Improving ATLAS hadronic object performance with ML/AI

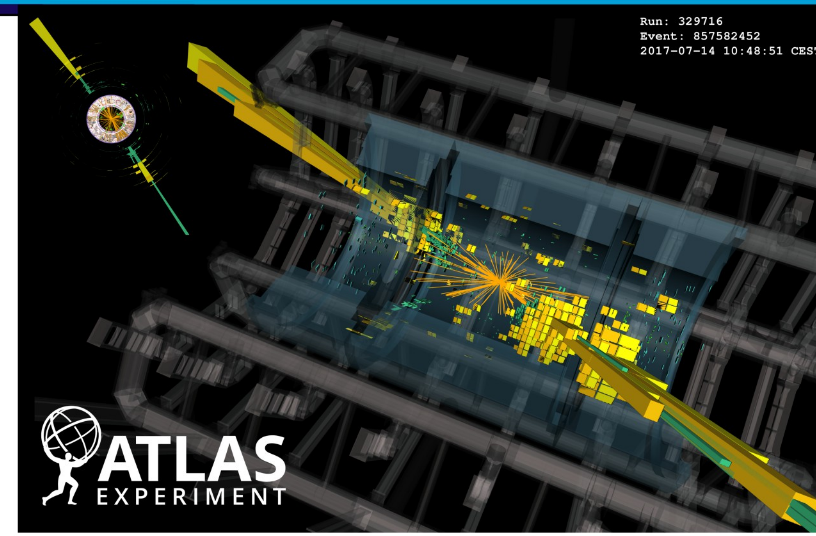
Reina Camacho Toro

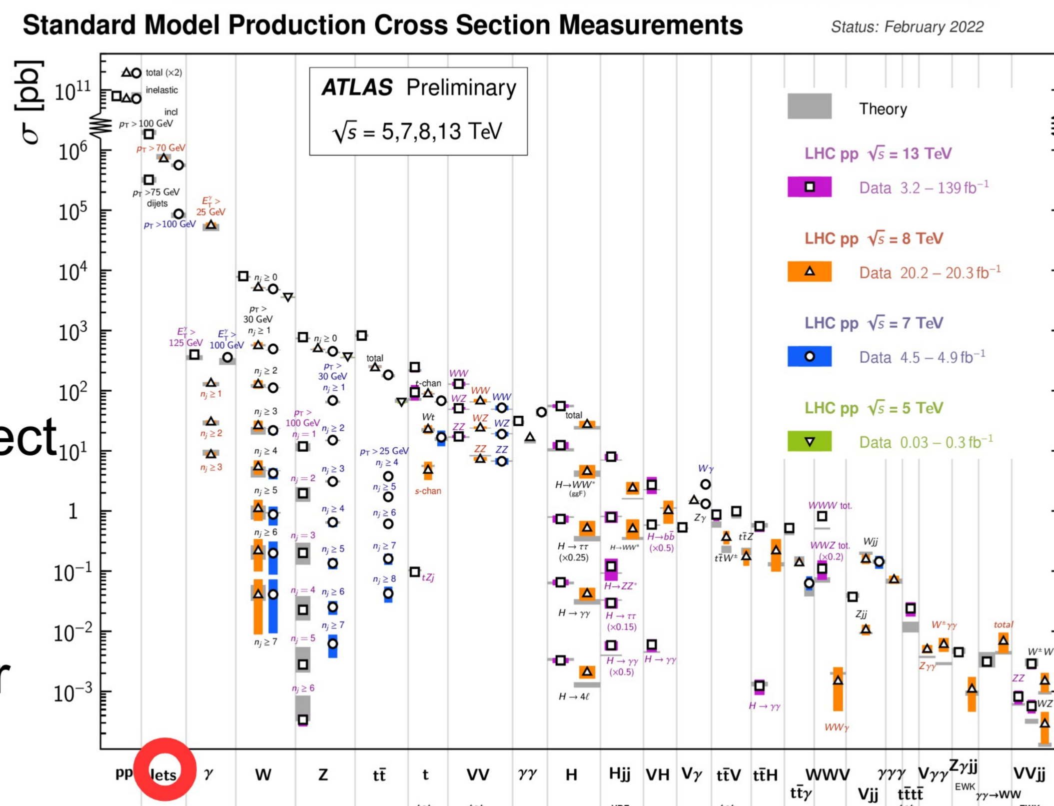
LPNHE/CNRS
On behalf of the ATLAS Collaboration

HEP2023, 8th international conference on High Energy Physics in the LHC Era
Universidad Técnica Federico Santa María (UTFSM), Valparaíso, Chile
January 9th-13th, 2023

The LHC is a very jetty place

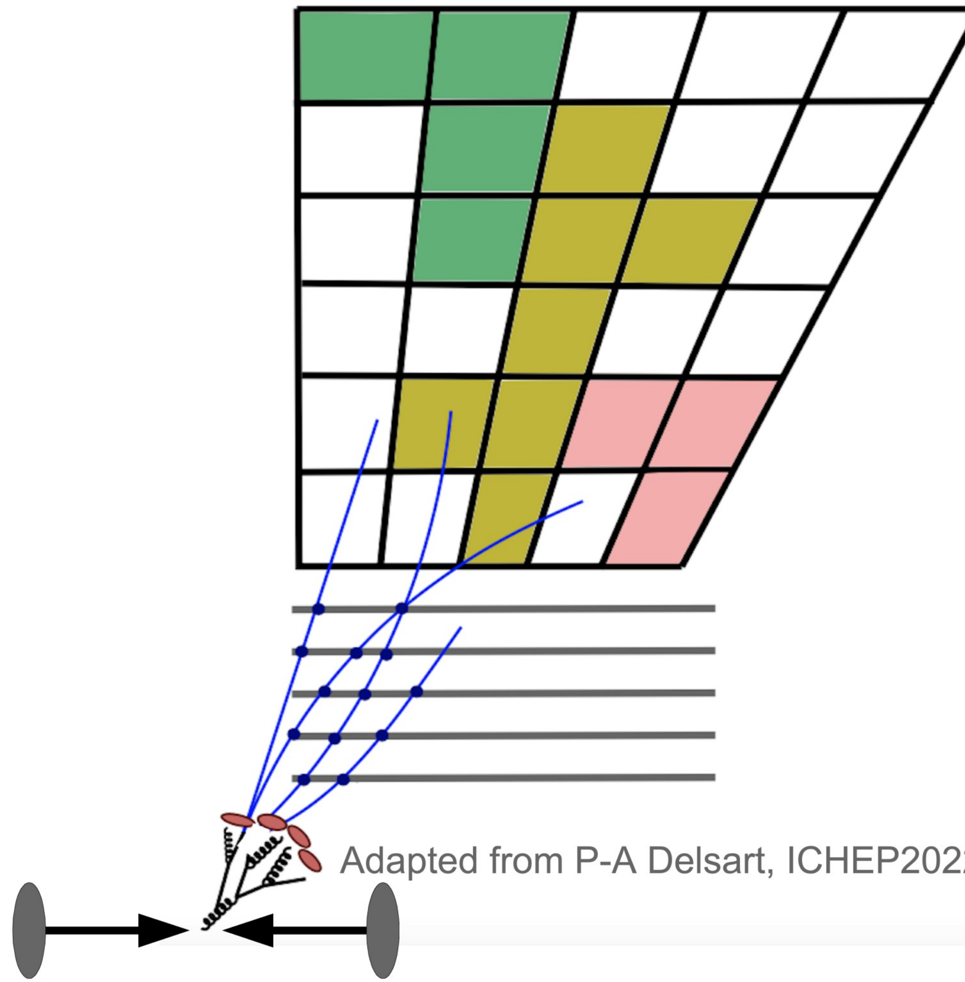
- Strongly interacting quarks and gluons produced in the LHC collisions hadronise and produce a cascade of particles, which can be collected using some specialised algorithms to build what we call “jets”
- Jets are ubiquitous in LHC analyses for new physics searches and Standard Model measurements
- Therefore a good hadronic object performance translates into physics precision
- There are many steps to take and experimental challenges to hadronic object reconstruction
- ATLAS has achieved percent level precision with Run-2 LHC data... and our optimisation work continues...





Hadronic object reconstruction in ATLAS: Jet and missing transverse energy

- Steps in jet building:
  - **Inputs/constituents:** jets are reconstructed from 4-vectors inputs representing the hadronic flow, such as tracks, calorimeter clusters, truth particles
  - **Reconstruction:** group constituents with a proper jet algorithm. Apply “grooming” (PU mitigation)
  - **Calibration:** to correct the jet energy (and also the mass in some cases) scale
  - **Tagging:** studying its substructure we can identify which particle is at the origin of the jet, e.g. is a quark or a gluon? Or rather a vector boson? Is it Higgs or a top quark?



Adapted from P-A Delsart, ICHEP2022

- Missing transverse momentum (p_T^{miss})
  - Calculated from the negative sum of the momenta the calibrated hard objects in the event; electrons, muons, τ-jets, photons, and jets
  - For soft energy; tracks from the PV not associated with hard objects are included

Hadronic object reconstruction in ATLAS: Jet and missing transverse energy

- Steps in jet building:
  - **Inputs/constituents:** jets are reconstructed from 4-vectors inputs representing the hadronic flow, such as tracks, calorimeter clusters, truth particles
  - **Reconstruction:** group constituents with a proper jet algorithm. Apply "grooming" (PU mitigation)
  - **Calibration:** to correct the jet energy (and also the mass in some cases) scale
  - **Tagging:** studying its substructure we can identify which particle is at the origin of the jet, e.g. is a quark or a gluon? Or rather a vector boson? Is it Higgs or a top quark?
- Missing transverse momentum (p_T^{miss})
  - Calculated from the negative sum of the momenta the calibrated hard objects in the event; electrons, muons, τ-jets, photons, and jets
  - For soft energy; tracks from the PV not associated with hard objects are included

All these steps are promising settings for cutting-edge machine learning and artificial intelligence algorithms at the LHC!

Selection of 4 ATLAS recent developments to improve hadronic object performance with ML

In 15mins I can't cover everything and there is much more information in the links

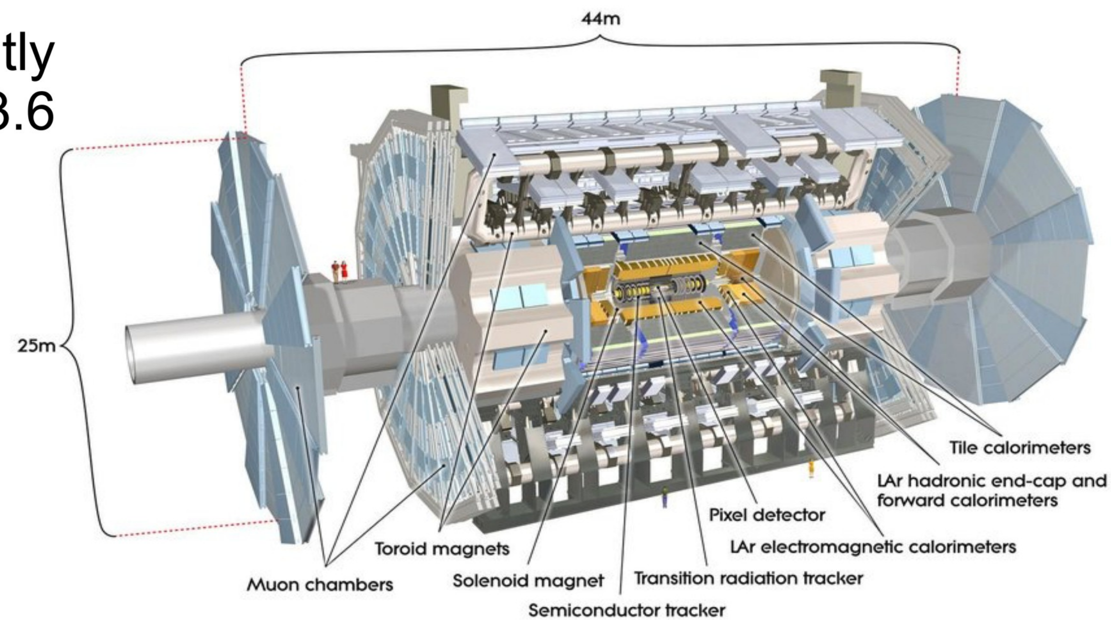
For more information about other developments, with a focus on small-radius jets, please see L. Ginabat's talk

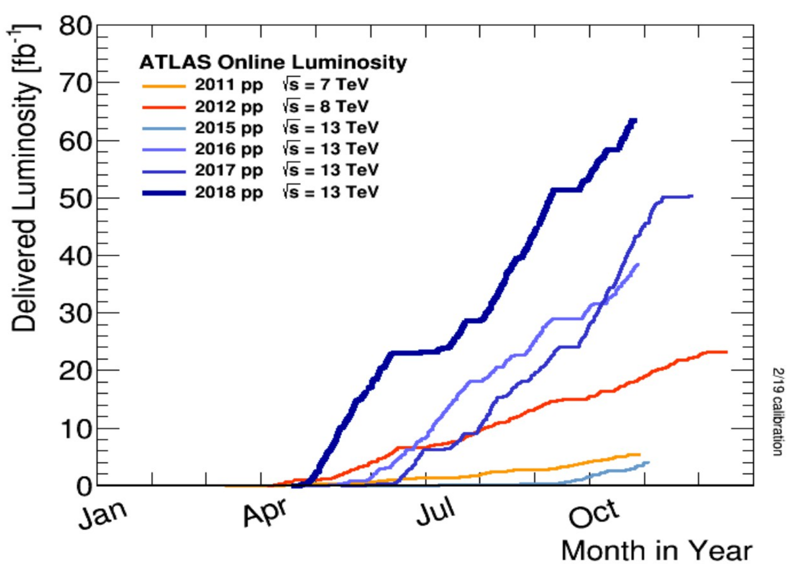
Our tools
The LHC and ATLAS

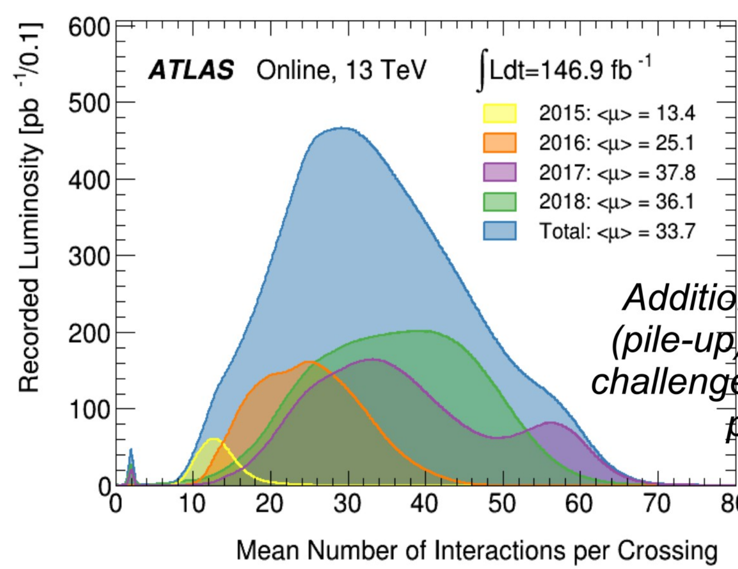
- A proton-proton collider of 27 Km circumference situated at CERN. Currently running at a center-of-mass energy of 13.6 TeV since 2022

- **Fantastic machines with capabilities beyond design**

- ATLAS is a non-specialized detector:
  - Excellent vertex and tracking systems
  - Large coverage for muon detection
  - Excellent calorimetry with extended coverage





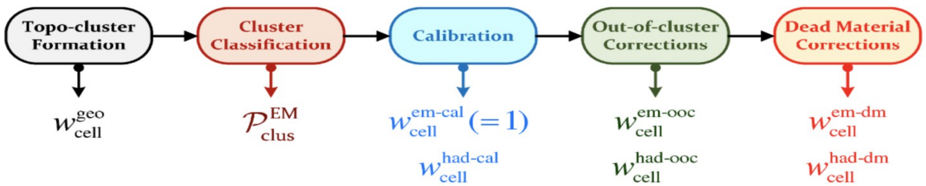


Additional pp interactions (pile-up) is one of the main challenges in hadronic object performance

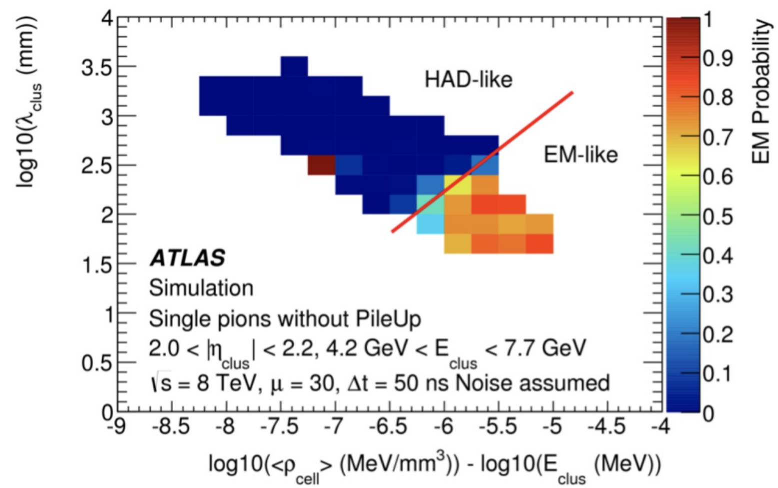
Classifying and calibrating clusters with ML

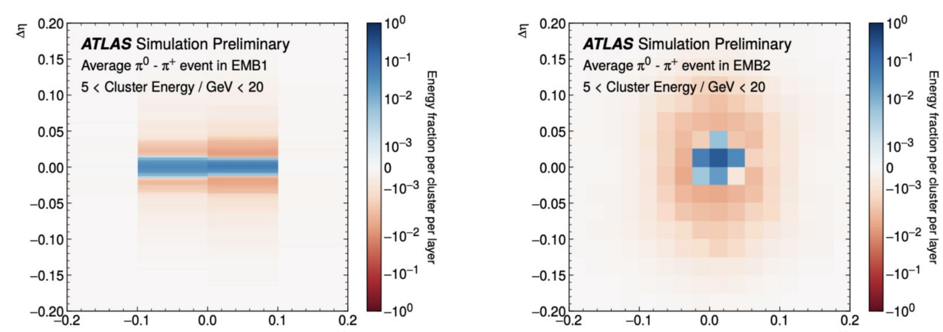
arXiv:1603.02934

- Our basic calorimeter energy reconstruction is based on **topoclusters**
- **Current approach:**
  - **Cluster classification (P^{EM}_{clus})** based on geometric and signal moments per-cluster
  - **Cluster calibration through Local Hadronic Cell Weighting (LCW)** based on local properties
  - All this in order to take into account the non-compensating nature of our calorimeter: different response for $\pi^{\pm},\pi^{0}$
- **New developments:**
  - **What if the many cells in topoclusters are represented in a different way using ML?** Could we improve the classification and calibration?
    - Images (1 cell = 1 pixel) → Convolutional NN (CNN)
    - Point clouds (1 cell = 1 point) → DeepSets/ParticleFlow Network (PFN)
    - Graphs (1 cell = 1 node) → Graph NN (GNN)



$$\mathcal{P}^{EM}_{clus}(E^{EM}_{clus}, \eta_{clus}, \rho_{clus}/E^{EM}_{clus}, \lambda_{clus})$$



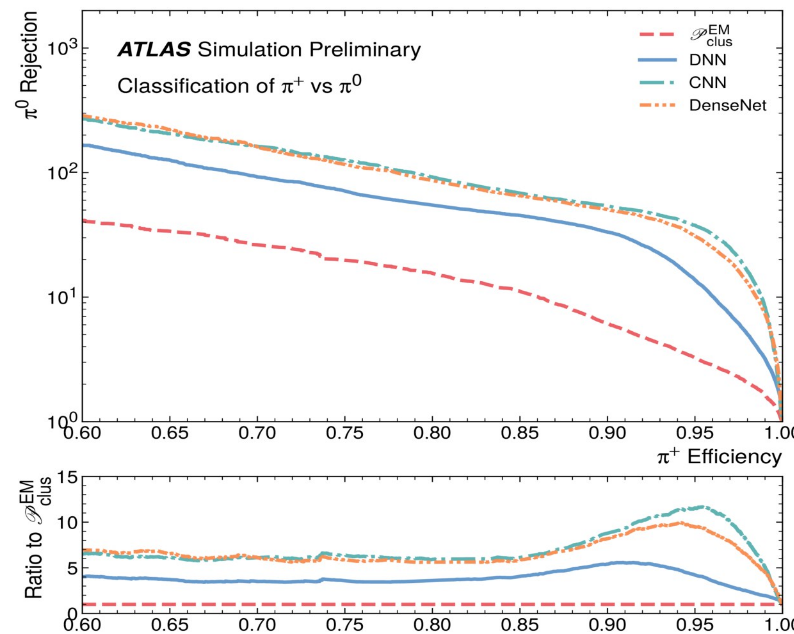


Inputs examples used for CNN

Classifying and calibrating clusters with ML

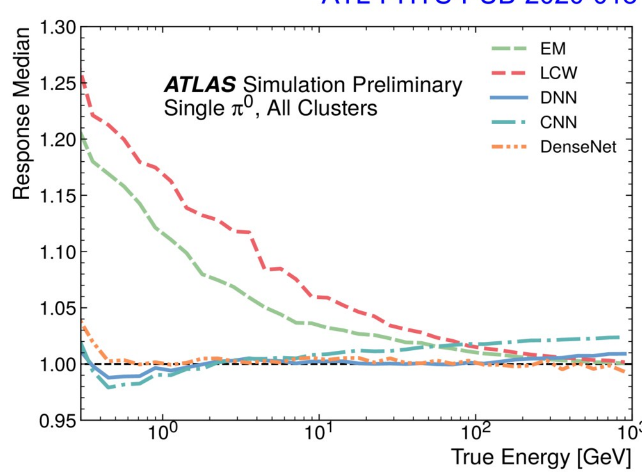
- **Classification of isolated charged vs neutral pions in simulation:**
  - DNN, CNN and PFN compared with ATLAS standard technique (P^{EM}_{clus})
  - ML improves rejection by factor >5

ATL-PHYS-PUB-2020-018

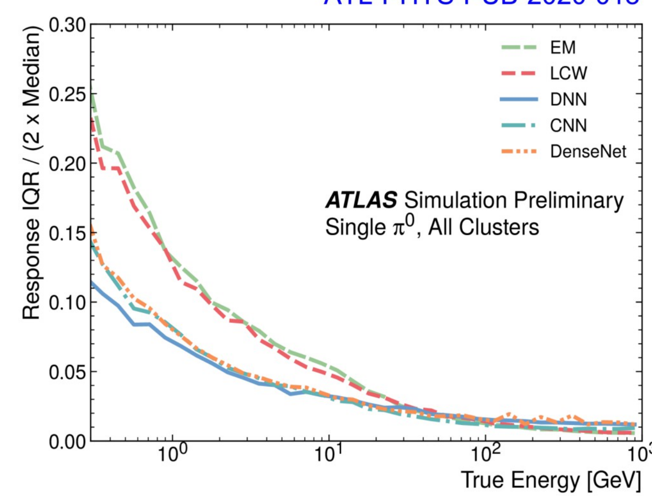


- **Calibrating the classified cluster energy response**
  - Compared with ATLAS uncalibrated (EM) and calibrated (LCW) clusters
  - ML improves the response and resolution
  - More results using also tracking information in PUB note link!

ATL-PHYS-PUB-2020-018



ATL-PHYS-PUB-2020-018



"This work demonstrates the potential of deep-learning based low-level hadronic calibrations to significantly improve the quality of particle reconstruction in the ATLAS Calorimeter"

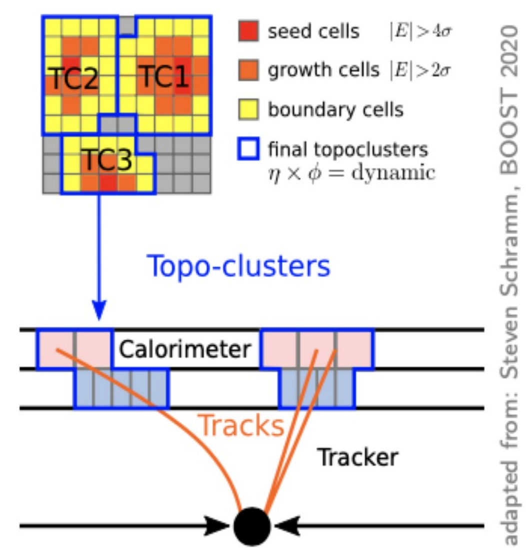
A new set of constituents/inputs: Unified Flow Objects

- During **Run-1 mostly calorimeter information** was used: topoclusters

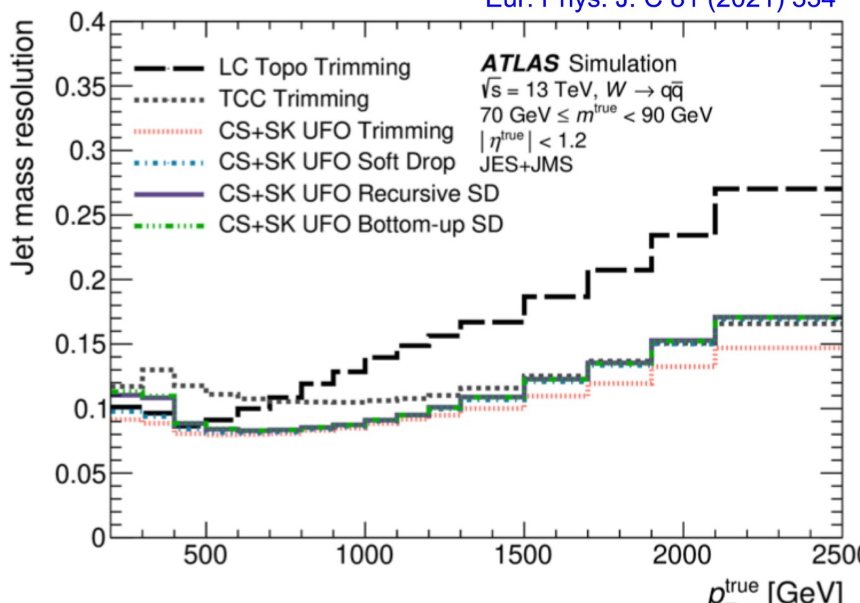
- In **Run-2 we started exploiting as well the information from the inner detector**:
  - **Particle Flow (PFlow) algorithm**: tracks with good momentum resolution extrapolated to calorimeter, cell-by-cell subtraction of their deposited energy. At high p_T tracks are ignored
  - **Track-calo cluster (TCC) algorithm**: effectively uses tracks to split up large clusters at high p_T, get energy from clusters but angles from tracks. At low p_T clusters-only are used

- Current state of the art is combination of TCC and PFlow
  - **Unified Flow Objects (UFO)**

- **UFO combines advantages**: good angular resolution of tracker and good energy resolution from calorimeter



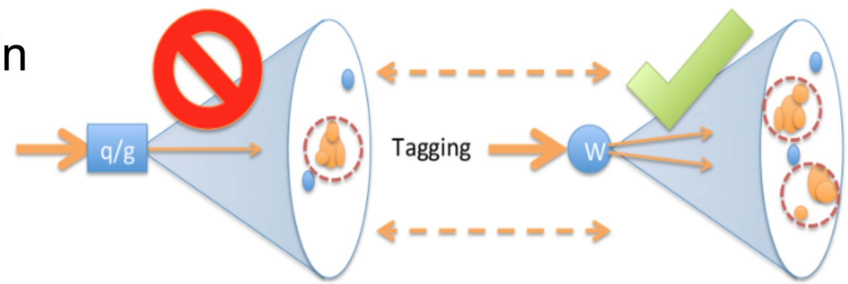
Eur. Phys. J. C 81 (2021) 334



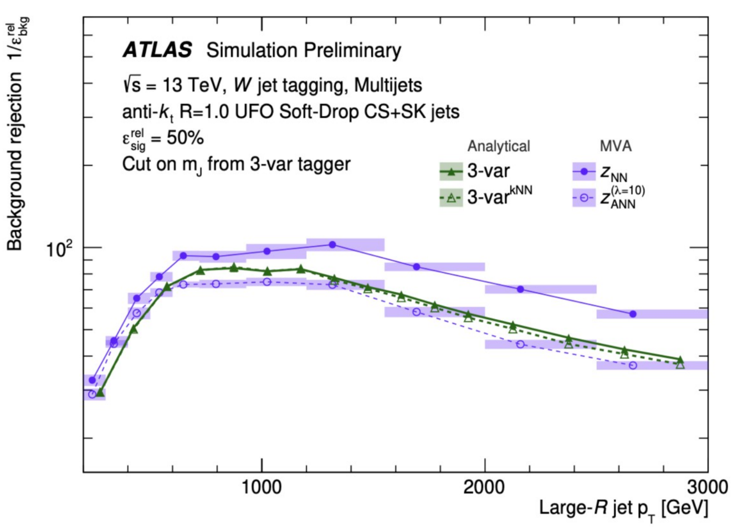
UFO currently being commissioned as baseline for large-radius jets

W/Z boosted boson tagging with UFO jets

- ATLAS has developed W/Z/H/top taggers in the last years taking advantage of the characteristic internal substructure of the large-R jets depending on their origin

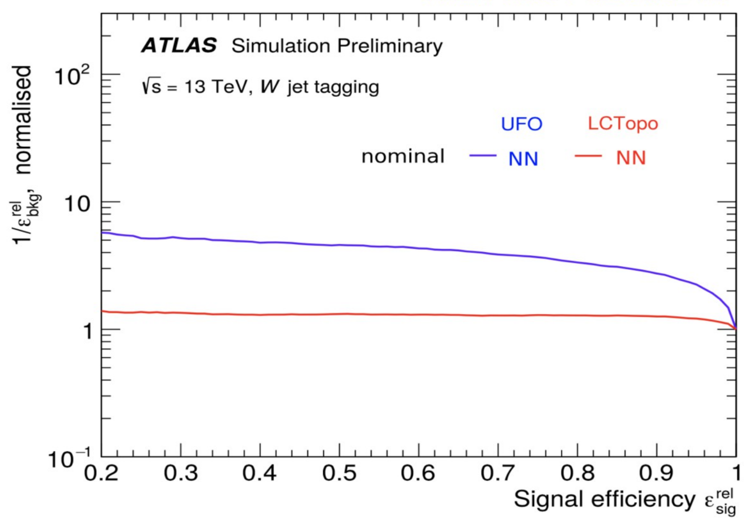


- **Two kind of multivariate taggers** being used at the moment
  - **Moment-based (3-var) taggers:** Identify jet substructure moments with good separation power, and apply cuts on them
  - **ML using high level features:** to exploit complex substructure correlations



ANN comparable with 3-variable tagger, but with decorrelation

ATL-PHYS-PUB-2022-39

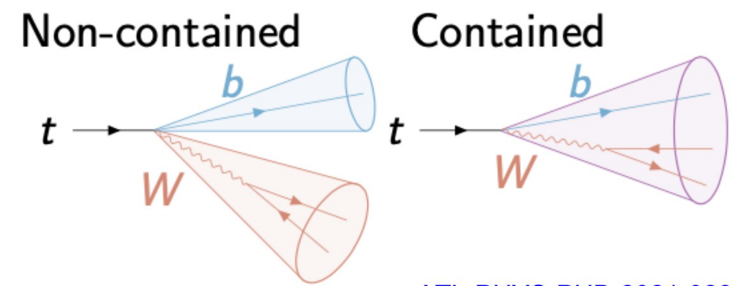


Background rejection improved by factor 2-3 in NN UFO tagger w.r.t .NN LCTopo tagger

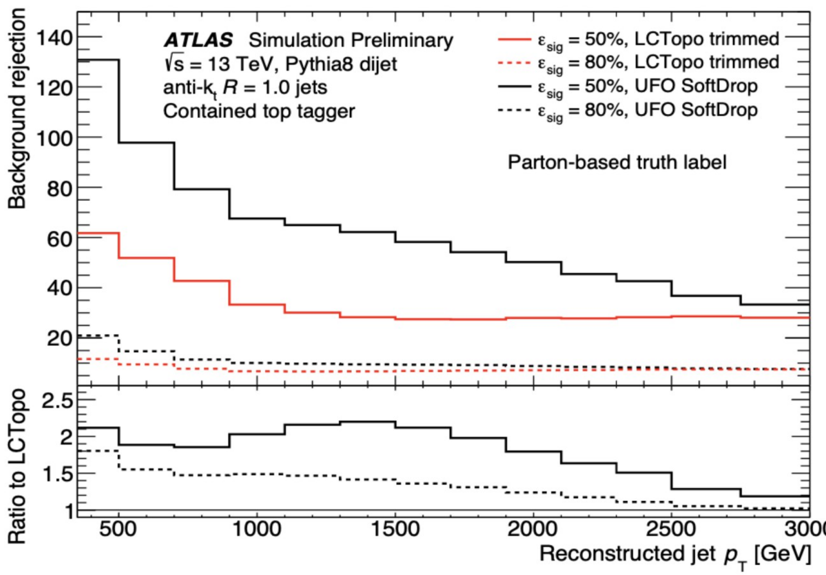
- 3-var: p_T-dependent cuts on:
  - Jet mass
  - Number of associated tracks: good quark/gluon discriminator
  - D_2 correlation function: exploit 2-prongedness of W/Z boson decay
- ML: Neural network with >10 different inputs
- **Decorrelated versions:**
  - Tagger may introduce unwanted mass shaping of the background
  - Adversarial Neural Network (ANN) successfully decorrelates

Boosted top tagging with UFO jets

- **Two Deep-NN based top taggers** defined:
  - **Contained and inclusive (any jet that contains parts of the decay) tops**
  - For 50% and 80% signal efficiency, p_T-dependent
  - 15 different jet substructure variables used as inputs (see backup for full list)
- **Some observations:**
  - 80% working point:
    - Inclusive: ~20% better rejection for p_T<1.5 TeV
    - Contained: Better over whole range wrt LCTopo taggers
  - 50% working point: clear improvement for inclusive and contained top taggers



ATL-PHYS-PUB-2021-028

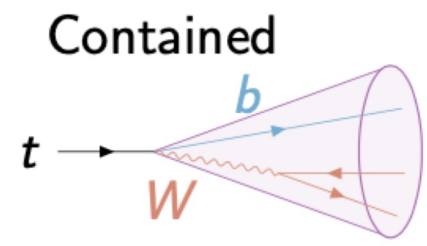


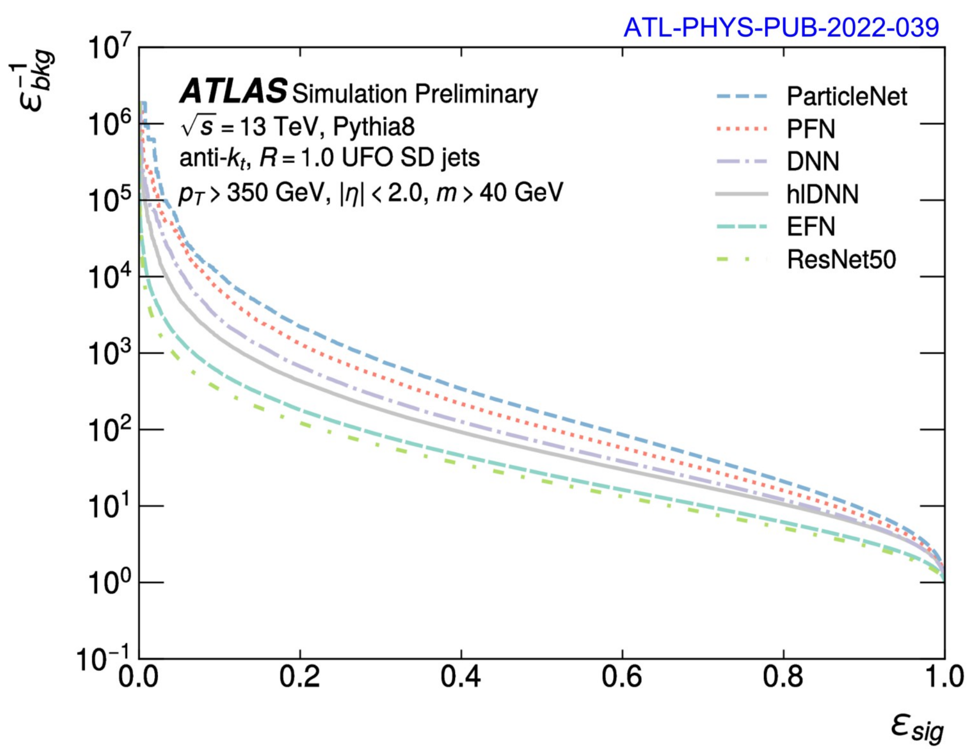
Boosted top tagging with UFO jets: constituent-based

- Another recent development are **constituent-based top taggers:**
  - Using low-level features based on 4-vectors of jet constituents. The information is combined with larger and more complex ML classifiers
  - Constituent level information pre-processed to exploit known symmetries
  - Contained boosted top case considered

- **What algorithms were studied?**
  - hlDNN: Baseline similar to DNN top tagger used by ATLAS in Run-2
  - DNN: Using constituent 4-momenta
  - EFN/PFN: Energy/Particle-flow networks
  - ResNet50: CNN using jet images
  - ParticleNet: Dynamic Graph-CNN

- **Some observations:**
  - ParticleNet and PFN show best performance
  - More details in next slide



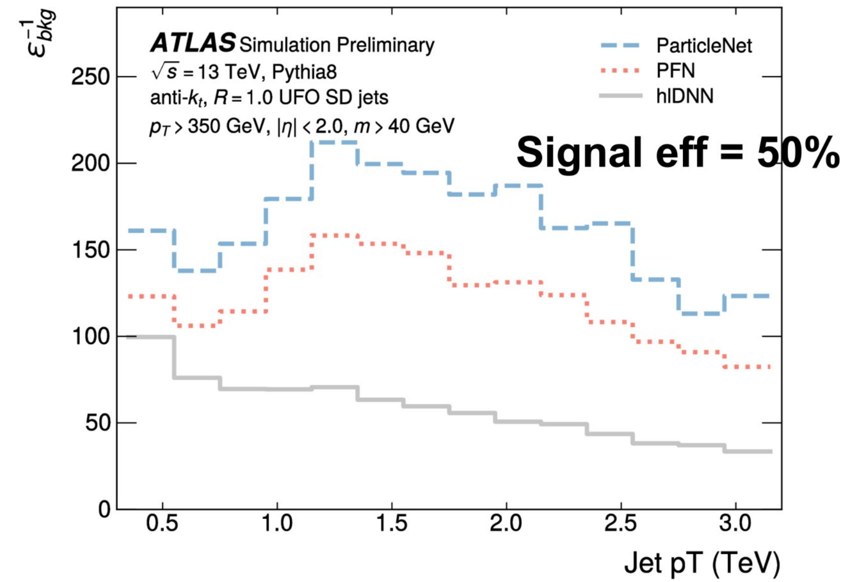


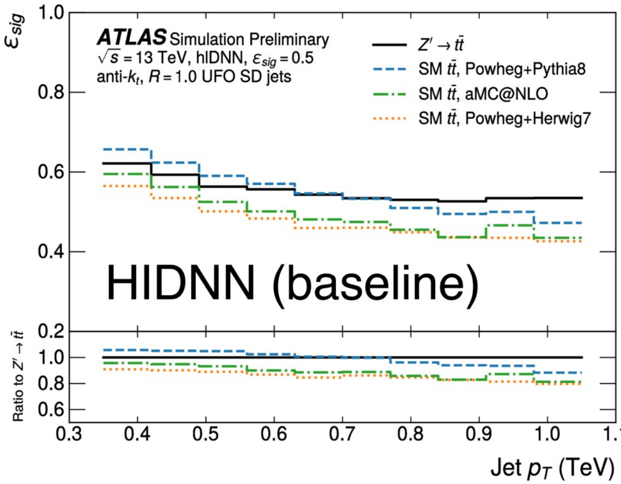
Simulated dataset used for training is public and documented:
http://opendata.cern.ch/record/15013

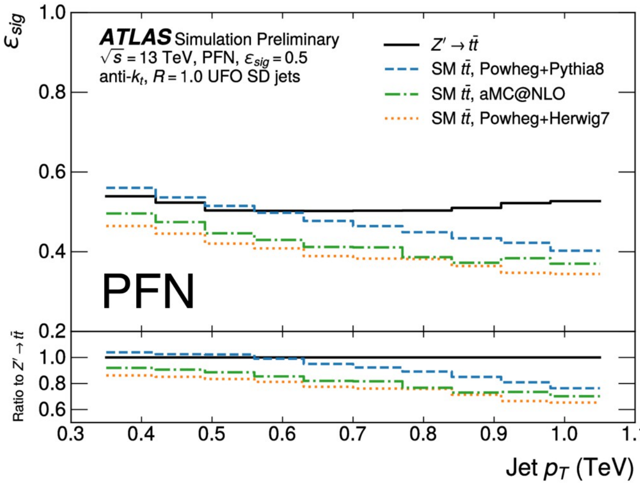
Boosted top tagging with UFO jets: constituent-based

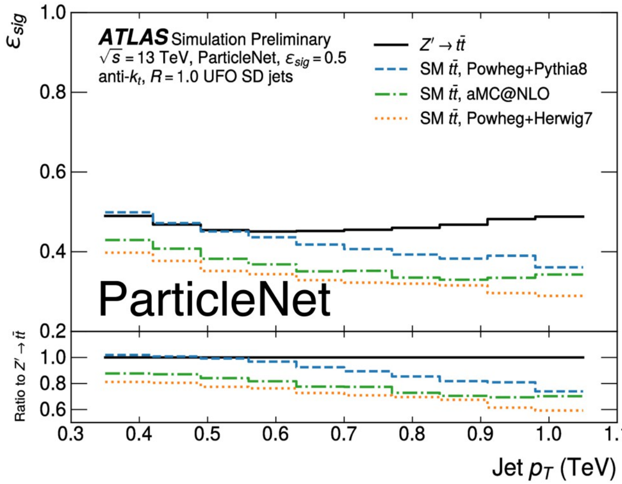
- ParticleNet and PFN achieve ~2-3x improvement in background rejection across kinematic range
- The **MC modelling dependence was also studied for these new taggers**:
  - PFN and Particle Net show increased model dependence wrt hIDNN (baseline)
  - Contributing to modelling uncertainties in physics analyses. *Important to understand the cause for future developments!*
  - To reduce the MC modelling uncertainty dedicated calibrations could be derived

ATL-PHYS-PUB-2022-039





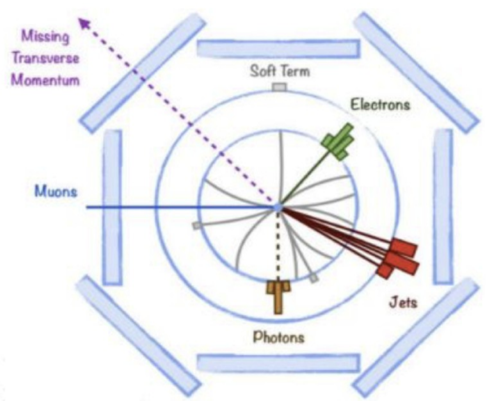




METNet: a MET combined working point using NN

- MET corresponds to the experimental proxy for the transverse momentum of undetected particles
  - Real MET: neutrinos, stable BSM particles, e.g. dark matter
  - Fake MET: smearing from pile-up, mis-measured objects, finite detector acceptance, etc

- **Current approach** in ATLAS: **Object based MET** reconstruction method
  - Calculated from the negative sum of the momenta the calibrated hard objects
  - Several different selections on jets are supported e.g. “tight”: higher p_T cuts on forward jets → **Different MET working points**



User-defined object selection

Specified by MET WP (Tight/Loose/Tenacious...)

$$E_T^{miss} = \underbrace{- \underbrace{\sum_{\substack{\text{selected}\\ \text{electrons}}} p_T^e}_{E_T^{miss,e}} - \underbrace{\sum_{\substack{\text{accepted}\\ \text{photons}}} p_T^\gamma}_{E_T^{miss,\gamma}} - \underbrace{\sum_{\substack{\text{accepted}\\ \tau\text{-leptons}}} p_T^{\tau_{had}}}_{E_T^{miss,\tau_{had}}} - \underbrace{\sum_{\substack{\text{selected}\\ \text{muons}}} p_T^\mu}_{E_T^{miss,\mu}} - \underbrace{\sum_{\substack{\text{accepted}\\ \text{jets}}} p_T^{jet}}_{E_T^{miss,jet}}}_{\text{hard term}} \underbrace{- \underbrace{\sum_{\substack{\text{unused}\\ \text{tracks}}} p_T^{track}}_{E_T^{miss,soft}}}_{\text{soft term}}$$

(+ internal overlap removal)

- **New developments:**
  - Currently analyses chose one MET working point to use for every event... But the optimal one for a given event depends on the pile-up and event topologies
  - **What if we could pick a different MET working point for each event?**
    - Could ML help us?

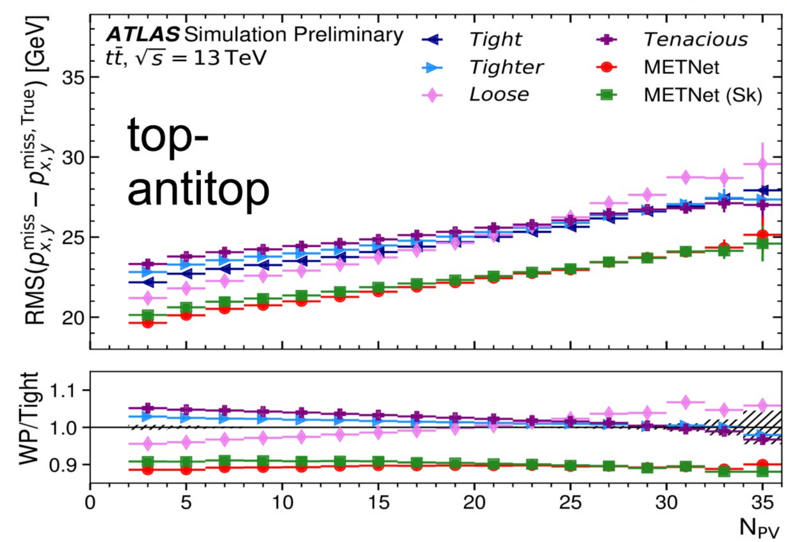
METNet: a MET combined working point using NN

- **METNet** is a **neural network** trained on simulation to perform a regression
  - Uses p_T^{miss} and event kinematics and conditions as inputs
  - Trained on a mix of topologies: top-antitop, WW and ZZ events

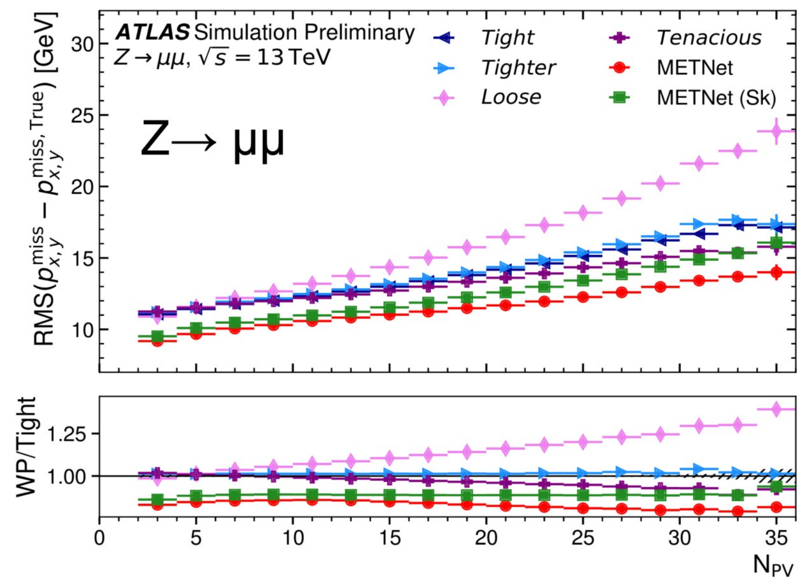
- Resulting in:
  - **Improved resolution when comparing with ATLAS current standard working points**
    - Even when studying processes not included in the training like single top and Z → μμ
  - Good p_T^{miss} response and distribution bias
    - Use of additional "Sinkhorn" loss can help with tails

- Promising studies indicate **potential to significantly improve p_T^{miss} resolution** using ML techniques
  - Further optimisations are possible, e.g. including tracking information

ATL-PHYS-PUB-2021-025



ATL-PHYS-PUB-2021-025



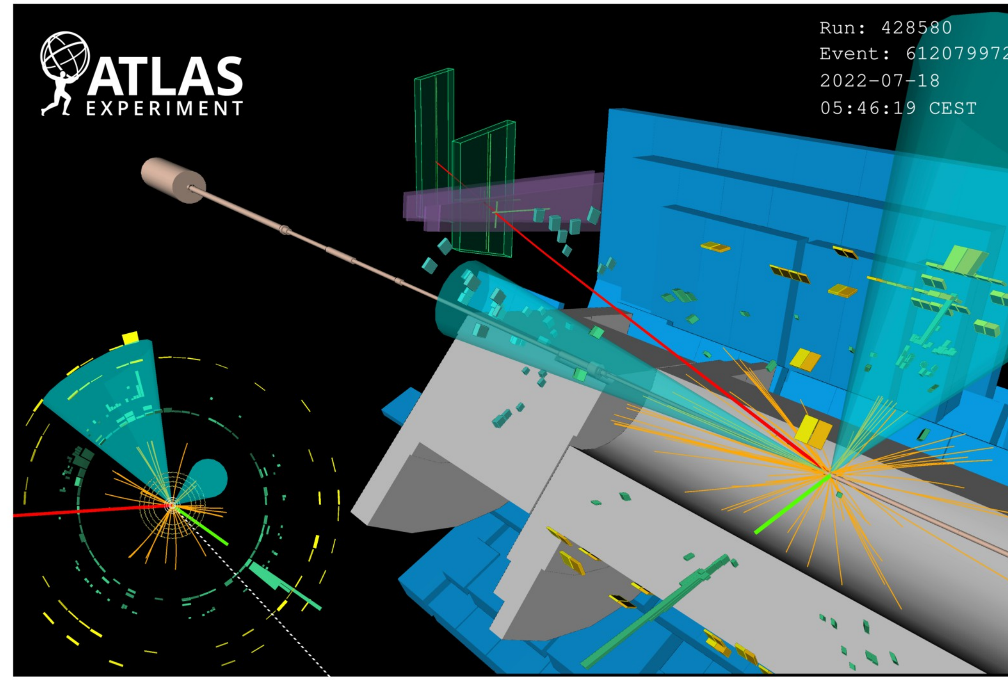
Summarising

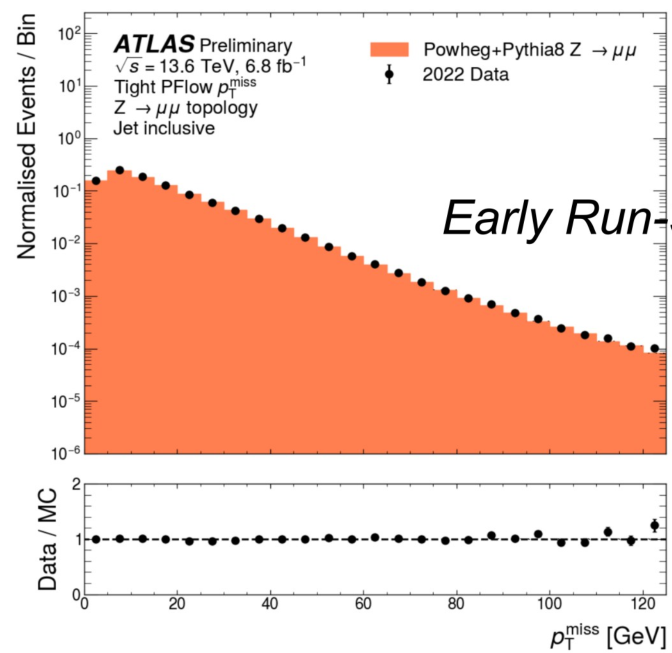
- Hadronic objects reconstruction, calibration and tagging important for precision measurements and BSM exploration at the LHC (and beyond)

- Great deal of improvements in the field during Run-2 and the LS2 from cutting-edge machine learning and artificial intelligence algorithms
  - Also exploring different techniques and phase-spaces (see L. Ginabat's talk)

- Run-3 just started, a great opportunity to exploit these developments and continue refining our strategies

  - Lots of work ongoing in that direction, stay tuned!





Early Run-3 data!

BACKUP

Substructure variables

W/Z tagger (NN/ANN)

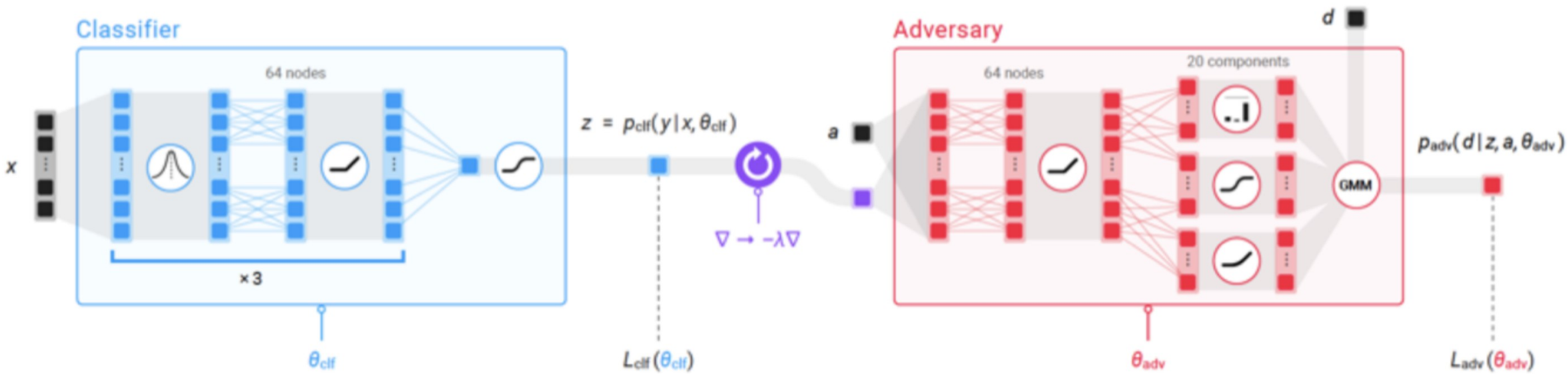
$D_2,\ C_2$	Energy correlation ratios
τ_{21}	N-subjettiness
R_2^{FW}	Fox-Wolfram moment
$\mathcal{P}$	Planar flow
a_3	Angularity
A	Aplanarity
Z_{cut}	$Z-$Splitting scales
$\sqrt{d_{12}}$	$d-$Splitting scales
$Kt\Delta R$	k_t-subjet ΔR
n_{trk}	number of tracks

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

Adversarial Neural Network

Adversarial Neural Network



Classifier Network

- Classifies signal vs. background jets by computing **z**.
- Feed-forwards **z** to adversary.
- Mass decorrelation by minimizing jet mass information in **z** and making it harder for the Adversary to infer jet mass.

Adversary Network

- Infers jet mass **d**, by constructing a Gaussian Mixture Model pdf and computing p(**d**).
- Loss function = - log p(**d**)
- Back-propagates loss via a gradient reversal channel to classifier (controlled by **λ**).

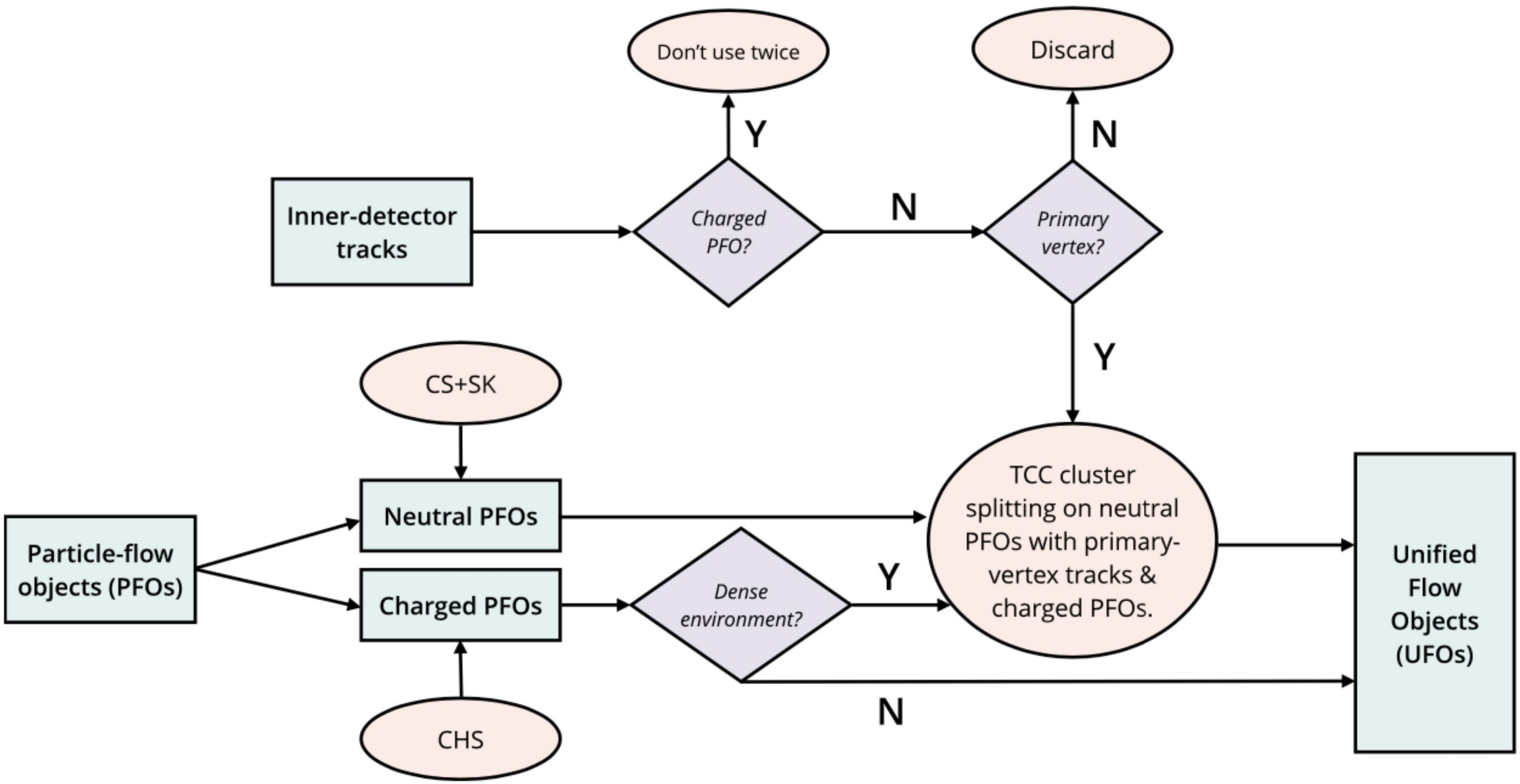
Joint Objectives: signal vs. background jet classification + minimal jet mass correlation.

Top constituent-based tagger

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

UFO algorithm

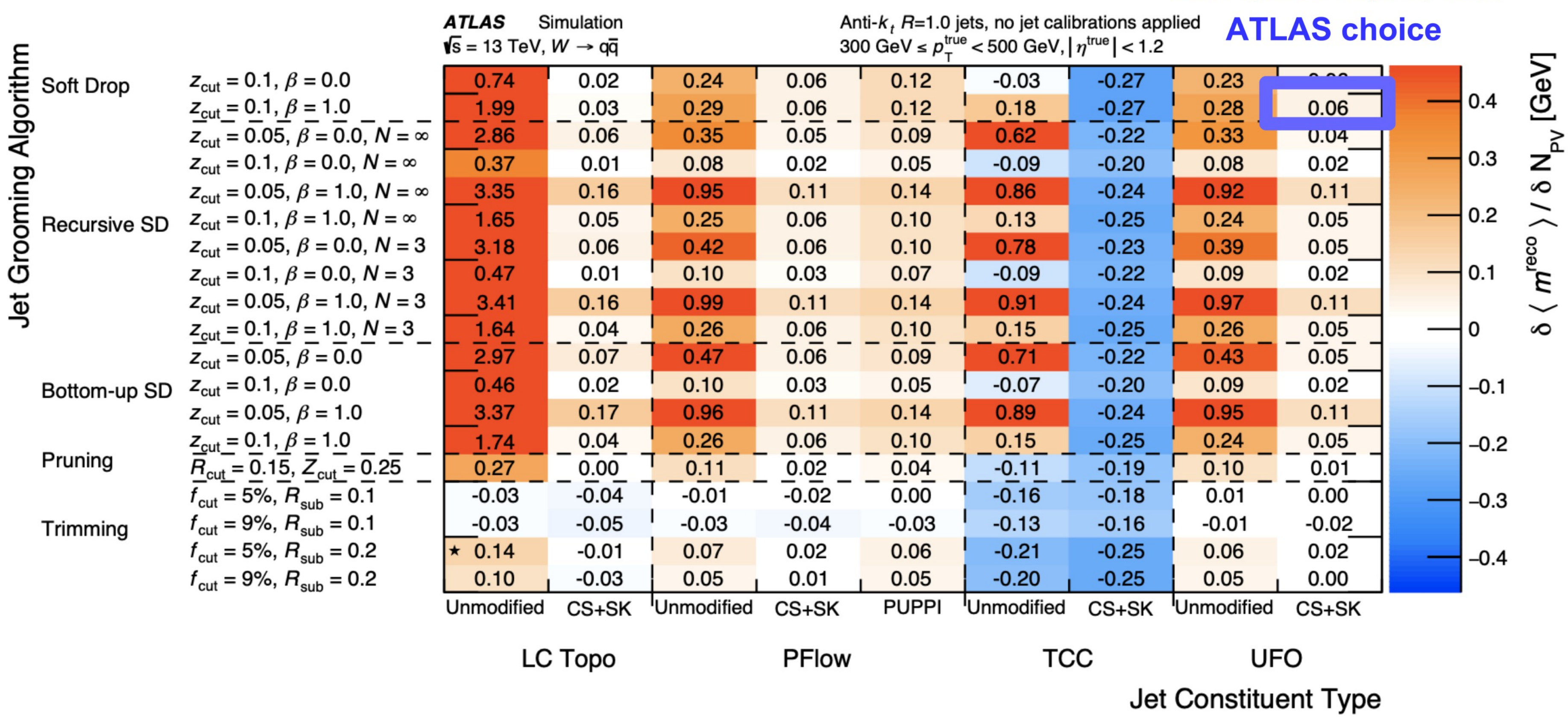


Eur. Phys. J. C 81, 334 (2021)

A new set of constituents/inputs: Unified Flow Objects

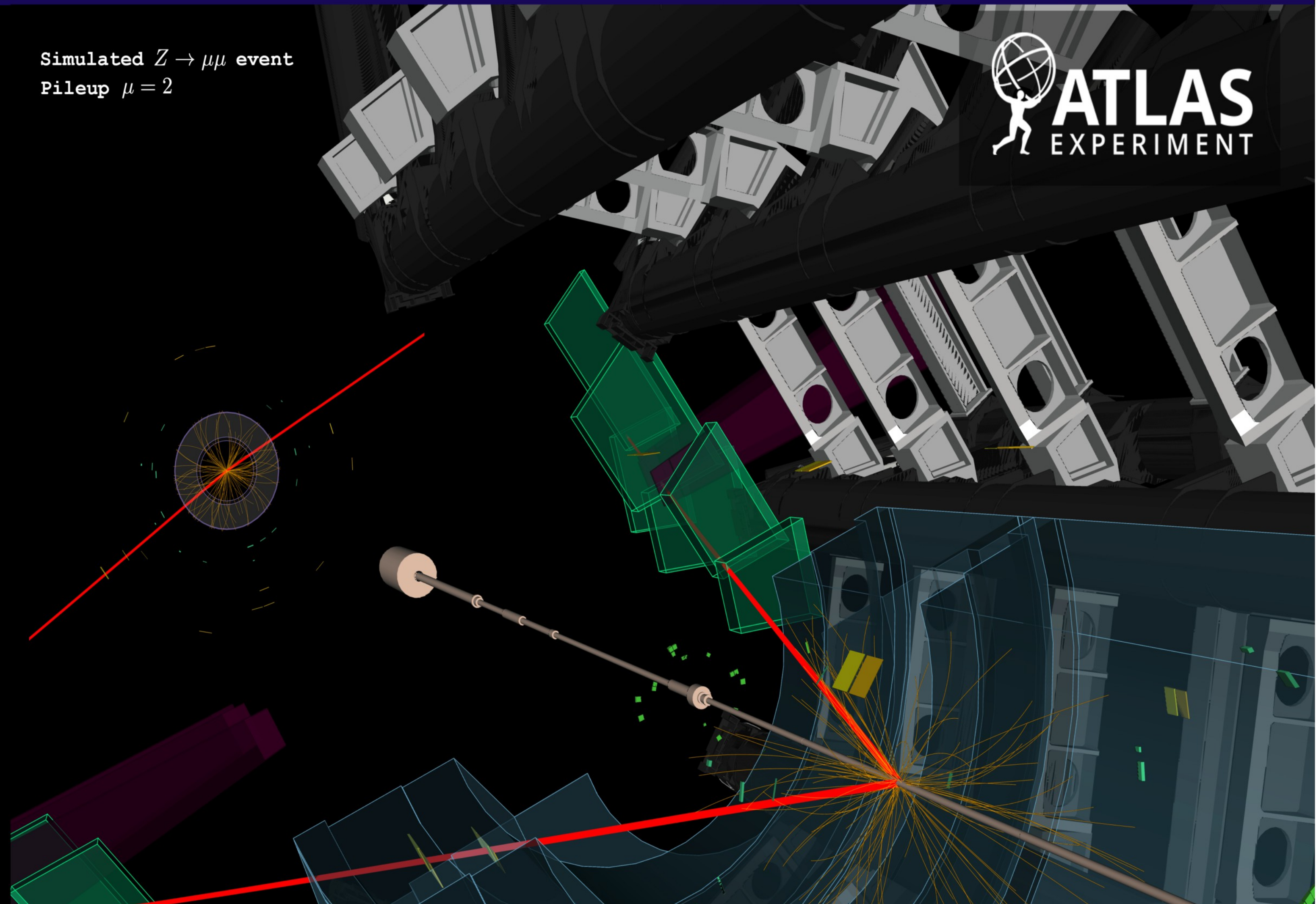
- **UFO currently being commissioned as baseline for large-radius jets**, following a long-term reoptimisation campaign during Run 2 and the long shutdown that followed
  - **Criteria used:** jet energy and mass resolution, pile-up stability, W/Z/Top tagging performance
  - **Choice:** UFO constituents, "Charged Subtraction"+"Soft Killer" pile-up mitigation and Soft Drop (Z=0.1,β=1) grooming technique (to remove energy from pile-up and possibly from underlying event)

Eur. Phys. J. C 81 (2021) 4, 334



Number Primary Vertex (NPV) impact on W-jet mass

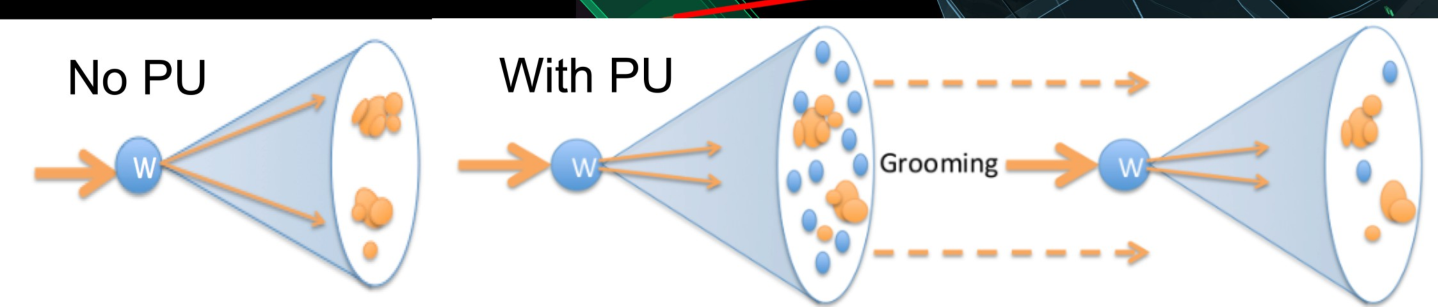
The LHC is a very jetty place



There are many challenges and the large number of additional interactions (pile-up) is one of them

PU treated for jets at different levels:

- in the calibration
- at the inputs
- using grooming techniques for large-radius jets

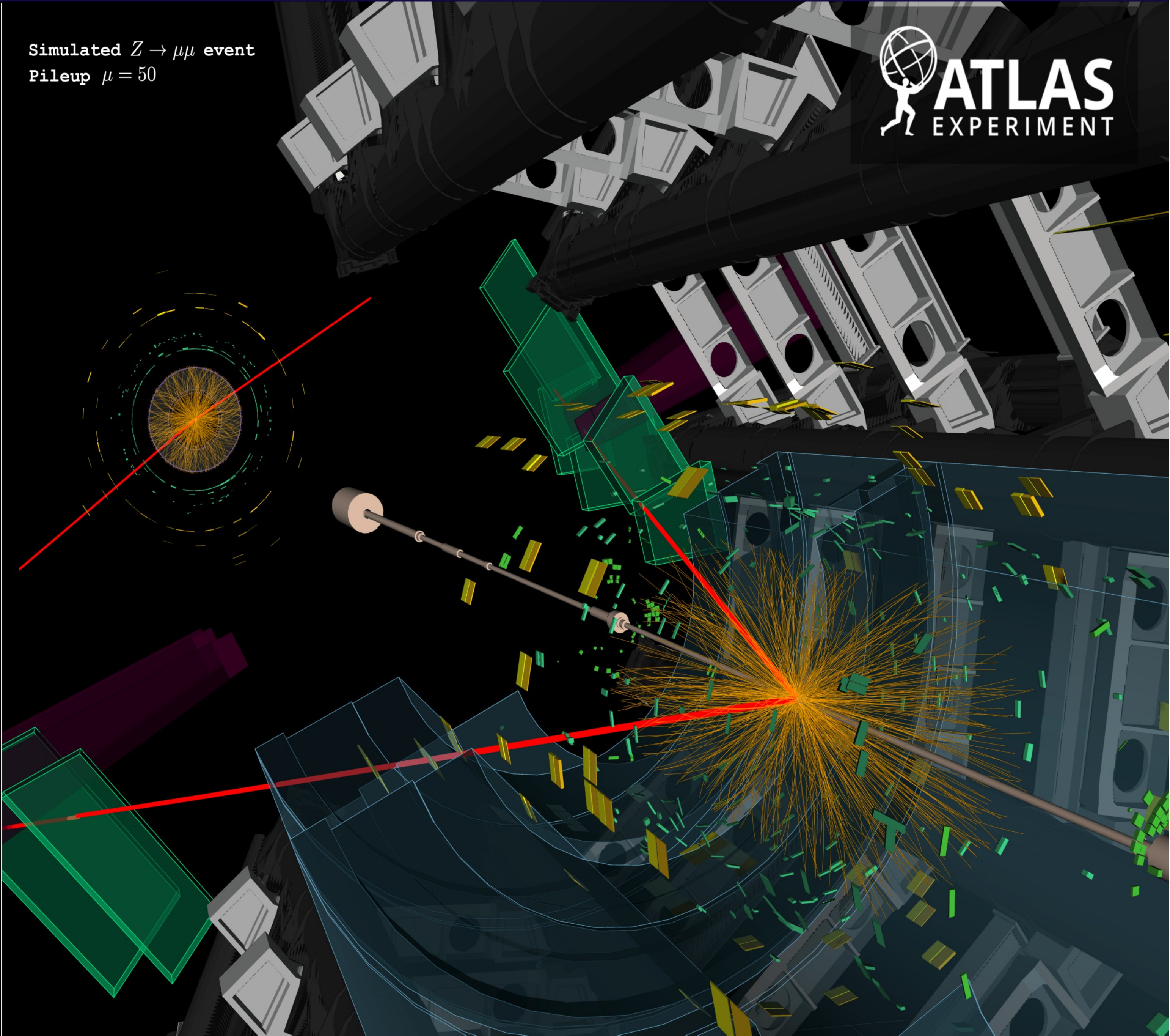


The LHC is a very jetty place

There are many challenges and the large number of additional interactions (pile-up) is one of them

PU treated for jets at different levels:

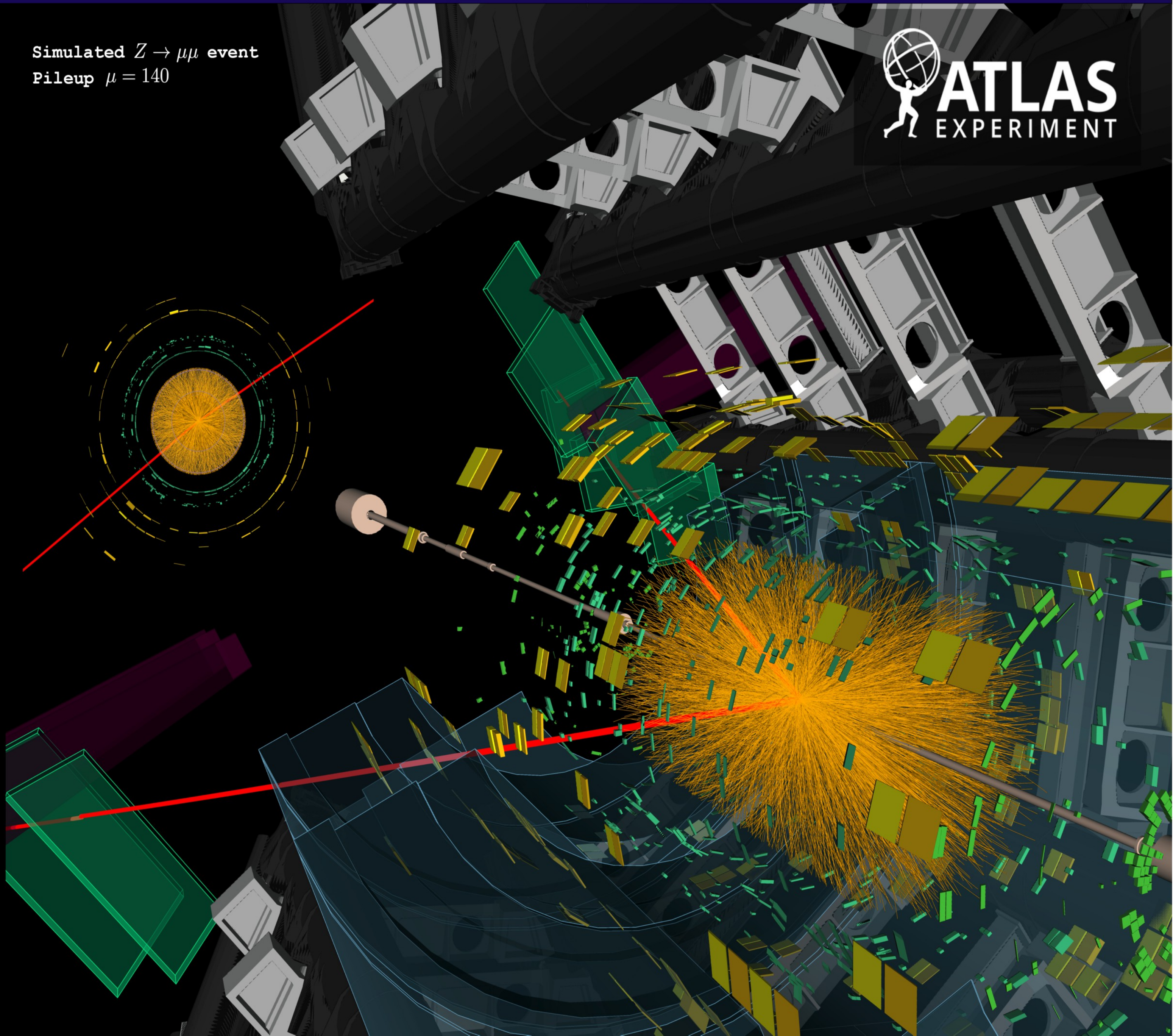
- in the calibration
- at the inputs
- using grooming techniques for large-radius jets



The LHC is a very jetty place

There are many challenges and the large number of additional interactions (pile-up) is one of them

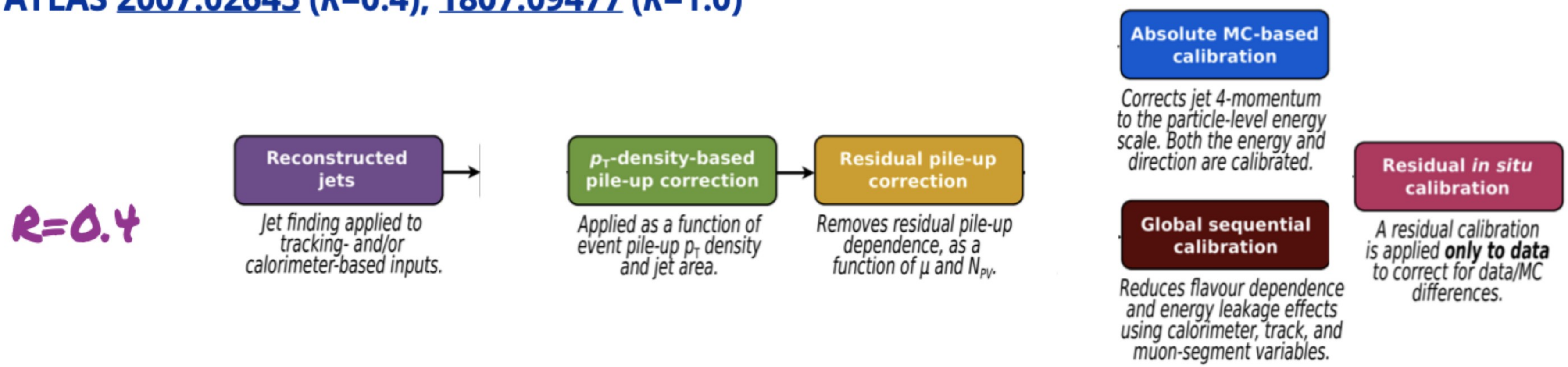
PU treated for jets at different levels:
- in the calibration
- at the inputs
- using grooming techniques for large-radius jets

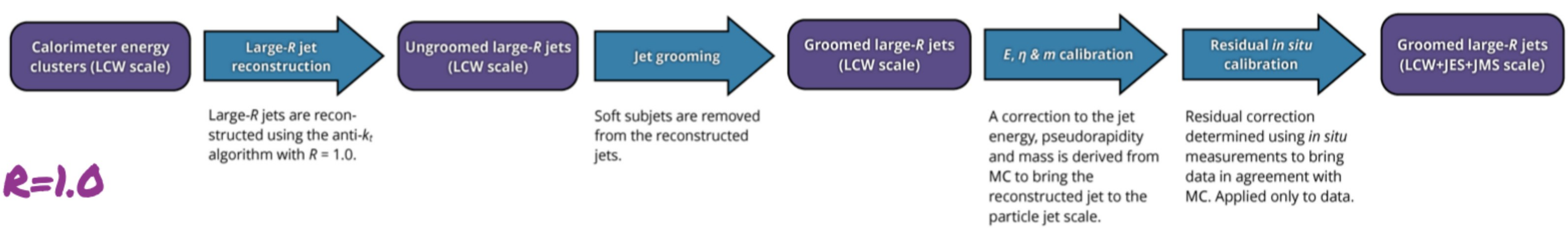


Jet calibration chain

From Matt LeBlanc Semi-Visible Jets Workshop @ ETH Zurich, July 2022

ATLAS 2007.02645 (*R*=0.4), 1807.09477 (*R*=1.0)



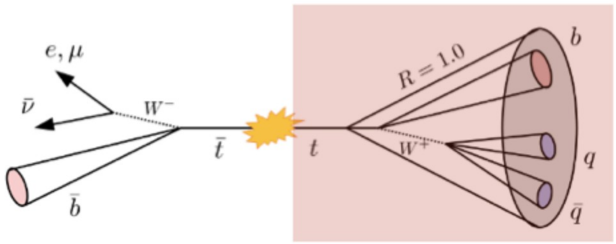


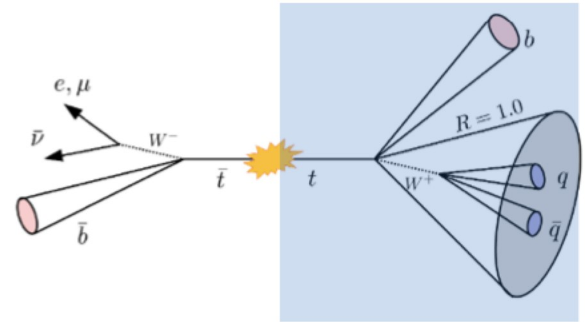
- The **jet calibration sequence** corrects for pile-up, restores $\langle p_T^{reco}/p_T^{true}\rangle = 1$ with MC-based correction, improves the resolution, and then corrects the response in data to match that in MC
- **Many steps which need to be performed sequentially!**

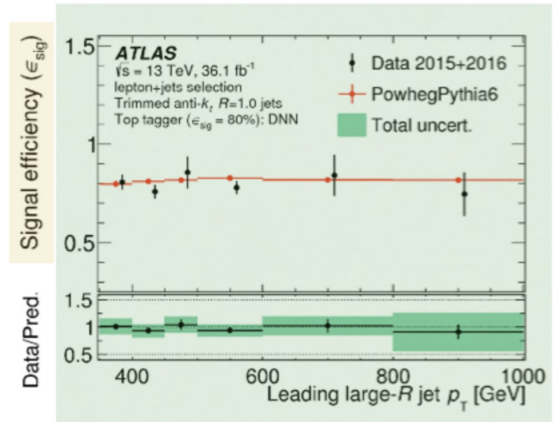
Jet tagging: how do we calibrate the taggers?

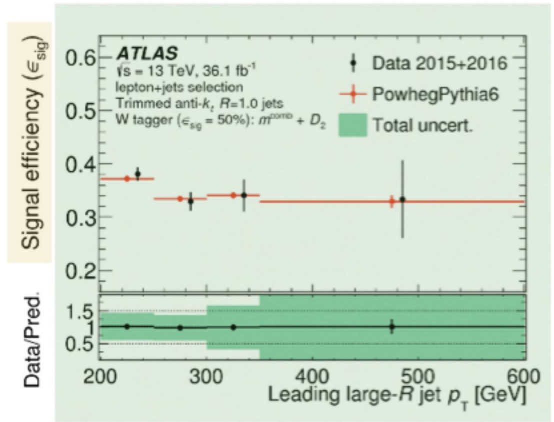
Eur. Phys. J. C 79 (2019) 375

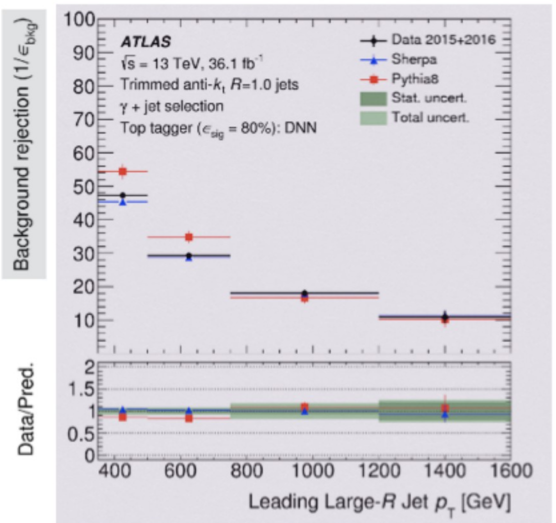
- **Correct the tagger efficiency in MC to match the one in data**
- **For W/top taggers**
  - Primarily look at top-antitop events for signal calibration
  - Dijet and photon+jet samples used for calibrating background rejection
- **For H→bb taggers:**
  - Z/gamma+jets for the signal
  - Top-antitop and g→bb for background

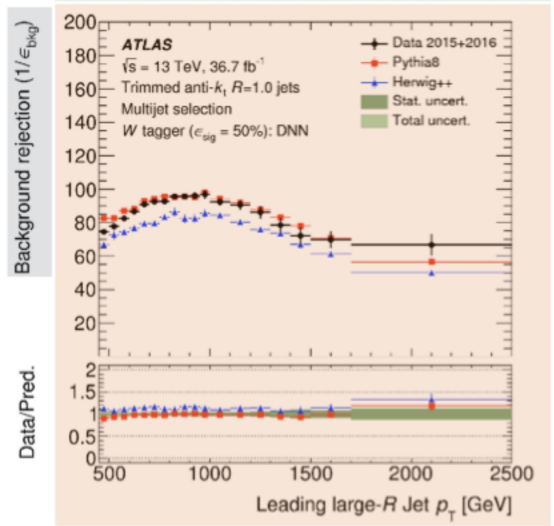












Adapted from Brian Le, ATLAS Flavour Tagging Workshop 2022

Missing transverse momentum

Table 2: Jet selections for the $p_{\mathrm{T}}^{\mathrm{miss}}$ working points used in this study.

	Selections			
	p_{T} [GeV] for jets with:			fJVT for jets with
Working point	$\lvert\eta\rvert < 2.4$	$2.4 < \lvert\eta\rvert < 4.5$	JVT for jets with $\lvert\eta\rvert < 2.4$	$2.5 < \lvert\eta\rvert < 4.5$ and $p_{\mathrm{T}} < 120$ GeV
Loose	> 20	> 20	> 0.5 for $p_{\mathrm{T}} < 60$ GeV jets	-
Tight	> 20	> 30	> 0.5 for $p_{\mathrm{T}} < 60$ GeV jets	< 0.4
Tighter	> 20	> 35	> 0.5 for $p_{\mathrm{T}} < 60$ GeV jets	-
Tenacious	> 20	> 35	> 0.91 for $20 < p_{\mathrm{T}} < 40$ GeV jets > 0.59 for $40 < p_{\mathrm{T}} < 60$ GeV jets > 0.11 for $60 < p_{\mathrm{T}} < 120$ GeV jets	< 0.5

4.1 Input and target features

The NN receives 60 event variables as input features, including:

1. $p_{\mathrm{T}}^{\mathrm{miss}}$ predictions and unique jet- and soft- terms for each working point (see Table 3).
2. Lepton $p_{\mathrm{T}}^{\mathrm{miss}}$ terms (see Table 4). These terms are independent of the working point used.
3. Additional variables which characterise the pile-up and topology of each event (see Table 5).

Input features for both training and testing data are passed through two pre-processing steps:

(1) Rotate each event such that $p_{\mathrm{T}}^{\mathrm{miss,\,Tight}}$ points along the x-axis by construction. This removes ϕ invariance from the inputs, increasing the statistical power of the training data.

(2) Standardise each input and output variable by subtracting the mean and dividing by the standard deviation. This is standard practice for deep learning regression problems.

Table 3: Input variables used for each of the *Tight*, *Tighter*, *Loose*, and *Tenacious* $p_{\mathrm{T}}^{\mathrm{miss}}$ working points. Note for the *Tight* working point the four soft $p_{\mathrm{T}}^{\mathrm{miss}}$ variables as well as $p_{\mathrm{x}}^{\mathrm{miss}}$ and $p_{\mathrm{y}}^{\mathrm{miss}}$ are excluded.

$p_{\mathrm{T}}^{\mathrm{miss}}$	$p_{\mathrm{x}}^{\mathrm{miss}}$	$p_{\mathrm{y}}^{\mathrm{miss}}$	$\sum p_{\mathrm{T}}$
$p_{\mathrm{T}}^{\mathrm{miss,\,jet}}$	$p_{\mathrm{x}}^{\mathrm{miss,\,jet}}$	$p_{\mathrm{y}}^{\mathrm{miss,\,jet}}$	$\sum p_{\mathrm{T}}^{\mathrm{jet}}$
$p_{\mathrm{T}}^{\mathrm{miss,\,soft}}$	$p_{\mathrm{x}}^{\mathrm{miss,\,soft}}$	$p_{\mathrm{y}}^{\mathrm{miss,\,soft}}$	$\sum p_{\mathrm{T}}^{\mathrm{soft}}$

Table 4: Lepton $p_{\mathrm{T}}^{\mathrm{miss}}$ term input variables.

$p_{\mathrm{T}}^{\mathrm{miss},\,e}$	$p_{\mathrm{x}}^{\mathrm{miss},\,e}$	$p_{\mathrm{y}}^{\mathrm{miss},\,e}$	$\sum p_{\mathrm{T}}^{e}$
$p_{\mathrm{T}}^{\mathrm{miss},\,\mu}$	$p_{\mathrm{x}}^{\mathrm{miss},\,\mu}$	$p_{\mathrm{y}}^{\mathrm{miss},\,\mu}$	$\sum p_{\mathrm{T}}^{\mu}$

Table 5: Additional input variables.

$<\mu>$	Mean number of interactions per bunch crossing
N_{PV}	Number of primary vertices with at least 2 associated tracks
N_{PV^2}	Number of primary vertices with at least 2 associated tracks, excluding the hard-scatter vertex[5]
N_{PV^4}	Number of primary vertices with at least 4 associated tracks
N_{trk}	Number of ID tracks associated with the primary vertex
N_e^b	Number of baseline electrons
N_μ^b	Number of baseline muons
N_e	Number of signal electrons
N_μ	Number of signal muons
N_J	Number of signal jets

ATLAS calorimeters

ATLAS Calorimeters

- EM: $|\eta| < 3.2$,
  - Pb/LAr calorimeter,
  - 22-26 X_o, 1.2 λ,
  - 3 longitudinal sections,
  - $\Delta\eta \times \Delta\Phi = 0.025 \times 0.025 - 0.1 \times 0.1$
  - $\sigma/E \simeq 10\%/\sqrt{E}$.
- Central Hadronic: $|\eta| < 1.7$,
  - Fe/Scintillator sampling calorimeter
  - 7.4 λ,
  - 3 longitudinal sections,
  - $\Delta\eta \times \Delta\Phi = 0.1 \times 0.1 - 0.2 \times 0.1$,
  - $\sigma/E \simeq 50\%/\sqrt{E} \oplus 0.03$.
- EndCap Hadronic: $1.7 < |\eta| < 3.2$,
  - Cu/LAr sampling calorimeter,
  - 4 longitudinal sections,
  - $\Delta\eta \times \Delta\Phi = 0.1 \times 0.1 - 0.2 \times 0.2$
- FCAL: $3 < |\eta| < 4.9$,
  - EM: Cu/LAr, HAD: W/LAr calorimeter,
  - 10 λ,
  - 1 EM + 2 HAD longitudinal sections,
  - $\Delta\eta \times \Delta\Phi = 0.75 \times 0.65 - 5.4 \times 4.7$