Recent ATLAS inclusive-jet and heavy-flavor-jet measurements

Benjamin Nachman
Lawrence Berkeley National Laboratory

on behalf of the ATLAS experiment

SantaFe 2020
Feb. 3, 2020
A useful way to think about measurements of jet structure:

\[ m_{\text{jet}}^2 \sim \sum z \Delta R^2 \]

*(diagonal line in the plane)*

Splitting functions

\[ g \rightarrow \text{bb + inclusive} \]

(1) \hspace{1cm} (2) \hspace{1cm} (3)

(4) Measure the **full plane** at once!
Groomed jet (mass)

The mass of a jet is mostly determined by the hardest splitting.

Systematically removing jet constituents (soft drop grooming) renders the mass calculable and measurable.

[Marzani, Schunk, Soyez, EPJC 78 (2018) 96]
[Marzani, Schunk, Soyez, JHEP 08 (2017) 132]
[Frye, Larkoski, Schwartz, Yan, JHEP 07 (2016) 064]
[Frye, Larkoski, Schwartz, Yan, 1603.06375]
[Kang, Lee, Liu, Ringer, JHEP 10 (2018) 137]
[Larkoski, Marzani, Soyez, Thaler, JHEP 1405 (2014) 146]
[Dasgupta, Fregoso, Marzani, Salam, JHEP 09 (2013) 029]
A little over a year ago, we published a first measurement of the groomed jet mass.

This measurement used constituent-level uncertainties derived from a study of the single particle response.

In the region of perturbative control, there is excellent agreement with LL and beyond LL calculations.

[Marzani, Schunk, Soyez, EPJC 78 (2018) 96]
[Marzani, Schunk, Soyez, JHEP 08 (2017) 132]
[Frye, Larkoski, Schwartz, Yan, JHEP 07 (2016) 064]
[Frye, Larkoski, Schwartz, Yan, 1603.06375]
We now have a new measurement that extends the original one with the same dataset.

One extension is that we use charged-only in addition to calorimeter-based jet constituents. The charged-only is more precisely measured and the distribution-level differences are small.
Jet Mass

We have also added additional groomed observables.
Jet Mass (+ friends)

1912.09837 New!

**ATLAS**

$\sqrt{s} = 13$ TeV, $32.9$ fb$^{-1}$

anti-$k_t$ $R = 0.8$

Soft Drop, $z_{\text{cut}} = 0.1$, $\beta = 0$

$p_T^{\text{lead}} > 300$ GeV

Data, Track-based

Data, Calorimeter-based

Charged and neutral are consistent - smaller uncertainty for former.
Jet Mass (+ friends)  1912.09837  New!

\[ \frac{1}{\sigma} \frac{d\sigma}{dr_g} \]

**ATLAS**
\[ \sqrt{s} = 13 \text{ TeV}, \text{32.9 fb}^{-1} \]
Calorimeter-based, anti-\(k_t\), \(R = 0.8\)
Soft Drop, \(z_{\text{cut}} = 0.1, \beta = 0\)
\(p_T^{\text{lead}} > 600 \text{ GeV}\)

\[ \begin{array}{c}
\text{Data} \\
\text{NLL} \\
\text{Nonperturbative} \\
\text{Perturbative}
\end{array} \]

Ratio to Data

\[ \begin{align*}
7 \times 10^{-2} & \quad 10^{-1} \\
2 \times 10^{-1} & \quad 3 \times 10^{-1}
\end{align*} \]

[Kang, Lee, Liu, Neill, Ringer, 1908.01783]
Can exploit rapidity dependence of q/g fraction to extract q/g properties!

ATLAS
\[\sqrt{s} = 13\ \text{TeV},\ 32.9\ \text{fb}^{-1}\]
Track-based, anti-\(k_t\) \(R = 0.8\)
Soft Drop, \(z_{\text{cut}} = 0.1, \beta = 1\)

Jet Mass (+ friends)
Gluon splitting to bottom quarks gives us the only ~pure access to QCD splitting functions.

(and of course, this is a very important process for Higgs)

We have used **small-radius b-tagged track jets** to study properties of $g \rightarrow bb$ at small opening angles.
Gluon Fragmentation with $g \rightarrow bb$

**Biggest challenge:**
Removing the non $bb$ background.

**Solution:**
Fit the impact parameter significance using templates from data.
We find that the flavor fractions directly from Pythia are rather off; interestingly, **BL** and **L +C** are \textasciitilde inverted. The fit is done independently for each bin!
We measured $\Delta R$, $z$, $m/p_T$, and this observable, sensitive to the gluon polarization.
We measured $\Delta R$, $z$, $m/p_T$, and this observable, sensitive to the gluon polarization.

While $g \rightarrow bb$ is an excellent probe of fragmentation, we need other probes for generic quarks and gluons.

We have made a measurement of generic high $p_T$ jets using all their charged particles.
Quark and Gluon Fragmentation

As with the soft drop observables measurement, we can exploit rapidity differences to extract quark/gluon distributions.

ATLAS Simulation
\( \sqrt{s} = 13 \text{ TeV} \)
NNPDF2.3LO + Pythia 8.186 A14

Fewer gluons

(more valence quarks)
Quark and Gluon Fragmentation

We simultaneously measure jet $p_T$ and $X$ for both jets.

$X$ is $\zeta$, number of tracks, track radius, or track $p_T$ relative to the jet axis.
Results in bins of $p_T$

$\sqrt{s} = 13$ TeV, 33 fb$^{-1}$

All selected jets, $2000 < \text{Jet } p_T / \text{GeV} < 2500$

**ATLAS**

- Data (stat. uncert.)
- Stat. $\oplus$ syst. uncert.
- Pythia 8.186 A14
- Herwig++ 2.7
- Sherpa 2.1

Pythia-data-Herwig sandwich!
Results in bins of $p_T$

$\sqrt{s} = 13$ TeV, 33 fb$^{-1}$

All selected jets, $900 < \text{Jet } p_T / \text{GeV} < 1000$

**ATLAS**

$1/N_{\text{jet}} (1/N_{\text{ch}}) dN_{\text{ch}} / d\zeta$

- Data (stat. uncert.)
- Stat. $\oplus$ syst. uncert.
- Pythia 8.186 A14
- Herwig++ 2.7
- Sherpa 2.1

$p_T > 500$ MeV
Results in bins of $p_T$

**ATLAS**

$\sqrt{s} = 13$ TeV, 33 fb$^{-1}$

All selected jets, $900 < \text{Jet } p_T / \text{GeV} < 1000$

$p_T > 500$ MeV

(Data (stat. uncert.) + (Stat. + syst. uncert.)

- Pythia 8.186 A14
- Herwig++ 2.7
- Sherpa 2.1

$\frac{1}{N_{\text{ch}}}$ d$N_{\text{ch}}$/d$p_T^{\text{rel}}$
Inclusive $p_T$ dependence

$\sqrt{s} = 13$ TeV, 33 fb$^{-1}$
All selected jets

ATLAS

Data (stat. uncert.)
Stat. $\oplus$ syst. uncert.
Pythia 8.186 A14
Herwig++ 2.7
Sherpa 2.1

Pythia-data-Herwig sandwich!
Inclusive $p_T$ dependence

$\sqrt{s} = 13$ TeV, 33 fb$^{-1}$

All selected jets

$\frac{n_{ch}(\zeta < 1.5^{17} \approx 0.001)}{n_{ch}} \approx -1.5$

Data (stat. uncert.)

Stat. $\oplus$ syst. uncert.

Pythia 8.186 A14

Herwig++ 2.7

Sherpa 2.1

Difference enhanced at low momentum fraction
Quark and gluon results

**ATLAS**
\( \sqrt{s} = 13 \text{ TeV}, 33 \text{ fb}^{-1} \)

Can use q/g fractions for forward/central to extract q/g properties.

Can predict how moments scale with \( p_T \) (DGLAP-like evolution)

Interesting deviations from resummed prediction (and MC) for gluons at high \( p_T \)
Isospin + energy conservation makes this \( \sim \)constant and \( \sim 2/3 \)

\( \kappa = 0 \) is the previous slide (but pQCD breaks down)

Prediction born out in data: increases when moment (\( \kappa \)) is small & decreases when large.
There was a recent proposal for a clever way of extracting quark and gluon distributions without using fractions.

I don’t have time to go into the details here, but we have measured the quark and gluon jet topics for the first time - they may help solve the tension that we see at high $p_T$.

Categorizing all hard splittings at once

\[ j \]

\[ j_1 \]

\[ j_2 \]

\[ z = j_1 \text{ momentum fraction of } j \]

\[ \Delta R = \text{angle between } j_1 \text{ and } j_2 \]

[Dreyer, Salam, Soyez, JHEP 12 (2018) 064]
Categorizing all hard splittings at once

\( z = j_1 \) momentum fraction of \( j \)

\( \Delta R = \) angle between \( j_1 \) and \( j_2 \)

[Dreyer, Salam, Soyez, JHEP 12 (2018) 064]
Categorizing all hard splittings at once

\[ z = \text{j}_1 \text{ momentum fraction of \text{j}} \]

\[ \Delta R = \text{angle between \text{j}_1 \text{ and \text{j}_2}} \]

[Dreyer, Salam, Soyez, JHEP 12 (2018) 064]
Categorizing all hard splittings at once

\[ z = j_1 \text{ momentum fraction of } j \]

\[ \Delta R = \text{angle between } j_1 \text{ and } j_2 \]

[Dreyer, Salam, Soyez, JHEP 12 (2018) 064]
z = \text{j}_1 \text{ momentum fraction of } \text{j} \\
\Delta R = \text{angle between } \text{j}_1 \text{ and } \text{j}_2
Categorizing all hard splittings at once

Regions of the Lund jet plane

\[ P(\triangle) \sim \frac{dz}{z} \frac{d\Delta R}{\Delta R} \]

Factorize physical processes!

[Factorize physical processes!]

[Particle-level Emission]

[Detector-level Emission]

[Dreyer, Salam, Soyez, JHEP 12 (2018) 064]
Categorizing all hard splittings at once

Full Run 2!

First measurement of the Lund jet plane!

...powerful tool for isolating hadronization, parton shower effects, and fixed-order effects

Key experimental challenge: tracking inside dense environments

[Dreyer, Salam, Soyez, JHEP 12 (2018) 064]
First measurement of the Lund jet plane!

... powerful tool for isolating hadronization, parton shower effects, and fixed-order effects

Key experimental challenge:
tracking inside dense environments

[+ V.S. × hadronization
+ V.S. × parton shower

\[ \frac{1}{N_{\text{jets}}} \frac{dN}{d \ln(\frac{R}{\Delta R})} \]

\[ \text{Ratio to Data} \]

ATLAS Preliminary
\( \sqrt{s} = 13 \text{ TeV}, 139 \text{ fb}^{-1} \)

0.97 < ln(1/z) < 1.25

\( \text{resummation} \)

region

NP-region

\[ \ln(\frac{R}{\Delta R}) \]

\[ \text{Ratio to Data} \]
Conclusions and outlook

ATLAS has an active program measuring inclusive jet and heavy flavor jet properties. These measurements provide an important test of QCD and also offer unique inputs to MC tuning.
b-tagging inside jets

**ATLAS**

$\sqrt{s} = 13$ TeV, $L_{\text{int}} = 33$ fb$^{-1}$

$0.25 < \Delta R(b,b) < 0.3$

Total (lead, $j$) $\chi^2 = 13.5 (9.5) / 19$ DoF

### Component (pre-fit %, post-fit %)

- L+C (45%, 34%)
- B (34%, 50%)
- BB (20%, 17%)

**MC Uncertainty**

**Data**

### $R(b,b) < 0.3$

$\Delta 0.25 < \Delta 13.5 (9.5) / 19$ DoF

**Total (lead, $j$)**

$\chi^2 = 13.5 (9.5) / 19$ DoF

**ATLAS**

$\sqrt{s} = 13$ TeV, $L_{\text{int}} = 33$ fb$^{-1}$

<table>
<thead>
<tr>
<th>Flavor Fraction</th>
<th>Data (post-fit)</th>
<th>MC (pre-fit)</th>
</tr>
</thead>
<tbody>
<tr>
<td>BB</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>B</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>L+C</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

### Data/MC

- $\pi$/ppg, gbb
- $\theta$

### $s_{\text{sub}}^{j_1}$

- $s_{\text{d}_0}^{j_1}$

### $\Delta \theta_{\text{ppg,gbb}} / \pi$

- $\pi$
We match calorimeter-cell clusters to tracks extrapolated from our inner detector.
Per particle calibrations - cross-checks

Many cross-checks to ensure that the “bottom up” approach is valid (not spoiled by collective effects)

Energy scale by comparing to Z +jet balance

Mass resolution from W peak

Mass scale by comparing with track jets
Tracking inside jets

**NN1: classification**
- $n$ particle clusters

**NN2: regression**
- $n$ particle locations

**NN3: regression**
- $n$ particle residuals

“Number NN”

“Position NN”

“Error NN”
Confusion matrix:

<table>
<thead>
<tr>
<th></th>
<th>NN 1</th>
<th>NN 2</th>
<th>NN 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>true 1</td>
<td>96%</td>
<td>3%</td>
<td>1%</td>
</tr>
<tr>
<td>true 2</td>
<td>33%</td>
<td>33%</td>
<td>33%</td>
</tr>
<tr>
<td>true 3</td>
<td>5%</td>
<td>10%</td>
<td>85%</td>
</tr>
</tbody>
</table>
Step 2: Position

ATLAS Simulation Preliminary

PYTHIA8 dijet, $1.8 < p_T^{\text{jet}} < 2.5$ TeV
2 particles clusters
local Y direction

<table>
<thead>
<tr>
<th>Particle density / 0.016 mm</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.14</td>
</tr>
<tr>
<td>0.12</td>
</tr>
<tr>
<td>0.1</td>
</tr>
<tr>
<td>0.08</td>
</tr>
<tr>
<td>0.06</td>
</tr>
<tr>
<td>0.04</td>
</tr>
<tr>
<td>0.02</td>
</tr>
<tr>
<td>0.0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Truth hit residuals [mm]</th>
</tr>
</thead>
<tbody>
<tr>
<td>-0.4</td>
</tr>
<tr>
<td>-0.3</td>
</tr>
<tr>
<td>-0.2</td>
</tr>
<tr>
<td>-0.1</td>
</tr>
<tr>
<td>0</td>
</tr>
<tr>
<td>0.1</td>
</tr>
<tr>
<td>0.2</td>
</tr>
<tr>
<td>0.3</td>
</tr>
<tr>
<td>0.4</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Pitch and ToT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pitch = 0.4 mm, 8 bit ToT</td>
</tr>
<tr>
<td>Pitch = 0.25 mm, 4 bit ToT</td>
</tr>
</tbody>
</table>
Step 3: Uncertainty

ATLAS Simulation Preliminary

PYTHIA8 dijet, $1.8 < p_T^{\text{jet}} < 2.5$ TeV
local X direction

1 particle clusters
$\mu = 0.01, \sigma = 0.78$

2 particles clusters
$\mu = 0.00, \sigma = 0.76$

3 particles clusters
$\mu = 0.07, \sigma = 0.82$

residual / uncertainty estimate = Pull
hits in the pixel and strip detectors

build 3-point seeds

filter the seeds

combinatorial Kalman filter

track candidates

ambiguity solving

high resolution track fit
(use NN for cluster positions)

order tracks by score

assign a score per candidate

clusters, holes, $\chi^2$, $\log(p)$

reject track

$>2$ shared clusters

recover track candidate
(NN to identify merged clusters)

split clusters no longer counted as shared

clusters can be shared by $\leq 2$ tracks;
tracks can have $\leq 2$ shared clusters

tracks!
Tracking inside jets

ATLAS
\(\sqrt{s}=13\text{ TeV}, 3.2\text{ fb}^{-1}\)

Data 2015  
Statistical Uncertainty  
Total Uncertainty  
Simulation (Pythia 8)  
Total Uncertainty


ATLAS
\(\sqrt{s} = 13\text{ TeV}, 3.2\text{ fb}^{-1}\)

200 GeV < \(p_T^{\text{jet}}\) < 400 GeV

- Single-Track Template
- Multiple-Track Template

\(dE/dx\) [MeV g\(^{-1}\) cm\(^2\)]

Tracks
CTIDE Uncertainties: Efficiency

Alternative methods: track-to-calorimeter ratio (left), angular asymmetry (right)

Consistent picture with the dE/dx method.
CTIDE Uncertainties: Fakes

What do fakes inside high $p_T$ jets look like?

Fake rate increases with density

μ-dependence decreases with area

N.B. fakes are defined in the usual way: MCProb < 0.5 (basically weighted energy matching to truth)
Method to constrain: look in a “control region” enriched in fakes
CTIDE Uncertainties: Fakes

Various methods are consistent within ~30%

(really does look like they are well-described in shape but not in rate)
Response matrices

ATLAS Simulation
\( \sqrt{s} = 13 \text{ TeV}, 32.9 \text{ fb}^{-1} \)
Calorimeter-based, anti-\( k_T \), \( R = 0.8 \)
Soft Drop, \( z_{\text{cut}} = 0.1, \beta = 0 \)
Pythia 8.186
Response matrices

**ATLAS** Simulation

$\sqrt{s} = 13$ TeV, 32.9 fb$^{-1}$
Track-based, anti-$k$, $R = 0.8$
Soft Drop, $z_{\text{cut}} = 0.1$, $\beta = 0$
Pythia 8.186

![Heatmap of response matrices](image)
**ATLAS**

\( \sqrt{s} = 13 \text{ TeV}, \ 32.9 \text{ fb}^{-1} \)

- Track-based
- \( \text{anti-}k_t, \ R = 0.8 \)
- Soft Drop, \( z_{\text{cut}} = 0.1, \ \beta = 0 \)
- \( p_T^{\text{lead}} > 300 \text{ GeV} \)

**Relative Uncertainty**

- **Total Uncertainty**
- Data statistical error
- Unfolding Nonclosure
- Fragmentation Modeling
- Efficiency within jets
- Fake rate
- Cluster energy scale
- Other

**r_g**

- \( 7 \times 10^{-2} \)
- \( 10^{-1} \)
- \( 2 \times 10^{-1} \)
- \( 3 \times 10^{-1} \)
Uncertainties

\[ \sqrt{s} = 13 \text{ TeV}, \ 32.9 \text{ fb}^{-1} \]

Calorimeter-based anti-\( k_t \) \( R = 0.8 \)

Soft Drop, \( z_{\text{cut}} = 0.1, \ \beta = 0 \)

\( p_T^{\text{lead}} > 300 \text{ GeV} \)

**ATLAS**

- Total Uncertainty
- Data statistical error
- Unfolding Nonclosure
- Fragmentation Modeling
- Cluster energy scale
- Cluster energy resolution
- Pileup modeling
- Other
Uncertainties

$\sqrt{s} = 13$ TeV, 32.9 fb$^{-1}$

- Track-based $R = 0.8$
- Soft Drop, $z_{\text{cut}} = 0.1$, $\beta = 0$
- $p_T^{\text{lead}} > 300$ GeV

Total Uncertainty
- Data statistical error
- Unfolding Nonclosure
- Fragmentation Modeling
- Efficiency within jets
- Fake rate
- Cluster energy scale
- Other
Uncertainties

$\sqrt{s} = 13$ TeV, 32.9 fb$^{-1}$

Calorimeter-based
anti-$k_t$ $R = 0.8$

Soft Drop, $z_{\text{cut}} = 0.1$, $\beta = 0$
$p_{T}^{\text{lead}} > 300$ GeV

Total Uncertainty
Data statistical error
Unfolding Nonclosure
Fragmentation Modeling
Cluster energy scale
Cluster energy resolution
Pileup modeling
Other
Uncertainties

$\sqrt{s} = 13$ TeV, 32.9 fb$^{-1}$

- Track-based anti-$k_t$ $R = 0.8$
- Soft Drop, $z_{\text{cut}} = 0.1$, $\beta = 0$
- $p_T^{\text{lead}} > 300$ GeV

- Total Uncertainty
- Data statistical error
- Unfolding Nonclosure
- Fragmentation Modeling
- Efficiency within jets
- Fake rate
- Cluster energy scale
- Other
Uncertainties

\[ \rho \]

\[ 0.05 \]

\[ 0.1 \]

\[ 0.15 \]

\[ 0.2 \]

\[ 0.25 \]

\[ 0.3 \]

\[ 0.35 \]

\[ 0.4 \]

\[ 0.45 \]

\[ 0.5 \]

\[ -4.5 \]

\[ -4 \]

\[ -3.5 \]

\[ -3 \]

\[ -2.5 \]

\[ -2 \]

\[ -1.5 \]

\[ -1 \]

\[ -0.5 \]

\[ ATLAS \]

\[ \sqrt{s} = 13 \text{ TeV}, 32.9 \text{ fb}^{-1} \]

Calorimeter-based

anti-\( k_t \) \( R = 0.8 \)

Soft Drop, \( z_{\text{cut}} = 0.1, \beta = 0 \)

\( p_T^{\text{lead}} > 300 \text{ GeV} \)

Total Uncertainty

Data statistical error

Unfolding Nonclosure

Fragmentation Modeling

Cluster energy scale

Cluster energy resolution

Pileup modeling

Other