QJets

Tuhin S. Roy
University of Washington

with Steve Ellis, Andrew Hornig, David Krohn and Matt Schwartz

arXiv:1201.1914
work-in-progress

Monday, October 29, 12
QJets

An event → calorimeter cells → traditional jet clustering algorithm such as C/A, $k_T$ → jets
particle flow → jet-observables

Monday, October 29, 12
QJets

An event → calorimeter cells → particle flow → traditional jet clustering algorithm such as C/A, k_T → jets

[Graph showing distribution of jet masses with lines for CA and KT]
An event → calorimeter cells → particle flow → traditional jet clustering algorithm such as C/A, $k_T$ ➔ variations of observables as new observables.
Qjets is an idea that explores the dimension of clustering history.
Qjets is an idea that explores the dimension of clustering history.

It is a challenging task -- let us start with something simpler.
QJets

Qjets is an idea that explores the dimension of clustering history

Start with the constituents of a jet → jet algorithm

it is a challenging task -- let us start with something simpler
Boosted Jets and Substructure Analysis

- Applications in Higgs Search
- Pruning

Clustering vs QClustering

- QPruning
  - Applications
Recipe for boosted resonance search:

(if you know what you are looking for)

- Look for “boosted” jets
- Identify “interesting” jets
- Clean jets
Recipe for boosted resonance search:

(if you know what you are looking for)  \[ \text{ex. } h \rightarrow bb \]

- Look for “boosted” jets

the angular separation of the decay products \[ \Delta R \sim 2 m_h/p_{Th} \]

“boosted jets” refer to jets containing four-vectors separated by \[ \Delta R \sim 1.0 \] and with \[ p_T > 2 m_h \]
Recipe for boosted resonance search:

(if you know what you are looking for)  \( \text{ex. } h \rightarrow bb \)

- Identify “interesting” jets

Higgs jets should have “mass-drop”
Higgs jets should be double b-tagged
Boosted jets and substructure analysis

Recipe for boosted resonance search:

(if you know what you are looking for)  ex. h -> bb

- Clean jets

- signal jets contain ISR + UE + pile-up other than the decay products

- cleaning a jet involves guessing which components are not due to decay + FSR and getting rid of these
  - ex: filtering, pruning, trimming etc.
Recently, a new technique for light Higgses in associated production of Higgs + Z, W:

\[ \text{(Butterworth, Davison, Rubin, Salam '08)} \]

significance for \[ L = 30 \text{ fb}^{-1} \]

obtained by focusing on boosted Higgses, \[ p_{T,h} > 200 \text{ GeV} \]

significant of \[ 4.2 \sigma \] at \[ L = 30 \text{ fb}^{-1} \]

using jet-substructure for jets with \[ p_{T,h} > 200 \text{ GeV} \]

Jet with substructure

- subjets are significantly lighter than the jet
- splitting is not too asymmetric
- jet is double b-tagged

filtered
Ex. Higgs from top partners

\[ \sqrt{s} = 14 \text{ TeV}, \mathcal{L} = 10 \text{ fb}^{-1} \]
**Ex.**  Higgs from top partners

\[ \sqrt{s} = 14 \text{ TeV}, \mathcal{L} = 10 \text{ fb}^{-1} \]

\[ M_T = 800 \text{ GeV} \]

\[ S/\sqrt{B} = 5.2 \]
Recipe for boosted resonance search:

(if you don’t know what you are looking for)

- Look for “boosted” jets
- Identify “interesting” jets
- Clean jets

more important than ever

rest of the talk will be on how pruning can be made a more effective groomer.
Pruning

Start with the constituents of a given jet and rebuild the jet along C/A or $k_T$
Pruning

At every step of clustering check whether the branch to be added is soft and wide angled.

- if yes discard the softer four-vector.

\[
\text{soft if: } \frac{\min(p_{T_i}, p_{T_j})}{|p_{T_i} + p_{T_j}|} < z_{\text{cut}}
\]

\[
\text{wide-angled if: } \Delta R_{ij} > D_{\text{cut}}
\]
Pruning

Pruned Jet

Monday, October 29, 12
Pruning

(a) unpruned QCD jets
(b) pruned QCD jets
(c) unpruned top jets
(d) pruned top jets

arXiv:0912.0033v1

arXiv:1204.2488v1
Pruning

- Four-vectors that are pruned are actually branches of the tree.

- Pruned jets depend crucially on the tree-structure or the clustering algorithm used to construct the jet.

but who ordered the clustering algorithm?
Clustering

# of four-vectors/set

distinct sets of four-vectors
Clustering

distinct sets of four-vectors

# four-vectors/set

1
2
3
...
N-2
N-1
N

C/A

k_T
Clustering

Many paths remain unexplored

distinct sets of four-vectors

Monday, October 29, 12
Clustering

Many paths remain unexplored

A better formalism should explore all such paths

one needs to be clever since the total number of distinct trees is enormous

\[
\frac{(2N)!}{2^N N!}
\]

our prescription is QClustering
As in a sequential recombination algorithm, assign every pair of four-vectors a distance measure $d_{ij}$.

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

$$\Omega_{ij} = \frac{1}{N} \exp \left(-\alpha \frac{d_{ij}}{d_{\text{min}}} \right)$$

Repeat many (~100-1000) times, till the distribution stabilizes.
QClustering

\[ \Omega_{ij} = \frac{1}{N} \exp\left(-\alpha \frac{d_{ij}}{d_{\text{min}}} \right) \]

\(d_{ij}\): we take C/A or kT measure

\(\alpha \to \infty\) Classical regime: only path corresponding to \(d_{\text{min}}\) is selected

\(\alpha > 0\) physical regime: physical paths are preferred

\(\alpha \to 0\) democratic regime: all paths have same weight

\(\alpha < 0\) unphysical regime: physical paths are de-weighted
QClustering vs. Clustering

A collection of 4 vectors

traditional jet clustering algorithm such as C/A, $k_T$

one jet

QClustering N times

same jet with N tree-structure $= N$ Qjets
QClustering vs. Clustering

A collection of 4 vectors

clustering + pruning

one pruned jet

QClustering N times + pruning

N pruned Qjets

one pruned jetmass

N pruned jetmasses
QClustering vs. Clustering

Classical Pruning

A collection of 4 vectors

one pruned jet

QPruning

N pruned Qjets

one pruned jetmass

N pruned jetmasses
QClustering + Pruning

Ex. a hadronic W jet from WW events

The original jet is made from C/A algorithm with $R = 1.0$ and $p_T > 200\text{GeV}$
QClustering + Pruning = QPruning

Ex. a hadronic W jet from WW events

The original jet is made from C/A algorithm with $R = 1.0$ and $p_T > 200\text{GeV}$

How can this distribution be used?
Before we proceed, one comment about the choice of weight

$$\Omega_{ij} = \frac{1}{N} \exp\left(-\alpha \frac{d_{ij}}{d_{\text{min}}}\right)$$

Who ordered the choice of $d_{ij}$ and $\alpha$?
Before we proceed, one comment about the choice of weight

\[ \Omega_{ij} = \frac{1}{N} \exp \left( -\alpha \frac{d_{ij}}{d_{\text{min}}} \right) \]
Before we proceed, one comment about the choice of weight

\[ \Omega_{ij} = \frac{1}{N} \exp\left(-\alpha \frac{d_{ij}}{d_{\text{min}}}\right) \]

For \( 0.1 > \alpha > 0 \) our results are insensitive to the choice of \( \alpha \) and the form of \( d_{ij} \)
QPruning vs. Pruning

Let us take a sample jet

How can this distribution be used?

Simply use the shape of the distribution to discriminate signal from background

Use the distribution to reduce statistical fluctuations in measurements

Application in signal discovery

Application in determination of cross-section, mass etc.
**QPruning vs. Pruning**

Let us take a sample jet

How can this distribution be used?

Simply use the shape of the distribution to discriminate signal from background.

Use the distribution to reduce statistical fluctuations in measurements.

Application in determination of cross-section, mass etc.

**Application in signal discovery**
- When there is an intrinsic mass scale for a jet, the pruned jetmass is more or less robust under variation of paths.

- Signal jets with decay products of massive resonances have intrinsic mass scales.

- Even QCD jets with \( m/p_T \sim 1 \) have hard splittings and hence intrinsic mass scales.

- But background is dominantly due to QCD jets with \( m/pt < 1/2 \) - whose masses are highly volatile.
Application 1: discovery of $W$

When there is an intrinsic mass scale for a jet, the pruned jet mass is more or less robust under variation of paths.

W jet

QCD jet with $m/p_T < 1/2$
Application 1: discovery of $W$

volatility of a jet \[ \nu = \frac{\omega_p}{m_p} \]

$\omega_p = \text{width of jetmass distribution}$

$m_p = \text{averaged pruned jetmass}$

\[ \alpha = 0.01 \]

- W-Jets
- QCD-Jets
**Application 1: discovery of W**

A cut on \( \mathcal{V} \) decreases background significantly.
Application 1: discovery of $W$

a cut on $\mathcal{V}$ decreases background significantly
Application 1: discovery of $W$

a cut on $\mathcal{V}$ decreases background significantly

Unofficial comparisons
QPruning vs. Pruning

Let us take a sample jet

How can this distribution be used?

Simply use the shape of the distribution to discriminate signal from background

Use the distribution to reduce statistical fluctuations in measurements

Application in determination of cross-section, mass etc.

Application in signal discovery

Monday, October 29, 12
QPruning vs. Pruning

Consider candidates for a W jet

- Classical pruned jet mass distribution
- QPruned jet mass distribution

- Pruned mass is either in or out of the bin
- Tagging efficiency is either 0 or 1

Mass window for W

- Tagging efficiency is a number between 0 to 1
QPruning vs. Pruning

Consider candidates for a W jet

Mass window for W

Pruning -> QPruning

A transition from a discrete (binomial distribution) to a continuous distribution
QPruning vs. Pruning

Pruning --> QPruning

A binomial distribution --> a continuous distribution

Use the distribution to reduce statistical fluctuations in measurements
QPruning vs. Pruning

Pruning --> QPruning

A binomial distribution --> a continuous distribution

Use the distribution to reduce statistical fluctuations in measurements

How to measure statistical fluctuations?
### Statistical Fluctuation

Consider a large number of pseudo-experiments. The number of jets in these pseudo-experiments vary according to the Poisson distribution.

<table>
<thead>
<tr>
<th>Experiments</th>
<th>Measurements (e.g. # tagged jet)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expt #1</td>
<td>N₁</td>
</tr>
<tr>
<td>Expt #2</td>
<td>N₂</td>
</tr>
<tr>
<td>Expt #3</td>
<td>N₃</td>
</tr>
<tr>
<td>Expt #4</td>
<td>N₄</td>
</tr>
<tr>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>.</td>
<td>.</td>
</tr>
</tbody>
</table>

Measure the fluctuation $\delta N$.
**Application 2: CS measurement**

\[ \frac{\delta N}{\sqrt{N}} = \sqrt{\langle \tau \rangle} + \frac{\text{var}(\tau)}{\langle \tau \rangle} \]

\[ \langle \tau \rangle = \text{average of the distribution} \]

\[ \text{var}(\tau) = \text{variance of the distribution} \]
Application 2: CS measurement

- As an example, take a sample of \sim 10 boosted QCD jets and ask for number of jets in a mass bin.

- The uncertainty associated with cross-section measurement decreases from classical pruning to QPruning.

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

\[ \alpha = 10^{-2} \]

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>( \frac{\delta N}{\sqrt{N}} )</th>
<th>Relative luminosity required</th>
</tr>
</thead>
<tbody>
<tr>
<td>prune with C/A</td>
<td>\sim 1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>QPrune</td>
<td>0.72</td>
<td>0.52</td>
</tr>
</tbody>
</table>
**Application 3: mass measurement**

\[
\frac{\delta m_{\text{exp}}}{m_{\text{exp}}} = \frac{1}{\sqrt{N}} \sqrt{\frac{\text{var}(\tau)}{\langle \tau \rangle^2} + \frac{\text{var}(\mu \tau)}{\langle \mu \tau \rangle^2} + \frac{\text{cov}(\tau, \mu \tau)}{\langle \tau \rangle \langle \mu \tau \rangle} + O \left( \frac{1}{N} \right)}
\]

probability distribution of W jets as a function of \( \tau \) and \( \mu \)
**Application 3: mass measurement**

- As an example, take a sample of ~10 boosted W jets and ask for average jet mass.
- The uncertainty associated with mass measurement decreases from classical pruning to QPruning.
- Need less than half the luminosity to make a measurement of the same precision.

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Mass uncertainty [GeV]</th>
<th>Relative luminosity required</th>
</tr>
</thead>
<tbody>
<tr>
<td>prune with C/A</td>
<td>3.2</td>
<td>1.0</td>
</tr>
<tr>
<td>QPrune</td>
<td>2.4</td>
<td>0.58</td>
</tr>
</tbody>
</table>

$$\alpha = 10^{-2}$$
Future Directions

- In substructure physics, it still remains to be seen whether QClustering can be applied to other quantities such as mass-drop, $Y_{23}$ etc.

- QClustering has been done on the elements of a jet. We intend to extend it to an entire event.

- We need to find a formalism towards analytical calculations.
Works in progress

- QClustering has been done on the elements of a jet. We intend to extend it to an entire event.

work in progress with Ellis
also Kahawala, Krohn, Schwartz

Q-Anti-$k_T$ Clustering

Mass window for $W$

QClustering with $\alpha=0.1$

QClustering with $\alpha=0.01$

all di-jet masses in a $W$+jet event
Works in progress

- QClustering has been done on the elements of a jet. We intend to extend it to an entire event.

work in progress with Ellis
also Kahawala, Krohn, Schwartz

Q-Anti-$k_T$ Clustering

Mass window for $W$

QClustering with $\alpha=0.1$

QClustering with $\alpha=0.01$

helps in suppressing combinatorial background

all di-jet masses in a $W+2$jet events
- QPruning extended to an event (tt event)

Mass window for top

QPruning around CA with $\alpha=0.1$

Pruned jetmass for a top candidate (3-jet resonance)
Towards analytically calculation for Qclustering

(Hornig & Schwartz)