Boosted object identification and their usage in searches for new physics in ATLAS

Yi-Lin Yang

on behalf of the ATLAS Collaboration

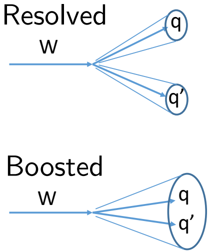
31.07.2020

Introduction

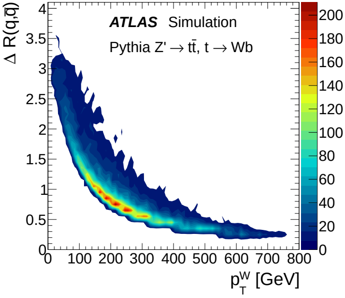
- Top quarks, W and Z bosons can decay hadronically. The angular separation between the decay products can be approximated by:

$$\Delta R \sim \frac{2m}{p_{\mathrm{T}}}$$

- At high pT, the daughter particles from the boosted parent start overlapping.
  - Large-R jets ($R = 1.0$) are used instead.
- Taggers using jet substructure information are used to separate signal jets from q/g initiated jets.
  - The performance of the taggers are calibrated with Run2 data.
- Developed taggers are applied to searchs for new physics.



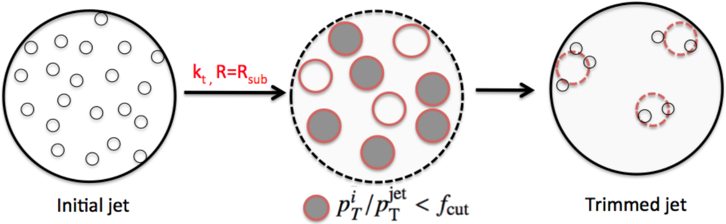
JHEP09 (2013) 076



Large-R jet reconstruction

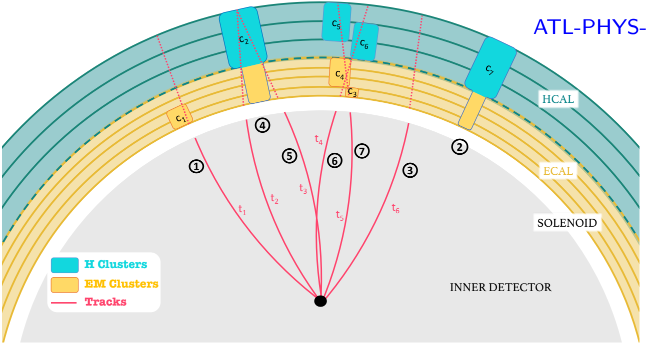
- Input types
  - Topo-cluster
  - Track-calo cluster
  - Particle flow objects (used for small-R jet)
  - Unified flow objects New!
- Topo-clusters (LCTopo)
  - clustered from calorimeter cell signals with local hadronic-cell weighting (LC)
- R = 1.0 jets are clustered by using anti-kt algorithm
- Trimming : remove the pileup and soft radiation (see Eva's talk for more details)
  - $f_{cut} = 0.05$, $R_{sub} = 0.2$

JHEP09 (2013) 076



Large-R jet reconstruction : Track-calo cluster (TCC) jets

- TCC objects
  - combinations of tracks and topo-clusters with ΔR matching
- Jet reconstruction is the same as topo-cluster jets but topo-clusters are replaced by TCC objects.
- Tracking provides better spatial resolution
  - Better mass resolution for $p_{\mathrm{T}} > 2$ TeV and D_2 resolution across full p_{T} range. (D_2 will be introduced later)



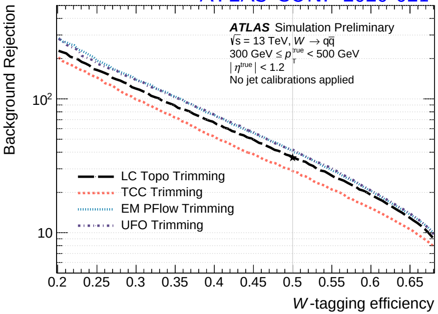
ATL-PHYS-PUB-2017-015

Large-R jet reconstruction : Unified Flow Object (UFO) jets

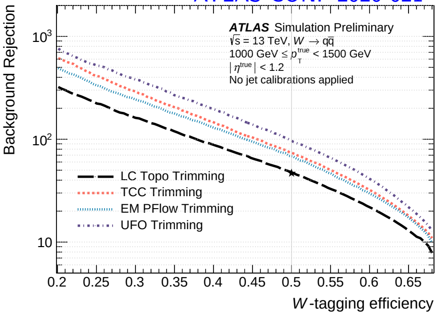
New jet definition!!

- Among LCTopo, PFO(Particle-flow object) and TCC, there is no jet reconstruction which has optimal performance across entire p_{T} range.
- UFOs : the combination of PFOs and TCCs
  - Selected tracks are used to subtract the energy of the matched topo-clusters. The neutral objects (topo-clusters with no matched tracks) are further split by the matched tracks which are not used in the charged objects.

ATLAS-CONF-2020-021



ATLAS-CONF-2020-021

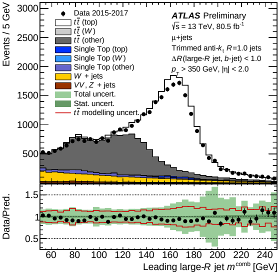


Jet substructure

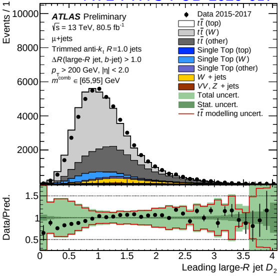
- The features inside a the jet depend on the originating particle.
- Jet Mass : combined mass
  - $m^{comb} = a \times m^{calo} + b \times m^{TA}, \quad m^{TA} = \frac{p_{\mathrm{T}}^{calo}}{p_{\mathrm{T}}^{track}} m^{track}$

    a and b are the weights based on the resolution
- Jet substructure moments
  - N-subjettiness and its ratios (τ_{32}, τ_{21}), splitting scales, energy correlation function and its ratios (C_2, D_2, e_3) and others
  - τ_{32} is used for identifying 3-prong (used for top quarks), D_2 is used for the separation 1-prong and 2-prong (used for W bosons)

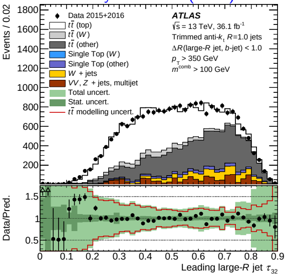
ATL-PHYS-PUB-2020-017



ATL-PHYS-PUB-2020-017



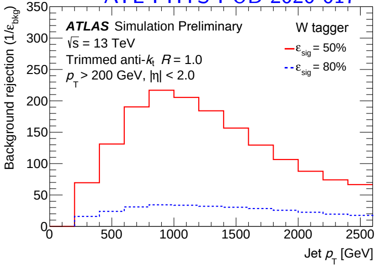
Eur. Phys. J. C 79 (2019) 375



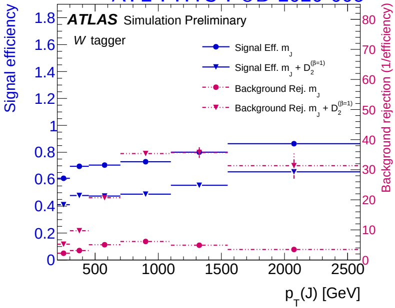
Boosted Jet Taggers: W/Z taggers

- Taggers are defined by imposing cuts on various substructure variables
- LCTopo jets
  - Variables :
    - ★ n^{trk}, m^{comb}, D_2
  - flat 50% and 80% efficiency
- TCC jets
  - Variables :
    - ★ m, D_2
  - significance based efficiency
    - ★ mainly used for the heavy diboson resonance search in the semileptonic final states to enhance the signal efficiencies due to low backgrounds. Refer to arXiv:2004.14636

ATL-PHYS-PUB-2020-017



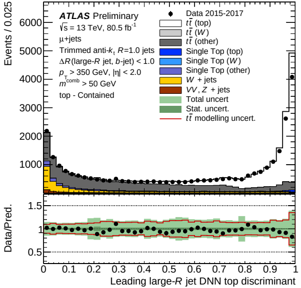
ATL-PHYS-PUB-2020-008



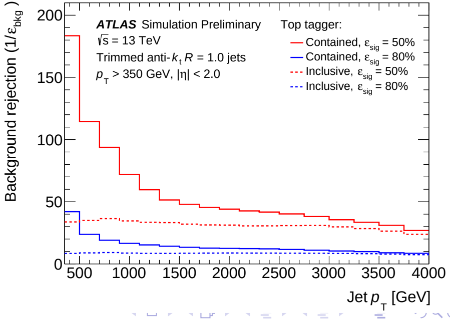
Boosted Jet Taggers : top taggers

- Deep Neutral Network (DNN) : a multivariate algorithm for separation of signals and backgrounds
- Inputs : 13 jet moments of LCTopo jets
  - p_{T}, m^{comb}, energy correlation ratios, N-subjettiness and its ratio, splitting factors, W reconstruction.
- "Inclusive" top and "contained" top (the truth jet includes most of the decay products) with flat 50% and 80% efficiency

ATL-PHYS-PUB-2020-017



ATL-PHYS-PUB-2020-017



Calibration in data

- The tagger performance in MC is calibrated to match the performance in data by deriving scale factors and uncertainties:

$$SF = \frac{\epsilon_{data}}{\epsilon_{MC}}$$

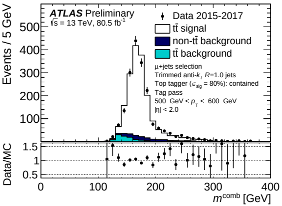
- Calibration of signal jets
  - **semileptonic $t\bar{t}$ event (muon channel)** : $\Delta R(\text{leading large}-R \text{ jet, b}-\text{tagged VR jet})$
    - ⋆ < 1.0 (top taggers)
    - ⋆ > 1.0 (W taggers)
- Calibration of background jets
  - **multijet event** : leading large-R jet $p_{\mathrm{T}} > 500$ GeV
  - γ**+jet event** : photon $p_{\mathrm{T}} > 155$ GeV, leading large-R jet $p_{\mathrm{T}} > 200$ GeV
  - γ+jet for low p_{T} range (< 500 GeV), multijet for high p_{T} range (> 500 GeV)

Signal scale factors

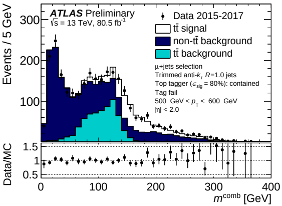
- Efficiency $\epsilon = \frac{N^{tagged}}{N^{tagged} + N^{not\ \ tagged}}$
  - Data efficiencies are derived from the simultaneous fit to tagged and not tagged leading jets for each p_{T} bin.
- p_{T} range : 350 GeV to 1000 GeV (top), 200 GeV to 600 GeV (W)
- Dominant systematics : $t\bar{t}$ modelling, heavy flavor tagging, large-R jet JES
- High p_{T} extrapolation : up to 4 TeV(top), 2.5 TeV(W)

  Use same scale factors in last p_{T} bin but derive uncertainties with variations of alternative generators, Geant4 configurations and ATLAS geometry.

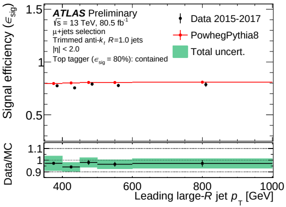
ATL-PHYS-PUB-2020-017



ATL-PHYS-PUB-2020-017



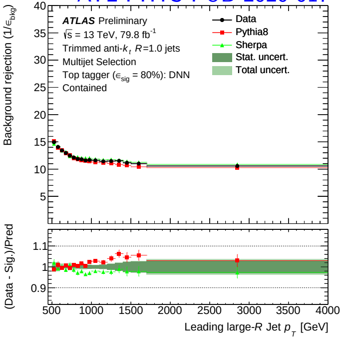
ATL-PHYS-PUB-2020-017



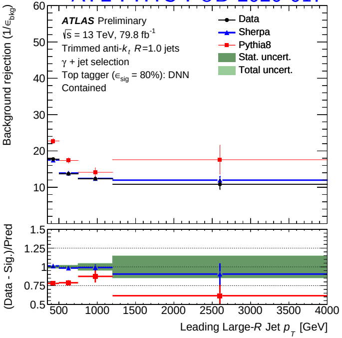
Background scale factors

- Efficiency $\epsilon_{MC} = \frac{N^{tagged}_{multijet/\gamma+jet}}{N^{tagged}_{multijet/\gamma+jet} + N^{not\ tagged}_{multijet/\gamma+jet}}$, $\epsilon_{data} = \frac{N^{tagged}_{data} - N^{tagged}_{other\ MC}}{N^{total}_{data} - N^{total}_{other\ MC}}$
  - SFs are derived as a function of p_{T} and $\ln(m/p_{\mathrm{T}})$
- p_{T} range : 350 GeV to 4000 GeV(top taggers), 200 GeV to 2500 GeV(W taggers)
- Dominant systematics : Large-R jet JES and generator difference.

ATL-PHYS-PUB-2020-017



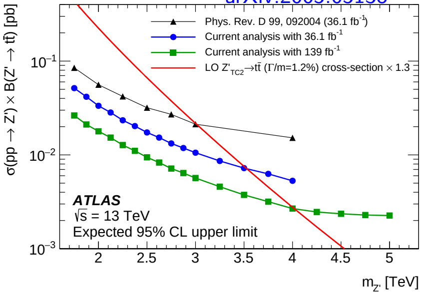
ATL-PHYS-PUB-2020-017



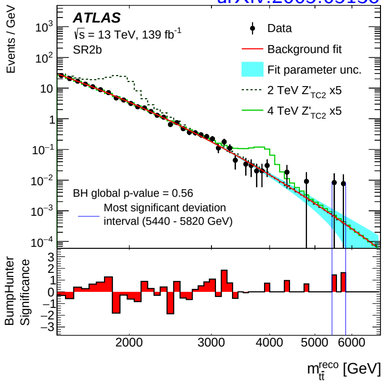
$t\bar{t}$ resonances

- Z'$\rightarrow t\bar{t}$ in fully hadronic final states predicted by topcolor-assisted-technicolor with full Run2 dataset is reported in arXiv:2005.05138
- Z' mass in mass range from 1.75 TeV to 5 TeV is searched.
  $\rightarrow$ boosted top $\rightarrow$ large-R LCTopo jet is used.
- DNN contained top tagger with 80 % efficiency is used.
  - Likelihood ratios based on τ_{32} are used in previous results with 36 fb^{-1}.

arXiv:2005.05138



arXiv:2005.05138

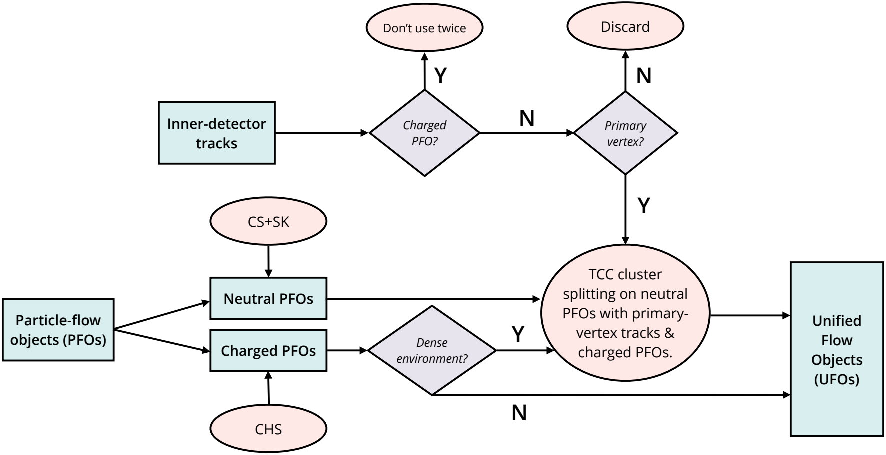


Summary

- Jet reconstruction
  - LCTopo jets, TCC jets and UFO jets with $R = 1$ are used for the identification of boosted top quarks and W/Z bosons.
- Taggers
  - W/Z taggers are developed for both LCTopo jets and TCC jets
  - Top taggers are developed by using DNN algorithm for LCTopo jets.
- Calibration
  - The scale factors of both signal and background are estimated and finalized.
- Physics analysis
  - DNN top tagger with the calibration is successfully applied to $t\bar{t}$ resonance.

Backup

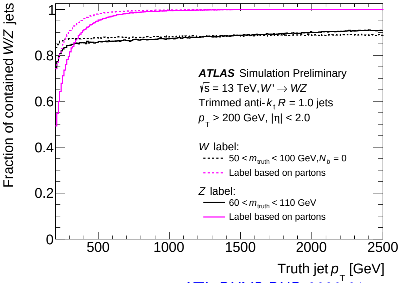
Unified flow objects



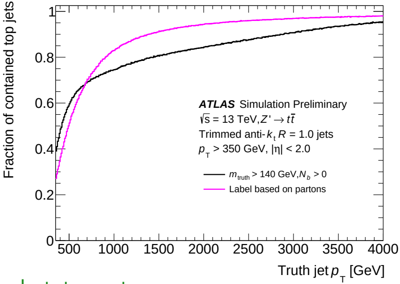
Jet truth labeling in Monte Carlo simulation

- Truth jet matching
  - $\Delta R(reco.jet, truthjet) < 0.75$
- W/Z jet
  - W/Z boson matches the truth jet by $\Delta R < 0.75$
  - $m^{truth} \in [50,\ 100]$ GeV for W jet
    $m^{truth} \in [60,\ 110]$ GeV for Z jet
- top jet
  - top quark matches the truth jet by $\Delta R < 0.75$
  - Contained top
    - ★ One ghost-associated B hadron, $m^{true} > 140$ GeV
  - Inclusive top
    - ★ No further requirements

ATL-PHYS-PUB-2020-017



ATL-PHYS-PUB-2020-017

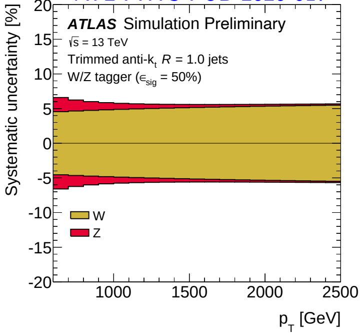


* Green lines denote the jet matched by using the final state partons.

High p_{T} extrapolation

- Due to the lack of the statistics in the high p_{T} region, high p_{T} extrapolation is required.
- The last bin derived from $t\bar{t}$ method is defined as the reference.
- Use same scale factors in last p_{T} bin but derive uncertainties with variations of alternative generators, Geant4 configurations and ATLAS geometry.
  - $SF^i(p_{\mathrm{T}}) = SF(p_{\mathrm{T}}^{ref}) \times R^i(p_{\mathrm{T}}),\ R^i(p_{\mathrm{T}}) = \frac{\epsilon^i_{MC}(p_{\mathrm{T}})/\epsilon^i_{MC}(p_{\mathrm{T}}^{ref})}{\epsilon_{MC}(p_{\mathrm{T}})/\epsilon_{MC}(p_{\mathrm{T}}^{ref})}$

ATL-PHYS-PUB-2020-017



ATL-PHYS-PUB-2020-017