Qjets

David Krohn (Harvard)

Rutgers 4/17/12

Outline

- Review of Jets & Jet Substructure
- Introduction to Qjets
- Example: Jet Pruning
- Future Directions
- Conclusions
Many jet substructure analyses employ trees

But, more than one tree can plausibly be associated with a jet

Typically, we use $k_T$ or $C/A$ to choose the “best” tree

However, if we force ourselves to only consider a single tree for each jet, we make ourselves more susceptible arbitrary choices of the jet algorithm

By looking at many trees for each jet, we can decrease random fluctuations and create a more powerful analysis
Review of Jets & Jet Substructure
Types of Algorithms

- There are two main classes of jet algorithm
  - Sequential recombinations
    - Combine four-momenta one by one
  - Cone algorithms
    - Stamp out jets as with a cookie cutter

Focus on these
Sequential Recombination

* Define a distance measure between every pair of four-momenta in an event (jet-jet distances)
  \[ d_{ij} \]

* Define a distance measure for each four-momenta individually (jet-beam distances)
  \[ d_{iB} \]
If smallest distance at any stage in clustering is jet-jet, add together corresponding four-momenta

- Otherwise take jet with smallest jet-beam distance and set it aside

- Repeat till all jets are set aside

- In this way, jets are constructed by pairwise recombinations - get a tree-like sequence at the end.
Coordinate System

\[ \eta = -\ln \left[ \cot \left( \frac{\theta}{2} \right) \right] \]
\[ d_{12} < d_{13} < d_{23} < d_{(1,2,3)B} < d_{i4} \]
$d_{12} < d_{(1,2)B} < d_{i4}$
Done!
Standard Recombination Algorithms

* $k_T$ 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-$k_T$ algorithm

$$d_{ij} = \min(p_{Ti}^{-2}, p_{Tj}^{-2}) \left( \frac{\Delta R}{R_0} \right)^2, \quad d_{iB} = p_{Ti}^{-2}$$
Approximate Jet Behavior:

\[ p_{TA} > p_{TB} \]

- **Hard to Soft**
  - anti-\(k_T\)

- **Near to Far**
  - C/A

- **Soft to Hard**
  - \(k_T\)
Tradeoffs

* $k_T \& C/A$

  * Pro: Cluster near to far (both) & soft to hard ($k_T$). Allows us to use parton shower heuristics to understand behavior.

  * Con: Jets can have perverse shapes, weird areas

* anti-$k_T$

  * Pro: Jets are cone-like. Area relatively well defined.

  * Con: The ordering of the shower has little or no physical significance.
Jet-Parton Correspondence

- Jets allow us to make the connection between what we calculate (feynman diagrams) and what we measure in the detector.

- For instance, we’d expect to see two jets for each h->b bbar decay.
Jets make this correspondence

What we calculate

What we measure

Hadrons

Hadrons

Event picture from http://atlas.ch/photos/events.html
However, this heuristic correspondence between jets and partons breaks down when things become collimated.
Kinematics of Boosted Particles

• The cone containing the decay products of a particle scales as

\[ R \sim \frac{2m_X}{p_T} \]

• At LHC energies, even the heaviest particles we know of (Top, W, Z, Higgs) become can become collimated.

• When this happens we say that they’re “boosted”.

• So we find that EW scale particles are clustered as a single jet as soon as their \( p_T \) exceeds a few hundred GeV.
Here one can see the effect - as we boost more and more (i.e. go to higher pT), the particles become more collimated.
Unboosted t-tbar pair

Boosted t-tbar pair

All three decay products of the top go into one jet

Figure source: http://www.pha.jhu.edu/groups/particle-theory/seminars/talks/F08/Yumiceva.pdf
Boosted Collider Physics

- This can be a problem!

- Most new physics models include heavy states at the TeV scale
  - If these decay down to $W/Z/t$, what do we do if everything’s collimated?

- Traditional answer: use the leptonic decays to avoid this mess.

- Modern answer: look inside the jet and make use of QCD to see if the jet came from a boosted heavy object.
Tools

- QCD jets look really different than the jets of boosted heavy objects.
- QCD has soft/collinear singularities.
- If we start with a high energy gluon/quark, it wants to emit soft/collinear gluons:
  \[
  P_{q \to qg}(z) = C_F \frac{1 + z^2}{1 - z}, \\
  P_{g \to gg}(z) = C_A \left[ \frac{1 - z}{z} + \frac{z}{1 - z} + z(1 - z) \right], \\
  P_{g \to q\bar{q}}(z) = T_R \left[ z^2 + (1 - z)^2 \right],
  \]
- Here \( P(z) \) measures how much a particle wants to emit another with energy fraction “\( z \)” (Altarelli-Parisi splitting fcns.).
* However, a high energy heavy particle (W/Z/t/h) just decays - it has no singularity.

Boosted Heavy Particle

- Hard splitting, energy shared equally

QCD Jet

- Softer splittings. Unequal sharing of energy (note only one hard center)
Moreover, QCD jets have a continuum mass distribution, while the jets of boosted heavy particles have a fixed mass.

These will form our main tools.

1. Jet radiation distribution

2. Jet mass

Qjets
Two Basic Approaches to Substructure

1. Consider only the two-dimensional distribution of energy in a jet
   ✤ Examples: Trimming & Filtering, N-Subjettiness, Jet substructure w/o trees

2. Try to associate a tree structure with a jet
   ✤ Allows one to use heuristic pictures of parton shower & decay chains.
   ✤ Examples: Pruning, energy sharing variables, mass drop
   ✤ However, the current procedure for constructing a tree is not ideal.
Mapping Jets to Trees

The energy distribution for a particular tree is unambiguous

![Energy distribution diagram]

But, more than one tree can correspond to the same energy distribution
How do we assign a particular tree to an energy distribution?

Standard answer: Use a well motivated algorithm like C/A or kT

Ideally, since both are well motivated algorithms they’ll give the same answer:
• However, sometimes the answers are very different.

• Considering only the kT or C/A tree introduces an element of randomness into this process, resulting in unnecessary fluctuations in the final state observable.

• Intuitively it makes sense that defining an observable in a way which reflects the ambiguity of this clustering should yield better results.
Solution: Sum over Trees

- We propose that rather than assigning a single number to each event, instead each event should contribute a distribution obtained by summing the observable over many trees.
- When we sum these together, the result is much more stable than the histogram we would have had if we just considered one number per event.
\[ = \sum \{ \text{various shapes} \} \]
Weights

- The only question is: when we add together the result obtained from different trees, how should we weight each tree’s contribution?

- Surely they should not all count equally. If they did, then why would we use kT or C/A to find our trees in the first place?

- In theory, one could weight each tree by the product of splitting functions and Sudakovs one would obtain from a parton shower.

- Work in progress.
Implementation

- Instead, we find a simpler Monte-Carlo procedure works quite well.

- As in a sequential recombination algorithm, assign every pair of proto-jets a distance measure $d_{ij}$.

- However, unlike a normal sequential algorithm (where the pair with the smallest measure is selected clustered), here we suggest that a given pair be randomly selected for merging with probability

$$
\Omega_{ij} \equiv \frac{1}{\Omega} \exp \left( -\alpha \frac{d_{ij}}{d_{ij}^{\min}} \right), \quad \alpha = \text{rigidity parameter}
$$

- Thus, paths which deviate from the CA or kT behavior are less likely to occur

- Repeat many (~100) times, till the distribution stabilizes
• The result is that you get many trees

• The probability of finding a given tree decreases as it becomes less $k_T$ or C/A like

• Available as a Fastjet plugin:

  http://jets.physics.harvard.edu/Qjets
IR/Collinear Safety

- As long as the rigidity variable (alpha) is non-zero, then infinitely soft or collinear particles will not change the observable at hand.

- How will this affect real analytical calculations?
  - Still unknown

- Perhaps there is a better, more theory-friendly weight?
Example: Boosted W-Jets with Pruning
Pruning

* Pruning was introduced to look for boosted heavy objects (e.g., tops, higgses, W’s, etc) by cleaning up their mass.

* Intuition: QCD has soft/collinear singularities. Wide-angle emissions should come from hard decays.

  * Remove all parts of the jet which are both soft and wide angle.

* Two main advantages:

  * Boosted objects see their mass reconstruction improved
  * Massive QCD jets (a large background) see their mass substantially decreased -> lower backgrounds
Pruning in Practice

- To run pruning:
  - Take the constituents of an ordinary jet (formed using any algorithm).
  - Recluster them using a modified version of C/A and $k_T$
  - When $C/A$ or $k_T$ says that a pair of subjets should be recombined, ask: are the two subjets separated by more than a fixed amount ($d_{cut}$) and is one much softer than the other ($p_{T1} / p_{T2} < z_{cut}$)?
  - If so, set aside the softer particle and don’t merge it with the main jet.
A Pruned Tree
Defining Reconstructed Tops – Search Mode

- A jet reconstructing a top will have a mass within the top mass window, and a primary subjet mass within the W mass window. Call these jets top jets.

- Defining the top, W mass windows:
  - Fit the jet mass and subjet mass distributions with (asymmetric) Breit-Wigner plus continuum.
  - The top and W windows are defined separately for pruned and not pruned – test whether pruning is narrowing the mass distribution

Figure source: [http://www.phys.washington.edu/users/ellis/USATLAS.pdf](http://www.phys.washington.edu/users/ellis/USATLAS.pdf)
Let’s see what happens when we modify pruning so that it runs over trees generated via the Qjet procedure.
**Example 1/3: Mass Measurement**

- As an example, let’s take a sample of boosted $W$ jets (pt>500), clean them up via jet pruning, and ask for the average jet mass.

- The uncertainty in this measurement goes down by ~1/3 when the technique described is applied.

- Need roughly half the luminosity to make a measurement of the same precision

\[ \delta\langle m \rangle \propto 1/\sqrt{N} \]

| $\alpha$ | $\frac{\delta\langle m \rangle|_{cl}}{\delta\langle m \rangle|_Q}$ |
|---------|-------------------|
| 0       | 1.32              |
| 0.01    | 1.31              |
| 0.1     | 1.25              |
| 1.0     | 1.10              |
| 100     | 1.03              |
**Example 2/3: Signal Discovery & Exclusion**

- **Signal** = boosted W-jets, $p_T > 500$
- **BG** = light QCD jets, $p_T > 500$
- Measure the signal size in a bin (here 70-90 GeV) and compare it to the size of the BG fluctuations (Poisson stats included)
- Need only $\sim 70\%$ the luminosity to have the same significance
  \[
  S/\delta B \propto \sqrt{N}
  \]

| \(\alpha\) | \(\langle S\rangle/\delta B|_Q\) | \(\langle S\rangle/\delta B|_{cl}\) |
|---|---|---|
| 0.0 | | 1.07 |
| 0.01 | | 1.13 |
| 0.1 | | 1.18 |
| 1.0 | | 1.14 |
| 100 | | 1.06 |
with

in

and hadronized events are grouped into 0

results with pileup, we have investigated its e

assuming a 14 TeV LHC. We employ the “DW” tune for
terminating the equivalent decrease in luminosity.

fitting our results to these parametric scalings, and de-
ing improvement obtained by employing path-integral jets by
improvement that the QCD parton shower would produce it.
selecting one of the candidate trees based upon the prob-
sically.

radiation does not map onto a single tree, not even clas-
a given tree yields a fixed distribution of radiation in a
allowing one to construct substructure observables by
thinking of jets as tree-like structures. However, while

Here we consider only signal events and look at the av-
find the mean mass using only a limited set of events.

Numerical results are presented in Table I, where one
can see that the decrease found in
implies that an
ect of the path-

Typically one partially overcomes this ambiguity by

The parton shower provides a very useful heuristic,

First, let us consider a measurement of
S/
R

J

kT

J

BG, kT

BG, PI

S/

m

B

PI

80

40

20

60

80

100

120

140

Jet Mass [GeV]

Cross Section [A.U.]

0.1

0.08

0.06

0.04

0.02

Jet Mass [GeV]

0.01

0.005

0.002

0.001

Jet Mass [GeV]

0

20

40

60

80

100

120

140

Jet Mass [GeV]

0

20

40

60

80

100

120

140

Jet Mass [GeV]

0

20

40

60

80

100

120

140

Jet Mass [GeV]

0

20

40

60

80

100

120

140

Jet Mass [GeV]

0

20

40

60

80

100

120

140

Jet Mass [GeV]
Example 3/3: Signal vs. Background Discriminant

- When there’s a “right answer” for a jet’s mass, most of the trees tend to center around that value.

- There’s a “right answer” for the pruned mass of a boosted particle’s jet, but not for a background QCD jet.

- The width of a mass distribution serves as a good signal to background discriminant!
Width to Mass Distribution

- volatility = width of pruned mass distribution

![Graph showing width to mass distribution](image)

<table>
<thead>
<tr>
<th>Volatility</th>
<th>W-Jets</th>
<th>QCD-Jets</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.01</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- The graph illustrates the width to mass distribution for W-Jets and QCD-Jets.
- The volatility cut is set as \( \alpha = 0.01 \).
- The graph shows a lower volatility for W-Jets compared to QCD-Jets, indicating a narrower mass distribution.
Future Directions

- Perhaps we should consider “summing” over multiple parameters, not just trees.
  - Jet radii, trimming parameters, etc.

- We’ve only looked at considering multiple tree structures for the radiation inside a jet.
  - Can this procedure be extended to an entire event?
  - Could this help with precision quantities like $y_{23}$?
Qanti-kT

- Work in progress (w/ D. Kahawala, M. Schwartz)
- Take anti-kT and perturb around it as with Qjets
- Final state is now different
  - Different jet four-momenta
  - Different jet multiplicities
eta, phi vs frequency, pT, 1TeV scalar, alpha= 0.001 akt m12= 794.047
eta, phi vs frequency, pT, 1TeV scalar, alpha = 0.01 akt m12 = 794.047
eta, phi vs frequency, pT, 1TeV scalar, alpha = 0.1, m12 = 794.047
eta, phi vs frequency, pT, 1TeV scalar, alpha= 1
akt m12= 794.047
eta, phi vs frequency, pT, 1TeV scalar, alpha = 100, m12 = 794.047
Significant Improvement in Stability

- $S/\Delta(B)$ is much larger than with traditional anti-$kT$.
- Still have more optimizations to play with
- Larger improvements as jet multiplicity increased
- Can make discoveries/exclusions much sooner!
Conclusion

- When we use C/A or k_T to associate a tree with a jet this is really just our “best guess” for the showering history.

- Sometimes these two algorithms return very different answers for the event at hand.
  - By choosing, e.g. the k_T answer over the C/A one, we introduce randomness into the picture, and the statistics are degraded.

- We propose that all trees be considered, each with a set weight, and a distribution obtained for each event (rather than a single number).
  - The results obtained from this are much less susceptible to unwanted fluctuations: equivalent to a ~2x increase in luminosity.
Backup
**Classical Stats: Poisson + Binomial**

\[
P_N(n) \equiv e^{-N} \frac{N^n}{n!} \quad B_\epsilon(n; r) \equiv nC_r \epsilon^r (1 - \epsilon)^{n-r}
\]

\[
F_\epsilon, N(r) \equiv \sum_{n=r}^{\infty} F_\epsilon, N(r|n) = e^{-N\epsilon} N^r \frac{\epsilon^r}{r!} \equiv P_N\epsilon(r)
\]

\[
\sigma_{cl} \equiv \langle r \rangle = \sum_{r=0}^{\infty} rF_{\epsilon_{cl}, N}(r) = N\epsilon_{cl}
\]

\[
\delta\sigma_{cl}^2 \equiv \langle (r - \langle r \rangle)^2 \rangle = \sum_{r=0}^{\infty} (r - N\epsilon_{cl})^2 F_{\epsilon_{cl}, N}(r) = N\epsilon_{cl}
\]

\[
\frac{\delta\sigma_{cl}}{\sigma_{cl}} = \frac{1}{\sqrt{N\epsilon_{cl}}}
\]
f_n(x) = \left( \prod_{i=1}^{n} \int_{0}^{1} dx_i f_1(x_i) \right) \delta \left( x - \frac{x_1 + \cdots + x_n}{n} \right)

\int_{0}^{1} dy y f_n(y) = \int_{0}^{1} dy y f_1(y) \equiv \epsilon_Q \quad \sigma_n^2 \equiv \int_{0}^{1} dy (y - \epsilon_Q)^2 f_n(y) = \frac{1}{n} \sigma_1^2

\sigma_Q \equiv \langle yn \rangle = \sum_{n=0}^{\infty} P_N(n) \int_{0}^{1} dy y n f_n(y) = \epsilon_Q N

\delta \sigma_Q^2 \equiv \langle (yn - \langle yn \rangle)^2 \rangle = \sum_{n=0}^{\infty} P_N(n) \int_{0}^{1} (y n - \epsilon_Q N)^2 f_n(y) = (\epsilon_Q^2 + \sigma_1^2) N

\frac{\delta \sigma_Q}{\sigma_Q} = \sqrt{\frac{1 + (\sigma_1/\epsilon_Q)^2}{N}}
Properties of Jets

What properties do we want our jets to have? Jets should be, at least,

1. *Boost invariant*

   Easily done - cluster using rapidity/phi coordinate system:

   \[ \eta = - \ln \left[ \cot \left( \frac{\theta}{2} \right) \right] \]

2. *IR/Collinear safe*
Want to make jets in a way that is insensitive to soft and collinear radiation (IR & Collinear Safe)

Necessary if we’re going to employ higher order corrections.

If jet algorithm is not IRC safe then cancellations between real and virtual diagrams will not take place

Example of an Unsafe Algorithm

* If we use calorimeter cells as seeds then even an infinitely soft emission can change the clustering behavior in a significant way:
Most models of new physics use the top quark in a special way.

Identifying energetic tops from new physics processes will be crucial in understanding BSM phenomena at the LHC.

If there are heavy states, the top will often be boosted.
Much work on Boosted Tops

Many approaches

1. Use jet shapes, analogous to event shapes (e.g. thrust & sphericity), to quantify how top-likeness of a jet.

Measure the radiation pattern.


\[ P_f = \frac{4\lambda_1 \lambda_2}{(\lambda_1 + \lambda_2)^2} \]

\[ I_{w}^{kl} = \sum_{i} E_i \frac{p_{i,k}}{E_i} \frac{p_{i,l}}{E_i} \]

Other approaches

2. Try to find subjets inside each top jet and impose kinematical constraints (using helicity structure, etc)

Tailor made analysis

3. See if first splitting in jet was QCD-like (soft emission) or top-like (hard emission)


Try to split a jet by running an algorithm backward

4. Take jet, work hard to clean it up, see if has a mass near the top’s.
