Boosted Hadronically Decaying W and Z bosons

Becky Chislett

IoP conference
9th April 2013
Aim: To measure the production cross section of a boosted, hadronically decaying W/Z boson and use this peak to study the effect of various grooming and substructure techniques

- Reconstructing hadronic decays at the LHC is difficult due to the large backgrounds. Look at high $p_T$ where this background rapidly drops off.

- Many beyond the standard model theories predict new particles with TeV masses which can decay into W/Z bosons. Expect to see boosted (high $p_T$) W/Z bosons.

- Reconstructing the known W/Z resonance is an important step in developing these boosted techniques.

- The observation of this peak provides a standard candle to study the effects of techniques to remove pileup and the underlying event. Offers a way to control the jet mass in pileup conditions. Particularly important as the levels of pileup increase.
Event Selection

Look for events where the decay products of the W/Z are contained within a single jet.

**Rule of Thumb:** separation between decay products: \( R \sim \frac{2m_X}{\rho_{XT}} \)

Look at Anti-\( k_T \) \( R=0.6 \) jets, \( p_T > 320 \text{ GeV} \)

The QCD background is a couple of orders of magnitude higher than the signal…

Look at the distribution of energy in the jet

Boost the jet back into its centre of mass frame, we expect:
- W/Z jet to have a back to back topology
- QCD jet to be more isotropic

Use jet shapes to distinguish signal from background

**Sphericity, Aplanarity, Minor Thrust**
Jet Shape Variables

**Sphericity**
A measure of how well the event lies on a single axis

**Aplanarity**
A measure of how well the event is contained within a plane

**Minor Thrust**
A measure of how much of the momentum lies on the axis minimising the longitudinal momentum

**Highly directional events:**  S, A, T $\rightarrow 0$

**Isotropic events:**  S, A, T $\rightarrow 1$
Extracting the Cross Section

Combine the variables into a likelihood and cut on this: 
\[ \text{LH} > 0.15 \]
Reject \(~90\%\) background, keep \(~55\%\) signal

The resulting mass distribution:

Work is in progress to perform a fit and extract the signal yield and hence perform a cross section measurement.
Jet Grooming and Substructure

Take this peak and consider the effect of applying various grooming and substructure techniques:

- Pruning
- Trimming
- Area Subtraction
- Splitting/Filtering

Three main questions – does applying grooming:

Improve the agreement between data and MC?
Improve the s/b that can be achieved?
Help decrease the pileup dependence?

Pileup: background from additional proton-proton interactions.
Pruning

Grooming techniques aim to remove the parts of the jet coming from pileup and the underlying event and leave only the hard structure


Recluster the jet (using the $k_T$ or C/A algorithm). Remove constituents that are

- **Wide angle:** $\Delta R_{j_1, j_2} > R_{cut}$
- **Soft:** $\frac{p_{T}^{j_2}}{p_{T}^{j_1+j_2}} < z_{cut}$

Consider $R_{cut} = 0.3$, $z_{cut} = 0.05$
Pruned variables

Look at the variables using the pruned constituents:

- Improvement in data-MC agreement
- Good signal/background discrimination
Pruned variables

Look how data compares with other MCs:

Generally better agreement, particularly for low values of sphericity across all MCs, except Herwig
Pruned mass

Form a likelihood discriminator from these three pruned variables:

Cut such that 90% of the background is rejected (for comparison to the default)

Flatter background distribution
Signal peak shifted to lower mass
Better agreement between data and MC
Pruning – pileup dependence

Pileup is the contribution from other proton-proton interactions

assess by looking at the distribution as a function of the number of vertices

Pileup dependence much reduced
Pileup dependence

Summary for all the grooming/substructure techniques studied:

Compare the ratio between low and high numbers of vertices for each technique.

Grooming in general reduces the pileup dependence, except in the case of splitting/filtering.
Look at the s/b as a function of the jet mass:

**Improvement in s/b, particularly in the case of trimming and splitting and filtering**
• The production cross section for a boosted hadronically decaying W/Z boson is in progress.

• The impact of various grooming techniques (Pruning, Trimming, Area Subtraction, Splitting/filtering) were studied using this peak. In general:
  
  • Grooming improves the agreement between data and MC
  • An improvement in the s/b can be made
  • Pileup dependence is reduced

• Grooming techniques look promising for controlling the jet mass in future analyses, especially as pileup become more important.
Recluster the jet using the $k_T$ algorithm with a smaller radius $R_{sub}$ and remove any subjets with $p_T^{subjet}/p_T^{jet} < f_{cut}$

Consider $R_{sub} = 0.2$, $f_{cut} = 0.03$

Look at the variables using the trimmed constituents:

good signal/background discrimination
better agreement between MC and data w.r.t before grooming
Form a likelihood discriminator from these three trimmed variables:

Cut such that 90% of the background is rejected (for comparison to the default)

Background slight distorted
Signal peak shifted to lower mass
Better agreement between data and MC

Look at variation with number of vertices:
Pileup dependence slightly improved
Area Subtraction

Calculate the average level of the background in an event \( (\rho = \text{median} \left( \frac{p_T,j}{A_j} \right)) \) and subtract this from the hard jets based on their area \( (p_{T,j}^{\text{(sub)}} = p_{T,j} - \rho A_j) \)

Where \( A_j \) is the area of a jet calculated by adding a uniform background of ghost particles and reclustering. The area of a jet is proportional to the number of ghosts it contains.

Apply this to each of the jets in the original selection:

Very good agreement between data and MC
No distortion to the background
Pileup dependence reduced a lot
CA split/filtering

Look for hard structure in a jet by
- Undoing the last clustering step in a C/A jet
- Require a sufficient mass drop and symmetric splitting
- Recluster with a smaller radius keeping the largest 3 subjets

Look at applying splitting/filtering procedure to the jets selected by the default analysis:
- Causes the background to peak in the same place as the signal
  - Due to the choice of $p_T$ cut and jet radius

1. Undo the last clustering step: jet \( j \) \( \rightarrow \) subjets \( j_1, j_2 \) with mass \( m_{j_1} > m_{j_2} \)
   Require: 
   \[ \delta R_{j_1,j_2} = \sqrt{\delta y_{j_1,j_2}^2 + \delta \phi_{j_1,j_2}^2} > R_{\text{split}} \]

2. If this has mass drop: \( \frac{m_{j_1}}{m_j} < \mu \) and fairly symmetric splitting: 
   \[ y_2 = \frac{\min(p_{tj_1}, p_{tj_2})^2}{m_j^2} \delta R_{j_1,j_2}^2 > y_2,\text{cut} \]
   continue, otherwise set \( j = j_1 \) and go back to 1

3. Recluster \( j \) with Cambridge Aachen radius 
   \[ R_{\text{filt}} = \min(R_{\text{min}}, \delta R_{j_1,j_2}/2) \]

4. Redefine \( j \) as the sum of the subjets \( s_i \)
   \[ j = \sum_{i=1}^{\min(n,3)} s_i \]
Consider the effect pruning has on the mass distribution using the ungroomed LH:

**Default LH cut > 0.16**

**Tighter LH cut > 0.25**
(reject 95% background)

Pruning has a large effect on the background shape, due to:
- the $p_T$ cut
- the choice of jet radius
Pileup dependence

Investigate how much the mass is affected by pileup by looking at the mass as a function of the number of primary vertices:

Slight reduction in pileup dependence in the case of pruning and trimming
Pileup dependence

Investigate how much the mass is affected by pileup by looking at the mass as a function of the number of primary vertices:

Pileup dependence much reduced by area subtraction
**Shapes**

**Sphericity** is defined as: 
\[ S = \frac{3}{2}(\lambda_2 + \lambda_3) \] 
where \( \lambda_3 \) is the largest eigenvector of the sphericity tensor:

\[ S^{\alpha\beta} = \frac{\sum_i p_i^\alpha p_i^\beta}{\sum_i |p_i|^2} \]

Highly directional events: sphericity \( \to 0 \)
Isotropic events: sphericity \( \to 1 \)

**Thrust:** 
\[ T = \frac{\sum_i |\hat{T}.p_i|}{\sum_i |p_i|} \]
where the major (minor) thrust axis, \( T \), is the direction which maximises (minimises) the sum of the longitudinal momenta of particles

Directional event: \( T_{\text{major}} \to 1, T_{\text{minor}} \to 0 \)
Isotropic event: \( T_{\text{major}} \to 0, T_{\text{minor}} \to 1 \)

**Aplanarity** is defined as: 
\[ A = \frac{3}{2} \lambda_3 \]
where \( \lambda_3 \) is the largest eigenvector of the sphericity tensor:

\[ S^{\alpha\beta} = \frac{\sum_i p_i^\alpha p_i^\beta}{\sum_i |p_i|^2} \]

Planar events: aplanarity \( \to 0 \)
Isotropic events: aplanarity \( \to 1 \)