Comparative performance of ATLAS boosted W taggers using different AI/ML algorithms

M. LeBlanc (Manchester), on behalf of the ATLAS Collaboration
@TopPhysicist, matt.leblanc@cern.ch

BOOST 2023 (31 July - 4 August), Lawrence Berkeley National Lab, California, USA

MANCHESTER 1824
The University of Manchester

ATLAS EXPERIMENT

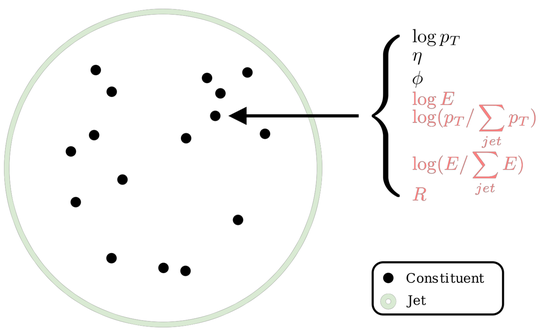
Introduction

There are many ways to represent the information in a jet, *e.g.*:

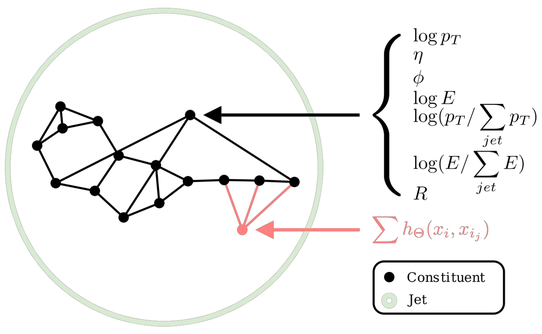
- A set of **precomputed features** (e.g. JSS observables)
- An **unordered set of particles** ('point-cloud')
- A **structured graph**, based on the constituent kinematics, or the **jet clustering history** (e.g. Lund jet plane, 'LJP')

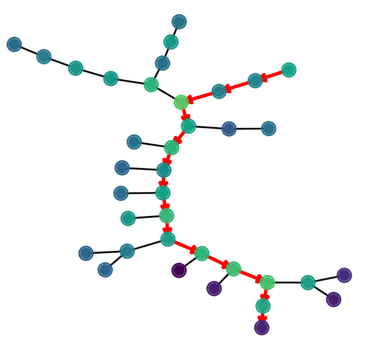
What are the advantages and disadvantages of each for W tagging? What gives the best performance?

Point Cloud (EFN/PFN)



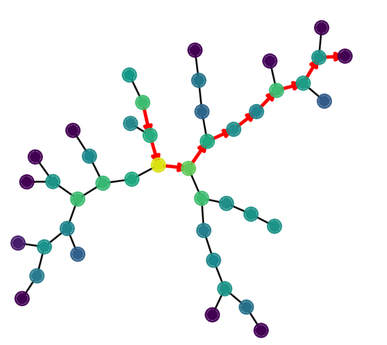
Graph (ParticleNet)



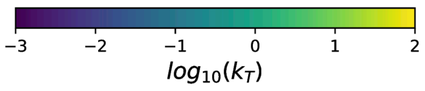


LundNet graphs

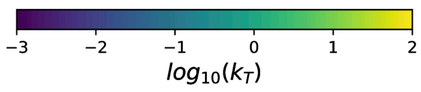
Constructed using information from the jet's angle-ordered (C/A) clustering history



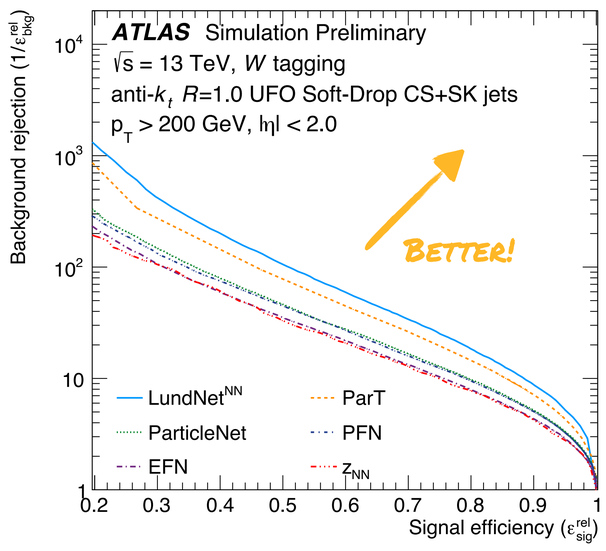
BACKGROUND (Q/G)

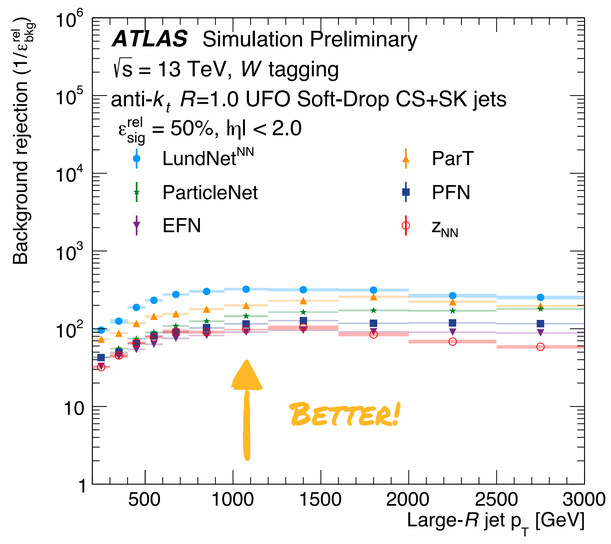


SIGNAL (W)



Classification Performance

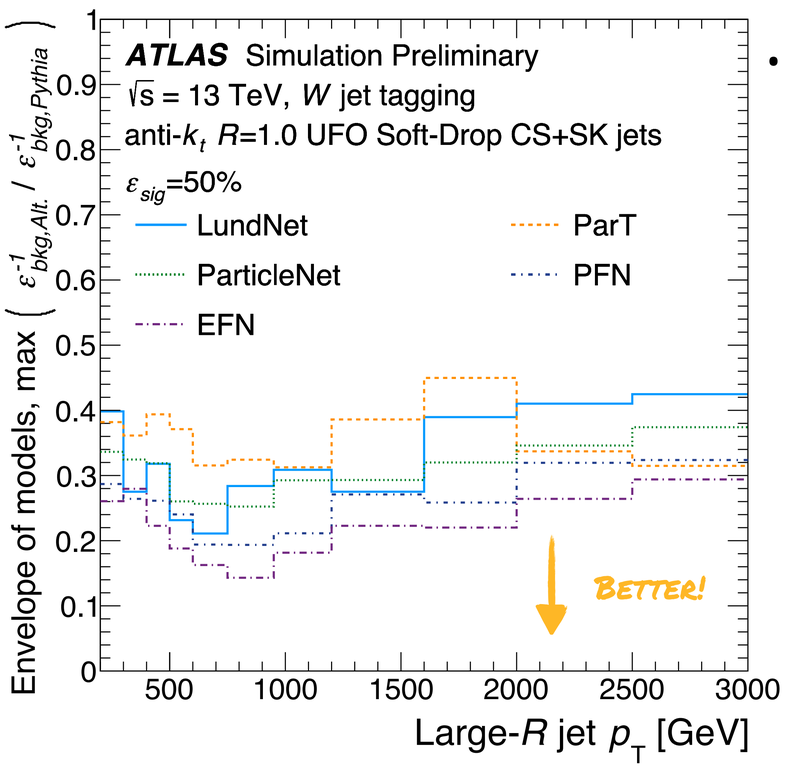




- Complex architectures often result in better classification performance.
  - Full ATLAS detector simulation is used: **R=1.0 CS+SK Soft-Drop (β=1, z_{cut}=0.1) UFO jets**
  - Consistent picture as a function of p_T
- **LJP network outperforms others** despite having fewer trainable parameters ~ O(10^5).
  - **Physics-driven pre-processing** provides clean picture for ML.
- Similar conclusions when studying the background rejection at a fixed tagging efficiency.

Model	AUC	ACC	ε_{bkg}^{-1} @ ε_{sig} = 0.5	ε_{bkg}^{-1} @ ε_{sig} = 0.8	# Params	Inference Time
EFN	0.920	0.835	35.1	7.95	56.73k	0.065 ms
PFN	0.931	0.853	44.7	9.50	57.13k	0.11 ms
ParticleNet	0.933	0.826	46.2	9.76	366.16k	0.36 ms
ParticleTransformer	0.951	0.880	77.9	14.6	2.14M	0.28 ms

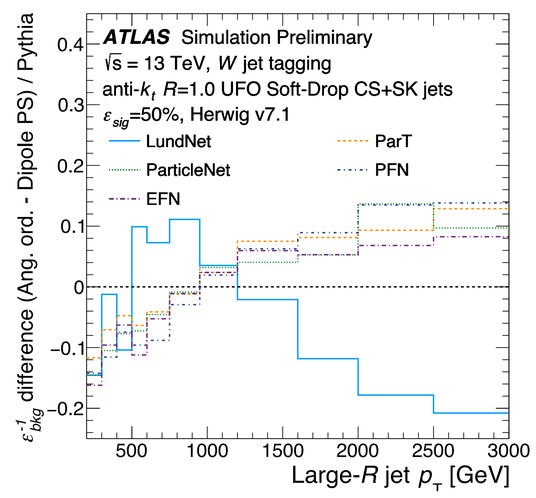
Model-dependence



- While more complicated taggers are more accurate, **they can also be more model-dependent**!
- Important to understand the origin of model-dependence, or else **improved classification performance may not generalise to data!**
  - Large data-to-MC scale factors, uncertainties can erode ML gains.

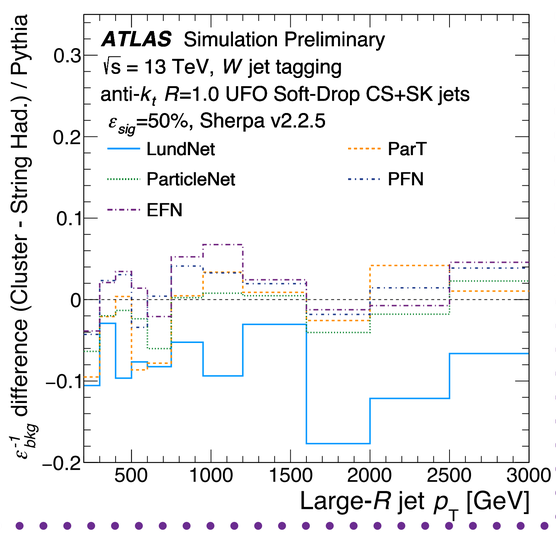
Parton Shower Modelling

- Varying the Parton Shower algorithm in Herwig 7 MC samples causes large variations.
- W-tagging algorithms sensitive to modelling of **perturbative emissions**.
  - **LundNet differences anti-correlated** from others.



Hadronisation Modelling

- AHADIC Cluster *vs.* Lund string hadronisation in Sherpa v2.2.5
  - **Most W-taggers not sensitive to soft emissions.**
- **LundNet tagger more sensitive:** does the LJP picture over-emphasise soft emissions?



ATLAS, *Tagging boosted W bosons with the Lund jet plane in ATLAS* ATL-PHYS-PUB-2023-017
ATLAS, *Constituent-Based W-boson Tagging with the ATLAS Detector* ATL-PHYS-PUB-2023-020