - QCD backgrounds are mostly gluon-heavy
  - Quark and gluon jets have different jet substructures

**Being able to distinguish quarks from gluons is very important**

- Gluino decay chain with a quark-heavy final state
- Different higgs production mechanisms with quark or gluon jets in the final state

![Diagram of gluino decay chain](image1)
![Diagram of higgs production mechanisms](image2)

Yang-Ting Chien
Multiple Event Interpretations in Jet Physics
Jet physics in a nutshell

- Jets are a manifestation of the underlying colored partons
  - Partons emit soft and collinear radiation
    - To reconstruct the hard process it is necessary to strip off the complication from QCD
  - Define jets and measure some jet observables
  - Tools: analytic calculations and numerical simulations

\[ H(\rightarrow b\bar{b}) + Z(\rightarrow \mu\bar{\mu}) \] production at parton and hadron levels

Key words:
pQCD, SCET, PYTHIA, MadGraph, Herwig, FastJet, jet shape, jet substructure, jet superstructure, ...
Jets at LEP and the LHC

- Jets are distinct, localized structure in calorimeter
  - Jets can have different widths, can have substructure, ..., etc.

- The first hadronic Z decay observed at OPAL with back-to-back dijets
- A multi-jet event at the 7 TeV LHC
What is a jet more precisely?

- Identifying (defining) jets: jet algorithms with a parameter $R$
  - The parameter $R$ sets the *artificial* jet size
    - jet constituents are those particles within an angular scale $R$ away from the jet direction
  - There are three types of angular scales in the problem
    - $R$, jet kinematics (the angles between jets) and jet widths
    - The jet width is a dynamically generated angular scale

**Interrogating jets!**

- With jets, we want to know all the information about them
  - which parton gives which jet? quarks v.s. gluons
  - can we measure the charge and spin of a parton?
  - what is the underlying scattering process?
Clustering algorithms

- Idea: merge the pair of particles with the shortest \textit{distance} until the particles are away from one another farther than $R$
- The distance measure $d_{ij}$ between the particles $i$ and $j$ is defined by

$$d_{ij} = \min(p_{ti}^{2\beta}, p_{tj}^{2\beta}) \Delta R_{ij}^2 / R^2, \quad d_{iB} = p_{ti}^{2\beta} \quad B: \text{beam}$$

\[ \beta = 1: k_T \]
\[ \beta = 0: \text{Cambridge/Aachen} \]
\[ \beta = -1: \text{anti-}k_T \]

\(\beta\) controls the merging priority in energy

- In the end we get a tree for each jet
- Different algorithms reveal different aspects of the jet structure
Q-jets: non-deterministic clustering algorithms (Ellis et al.)

- Idea: merge particles probabilistically according to a weight
  \[ w_{ij}^{(\alpha)} = \exp \left\{ -\alpha \frac{d_{ij} - d_{\text{min}}}{d_{\text{min}}} \right\}, \quad d_{\text{min}} = \min d_{ij} \]

- There is still one single parameter \( R \)
- \( \alpha \) controls the deviation from the classical, deterministic clustering
- Q-jets gives different trees and jet constituents in each reconstruction
- Each jet observable turns from a single number to a distribution

**Nice performance in boosted \( W \)-tagging with pruning**

- Pruned jet mass for a single QCD-jet
- Volatility distributions of \( W \) and QCD jets
Q-events: Q-jets applied to the whole event (Schwartz et al.)

- Each run gives a different reconstruction, or interpretation, of an event
  - Every time jets are *sampled* from the event differently

**Nice performance in** $pp \rightarrow \phi, \phi\phi, Z\phi$ **and** $ZH$ **searches using** Qanti-$k_T$

- The frequency with which a calorimeter cell is clustered into one of the hard jets in a simulated $pp \rightarrow \phi\phi \rightarrow gggg$ event at the LHC

- The jet area becomes a distribution, instead of a single number $\pi R^2$
Then why use only one fixed $R$ for all the jets?

- There is no reason for jets to have the same size $R$
  - Jet formation is quantum mechanical

- Two $b$ jets with the same partonic kinematics but different widths

- Distinguishing the width of the localized energy distribution of a jet from the parameter $R$
Telescoping jets: T-jets (Chien)

- Idea: use whatever jet algorithm you like with multiple $R$’s
  - Each choice of $R$ gives a distinct interpretation of an event
  - Probe the structure of an event many times with different angular resolutions

- Demonstration: the higgs search in $ZH$ production with $H \rightarrow b\bar{b}$
  - With a $p_T^Z > 120$ GeV cut
  - Perform a counting experiment with a dijet invariant mass window

**T-jets recipe**

- Determine the $b$-jet axes $n_1$ and $n_2$: use the anti-$k_T$ algorithm with $R_{core}$ to reconstruct the cores of the two hardest jets

- Telescope around the axes: define the $i$-th jet to be the particles within a distance $R$ away from $n_i$

\[
\text{jet}_R^i = \{ \vec{p} \mid (\eta_\vec{p} - \eta_{n_i})^2 + (\phi_\vec{p} - \phi_{n_i})^2 < R^2 \}\]

- If jets overlap, assign particles to the jet with the nearest jet axis
Each event is counted by the fraction of interpretations $z$ passing the cuts, instead of 0 or 1 in a conventional analysis:
- $110 \text{ GeV} < m_{jj} < 140 \text{ GeV}$
- $N$ $R$’s ranging from $R_{\text{min}}$ to $R_{\text{max}}$

$m_{jj}$ becomes a distribution for a single $ZH$ event with multiple interpretations.

Signals are more robust in multiple event interpretations than backgrounds.

In general, we can count with a weight function $w(x) = \frac{\rho_S(x)}{\rho_B(x)}$ associated with the distribution $\rho(x)$ of any observable $x$. 

![Graph showing $m_{jj}$ distribution for $ZH$ event with multiple interpretations](image)
Merged $m_{jj}$ distribution

- The signal mass peak gets broader, but the statistical stability of observables is increased so that background fluctuations shrink considerably, which is the key for $S/\delta B$ improvement

\[ \frac{S}{\delta B} = \frac{\text{expected signal in the window}}{\text{fluctuation of background in the window}} \]

<table>
<thead>
<tr>
<th>$R$ range</th>
<th>$N$</th>
<th>algorithm</th>
<th>weight</th>
<th>$S/\delta B \uparrow$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.4 and 1.0</td>
<td>2</td>
<td>cone</td>
<td>$z$</td>
<td>14%</td>
</tr>
<tr>
<td>0.4 to 1.0</td>
<td>7</td>
<td>cone</td>
<td>$z$</td>
<td>20%</td>
</tr>
<tr>
<td>0.4 to 1.5</td>
<td>12</td>
<td>cone</td>
<td>$z$</td>
<td>26%</td>
</tr>
<tr>
<td>0.2 to 1.5</td>
<td>100</td>
<td>anti-$k_T$</td>
<td>$z$</td>
<td>20%</td>
</tr>
<tr>
<td>0.2 to 1.5</td>
<td>100</td>
<td>cone</td>
<td>$z$</td>
<td>28%</td>
</tr>
<tr>
<td>0.4 to 1.5</td>
<td>12</td>
<td>cone</td>
<td>$\rho_S/\rho_B$</td>
<td>38%</td>
</tr>
<tr>
<td>0.2 to 1.5</td>
<td>100</td>
<td>cone</td>
<td>$\rho_S/\rho_B$</td>
<td>46%</td>
</tr>
</tbody>
</table>

- The $S/\delta B$ improvement is significant
Extended maximum likelihood fit and multivariate analysis

- Cut-and-count pre-assumes $\sigma_B$, while EML fits $\sigma_S$ and $\sigma_B$ simultaneously
- We can fit the 1-D merged distribution or do the multi-dimensional fit
- The numbers of interpretations and bins can make the multi-dimensional fit impractical, and we replace the exact likelihood with boosted decision trees (BDT)

2-D mass distributions and correlation matrices for signals and backgrounds
What extra information do we gain? Can they be equally captured by the standard kinematic variables? The answer is no.

<table>
<thead>
<tr>
<th>Improvement (over R=0.5 only)</th>
<th>$p_T^Z &lt; 120$ GeV</th>
<th>$p_T^Z &gt; 120$ GeV</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>xs-based</td>
<td>EML</td>
</tr>
<tr>
<td>1 $R$, fraction in window</td>
<td>0.83</td>
<td>0.74</td>
</tr>
<tr>
<td>12 $R$'s, fraction in window</td>
<td>0.98</td>
<td>0.97</td>
</tr>
<tr>
<td>1 $R$, $m_{b\bar{b}}$</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>5 $R$'s, $m_{b\bar{b}}$ merged</td>
<td>0.94</td>
<td>1.08</td>
</tr>
<tr>
<td>4 $R$'s, 3 bins</td>
<td>0.99</td>
<td>1.00</td>
</tr>
<tr>
<td>2 $R$'s, full</td>
<td>1.10</td>
<td>1.14</td>
</tr>
<tr>
<td>2 $R$'s, BDT</td>
<td>1.04</td>
<td>1.08</td>
</tr>
<tr>
<td>12 $R$'s, BDT</td>
<td>1.19</td>
<td>1.30</td>
</tr>
<tr>
<td>12 kinematic</td>
<td>1.33</td>
<td>1.50</td>
</tr>
<tr>
<td>12 kinematic + 12 $R$'s</td>
<td>1.39</td>
<td>1.68</td>
</tr>
</tbody>
</table>

The 12 kinematic variables are the nearly optimal set in MVA (Schwartz et al.)
Conclusions and future work

- Multiple event interpretations extract more information out of an event
  - wide-angle radiation turns out to be important

- The T-jets implementation is a simple, powerful, and experiment-friendly method, which hopefully will become a standard analysis procedure

- Future work
  - Deal with the issue of pile-up
  - Analytic understanding of new observables with multiple interpretations
    e.g. correlations between observables ($m_{R_1}$ and $m_{R_2}$)
  - T-jets in boosted-particle tagging
  - Applications to BSM searches