Jets and boosted objects at the LHC

Mario Campanelli/ UCL
At the end of the first data-taking period of the LHC, after > 20/fb of data, and the discovery of a Higgs-like particle, why do I come here to talk about jets? Isn't it old, boring physics?

QCD may be old (well, not more than the Higgs mechanism!), but its experimental study can be extremely creative, and has been able to reinvent itself over the years (yes it would have been boring if we were using the same techniques as Tevatron Run I)

In the following I will highlight some aspects of jet physics through the lens of new ideas (some of which I had, most of which I would have like to have)
More specifically, I will talk about

- Algorithms, triggers, calibration
- Inclusive and dijet cross section
- Hadronic decays of boosted objects as a window to new physics
- Jet grooming techniques
- Boosted top
- Quark/gluon jet separation
Jet algorithms

Most of jet measurements at the Tevatron have been made with not fully adequate jet clustering.

Trigger was including in a jet all calorimeter clusters within a given radius of a major energy deposition.

Offline clustering followed the same technique, with some refinement in the case of overlapping jets.

Recombination algorithms too slow for either trigger or offline.

But cone algorithms were shown to be all infrared unsafe: the result critically depends on the threshold at which clusters can initiate jets.
An infrared-unsafe algorithm does not guarantee cancellation of infinities performed by high-order theory calculations, therefore no meaningful comparison with higher orders of theory is possible. All LHC publication now use the infrared-safe Anti-Kt algorithm, (Salam, Soyez), it looks trivial now but it was a very big step forward.
Infrared-safe triggering: breaking the ATLAS trigger strategy

Based on the concept of region of interest (ROI) L1 (calo or muon) tells L2 which parts of the detector had interesting signals, and only those will be read. Same between L2 and EF.

This approach is infrared-unsafe, and it was shown that a single cluster passing or not noise thresholds leads to different ROI configurations, therefore different triggered jets.

Not much could be done on the hardware-only L1 system, but we introduced full-scan clustering at EF (all events) and L2 (selected chains).
Full-scan anti-\(k_T\) jets from L1Calo towers
- read L1Calo instead of Calo ROS yielding increased L2-input rate
- included since run 193834

Adds flexibility & acceptance (multijets, fat jets, jet+X)
- reduces inefficiency associated with L1 sliding windows & L1/L2 ROI

We did our best to produce the most luxurious bread-and-butter you can think of.

Measure jets clustered with \(\text{akt04} \) and \(\text{akt06} \) almost to kinematic limit.

For very forward jets NLO calculation becomes negative, and we had to use a special scale definition to make sense of it.
Two-jet trigger strategy

Central and forward jet triggers are independent and with different turn-on and prescales.

No single-jet trigger strategy was optimal to select central-forward nor forward-forward events.

We divided dijet phase-space into hundreds of regions according to jet pt and eta. At least one jet was at trigger plateau. Equivalent luminosities for each region were computed using trigger prescales, and everything was combined back again accounting for overlaps.

This technique maximised the dijet output for every kinematical configuration.
JES uncertainty in 2010

Largest source of systematics for any steeply-falling distribution, and a lot of effort to reduce it. Dominated by statistics of single particle response. Now we use a much more refined calibration.
Without tracking, the only way to cross-check the JES in the forward region is central-forward jet balance, in the limit of vanishing third jet. Discrepancies $O(10\%)$ have been found, going in opposite directions between Pythia and Herwig (Alpgen + Herwig) showering. Additional systematic uncertainties applied.
Non-perturbative corrections

|\(|y| < 0.3\) |
|---|
|\(|y| < 4.4\) |

**ATLAS**

<table>
<thead>
<tr>
<th>(p_T) [GeV]</th>
</tr>
</thead>
<tbody>
<tr>
<td>20 30 2\times10^{2} 10^{3}</td>
</tr>
</tbody>
</table>
Systematic uncertainties: magnitude and correlations

$\int L \, dt = 37 \, pb^{-1}, \ \sqrt{s} = 7 \, TeV$

$|y| < 0.3$

$\text{ATLAS}$
All bin correlations have been provided in the form of nuisance parameters to maximise information to be used in Pdf fits.
Inclusive jet cross section for antikt 06 jets after detector unfolding. Pt range from 60 to 600 GeV, rapidity < 4.4
Ratio with NLO + soft corrections comparison between jet sizes (central)
Ratio with NLO + soft corrections comparison between jet sizes (forward)

No big differences between cone sizes once soft corrections are applied
Comparison with Powheg (0.4)

\[ \int L \, dt = 37 \text{ pb}^{-1} \]
\[ \sqrt{s} = 7 \text{ TeV} \]
anti-\( k_t \) jets, \( R = 0.4 \)

Data with statistical error
Systematic uncertainties
NLOJET++
\((\text{CT10}, \mu = \rho_T^{\text{max}}) \times \) Non-pert. corr.
POWHEG
\((\text{CT10}, \mu = \rho_T^{\text{Born}}) \odot \) PYTHIA AUET2B
POWHEG
\((\text{CT10}, \mu = \rho_T^{\text{Born}}) \odot \) PYTHIA Perugia201
POWHEG
\((\text{CT10}, \mu = \rho_T^{\text{Born}}) \odot \) HERWIG AUET2
POWHEG fixed order
\((\text{CT10}, \mu = \rho_T^{\text{Born}}) \times \) Non-pert. corr.
Dijet cross-section and ratio (0.6)
Comparison with Powheg
The 2011 dijet measurement
The big issue is controlling pileup, first measurement performed in the central region (|\eta|<2.8) to have tracking, and with \(pt>60\) GeV

More MC statistics and different tunes
ATLAS Preliminary
anti-$k_t$ jets, $R = 0.6$
$s = 7$ TeV, $\int L \, dt = 4.8$ fb$^{-1}$
2011 Data

Systematic uncertainties

NLOJET++
(CT10, $\mu = \rho_T \exp(0.3 \, y^*)$) × Non-pert. corr.

$\frac{d^2 \sigma}{d m_1 d y^*}$ [pb/TeV]

$2.0 \leq y^* < 2.5 \times 10^8$
$1.5 \leq y^* < 2.0 \times 10^6$
$1.0 \leq y^* < 1.5 \times 10^4$
$0.5 \leq y^* < 1.0 \times 10^2$
y$^* < 0.5 \times 10^0$
Boosted decays of heavy objects

At the LHC for the first time EW-scale particles can have such large momenta that their hadronic decays can be reconstructed as a single “fat” jet

The real excitement came when Butterworth Davison Rubin and Salam (PRL 100, 242001 2008) have shown that the neglected \( W/Z + H \rightarrow bb \) could be seen in the boosted configuration.

They suggested a jet substructure technique, now known as filtering:

- Start from a fat jet reconstructed with Cambridge-Aachen (only angular information used), then deconstruct the last steps of clustering
- Require the 2-jet configuration to have a large mass drop with respect to the fat jet
- Go back one more step, and keep the three hardest subjets (to remove soft components)
Other jet grooming techniques

- **Pruning** (Ellis Vermillon Walsh 0912.0033)
  Take a fat jet, and recombine constituents after removing wide angle and soft ones

- **Trimming** (Krohn, Thaler, Wang 0912.1342)
  build fixed-radius subjets using $k_t$, then remove those with $p_T/p_T^{jet} < f_{cut}$
Simulated grooming performances

ATLAS Preliminary - Simulation
anti-k, LCW jets with R=1.0, Dijets (POWHEG+Pythia)

ATLAS Preliminary - Simulation
C/A, R=1.2

ATLAS Preliminary - Simulation
anti-k, LCW jets with R=1.0, Dijets (POWHEG+Pythia)

ATLAS Preliminary - Simulation
C/A, LCW jets with R=1.2, Dijets (POWHEG+Pythia)
Effect of grooming on data (QCD)

**ATLAS Preliminary**

- **Data 2011, Lct = 1 fb⁻¹**
- **anti-kt LCG jets with R=1.0**
- **600 ≤ p_T ≤ 800 GeV, |η| < 0.8**

- No jet grooming
- \( R_{cut} = 0.01, R_{sub} = 0.3 \)
- \( R_{cut} = 0.03, R_{sub} = 0.2 \)
- \( R_{cut} = 0.05, R_{sub} = 0.2 \)

**ATLAS Preliminary**

- **Data 2011, Lct = 1 fb⁻¹**
- **anti-kt LCG jets with R=1.0**
- **600 ≤ p_T ≤ 800 GeV, |η| < 0.8**

- No jet grooming
- \( R_{cut} = 0.10, z_{cut} = 0.05 \)
- \( R_{cut} = 0.20, z_{cut} = 0.05 \)
- \( R_{cut} = 0.30, z_{cut} = 0.05 \)

**ATLAS Preliminary**

- **Data 2011, Lct = 4.7 fb⁻¹**

- Dijets (Pythia)
- Dijets (POWHEG+Pythia)

**ATLAS Preliminary**

- **anti-kt LCG jets with R=1.0**
- **Trimmed (R_{cut} = 0.05, R_{sub} = 0.3)**
- **500 ≤ p_T ≤ 800 GeV, |η| < 0.8**

- Dijets (Pythia)
- Dijets (POWHEG+Pythia)

**Active jet area**

**Mass dependence on pileup**
Effect of grooming on data (top)
W mass peak from semileptonic top

This is an old plot, a newer one is not available yet.

The W peak (obtained by BDRS filtering) is used as a standard candle for mass scale and resolution studies.

An analysis on inclusive hadronic W's and Z's (not from top) is also very advanced.
N-subjettiness (Thaler-ValTilburg 1011.2268)

\[ \tau_N = \frac{1}{d_0} \sum_k p_{T,k} \min \{ \Delta R_{k,1}, \Delta R_{k,2}, \ldots, \Delta R_{k,N} \} \]

A measure of how much a jet can be naturally split in N subjets: 0 if \( \leq N \), 1 if \( >N \)

Typically ratios are used to tag 3-body vs 2-body decays

In a similar way, \( t_2/t_1 \) can discriminate W/Z from QCD

Can be calculated theoretically (Fiege et al. 1204.3898)

Can be used as a jet algorithm

Can be combined with grooming techniques
N-subjettiness in data (top)
3-body final states: HepTopTagger

Plehn, Spannowsky, Takeuki, Zerwas 1006.2833
HepTopTagger results in ATLAS

3-body tagging improves S/B for top searches, narrows the top mass peak and allows the search for heavy resonances decaying into boosted top
Maximal information approach to separate various types of events using Feynman rules.

\[
\chi\left(\{p, t\}_N\right) = \frac{P\left(\{p, t\}_N|S\right)}{P\left(\{p, t\}_N|B\right)} = \frac{\sum_{\text{histories}} H_{\text{ISR}} e^{-S_{11}} \cdots \sum_{\text{histories}} H_{\text{ISR}} e^{-S_{s1}} H_{bg} e^{-S_{s2}} \cdots}{\sum_{\text{histories}} H_{\text{ISR}} e^{-S_{11}} \cdots \sum_{\text{histories}} H_{gbb} e^{-S_{b1}} H_{bg} e^{-S_{b2}} \cdots}
\]
Performance of shower deconstruction

In principle, the most powerful method since it uses full matrix element
in practice relies on finding (and calibrating!) micro-jets
An ATLAS implementation currently underway for top and hadronic W's
Quark-gluon separation

Several times it was proposed to use the fact that gluon jets tend to radiate more than quark jets as a tool to discriminate between them.

At LHC, a simple combination of two variables is not much less powerful than the combination of 10 or more (Gallicchio, Schwartz 1106.3076)

The new thing at the LHC is that it is possible to get pure samples of quark or gluon jets from data, overcoming the main problem that these approaches had in the past

Caveat: there is no obvious rigorous definition of quark or gluon jet! For what follows, people just used the first parton of Pythia event record, waiting for a more robust approach.
We already have thousands of pure jets to calibrate/cross-check tool.
Big model-dependence of basic variables, especially for gluon jets; an approach based on data is mandatory.

Properties of quark and gluon jets are extracted from data using QCD and $\gamma$+jet samples, weighting them with the expected quark/gluon fractions in Pythia.

After this procedure, a 10% agreement between n. of tracks and jet width is found on the pure samples (there was not enough statistics yet to use the pure samples as templates rather than as a cross-check).
Perspectives and applications of quark-gluon separation

With more collected data, it will be possible to extract templates directly from pure samples.

With a proper signal definition, we aim at measuring quark-like and gluon-like inclusive jet and dijet cross-section.

Tagger information will be used to improve calibration.

Will be used to discriminate Contact Interactions from harder high-x Pdf's.
Conclusions

QCD is everywhere at the LHC, and should be studied with the best possible tools. We are taking “traditional” measurements to new quality standards.

The field received a big boost a few years ago, when it was realised that jet substructure could be used to highlight hadronic decays of heavy objects.

Since then, several techniques has been developed and proposed, and the field is still booming.

First results on data are coming out now, big hopes for a boosted discovery with the 13 TeV data.