Jet Substructure in the Pre-ILC Era

Brock Tweedie
Boston University
18 March 2011
We’re Making Progress

• Everything “heavy” becomes “light” eventually
  – 1970’s: Taus, c-quarks, and b-quarks freshly discovered, exploring the multi-GeV mass scale!
  – 1980’s: Tau-jets, c-jets, and b-jets are secondary signals of interesting new things
    • jet substructure with small # hadrons
We’re Making Progress

• Everything “heavy” becomes “light” eventually
  – 1980/90’s: W-bosons, Z-bosons, and top-quarks freshly discovered, exploring the ~100-GeV mass scale!
  – 2010’s: W-jets, Z-jets, and top-jets (and Higgs-jets if we’re lucky) are secondary signals of interesting new things
    • jet substructure with large # hadrons
The New Jets

$E_{CM} \sim 2m_W$

$E_{CM} \gg 2m_W$

jet

W-jet

Friday, March 18, 2011
The New Jets

\[ E_{CM} \sim 2m_W \]

\[ E_{CM} >> 2m_W \]

LHC

Friday, March 18, 2011
The New Jets

\[ E_{CM} \sim 2m_W \]

\[ E_{CM} \gg 2m_W \]

Friday, March 18, 2011
Some Relevant Angular Scales

- **W->qq**
  - $\Delta R \sim 2m/p_T$
  - $p_T = 400$ GeV $\Rightarrow \Delta R \sim 0.4$

- **Typical LHC jet size**
  - $\Delta R \sim 0.4$

- **HCAL cells**
  - $\Delta R \sim 0.1$

- **ECAL cells**
  - $\Delta R \sim 0.02$

- **Tracker**
  - $\Delta R \sim 0.001$
Angular Scales in Top Decay
QCD is our enemy

- collinear splittings => easy to get several “hard” partons emitted into a small solid angle
- jet mass \( \sim \sqrt{\alpha_s} \cdot p_T \cdot R \sim 100 \text{ GeV} \)
So What?

• QCD is our enemy
  – All jets are full of junk!
  – Size too large compared to $\Delta R \Rightarrow$ messed-up jet-mass measurements
Goals of “All-Hadronic” Substructure

• Figure out the relevant $\Delta R$ scales on an event-by-event basis
  – jets $\rightarrow$ subjets

• Discrimination against QCD parton splittings
  – multibody kinematics at small angles

• Keep the radiation we want, toss the junk
Brief Early History (2-Body)

• 1994: Seymour
  – 600 GeV Higgs->WW
• 2002: Butterworth Cox Forshaw
  – TeV-scale WW scattering
• 2007: Butterworth Ellis Raklev
  – W-jets from SUSY decays
• 2008: Butterworth Davison Rubin Salam
  – (semi)boosted Higgs->bb in LHC Higgsstrahlung
Butterworth, Davison, Rubin, Salam (BDRS)

\[ R = 1.2 \text{ "fat jet"} \]

\[ p_T(V) \sim p_T(h) \sim 200 \text{ GeV} \ (\Delta R \sim 1.0) \]

high-\( p_T \) kills backgrounds faster than signal

also: Agrawal, Bowser-Chao, Cheung, Dicus, DPF Conf.1994:488-492
Butterworth, Davison, Rubin, Salam (BDRS)

**ATLAS TDR**
30/fb, $m_H = 100$ GeV

$Wh \rightarrow (l\nu)(bb)$

Zh $\rightarrow (l^+l^-)(bb)$
Zh $\rightarrow (\nu\bar{\nu})(bb)$

High $p_T$

S/sqrt(B) $\sim 4.5$ at 30/fb LHC14
Butterworth, Davison, Rubin, Salam (BDRS)

$Wh \rightarrow (\ell\nu)(bb)$

New ATLAS: $\sim 2.6\sigma$

1/fb LHC7 Higgs exclusion
Butterworth, Davison, Rubin, Salam (BDRS)

R=1.2 fat-jet  subjets  refined subjets

Friday, March 18, 2011
Cambridge/Aachen Clustering History
Jet Declustering
Jet Declustering

Friday, March 18, 2011
Jet Declustering

SUBJET #1

SUBJET #2

“hard” split
“soft” split
BDRS Declustering Criteria

- Mass-drop: \( \frac{\text{Max}(m_1,m_2)}{m_{1+2}} < 0.67 \)
- Symmetry: \( \frac{\text{Min}(\rho_{T1},\rho_{T2}) \cdot \Delta R_{12}}{m_{1+2}} > 0.3 \)
  - or, \( \frac{\text{Min}(\rho_{T1},\rho_{T2})}{\rho_{T1+2}} > 0.1 \)
BDRS Filtering

Recluster subjet constituents using $R \sim \Delta R_{12}/2$ to get rid of remaining junk (these are still big subjets!)
More “Jet Grooming” Strategies

• Pruning  (Ellis, Vermilion, Walsh)
  – jet is selectively reclustered, ignoring the junk (“bottom-up” approach)
• Trimming  (Krohn, Thaler, Wang)
  – remove regions of the jet with low energy density
    • recluster jet with miniature R (~0.2), throw away too-soft minijets
More “Jet Grooming” Strategies

• **Pruning**  (Ellis, Vermilion, Walsh)
  – jet is selectively reclustered, ignoring the junk (“bottom-up” approach)

• **Landscaping**  (Krohn, Thaler, Wang)
  – remove regions of the jet with low energy density
    • recluster jet with miniature R (~0.2), throw away too-soft minijets
Grooming Comparison on Top-Jets

![Graphs showing jet mass distributions for different grooming techniques.](image)

**BOOST 2010, arXiv:1012.5412**

complimentarity study: **Soper & Spannowsky, arXiv:1005.0417**

Friday, March 18, 2011
BDRS on Steroids:
$Z' \rightarrow$ Electroweak Bosons

Katz, Son, Tweedie, arXiv:1010.5253

Related study (heavy $h \rightarrow ZZ$): Hackstein & Spannowsky, arXiv:1008.2202

Friday, March 18, 2011
What is this Buying Us?: Substructure vs Traditional Jets

- Case study using particle-level $Z' \rightarrow WW \rightarrow (l\nu)(qq')$ and $W + \text{jets} \rightarrow (l\nu) + \text{jets}$ background events
  - Hunt for the hadronic $W$ using either decomposed fat-jet or hybrid of traditional monojet and dijet searches
Substructure vs Jets: Signal

![Graphs showing signal rates for different masses and substructure methods.]

- **BDRS**
- traditional dijet
- traditional monojet
- traditional combined

Friday, March 18, 2011
Substructure vs Jets: Background

PYTHIA 6.4 $Wq/Wg$, showered

<table>
<thead>
<tr>
<th>BDRS</th>
<th>traditional dijet</th>
<th>traditional monojet</th>
<th>traditional combined</th>
</tr>
</thead>
</table>

Friday, March 18, 2011
Internal Kinematics

\[ R = 0.4 \]

- **W-like**
- **QCD-like**

Friday, March 18, 2011
Internal Kinematics

W-like

QCD-like

R=0.4
Internal Kinematics

$R=0.4$

$W$-like

QCD-like
### W/Higgs Tag Rates from Substructure

**SIGNAL**

<table>
<thead>
<tr>
<th></th>
<th>$p_T \approx 500$ GeV</th>
<th>$p_T \approx 1000$ GeV</th>
<th>$p_T \approx 1500$ GeV</th>
</tr>
</thead>
<tbody>
<tr>
<td>$W$</td>
<td>76%</td>
<td>77</td>
<td>72</td>
</tr>
<tr>
<td>$h$</td>
<td>59</td>
<td>61</td>
<td>62</td>
</tr>
</tbody>
</table>

**BACKGROUND**

<table>
<thead>
<tr>
<th></th>
<th>$p_T \approx 500$ GeV</th>
<th>$p_T \approx 1000$ GeV</th>
<th>$p_T \approx 1500$ GeV</th>
</tr>
</thead>
<tbody>
<tr>
<td>quark $\to W$</td>
<td>6.5%</td>
<td>6.5</td>
<td>5.9</td>
</tr>
<tr>
<td>quark $\to h$</td>
<td>6.8</td>
<td>5.6</td>
<td>5.8</td>
</tr>
<tr>
<td>gluon $\to W$</td>
<td>10.4</td>
<td>8.3</td>
<td>7.4</td>
</tr>
<tr>
<td>gluon $\to h$</td>
<td>10.5</td>
<td>8.8</td>
<td>7.4</td>
</tr>
</tbody>
</table>

Counting events in windows [65,95] GeV for $W$ and [100,140] GeV for Higgs (simple detector model applied)
Why are the Mistag Rates Decreasing with $p_T$?

a hint of sensitivity to global color connections
Model-Independent $Z' \rightarrow WW$ Discovery Reach (LHC 14)

Arrows indicate custodial RS model
Line indicates S/B=1

Earlier result (arXiv:0709.0007): Need 1000/fb

Friday, March 18, 2011
Data-Driven W Mistag (CMS)

using pruning + BDRS style cuts on the final merging

Guofan Hu
Friday, March 18, 2011
Tag Rate and Tag vs Mistag

Monte Carlo signal efficiency for W tagging

Monte Carlo mistag rate vs tag efficiency
Boosted Tops

CMS Preliminary
36 pb\(^{-1}\) at \(\sqrt{s} = 7\ TeV\)
e, \(\geq 4\) jets, = 1 b-tag

Events / 50 GeV/c\(^{2}\)
Top Tagging

• Tear the jet down one more layer (or rebuild it from bottom-up)
  – 3 or 4 subjets
• >2 body kinematics
  – subjet pairwise invariant masses (look for the W, veto small-mass pairs)
  – reconstruct top and W decay angles
• Groom as needed
1 TeV Top-Jet Gallery
Data-Driven Top Mistag (CMS)

Hopkins top-tagger on (bad) simulation

CMS top-tagger on data

Jim Dolan
Friday, March 18, 2011
CMS Tag vs Mistag

Simulation study
Many Other Techniques...

- ATLAS: Brooijmans
- Thaler/Wang: Thaler, Wang
- + HEPtagger: Plehn, Spannowsky, Takeuchi, Zerwas

**BOOST 2010, arXiv:1012.5412**

Friday, March 18, 2011
Seeing W-Jets in Color... in Many Dimensions

Many weakly color-sensitive discriminants

One powerful discriminant (Boosted Decision Tree)

Cui, Han, Schwartz, arXiv:1012.2077

Friday, March 18, 2011
Seeing W-Jets in Color...in Many Dimensions

$p_T \sim 500$ GeV
Top-Tag with Dipolarity

Hook, Jankowiak, Wacker,
arXiv:1102.1012

Friday, March 18, 2011
Shower Deconstruction

Soper & Spannowsky, arXiv:1102.3480
Friday, March 18, 2011
Jet Shape Ideas

- N-subjettiness: Kim, Thaler & Van Tilberg
  - More amenable to PQCD calculations?
- Jet-angularities, planar flow, etc: Almeida, Lee, Perez, Sterman, Sung, Virzi
  - Ditto
- Template overlap: Almeida, Lee, Perez, Sterman, Sung
More Complicated Processes

• Top-Higgsstrahlung: Plehn, Spannowsky, Salam
• Boosted Higgs in SUSY: Adam Martin + UO
New Physics Substructure

• RPV neutralino: Butterworth, Ellis, Raklev, Salam
  – an alternative 3-quark final-state
• Unburied Higgs: Falkowski, Krohn, Shelton, Thalapillil, Wang; Chen, Nojiri, Sreethawong
  – $h \rightarrow aa \rightarrow (gg)(gg)$
  – exploit color isolation (analogous to taus)
Jets With Embedded Leptons

- Semileptonic boosted tops: Thaler & Wang, Rehermann & Tweedie
- Boosted $h \rightarrow \tau \tau$: Katz, Son, Tweedie
Substructure at ILC or CLIC?

• No UE, no ISR
  – no trimming, no pruning, no filtering (?)
• Color singlet machine => “lepton+jets” backgrounds are of a very different quality
  – e.g., Wq, Wg absent
  – W+jets still there, but much lower rate
• Clearly useful for all-hadronic search channels (q-qbar backgrounds)
• Boosted SUSY?
Summary

• Lots of ideas to beat QCD and identify EW boson jets & top jets, and all kinds of other good stuff
• So far mistag rates look sane in data
  – Stick around for Marcel’s talk for an actual experimentalists’ perspective on all of this
• Relevance for ILC1000 or CLIC?
Extras
CDF Discovery of Semi-Leptonic WW/WZ

CDF arXiv:1008.4404 (4.6/fb of data)
CDF Exclusion of WW/WZ Resonance

CDF arXiv:1004.4946 (2.9/fb of data)

(Warped KK Z’ at LHC energy)

(Warped KK Z’ at Tevatron energy)

Friday, March 18, 2011