Inputs and Procedures of Jet Reconstruction in ATLAS

Chaowaroj (Max) Wanotayaroj on behalf of the ATLAS Collaboration

22 July 2019 BOOSTON
• Jets are important for almost all ATLAS analyses
• Most use anti-\(k_T\) jets with \(R=0.4\) (small-\(R\)) or \(R=1.0\) (large-\(R\))
  – Some use both
  – Also looking at other sizes in some cases
• How we build and calibrate jets has wide-reaching impact
  ➢ What is(are) the best choice(s)?
• Jet energy scale uncertainty and jet energy resolution dominate many searches and measurements
  ➢ Work to reduce these uncertainties as much as possible
1. **Inputs**
2. **Jet reconstruction**
3. **Pile-up corrections**
4. **MC-based calibration**
   - Match truth to reco jets, then calculate reco/truth “response”
   - Correct the jet 4-momentum to truth level
5. **In-situ calibration for data**
Jet Reco and Calibration Chain

1. Inputs
2. Jet reconstruction
3. Pile-up corrections
4. MC-based calibration
   - Match truth to reco jets, then calculate reco/truth “response”
   - Correct the jet 4-momentum to particle level
4.1 (small-R) Global sequential calibration
   - Reduce flavor dependence and correct various detector effects
5. In-situ calibration for data
1. Small-R, update on jet energy scale/resolution (JES/JER)
2. Large-R, in-situ calibration with 80 fb^{-1}
3. Large-R, alternative inputs and grooming strategies
Small-R JES/JER

Taking advantage of particle flow jets
Jet Inputs

• **Topocluster: calorimeters only**
  - 4σ-above-noise seed cells, iteratively add all 2σ neighbors and cells surrounding them
  - These are called EM-scale topoclusters
  - Topocluster are corrected to point at the primary vertex (origin correction)

- **Particle flow (PFlow):** Subtract matched tracks’ momentum from topoclusters
  - Remaining clusters and tracks form PFlow objects
  - Only PFlow objects matched to primary vertex tracks are used for jet building
  - Great performance for low $p_T$
• ATLAS jet usage is moving toward PFlow
• It is now the primary option in ATLAS
  – Better response
  – Better resolution at low $p_T$, and comparable to EM-scale at high $p_T$
• EM-scale topocluster jets still used by some analyses
PFlow jets have comparable JES uncertainty, and much better JER uncertainty at low \( p_T \) compared to topocluster jets.
Large-R, In-situ Calibration

Updated with 80 fb$^{-1}$
In-situ Calibration—JES

- Derived from a jet recoiling against a well-measured object:
  1. Z boson or $\gamma$ as reference objects
  2. Several lower-$p_T$ small-R jets for high-$p_T$ jets

- Combining three techniques to cover the full $p_T$ range
  - $Z+\text{jet}$ method runs out of statistics $\sim 450$ GeV
  - $\gamma+\text{jet}$ is used until $\sim 1$ TeV
  - Use multijet method above that threshold

- Combine methods in overlapping regions
  - Overall uncertainty can be reduced as overlapping regions agree and each method's uncertainties are mostly independent
Large-R Jet In-situ Calibration

**JES Uncertainty**

- Large reduction compared to without in-situ calibration
  - Total uncertainty depends on the assumption of the topology (W, Z, or top) and flavor (quark or gluon initiated) composition of the jets

![Graph showing fractional JES uncertainty vs. jet transverse momentum](image-url)
In-situ Calibration—JMS/JMR

- For large-R jets, we also need in-situ calibration for the mass
- Two methods for in-situ mass calibration:
  1. Forward-folding: Use high purity top sample, shift and stretch jet mass resolution function so that the simulation matches the data
  2. $R_{\text{track}}^m = \frac{m_{\text{track}}}{m_{\text{calo}}}$: Tracker provides an independent (charged only) measurement of a jet, so any deviation of the double ratio, $\frac{R_{\text{MC}}}{R_{\text{Data}}}$, from 1 provides an estimate of the scale uncertainty

- Forward-folding has smaller uncertainty, but $R_{\text{track}}$ covers a much broader $p_T$ and mass range

---

**Graph Description:**
- Axes:
  - X-axis: $\text{Large-R jet } p_T$ [GeV] with range from $3 \times 10^2$ to $2 \times 10^3$
  - Y-axis: Mass response ratio, $\frac{R_{\text{Data}}}{R_{\text{MC}}}$
- Data points and trend line:
  - Yellow shaded area: Total uncertainty
  - Blue line: Statistical component
  - Red triangle: Forward Folding
- Additional Information:
  - $\sqrt{s} = 13 \text{ TeV, 36.2 fb}^{-1}$
  - Trimmed $R = 1.0$ anti-$k_t$ (LCW+JES+JMS)
  - $120 < m_{\text{jet}} < 300 \text{ GeV}$
Large-R, Alternative Inputs and Jet Grooming

Inputs, constituent-level pile-up suppressions, and grooming scan
Large-R Jet Inputs

- **Topocluster:** calorimeters only
  - $4\sigma$-above-noise seed cells, add all $2\sigma$ neighbors and cells surrounding them
  - Calibrate to account for EM and HAD differences, dead material and out-of-cluster deposits to get Local Cell Weighting (LCW) topoclusters
  - Topocluster are corrected to point at the primary vertex (origin correction)

- **Particle flow (PFlow):** Subtract matched tracks’ momentum from topoclusters
  - Remaining clusters and hard-scatter tracks form PFlow objects
  - Remove PFlow object matched to non-primary vertex tracks
  - Great performance for low $p_T$

- **Track-CaloCluster (TCC):** Use energy from topoclusters and angle from tracks
  - For multiple-to-multiple matching, energy is shared among tracks to create multiple TCC objects
  - Remove TCC objects with non-primary vertex tracks
  - Great performance for high $p_T$
Pile-up Correction

• From inputs:
  – Topocluster: noise suppression
  – PFlow/TCC: charged objects not associated with primary vertex are rejected

• Constituents-level (topocluster) correction:
  – Voronoi Subtraction (VS): correct constituent's energy by $\rho \cdot A_{\text{Voronoi}}$
    • $\rho$: transverse momentum density
  – Constituent Subtraction (CS): Add ghosts with $p_T^g = \rho \cdot A_g$ then:
    \[
    \begin{align*}
    \text{If } p_{T,i} & \geq p_{T,k}^g: & p_{T,i} & \rightarrow p_{T,i} - p_{T,k}^g, \\
    & & p_{T,k}^g & \rightarrow 0 \text{ GeV}; \\
    \text{otherwise:} & & p_{T,k}^g & \rightarrow p_{T,k}^g - p_{T,i}, \\
    & & p_{T,i} & \rightarrow 0 \text{ GeV},
    \end{align*}
    \]
  – SoftKiller (SK): $p_T$ cut so half of $\eta - \phi$ grid spaces are empty
  – PUPPI: $p_T$ cut based on information from nearby constituents

• Jet-level: Grooming

<table>
<thead>
<tr>
<th>Grooming Algorithm</th>
<th>Name</th>
<th>Parameters Tested</th>
</tr>
</thead>
<tbody>
<tr>
<td>Soft Drop</td>
<td>SD</td>
<td>$(\zeta_{\text{cut}}, \beta) \in [0.1] \times [0, 0.5, 1]$</td>
</tr>
<tr>
<td>Bottom-up Soft Drop</td>
<td>BUSD</td>
<td>$(\zeta_{\text{cut}}, \beta) \in [0.05, 0.1] \times [0, 0.5, 1]$</td>
</tr>
<tr>
<td>Recursive Soft Drop</td>
<td>RSD</td>
<td>$(\zeta_{\text{cut}}, \beta, \eta) \in [0.05, 0.1] \times [0, 0.5, 1] \times {2, 3, 5, \text{cut}}$</td>
</tr>
<tr>
<td>Pruning</td>
<td>Pruned</td>
<td>$(\zeta_{\text{cut}}, R_{\text{cut}}) \in [0.15] \times [0.25]$</td>
</tr>
<tr>
<td></td>
<td>Trimmed</td>
<td>$(f_{\text{cut}}, R_{\text{sub}}) \in [5, 9]% \times [0.1, 0.2]$</td>
</tr>
</tbody>
</table>

New! Based on previous study (ATL-PHYS-PUB-2017-020)
Large-R Jet Performance Study

• Three inputs (Topocluster, PFlow, TCC)
• Many pile-up suppression techniques
  – Some of them can also be combined
  – Based on previous study (ATL-PHYS-PUB-2017-020), we choose:
    • Constituent Subtraction + SoftKiller (CS+SK)
    • Voronoi Subtraction + SoftKiller (VS+SK)
    • PUPPI for PFlow only

Two methods:
1. Compare the impact of different pile-up mitigation techniques on individual clusters with and without pile-up included in simulation (“DigiTruth”)
2. Scan over choices of input, constituent-level pile-up suppression, and grooming, and compare them with ATLAS’ standard trimmed jet
   – Compare the performance. Specifically:
     • Pile-up stability
     • Topology dependence
     • Tagging performance
• Showing $E^{\text{Digitruth}}$ (no pile-up) over $E^{\text{reco}}$ (with pile-up)
  – $E^{\text{Digitruth}}$ does include underlying event → 1 is not necessary the ideal value

• Majority of clusters tend to be dominated by hard-scatter (HS) or pile-up (PU)
  – So we will call $\frac{E^{\text{Digitruth}}}{E^{\text{reco}}}>0.5$ a HS cluster and <0.5 a PU cluster

• CS+SK is removing pile-up; HS clusters are more pronounced
DigiTruth Study

All algorithms remove more pile-up than hard-scatter clusters

- Standard ATLAS trimming is doing well

Hard-scatter cluster efficiency (1 is better)  Pile-up cluster efficiency (0 is better)
Quantifying the effect of pile-up on the W mass:

1. Take the mass distribution of W-jet sample in a $N_{pv}$ bin
2. Fit a Gaussian on the W mass peak
3. Plot either the central value or the width as a function of $N_{pv}$
4. Fit a line and measure the slope
### W Mass Peak Values Slope

#### ATL-PHYS-PUB-2019-027

**Jet Grooming Method**

- **Soft Drop**
  - $z_{cut} = 0.1, \beta = 0.0$
  - $z_{cut} = 0.1, \beta = 1.0$
  - $z_{cut} = 0.05, \beta = 0.0, N = \infty$
  - $z_{cut} = 0.05, \beta = 0.0, N = 3$
  - $z_{cut} = 0.1, \beta = 0.0, N = \infty$
  - $z_{cut} = 0.1, \beta = 0.0, N = 3$
  - $z_{cut} = 0.1, \beta = 0.0, N = 5$
  - $z_{cut} = 0.1, \beta = 0.0, N = 10$

- **Recursive Soft Drop**
  - $z_{cut} = 0.1, \beta = 0.0, N = 3$
  - $z_{cut} = 0.05, \beta = 0.0, N = 3$
  - $z_{cut} = 0.05, \beta = 1.0, N = 3$
  - $z_{cut} = 0.1, \beta = 1.0, N = 3$
  - $z_{cut} = 0.05, \beta = 0.0, N = 5$
  - $z_{cut} = 0.1, \beta = 0.0, N = 5$
  - $z_{cut} = 0.5, \beta = 0.0, N = 10$
  - $z_{cut} = 0.1, \beta = 0.0, N = 10$

- **Bottom-up**
  - $z_{cut} = 0.1, \beta = 1.0$
  - $z_{cut} = 0.1, \beta = 0.0$

- **Pruning**
  - $R_{clus} = 0.15, \beta = 0.2$
  - $R_{clus} = 0.1$

- **Trimming**
  - $f_{clus} = 5\%, R_{clus} = 0.1$
  - $f_{clus} = 9\%, R_{clus} = 0.1$
  - $f_{clus} = 5\%, R_{clus} = 0.2$
  - $f_{clus} = 9\%, R_{clus} = 0.2$

**Simulation Preliminary**

- $\sqrt{s} = 13\text{ TeV}, W$ jets

**Anti-$k_t$, $R=1.0$ jets, no JES or JMS calibration applied**

- $300\text{ GeV} \leq p_T^{\text{true}} < 500\text{ GeV}$,
  - $|\eta^{\text{true}}| < 1.2$

#### Table:

<table>
<thead>
<tr>
<th>Row Description</th>
<th>LCTopo Slope</th>
<th>PFlow Slope</th>
<th>TCC Slope</th>
</tr>
</thead>
<tbody>
<tr>
<td>Soft Drop</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$z_{cut} = 0.1, \beta = 0.0$</td>
<td>0.32</td>
<td>0.02</td>
<td>0.04</td>
</tr>
<tr>
<td>$z_{cut} = 0.1, \beta = 1.0$</td>
<td>0.68</td>
<td>0.03</td>
<td>0.05</td>
</tr>
<tr>
<td>$z_{cut} = 0.05, \beta = 0.0, N = \infty$</td>
<td>2.91</td>
<td>0.03</td>
<td>0.06</td>
</tr>
<tr>
<td>$z_{cut} = 0.05, \beta = 0.0, N = 3$</td>
<td>0.14</td>
<td>0.02</td>
<td>0.06</td>
</tr>
<tr>
<td>$z_{cut} = 0.1, \beta = 0.0, N = \infty$</td>
<td>3.35</td>
<td>0.09</td>
<td>0.15</td>
</tr>
<tr>
<td>$z_{cut} = 0.05, \beta = 0.0, N = 3$</td>
<td>0.87</td>
<td>0.03</td>
<td>0.06</td>
</tr>
<tr>
<td>Recursive Soft Drop</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$z_{cut} = 0.1, \beta = 0.0, N = 3$</td>
<td>0.15</td>
<td>0.01</td>
<td>0.03</td>
</tr>
<tr>
<td>$z_{cut} = 0.05, \beta = 0.0, N = 3$</td>
<td>3.61</td>
<td>0.09</td>
<td>0.14</td>
</tr>
<tr>
<td>$z_{cut} = 0.05, \beta = 1.0, N = 3$</td>
<td>0.91</td>
<td>0.03</td>
<td>0.06</td>
</tr>
<tr>
<td>$z_{cut} = 0.1, \beta = 1.0, N = 3$</td>
<td>3.14</td>
<td>0.03</td>
<td>0.06</td>
</tr>
<tr>
<td>$z_{cut} = 0.05, \beta = 0.0, N = 5$</td>
<td>0.14</td>
<td>0.01</td>
<td>0.02</td>
</tr>
<tr>
<td>$z_{cut} = 0.1, \beta = 0.0, N = 5$</td>
<td>3.57</td>
<td>0.09</td>
<td>0.14</td>
</tr>
<tr>
<td>$z_{cut} = 0.5, \beta = 0.0, N = 10$</td>
<td>0.39</td>
<td>0.03</td>
<td>0.06</td>
</tr>
</tbody>
</table>

**Notes:**

- **Row:** grooming algorithm (e.g. **standard ATLAS trimming**)
- **Column:** jet constituent type
- **Number:** Slope (also color coded)
- **Prefer zero** (pile-up stable)
PFlow is more stable than topocluster, even unmodified.

Also reduce the width.
W Mass Peak Values Slope

But still benefit from constituent pile-up suppression
W Mass Peak Values Slope

- Peak values increase with pile-up in PFlow
- The opposite is true for TCC
  - TCC over-subtract the pile-up
    - TCC use all tracks for cluster splitting, so more clusters are removed by matching with pile-up tracks
Since TCC is already over-subtracting, adding constituent pile-up suppressions make it worse.
Quantifying the mass scale calibration’s dependence on jet topology

Take the ratio of the average mass response \( \frac{m_{\text{reco}}}{m_{\text{truth}}} \) of W-jets over QCD jets
Alternative Inputs and Jet Grooming

Topology Dependence

ATL-PHYS-PUB-2019-027

- Prefer one (no topology dependence)
- Constituent pile-up suppression make it worse, but there are good option available, especially for PFlow
Tagging Performance

• Perform a simple two-variable tagger:
  – 68% signal mass window cut
  – One-side cut on $D_2$ (W) or $\tau_{32}$ (top)

• Compare background rejection (1/background efficiency) vs. signal efficiency

Look forward to Steven’s talk on tagging this afternoon!
Summary

• ATLAS is moving towards particle flow for small-R jets
  – Better pile-up suppression, better resolution

• In-situ JES calibration for large-R jets is done with 80 fb\(^{-1}\) of data
  – Reduce the JES uncertainty significantly (from 8% to 1% !)

• Study the impact on the large-R jet performance with various choices of inputs, constituent pile-up suppression, and grooming algorithms
  – PFlow jets outperform LCW topocluster jets across the board
  – TCC can be better than PFlow for high-\(p_T\) jets, but poor performance at low-\(p_T\)
  – Both PFlow and topocluster benefit from pile-up suppression at constituent level
    • Among the choices in this study, CS+SK did best
  – Standard ATLAS trimming does well, but some SoftDrop configurations show possible improvement
Backup
ATLAS Detector

- Inner tracker
  - Inside solenoid magnet
- EM+Hadronic calorimeters
- Muon spectrometer
  - With toroid magnet

Introduction
In-situ Calibration — Small-R

ATLAS Preliminary
\( \sqrt{s} = 13 \text{ TeV}, 80 \text{ fb}^{-1} \)
Anti-\(k_T, R = 0.4 \) (PFlow+JES)

- \( \gamma + \text{jet} \)
- \( Z + \text{jet}, Z \rightarrow ee \)
- \( Z + \text{jet}, Z \rightarrow \mu\mu \)
- Multijet

Relative weight in combination vs. \( p_T^{\text{jet}} \) [GeV]

\( R_{\text{data}} / R_{\text{MC}} \) vs. \( p_T^{\text{jet}} \) [GeV]

- Total uncertainty
- \( Z \rightarrow ee + \text{jet} \)
- \( Z \rightarrow \mu\mu + \text{jet} \)
- Multijet

ATLAS Preliminary
\( \sqrt{s} = 13 \text{ TeV}, 80 \text{ fb}^{-1} \)
Anti-\(k_T, R = 0.4 \) (PFlow+JES)
In-situ Calibration

- PFlow jet $p_T$ response derived from $Z \rightarrow ee$, $Z \rightarrow \mu\mu$, $\gamma$, and multijet using MPF technique

• Total uncertainty depends on the assumption of the topology (W, Z, or top) and flavor (quark or gluon initiated) composition of the jets
• Large reduction compared to without in-situ calibration
In-situ Calibration—JMS/JMR

• For large-R jets, we also need in-situ calibration for the mass

• Two methods for in-situ mass calibration:
  1. Forward-folding: Use high purity top sample, shift and stretch jet mass resolution function so that the simulation matches the data
  2. $R^{\text{track}}_{\text{MC}} = \frac{m_{\text{track}}}{m_{\text{calo}}}$: Tracker provides an independent (charged only) measurement of a jet, so any deviation of the double ratio of $R_{\text{track,MC}}/R_{\text{track,Data}}$ between data and MC from 1 provides an estimate of the scale uncertainty

• Forward-folding has smaller uncertainty, but $R_{\text{track}}$ covers a much broader $p_T$ and mass range
Quantifying the effect of pile-up on tagging observable ($D_2$):

1. Take the $D_2$ distribution of W-jet sample in the low pile-up ($N_{PV} < 15$) bin
2. Find a cut with 50% efficiency
3. Apply the cut in bins of $N_{PV}$ and plot the efficiencies
4. Fit a line and measure the slope
Shown only W-jet with $D_2$, but result are consistent with top and $\tau_{32}$

Prefer zero (pile-up stable)

Generally, efficiency decreases with pile-up, except TCC
Pile-up Stability—Jet Topology

- Prefer one
- Constituent pile-up suppression make it worse, but there are good option available, especially for PFlow