Jet substructure studies at the ATLAS experiment

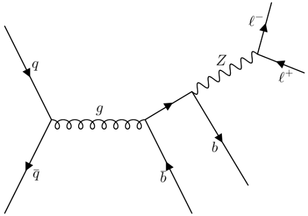
Flavoured Jets at the LHC

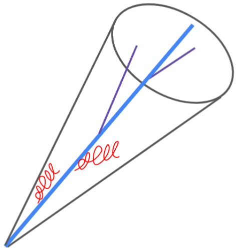
Radosław Grabarczyk, <u>Alberto Rescia</u>, Federico Sforza

11 June 2024

Motivation & Aim

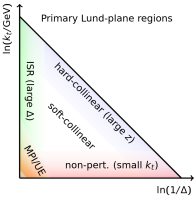
- Aim to measure JSS observables on jets in $Z(\ell^+\ell^-) + b\bar{b}$ events
  - Primary Lund Jet Plane
- Radiation pattern around b-quarks is largely unknown!
- Better understanding of radiation can lead to the development of state-of-the-art taggers
- Measurement provides data for MC tuning of HF effects
- Process important background for many Higgs and BSM analyses
- Flavour-aware algorithms will be used to define truth-level jets for unfolding in this analysis





Observables

Lund Jet Plane

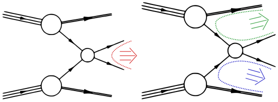


- An angular-ordered representation of the clustering history of a jet
- $(\ln(1/\Delta), \ln(k_t))$ coordinates are plotted in a plane
- Kinematic regions within a jet are factorised

arXiv:1807.04758

D_2

- Ratio of n-point energy correlator functions
- Discriminator of 1-prong vs. 2-prong jets
- Defined only in boosted regime



arXiv:1001.5027

Other observables

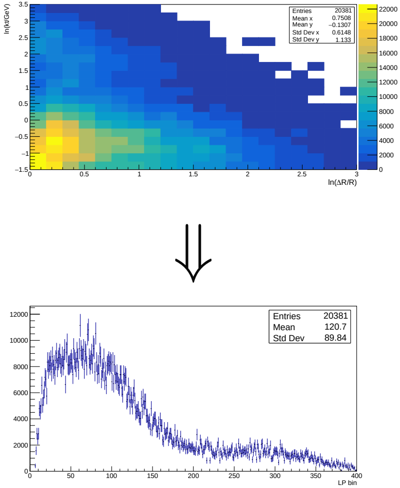
- Jet mass
- Leading jet p_T

A note on the Lund Jet Plane

- To ease the comparison of algorithms, the “unrolled” Lund Jet Plane will be presented
- Starting from bottom row, plot entries in each bin row-by-row, moving from left to right and proceeding from top to bottom

$$LP_{bin} = x_{bin} + (y_{bin} * n_{bins})$$

- Additionally, a cut of $\ln(k_t) > -1.5$ is applied to limit ourselves to the perturbative regime



Details & Selections

- Analysis has 4 signal regions (SRs):
  - Double b-tag and double anti b-tag
  - Boosted and resolved topologies

Analysis selections:

- Require 2 same-flavour opposite sign leptons
- $p_{T,\ell\ell} > 27$ GeV & $m_{\ell\ell} \in [76, 106]$ GeV

Resolved:

- Identify 2 (anti) *b*-tagged R = 0.4 jets with $p_T > 20$ GeV & $|y| < 2.5$
- Measure JSS observables on these jets

Boosted:

- 1 R = 1.0 jet with $p_T > 200$ GeV & $|y| < 1.5$
- Large-R jet double-*b*-tagged when 2 *b*-jets associated to it
- Measure JSS observables on this jet

- Jets selected with high purity *b*-tagging algorithm (70% efficiency) in boosted and resolved analysis regions
- Double anti-*b*-tagged selection requires 0 b-tagged jets

Details & Selections

- Analysis has 4 signal regions (SRs):
  - Double b-tag and double anti b-tag
  - Boosted and resolved topologies

Analysis selections:

For this study:

- Compare JSS observables obtained from jets clustered with different flavour-aware algorithms to those obtained from experimental labels
- **NO ATLAS DATA** — all results at truth (hadron) level!
- Study performed analysing a Rivet routine with same analysis selections (Rivet v3.1.7)
- Rivet routine basically finalised — can debate today!

ets

boosted and resolved analysis regions

- Double anti-b-tagged selection requires 0 b-tagged jets

Generation Details

- Generate 500k events in each SR at LO w/ MG5_aMC v3.5.0
  Resolved + Boosted
  1. $pp \to Z(\ell^+\ell^-) + b\bar{b}$
  2. $pp \to Z(\ell^+\ell^-) + jj$, $j = u,\ d,\ s,\ \bar{u},\ \bar{d},\ \bar{s},\ g$
- In boosted case, apply cut $p_{T,\ell\ell} > 150$ GeV
- Shower with Pythia v8.309 with hadronisation and MPI

Jet reconstruction

1. Cluster all[1] visible, final state particles into jets
   - **Instructions from theorists for flavour-aware algorithms:** cluster jets with undecayed B-hadron
2. Apply cuts and flavour selection on jets
3. Identify charged constituents which satisfy p_T cuts
4. Identify constituents from B-hadron decays and construct B-hadron charged momentum
   - **Boosted case:** account for the fact that descendants of multiple hadrons may end up in jet
   - Separately reconstruct charged momentum of each hadron with available constituents
5. Recluster suitable constituents using *same jet algorithm*
6. Require *at least* as many jets as originally found, forgoing a p_T cut
   - If more jets reclustered than input, consider only *leading jets*
7. Calculate JSS observables on these jets

[1] Following prescription for appropriate ATLAS truth jet collection, in this case AntiKt4TruthWZJets and AntiKt10TruthJets

Experimental considerations

1. *Detector resolution constrains us to charged component of jet!*
   - Inner tracker resolution $\sim$0.01 rad
   - Allows for finer reconstruction of Lund Plane
   - Reconstructed track have $p_T > 500$ MeV
2. *Need to include b-hadron decay products*
   - Currently truth-level jets are defined with the visible final state in experiments
   - Decay products must still be removed *a posteriori* as they would "contaminate" some regions of Lund Plane where effects such as dead cone could be visibile
3. For flavour labelling purposes in ATLAS, *only heavy flavour hadrons with $p_T > 5$ GeV are considered*
4. In ATLAS, a jet is labelled as a *b*-jet if a B hadron is found within $\Delta R(B, jet) < 0.3$

Flavour algorithms

Algorithms considered

1. Anti-k_t jets with CMS-style ghost flavour labelling
2. Anti-k_t jets with ATLAS-style ΔR flavour labelling
3. IFN algorithm
4. CMP algorithm
5. GHS algorithm
6. SDFlav algorithm

- **ATLAS & CMS:** require match to (2) 1 B-had in (boosted) resolved regime
- **Flavour-aware algorithms**
  - *Resolved regime:* require single label per jet, make no distinction between b and $\bar{b}$
  - *Boosted regime:* require large-R jet to have no overall flavour, but check if $b\bar{b}$ cancellation occurs within jet

Algorithm settings

1. IFN
   - $\alpha = 2.0$
   - $\omega = 3.0 - \alpha$
   - FlavSummation = FlavRecombiner::net
2. CMP
   - $a = 0.1$
   - CorrectionType = OverAllCoshyCosPhi_a2
   - ClusteringType = DynamicKtMax
3. GHS
   - $\alpha = 1.0$
   - $\omega = 2.0$
   - p_T cut = 20.0 GeV
4. SDFlav
   - z cut = 0.1
   - $\beta = 1.0$

Results

Legend

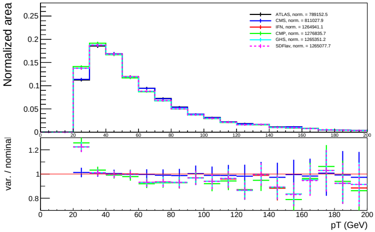
BLACK — ATLAS
BLUE — CMS
RED — IFN
GREEN — CMP
CYAN — GHS
MAGENTA — SDFlav

Disclaimer: all work in progress!

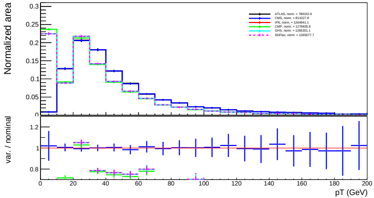
$Z + b\bar{b}$ p_T distribution

- No large differences in shape of p_T distribution of leading jet
- At low p_T some differences arise between ATLAS/CMS algorithms and flavour-aware clustering algorithms
- Large differences seen in charged jet p_T distribution
  - Effect due to different constituents
  - ATLAS/CMS include decay products of B-hadrons
  - Flav. algs. include undecayed hadrons

Leading jet

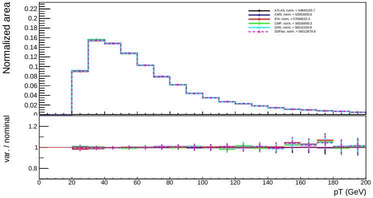


Leading charged jet

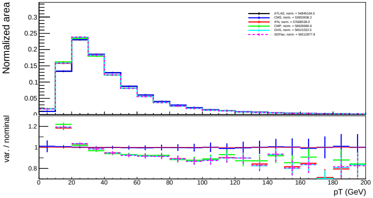


$Z + jj$ p_T distribution

Leading jet



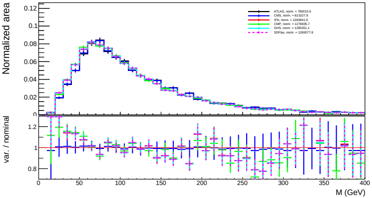
Leading charged jet



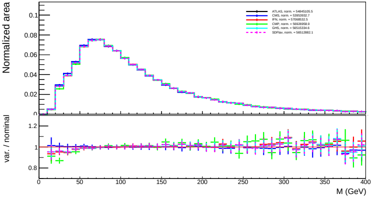
All algorithms show same shape

Dijet mass

$Z + b\bar{b}$



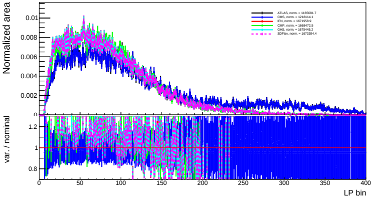
$Z + jj$



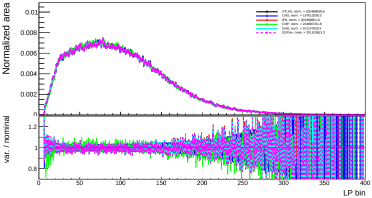
Shapes line up nicely
Some differences observed in *b*-jets due to differeing constituents between labelling schemes

Lund Jet Plane

$Z + b\bar{b}$ leading



$Z + jj$ leading

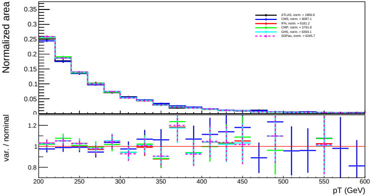


LJP seems to be mostly independent of choice of algorithm
Differences again present for *b*-jets due to different constitutents in experimental/theoretical labelling schemes

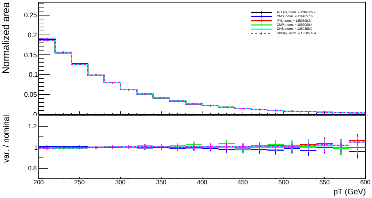
Boosted results

Large-R jet p_T distribution

$Z + b\bar{b}$



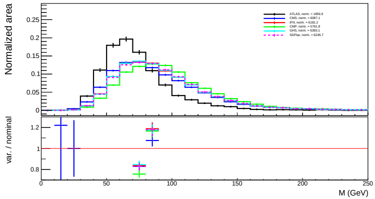
$Z + jj$



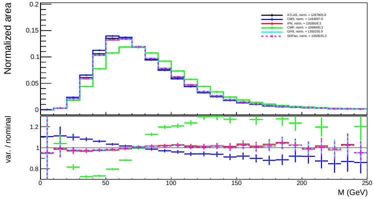
Shapes mostly consistent

Large-R jet mass distribution

$Z + b\bar{b}$



$Z + jj$

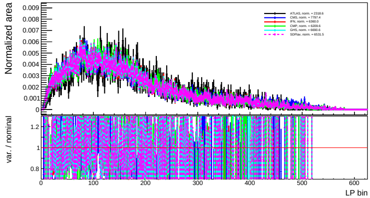


ATLAS deviates significantly due to small cone size used for labelling
Some deviations also shown by CMP for light jets

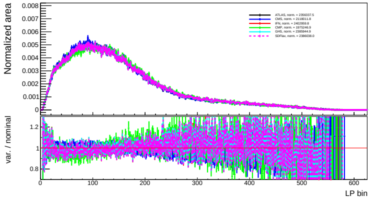
Lund Jet Plane

Jet $p_T \in [300, 400]$ GeV

$Z + b\bar{b}$



$Z + jj$



Again no major shape differences

Conclusions

- Major differences arise due to different definitions of input particles between theory/experiment
  - *Need to harmonise definitions if we wish to apply new algorithms experimentally!*
- CMS and especially ATLAS labelling for large-R jets needs improvement
  - Need to clearly define what it means to double-tag a large-R jet
- All algorithms besides CMP behave similarly for large-R jets

Work in progress! More studies needed