

Improving ATLAS Hadronic Object Performance with ML/AI Algorithms

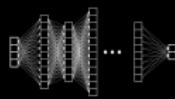
SUSY2023

Tobias Fitschen on behalf of the ATLAS Collaboration

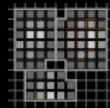
17 July 2023

University of Manchester

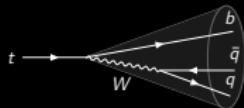




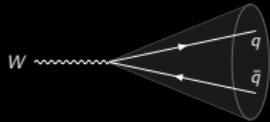
Neural Networks



Calo Clusters



Top Tagging

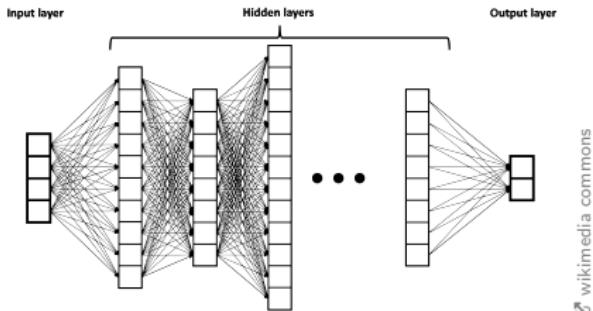


W/Z Tagging



Lund Jet Plane Tagger

Multilayer Perceptron (MLP)

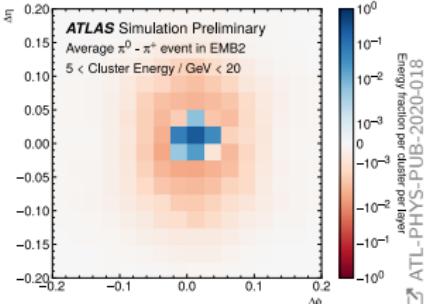


- **Most simple architecture**
- Fully connected layers
- Feed forward
- "Deep Neural Net" usually means this
- One or two outputs
- Binary or multi-classification

Convolutional Neural Net (CNN)

- Developed for image processing/classification
 - Input has to be projected into images (loss of information)
- Take advantage of hierarchical pattern in data
- **Identifies spatially localized features**

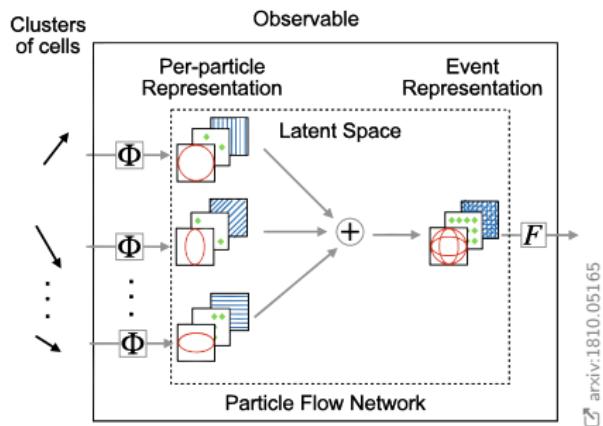
π^0/π^\pm calorimeter shower as image



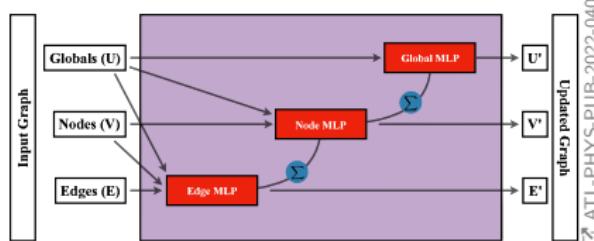
Deep Sets

☒ (Energy/Particle Flow Network)

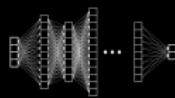
- **Unordererd, variable length input**
- E.g. of jet constituent momenta
- Permutation invariant



☒ Graph Neural Net (GNN)



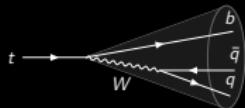
- **Structured representation:** nodes V , edges E
- Pairwise message passing
 - Nodes iteratively updated by exchanging information with neighbors
- Permutation invariant
- E.g. neighboring calo cells connected via edges in GNN



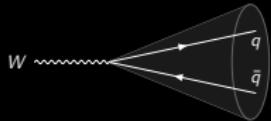
Neural Networks



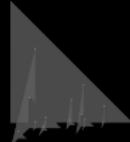
Calo Clusters



Top Tagging



W/Z Tagging



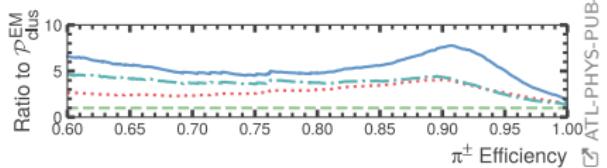
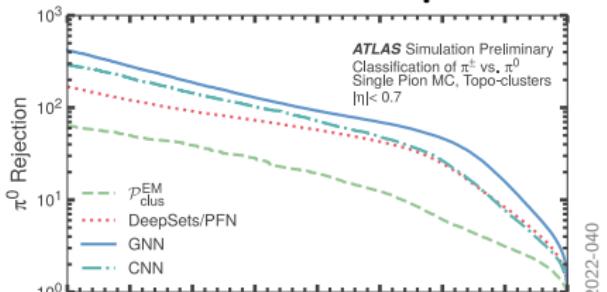
Lund Jet Plane Tagger

π^0 vs π^\pm Shower Classification

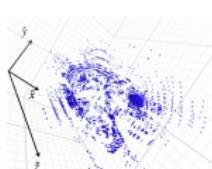
Non-compensating ATLAS calorimeter requires different calibrations for neutral/charged clusters

First step in cluster calibration: Differentiate EM from hadronic clusters

π^0 vs π^\pm classification performance

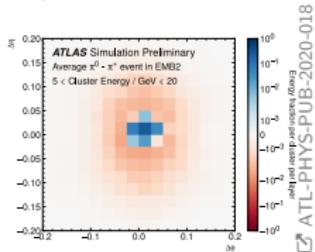


Point cloud of energy deposits in calorimeter cells



ATL-PHYS-PUB-2022-040

Image: $\pi^0 - \pi^\pm$ difference



Baseline used in LCW: $\mathcal{P}_{\text{clus}}^{\text{EM}}$

- Binned EM-scale cluster variables
 - Total cluster energy $E_{\text{cluster}}^{\text{EM}}$
 - Pseudorapidity η
 - Longitudinal depth λ_{clus}
 - 1st cell energy density moment $\langle \rho_{\text{cell}} \rangle$
- Combined into likelihood $\mathcal{P}_{\text{clus}}^{\text{EM}}$

Individual calorimeter cell signals

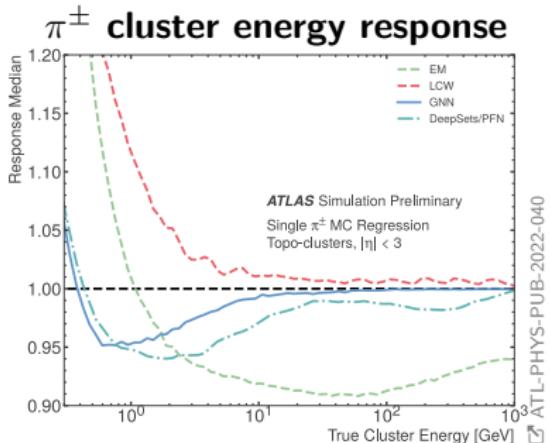
→ As point clouds (**GNN**, **PFN**)

→ Or projected on images (**CNN**)

Observations

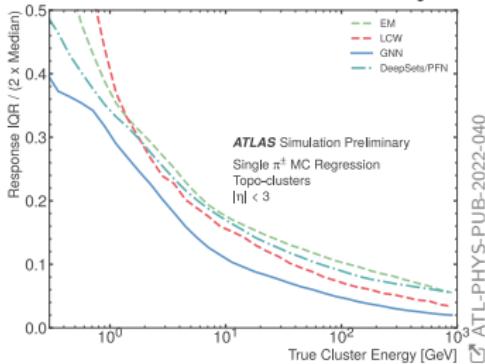
- All point cloud methods significantly outperform baseline $\mathcal{P}_{\text{clus}}^{\text{EM}}$

Energy Regression



Interquantile range IQR (measure for spread)

Calorimeter clusters only



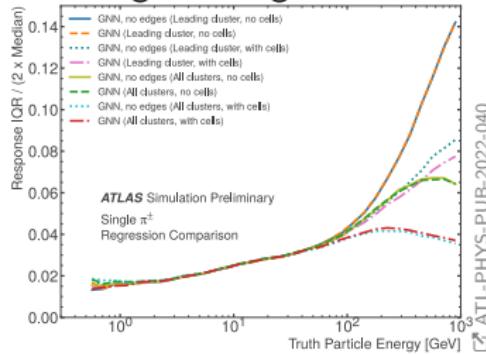
Second step: Energy Calibration

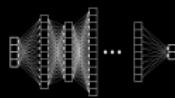
Here: charged (π^\pm) clusters

Observations

- **GNN** performs best wrt. response and width
- Followed by **Deep Sets**
- Similar for neutral π^0 clusters (see backup)

Including tracking information

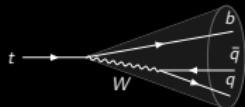




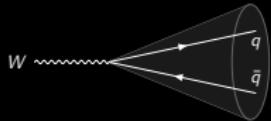
Neural Networks



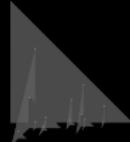
Calo Clusters



Top Tagging



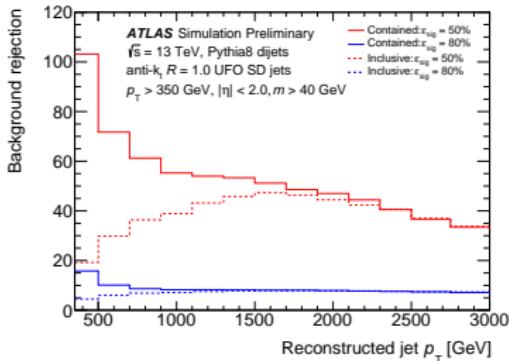
W/Z Tagging



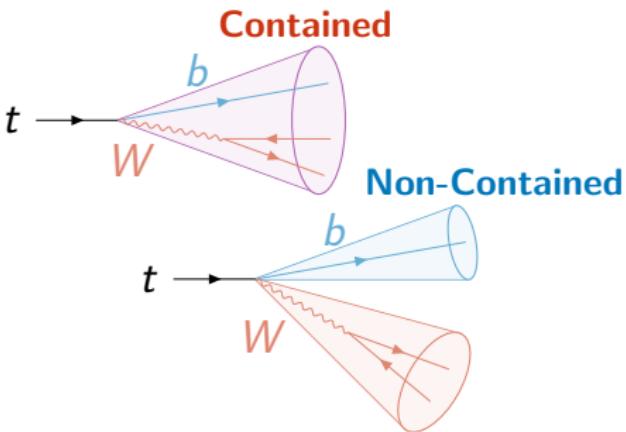
Lund Jet Plane Tagger

2 taggers: Inclusive and contained

- Fixed working points: 50% and 80%
 - Defined as function of p_T
- For Unified Flow Objects (UFO) jets:
 - Combining information from calorimeter + tracker



ATL-PHYS-PUB-2021-028



MLP on Substructure Variables

| | |
|--|--------------------------------|
| $\tau_1, \tau_2, \tau_3, \tau_4$ | N -subjettiness |
| $\sqrt{d_{12}}, \sqrt{d_{23}}$ | Splitting scales |
| $\text{ECF}_1, \text{ECF}_2, \text{ECF}_3$ | Energy correlation functions |
| C_2, D_2 | Energy correlation ratios |
| L_2, L_3 | Generalised energy correlation |
| Q_W | Invariant mass / virtuality |
| T_M | Thrust major |

(see backup for definitions)

ATL-PHYS-PUB-2022-39: Constituent-Based Top-Quark Tagging

DNN top tagger (prev. slides):

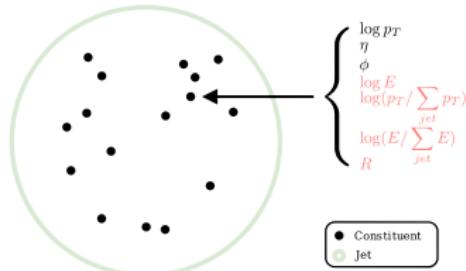
- Set of high-level features (substructure variables)
- Used as baseline (hIDNN)

Constituent-based taggers:

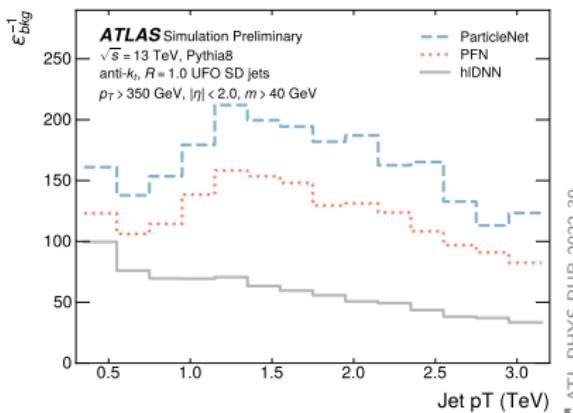
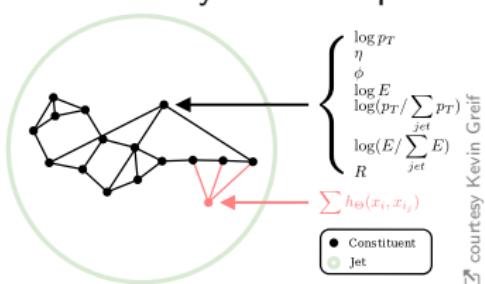
- Low-level features based on 4-vectors of jet constituents

→ Up to $\times 2$ improvement over baseline (hIDNN)!

PFN: Particle Flow Network



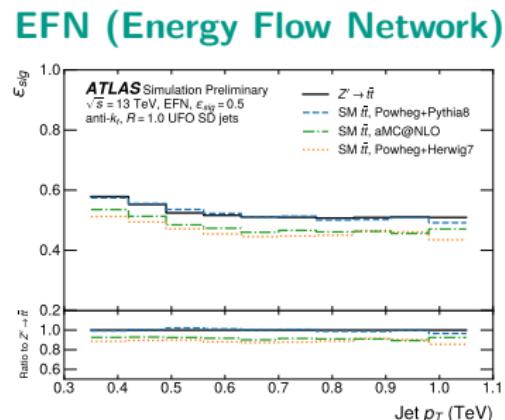
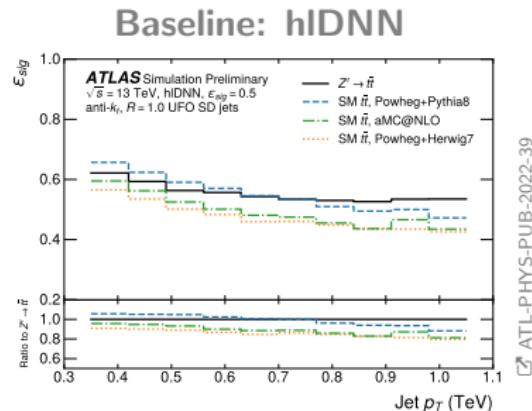
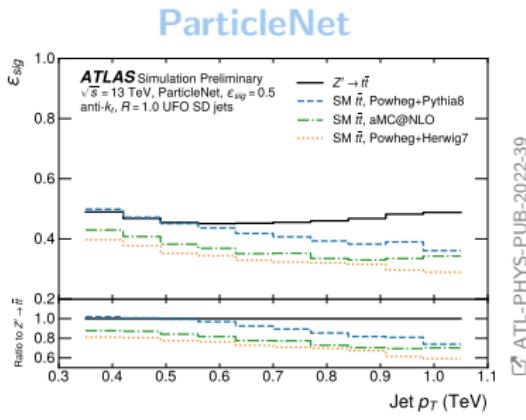
ParticleNet: Dynamic Graph-CNN

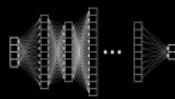


Constituent-Based Top Tagger

Model dependence

- Different parton shower and hadronisation models
- ϵ^{sig} measured at threshold for $\epsilon^{\text{sig}} = 50\%$ in nominal sample
- **ParticleNet**: More model dependent than hIDNN
- **EFN**: Less model dependent

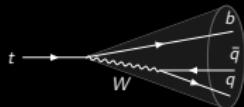




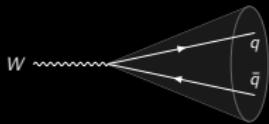
Neural Networks



Calo Clusters



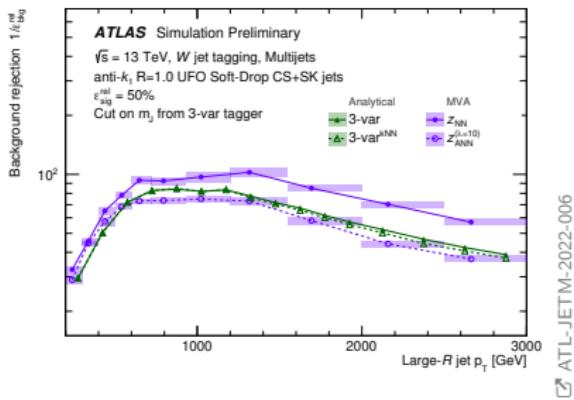
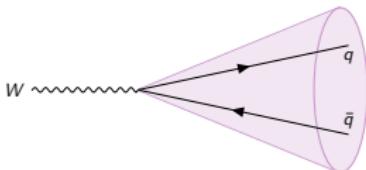
Top Tagging



W/Z Tagging



Lund Jet Plane Tagger



3-variable cut based:

- D_2 Energy correlation ratios
- N_{trk} Number of associated tracks
- m Jet mass

Machine learning based (NN):

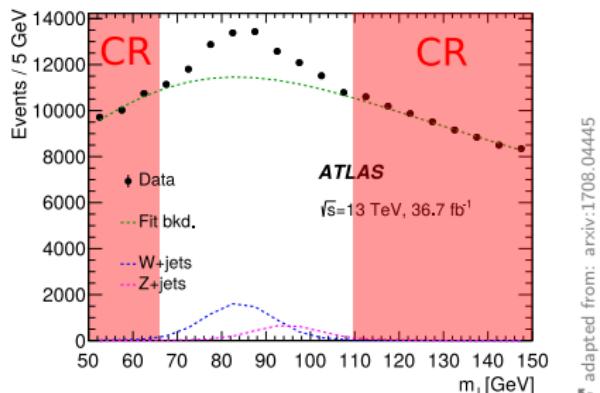
- Various substructure variables

- D_2, C_2 Energy correlation ratios
- T_{21} N -subjettiness
- R_2^{FW} Fox-Wolfram moment
- \mathcal{P} Planar flow
- a_3 Angularity
- A Aplanarity
- Z_{cut} Z -Splitting scales
- $\sqrt{d_{12}}$ d -Splitting scales
- $Kt\Delta R$ k_t -subjett ΔR
- N_{trk} Number of associated tracks
(see backup for definitions)

- **NN** tagger significantly outperformed **cut based 3-var** tagger
- Even mass-decorrelated version **ANN** shows similar performance to **cut based 3-var** using m

The Need for Mass Decorrelation

Data-driven background estimates:



- Define mass side-bands as **control regions (CR)**
- Fit smooth function to data from left to right side-band
- Estimate **background** in signal region from fit

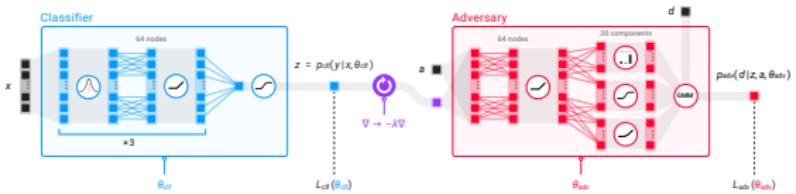
Problem: Tagger may introduce unwanted shaping of background, de-populating the sideband regions

Solution: Decorrelate tagger decision from m_j :

- Adversarial neural networks (ANN) for NN tagger

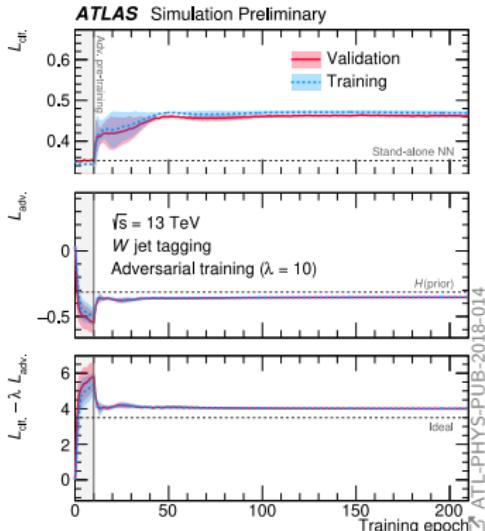
Architecture

- **Classifier** MLP for tagging task
- **Adversary** for decorrelation
(e.g. of mass)
 - Predicts mass based on classifier output (+auxiliary variables)
- **Gradient reversal layer:** During back-propagation penalise Classifier if Adversary predicts mass too well
 - Final tagger only consists of Classifier
 - Tagger decision decorrelated to mass



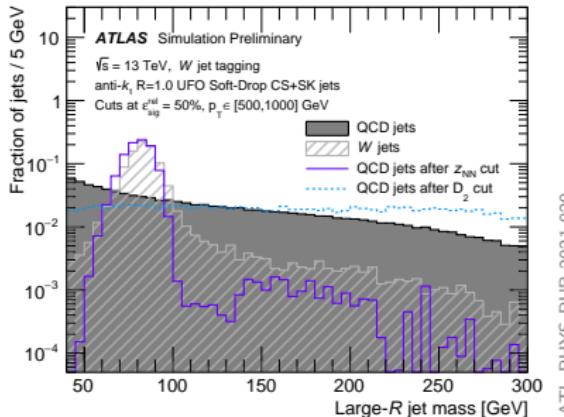
Training schedule

1. **Classifier alone**
2. **Adversary alone**
3. **Both together**

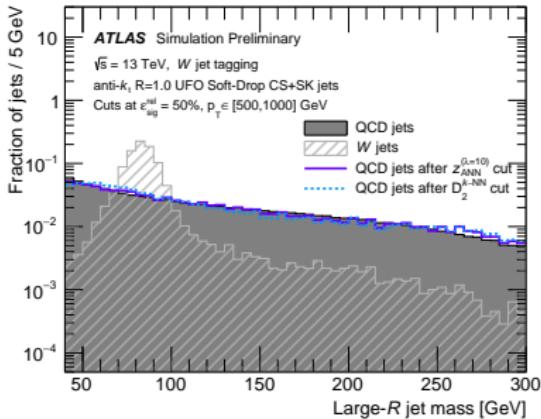


Mass Decorrelation with Adversarial Training

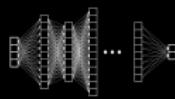
NN: Correlated to m_j



ANN: Active decorrelation



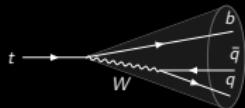
- **Background** mass distribution shaped according to signal by NN
- **Adversarial Neural Network (ANN) successfully decorrelates**



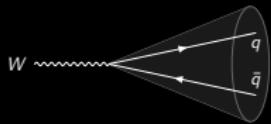
Neural Networks



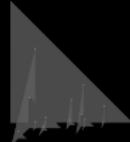
Calo Clusters



Top Tagging



W/Z Tagging

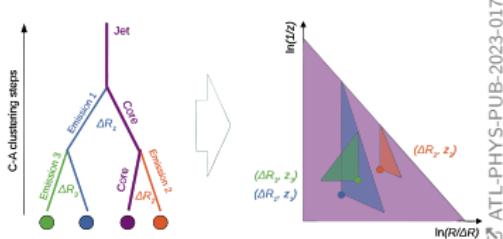


Lund Jet Plane Tagger

The Lund Jet Plane (LJP)

Lund Plane

- Study clustering history of jet algorithm in 2D(x,y) space[†]:
 - x: Opening angle ΔR_{ij}
 - y: p_T -fraction $k_T = p_T^j \Delta R_{ij}$
 - of emission j wrt emitter i
- C/A: Angle-ordered clustering

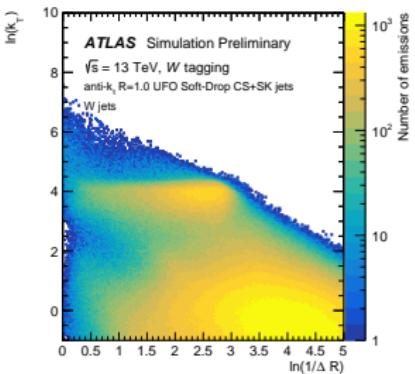


Lund Jet Plane (LJP)

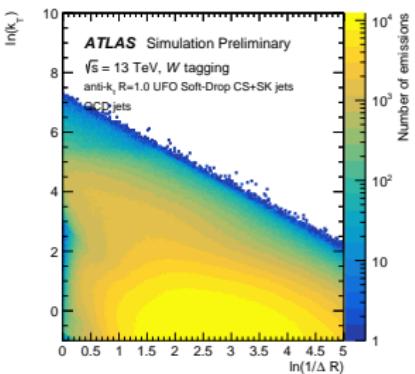
- Jet constituents as proxy for emissions
 - Reco-level observable
 - Usable for tagging!

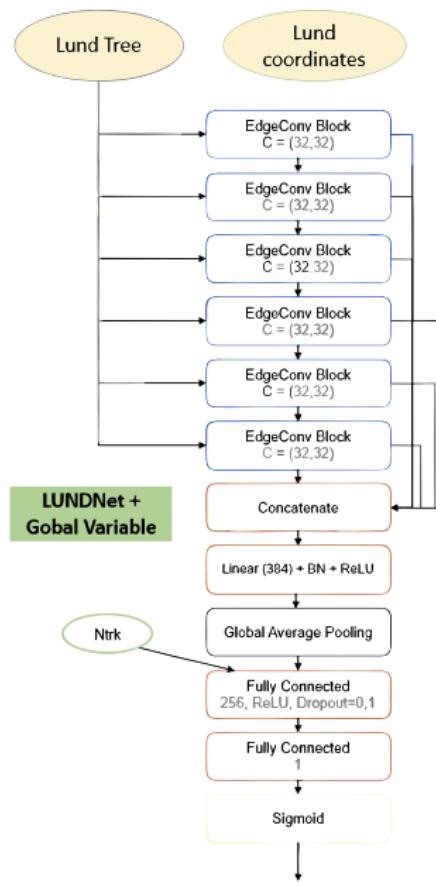
[†]: Tagger additionally uses momentum fraction $z = p_T^j / (p_T^j + p_T^i)$

Primary LJP: W jets



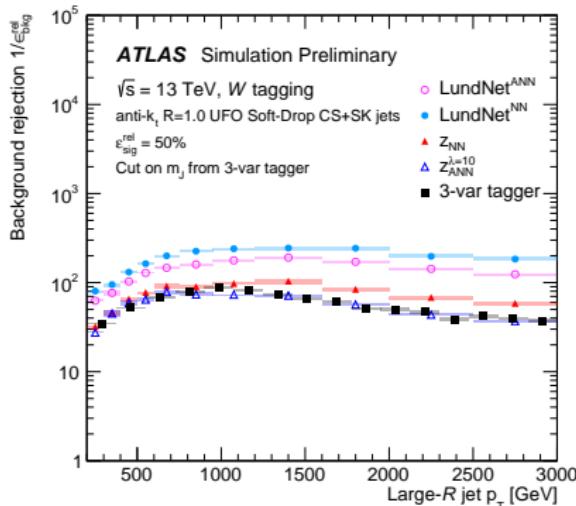
Primary LJP: QCD jets



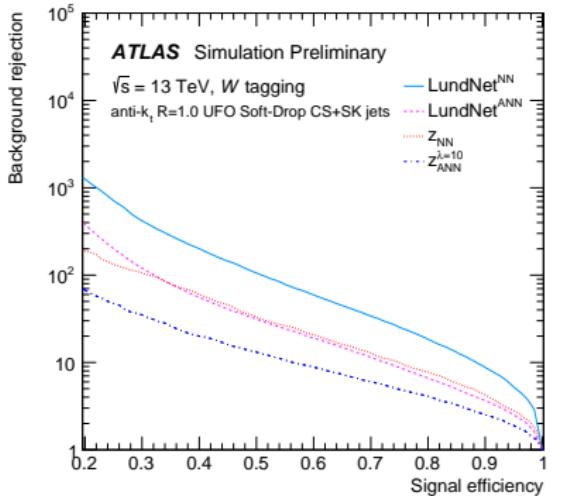


The LJP Tagger

- W tagger based on  LundNet
- **Graph NN** acting on LJP variables
- And n_{trk} as additional global variable
- ANN mass-decorrelated version available: **LundNet^{ANN}**

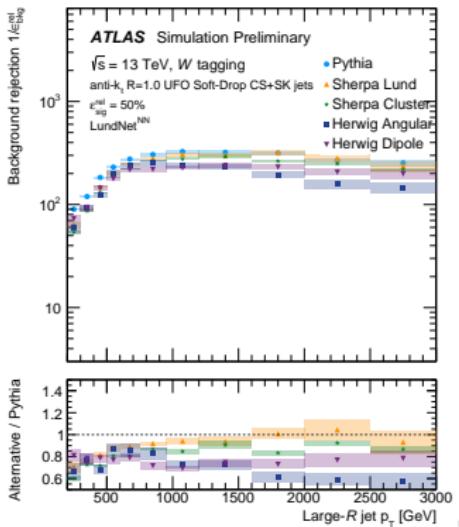


LJP outperforms SSV by far



ATL-PHYS-PUB-2023-017

but sensitive wrt modelling

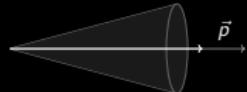


ATL-PHYS-PUB-2023-017

- LJP makes use of low-level information lost in substructure variables (SSV)
 - **Better background rejection**
 - Decorrelated LJP shows similar efficiency as nominal SSV tagger!
- But **low-level information more dependent on modelling**
 - LJP should be less model-dependent compared to using raw constituents

...And Many More

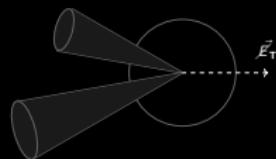
More detail: Michael Holzbock's talk



Jet Energy Scale

GNNC: New Techniques for Jet Calibration

[More in appendix](#)



Missing E_T

METNet: ML-Based \cancel{E}_T Reconstruction

[More in appendix](#)

Summary

Many ML applications for hadronic objects in ATLAS

- Calorimeter cluster classification and energy regression
- W/Z and top tagging with and without mass decorrelation
- LundNet
- Jet energy scale calibration
- MET calibration
- Many more...

Constituent based methods perform best in all domains

- In most cases: GNN > Deep Sets > CNN > MLP > BDT > cuts
- **Important:** Better ROC curves are great, but data/MC agreement & model independence should not be neglected!

Appendix

Calo Clusters

Jet Energy Calibration

\cancel{E}_T Calibration

Top Taggers

W/Z Taggers

Lund Jet Plane Tagger

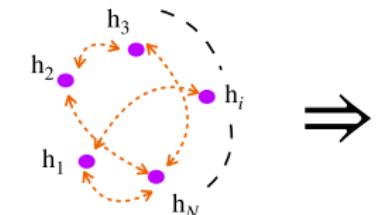
Substructure Variables

UFO Jets

Calo Clusters

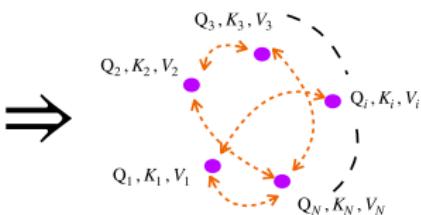
Appendix

Transformer for Graph Updates



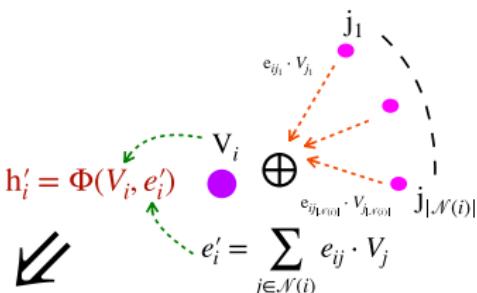
Start with a graph G having N nodes with node-features h_i on the i -th node.

$$\begin{aligned} Q_i &= \Theta_1(h_i) \\ K_i &= \Theta_2(h_i) \\ V_i &= \Theta_3(h_i) \end{aligned}$$

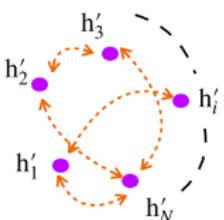


Form the query, key and value features using three MLP.

$$e_{ij} = \sigma\left(\frac{Q_i \cdot K_j^T}{\sqrt{d}}\right)$$

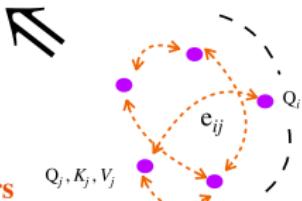


Perform the node aggregation through sum pooling and compute the new node features h'_i .

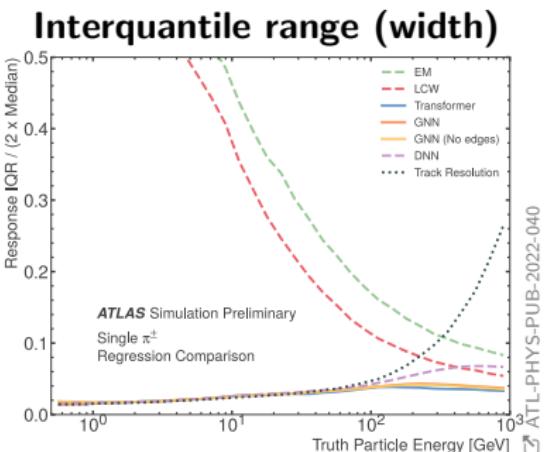
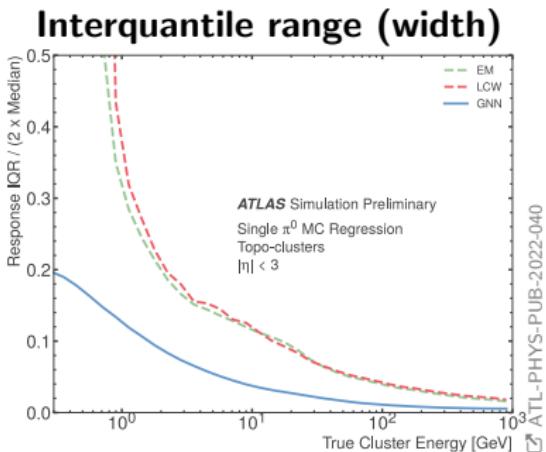
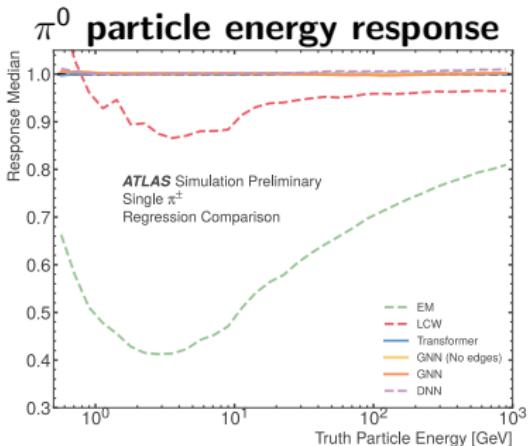
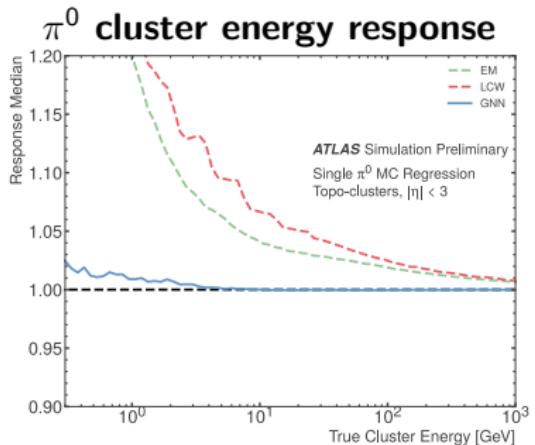


$\Theta_1, \Theta_2, \Theta_3, \Phi$ has trainable parameters

Create edge data using attention mechanism



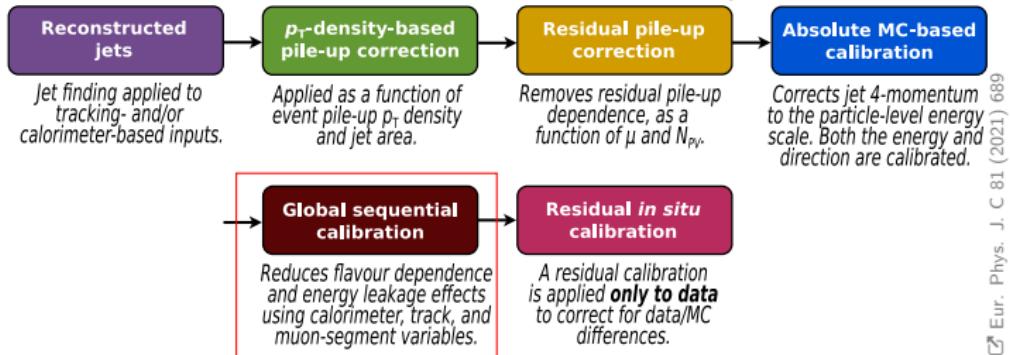
Performance



Jet Energy Calibration

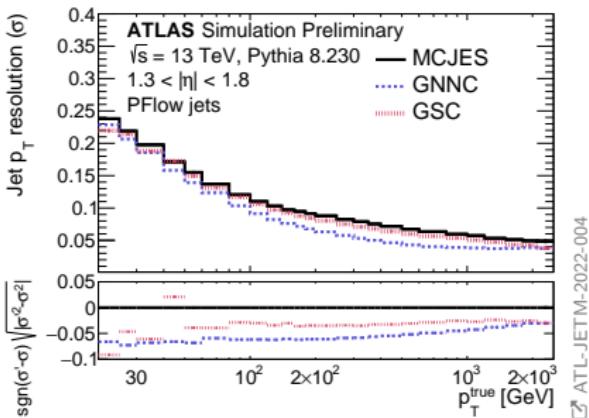
Appendix

ATL-JETM-2022-004: ML based Global Sequential Calibration (GSC)



Eur. Phys. J. C 81 (2021)

After energy scale calibrated on average, GSC corrects for small differences for different jet flavours



- **GSC sequentially** corrects for each variable
 → Does not exploit correlations
- New method (**GNNC**) uses MLP trained to predict p_T response
 → Improvement over full p_T range

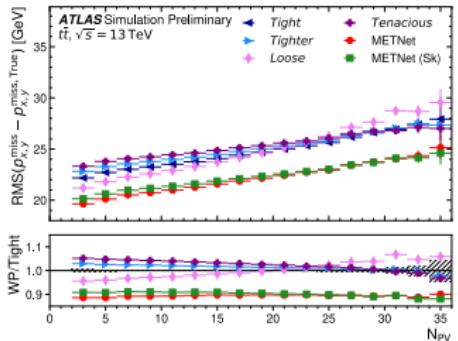
E_T Calibration

Appendix

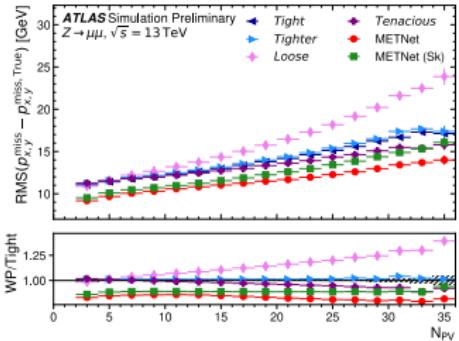
METNet: A combined p_T^{miss} working point (☒ ATL-PHYS-PUB-2021-025):

- p_T^{miss} in ATLAS: Negative sum of calibrated momenta of hard objects (e , μ , τ -jets, γ , jets)
- Plus soft term: Tracks from PV not associate to hard objects
- Different working points (WPs) defined for various pileup conditions
 - E.g. "tight": Higher p_T cuts on forward jets
- MetNet: MLP combining p_T^{miss} values from different WPs
 - Based on event kinematics and conditions
- Overall better performance than any WP alone

Trained among others on $t\bar{t}$



Extrapolates well to $Z \rightarrow \mu\mu$



→ E_T^{miss} definition depends on process but MetNet performs best for all

Top Taggers

Appendix

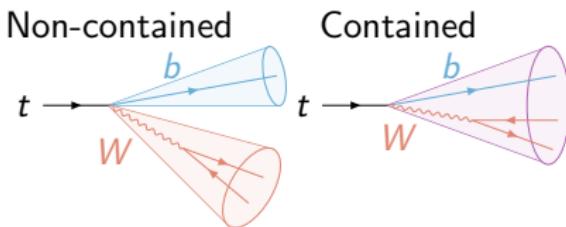
Two types of DNN-based top taggers defined:

Inclusive: Purely defined by ΔR matching:

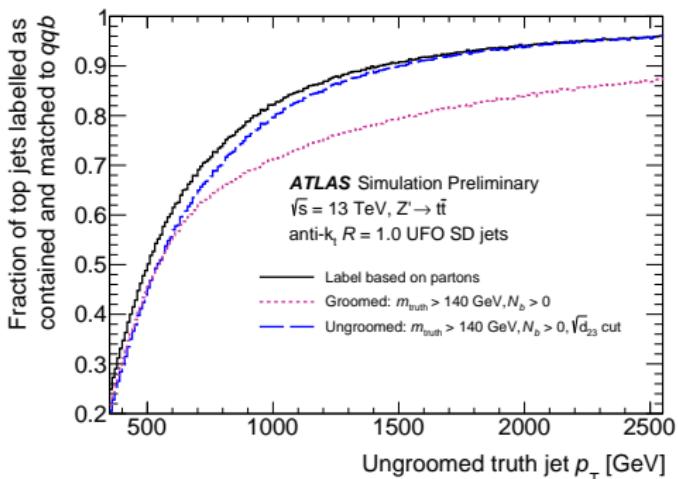
- $\Delta R(j^{\text{reco}}, j^{\text{truth}}) < 0.75$ and $\Delta R(j^{\text{truth}}, t^{\text{truth}}) < 0.75$

Contained: Extra cuts to ensure $t \rightarrow qqb$ fully contained within jet:

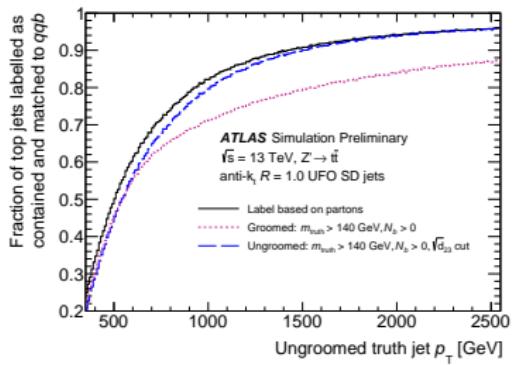
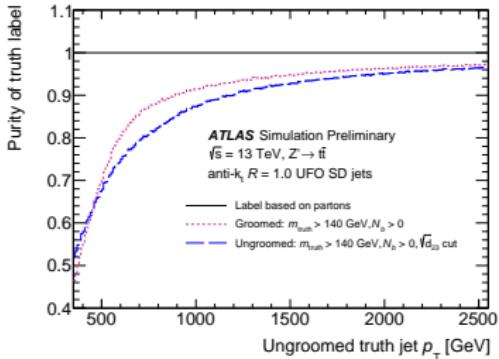
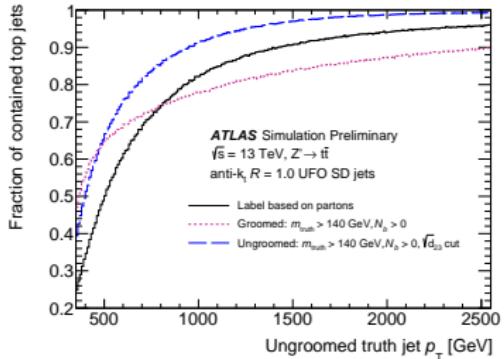
- Same ΔR requirement
- $N_B \geq 1$
- $m_{\text{ungroomed}}^{\text{truth}} > 140 \text{ GeV}$
- $\sqrt{d_{23}} > 27 e^{-\frac{p_T}{1433 \text{ GeV}}} \text{ GeV}$



Contained top labelling efficiency:

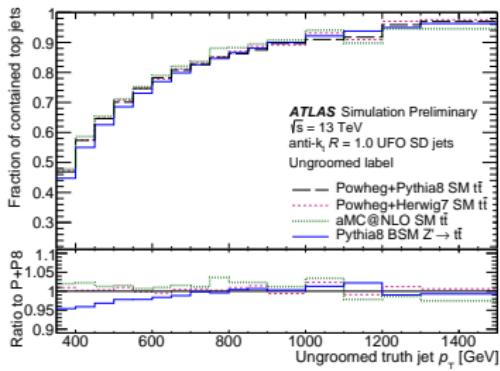
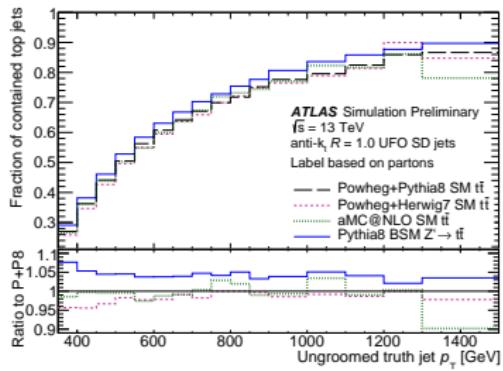


DNN Top Tagger: Truth Labelling

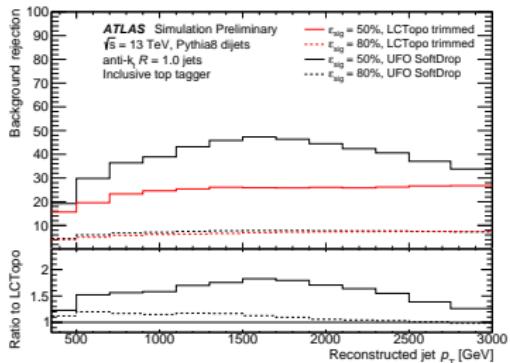
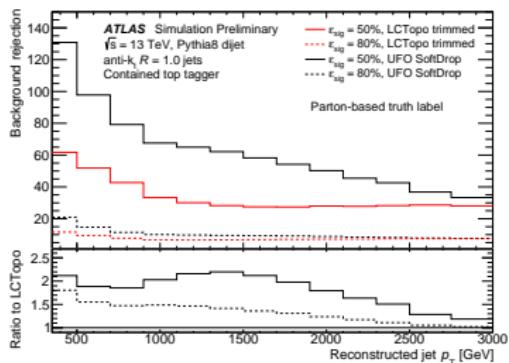
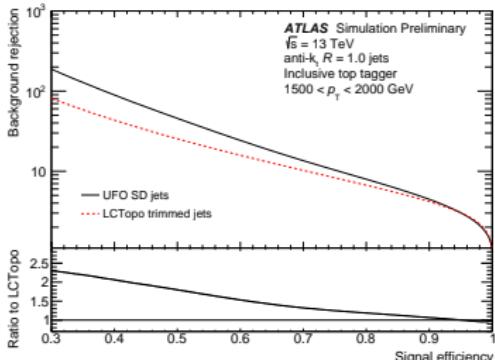
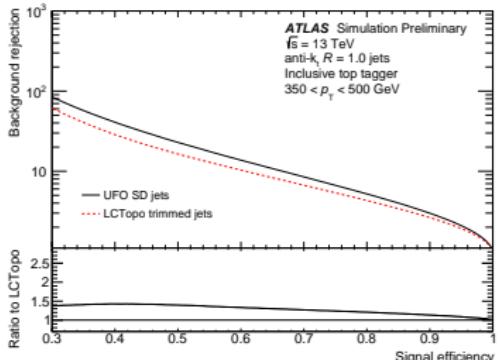


ATL-PHYS-PUB-2021-028

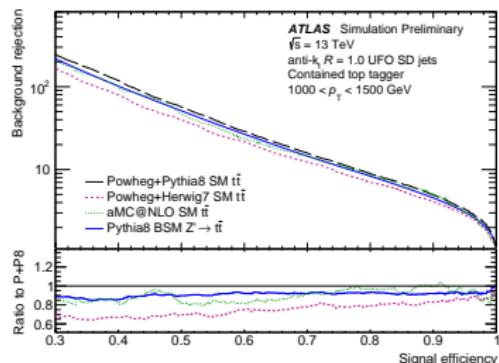
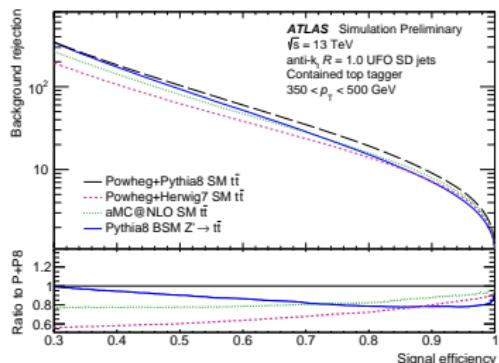
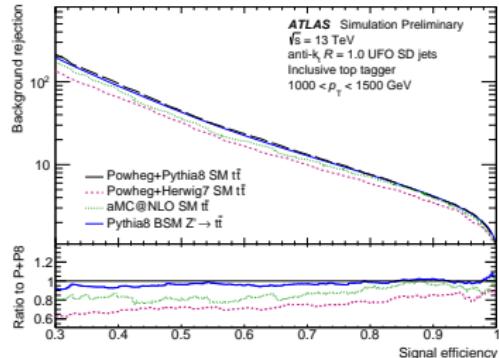
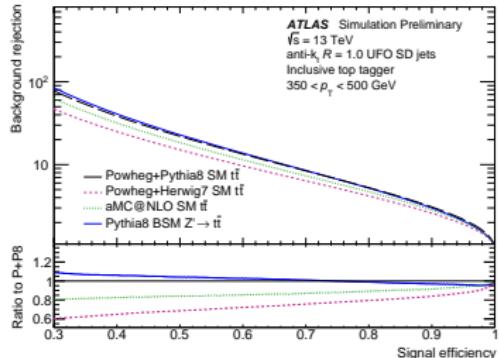
DNN Top Tagger: Truth Labelling



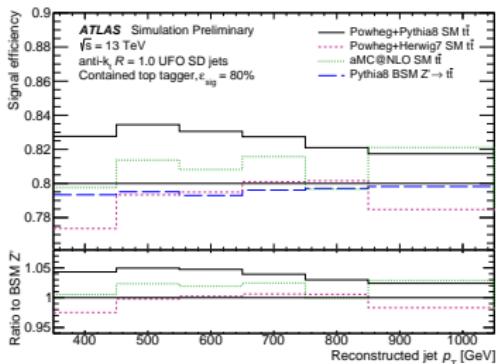
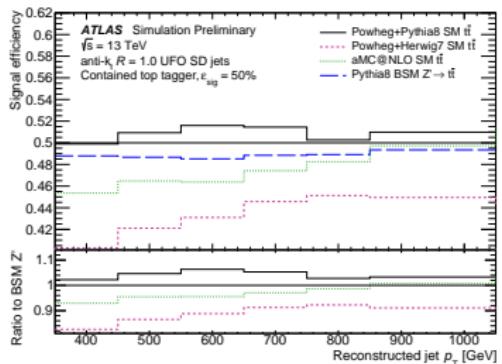
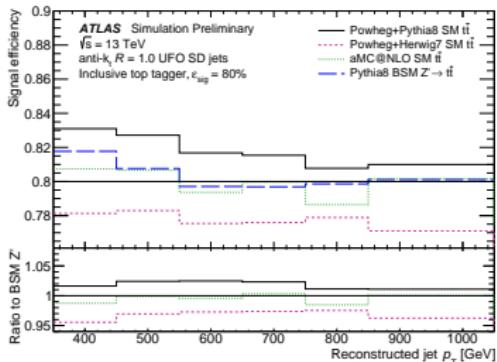
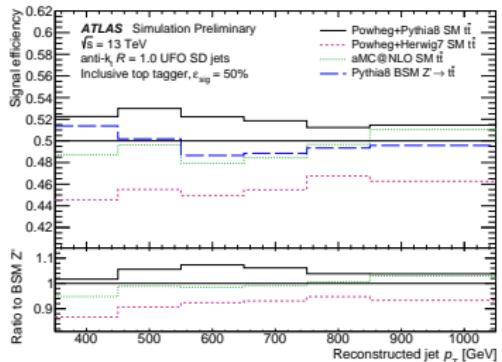
DNN Top Tagger: ROC Curves



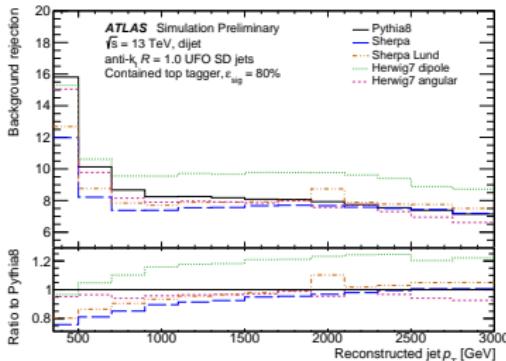
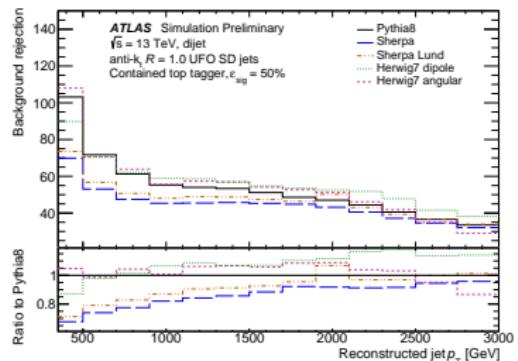
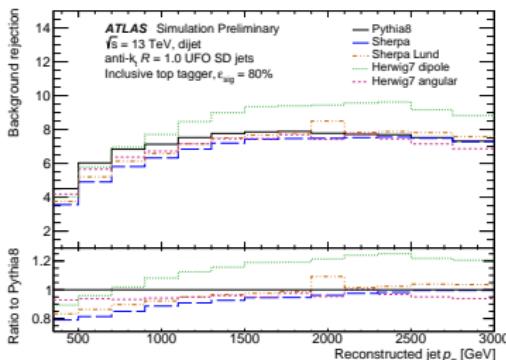
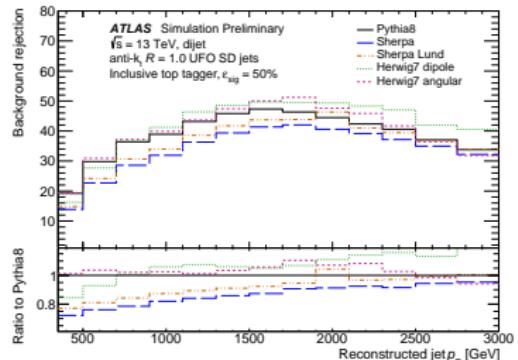
DNN Top Tagger: Modelling



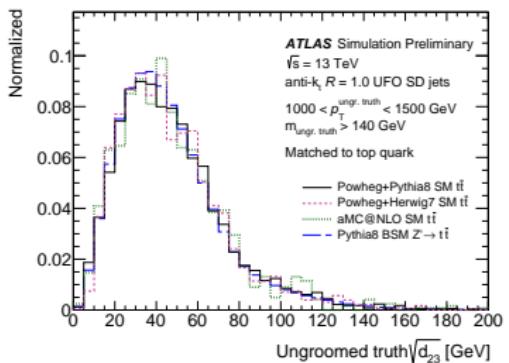
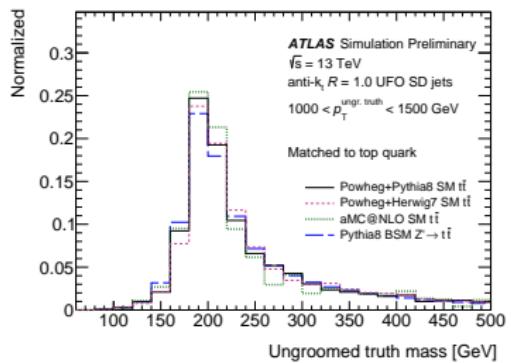
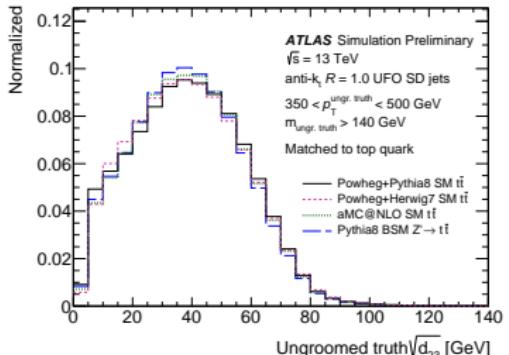
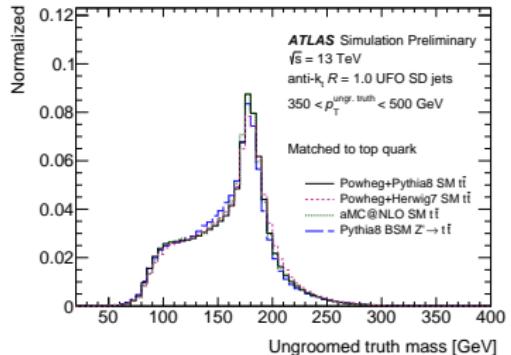
DNN Top Tagger: Modelling



DNN Top Tagger: Modelling



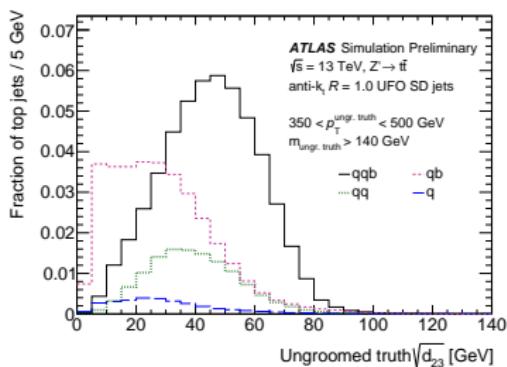
DNN Top Tagger: Modelling for Truth Labelling



DNN Top Tagger: Truth Labelling

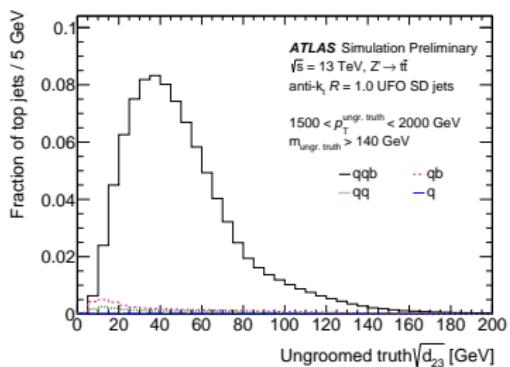
Low p_T :

→ optimal cut: $\sqrt{d_{23}} \approx 21$ GeV



High p_T :

→ optimal cut: $\sqrt{d_{23}} \approx 7$ GeV



Constituent-Based Top Tagger: Performance

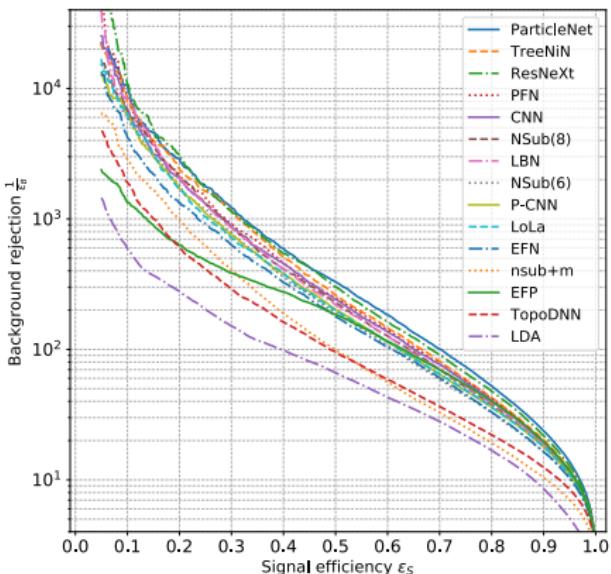


| Model | AUC | ACC | ε_{bkg}^{-1} @ $\varepsilon_{sig} = 0.5$ | ε_{bkg}^{-1} @ $\varepsilon_{sig} = 0.8$ | # Params | Inference Time |
|-------------|-------|-------|--|--|-----------|----------------|
| ResNet 50 | 0.885 | 0.803 | 21.4 | 5.13 | 1,486,209 | 9 ms |
| EFN | 0.901 | 0.819 | 26.6 | 6.12 | 1,670,451 | 4 ms |
| hIDNN | 0.938 | 0.863 | 51.5 | 10.5 | 93,151 | 3 ms |
| DNN | 0.942 | 0.868 | 67.7 | 12.0 | 876,641 | 3 ms |
| PFN | 0.954 | 0.882 | 108.0 | 15.9 | 689,801 | 4 ms |
| ParticleNet | 0.961 | 0.894 | 153.7 | 20.4 | 764,887 | 38 ms |

The Machine Learning Landscape of Top Taggers

doi:10.21468/SciPostPhys.7.1.014

- Comparison of many modern ML techniques applied to the top tagging task
- Simplified detector simulation with Delphes + ATLAS card
- Calorimeter information only
→ No tracking as in UFO jets



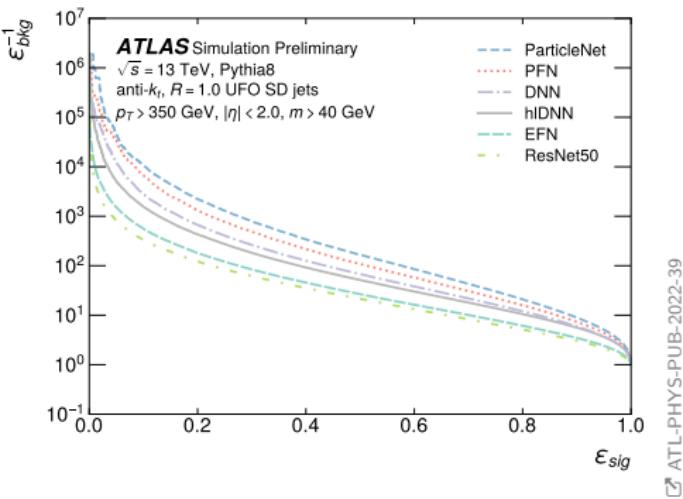
→ Best performance with ParticleNet and ResNeXt

Constituent-Based Top Tagger



ATL-PHYS-PUB-2022-39

New: How do these algorithms perform on ATLAS simulated UFO jets?



- **hIDNN:** Baseline similar to DNN top tagger
→ [ATL-PHYS-PUB-2021-028](#)
- **DNN:** Using constituent 4-momenta
→ [arxiv:1704.02124](#)
- **EFN/PFN:** Energy/Particle-flow networks
→ [arxiv:1810.05165](#)
- **ResNet50:** CNN using jet images
→ [arxiv:1512.03385](#)
- **ParticleNet:** Dynamic Graph-CNN
→ [arxiv:1902.08570](#)

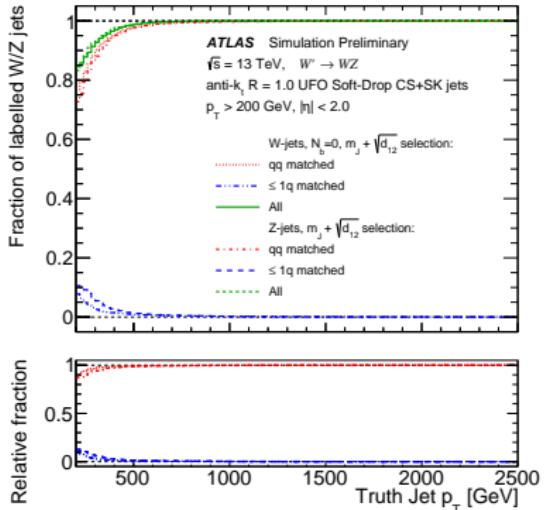
- ParticleNet and PFN show best performance

Simulated data made public for ML experts along with PUB-note!

W/Z Taggers

Appendix

Truth Labelling



Truth jet definition:

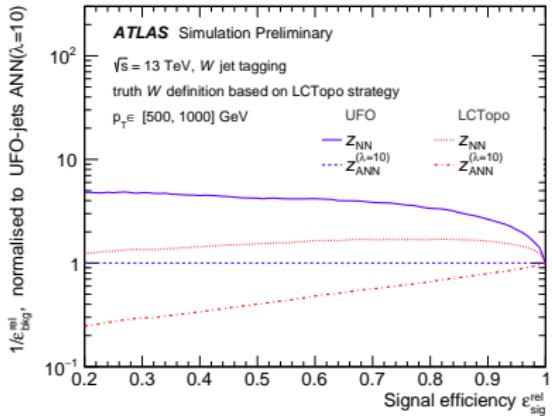
- Parton level MC to label jets as:
 - Signal: Containing full $W \rightarrow qq'$
 - Background: From single q/g
- Truth jets: Reconstructed from stable particles, anti- k_T $R = 1.0$
 - No grooming applied to ensure independence of grooming algorithm
- Requirements for truth signal jets:
 - Truth W/Z within $\Delta R < 0.75$
 - $m_J > 50$ GeV
 - $\sqrt{d_{12}} > 55.25 + e^{-2.35 \frac{p_T}{\text{GeV}}} \text{ GeV}$
 - $N_B = 0$ for W jets to reduce top contamination
- Matched to UFO jets with $\Delta R < 0.75$
- Optimised for $\epsilon_{\text{sig}} = 85\%$ at $p_T = 200$ GeV, 100% at $p_T = 300$ GeV

Matching to daughter quarks q, q' :

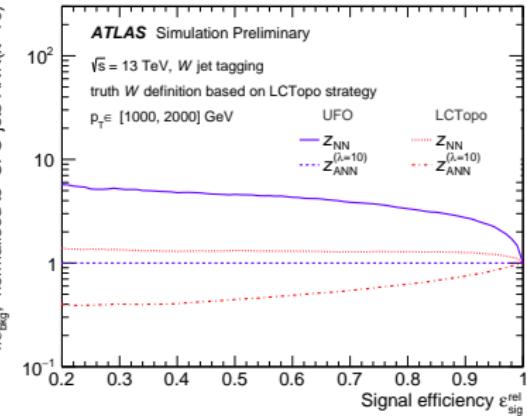
- Fraction of W (Z) containing both q, q' within $\Delta R < 0.75$:
 $> 98\%$ (96%) at $p_T > 300$ GeV
 (previously: 90%)

Tagger Performance: UFO vs LCTopo

Low p_T



High p_T



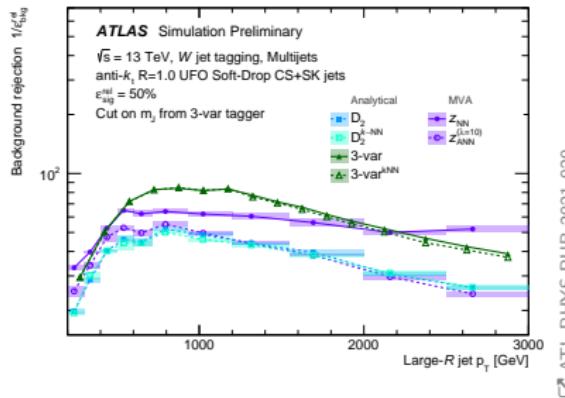
- Bkg rejection with **UFO** improved by a factor of 2-3 w.r.t. **LCTopo**!
- Mass decorrelation (ANN) comes with tradeoff of reduced efficiency
 - But may be better option for data-driven background estimates on m_J distribution

NN/ANN Tagger with n_{trk}



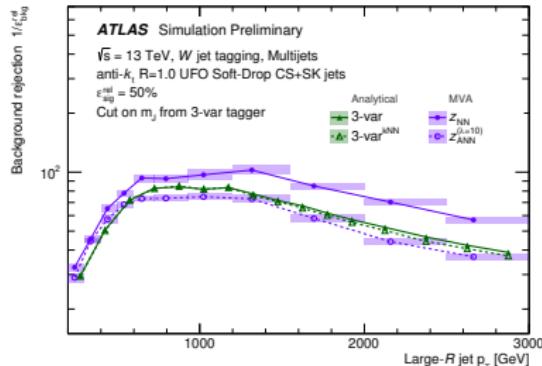
Previously:

ATL-PHYS-PUB-2021-029



Follow up publication:

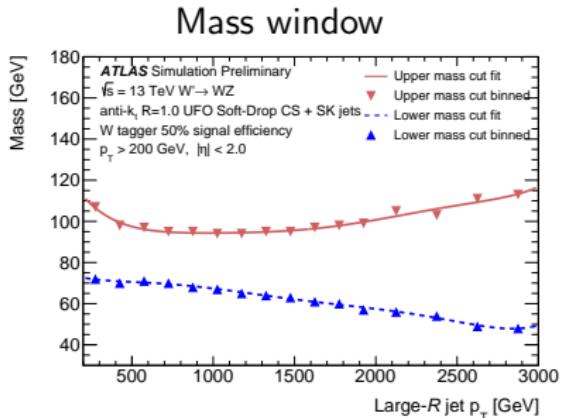
With n_{trk} as additional feature



ATL-JETM-2022-006

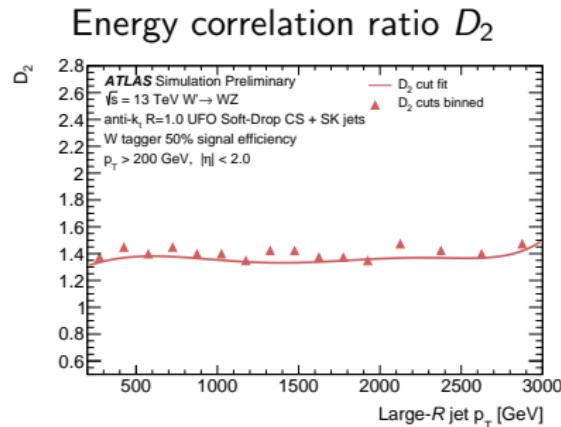
- Previously: 3-variable tagger showed better performance than NN
- Now: NN much better after including n_{trk} as additional feature!
 - ANN comparable with 3-variable tagger, but with decorrelation!
- Reason for such strong improvement:
 - Most other features exploit 2-prongedness of W/Z decay
 - n_{trk} is good quark/gluon discriminator

3-Variable Tagger

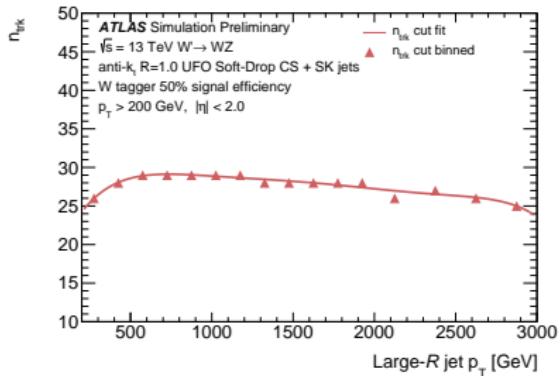


3-variable tagger: W

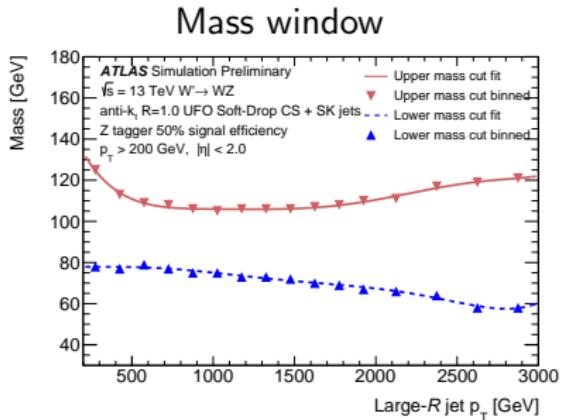
- p_T -dependent cuts on 3 features
- Maximise background rejection while keeping 50% signal efficiency per bin
- D_2 nearly flat in p_T
 - Thanks to angular resolution of UFO constituents
 - Fixed D_2 cut possible



Number of tracks n_{trk}

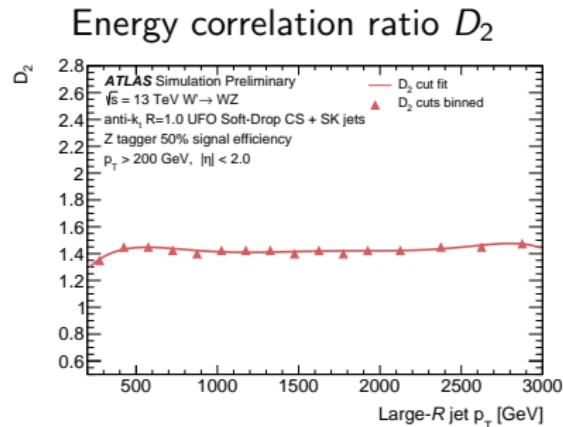


3-Variable Tagger

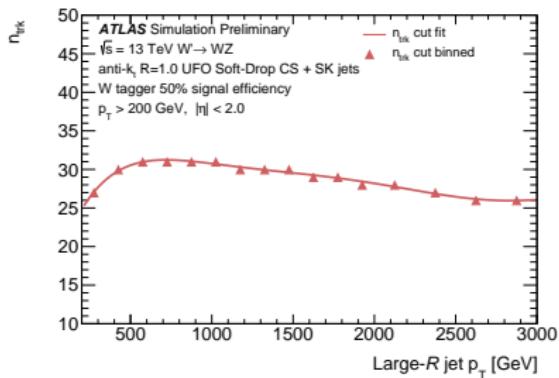


3-variable tagger: Z

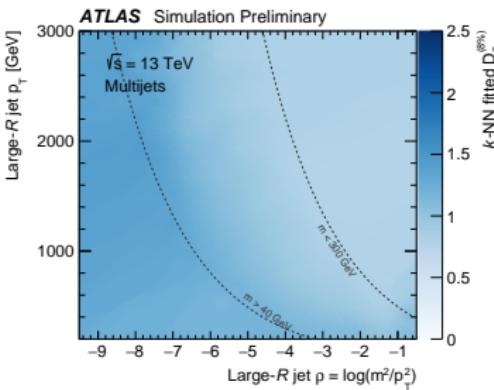
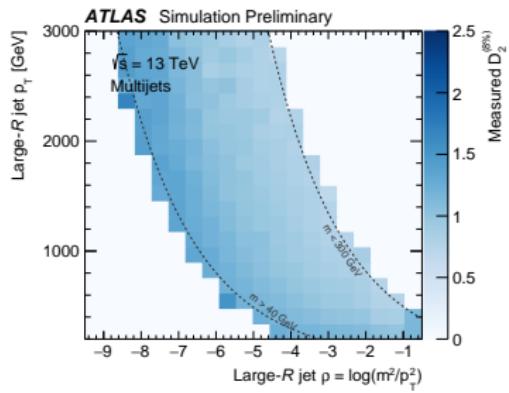
- p_T -dependent cuts on 3 features
- Maximise background rejection while keeping 50% signal efficiency per bin
- D_2 nearly flat in p_T
 - Thanks to angular resolution of UFO constituents
 - Fixed D_2 cut possible



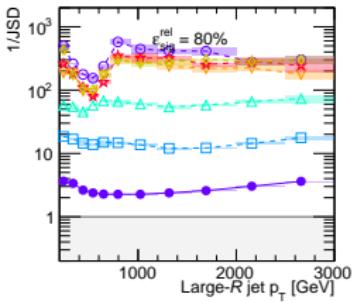
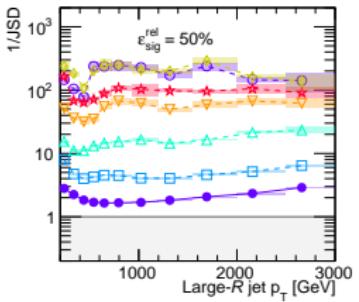
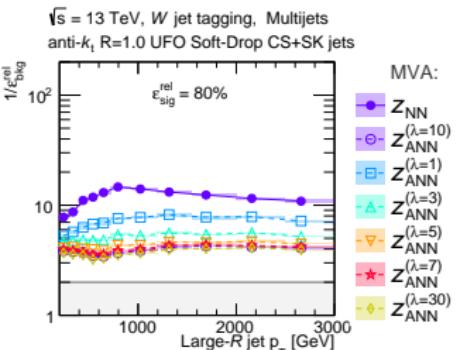
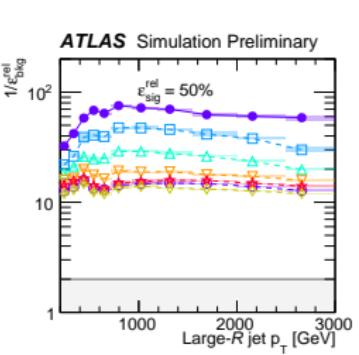
Number of tracks n_{trk}



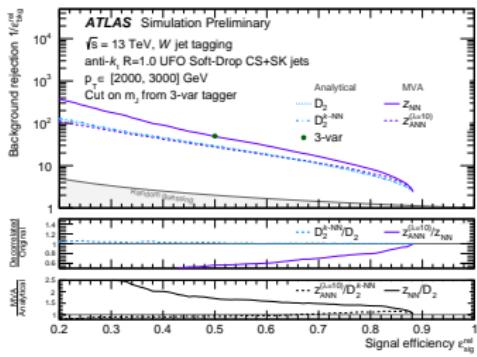
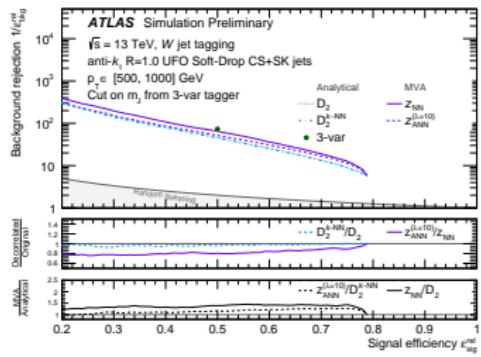
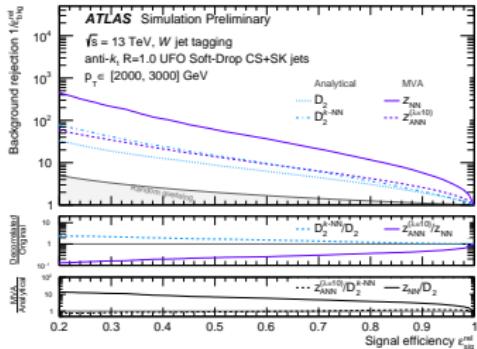
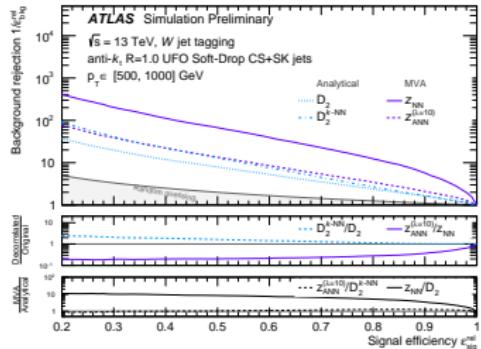
W/Z Tagger: k-NN Method



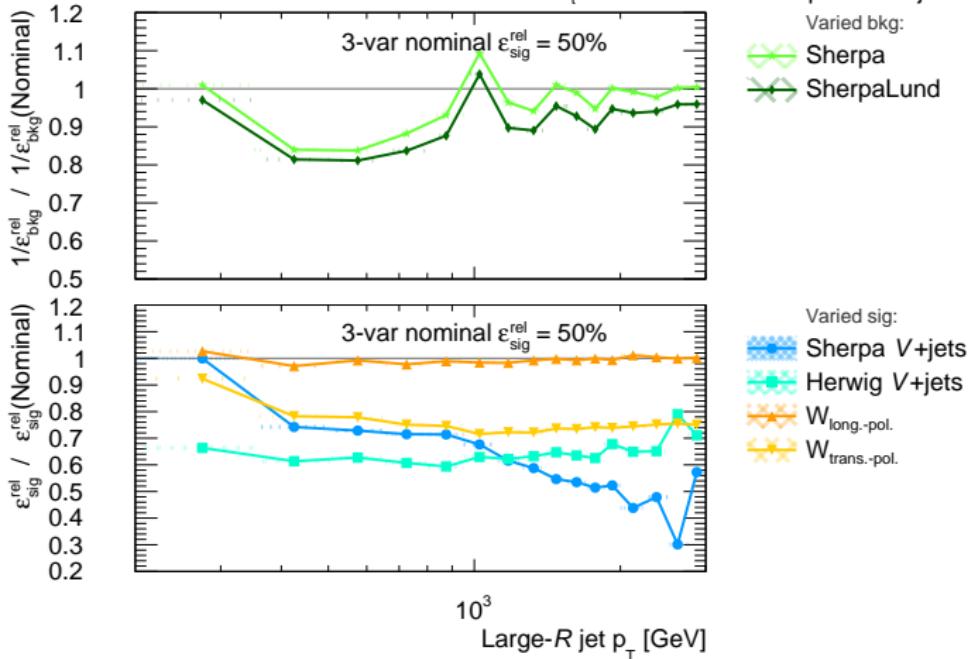
W/Z Tagger: Effect of λ



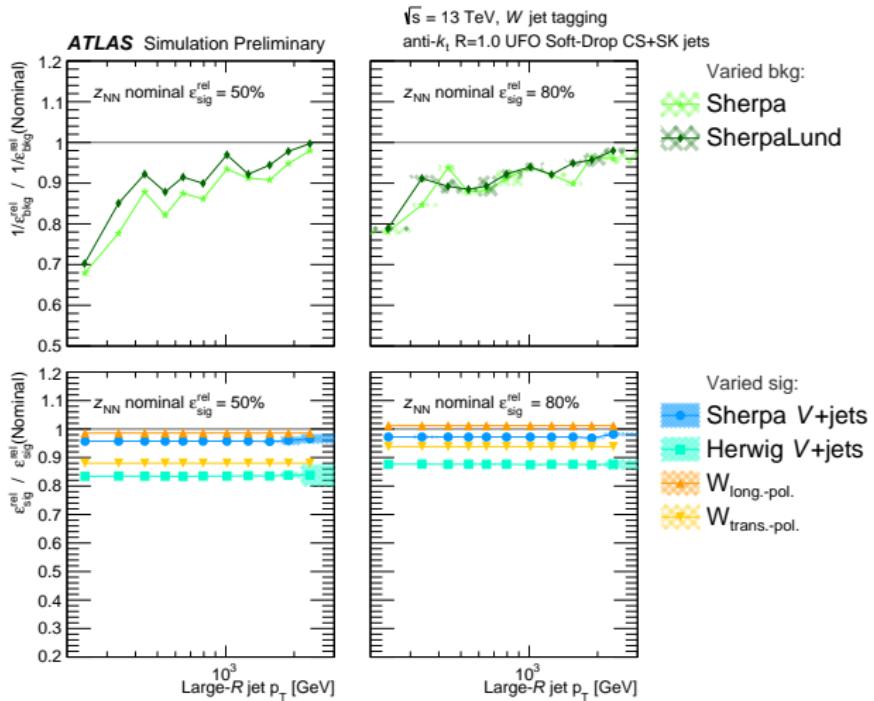
W/Z Tagger: ROC Curves



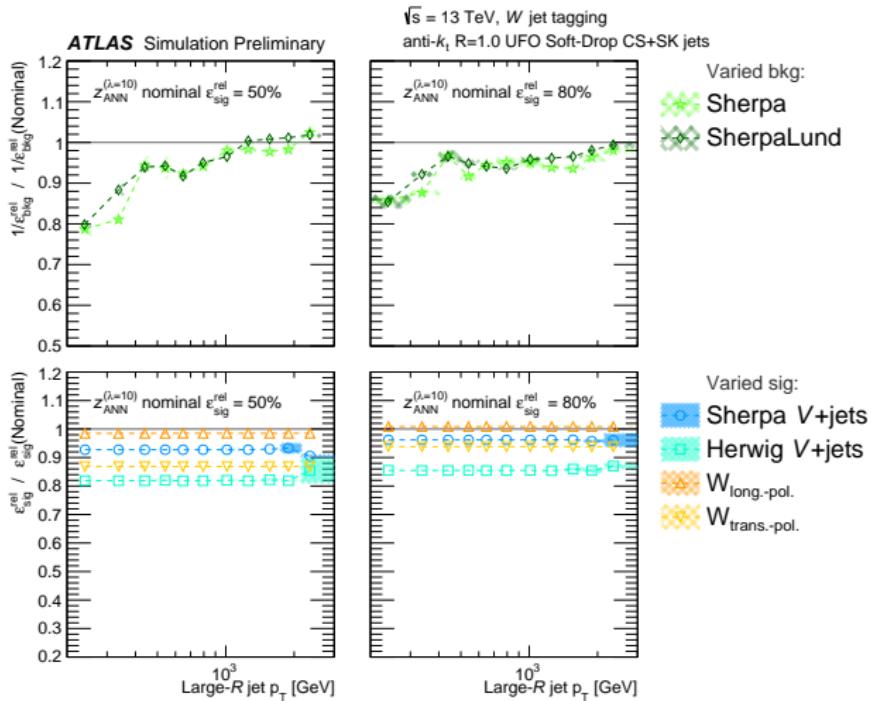
ATLAS Simulation Preliminary $\sqrt{s} = 13 \text{ TeV}$, W jet tagging
 anti- k_t R=1.0 UFO Soft-Drop CS+SK jets



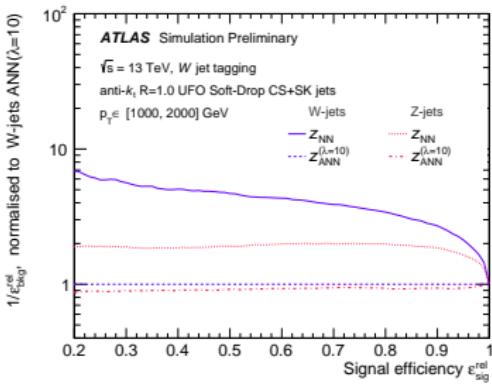
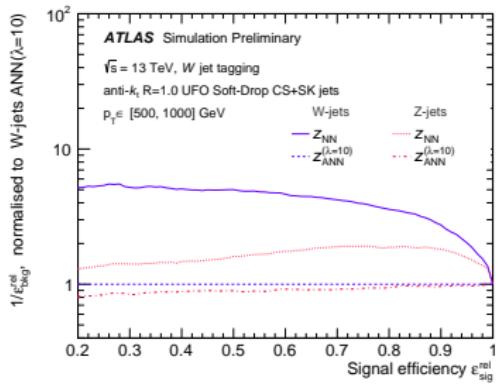
W/Z Tagger: Modelling



W/Z Tagger: Modelling

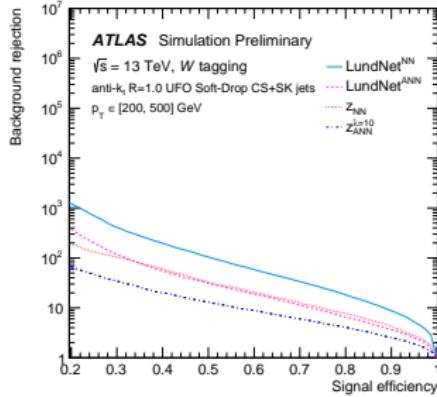


W/Z Tagger: Z vs W

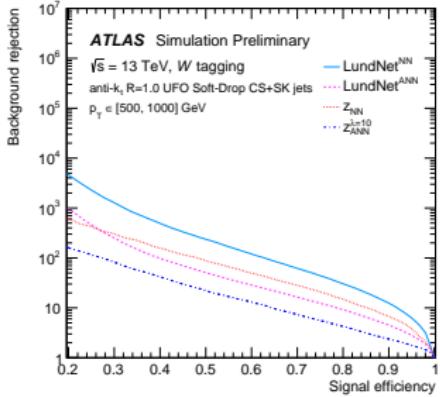


Lund Jet Plane Tagger

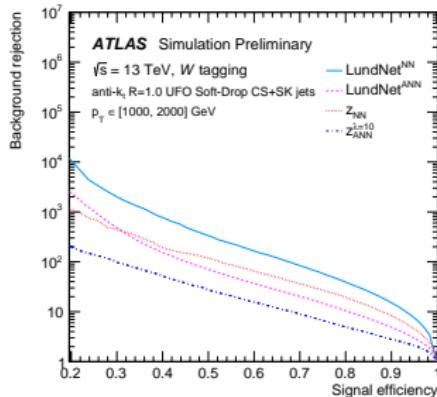
Appendix



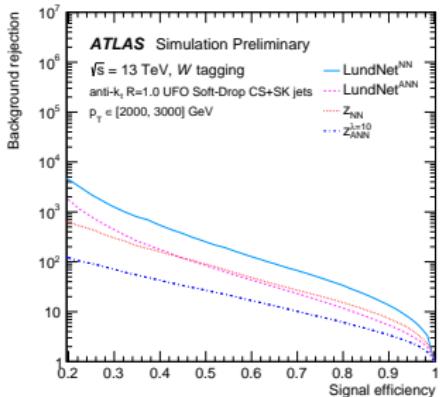
ATL-PHYS-PUB-2023-017



ATL-PHYS-PUB-2023-017

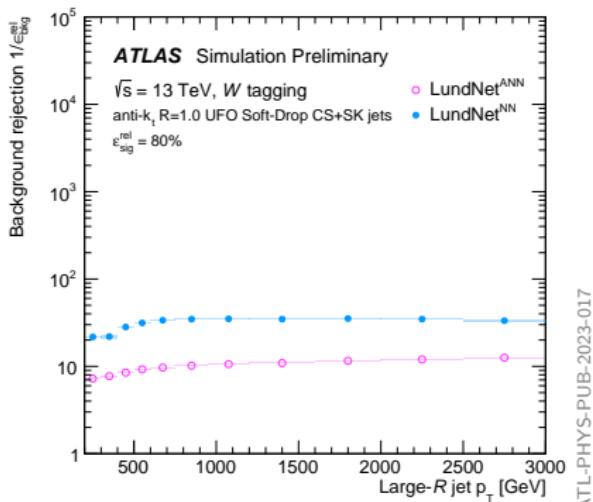


ATL-PHYS-PUB-2023-017



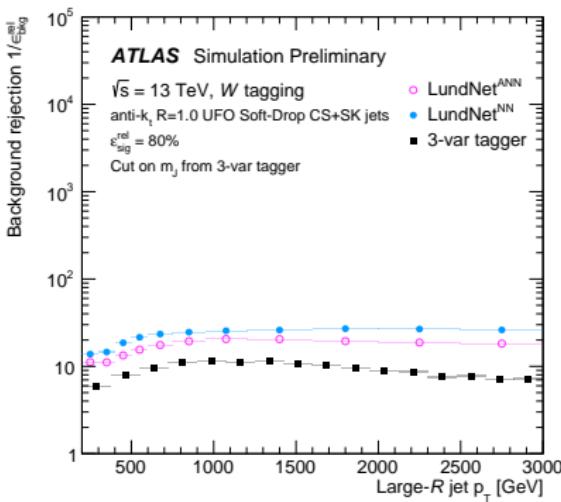
ATL-PHYS-PUB-2023-017

Inclusive



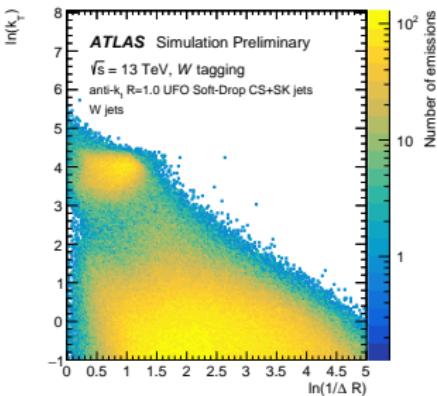
ATL-PHYS-PUB-2023-017

W mass window



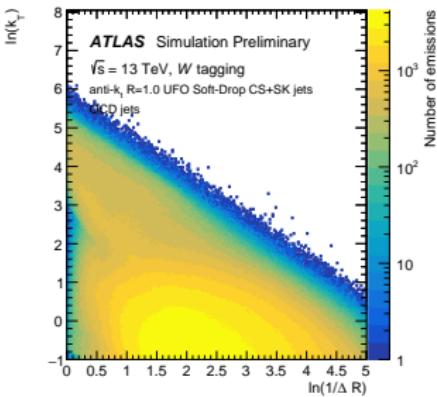
ATL-PHYS-PUB-2023-017

W jets



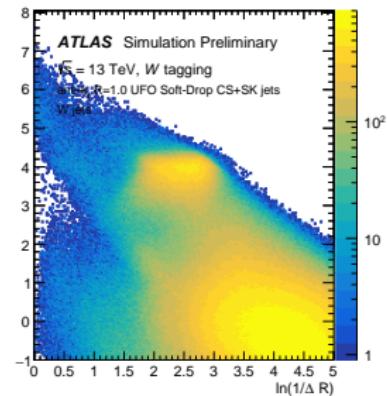
✉ ATL-PHYS-PUB-2023-017

QCD jets

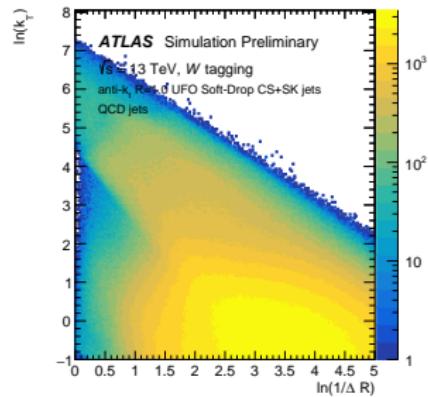


✉ ATL-PHYS-PUB-2023-017

$p_T \in [1000, 3000]$ GeV



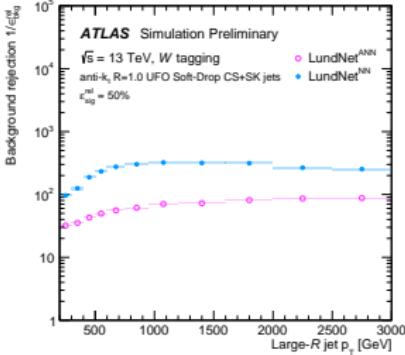
✉ ATL-PHYS-PUB-2023-017



✉ ATL-PHYS-PUB-2023-017

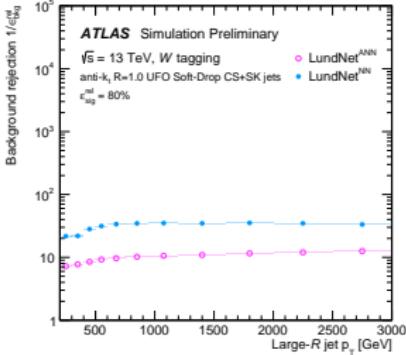
$$\epsilon_{\text{sig}}^{\text{rel}} = 50\% \text{ WP}$$

inclusive



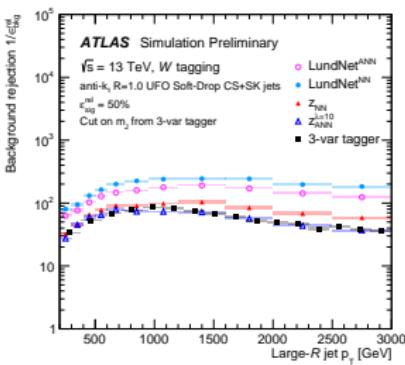
ATL-PHYS-PUB-2023-017

$$\epsilon_{\text{sig}}^{\text{rel}} = 80\% \text{ WP}$$

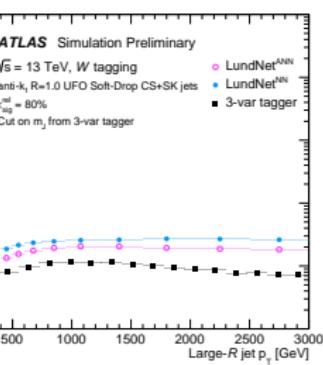


ATL-PHYS-PUB-2023-017

W mass-window



ATL-PHYS-PUB-2023-017

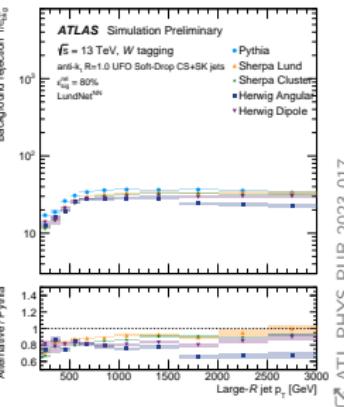
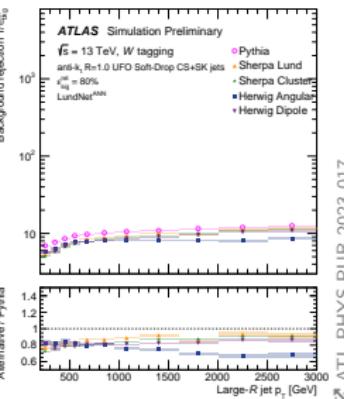
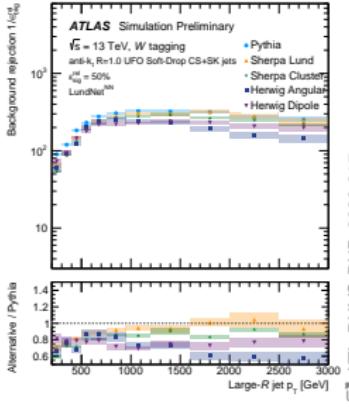
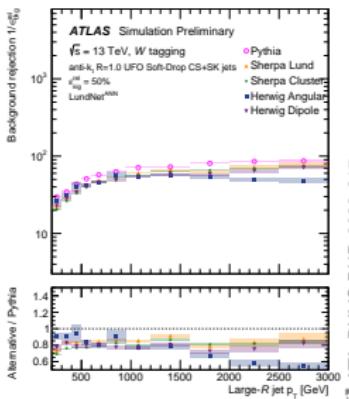


ATL-PHYS-PUB-2023-017

LJP Tagger



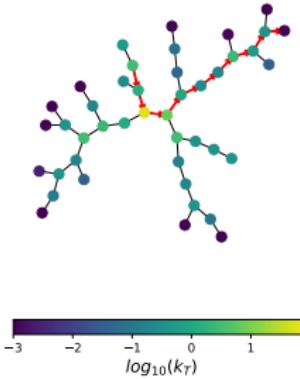
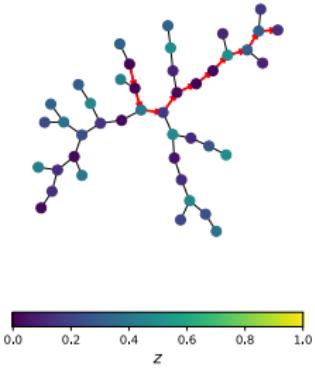
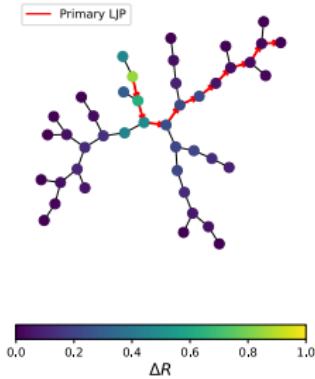
LundNet^{NN}



ATL-PHYS-PUB-2023-017

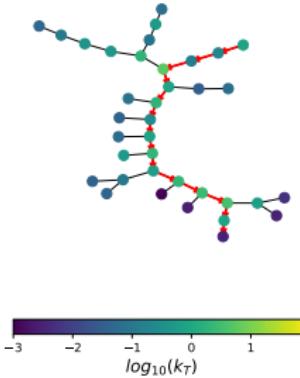
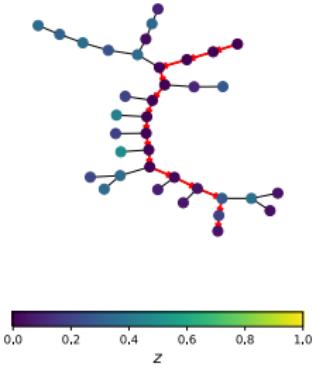
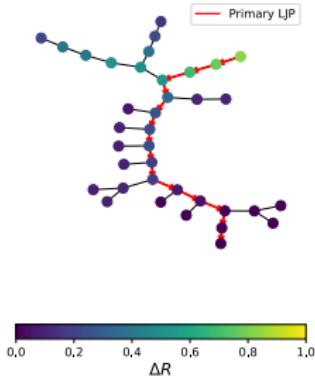
LJP Tagger: Full Declustering History

W jet



ATL-PHYS-PUB-2023-017

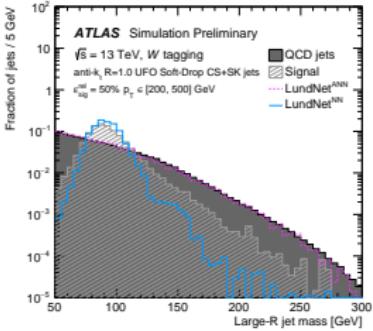
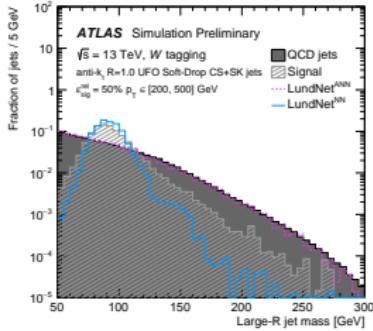
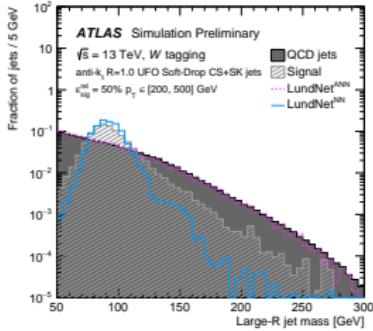
QCD jet



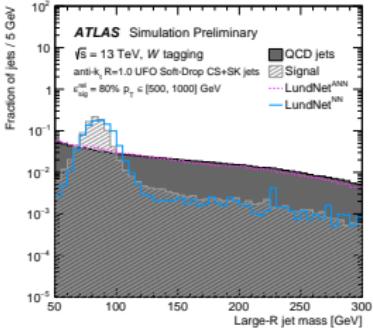
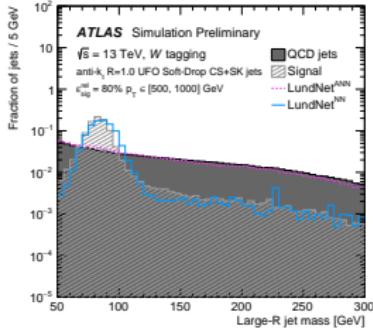
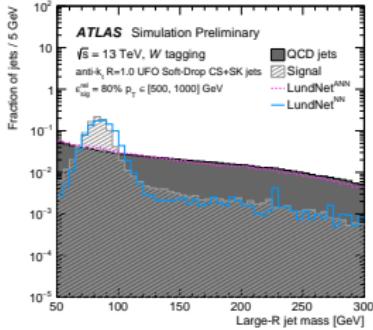
ATL-PHYS-PUB-2023-017

LJP Tagger: Mass Decorrelation

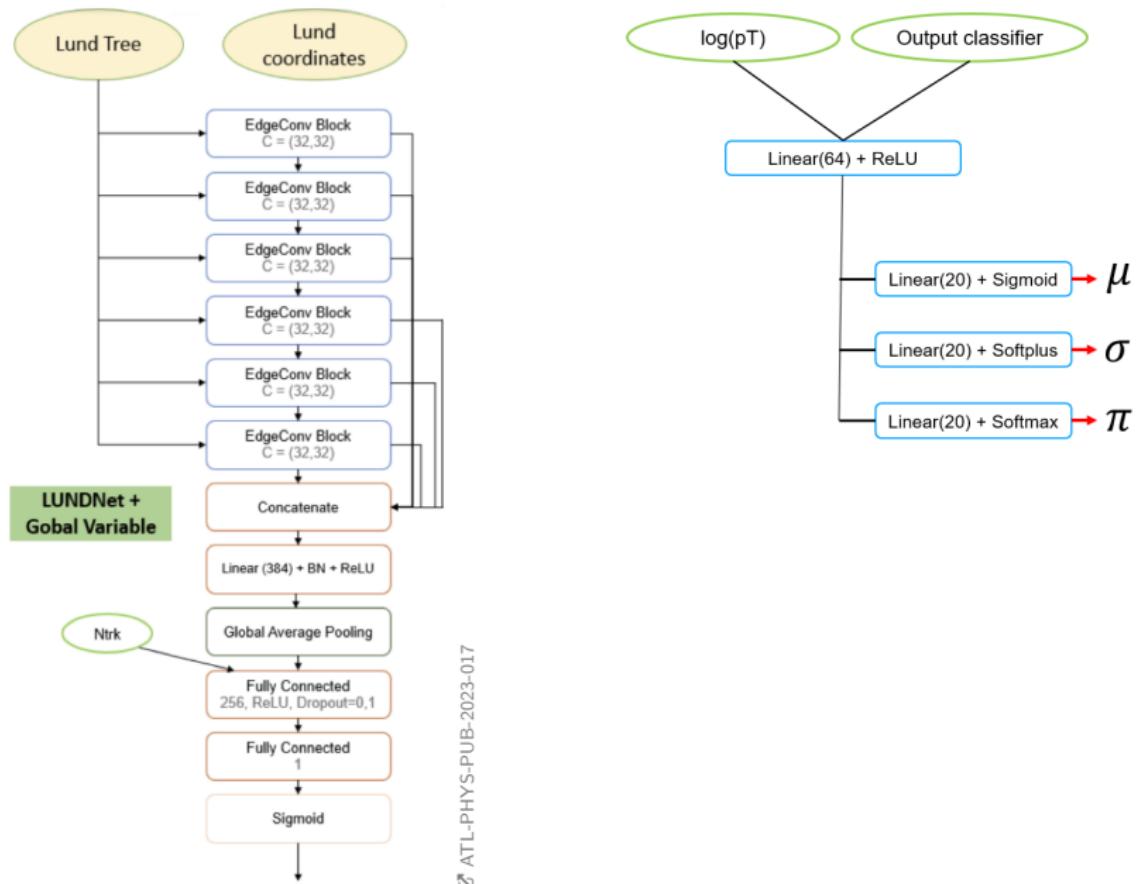
$$\epsilon_{\text{sig}}^{\text{rel}} = 50\% \text{ WP}$$



$$\epsilon_{\text{sig}}^{\text{rel}} = 80\% \text{ WP}$$



LJP Tagger: Architecture



Substructure Variables

Appendix

| W/Z tagger (NN/ANN) | | Top tagger (DNN) | |
|---------------------|---------------------------|--|-----------------------------------|
| D_2 , C_2 | Energy correlation ratios | τ_1 , τ_2 , τ_3 , τ_4 | N -subjettiness |
| τ_{21} | N -subjettiness | $\sqrt{d_{12}}$, $\sqrt{d_{23}}$ | Splitting scales |
| R_2^{FW} | Fox-Wolfram moment | ECF_1 , ECF_2 , ECF_3 | Energy correlation (EC) functions |
| \mathcal{P} | Planar flow | C_2 , D_2 | EC ratios |
| a_3 | Angularity | L_2 , L_3 | Generalised EC ratios |
| A | Aplanarity | Q_W | Invariant mass / virtuality |
| Z_{cut} | Z -Splitting scales | T_M | Thrust major |
| $\sqrt{d_{12}}$ | d -Splitting scales | | |
| $Kt\Delta R$ | k_t -subjjet ΔR | | |
| n_{trk} | number of tracks | | |

arxiv.org/abs/1305.0007

$$\text{ECF}(N, \beta) = \sum_{i_1 < i_2 < \dots < i_N \in J} \left(\prod_{a=1}^N p_{T,i_a} \right) \left(\prod_{b=1}^{N-1} \prod_{c=b+1}^N R_{i_b i_c} \right)^\beta$$

N constituents i of the jet J with Euclidian distance:

$$R_{i_b i_c} = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$$

- IRC (infrared & collinear) safe $\forall \beta > 0$
 - Goes to $\rightarrow 0$ in infrared/collinear limit

Here: ECF₁, ECF₂, ECF₃;

Energy Correlation Ratios C_2 , D_2

[doi.org/10.1007/JHEP12\(2014\)009](https://doi.org/10.1007/JHEP12(2014)009)

Normalised ECFs e_n^β :

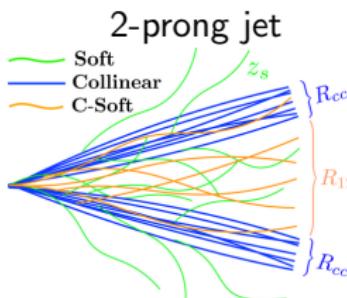
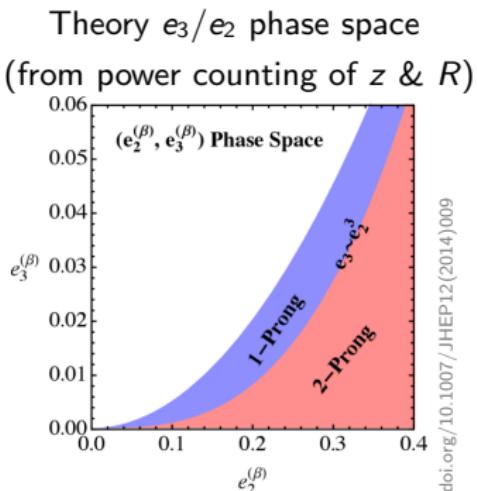
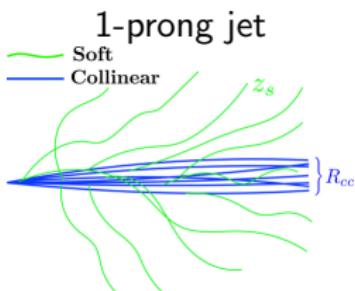
$$e_n^\beta = \frac{\text{ECF}(n, \beta)}{\text{ECF}(1, \beta)^n}; \quad z_i = \frac{p_{T_i}}{p_{T_J}}$$

$$\Rightarrow e_2^\beta = \sum_{1 \leq i \leq j \leq n_J} z_i z_j R_{ij}^\beta$$

$$\Rightarrow e_3^\beta = \sum_{1 \leq i \leq j \leq k \leq n_J} z_i z_j z_k R_{ij}^\beta R_{ik}^\beta R_{jk}^\beta$$

Ratios of e_n^β :

$$C_2 = \frac{e_3^\beta}{\left(e_2^\beta\right)^2}, \quad D_2 = \frac{e_3^\beta}{\left(e_2^\beta\right)^3}$$



→ C_2 and D_2 Separate 1- and 2-prong jets on the e_3/e_2 plane

N-Subjettiness τ_N

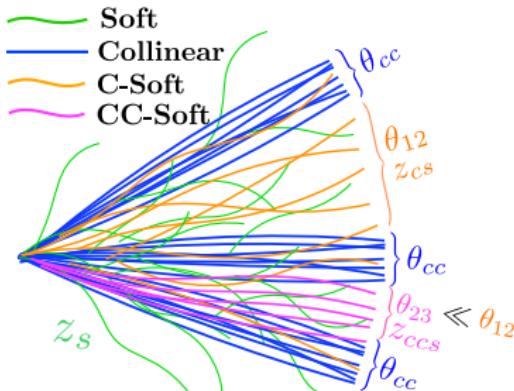
<https://arxiv.org/abs/1011.2268>

$$\tau_N^\beta = \sum_{1 \leq i \leq n_J} z_i \min \left\{ R_{i1}^\beta, \dots, R_{iN}^\beta \right\}$$

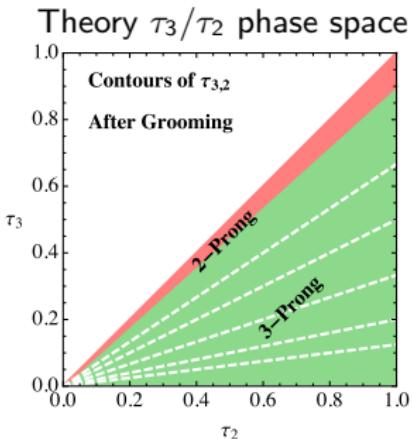
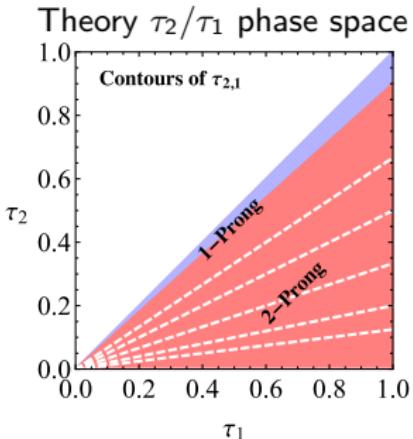
$$\text{with } z_i = \frac{p_{T_i}}{p_{T_J}}$$

$$\rightarrow \tau_{21}^\beta = \frac{\tau_2^\beta}{\tau_1^\beta}$$

3-prong jet



arxiv.org/abs/1609.07483



arxiv.org/abs/1609.07483

$${}_\nu e_N^\beta = \sum_{i_1 < i_2 < \dots < i_N \in J} \left(\prod_{a=1}^N p_{T_{i_a}} \right) \left(\prod_{m=1}^{\nu} \min_{s < t \in \{i_1, i_2, \dots, i_N\}} R_{st} \right)^\beta$$

Where $\min_X^{(m)}$ denotes the m th smallest element in the set X

Reduces to nominal ECF in the case $\nu = \binom{N}{2}$:

$$\text{ECF}(N, \beta) = \sum_{i_1 < i_2 < \dots < i_N \in J} \left(\prod_{a=1}^N p_{T_{i_a}} \right) \left(\prod_{b=1}^{N-1} \prod_{c=b+1}^N R_{i_b i_c} \right)^\beta$$

→ ${}_\nu e_N^\beta$ are sensitive to hierarchy of distinct angular (R) scales m in jet

- ECF average over them

Ratios to separate 2- & 3-prong jets: $L_2 = \frac{{}_3 e_3^{\beta=1}}{{}_1 e_2^{\beta=2}}^{\frac{3}{2}}$, $L_3 = \frac{{}_1 e_3^{\beta=1}}{{}_3 e_3^{\beta=1}}^{\frac{1}{3}}$

Number of Ghost-Associated Tracks n_{trk}



JHEP04(2008)005

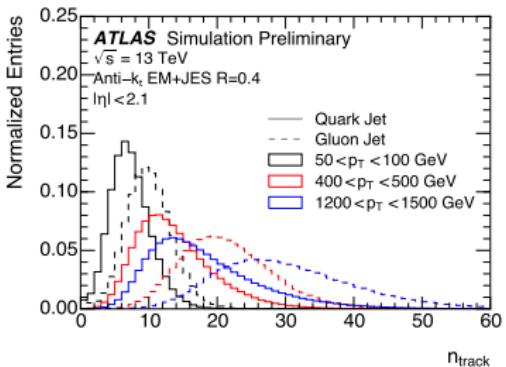
n_{trk} : number of tracks

- With $p_T > 500 \text{ MeV}$
 - Ghost-associated to jet
- Powerful q/g discriminant

Ghost-associated jet area

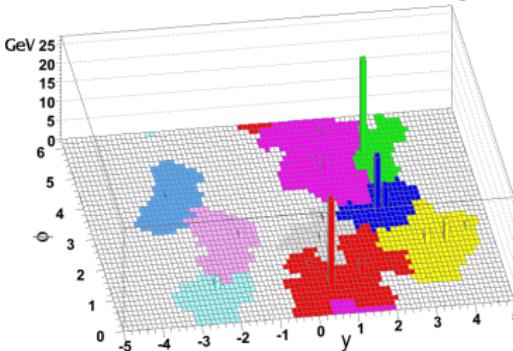
- Add dense coverage of 'infinitely' soft 'ghost' constituents
- Count how many are clustered within the jet

n_{trk} as q/g discriminant



ATL-PHYS-PUB-2017-009

Ghost associated areas of k_t jets



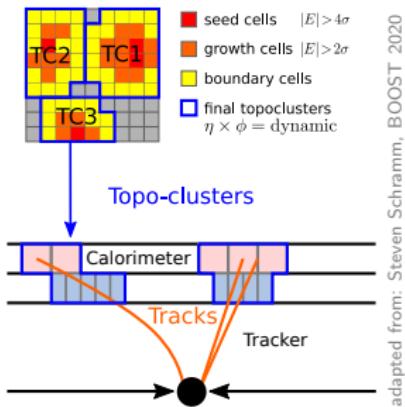
JHEP04(2008)005

UFO Jets

Appendix

Calorimeter only:

- **LCTopo**: Topological calorimeter clusters



Combined with tracking:

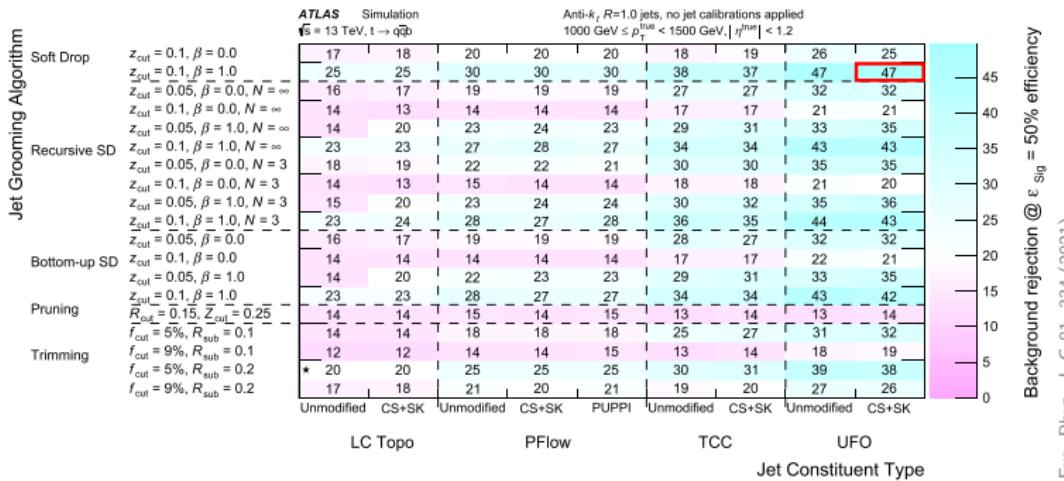
- **PFlow**: Particle Flow Objects
 - Low p_T : Use **track** 4-vector for charged particles, subtract energy from **cluster** 4-vectors
 - High p_T : Use **cluster** 4-vectors, ignore **tracks**
- **TCC**: Track Calo Clusters
 - Low p_T : Use **cluster** 4-vectors, ignore **tracks**
 - High p_T : Split **clusters** using **tracks**, get energy from **clusters** but angles from **tracks**

Combining PFlow and TCC:

- **UFO** combines **TCC** and **PFlow** to achieve optimal performance over a broad kinematic (p_T) range

Background rejection for various pileup mitigations and groomings:

Here: 2-variable top tagger, high- p_T range
(plots for W and low- p_T in backup)



Best background rejection with:

- $R = 1.0$ anti- k_T **UFO** jets
- Pileup Mitigation: Constituent Subtraction + SoftKiller (**CS+SK**)
- Grooming: Soft Drop (**SD**) with $\beta = 1.0$ $z_{\text{cut}} = 0.1$

Other factors: Good pileup stability, mass resolution, ...