Improving ATLAS Hadronic Object Performance with ML/AI Algorithms

Lake Louise Winter Institute 2023

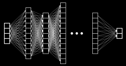
Tobias Fitschen

20 Feb 2023

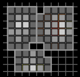
University of Manchester

Funded by the European Union

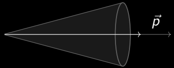
European Research Council
Established by the European Commission



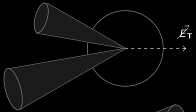
Neural Networks



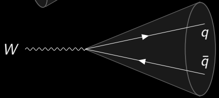
Calo Clusters



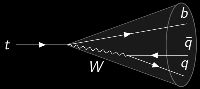
Jet Energy Scale



Missing E_T

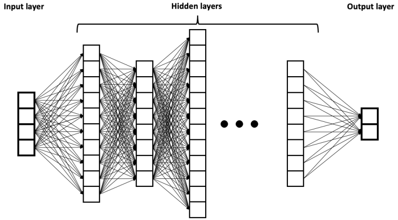


W/Z Tagging



Top Tagging

Multilayer Perceptron (MLP)



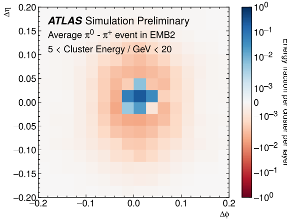
wikimedia commons

- 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

$\pi^0/\pi^\pm$ calorimeter shower as image

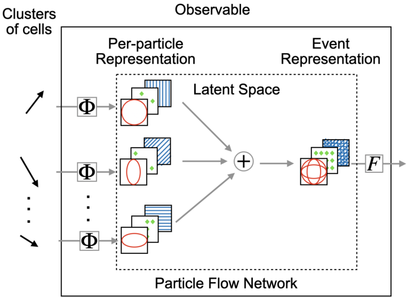


ATL-PHYS-PUB-2020-018

Deep Sets

(Energy/Particle Flow Network)

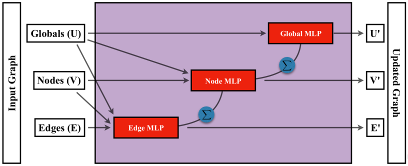
- Unorderd, variable length input
- E.g. of jet constituent momenta
- Permutation invariant



https://arxiv.org/abs/1810.05165

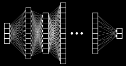
Graph Neural Net (GNN)

More info in previous talk by Alexander Khanov

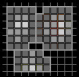


ATL-PHYS-PUB-2022-040

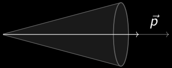
- Structured representation: nodes V, edges E
- Pairwise message passing
  - Nodes iteratively updated by exchanging information with neighbors
- Permutation invariant
- E.g. neighboring calo cells connected via edges in GNN



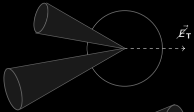
Neural Networks



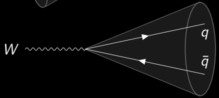
Calo Clusters



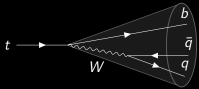
Jet Energy Scale



Missing E_T



W/Z Tagging



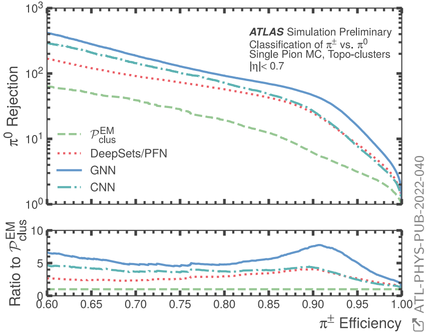
Top Tagging

π^0 vs $\pi^\pm$ 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 $\pi^\pm$ classification performance



Point cloud of energy deposits in calorimeter cells

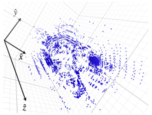
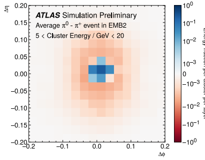


Image: $\pi^0 - \pi^\pm$ difference



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

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

Individual calorimeter cell signals

→ As point clouds (GNN, PFN)

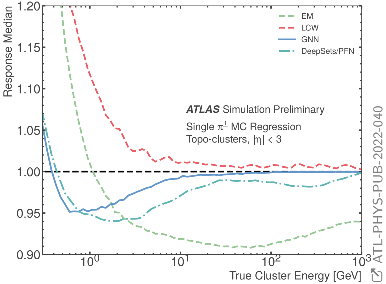
→ Or projected on images (CNN)

Observations

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

Energy Regression

$\pi^{\pm}$ cluster energy response



Second step: Energy Calibration

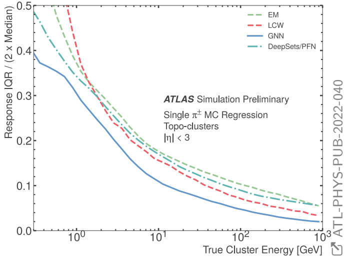
Here: charged ($\pi^{\pm}$) clusters

Observations

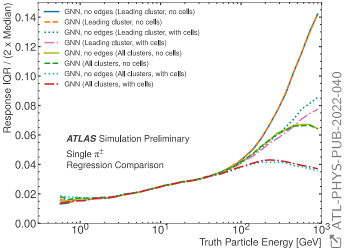
- GNN performs best wrt. response and width
- Followed by Deep Sets
- Similar for neutral π^{0} clusters (see backup)

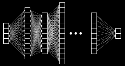
Interquantile range IQR (measure for spread)

Calorimeter clusters only

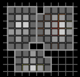


Including tracking information

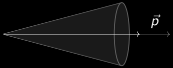




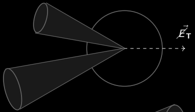
Neural Networks



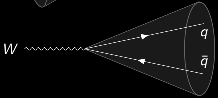
Calo Clusters



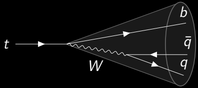
Jet Energy Scale



Missing E_T

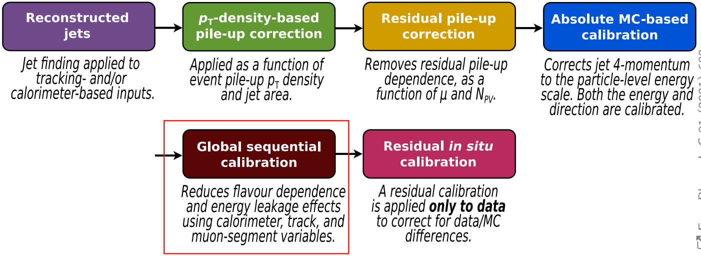


W/Z Tagging



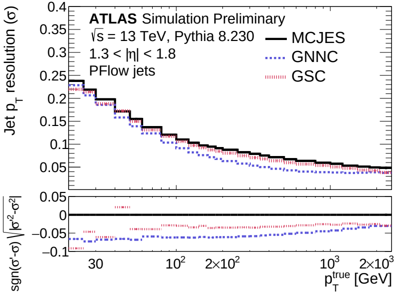
Top Tagging

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



Eur. Phys. J. C 81 (2021) 689

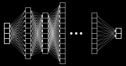
After energy scale calibrated on average, GSC corrects for small differences for different jet flavours



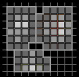
ATL-JETM-2022-004

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

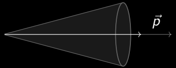
→ Improvement over full p_T range



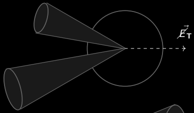
Neural Networks



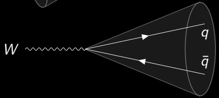
Calo Clusters



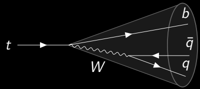
Jet Energy Scale



Missing E_T



W/Z Tagging



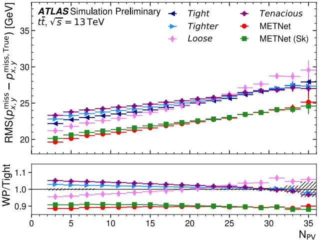
Top Tagging

ATLAS EXPERIMENT

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

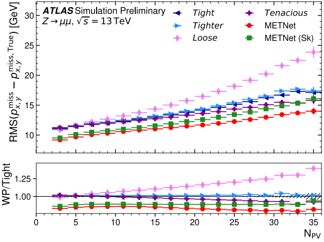
- p_T^{miss} in ATLAS: Negative sum of calibrated momenta of hard objects (e, μ, τ-jets, γ, jets)
- Plus soft term: Tracks from PV not associate to hard objects
- Different working points (WPs) defined for various pileup conditions
  - E.g. "tight": Higher p_T cuts on forward jets
- MetNet: MLP combining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

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



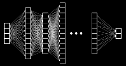
ATL-PHYS-PUB-2021-025

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

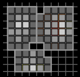


ATL-PHYS-PUB-2021-025

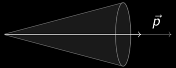
$\rightarrow$ E_T^{miss} definition depends on process but MetNet performs best for all



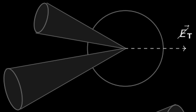
Neural Networks



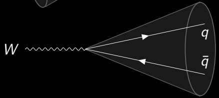
Calo Clusters



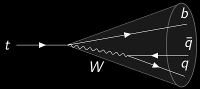
Jet Energy Scale



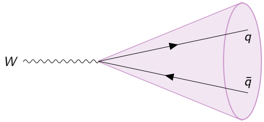
Missing E_T

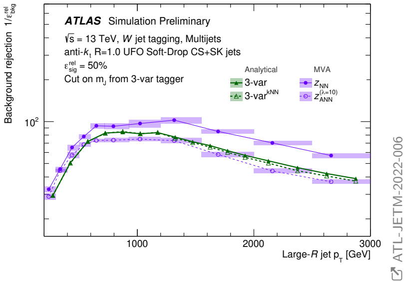


W/Z Tagging



Top Tagging





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
τ_{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-subjet ΔR
N_{trk}	Number of associated tracks

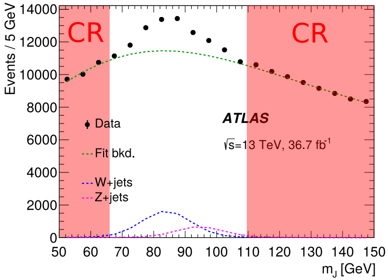
(see backup for definitions)

$\rightarrow$ NN tagger significantly outperformed cut based 3-var tagger

$\rightarrow$ Even mass-decorrelated version ANN shows similar performance to cut based 3-var using m

The Need for Mass Decorrelation

Data-driven background estimates:



adapted from: arxiv.org/abs/1708.04445

- 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

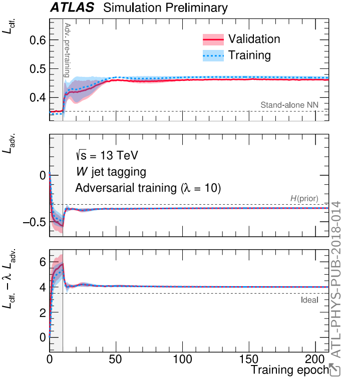
- Classifier MLP for tagging task
- Adversary for decorrelation (e.g. of mass)
  - → Predicts mass based on classifier output (+auxiliary variables)
- Gradient reversal layer: During back-propagation penalise Classifier if Adversary predicts mass too well

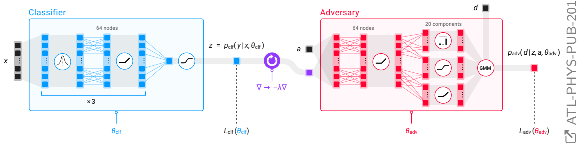
→ Final tagger only consists of Classifier

→ Tagger decision decorrelated to mass

Training schedule

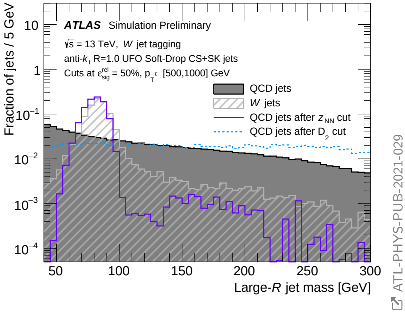
1. Classifier alone
2. Adversary alone
3. Both togeher



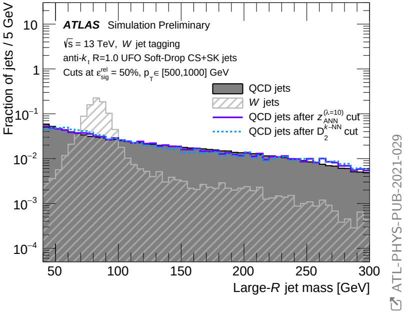


Mass Decorrelation with Adversarial Training

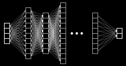
NN: Correlated to m_J



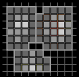
ANN: Active decorrelation



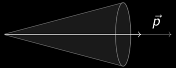
- Background mass distribution shaped according to signal by NN
- Adversarial Neural Network (ANN) successfully decorrelates



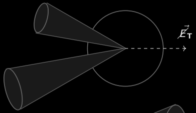
Neural Networks



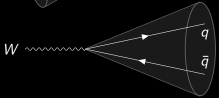
Calo Clusters



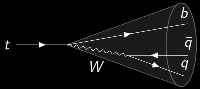
Jet Energy Scale



Missing E_T



W/Z Tagging



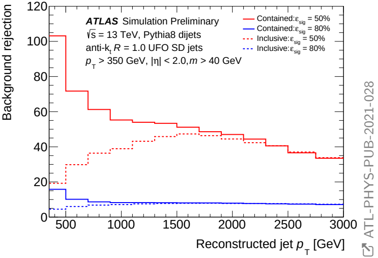
Top Tagging

2 taggers: Inclusive and contained

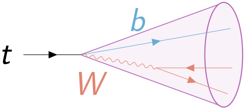
- Fixed working points: 50% and 80%
  - Defined as function of p_T
- DNN features optimised for UFO jets:

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

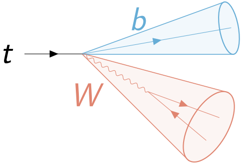
(see backup for definitions)



Contained



Non-contained



Constituent-Based Top Tagger

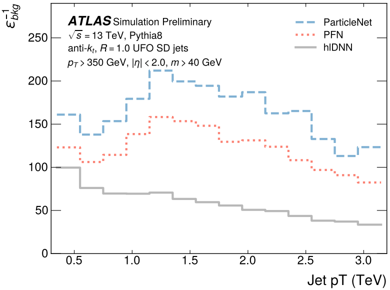
ATL-PHYS-PUB-2022-39: Constituent-Based Top-Quark Tagging

DNN top tagger (prev. slides):

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

Constituent-based taggers:

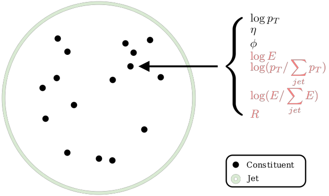
- Low-level features based on 4-vectors of jet constituents



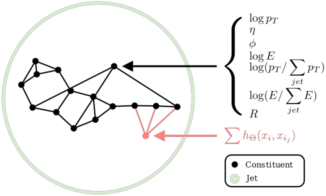
ATL-PHYS-PUB-2022-39

→ **Up to ×2 improvement over baseline (hlDNN)!**

PFN: Particle Flow Network



ParticleNet: Dynamic Graph-CNN



courtesy Kevin Greif

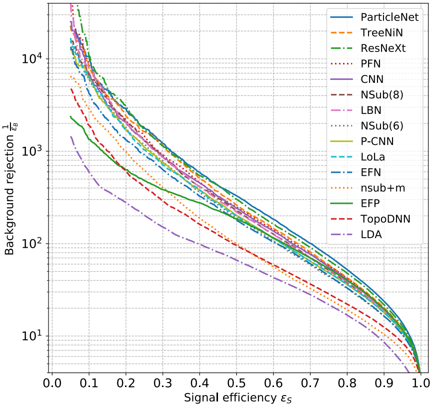
Based on: The Machine Learning landscape of top taggers

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



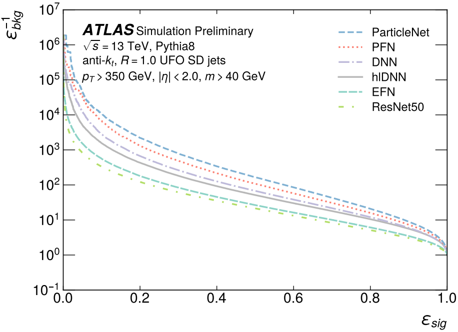
doi:10.21468/SciPostPhys.7.1.014

→ **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?



ATL-PHYS-PUB-2022-39

- **hlDNN**: Baseline similar to DNN top tagger
  → ATL-PHYS-PUB-2021-028
- **DNN**: Using constituent 4-momenta
  → arxiv.org/abs/1704.02124
- **EFN/PFN**: Energy/Particle-flow networks
  → arxiv.org/abs/1810.05165
- **ResNet50**: CNN using jet images
  → arxiv.org/abs/1512.03385
- **ParticleNet**: Dynamic Graph-CNN
  → arxiv.org/abs/1902.08570

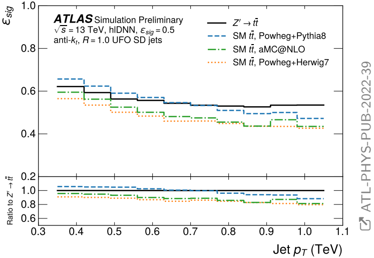
- ParticleNet and PFN show best performance
- ResNet50 & EFN underperform → Do not translate well from Delphes study

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

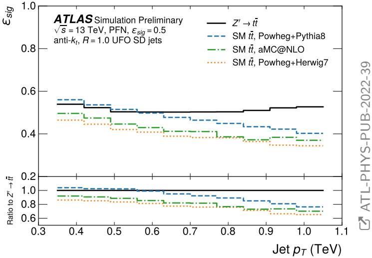
Model dependence

- Different parton shower and hadronisation models
- ϵ^{sig} measured at threshold for $\epsilon^{\text{sig}} = 50\%$ in nominal sample
- PFN and ParticleNet: Slightly more model dependent than baseline hlDNN

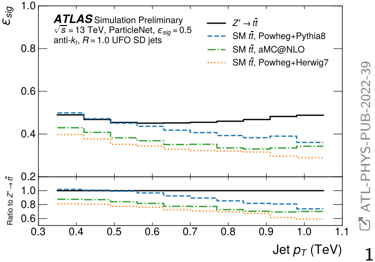
baseline: hlDNN



PFN (Particle Flow Network)



ParticleNet



Summary

Summary

Many ML applications for hadronic objects in ATLAS

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

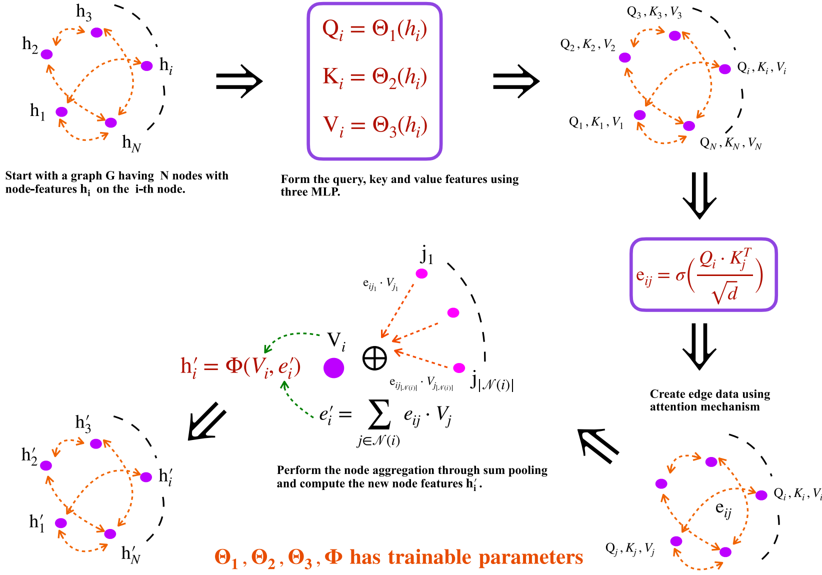
Constituent based methods perform best in all domains

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

Appendix

ML for Calo Clusters

Transformer for Graph Updates

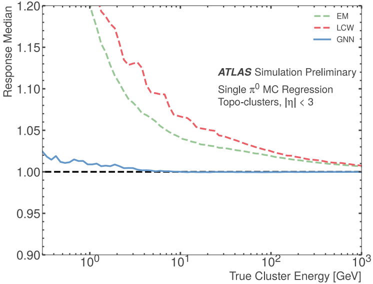


ATL-PHYS-PUB-2022-040

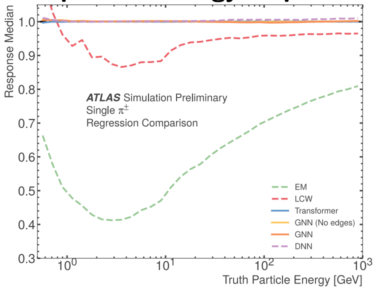
Performance

ATLAS EXPERIMENT

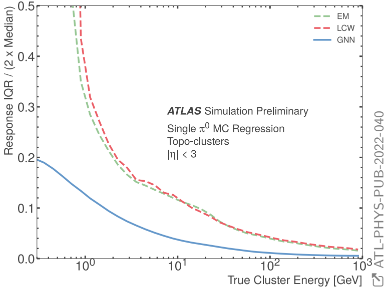
π^0 cluster energy response



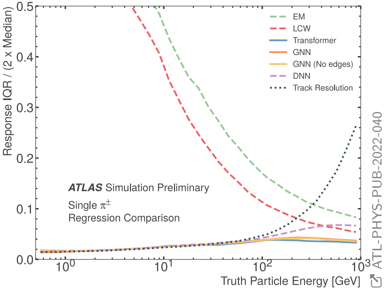
π^0 particle energy response



Interquantile range (width)

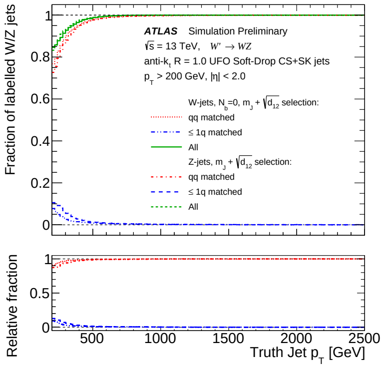


Interquantile range (width)



W/Z Taggers

Truth Labelling



ATL-PHYS-PUB-2021-029

Matching to daughter quarks q, q':

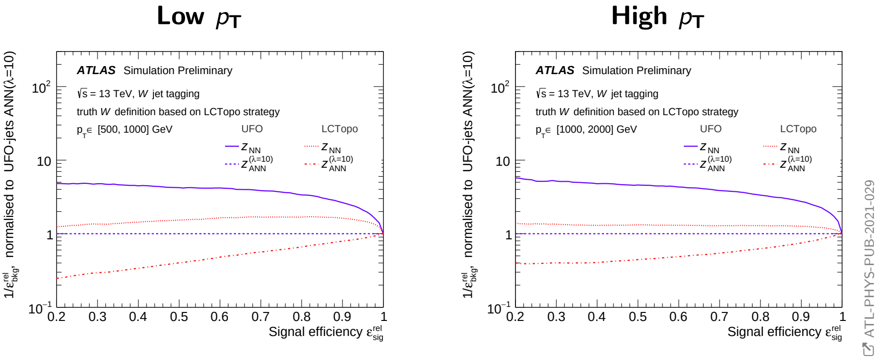
- Fraction of W (Z) containing both q, q' within $\Delta R < 0.75$:

  $> 98\%$ (96%) at $p_T > 300$ GeV (previously: 90%)

Truth jet definition:

- Parton level MC to label jets as:
  - Signal: Containing full $W \to qq'$
  - 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
  - $\sqrt{d_{12}} > 55.25 + e^{-2.35 \frac{p_T}{\text{GeV}}}$ 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

Tagger Performance: UFO vs LCTopo

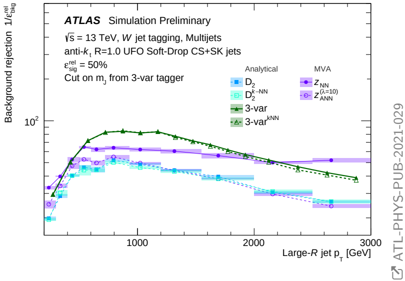


- Bkg rejection with UFO improved by a factor of 2-3 w.r.t. LCTopo!
- Mass decorrelation (ANN) comes with tradeoff of reduced efficiency
    - But may be better option for data-driven background estimates on m_J distribution

NN/ANN Tagger with n_{trk}

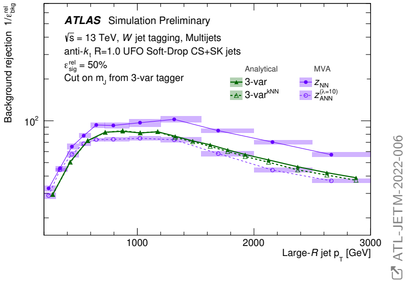
Previously:

ATL-PHYS-PUB-2021-029



Folow up publication:

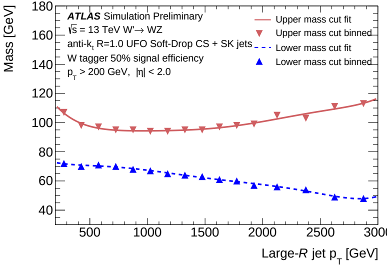
With n_{trk} as additional feature



- **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

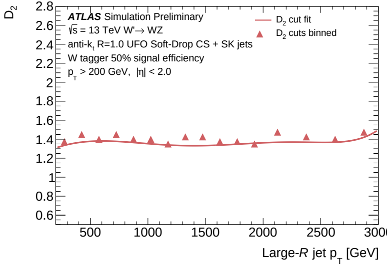
Mass window



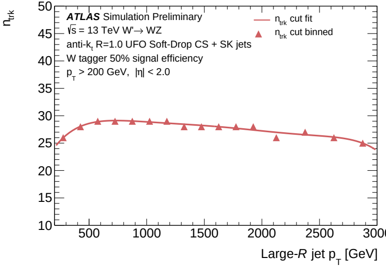
3-variable tagger: W

- 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

Energy correlation ratio D_2



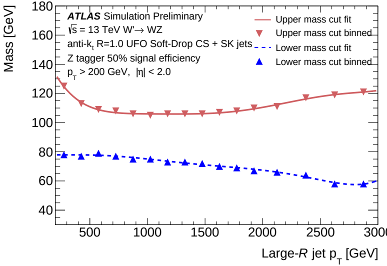
Number of tracks n_{trk}



ATL-PHYS-PUB-2021-029

3-Variable Tagger

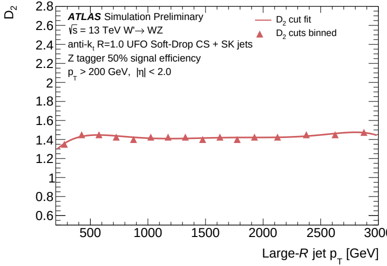
Mass window



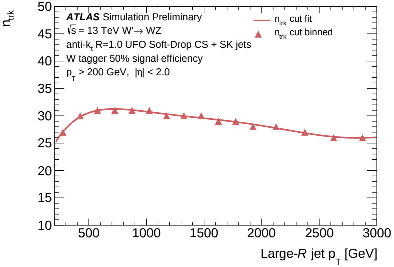
3-variable tagger: Z

- 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

Energy correlation ratio D_2

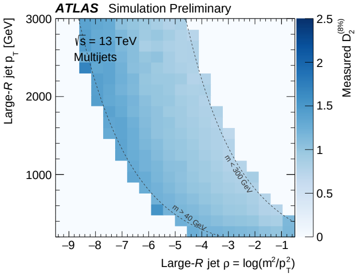


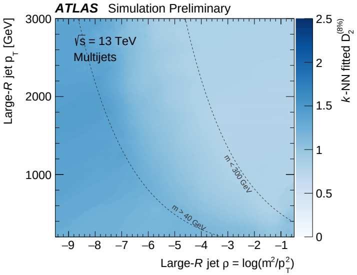
Number of tracks n_{trk}



ATL-PHYS-PUB-2021-029

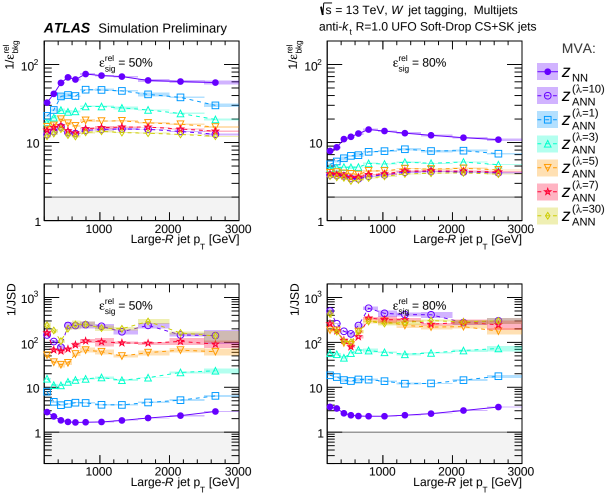
W/Z Tagger: k-NN Method





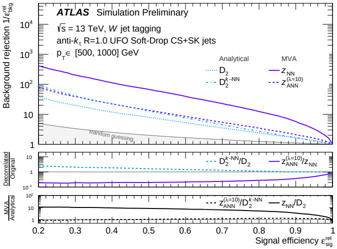
ATL-PHYS-PUB-2021-029

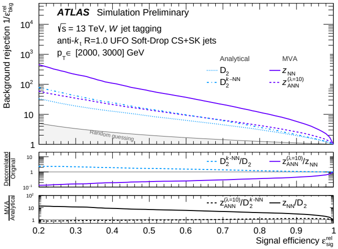
W/Z Tagger: Effect of λ

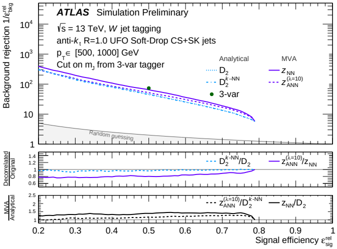


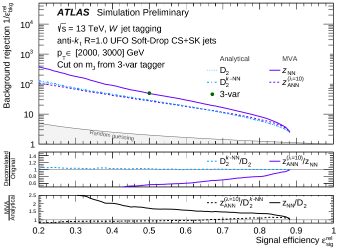
ATL-PHYS-PUB-2021-029

W/Z Tagger: ROC Curves



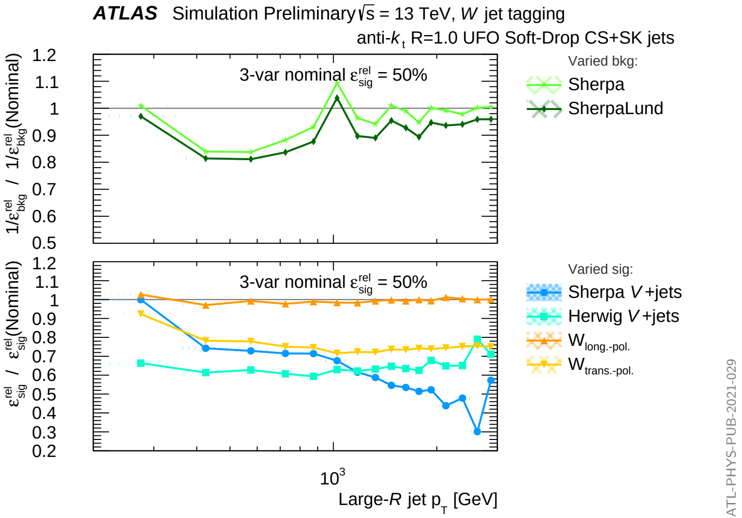




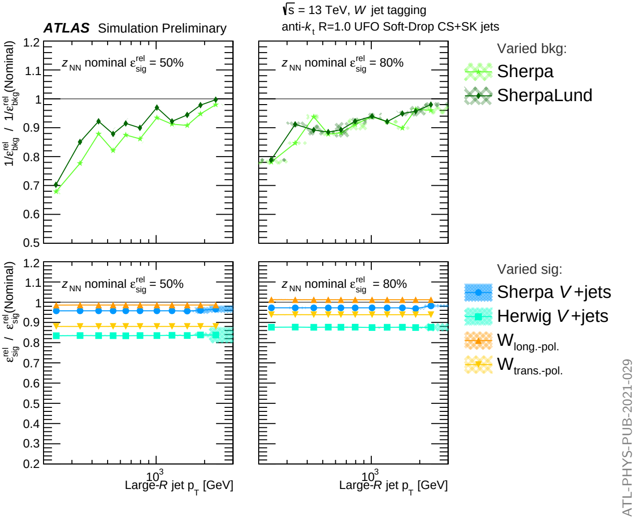


ATL-PHYS-PUB-2021-029

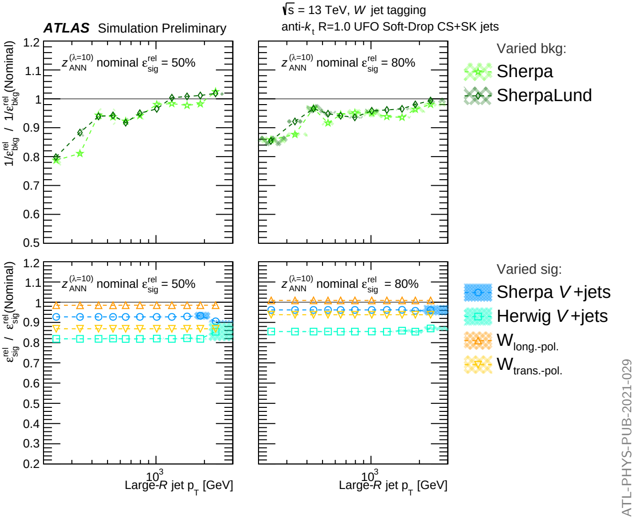
W/Z Tagger: Modelling



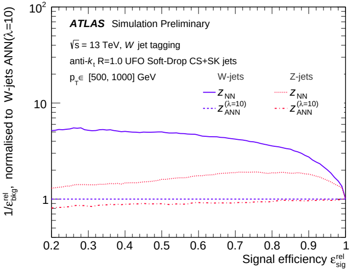
W/Z Tagger: Modelling

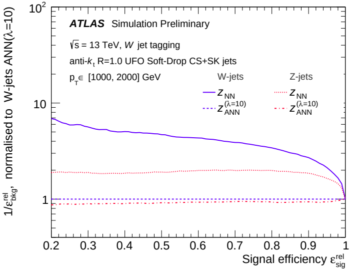


W/Z Tagger: Modelling



W/Z Tagger: Z vs W





ATL-PHYS-PUB-2021-029

Top Taggers

Truth Labelling: Inclusive/Contained

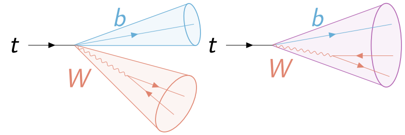
Two types of DNN-based top taggers defined:

Inclusive: Purely defined by ΔR matching:

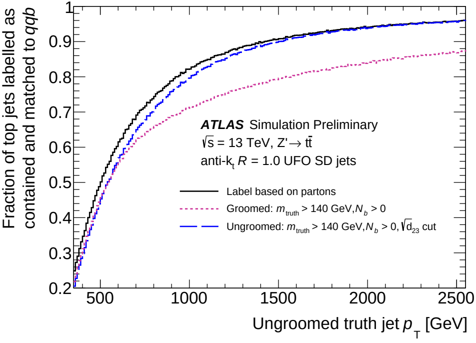
- $\Delta R(j^{\text{reco}},\ j^{\text{truth}}) < 0.75$ and $\Delta R(j^{\text{truth}},\ t^{\text{truth}}) < 0.75$

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

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

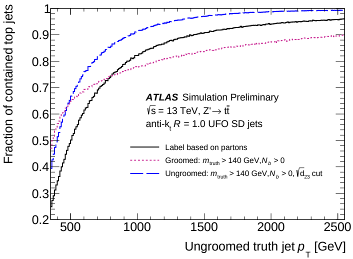


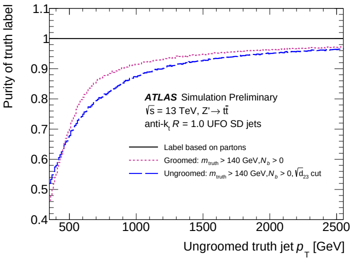
Contained top labelling efficiency:

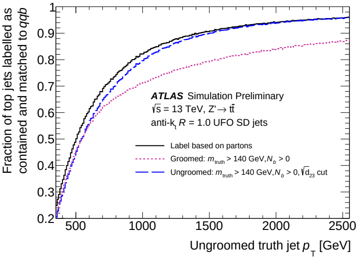


ATL-PHYS-PUB-2021-028

DNN Top Tagger: Truth Labelling

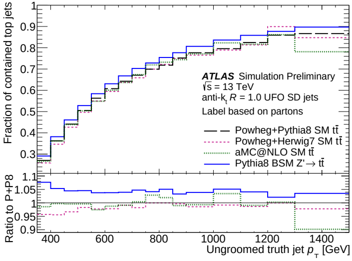


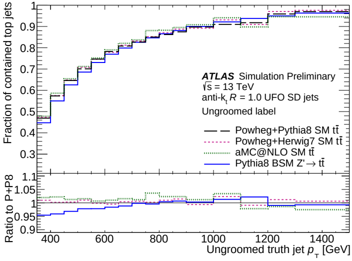




ATL-PHYS-PUB-2021-028

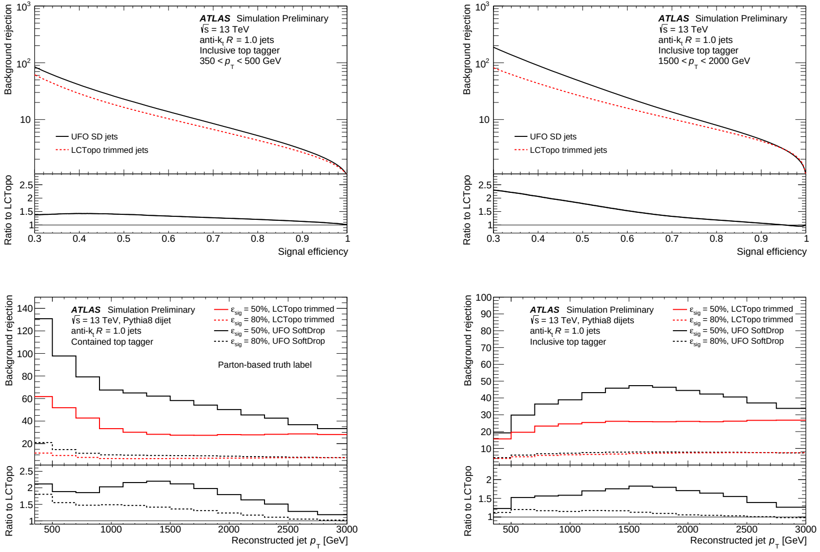
DNN Top Tagger: Truth Labelling





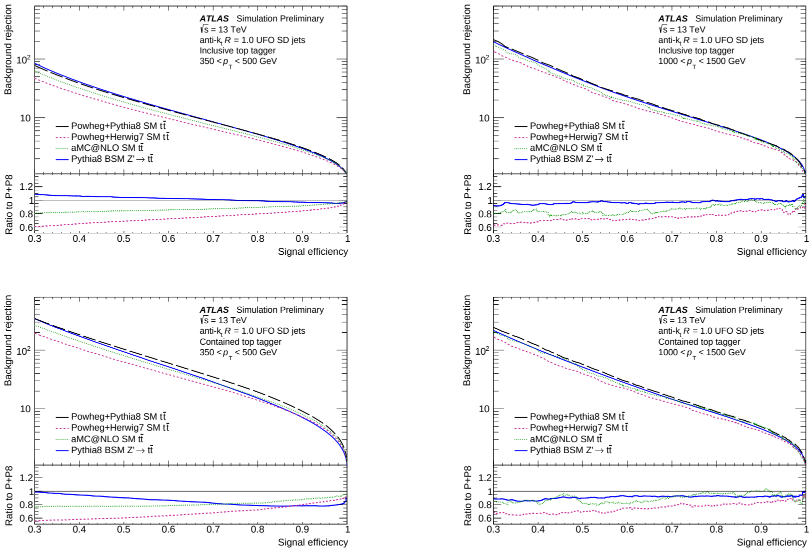
ATL-PHYS-PUB-2021-028

DNN Top Tagger: ROC Curves



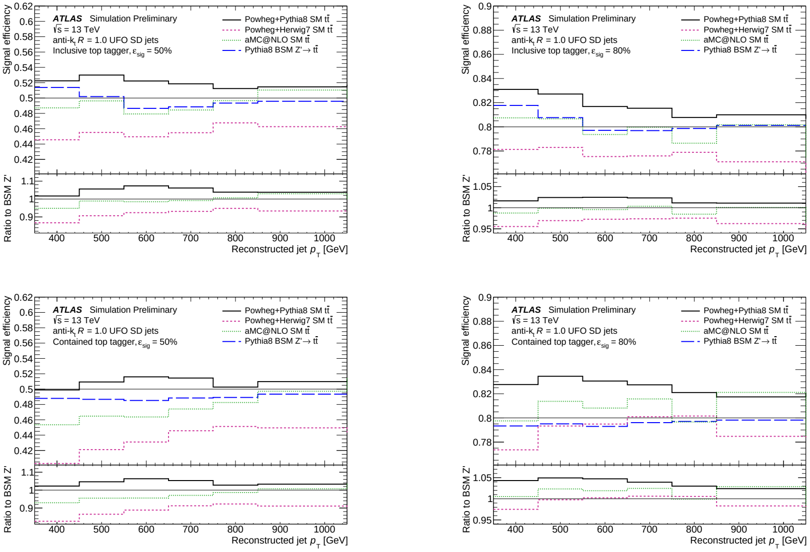
ATL-PHYS-PUB-2021-028

DNN Top Tagger: Modelling



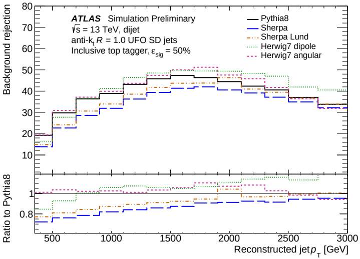
ATL-PHYS-PUB-2021-028

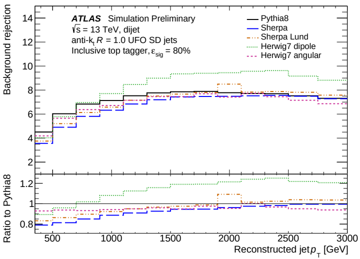
DNN Top Tagger: Modelling

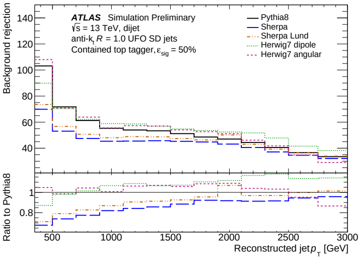


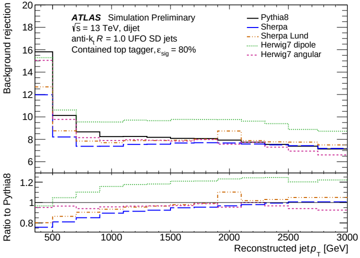
ATL-PHYS-PUB-2021-028

DNN Top Tagger: Modelling



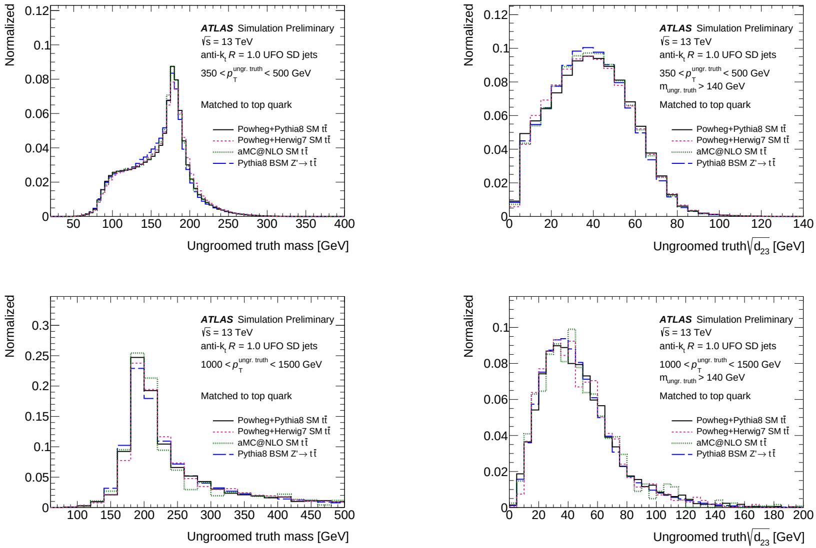






ATL-PHYS-PUB-2021-028

DNN Top Tagger: Modelling for Truth Labelling

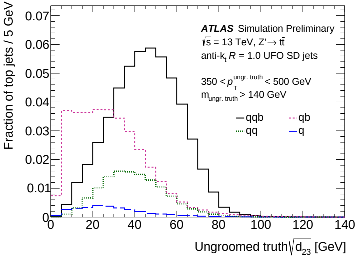


ATL-PHYS-PUB-2021-028

DNN Top Tagger: Truth Labelling

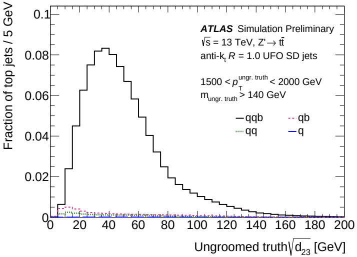
Low p_T :

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



High p_T :

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



ATL-PHYS-PUB-2021-028

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
hlDNN	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

ATL-PHYS-PUB-2022-39

Substructure Variables

Substructure Variables

W/Z tagger (NN/ANN)

$D_2,\ C_2$	Energy correlation ratios
τ_{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-subjet ΔR
n_{trk}	number of tracks

Top tagger (DNN)

$\tau_1,\ \tau_2,\ \tau_3,\ \tau_4$	N-subjettiness
$\sqrt{d_{12}},\ \sqrt{d_{23}}$	Splitting scales
$\mathrm{ECF}_1,\ \mathrm{ECF}_2,\ \mathrm{ECF}_3$	Energy correlation (EC) functions
$C_2,\ D_2$	EC ratios
$L_2,\ L_3$	Generalised EC ratios
Q_{W}	Invariant mass / virtuality
T_{M}	Thrust major

arxiv.org/abs/1305.0007

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

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: ECF_1, ECF_2, ECF_3;

Energy Correlation Ratios C_2, D_2

doi.org/10.1007/JHEP12(2014)009

Normalised ECFs e_n^β:

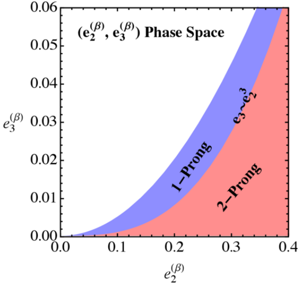
$$e_n^\beta = \frac{\mathrm{ECF}(n,\ \beta)}{\mathrm{ECF}(1,\ \beta)^n}; \qquad z_i = \frac{p_{T_i}}{p_{T_J}}$$

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

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

Theory e_3/e_2 phase space

(from power counting of z & R)

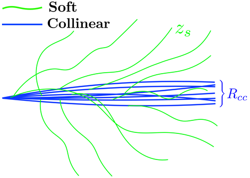


doi.org/10.1007/JHEP12(2014)009

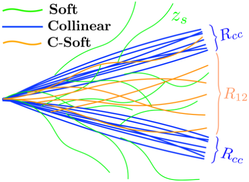
Ratios of e_n^β :

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

1-prong jet



2-prong jet



doi.org/10.1007/JHEP12(2014)009

$\rightarrow$ C_2 and D_2 Separate 1- and 2-prong jets on the e_3/e_2 plane

N-Subjettiness τ_N

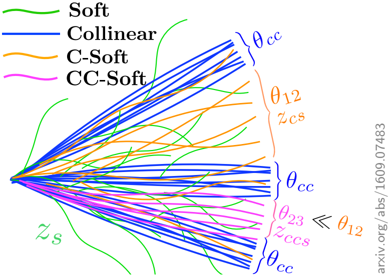
https://arxiv.org/abs/1011.2268

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

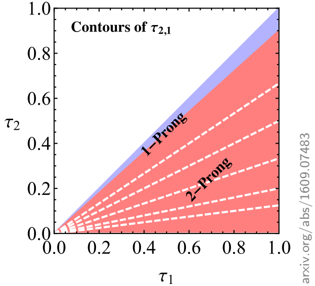
$$\text{with } z_i = \frac{p_{T_i}}{p_{T_J}}$$

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

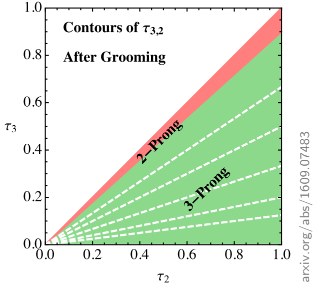
3-prong jet



Theory τ_2/τ_1 phase space



Theory τ_3/τ_2 phase space



Generalised Energy Correlation Functions ${}_{\nu}e_N^{\beta} \rightarrow L_2,\ L_3$

arxiv.org/abs/1609.07483

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

Where $\min\limits_{X}^{(m)}$ denotes the mth smallest element in the set X

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

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

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

- ECF average over them

Ratios to separate 2- & 3-prong jets: $L_2 = \frac{{}_{3}e_3^{\beta=1}}{\left({}_{1}e_2^{\beta=2}\right)^{\frac{3}{2}}},\ L_3 = \frac{{}_{1}e_3^{\beta=1}}{\left({}_{3}e_3^{\beta=1}\right)^{\frac{1}{3}}}$

Number of Ghost-Associated Tracks n_{trk}

JHEP04(2008)005

n_{trk}: number of tracks

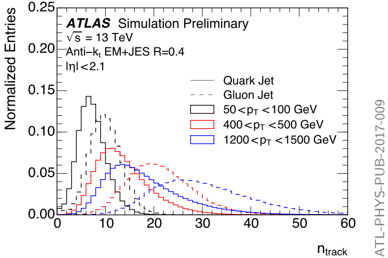
- With $p_{\mathrm{T}} > 500$ MeV
- Ghost-associated to jet

$\rightarrow$ 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



Ghost associated areas of k_t jets