Extracting Hadronic Resonances using Jet Ensembles

LHC New Physics Signatures Workshop
January 10, 2008

Eva Halkiadakis
In collaboration with: Shan-Huei Chuang, Amit Lath, Scott Thomas, Rouven Essig
Rutgers, the State University of NJ
Motivation

- *Exotic* searches at colliders always involve MET or leptons/photons.
  - Strong production
  - ElectroWeak decays
  - Backgrounds suppressed

- New physics $\rightarrow$ Jets?
  - Strong production cross-section
  - Strong decays (multi-jet)
  - Backgrounds severe
New Physics in Multi-Jets

- Look for new physics in multi-jets
- Studying $pp \rightarrow QQ \rightarrow 3j+3j = 6j$
  - $Q = \tilde{g} = SU(3)C$ Adjoint Majorana Fermion
- Challenging
  - Large backgrounds!
  - Magnitude of multi-jet backgrounds from higher order processes difficult to calculate a priori ($\alpha_s^n$).
  - But possible new physics may be hidden in jets!
- Get guidance from all-hadronic top studies
- Make use of kinematic features and correlations
- Use an ensemble of jet combinations
- Techniques may also be useful for jets produced with leptons, MET, photons and we can study this later
Signal & Background

Signal:

pp → QQ

6 Jet Background:

Hadronic Top Background:

Detector Simulator:

Analysis:

Pythia
Model it as RPV (uud Yukawa) gluino
MSbar masses 200 GeV (real mass 290 GeV) and up

ALPGEN → Pythia

PGS (fast simulation)
Studies with full simulation CMSSW ongoing

(Ch)Root
“Bump hunt”

No leptons, No MET, No W resonance, No b
Cuts: Trigger Level

- $|\eta| < 3$ of the first 6 jets
  (PGS requires a "jet" to have at least 5 GeV of $p_T$)

- $(1\text{st jet} > 400 \text{ GeV} \text{ OR} \ 2\text{nd jet} > 350 \text{ GeV} \text{ OR} \ 3\text{rd jet} > 195 \text{ GeV} \text{ OR} \ 4\text{th jet} > 80 \text{ GeV} \text{ OR} \ \text{sum had} > 1000 \text{ GeV})$

This is dominated by the 4th jet trigger.
Adding the rest adds only a few percent.
Cuts: Analysis Level

- Cut on sum $p_T$ of the 1st 6 jets:
  \[ \sum_{i=1}^{6} p_{T,i} > \]
  - gluino200 600 GeV
  - gluino300 700 GeV
  - gluino500 1100 GeV
  - gluino700 1500 GeV

- Cut on the 6th jet $p_T >$
  - 30 GeV, 60 GeV, 90 GeV, 120 GeV

We are trying different cuts to optimize signal vs. background
Sum $P_{T,jets}$ vs. $P_{T,6th}$ jet

Example

Cuts: $\sum_{i=1}^{6} p_{T,i} > 600$ GeV, $p_{T,6th}$ jet > 90 GeV

$pp \rightarrow QQ \rightarrow jjjjjj$

$N_{jet} \geq 6$

Signal

$M_Q = 290$ GeV

$\sum_{i=1}^{6} p_{T,i} > 600$ GeV, $p_{T,6th}$ jet > 90 GeV

Hadronic Background

Signal Efficiency $\sim 0.02$
How To Select Jet Triplets?

\[ pp \rightarrow QQ \rightarrow jjjjjj \]

Assume six hardest jets come from QQ decay

Two Three-Body Resonances \( M_{jjj} = M_{jjj} \)

Choose Pair of Jet triplets (from 10 possibilities) with smallest \( |M_{jjj} - M_{jjj}| < 60 \text{ GeV} \) (plus other kinematic cuts)

- Tail much larger than jet resolution
- Mismatching of jet triplets
- Combinatoric background within signal

**Signal**

\[ M_Q = 290 \text{ GeV} \]

**Hadronic Background**

\[ 250 \text{ pb}^{-1} \]

January 10, 2008

Extracting Hadronic Resonances using Jet Ensembles, E. Halkiadakis
Selecting Jet Triplets: 
Ensemble of Jet Combinations

Which Combination?
There are 20 possible triplets among 6 jets.

Use MC matching info to find which triplets are correct most often:

<table>
<thead>
<tr>
<th>Combo</th>
<th>%correct</th>
<th>Combo</th>
<th>%correct</th>
</tr>
</thead>
<tbody>
<tr>
<td>235</td>
<td>5.3</td>
<td>246</td>
<td>3.0</td>
</tr>
<tr>
<td>234</td>
<td>4.7</td>
<td>135</td>
<td>2.9</td>
</tr>
<tr>
<td>236</td>
<td>4.3</td>
<td>345</td>
<td>2.6</td>
</tr>
<tr>
<td>245</td>
<td>4.3</td>
<td>256</td>
<td>2.0</td>
</tr>
<tr>
<td>145</td>
<td>4.2</td>
<td>134</td>
<td>2.0</td>
</tr>
<tr>
<td>146</td>
<td>4.0</td>
<td>126</td>
<td>1.7</td>
</tr>
<tr>
<td>156</td>
<td>3.4</td>
<td>346</td>
<td>1.6</td>
</tr>
<tr>
<td>136</td>
<td>3.2</td>
<td>125</td>
<td>1.4</td>
</tr>
</tbody>
</table>

Jets ordered in $p_T$ (e.g. 123 are three highest $p_T$ jets)

Combo 235 is correct 5.3% of the time.

Combo 123 is not among the top 16 combinations.
Extracting Hadronic Resonances using Jet Ensembles, E. Halkiadakis

Using Kinematic Correlations: Mass vs. Sum $P_T$

Extract Kinematic Features from Combinatoric Confusion

Best 16 Triplets (16 entries/event)

Want to isolate good triplets

Horizontal Branch: Region of high signal to combinatoric background contrast

$M_Q = 290$

$\sum_j |p_{T,j}|$ GeV

Signal $250 \text{ fb}^{-1}$
Using Kinematic Correlations:
Mass vs. Sum $P_T$

Extract Kinematic Features from Combinatoric Confusion

Best 16 Triplets (16 entries/event)

Signal
$M_Q=290$

Cut: $M_{jjj} < \sum |p_{T,j}| - offset$

Want to isolate good triplets

Horizontal Branch:
Region of high signal to combinatoric background contrast

$250 \text{ fb}^{-1}$
Mass vs. Sum $P_T$ for Backgrounds

**Hadronic Top Background**

**Hadronic Background**

\[ \Sigma_j |p_{T,j}| \]

250 pb$^{-1}$
Cuts: Analysis Level

- For ANY triplet of jets from the set of the “best” 16 require:
  \[ M_{jjj} < \sum |p_{T,j}| - \text{offset} \]
  - where \( M_{jjj} \) is the invariant mass of the 3 jets
  - \( \sum |p_{T,j}| \) is the scalar sum \(|p_T|\) of the 3 jets
  - offset is either infinity (i.e. no cut) or 0 GeV, 100 GeV, 200 GeV, or 300 GeV

This cut isolates the "horizontal branch" with the "correct" invariant mass, and removes a lot of background and combinatoric background within the signal.
Now Fit and Optimize Cuts

Signal: \( m = 200 \) \((m_Q = 290)\)
+ Alpgen 6j background
+ ttbar background

Fit: Gaussian + Landau

Luminosity: 1 fb\(^{-1}\)

Note: Number of entries per event is between 0 and 16.
- Avg. for signal: 0.6 entries/event
- Avg. for bkg: 0.3 entries/event

Cuts: 600, Sum6|\( P_T \)|, 90, 0, 100, \( P_{T,6\text{th jet}} \), SumVector\( P_T \), Diagonal cut offset
### Fit Results

<table>
<thead>
<tr>
<th>MSbar mass</th>
<th>cut set</th>
<th>eff_SG</th>
<th>eff_BG</th>
<th>peak mass</th>
<th>(2σ)</th>
<th>nSG</th>
<th>nBG</th>
<th>S/B</th>
<th>S/Sqrt(B)</th>
</tr>
</thead>
<tbody>
<tr>
<td>200 600__30_0_100</td>
<td>0.0960 0.0465</td>
<td>284.8</td>
<td>11514</td>
<td>1097865</td>
<td>0.01</td>
<td>10.99</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>200 600__60_0_100</td>
<td>0.0874 0.0477</td>
<td>285.3</td>
<td>4568</td>
<td>56161</td>
<td>0.08</td>
<td>19.27</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>200 600__90_0_100</td>
<td>0.1037 0.0642</td>
<td>285.2</td>
<td>2070</td>
<td>5558</td>
<td>0.37</td>
<td>27.77</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| 200 600__30_0_200 | 0.0266 0.0095 | 286.0 | 3433  | 199651  | 0.02 | 7.68  |
| 200 600__60_0_200 | 0.0261 0.0121 | 286.6 | 1821  | 56161   | 0.14 | 19.27 |
| 200 600__90_0_200 | 0.0391 0.0224 | 288.2 | 1017  | 5558    | 0.37 | 27.77 |

| 200 600__60_0_300 | 0.0088 0.0033 | 287.8 | 891   | 3592    | 0.25 | 14.87 |
| 200 600__90_0_300 | 0.0145 0.0076 | 289.1 | 429   | 574     | 0.75 | 17.92 |

| 500 1100__60_0_200 | 0.0585 0.0378 | 638.5 | 185   | 10305   | 0.02 | 1.82  |
| 500 1100__90_0_200 | 0.0547 0.0341 | 643.0 | 110   | 1646    | 0.07 | 2.72  |
| 500 1100__120_0_200| 0.0558 0.0374 | 646.0 | 52    | 344     | 0.15 | 2.78  |

| 500 1100__60_0_300 | 0.0212 0.0126 | 643.9 | 83    | 3019    | 0.03 | 1.52  |
| 500 1100__90_0_300 | 0.0208 0.0126 | 646.5 | 50    | 581     | 0.09 | 2.06  |
| 500 1100__120_0_300| 0.0233 0.0158 | 650.2 | 26    | 125     | 0.20 | 2.28  |

| 500 1100__60_0_400 | 0.0081 0.0042 | 646.8 | 37    | 1018    | 0.04 | 1.17  |
| 500 1100__90_0_400 | 0.0084 0.0045 | 648.6 | 23    | 209     | 0.11 | 1.60  |
| 500 1100__120_0_400| 0.0101 0.0064 | 652.2 | 13    | 46      | 0.29 | 1.95  |

S/sqrt(B) looks good, can be optimized further.

Not so rosy for higher masses...
For MSbar mass 200 GeV. Invariant mass of correct triplet (MC matched).

CMSSW is broader than PGS, but not too bad, can pull it out of background.
Systematic Uncertainty: Jet Resolution

- A good understanding of jets is important in this analysis.
- There are uncertainties in the jet resolution.
- Procedure (Ref. CMS Physics TDR):
  - Add an additional smearing to the jet energy which broadens the overall jet resolution by 10%.
  - Done by throwing a Gaussian random number and adding an energy term which is 46% of the jet resolution (to get overall widening of 10%).
- Jet-by-jet, event-by-event smearing:
  \[ E_T^{\text{jet}} = E_T^{\text{jet}} + \text{Gaus}[0, 0.46 \times \sigma(E_T, \eta)] \]

Reference jet resolution the central jets:
\[ \sigma(E_T^{\text{jet}}, |\eta| < 1.4) = (5.8 \text{ GeV}) \pm (1.25 \times \sqrt{E_T^{\text{jet}}}) \pm 0.033 \times E_T^{\text{jet}} \]

Reference jet resolution the forward jets:
\[ \sigma(E_T^{\text{jet}}, 1.4 < |\eta| < 3.0) = (4.8 \text{ GeV}) \pm (0.89 \times \sqrt{E_T^{\text{jet}}}) \pm 0.043 \times E_T^{\text{jet}} \]

In CMS MC Simulation

In PGS:
\[ \sigma(E_T^{\text{jet}}) \propto 0.8 \sqrt{E_T^{\text{jet}}} \text{ (for HCAL)} \]
Jet Smearing

Same as before but now with more jet smearing:

- PGS smeared to match CMSSW
  - $\sigma \approx 26$ GeV
  - $\sigma \approx 35$ GeV

- PGS smeared to match CMSSW + systematic smearing
  - $\sigma \approx 33$ GeV
  - $\sigma \approx 37$ GeV
Jet Smearing With Cuts

Now with all the analysis cuts:

- PGS (poor statistics)
- CMSSW
- PGS smeared to match CMSSW
- PGS smeared to match CMSSW + systematic smearing
Summary

- We *can* do searches with jets
- Full CMSSW studies underway, *encouraging*
  - Responsible for large ALPGEN n-jet samples within CMS
  - Generating home-brewed CMSSW signal samples
  - Beginning studies on systematic uncertainties
- Ensemble techniques useful for other analyses as well
- Working closely with theorists can pay off handsomely!
Backup
More Kinematic Correlations

Note “Correct” Triplets – Horizontal Branch

Background is Very Similar to Combinatoric Confusion

Note Shaping from Cuts