Boosted Top

Salvatore Rappoccio

Johns Hopkins University / CMS / CDF
Fine Print

• This talk is meant to stimulate discussion

• Very active area of HEP, many new and exciting developments
  – If I missed something don’t be offended!

• Thanks to the many people who have worked on this!
Assorted References

- **Theoretical motivations:**
  - arXiv:0803.1160

- **Algorithmic Developments:**
  - arXiv:0806.0023
  - arXiv:0802.2470
  - arXiv:0806.0848v2
  - arXiv:0903.5081

- **Tevatron Developments:**
  - arXiv:0902.3276
  - arXiv:0709.0705v1
  - arXiv:0804.3664v1
  - arXiv:0803.3256v1

- **LHC Developments**
  - ATL-COM-PHYS-2008-001
  - ATL-PHYS-PUB-2006-002
  - ATL-PHYS-PUB-2006-033
  - CMS-TOP-2009-009
  - CMS-JME-2009-001
  - CMS-EXO-2009-002
  - CMS-EXO-2009-008
Motivation

- New physics scenarios often involve ttbar resonances
- Are we sensitive to hadronic decays of top?
  - Can we suppress the huge dijet background?
- What about tricks with leptonic decays also?
  - “Traditional” methods are (eventually) ineffective due to crowded environment

hep-ph/0612015v1
SBM-07-001
Experimental Status

CDF

$M_{Z'} > 700 \text{ GeV}/c^2$

20 Nov 09  Top@Tevatron4LHC
Experimental Status

Reliant on “traditional” reconstruction methods:
- semileptonic mode
- b-tagging
- lepton isolation
- jet counting

\[ M_{Z'} > 700 \text{ GeV/c}^2 \]
Semileptonic Sensitivities at LHC

CMS Simulation

$\sigma_Z \times Br (pb)$

$100 \text{ pb}^{-1}$

CMS-TOP-2009-009

CMS preliminary

$\sigma \times Br (Z^0 \rightarrow \tilde{t} \tilde{t})$ [pb]

CMS-EXO-2009-008

$g^* \rightarrow tt$

$30 \text{ fb}^{-1}$

ATL-PHYS-PUB-2006-002

$\sigma^2 Br (fb)$

ATL-PHYS-PUB-2006-002

@Tevatron
Semileptonic Sensitivities at LHC

100 pb⁻¹

Kinematic Fit

CMS-TOP-2009-009

ΔR, p_T^{Rel}, H_T^{lep}

CMS-EXO-2009-008

Double B-tagging

ATL-PHYS-PUB-2006-002

Still reliant on “traditional” reconstruction methods...
Semileptonic Sensitivities at LHC

CMS Simulation

$\sigma \times BR (pb)$

100 pb$^{-1}$

Kinematic Fit

$\Delta R, p_T^{\text{Rel}}, H_T^{\text{lep}}$

Is there anything else we can think of?

Double B-tagging

Still reliant on “traditional” reconstruction methods…
New Idea: Top Jets

- Even moderate parent masses will result in collimated top “jets”
- Substructure can still be resolved
  - Subjets
- Two mass scales involved:
  - Top mass
  - W mass
  - Affects angular distribution of subjets
- For QCD, only gluon emission scales
  - Tend toward “zero” mass, smaller angular separation

20 Nov 09  Top@Tevatron
New Idea: Top Jets

- Even moderate parent masses will result in collimated top "jets"
- Substructure can still be resolved
  - Subjets
- Two mass scales involved:
  - Top mass
  - W mass
    - Affects angular distribution of subjets
- For QCD, only gluon emission scales
  - Tend toward "zero" mass, smaller angular separation

Much Work in the Area of Subjets!

Examine this in detail
New Idea: Top Jets

- How to get at substructure?

- Cone-like algorithms are not very good
  - Start with “large pt” and then incorporate “small pt”
  - This is going to wash out substructure

arXiv:0806.0848v2
New Idea: Top Jets

• How to get at substructure?

• Cone-like algorithms are not very good
  – Start with “large pt” and then incorporate “small pt”
  – This is going to wash out substructure

arXiv:0806.0848v2
New Idea: Top Jets

• How to get at substructure?

• Cone-like algorithms are not very good
  – Start with “large pt” and then incorporate “small pt”
  – This is going to wash out substructure
New Idea: Top Jets

• How to get at substructure?

• Cone-like algorithms are not very good
  – Start with “large pt” and then incorporate “small pt”
  – This is going to wash out substructure

arXiv:0806.0848v2

20 Nov 09 Top@Tevatron4LHC 15
New Idea: Top Jets

• How to get at substructure?

• Sequential combination algorithms produce “subjets” naturally in the course of the algorithm!
  – Exploit the clustering sequence

arXiv:0806.0848v2
New Idea: Top Jets

• How to get at substructure?

• Sequential combination algorithms produce “subjets” naturally in the course of the algorithm!
  – Exploit the clustering sequence

arXiv:0806.0848v2
New Idea: Top Jets

• How to get at substructure?

• Sequential combination algorithms produce “subjets” naturally in the course of the algorithm!
  – Exploit the clustering sequence

Look at sequential combination in more detail

arXiv:0806.0848v2
Sequential Combination

- Pairwise examination of input 4-vectors
- Calculate $d_{ij}$
  \[ d_{ij} = \min(n_k_{ti}, n_k_{tj}) \frac{\Delta R_{ij}^2}{R^2} \]
  - $N = 2$: $k_T$
  - $N = 0$: Cambridge Aachen
  - $N = -2$: anti-$k_T$
- Also find the “beam distance”
  \[ d_{iB} = k_{Ti}^n \]
- Find min of all $d_{ij}$ and $d_{iB}$
- If min is a $d_{ij}$, merge and iterate
- If min is a $d_{iB}$, classify as a final jet
- Continue until list is exhausted

fastjet manual has good overview
Substructure Finding

**Top-down**
- “Peel off” layers of jet clustering sequence
- Throw away soft and colinear clusters

4 \rightarrow 3 \rightarrow 2 \rightarrow 1
- arXiv:0806.0848v2

**Bottom-up**
- Start from “ground up” of clustering sequence
- Throw away soft and colinear clusters

1 \rightarrow 2 \rightarrow 3 \rightarrow 4
- arXiv:0810.0934

Comparable results for both
Substructure Finding

**Top-down**
- “Peel off” layers of jet clustering sequence
- Throw away soft and collinear clusters

4 -> 3 -> 2 -> 1
- arXiv:0806.0848v2

**Bottom-up**
- Start from “ground up” of clustering sequence
- Throw away soft and collinear clusters
1 -> 2 -> 3 -> 4
- arXiv:0903.5081

Comparable results for both

In the interest of time, discuss top-down only

20 Nov 09 Top@Tevatron4LHC
Substructure Finding

- Exploit kinematics!
  - Angular information
  - Mass of top, W
  - Mass "drop" from full jet
  - Energy scale at which decomposition occurs
  - ...

- All provide **very powerful** discrimination against generic QCD

---

ATL-COM-PHYS-2008-001

CMS-JME-009-01

arXiv:0806.0848v2
All-Hadronic Sensitivities at LHC

Comparable sensitivities to semileptonic mode for low mass,
Better sensitivity at high mass

NB: Y-splitter ATLAS analysis didn’t post sensitivities but the analysis is active

20 Nov 09 Top@Tevatron4LHC
The Devil is in the Cliche

- Improvements in reconstruction
- How to characterize top-taggers in data
- Combining information and channels
- Wide resonances
Improvements in Reconstruction

- Use of finer-grain angular resolution
  - “Particle flow”
  - Unclustered track hit information
- Subjet energy corrections
Top “Tag” Characterization

• **Efficiency**
  - Find a sample of top jets
  - Find a data-to-MC “scale factor”
  - Apply to other samples

• **Fake rate**
  - Find an unbiased signal-depleted sample of jets
  - Characterize fake rate as a function of jet characteristics

• **Candidate sample:**
  - Ttbar continuum in semileptonic sample
    - “Tag” leptonic top (with lepton)
    - “Probe” hadronic top

• **Candidate sample:**
  - QCD Dijets
    - “Anti-tag” one side
    - “Probe” away side
Combination Issues

- Data-driven **efficiency** in hadronic sample
  - From semileptonic signal region

- Data-driven **background** in semileptonic sample
  - From a sample with different gluon fraction
  - Needs a correction

- How to resolve these two issues in a data-driven way?

**Factor of > 2!**
Wide Resonances

- Interplay can happen between PDF’s and falling tail of wide resonances
- Manifests in a “dual hump” structure
- A “bump hunt” can rapidly turn into a “global excess” hunt
- Analyses need to be prepared to deal with this!
- Makes an accurate background model **absolutely critical**
  
  - Data-driven approaches are probably the only viable option for these cases
Discussion

• This was meant to stimulate discussion

• The results were arranged pedagogically instead of historically
  – May have left out some pieces, if so, please bring them up so we can discuss them!

• Now the conversation is supposed to organically grow