Tagging Boosted Jets with Machine Learning Techniques

Francesco Carnevali
INFN & Università degli Studi di Napoli Federico II

On behalf of the ATLAS & CMS Collaborations

QCD@LHC 2022

28 Nov - 2 Dec 2022

Francesco Carnevali

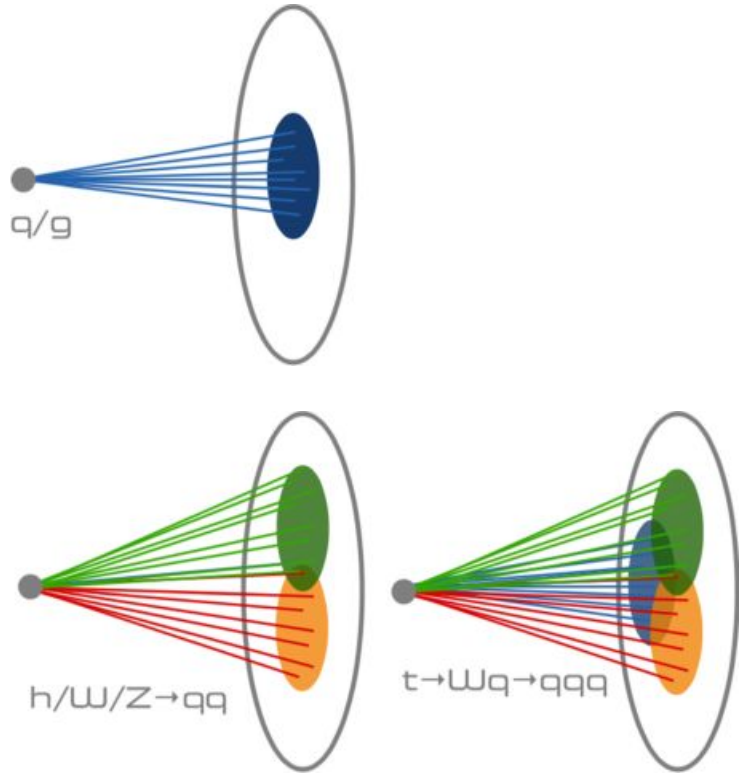
Introduction

- Jets are produced via hadronization of quarks and gluon.

- Heavy SM objects (W/Z/H/top) can decay through multiple quarks, resulting into one or more jets depending on the lorentz boost.

- Discriminating which particle a jet comes from is a key point for many measurements and searches at the Large Hadron Collider (LHC) experiments.

- Different Machine Learning algorithms have been developed in the ATLAS and CMS experiments for boosted jet tagging, taking into account jets substructures and its constituents.



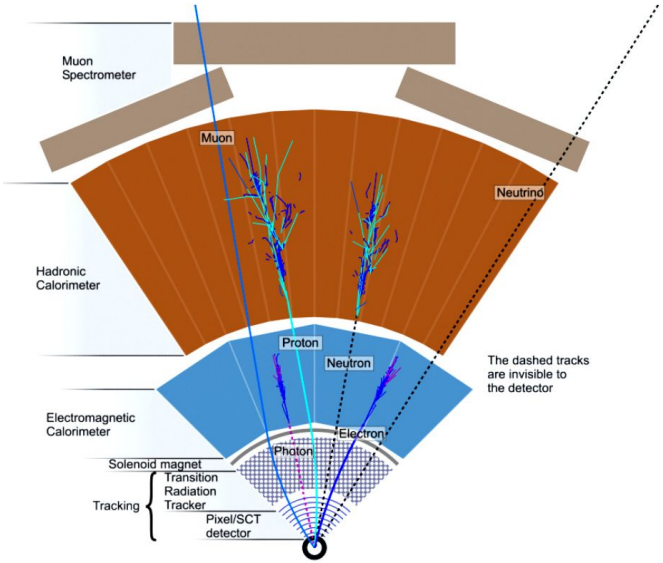
Francesco Carnevali

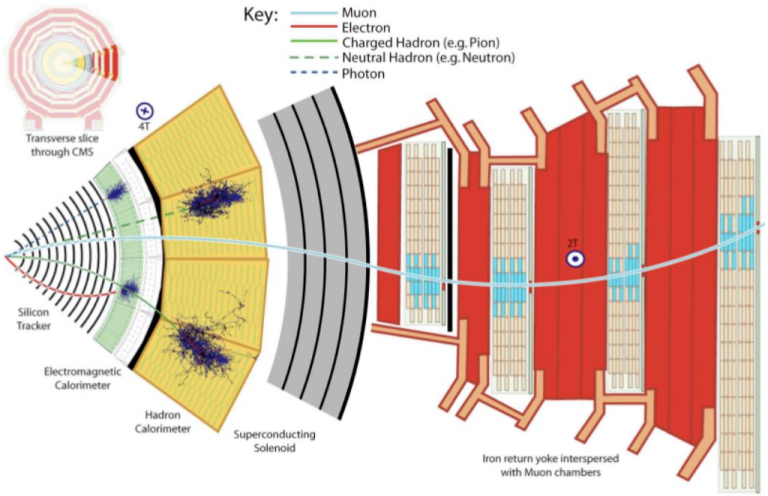
Jet Reconstruction at ATLAS and CMS experiments

- ATLAS and CMS are "general purpose" experiments at LHC: different subdetectors to reconstruct and identify particles produced in proton-proton collisions.

- Clusters of color-neutral hadrons arise from the fragmentation and hadronization of partons.

- Stable particles are detected showering and releasing energy in the hadronic calorimeters.





Francesco Carnevali

Anti-k_T Algorithm

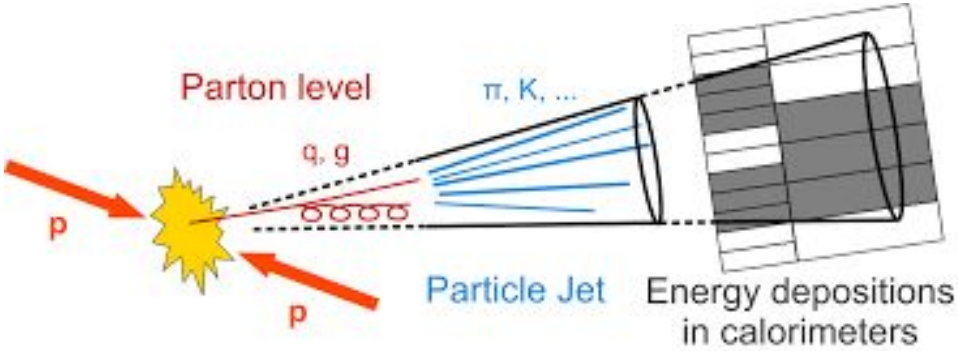
- The Anti-k_T (AK) algorithm is a powerful tool to cluster the hadronic shower in the calorimeters, generated by the constituents of the jets.

- The R parameters of the AK algorithm defines the size of the jet in the (φ,η) metrics.

- To represent more complex object, like W/Z/H boson or top quark, "large-R jets" are used:

  - R = 1.0, ATLAS

  - R= 0.8, CMS



This algorithm considers list of reconstructed final state objects for the clustering.

An element *i* and the pseudojet *j* are combined into a new pseudojet if the distance to the pseudojet axis (d_{ij}) is lower than the distance to the beam (d_{iB})

$$d_{ij} = min\left(\frac{1}{p_{T,i}^2}, \frac{1}{p_{T,j}^2}\right)\frac{(\Delta R_{ij})^2}{R^2}$$

$$d_{i,B} = \frac{1}{p_{T,i}^2}$$

Francesco Carnevali

Algorithms Described

- Boosted Event Shape Tagger:
https://journals.aps.org/prd/abstract/10.1103/PhysRevD.94.094027

- DeepAK8 tagger:
https://iopscience.iop.org/article/10.1088/1748-0221/15/06/P06005

- W/Z tagger with UFO Jets:
https://cds.cern.ch/record/2777009

- Energy Flow Network tagger:
https://link.springer.com/article/10.1007/JHEP01(2019)121

- ParticleNet tagger:
https://journals.aps.org/prd/abstract/10.1103/PhysRevD.101.056019
https://cds.cern.ch/record/2825328

For a talk on the analyses using tagging techniques also see the *"BSM searches with jet substructure"* contribution from Antimo

Francesco Carnevali

Boosted Event Shape Tagger (BEST)

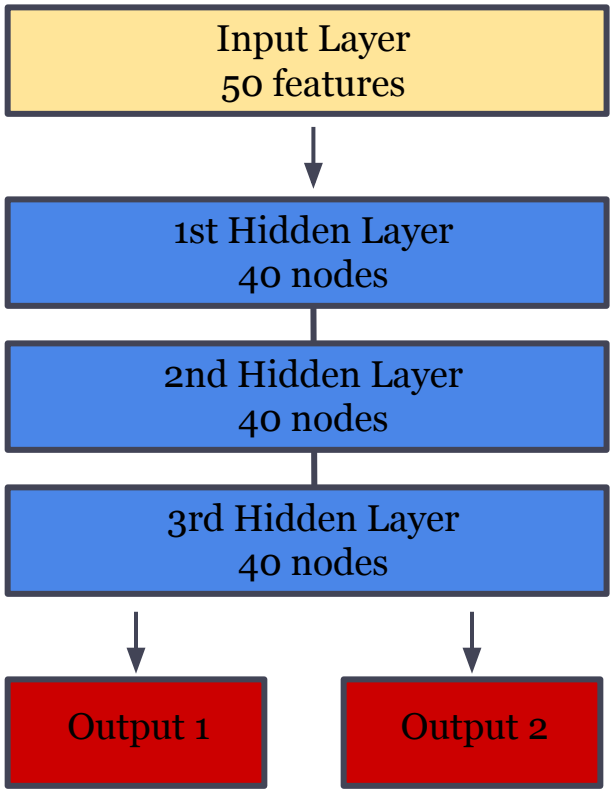
- The BEST algorithm is a MultiLayerPerceptron (MLP).

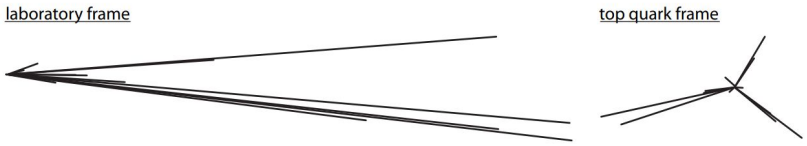
- Two output multiclassifier to distinguish between jets from b quark, W boson, and top quarks:

  1. b quark jet -> (0,0)
  2. W boson jet -> (0,1)
  3. t quark jet -> (1,0)

- Input features obtained boosting the jet constituents in 4 different framework, assuming for the boosting vector reconstructed transverse momentum and energy, and mass of W/Z/H/t .

- Event shape parameters, i.e. thrust, in the 4 different frameworks (boosted event shape) are associated to the 4 leading subjets.





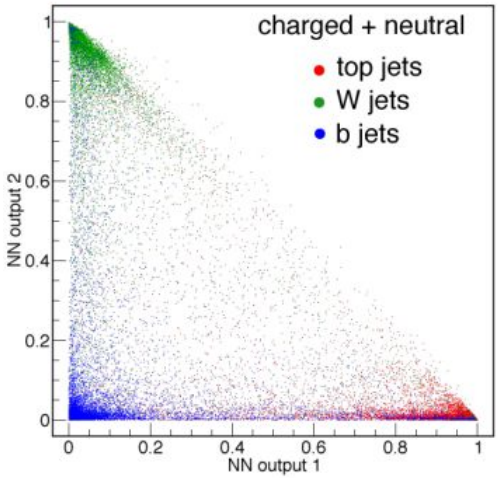
Francesco Carnevali

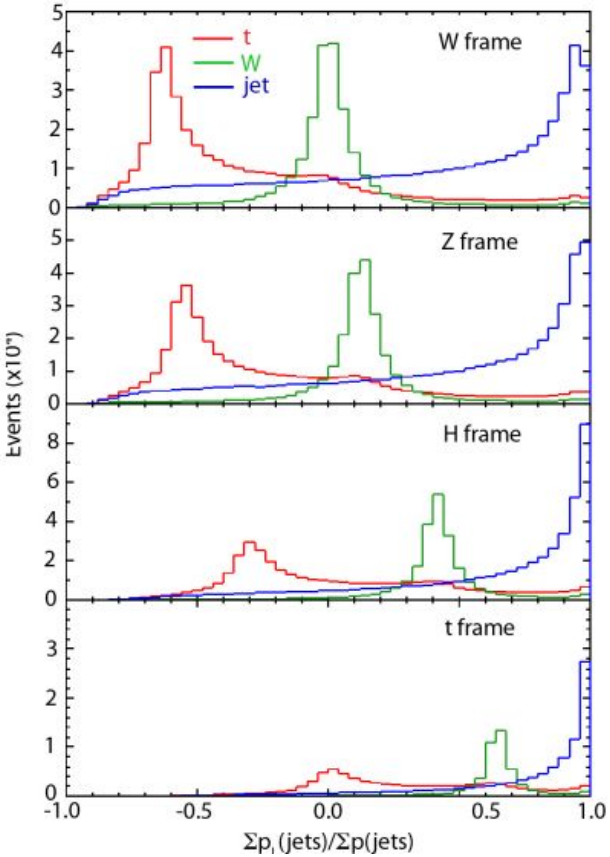
Boosted Event Shape Tagger (BEST)

- The BES features and kinematic variables in the boosted frameworks allow to increase the accuracy of the algorithm.

- Classification probability of the algorithm performed on a Z'(3 TeV) sample:

Generated:	t	W	b
Classified as t	84.2%	3.9%	10.2%
Classified as W	5.6%	86.8%	5.9%
Classified as b	10.4%	9.3%	83.9%





- A second version of the algorithm with 6 hiddens layers and 4 output has been used to perform also H and Z tagging

Francesco Carnevali

DeepAK8 Tagger

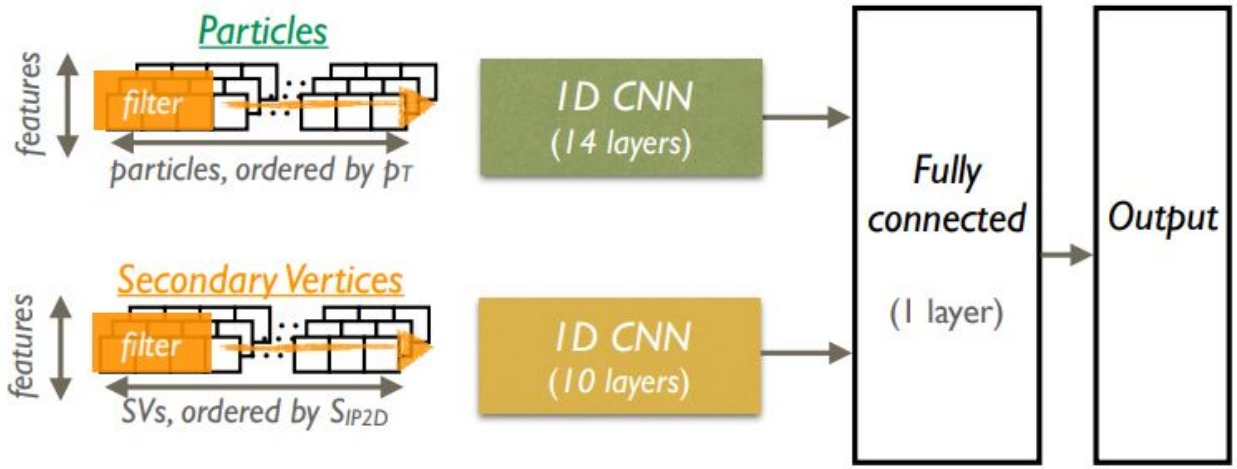
- The DeepAK8 algorithm is a multiclassifier to discriminate jets from W/Z/H/t/other, the main classes are subdivided to minor categories, e.g. Z->bb, Z->cc.

- The algorithm works in two different steps, in the first step constituents of the jets and SVs are processed separately with 1D CNNs.

- Two list of input for each jet:

1. First 100 jet constituents, order by decreasing transverse momentum, 15 features for each particle

2. Up to 7 SecondaryVertexes (SVs), ordered by 2D secondary ImpactParameter significance, 15 features for each SV



- **512 input of the 1 fully connected layer**

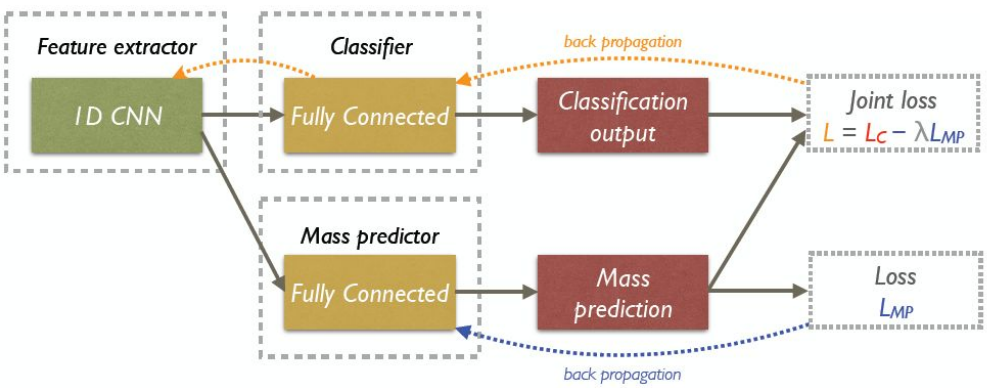
- **20% of Dropout to reduce the overtraining.**

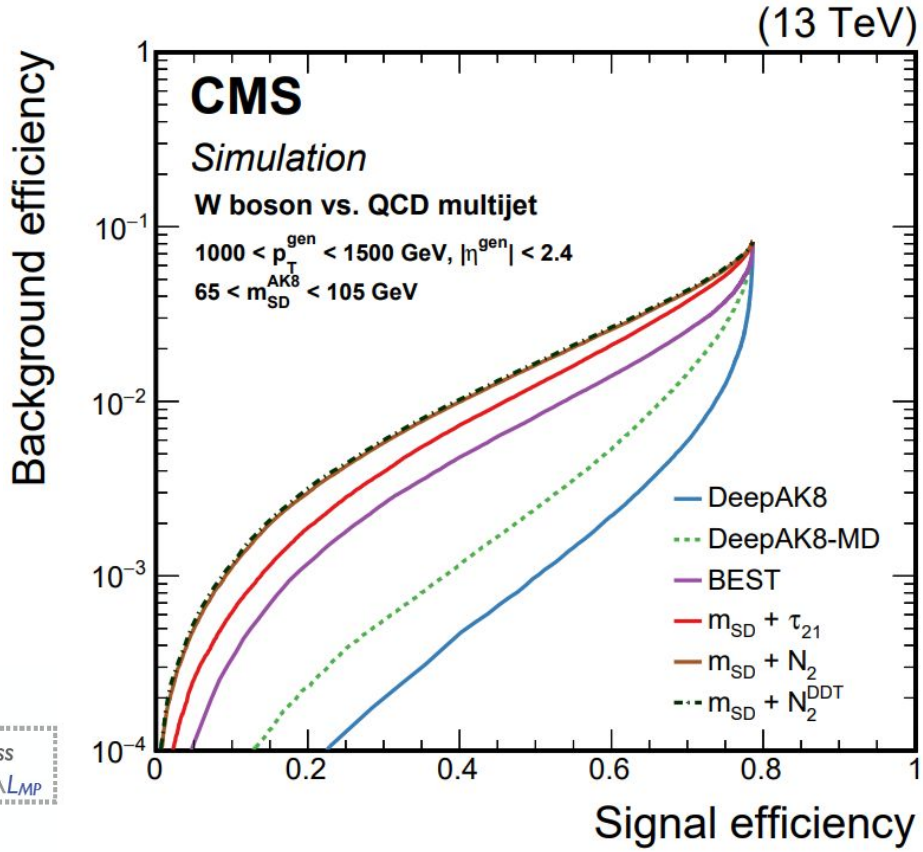
- **Cross-entropy loss minimized with Adam Optimizer**

Francesco Carnevali

DeepAK8 Tagger

- The DeepAK8 tagger provides improvements in boosted jet tagging.
- A second version of the tagger, **DeepAK8-MD**, to decorrelate the tagger output to the mass of the jet.
- The loss function from the mass predictor works as a penalty for the Classification tagger.





Francesco Carnevali

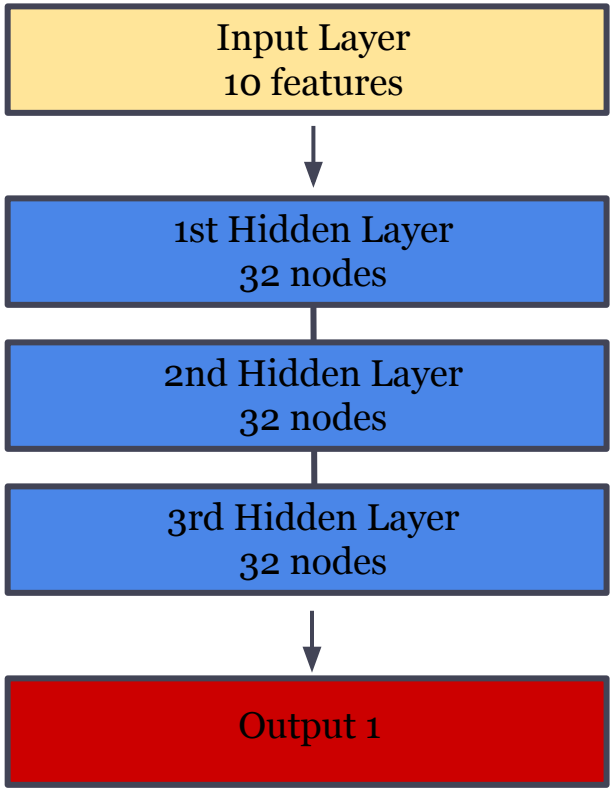
W/Z Tagger with UFO Jets in ATLAS

- To separate W jets from background jets the ATLAS experiment has developed a DeepNeuralNetwork (DNN) algorithm.

- High-level-quantities describing the jets substructure are chosen as input. “Unified Flow Objects” (UFO) Jet are reconstructed, thanks to a new particle-flow algorithm.

- Adam optimizer is used to minimize the binary cross-entropy loss function.

- Improving the DNN algorithm as well as the jet reconstruction has lead to decrease the background rejection by a factor of 2 to 4, depending on the transverse momentum of the jet and on the signal efficiency.



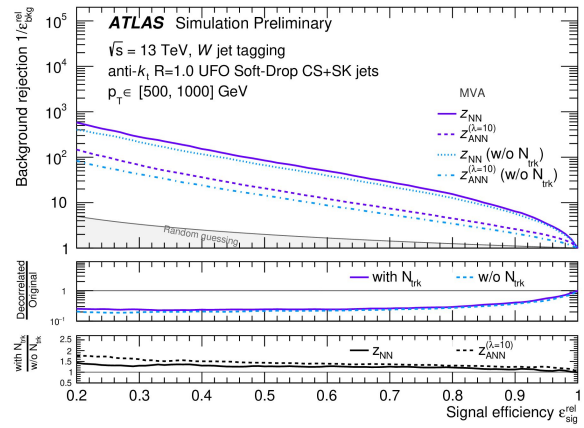
Francesco Carnevali

W/Z Tagger with UFO Jets in ATLAS

- Two different approaches have been used to decorrelate the tagger performance from the jet mass:
  1. Analytical mass decorrelation using a fixed-efficiency regression.
  2. Adversarial Neural Network (ANN).

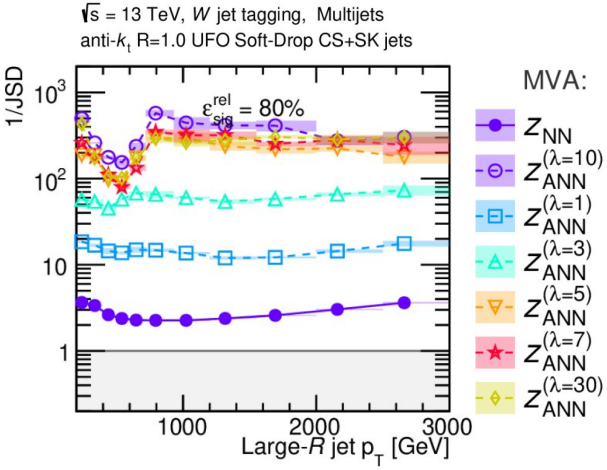
- In the second approach an adversary network is trained to compete with the DNN tagger. The goal of the ANN is to infer the jet mass from the DNN score.

- The ANN algorithm is a single fully connected layer with 64 nodes and a ReLU function as activation function. It is connected to an encoded Gaussian Mixture Model with 20 components and a single node output with a sigmoid activation function. The Jensen-Shannon divergence (JSD) evaluates the relative entropy between the jet mass distribution before and after the tagging.



Total Loss function:

$$\min_{\theta_{\text{clf}}} \max_{\theta_{\text{adv}}} L_{\text{clf}}(\theta_{\text{clf}}) - \lambda L_{\text{adv}}(\theta_{\text{clf}}, \theta_{\text{adv}})$$



Francesco Carnevali

Energy Flow Networks Tagger

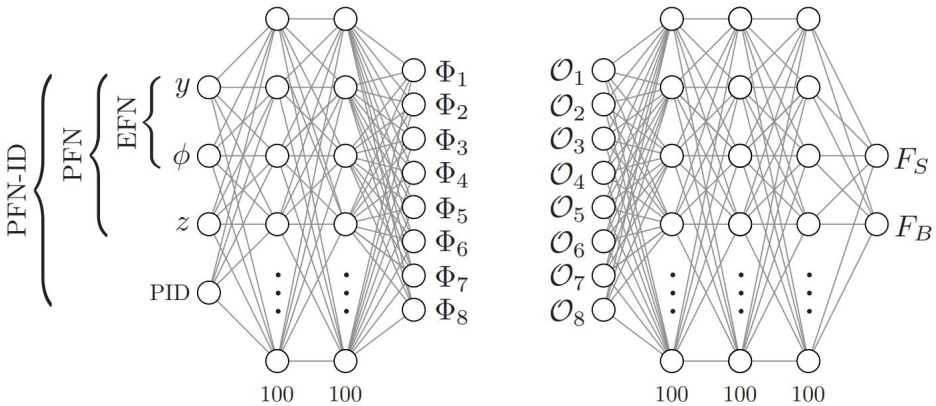
- The Energy Flow Network (EFN) is a top tagging algorithm. It uses DeepSets structure to represent the jet, ensuring invariance w.r.t. network inputs.

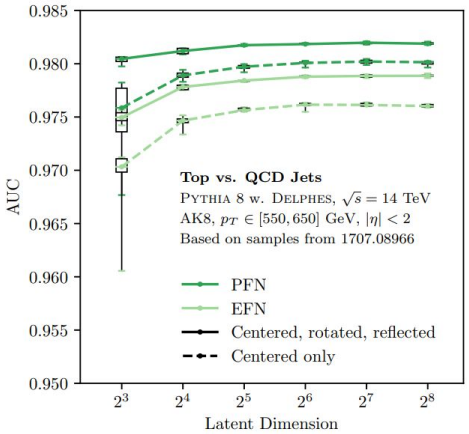
- The EFN uses constituent's angular coordinates and the logarithm of transverse momentum as input.

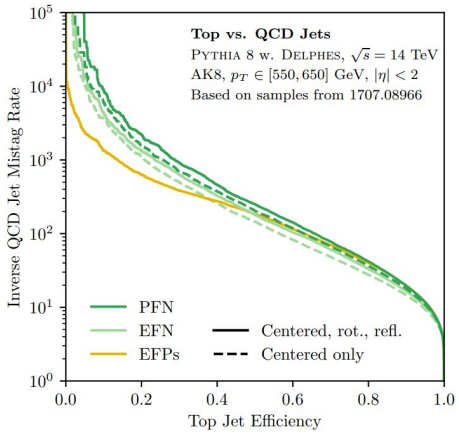
- The input is mapped in a l-dimensional latent space with a first DNN.

$$\mathcal{O}_a = \sum_i z_i \, \Phi_a(y_i, \phi_i)$$

- A second DNN is used to discriminate signal from background, its input is made up observables, function of the first algorithm's output:







ATLAS EFN performance
AUC = 0.901

Francesco Carnevali

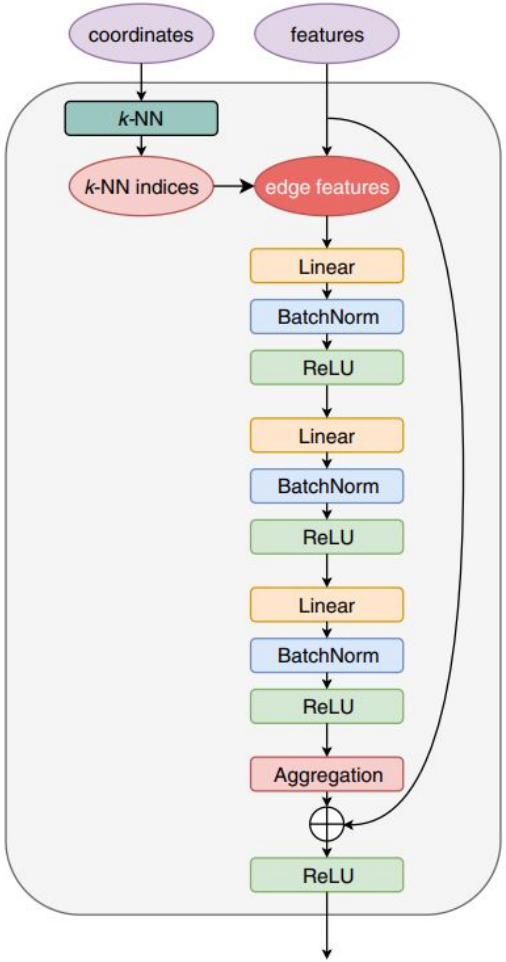
ParticleNet Tagger

- The ParticleNet algorithm is a Dynamic Graph Convolutional Neural Network.

- Mapping jet as image, made of energetic release of its constituents into the calorimeter, lead to a great number of pixels, many of which are empty.

- The jet can be seen as a **ParticleCloud**, unordered set of entities irregularly distributed. The elements are correlated and with a rich internal structure, taking into account non additive features.

- **EdgeConvolutional** layers are invariant per permutations of the point and the output combines the feature of each point to its **k-NearestNeighbours** ones.



Francesco Carnevali

ParticleNet Tagger

- Each EdgeConvulutional layer output is performed by a MLP algorithm.

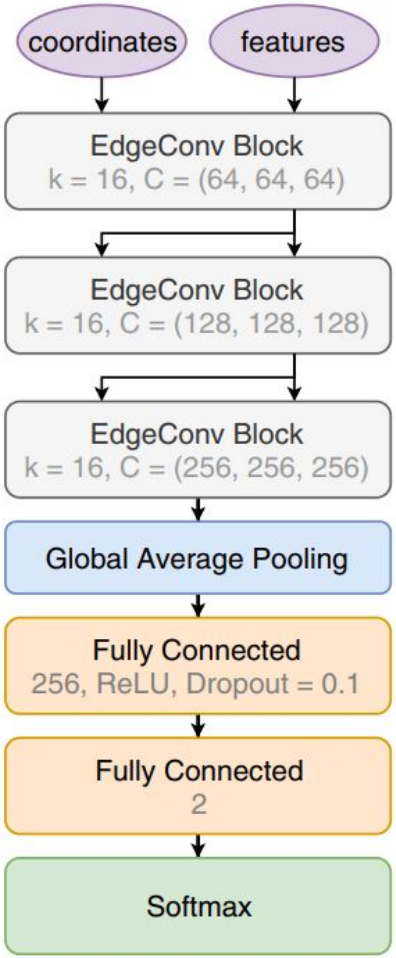
- Training is done with a batch size of 384 jets in order to reduce parametrization costs.

- Cross-entropy loss function is minimized with AdamW optimizer, with one cycle LearningRate:

  1. Initial LR = $3 \cdot 10^{-4}$
  2. Linearly increase for 8 epochs up to $3 \cdot 10^{-3}$
  3. Linearly decrease for 8 epochs to initial LR
  4. LR Cool Down for 4 epochs to $5 \cdot 10^{-7}$

- Input of 100 leading continents of the of the jet kinematic and ParticleId variables.

- (φ,η) metrics used for the NearestNeighbour.

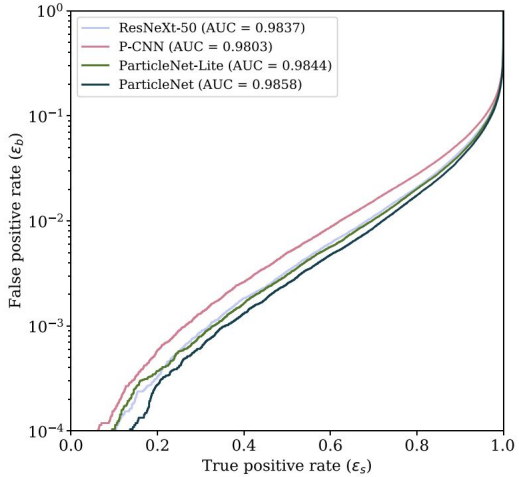


Francesco Carnevali

ParticleNet Tagger

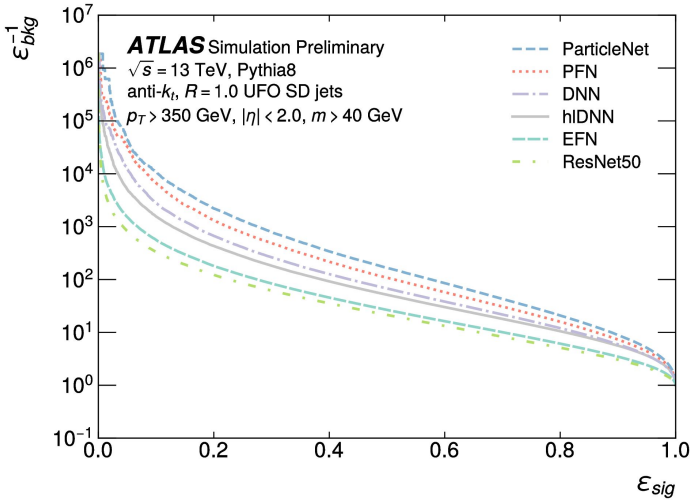
- The ParticleNet-Lite algorithm has been developed reducing the number of the EdgeConvolutional layers.

- Top tagger discriminate top quark from QCD dijet AK8 jet with an **accuracy of 0.94**.



- ATLAS experiment shows that the ParticleNet tagger has the strongest background rejection w.r.t to the other taggers with an **accuracy of 0.894** with jet of R=1.

- Moreover taggers using jets constituents variable performs better than high-level-quantity based taggers, i.e. ResNet 50.

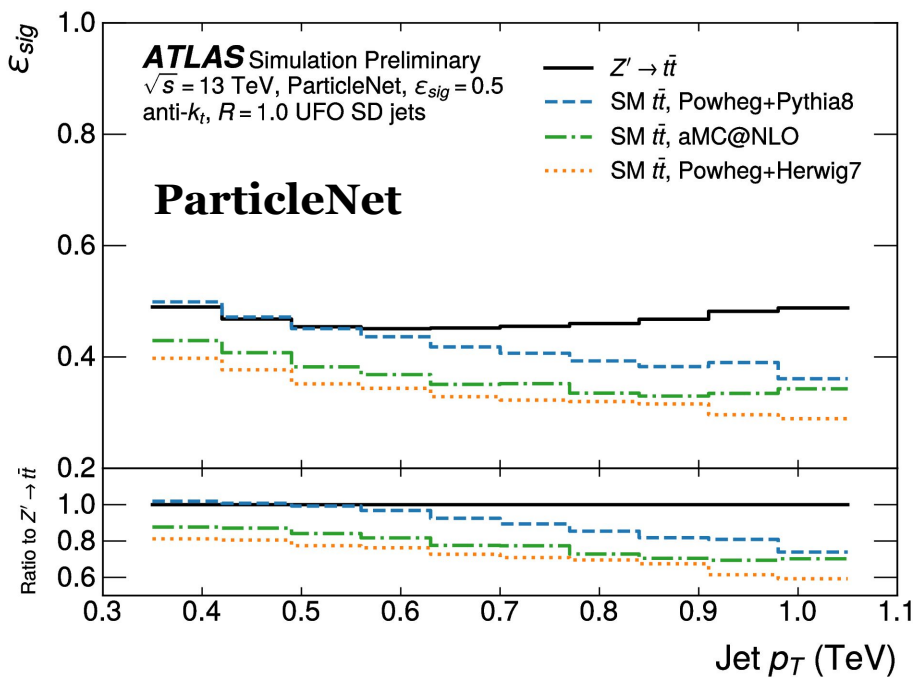


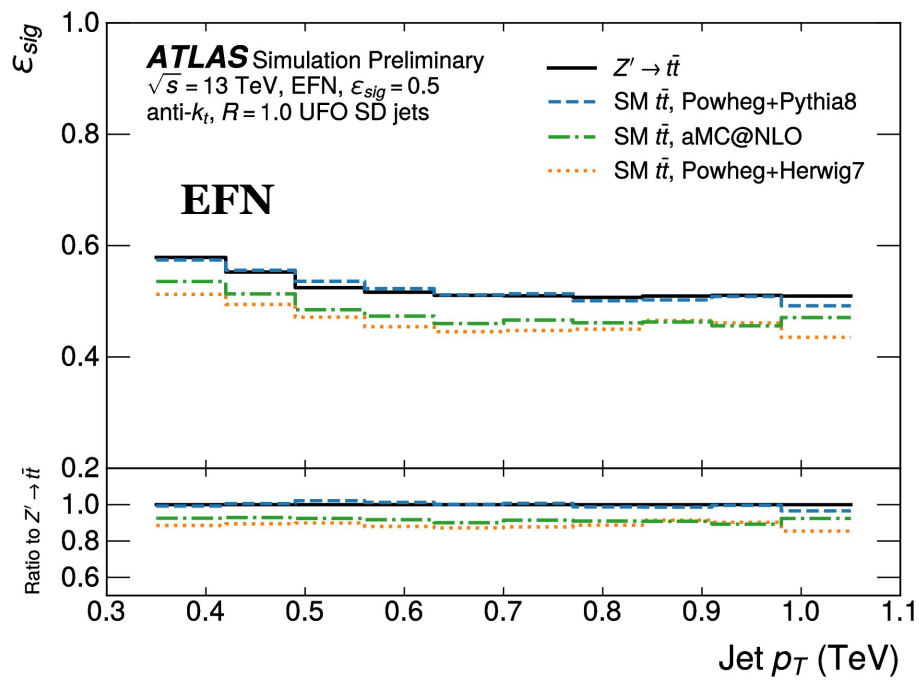
Francesco Carnevali

ParticleNet Tagger

- AUC is not the only important metric!
- Low-level-quantity based taggers are more sensitive to the differences between the simulated processes.

Model	AUC
ResNet 50	0.885
EFN	0.901
hlDNN	0.938
DNN	0.942
PFN	0.954
ParticleNet	0.961





Conclusions

- Discriminating which particle a jet comes from is a key point for many measurements and searches at the LHC experiments.

- Different Machine Learning algorithms have been developed in the ATLAS and CMS experiments for boosted jet tagging, taking into account jets substructures and its constituents.

- Improving jet reconstruction and representation, e.g. image jet or particle cloud, to take into account jet component and patterns in jets substructure can improve the signal efficiency of the taggers.

- Taggers mass decorrelation achieved by using different methods, for example with an ANN.

- Taggers using jets constituents variable show better performance better than high-level-quantity based taggers, however the first ones are more sensitive to the differences between the simulated processes.

THANK YOU!

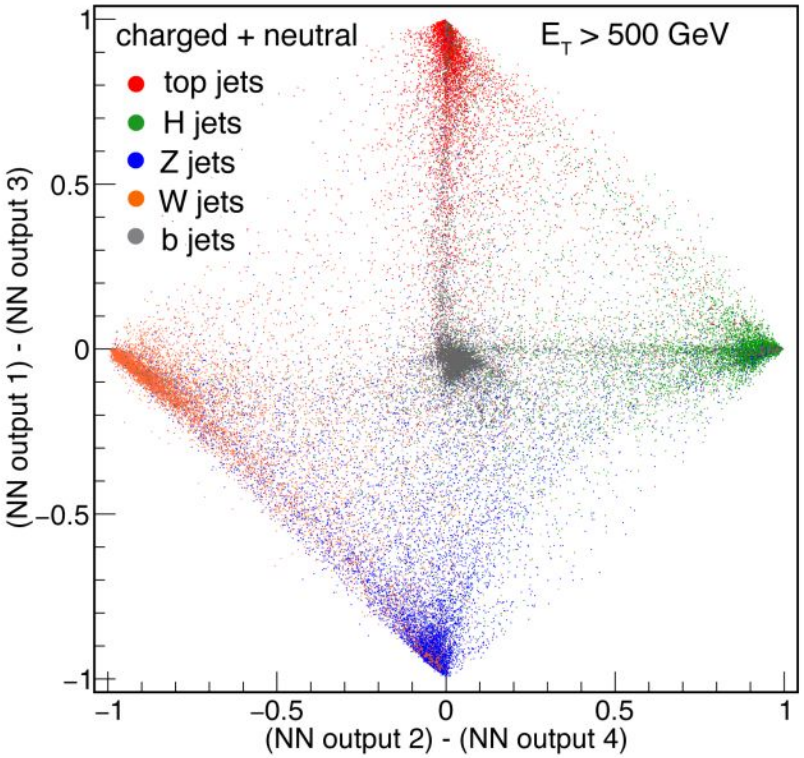
Francesco Carnevali

Boosted Event Shape Tagger (BEST)

- A second version of the algorithm with 6 hiddens layers and 4 output has been used to perform also H and Z tagging

- Four output multiclassifier to distinguish between:

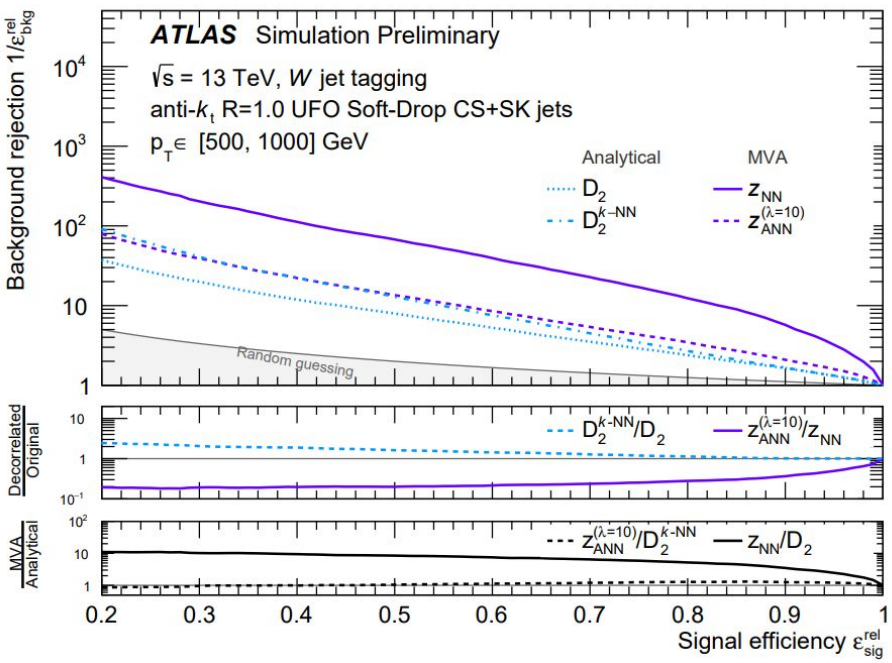
  1. b quark jet -> (0,0,0,0)
  2. W boson jet -> (0,0,0,1)
  3. Z boson jet -> (0,0,1,0)
  4. H boson jet -> (0,1,0,0)
  5. t quark jet -> (1,0,0,0)

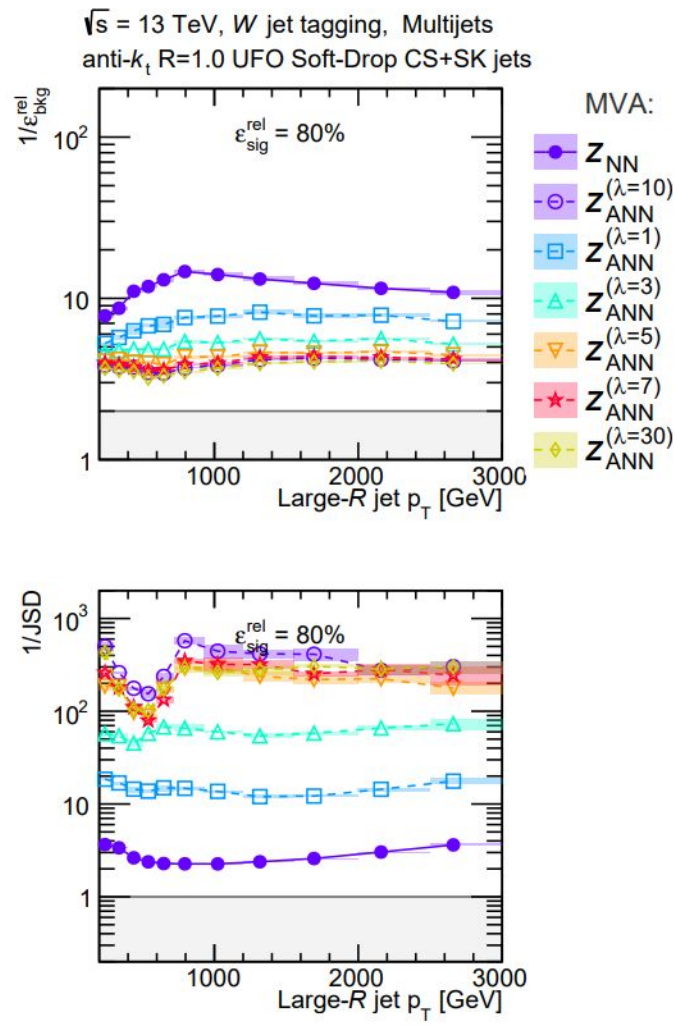


Francesco Carnevali

W/Z Tagger with UFO Jets in ATLAS

- Higher value of λ gives a better decorrelation but worse background rejection.
- D2 2-prong discriminant related to the energy correlation functions.





Francesco Carnevali

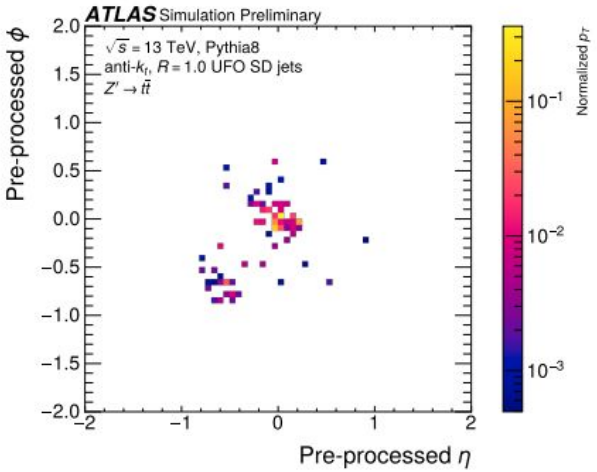
ResNet 50 Tagger

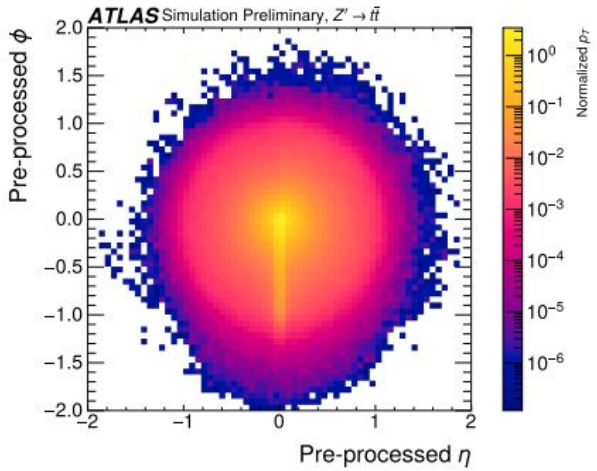
- The ResNet 50 algorithm is a large scale CNN for image classification, trained for top tagging.

- Each jet is mapped in a 64x64 pixels image in the (φ,η) space. Each pixel value is the sum of transverse momentum of the jet constituents normalized on the total transverse momentum of the jet.

- In order to take into account the lower transverse momentum patterns of the substructure the pixels intensities are prescaled by log(1+100x).

- The standard ResNet architecture has been modified to take into account "one color" image jet structure and the binary classification task for top tagging in the ATLAS experiment.





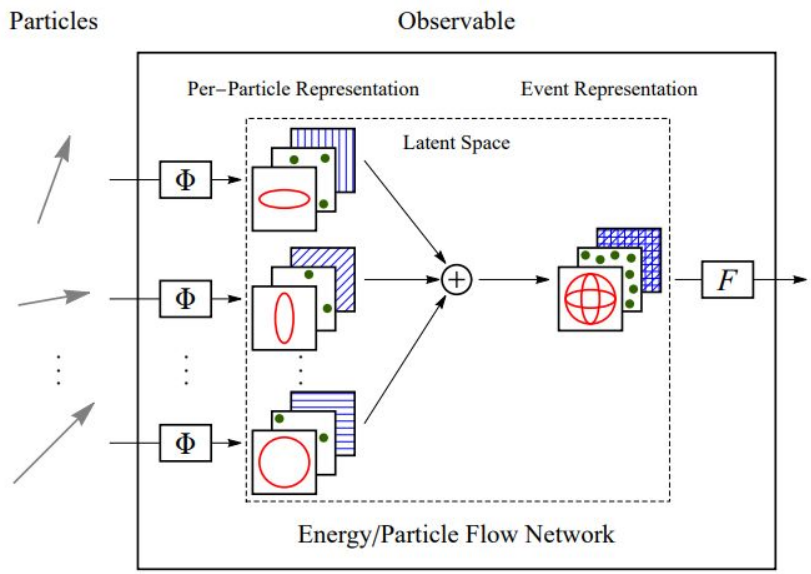
Model	AUC	ACC
ResNet 50	0.885	0.803

Francesco Carnevali

Energy Flow Networks Tagger

- The Energy Flow Network (EFN) is a top tagging algorithm. It uses DeepSets structure to represent the jet, ensuring invariance w.r.t. network inputs.

- The EFN uses constituent's angular coordinates and the logarithm of transverse momentum as input.



EFN Tagger:

$$\mathcal{O}_a = \sum_i z_i \, \Phi_a(y_i, \phi_i)$$

PFN-ID Tagger:

$$\mathcal{O}_a = \sum_i \Phi_a(y_i, \phi_i, z_i, \mathrm{PID}_i)$$