# Classify the Higgs decays with PFN and ParticleNet at CEPC & Plan for ATLAS boson tagger with ParticleNet

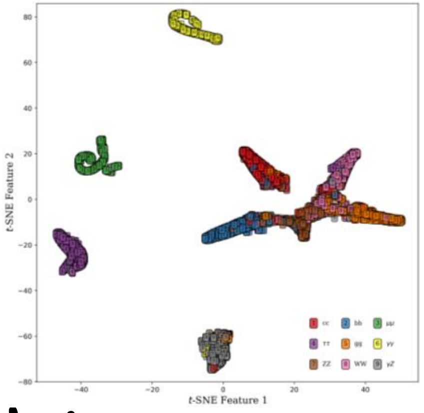
Shudong WANG

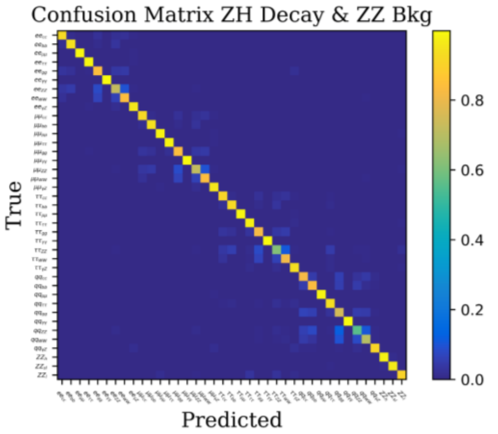
Experimental Physics Division

Institute of High Energy Physics, CAS

# Outline

- **Classify the Higgs decays with PFN and ParticleNet at CEPC**
  - A Brief Introduction on PFN and ParticleNet
  - ML model setup & Simulation samples
  - 9-category & 39-category classification results




- **Plan for ATLAS boson tagger with ParticleNet**
  - Motivation
  - Plan for boson tagger
  - Current situation & Expected performance

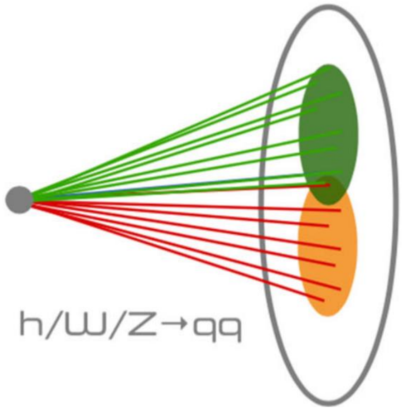



Image credit

# Classify the Higgs decays with PFN and ParticleNet at CEPC

Chinese Physics C > In Press > Article

**Classify the Higgs decays with the PFN and ParticleNet at electron–positron colliders**

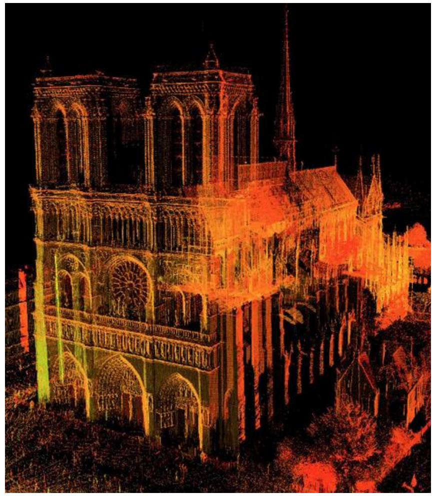
Gang Li [1,✉], Libo Liao [2,✉], Xinchou Lou [1], Peixun Shen [1], Weimin Song [3], Shudong Wang [1,4,✉], Zhaoling Zhang [3]

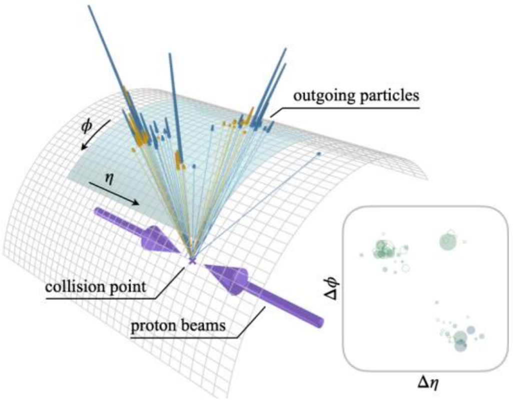
*link*

# A Brief Introduction on PFN and ParticleNet

- **Jet/Event as a point cloud**

This page is excerpted from H. Qu's talk






## Point cloud

From Wikipedia, the free encyclopedia

A **point cloud** is a set of data points in space. The points may represent a 3D shape or object. Each point position has its set of Cartesian coordinates (X, Y, Z).[1] Point clouds are generally produced by 3D scanners or by photogrammetry software, which measure many points on the external surfaces of objects around them. As the output of 3D

## Jet (Particle cloud)

From Wikipedia, the free encyclopedia

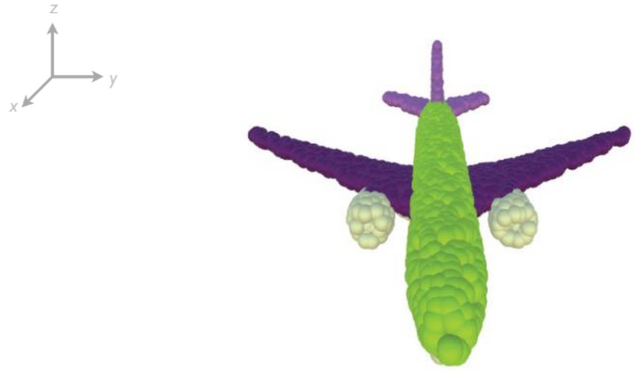
A **jet (particle cloud)** is a set of particles in space. Particle clouds are generally created by clustering a large number of particles measured by particle detectors, e.g., ATLAS and CMS. which measure

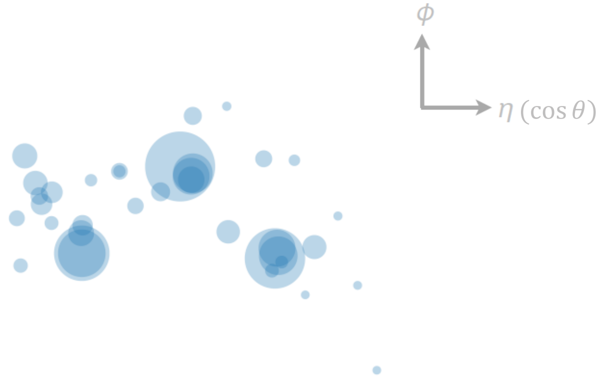
*Jet: An **unordered**, **variable length** collection of particles*

# A Brief Introduction on PFN and ParticleNet

- **Point clouds VS Particle clouds**

This page is excerpted from H. Qu's talk




## Point cloud

- points are intrinsically ***unordered***
- points are distributed in space
  - spatial coordinates (3D xyz) encode geometric structure information

## Particle cloud

- particles are intrinsically ***unordered***
- particles are distributed in space
  - spatial distribution (2D coordinates in the η(cos θ)-φ space) reflects radiation patterns

***But particles have more features:***

- energy/momenta/displacement/particle ID/etc.
- more interesting than a plain point cloud!

# A Brief Introduction on PFN and ParticleNet

- **The architecture of PFN (Particle Flow Network)**

P. T. Komiske, E. M. Metodiev and J. Thaler [JHEP01(2019)121]

## Deep Sets [1703.06114]

- Namespace for symmetric function parametrization
- A general permutation-symmetric function is additive in a latent space

Feature space | Variable length | Permutation invariance | Latent space

**Deep Sets Theorem** [63]. *Let $\mathfrak{X} \subset \mathbb{R}^d$ be compact, $X \subset 2^{\mathfrak{X}}$ be the space of sets with bounded cardinality of elements in $\mathfrak{X}$, and $Y \subset \mathbb{R}$ be a bounded interval. Consider a continuous function $f : X \to Y$ that is invariant under permutations of its inputs, i.e. $f(x_1, \ldots, x_M) = f(x_{\pi(1)}, \ldots, x_{\pi(M)})$ for all $x_i \in \mathfrak{X}$ and $\pi \in S_M$. Then there exists a sufficiently large integer $\ell$ and continuous functions $\Phi : \mathfrak{X} \to \mathbb{R}^\ell$, $F : \mathbb{R}^\ell \to Y$ such that the following holds to an arbitrarily good approximation:*[1]

$$f(\{x_1, \ldots, x_M\}) = F\left(\sum_{i=1}^{M} \Phi(x_i)\right).$$

General parametrization for a function of sets

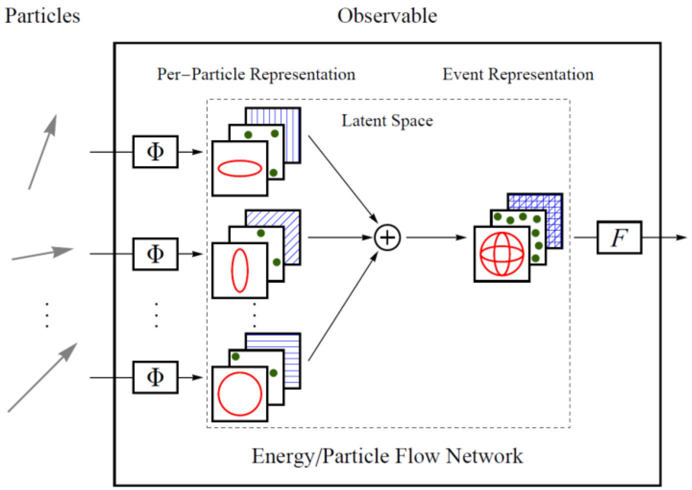
## EFN (Energy Flow Network) / PFN (Particle Flow Network)

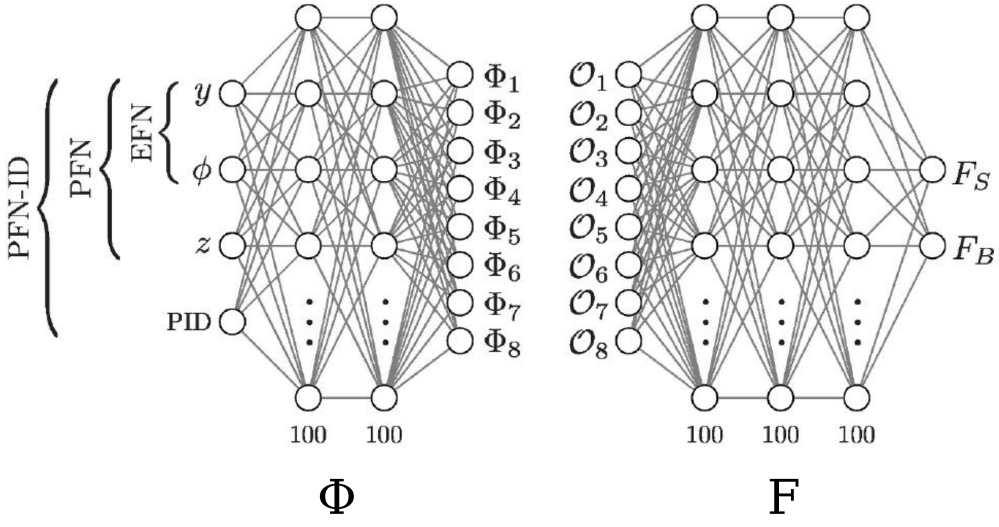
- EFN: IRC-safe latent space

$$\text{EFN:} \quad F\left(\sum_{i=1}^{M} z_i \Phi(\hat{p}_i)\right)$$

- PFN: Fully general latent space

$$\text{PFN:} \quad F\left(\sum_{i=1}^{M} \Phi(p_i)\right)$$




This page is excerpted from P. T. Komiske's talk

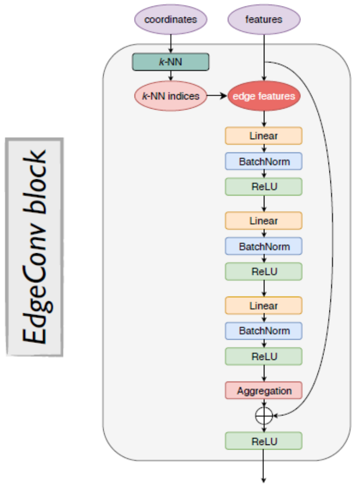
# A Brief Introduction on PFN and ParticleNet

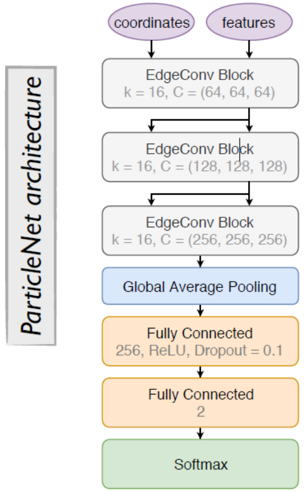
- **The architecture of ParticleNet**

H. Qu and L. Gouskos [Phys.Rev.D 101 (2020) 5, 056019]

## ParticleNet

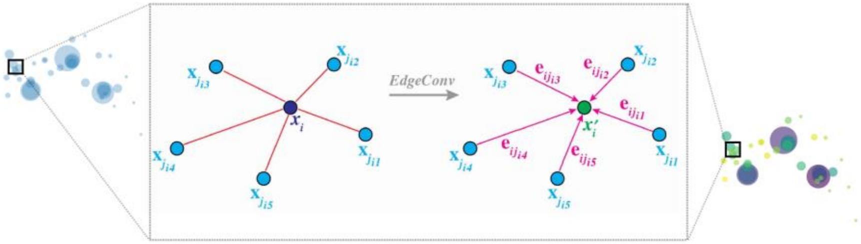
- customized graph neural network architecture for jet tagging with the point cloud approach, based on Dynamic Graph CNN (DGCNN) [Y. Wang et al., arXiv:1801.07829]
- explicitly respects the permutation symmetry of the point cloud




## Key building block: EdgeConv

- treating a point cloud as a graph: each point is a vertex
  - for each point, a local patch is defined by finding its k-nearest neighbors
- designing a permutation-invariant "convolution" function
  - define "edge feature" for each center-neighbor pair: $e_{ij} = h_\Theta(x_i,\ x_{i_j}) = \bar{h}_\Theta(x_i,\ x_{i_j} - x_i)$
    - same $h_\Theta$ for all neighbor points, and all center points, for symmetry
  - aggregate the edge features in a symmetric way: $x_i' = \square_{j=1}^{k} h_\Theta(x_i,\ x_{i_j}) = \frac{1}{k}\sum h_\Theta(x_i,\ x_{i_j})$



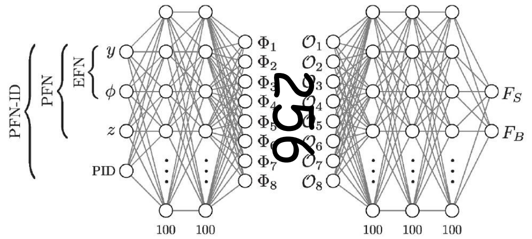

## EdgeConv can be stacked to form a deep network

- learning both local and global structures, in a hierarchical way

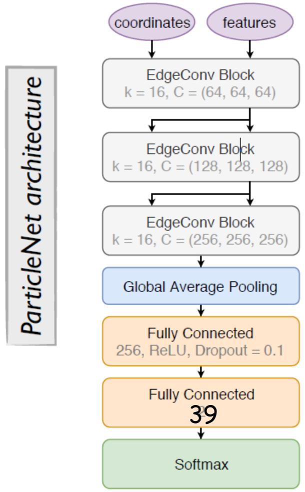
This page is excerpted from H. Qu's talk

# ML model setup & Simulation samples

- **ML model setup**
  - For PFN (used for 9-category classification)
    - For Φ, 3 dense layers, with 100, 100, and 256 nodes
    - For F , same configuration as the original paper, 3 dense layers, each with 100 nodes
    - Each dense layer uses the ReLU activation function and He-uniform parameter initialization (same as the original paper)
    - A 9-unit layer with a SoftMax activation function is the output layer




  - For ParticleNet (39-category classification)
    - Almost nothing changed w.r.t. the architecture in the original paper
    - The last fully connected layer has 39 units, since it's a 39-category classification




- **Simulation samples**
  - 4 production modes for the Higgs boson at 240 GeV to be analyzed: $e^+e^-H, \mu^+\mu^-H, \tau^+\tau^-H, q\bar{q}H$
  - For production mode, 9 decay modes are used: $H \to c\bar{c}, H \to b\bar{b}, H \to \mu^+\mu^-, H \to \tau^+\tau^-, H \to gg, H \to \gamma\gamma, H \to ZZ^*, H \to WW^*$ and $H \to \gamma Z$,
  - For 39-category classification: added $ZZ_l, ZZ_{sl}, ZZ_h$ as irreducible backgrounds
  - For each process, 400 000 events are generated with WHIZARD 1.9.5 and fed to PYTHIA6 for hadronization.
  - All particles are simulated in a simplified way according to the performance of the baseline detector in the CEPC CDR
  - Reconstruction efficiencies are assumed to be 100% in the simulation.
  - IPs and PIDs are directly taken from the truth of generation
  - Momenta, Energy, Cosθ, Φ, PID, IP of particles are used as inputs

# 9-category & 39-category classification results

- **9-category classification (done with PFN)**

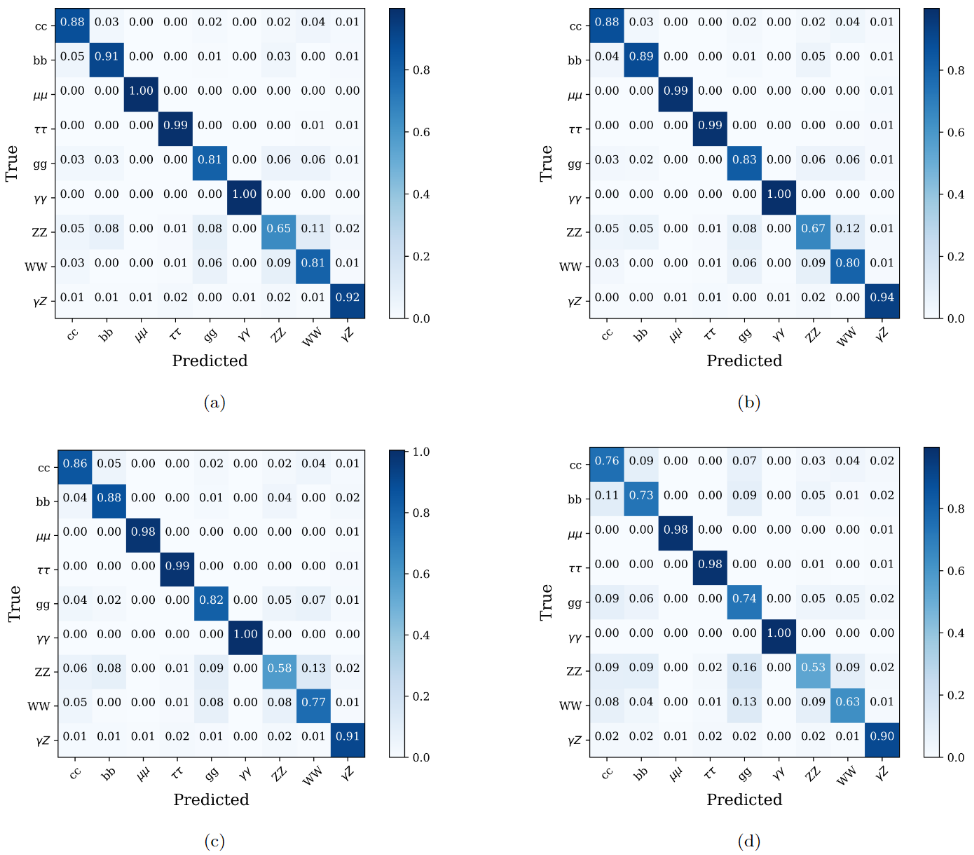



Fig. 4. Confusion matrices of (a): $e^+e^- \to e^+e^-H$, (b): $e^+e^- \to \mu^+\mu^-H$, (c): $e^+e^- \to \tau^+\tau^-H$, and (d): $e^+e^- \to q\bar{q}H$.

| Decay Mode | $e^+e^-H$ EFF | $e^+e^-H$ AUC | $\mu^+\mu^-H$ EFF | $\mu^+\mu^-H$ AUC | $\tau^+\tau^-H$ EFF | $\tau^+\tau^-H$ AUC | $q\bar{q}H$ EFF | $q\bar{q}H$ AUC |
|---|---|---|---|---|---|---|---|---|
| $H\to c\bar{c}$ | 0.880 | 0.991 | 0.882 | 0.991 | 0.857 | 0.987 | 0.755 | 0.966 |
| $H\to b\bar{b}$ | 0.908 | 0.994 | 0.893 | 0.994 | 0.877 | 0.991 | 0.733 | 0.972 |
| $H\to\mu^+\mu^-$ | 0.997 | 1.000 | 0.986 | 1.000 | 0.981 | 1.000 | 0.983 | 1.000 |
| $H\to\tau^+\tau^-$ | 0.993 | 0.999 | 0.985 | 0.999 | 0.985 | 0.999 | 0.982 | 0.999 |
| $H\to gg$ | 0.810 | 0.985 | 0.830 | 0.986 | 0.816 | 0.982 | 0.736 | 0.954 |
| $H\to\gamma\gamma$ | 0.997 | 1.000 | 0.999 | 1.000 | 1.000 | 1.000 | 0.997 | 1.000 |
| $H\to ZZ^*$ | 0.650 | 0.958 | 0.667 | 0.960 | 0.585 | 0.947 | 0.535 | 0.926 |
| $H\to WW^*$ | 0.806 | 0.981 | 0.801 | 0.981 | 0.771 | 0.974 | 0.632 | 0.952 |
| $H\to\gamma Z$ | 0.921 | 0.996 | 0.936 | 0.996 | 0.910 | 0.993 | 0.896 | 0.993 |

Table 2. Efficiencies (left) and AUCs (right) of four classifiers.

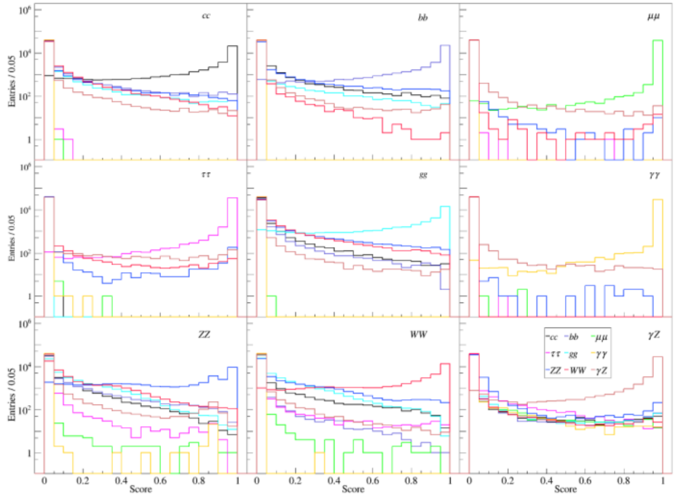
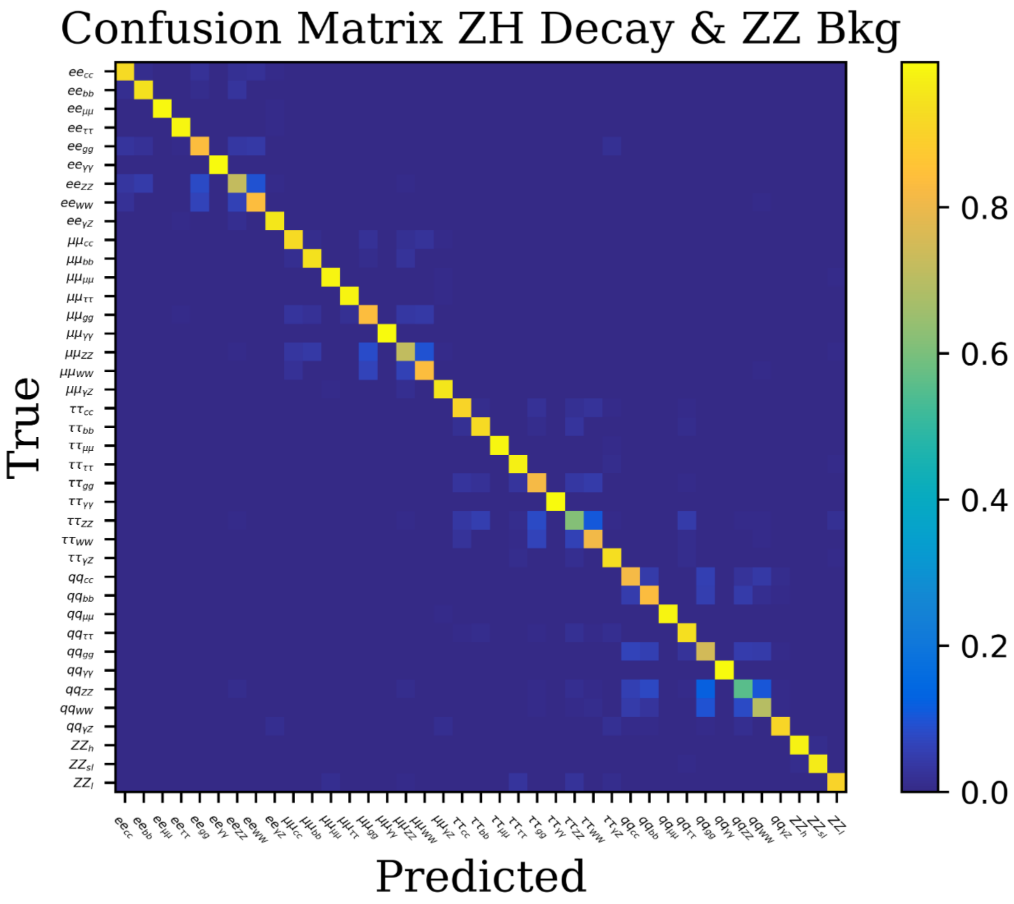



Fig. 2. The distributions of 9 outputs for each true category, taking $e^+e^-H$ as an example. Each score is calculated by assuming that the event belongs to that category.

# 9-category & 39-category classification results

- **39-category classification (done with ParticleNet)**



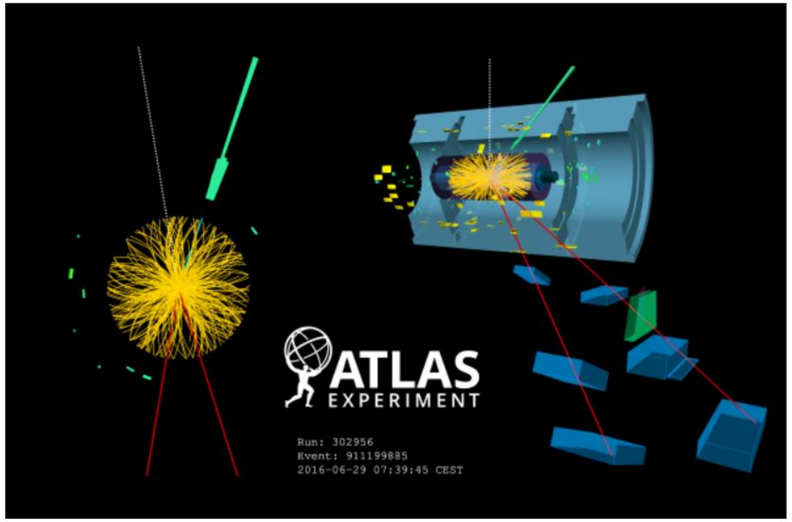

- *Avg. Acc. : 88%*

# Plan for ATLAS boson tagger

# Motivation

- Probing processes involving highly boosted massive particles, such as W and Z bosons and top quarks is important to searches for physics phenomena beyond the Standard Model, as well as Standard Model precision measurements.

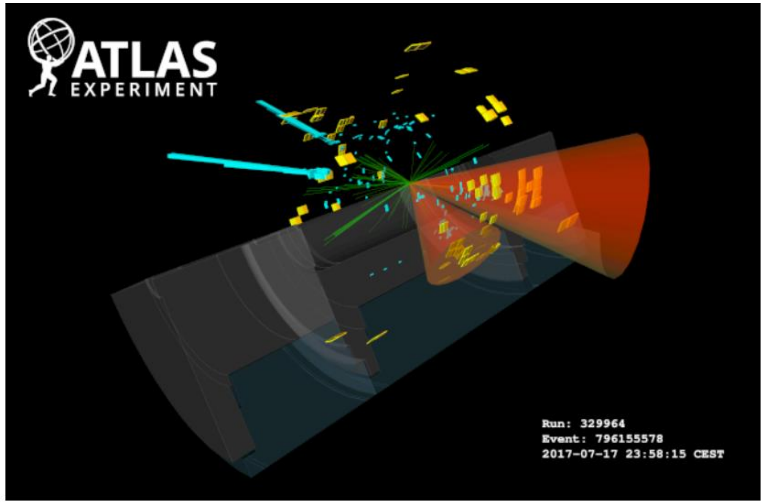
- To fully exploit these final states, it is important to reconstruct and accurately identify the hadronic decay modes of these massive particles which serve as an effective tool to reject events produced by background processes and improve the sensitivity in searches for physics beyond the Standard Model.




ATLAS Event Display: WZ polarization

# Plan for boson tagger

- Get in touch with jet tagging conveners (done)
  - Request for useful information (data samples, standard procedure, current situation, etc.)
- Get familiar with related works in the past (on-going)
  - boson taggers in the past used in both ATLAS & CMS
  - newly released ATLAS note about constituent-based top tagger
  - …
- Get started to work on it
  - Understand physics behind the samples and data structure of them
  - Use ParticleNet or even more powerful & state-of-the-art techs (e.g. Particle Transformer) to explore the performance of these tools on boson tagging tasks
  - …
- Get involved in ATLAS ML community
  - Participate the meetings, forums
  - …
- Get it used in $\gamma\gamma$WW(hadronic decay) / $\gamma\gamma$ML analyses




ATLAS event display: candidate pair of Higgs bosons decay in ATLAS

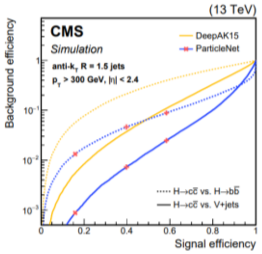
# Current situation & Expected performance

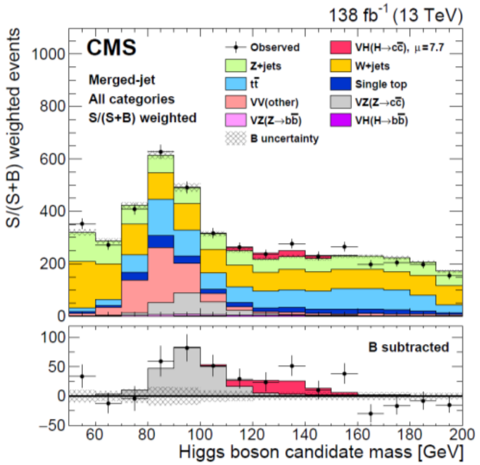
- **Current situation**

  - CMS has used these techniques in analyses and got pretty good results
    - for example, VH(cc) [2205.05550], better than ATLAS's result [Eur.Phys.J.C 82 (2022) 717]

Search for Higgs boson decay to a charm quark-antiquark pair in proton-proton collisions at $\sqrt{s} = 13$ TeV

The CMS Collaboration



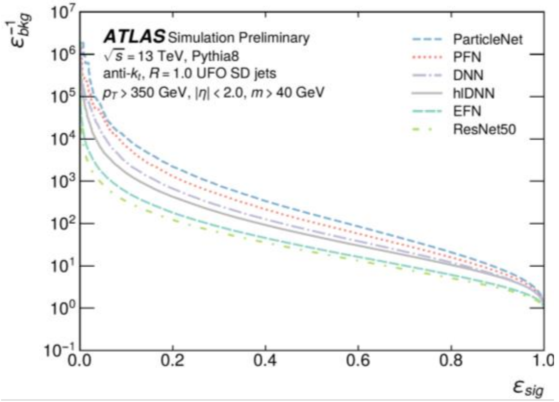

Most stringent limit on H-c coupling to date

  - Constituent-based top tagging with the ATLAS detector [ATL-PHYS-PUB-2022-039]
    - A factor of two increase in background rejection across the kinematic range was achieved using PaticleNet

ATLAS PUB Note
ATL-PHYS-PUB-2022-039
24th August 2022

**Constituent-Based Top-Quark Tagging with the ATLAS Detector**

The ATLAS Collaboration




- **Expected performance**

  - Must be better than taggers currently in use!