New jet tools for at the LHC

Lian-Tao Wang
University of Chicago

Implications of EWSB, Madison, May 2011
The importance of hadronic final state:

• “Everywhere” at hadron colliders. \( pp \), or, \( p\bar{p} \) initial state.

• Present in (almost) all new physics signals.
  - Many of them only have hadronic channels.
  - TeV new physics states can decay to SM “heavy” particles, e.g. \( t, W, Z \), often look like a cluster of hadrons.

• Understanding of basic structure of QCD and the properties of new physics has lead to the development of a set of modern tools which significantly enhanced the discovery potential.
Cover two aspects.

- Better characterization of QCD jets
  - Improving jet algorithms.
  - Finding ISR.

- Jet substructure and new physics searches.
  - Boosted tops.
  - Higgs search.

Boston Jet Workshop:
http://jets.physics.harvard.edu/workshop/Main.html

http://boost2011.org
Better QCD jet
Why is it hard?
Why is it hard?
Why is it hard?
Why is it hard?

• We would like to preserve $p_{\text{jet}} \sim p_{\text{parton}}$.
Why is it hard?
Why is it hard?

Multiple interaction, underlying events, pile-up
Why is it hard?

Multiple interaction, underlying events, pile-up

Overlapping jets.

"beam"
Why is it hard?

Multiple interaction, underlying events, pile-up

Overlapping jets.

ISR (beam) clustered
Why is it hard?

Overlapping jets.

Part of the beam?

Multiple interaction, underlying events, pile-up

ISR (beam) clustered

Overlapping jets.
Why is it hard?

- Overlapping jets.
- Proper "size" of jets.
- Part of the beam?
- ISR (beam) clustered
- Multiple interaction, underlying events, pile-up
- Overlapping jets.
Why is it hard?

- To best preserve $p_{jet} \approx p_{parton}$ we would like to:
  - Use “smart” jet shapes.
  - Control “contamination”.

Multiple interaction, underlying events, pile-up

ISR (beam) clustered

Overlapping jets.

Proper “size” of jets.

Part of the beam?
Begin with jet algorithm

- An algorithm of clustering together “close by” objects.
- Basic ingredients of a “sequential” jet algorithm.
- Two types of “distances”
  - Jet-jet distance: $d_{ij}$ “when to cluster”
  - Jet-beam distance: $d_{iB}$ “when to stop clustering”
- Pair wise comparison of all distances
  - If smallest distance at any stage in clustering is jet-jet, add together corresponding four-momenta, else take jet with smallest jet-beam distance and set it aside.
- Repeat till all jets are set aside.
Recombination Algorithms

• k_\text{T} algorithm

\[ d_{ij} = \min(p_{Ti}^2, p_{Tj}^2) \left( \frac{\Delta R}{R_0} \right)^2, \quad d_{iB} = p_{Ti}^2 \]

• C/A algorithm

\[ d_{ij} = \left( \frac{\Delta R}{R_0} \right)^2, \quad d_{iB} = 1 \]

• anti-k_\text{T} algorithm

\[ d_{ij} = \min(p_{T_i}^{-2}, p_{T_j}^{-2}) \left( \frac{\Delta R}{R_0} \right)^2, \quad d_{iB} = p_{Ti}^{-2} \]

\[ (\Delta R)^2 \equiv (\Delta \eta)^2 + (\Delta \phi)^2 \]
Recombination Algorithms

- **$k_T$ algorithm**
  \[ d_{ij} = \min(p_{T_i}^2, p_{T_j}^2) \left( \frac{\Delta R}{R_0} \right)^2, \ d_{iB} = p_{T_i}^2 \]

- **C/A algorithm**
  \[ d_{ij} = \left( \frac{\Delta R}{R_0} \right)^2, \ d_{iB} = 1 \]

- **anti-$k_T$ algorithm**
  \[ d_{ij} = \min(p_{-T_i}^2, p_{-T_j}^2) \left( \frac{\Delta R}{R_0} \right)^2, \ d_{iB} = p_{-T_i}^2 \]

\[ (\Delta R)^2 \equiv (\Delta \eta)^2 + (\Delta \phi)^2 \]

\[ p_T^A > p_T^B \]
Effect of contamination.

- Initial state radiation (ISR), multiple interaction (MI), underlying events (UE), pile-up (PU).
- An example of resonance reconstruction.

\[ gg \rightarrow \phi \rightarrow gg \quad m_\phi = 1 \text{ TeV} \]

\[ m^2 = (p_j^1 + p_j^2)^2 \approx m^2_\phi \]

FSR only
No contamination

Including ISR, MI, UE, pile-up

Room for improvement!
Jet trimming.  

D. Krohn, J. Thaler, LTW, arXiv:0912.1342

- Introducing a “cut” on soft radiation.
- Discard “stuff” below the cut after jet clustering.

- Our implementation.
  - Cluster all calorimeter data using any algorithm
  - Take the constituents of each jet and recluster with smaller radius $R_{sub}$ ($R_{sub} = 0.2$ seems to work well).
  - Discard the subjet $i$ if $p_{T i} < f_{\text{cut}} \cdot \Lambda_{\text{hard}}$  
    
    ISR argument.

- Best choice of the hard scattering scale and $f_{\text{cut}}$.
- Process dependent.
- Can be optimized experimentally.

Related but different “jet grooming” approaches:
Pruning: S. Ellis, C. Vermilion, J. Walsh, arXiv:0903.5081
Start
Cluster into subjets
Discard soft subjets
Reassemble
Simple test case: di-jet resonance

\[ gg \rightarrow \phi \rightarrow gg \]

We provide plugins fully compatible with Fastjet.

http://jthaler.net/jets/VR_Jets.html
http://jthaler.net/jets/Jet_Trimming.html
Finding ISR jet.  

• Looking for the different jet.

**Tag**

• Take three hardest jets. Look for those
  
  1. Distinguished in pT

**Check**

• Require the candidate ISR jet
  
  1. Not be central

  **AND**

  2. Remain somewhat isolated in rapidity

• And, require that the implicit FSR jets be
  
  1. Close in pT
Simple kinematical tagger works well.

<table>
<thead>
<tr>
<th>Spectrum</th>
<th>Efficiencies [%]</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Trigger</td>
<td>Mistag</td>
</tr>
<tr>
<td>$m_{\tilde{q}}/m_{\tilde{g}}$</td>
<td>$m_{\text{LSP}}$</td>
<td></td>
</tr>
<tr>
<td>500 GeV</td>
<td>100 GeV</td>
<td>42</td>
</tr>
<tr>
<td>500 GeV</td>
<td>450 GeV</td>
<td>42</td>
</tr>
<tr>
<td>1 TeV</td>
<td>100 GeV</td>
<td>41</td>
</tr>
<tr>
<td>1 TeV</td>
<td>950 GeV</td>
<td>41</td>
</tr>
<tr>
<td>500 GeV</td>
<td>100 GeV</td>
<td>13</td>
</tr>
<tr>
<td>500 GeV</td>
<td>400 GeV</td>
<td>15</td>
</tr>
<tr>
<td>1 TeV</td>
<td>100 GeV</td>
<td>12</td>
</tr>
<tr>
<td>1 TeV</td>
<td>900 GeV</td>
<td>16</td>
</tr>
</tbody>
</table>

- Further developments underway.
- Asymmetric topology?
- More inclusive?
Many potential applications.

- Reducing combinatorics.
  - SUSY decay chain, ttbar, ....

- ISR could be the main component of the signal.
  - Squeezed SUSY spectrum, ...

- Measuring mass
  - ISR spectrum is proportional to the scale of hard interaction.

- Even more directly:
Mass measurement

A system with invariant mass $M_0$ including visible+invisible

For any FSR with $p_T = \vec{p}_T \cdot \vec{p}_T^{\text{ISR}}$, and assuming $M_{\text{test}}$,

$$p_T \rightarrow \frac{p_T^{\text{ISR}}}{M_{\text{test}}} E_i + \sqrt{1 + \left(\frac{p_T^{\text{ISR}}}{M_{\text{test}}}\right)^2} p_T^i.$$
Example: squark pair production

- Produced near threshold, \( M_0 \sim 2m_{\tilde{q}} \)

\[
m_{\tilde{q}} = 500 \text{ GeV} \\
M_0 \sim 1.2 \text{ TeV} \\
m_{\tilde{q}} = 1 \text{ TeV} \\
M_0 \sim 2.5 \text{ TeV}
\]
Examples: squark pair production

\[ m_{\tilde{q}} = 500 \text{ GeV} \]
\[ M_0 \sim 1.2 \text{ TeV} \]

\[ m_{\tilde{q}} = 1 \text{ TeV} \]
\[ M_0 \sim 2.5 \text{ TeV} \]

\[ \sigma = +1 \ (-1) \text{ if } \sum_i p_{T_i} > 0 \ (<0) , \langle \sigma \rangle = \sum_{j=1}^{N} \sigma_j / N \]
QCD Jets: current and future

- Well established the new improved jet algorithms will be instrumental in new physics discovery.
  - Optimization.
  - LHC experimental groups are testing them.
- More flexible, more dynamical.
- Jet tagging.
  - ISR. q vs g, charge?
- Better theoretical understanding.
  - Factorization ....
Jet substructure, and applications in new physics searches.
Jet substructure.

- When produced at TeV-scale energies, they have a large boost.

Jets with substructure.

Challenge: distinguishing them from QCD jets (q and g).
When to consider substructure

- Have to consider the boosted objects.

  For example, boost tops
  
  Brooijmans; Lillie, Randall, LTW; Thaler, LTW;
  
  D. Kaplan, K. Reherman, M. Schwartz, B. Tweedie;
  
  L. Almeida, S. Lee, G. Perez, G. Sterman, I. Sung, J. Virzi
  
  ... 

- It is beneficial to consider the boosted objects.

  Lower combinatorics,
  
  SM background boost differently.
  
  Butterworth, Davidson, Bubin, Salam

For a summary of recent developments: C. Vermilion, 1001.1335
Example: boosted top tagging at the LHC

- Fully collimated tops look like QCD jets.
Example: boosted top tagging at the LHC

- Fully collimated tops look like QCD jets.

Basic distinction:
- QCD: radiation.
- Top decay: $t \rightarrow bW(\rightarrow qq')$ 3 hard objects.

Zooming in near the first splitting

QCD. Soft radiation: $z = \frac{\text{Min}(E_1, E_2)}{E_1 + E_2} \rightarrow 0$

Top. Decay: $z = \frac{\text{Min}(E_W, E_b)}{E_W + E_b} \rightarrow \text{finite}$
Example: boosted top tagging at the LHC

- Fully collimated tops look like QCD jets.

Basic distinction:
- QCD: radiation.
- Top decay: $t \rightarrow bW(\rightarrow qq')$ 3 hard objects.

Zooming in near the first splitting

**QCD.** Soft radiation:
$$z = \frac{\text{Min}(E_1, E_2)}{E_1 + E_2} \rightarrow 0$$

Jet mass:
$$d\sigma \propto \frac{1}{m_{\text{jet}}^2}$$

**Top.** Decay:
$$z = \frac{\text{Min}(E_W, E_b)}{E_W + E_b} \rightarrow \text{finite}$$

Jet mass:
$$m_{\text{jet}} \simeq m_{\text{top}}$$
Example: boosted top tagging at the LHC

- Fully collimated tops look like QCD jets.
- QCD: radiation.
- Top decay: \( t \to bW(\to qq') \) 3 hard objects.

Basic distinction:

Zooming in near the first splitting

**QCD.** Soft radiation: 
\[
z = \frac{\text{Min}(E_1, E_2)}{E_1 + E_2} \to 0
\]

Jet mass: 
\[
d\sigma \propto \frac{1}{m_{\text{jet}}^2}
\]

**Top.** Decay: 
\[
z = \frac{\text{Min}(E_W, E_b)}{E_W + E_b} \to \text{finite}
\]

Jet mass: 
\[
m_{\text{jet}} \simeq m_{\text{top}}
\]

Early splittings

microscope: jet substructure variables
Jet mass: new jet algorithm can help.

Effect of radiation contamination on the jet mass

\[ \langle \delta M^2 \rangle \simeq (\Lambda_{\text{soft}} + p^\text{ISR}_T)p^j_T \left( \frac{(\Delta R)^4}{4} + \ldots \right) \]

- **Trimming gives large improvement by reducing effective jet size significantly.**
Top jets vs QCD jets

"Following" the branching History

Recursive algorithm, e.g., $k_T^*$ close to an evolution variable.

$k_T$ clustering history $\sim$ inverse branching history.

Rough approximation of finite calorimetry: $\delta \eta \times \delta \phi = 0$.

QCD soft singularity is in effect regulated.

J. Thaler, LW, arxiv:0806.0023

Top jets vs QCD jets

- Combined cuts on jet mass and $z$ can enhance further the signal with respect to the background.

For a comprehensive study:
Boost 2010 proceeding, 1012.5421
More jet shape variables.

- Top decay is more like 3-body. Span a “plane” perpendicular to the jet axis.

- Transverse sphericity, or “planar flow”

\[
S_{\perp ij} = \frac{\sum_{\alpha \in \text{jet}} p_{\alpha i} p_{\alpha j}}{\sum_{\alpha \in \text{jet}} |p_{\alpha \perp}|}.
\]

\[\vec{p}_{\alpha \perp} : \text{w.r.t. jet axis, } i = 1, 2\]


\[\text{det } S_{\perp} = 0\]

\[\text{det } S_{\perp} \neq 0\]
Better reconstruction of the jet shape

Planar flow

Defined in

• Can be used to further improve top tagging. An additional factor of several possible.

• Interesting to compare with improved QCD calculation, using modern technologies such as SCET.
Hiding Higgs.

- Alternative decay channels can dramatically change Higgs search strategy.

  \[ h \rightarrow aa \rightarrow 4\tau, \ 4b, \ \bar{b}b\tau \]

  For example:
  - M. Carena, T. Han, G. Huang, C. Wagner, arXiv:0712.2466

  \[ h \rightarrow aa \rightarrow c\bar{c}c\bar{c}, \ \text{"charful"}? \]
  For example:

  \[ h \rightarrow aa \rightarrow gggg, \ \text{"buried"!} \]

- Why can new jet technology help?

  - Less radiation outside this cone
  - Jet substructure
Some preliminary results.

Higgs + Z signal

Z+jet background
Encouraging results.

<table>
<thead>
<tr>
<th>p_T(j) &gt; 200 GeV</th>
<th>( \sigma_{sig} ) (fb)</th>
<th>( \sigma_{bg} ) (fb)</th>
<th>( S/B )</th>
<th>( S/\sqrt{B} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>subject mass</td>
<td>16</td>
<td>30000</td>
<td>0.00052</td>
<td>0.9</td>
</tr>
<tr>
<td>Higgs window</td>
<td>12</td>
<td>19000</td>
<td>0.00062</td>
<td>0.9</td>
</tr>
<tr>
<td>( \alpha &gt; 0.7 )</td>
<td>7.1</td>
<td>400</td>
<td>0.018</td>
<td>3.6</td>
</tr>
<tr>
<td>( \beta &lt; 0.005, p_T^{min} = 1 \text{ GeV} )</td>
<td>0.67</td>
<td>0.74</td>
<td>0.90</td>
<td>7.8</td>
</tr>
<tr>
<td>( \beta &lt; 0.005, p_T^{min} = 5 \text{ GeV} )</td>
<td>2.9</td>
<td>2.6</td>
<td>0.11</td>
<td>5.7</td>
</tr>
</tbody>
</table>

\( > 5\sigma \) at 100 fb\(^{-1}\)

Substructure can be useful for

- Search for Higgs in difficult channels, such as W/Z+H.  
  Butterworth, Davidson, Rubin, Salam, 0802.2470
- From NP decay.
  - Resonance ttbar.
  - SUSY.
  - Top partner to Higgs.
  - Z’ to WW, Zh...  
  Talk by Kribs.
- Kribs, Martin, and Roy, 1012.2886
- Cui, Han, Schwartz, 1012.2077
  Katz, Son, Tweedie, 1010.5253
Substructure can be useful for

- Search for Higgs in difficult channels, such as W/Z+H.  
  Butterworth, Davidson, Rubin, Salam, 0802.2470
- From NP decay.
  - Resonance ttbar.
  - SUSY.
- Top partner to Higgs.
- Z’ to WW, Zh...
  
  Talk by Kribs.
  Kribs, Martin, and Roy, 1012.2886
  Cui, Han, Schwartz, 1012.2077
  Katz, Son, Tweedie, 1010.5253

- Any maybe
Substructure can be useful for

- Search for Higgs in difficult channels, such as $W/Z+H$.
  
  Butterworth, Davidson, Rubin, Salam, 0802.2470

- From NP decay.
  
  - Resonance $ttbar$.
  
  - SUSY.
  
  Talk by Kribs.

- Top partner to Higgs.
  
  Kribs, Martin, and Roy, 1012.2886

- $Z'$ to WW, Zh...
  
  Cui, Han, Schwartz, 1012.2077
  
  Katz, Son, Tweedie, 1010.5253

- Any maybe
  
  - boosted di-jet resonance of mass 150 GeV.
Substructure can be useful for

- Search for Higgs in difficult channels, such as $W/Z+H$.  
  Butterworth, Davidson, Rubin, Salam, 0802.2470

- From NP decay.
  - Resonance $t\bar{t}$.  
  - SUSY.
  - Top partner to Higgs.  
  - $Z'$ to $WW, Zh...$
    
  Talk by Kribs.
  Kribs, Martin, and Roy, 1012.2886
  Cui, Han, Schwartz, 1012.2077
  Katz, Son, Tweedie, 1010.5253

- Any maybe
  - boosted di-jet resonance of mass 150 GeV.
  - TeV axi-gluon decaying into boosted tops.
Substructures: current and future

• Boosted (almost) everything already.

• Rather simple techniques. Currently,
  • Combine, optimize, test.
  • LHC experimental groups have started using them.

• New substructure (jet shape).
  • Color flow...

• Better theoretical (QCD) understanding of the substructure.
  • Particularly for the QCD jets.
  • SCET....
Boost $W$ at early LHC

$\not{p}_{T}^{j} > 400$ GeV

$\sim 5\sigma$ at $1$ fb$^{-1}$

Adam Davison, talk at Boost 10
Conclusions

• Better handles on the hadronic final states are instrumental for discovery at the LHC.

• Based on consideration of QCD radiation, we proposed a set of carefully constructed new jet algorithms and substructure variables.
  
  • Much improved performance, jet mass, jet shape, etc.
  
  • They can significantly enhance new physics signals in many important new physics channels.
    
    • Boosted or “slow” hadronic tops, WW scattering, Higgs search, heavy squark...

• A promising direction. Stay tuned.
ISR tagging

- Pick: distinct PT
  \[ \frac{\max(p_{T_i}, p_{T_j})}{\min(p_{T_i}, p_{T_j})} > 2, \ i \neq j \]

- Or, distinct rapidity
  \[ |y_i - y_j| > 1.5, \ i \neq j \]

- Or
  \[ \frac{\max(\Delta_i, \Delta_j)}{\min(\Delta_i, \Delta_j)} > 1.5, \ \Delta_i = m_i/p_{T_i} \]

- And, not central
  \[ |y_i| > 1 \]

- Separated from others
  \[ |y_i - y_j| > 0.5 \]

- FSRs must be similar

  \[ \frac{p_{T_j}}{p_{T_k}} < \rho + \frac{0.5}{1 - \alpha} \]

  \[ p_{T_{j(k)}} = \max(\min\{p_{T_l}, l \neq i\}, \quad \rho = 2(3) \text{ for } N_f = 2(4), \]

  \[ \alpha = \frac{\min(p_{T_i}, E_T)}{\max(p_{T_i}, E_T)} \]
Why is it possible to gain?

- MI, UE, and pile-up are incoherent soft background. They can be effectively removed with a cut on soft radiation.

- Both FSR (want to keep) and ISR (want to discard) have soft radiation, but
  
  - ISR: \[ d\sigma \propto \frac{dp_T^{ISR}}{p_T^{ISR}} \]

  - FSR is controlled by both collinear and soft singularities:
    \[ d\sigma \propto \frac{d(\Delta R)}{\Delta R} \times \frac{dp_T^{FSR}}{p_T^{FSR}} \]

  - Therefore, a soft cut relative to the jet energy flow could enhance FSR relative to ISR.
Planar Flow

\[ I_{wl}^{kl} = \sum_i w_i \frac{p_i,k p_i,l}{w_i w_i} \]

\[ \lambda_1, \lambda_2 : 2 \text{ eigenvalues of } I_{wl}^{kl} \]

\[ \text{Pf} = \frac{4\lambda_1\lambda_2}{(\lambda_1 + \lambda_2)^2} \]
Top tagging: jet mass

- QCD jets also have mass.

\[
\langle M^2 \rangle \approx \int \frac{d\theta^2}{\theta^2} \int dz \ p_T^2 z (1 - z) \theta^2 \ \frac{\alpha_s(p_T)}{2\pi} P(z) \Theta(\Delta R - \theta)
\]

\[
\approx C \frac{\alpha_s}{\pi} p_T^2 (\Delta R)^2
\]

Useful. Additional variable?

QCD Dijet vs. Top Resonance sweeping \( Q_{\text{max}} \)

Jet mass.

Using jet mass only.
Boosted top is also hard to identify.

- Heavy resonance decay.

\[ E_{\text{top}} \approx \frac{M_X}{2} \]

- For \( m_{t\bar{t}} > 3 \text{ TeV} \), > 90\% events with at least one top fully collimated.
- Large fraction of events "2-object"-like. QCD \( b\bar{b}, jj \) background.
- A few \% with lepton isolation

B. Lillie, L. Randall, and LTW, hep-ph/0701166
Top tagging efficiency


Performance of different $z$ variables.

- $z$-variable gives an additional about factor of 2 enhancement in performance.

- Together with jet mass, an enhancement of 100 of S/B is possible.

Related studies:
Gustaaf H. Brooijmans, arXiv:0802.3715; CMS, CMS PAS JME-09-001
Boosted tops.

- Tops are interesting!
  - Top plays an important role in electroweak symmetry breaking.
  - Top generically couples to heavy new resonances which is an important part of TeV new physics.

- Examples.
  - Composite top couples strongly to other composite resonances.
    Many examples.

- New heavy scalars couple like Higgs.
  For example: A. Manohar and M. Wise, hep-ph/0606172

- A good example of subjet techniques.
\[ d_{12} < d_{13} < d_{23} < d_{(1,2,3)B} < d_{i4} \]
$d_{12} < d_{(1,2)B} < d_{i4}$
Done!
A closer look at the soft radiations

- ISR scale with the hard collision
  \[ (p_T^{\text{ISR}})_\text{max} \simeq |q|^2 \leq \Lambda_{\text{hard}}^2 \]
  \[ \Lambda_{\text{hard}} : \text{hard interaction scale} \]

\[ \frac{d\sigma}{d(p_T^{\text{ISR}})^2} \simeq \frac{1}{(p_T^{\text{ISR}})^2} \left( \alpha_s \log \left( \frac{\Lambda_{\text{hard}}^2}{(p_T^{\text{ISR}})^2} \right) + O(\alpha_s^2) \right) \]
\[ \langle p_T^{\text{ISR}} \rangle \propto (p_T^{\text{ISR}})_\text{max} \]

- MI, UE, and pileup “incoherent”, independent of the hard collision scale.

- A “universal” soft background.
  \[ \delta(p_T^i) \simeq \Lambda_{\text{soft}} \left( \frac{\Delta R^2}{2} + \ldots \right) \]
Substructure, z-finding

- Jet clustering history is approximately the inverse of parton shower.

- \( z \rightarrow 0 \) for QCD jet, \( z \) finite for top jet.

\[
z = \frac{\min(E_1, E_2)}{E_0}
\]

\( Q : \) jet mass

- Jet clustering history is approximately the inverse of parton shower.