# *Testing CP properties of the Higgs boson with heterogeneous graphs*

with W. Esmail, C. Schreb and M. Nojiri (2409.06132)


Ahmed Hammad

Theory center, KEK, Japan


MLPhYs 学術変革領域研究(A) 学習物理学の創成
Foundation of "Machine Learning Physics"

**The Frontier of Particle physics, 2025, Kyoto**

# Motivation

*Why are we interested in CP properties of Higgs boson?*

- *Test of SM: if we find the observed scalar is not pure CP-even eigenstate, that's a hint for new physics*

# Motivation

## How to test the CP violation at the LHC?

- *If extra scalar discovered, top associated production can be used (ttH coupling is sensitive to CP-even vs -odd contributions)*
- *Angular momentum correlation in* $H_i \to ZZ \to 4\mu$ *(hard/impossible at HL-LHC, because loop-induced CP-odd contributions suppressed)*
- *Three peaks in the four leptons channel (Only True in the 2HDM)*
- *Angular correlation between tau leptons decay planes*

# Methodology

*Parametrization of Higgs boson coupling to taus:*

$$\mathcal{L}_{H\tau\tau} = -\frac{m_\tau}{v}\kappa_\tau \bar{\tau}\left(\cos\theta_\tau + i\gamma_5 \sin\theta_\tau\right)\tau h$$

$\theta_\tau$ is the CP-mixing angle

- $\theta_\tau = 0°$ represents purely CP-even Higgs
- $\theta_\tau = 90°$ represent purely CP-odd Higgs

# Methodology

From the angle, define a variable to be measured at the LHC

Observable we use is the coplanarity angle between the Tau decay planes

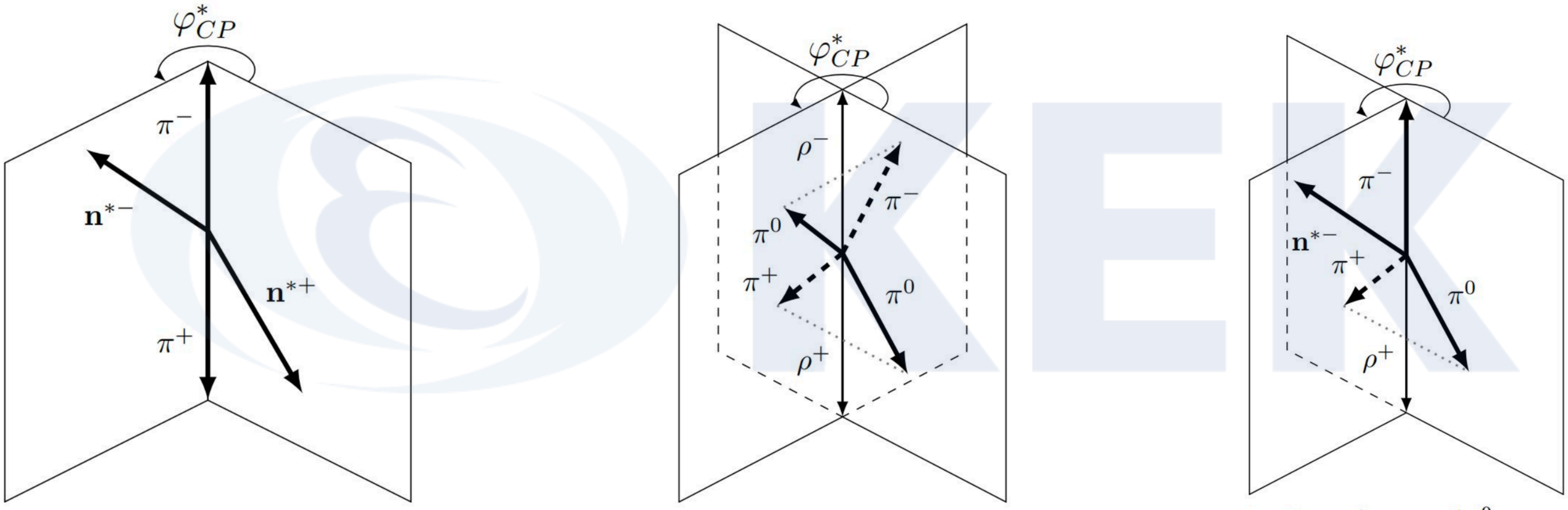
angular correlation in the decay width

arXiv:1408.0798 [hep-ph]

$$\frac{1}{\Gamma}\frac{\mathrm{d}\Gamma}{\mathrm{d}\Phi^*} = \frac{1}{2\pi}\left(1 - \frac{\pi^2}{16}(\cos^2\theta_\tau - \sin^2\theta_\tau)\cos\Phi^*\right)$$

# Methodology

acoplanarity plane can be reconstructed from decay products




(a) $H \to \tau^+\tau^- \to \pi^+\pi^- + 2\nu$

(b) $H \to \tau^+\tau^- \to \pi^+\pi^0\nu\pi^-\pi^0\nu$

(c) $H \to \tau^+\tau^- \to \pi^+\pi^0\nu\pi^-\nu$

figure taken from 2212.05833

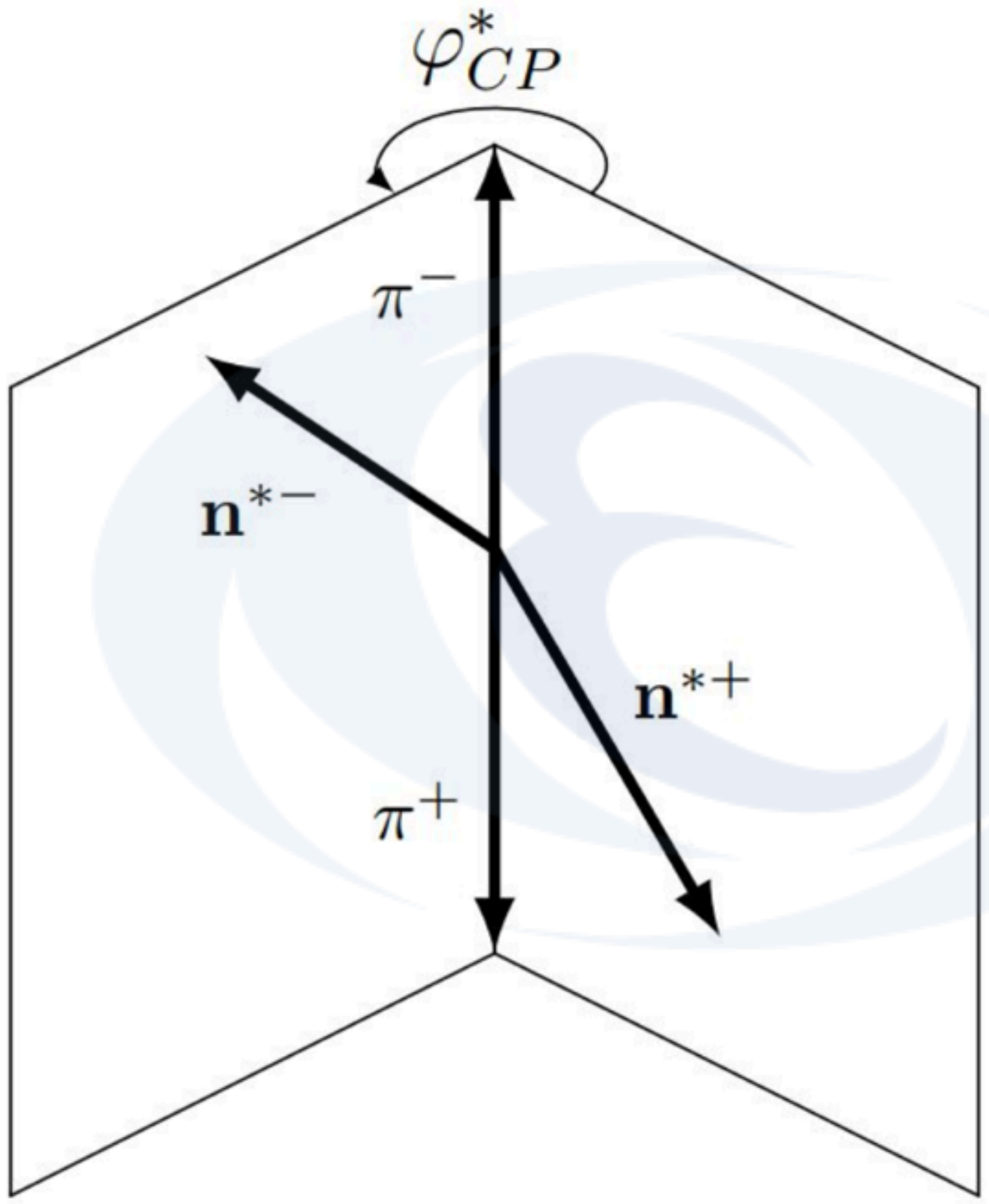
*One visible particle from the Tau decays*

*Two visible particle from the Tau decays*

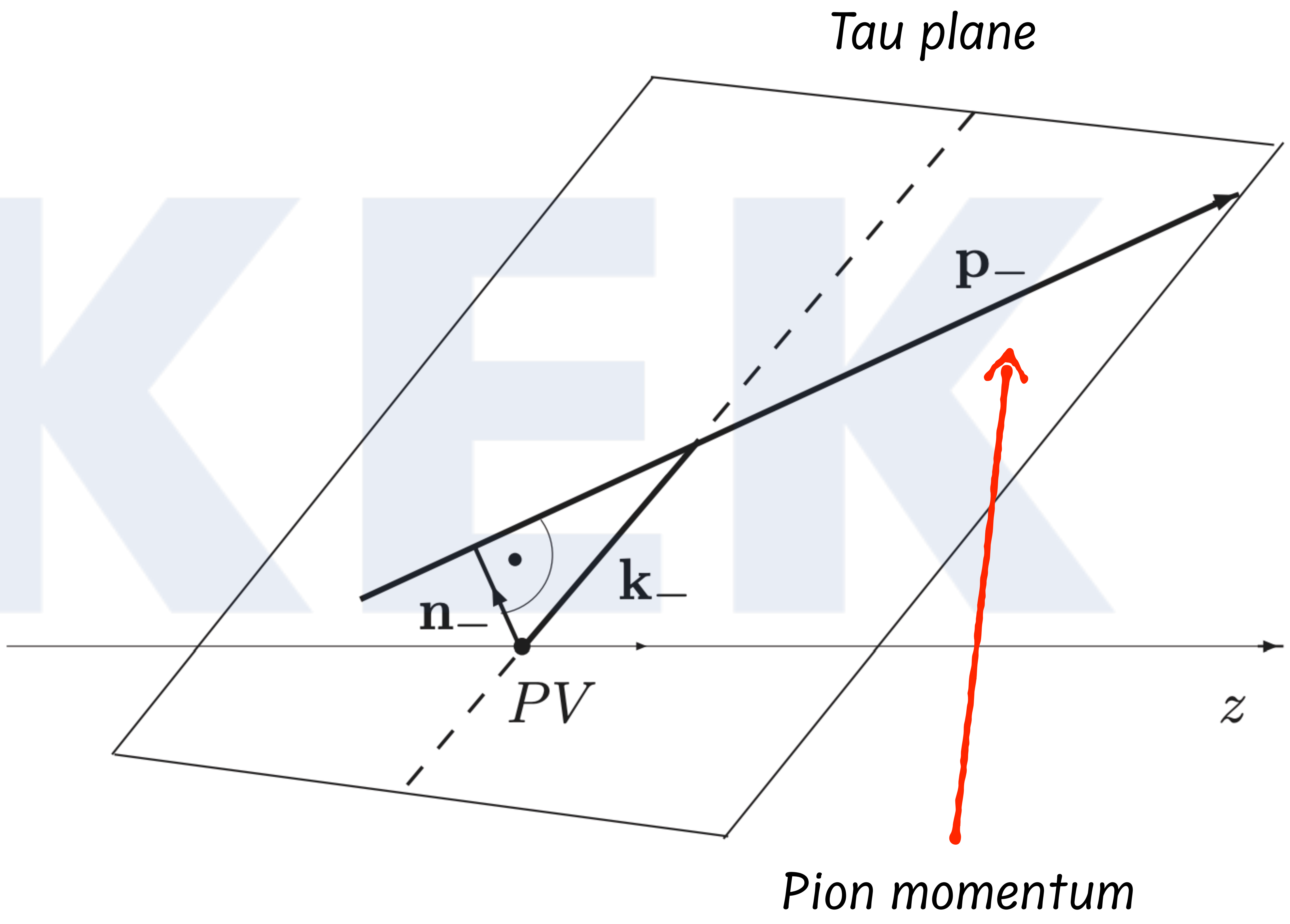
*Mixed Tau decays*

# Methodology

*Impact Parameter method is applied when Tau decays to a single visible pion*




$$\Phi^* = \arccos\left(\hat{n}_\perp^{*+} \cdot \hat{n}_\perp^{*-}\right) \times \mathrm{sgn}\left(\hat{p}^{*-}\left(\hat{n}_\perp^{*+} \times \hat{n}_\perp^{*-}\right)\right)$$




*Impact parameter vector is the closest distance between the primary vertex and the pion direction*

# Methodology

$$\Phi^* = \arccos\left(\hat{p}_\perp^{0+} \cdot \hat{p}_\perp^{0-}\right) \times \mathrm{sgn}\left(\hat{p}^- \cdot \left(\hat{p}_\perp^{0+} \times \hat{p}_\perp^{0-}\right)\right)$$

study three τ decay modes:

- $\tau^\pm \to \pi^\pm \nu_\tau$
- $\tau^\pm \to \rho^\pm \nu_\tau$
- $\tau^\pm (\to a_1^\pm \nu_\tau) \to \pi^+ 2\pi^0 \nu_\tau$

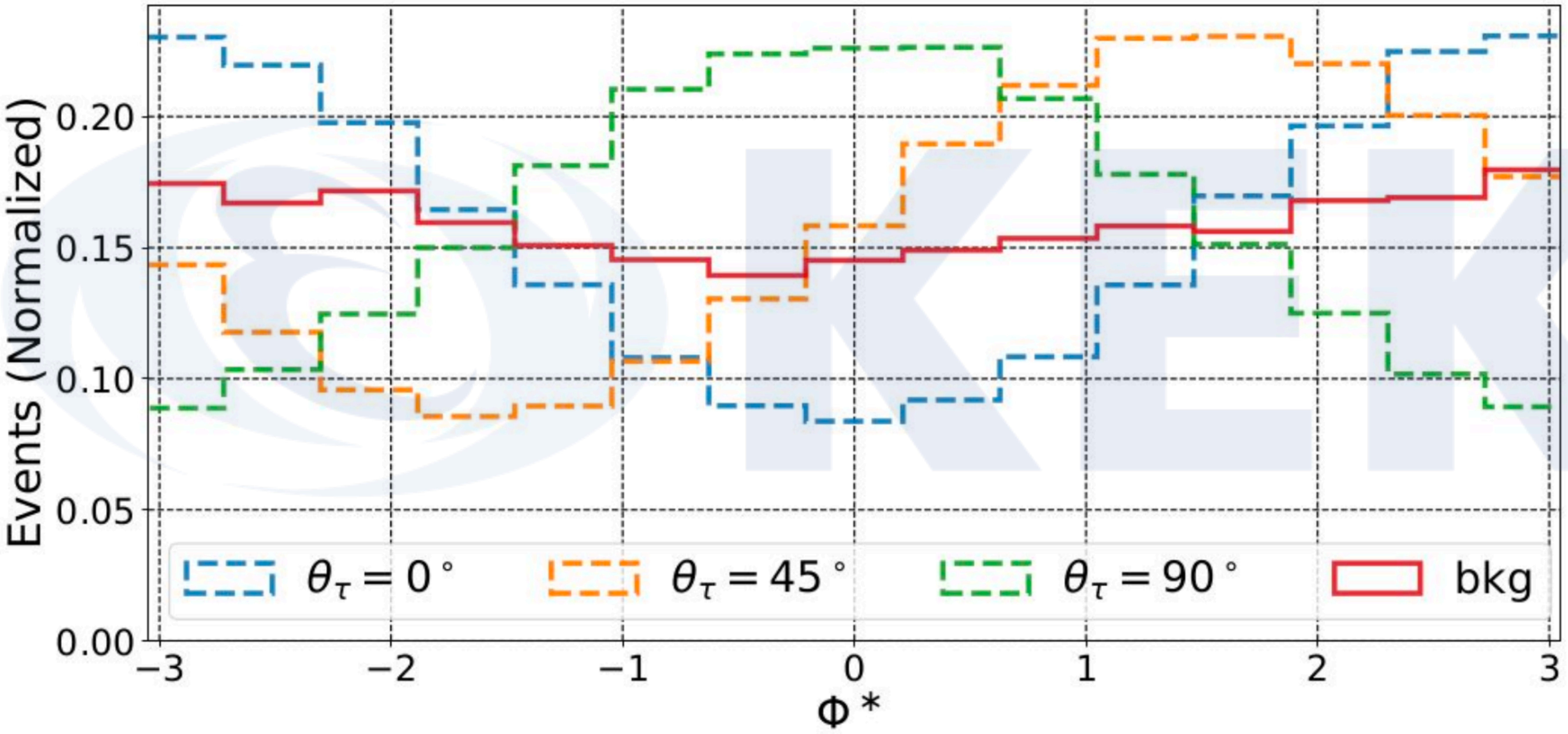
**main background:** Z/γ decays to τ pairs

all processes simulated using MadGraph + Pythia + Delphes and checked against constraints

$\mathrm{Br}(\tau^- \to \pi^- \nu) = 10.8\%$ $\mathrm{Br}(\tau^- \to \rho^- \nu) = 25.49\%$ $\mathrm{Br}(\tau^- \to \pi^- 2\pi^0 \nu) = 9.26\%$

# Methodology

## Acoplanarity angle distribution, MG+Pythia+Delphes




*To keep the angular correlation in the Tau decay products, TauDecay library is used (1212.6247)*

# Methodology

## *Current status of the CPV searches at the LHC*

ATLAS and CMS have both performed such analyses using traditional

**cut-and-count** and **boosted decision trees**

do well in **excluding purely CP-odd Higgs**

not great at actually measuring CP angle

Without constraining the $H \to \tau\tau$ signal strength to its expected value under the Standard Model hypothesis, the mixing angle $\phi_\tau$ is measured to be $9^\circ \pm 16^\circ$, with an expected value of $0^\circ \pm 28^\circ$ at the 68% confidence level. The pure $CP$-odd hypothesis is disfavoured at a level of 3.4 standard deviations. The results are compatible with the predictions for the Higgs boson in the Standard Model.

ATLAS 2212.05833

# Deep Learning analyses

use three Deep Learning Neural Networks to increase signal-to-background yield:

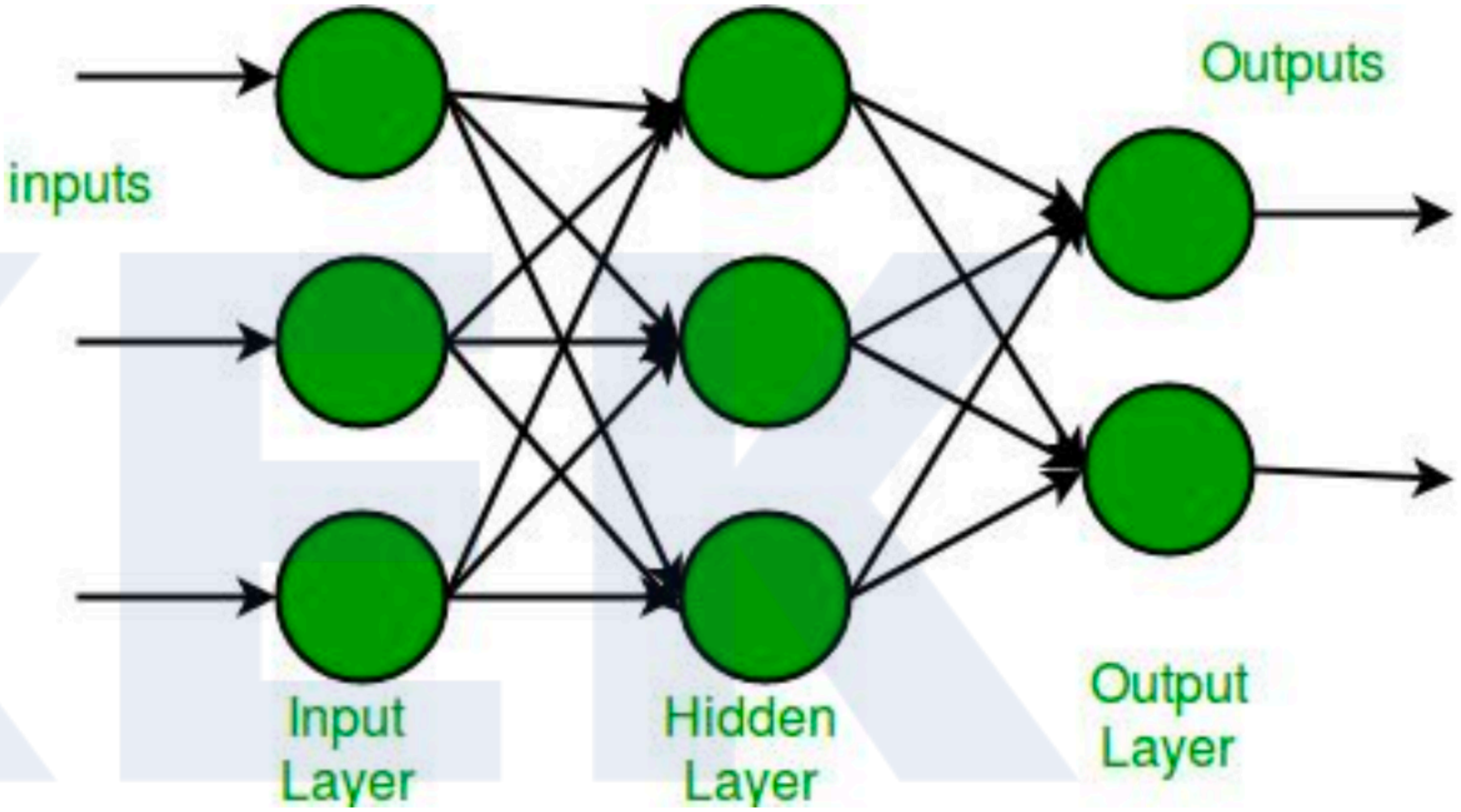
1. Multi-Layer Perceptron (MLP)
2. Graph Neural Network (GNN)
3. Graph Transformer Network (GTN)

*Our codes are made for public, for reproducibility purpose*

https://github.com/wesmail/HiggsCP

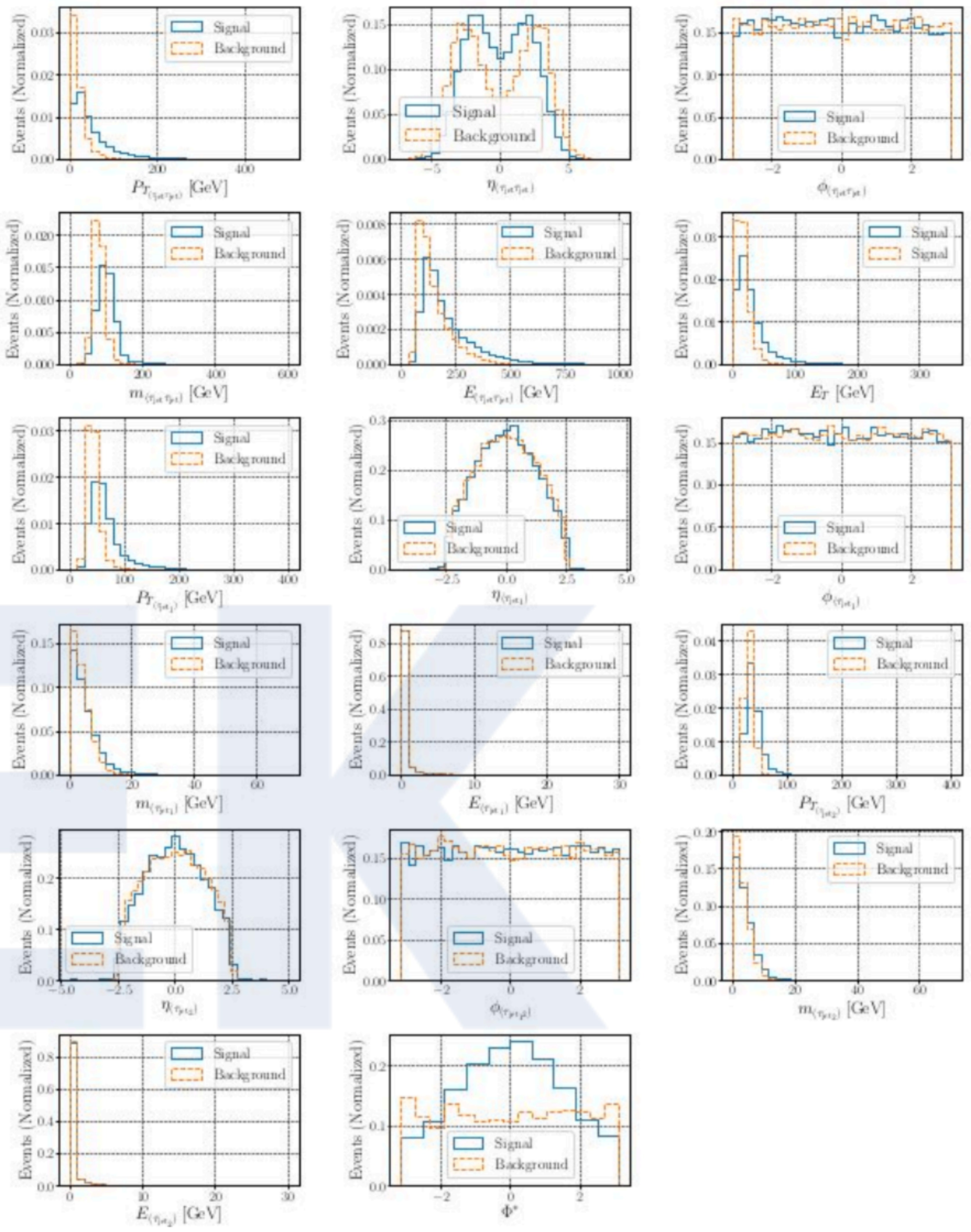
# Multi-Layers perceptron

- feed-forward neural network consisting of
  - input layer
  - one or more hidden layer(s)
  - output layer
- neurons are fully connected
- adept at analyzing high-level kinematic events



# Multi-Layers perceptron

- we use 17 input variables:
  - m, η, ϕ, E, pT of two leading τ jets and τ-pair
  - missing energy
  - acoplanarity angle $\Phi^*$
- stack signal and background separately
- train with 80000 signal and background events each
- use 20000 events for evaluation during training




Limitations:

identification performance limited due to fully connected neurons in each hidden layer

⟹ dilutes learned CP pattern by fully connecting it to kinematics

⟹ graph neural networks (GNN) can overcome this by using heterogeneous graphs

# Graph Neural Network

analyze graph-like structures, typically

- nodes represent final states, fully connected
- nodes weighted with 4-momenta of final states
- edges weighted with angular distance between final states

⟹ not easy to incorporate Higgs CP properties

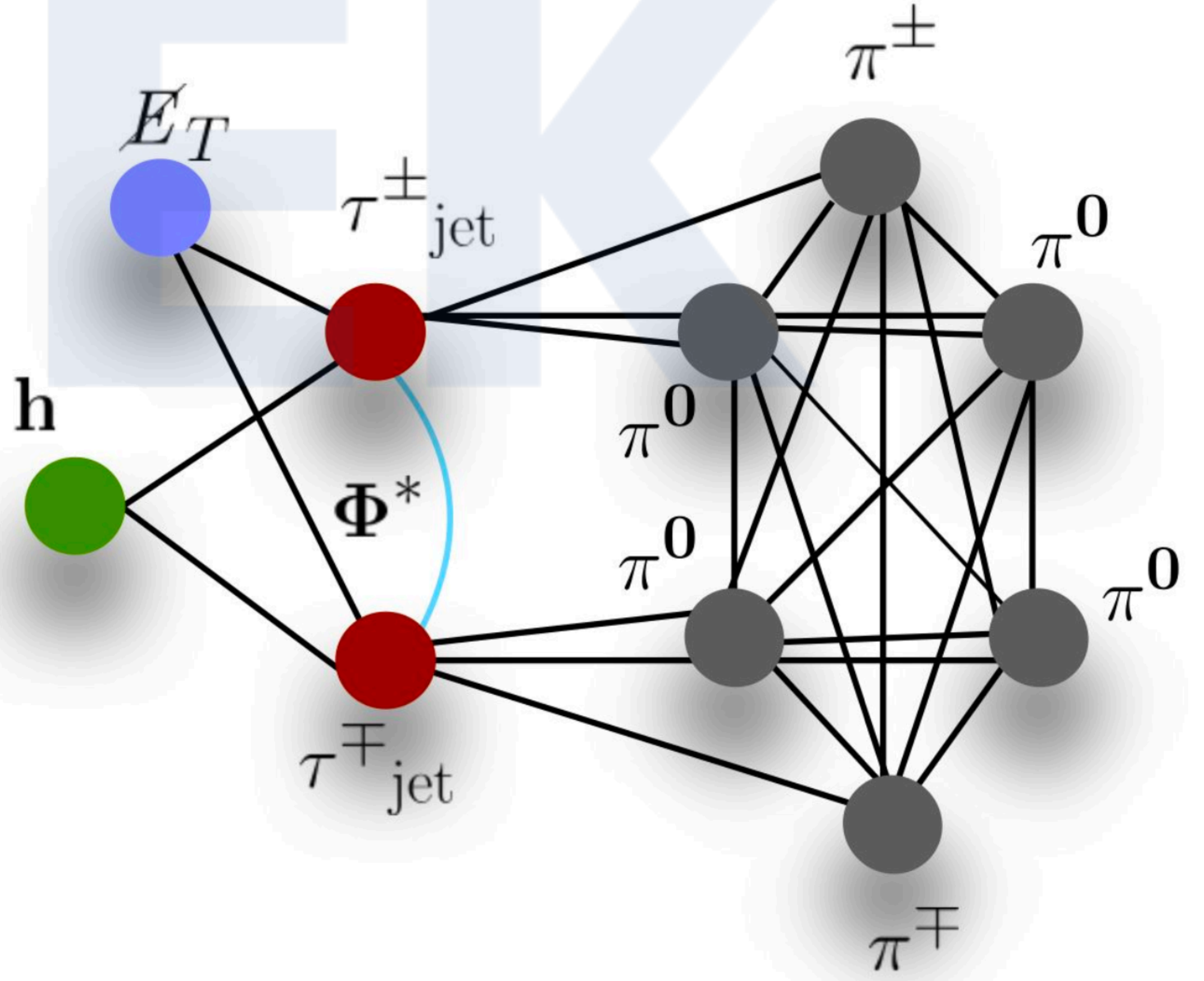
⟹ use heterogeneous graphs

**homogeneous graphs:**

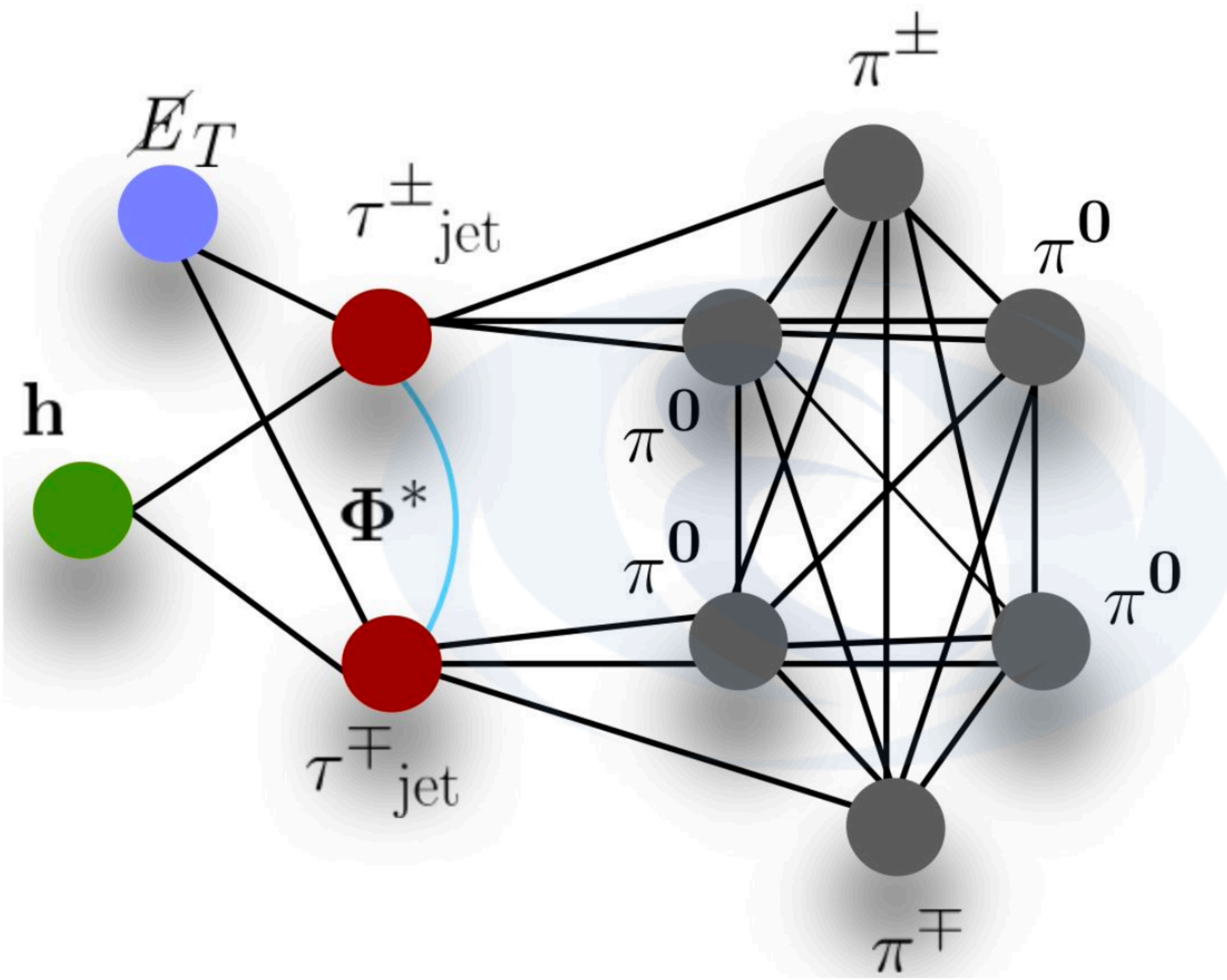
- single type of nodes and edges

**heterogeneous graphs:**

- multiple types of edges and nodes with different properties and relationships
- can accurately model complex systems, e.g. topology of Higgs decay events



# Graph Neural Network




| Node name | Features |
|---|---|
| $\pi^{\pm,0}$ | $\eta_\pi,\ \phi_\pi\ P_{T_\pi},\ \theta_{(\pi,\tau_{\text{jet}})},\ E_\pi,\ \log(P_{T_\pi}/P_{T_{\tau_{\text{jet}}}})$ |
| $\tau^{\pm}$ | $\eta_{\tau_{\text{jet}}},\ \phi_{\tau_{\text{jet}}}\ P_{T_{\tau_{\text{jet}}}},\ m_{\tau_{\text{jet}}},\ E_{\tau_{\text{jet}}}$ |
| $\not{E}_T$ | $\not{E}_T$ |
| $h$ | $\eta_h,\ \phi_h\ P_{T_h},\ m_h,$ where $P_h \equiv P_{\tau_{j_1}} + P_{\tau_{j_2}}$ |

| Edge name | Features |
|---|---|
| $\pi_i$ - $\pi_j$ | $\Delta R_{ij}$ |
| $\pi_i$ - $\tau_j$ | $\log(P_{T_i}/P_{T_{\tau_j}}),\ \theta_j$ |
| $\not{E}_T$ - $\tau$ | $\mathbb{I}$ |
| $h$ - $\tau$ | $\log(P_{T_\tau}/P_{T_h})$ |
| $\tau$ - $\tau$ | $\Phi^*$ |

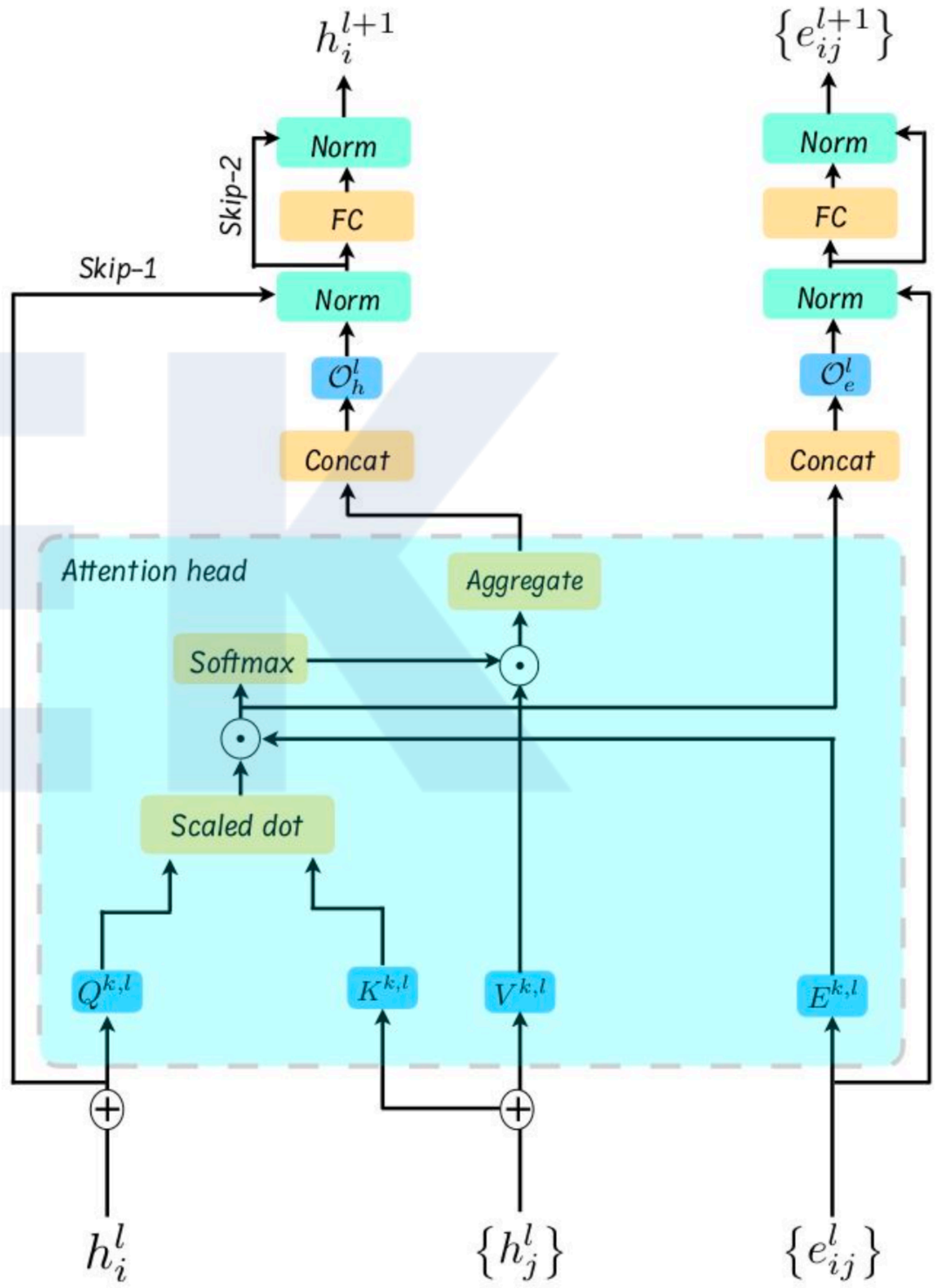
*This structure is considered as a prior to separate between the kinematic and CP observables.*

# Graph Transformer Network

combines traditional GNN and transformers

⟹ more powerful & flexible framework for analyzing graphs

attention mechanism enables selective focus on different nodes (e.g. here pion nodes vs tau nodes)




<u>self attention mechanism</u>

attention matrix $\alpha_{ij} = \frac{Q \cdot K^T}{\sqrt{d_k}}$ with query Q and key K

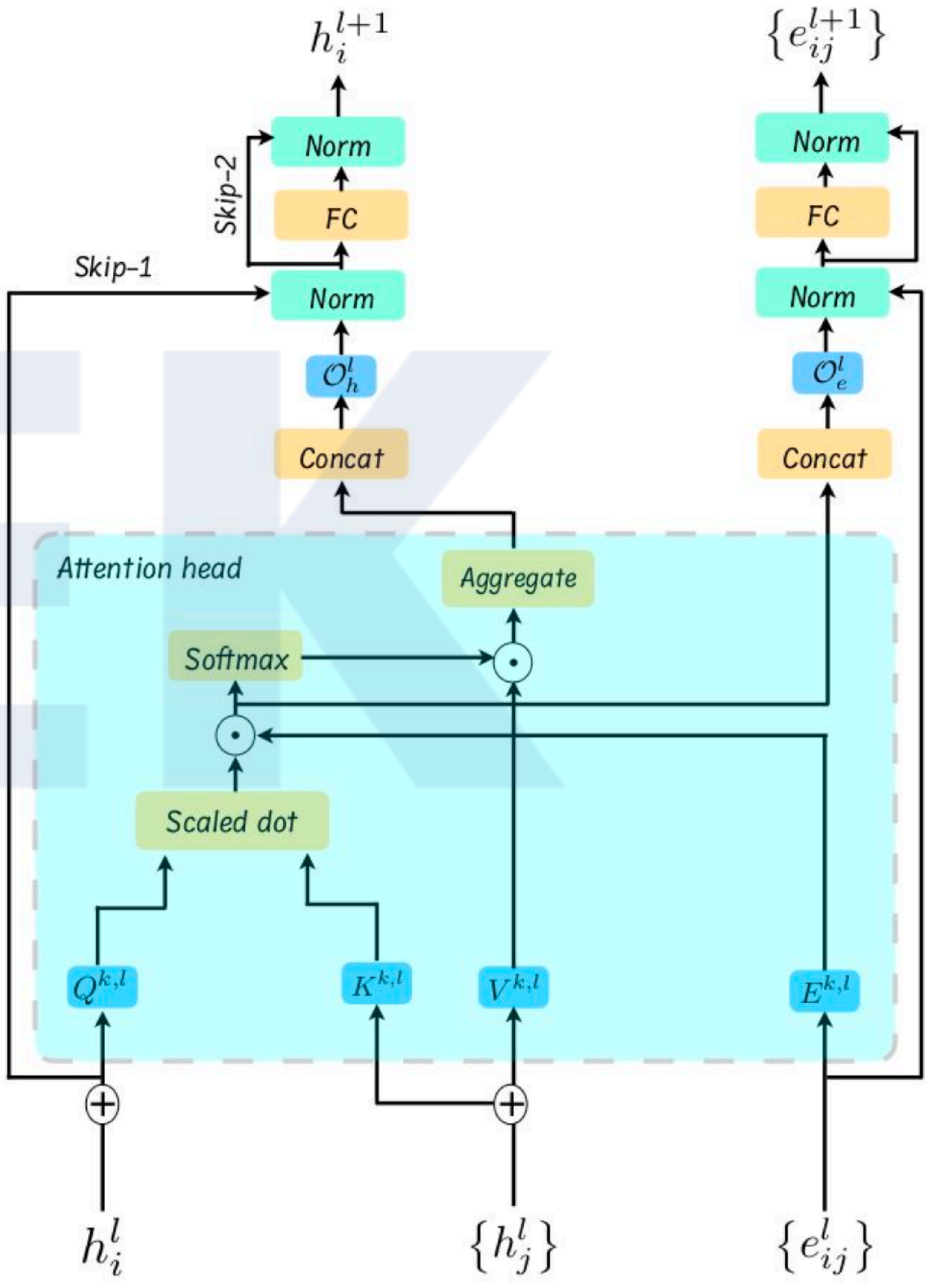
edge-wise attention mechanism

$$A_{ij} = \frac{\exp(\alpha_{ij} \cdot E_e^k)}{\sum_{k \in \mathcal{N}(i)} \exp(\alpha_{ij} \cdot E_e^k)}$$

with $E_e^k = e_{ij} \cdot W_E$

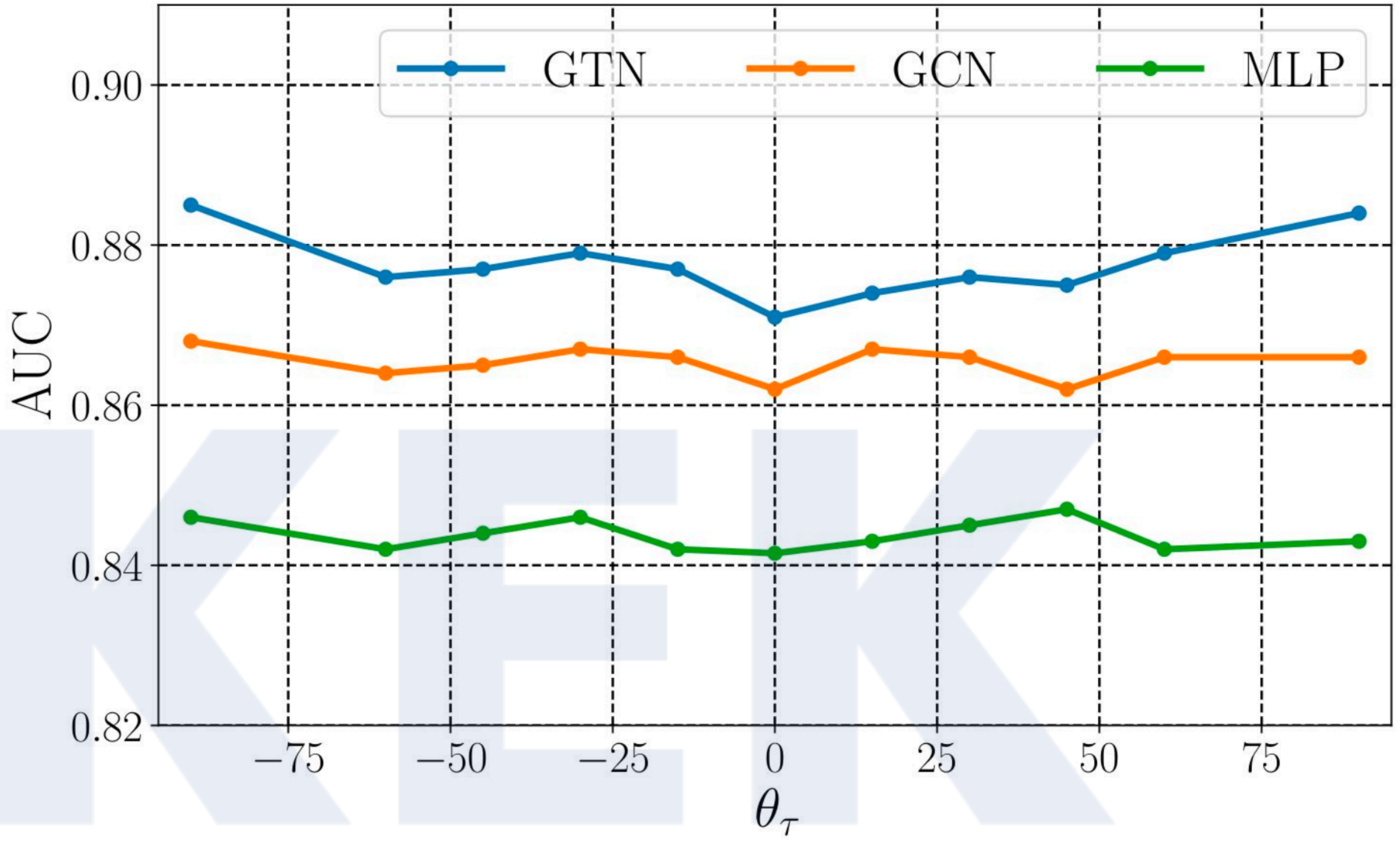
# Graph Transformer Network

*Multihead attention enables the model to learn rich, topological-aware representations by attending to different aspects of a node's neighborhood simultaneously.*
*Each attention head captures distinct structural or semantic relationships, allowing the model to weigh neighbor contributions dynamically based on their features. This leads to more expressive and flexible learning compared to traditional GNNs, and it enhances both performance and interpretability on graph-based tasks.*



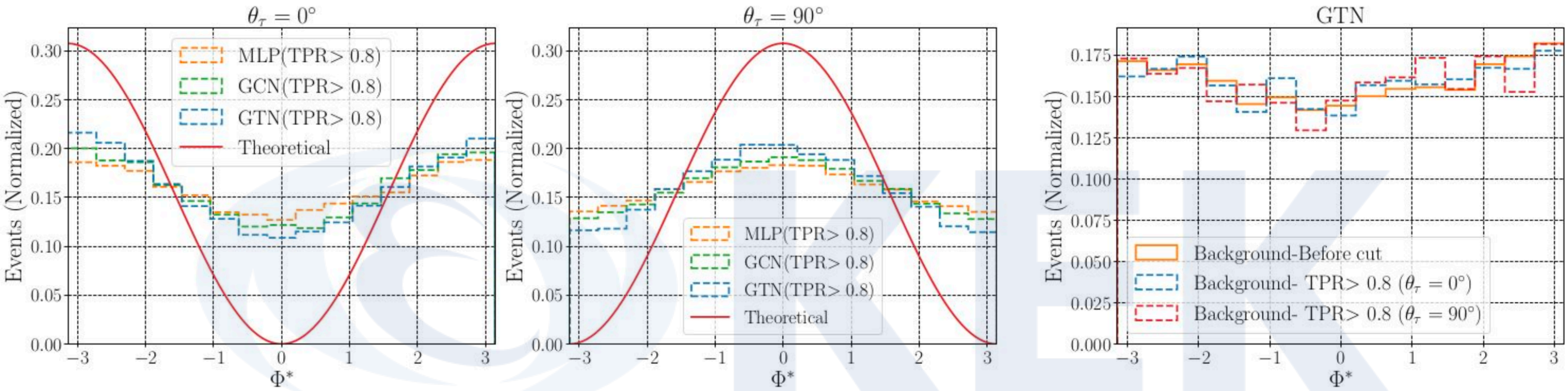
# Results

*AUC: to the area under the ROC curve. It is a metric used to evaluate the performance of binary classification models.*




| | Selection cuts | MLP(TPR> 0.8) | GCN(TPR> 0.8) | GTN(TPR> 0.8) |
|---|---|---|---|---|
| Background events | 872554 | 14982 | 8901 | 6169 |
| Signal events | 1102 | 703 | 705 | 708 |
| Signal significance | $2.9\sigma$ | $5.6\sigma$ | $7.2\sigma$ | $8.6\sigma$ |

# Results: Shape analysis




*Acoplanarity angle distribution after optimizing the cut over the network output probability.*

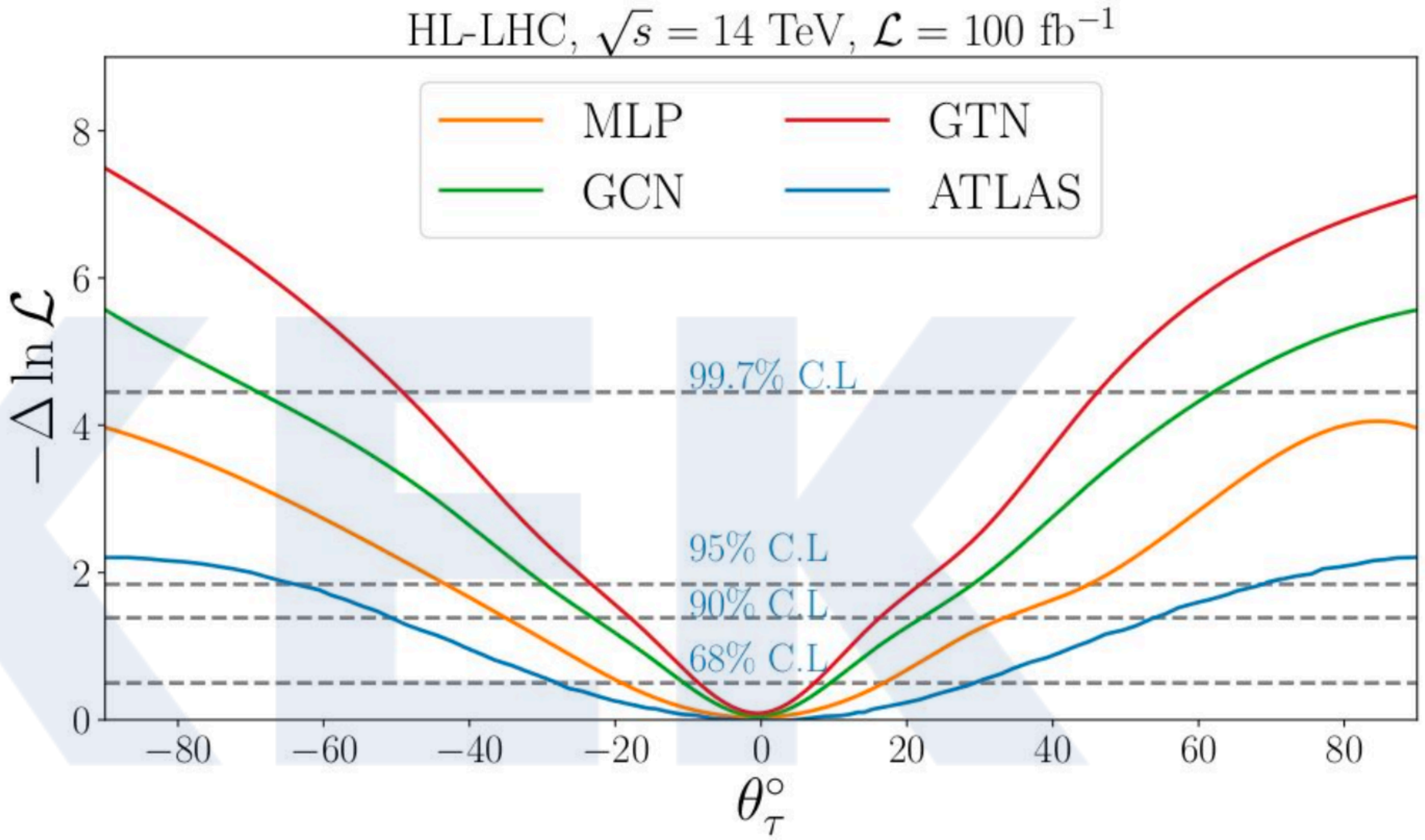
# Results: Shape analysis

log-likelihood test using

$$-\Delta \ln \mathcal{L} = -\sum_i \left[ n_i \log\left(\frac{n_i}{\nu_i}\right) + \nu_i - n_i \right]$$

$n_i$: bins for null hypothesis (purely CP-even Higgs)

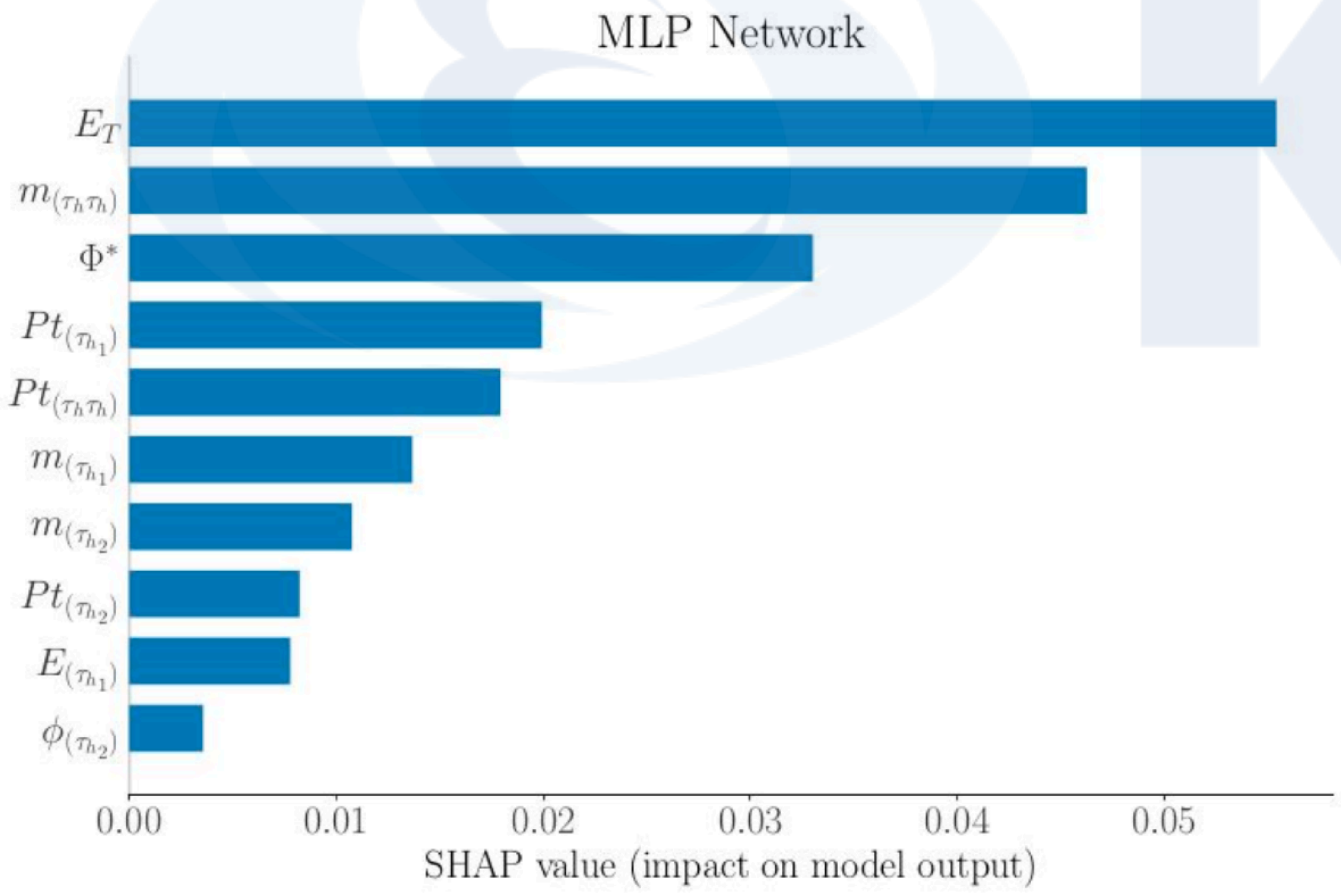
$\nu_i$: bins for alternate hypothesis (CP admixture Higgs)

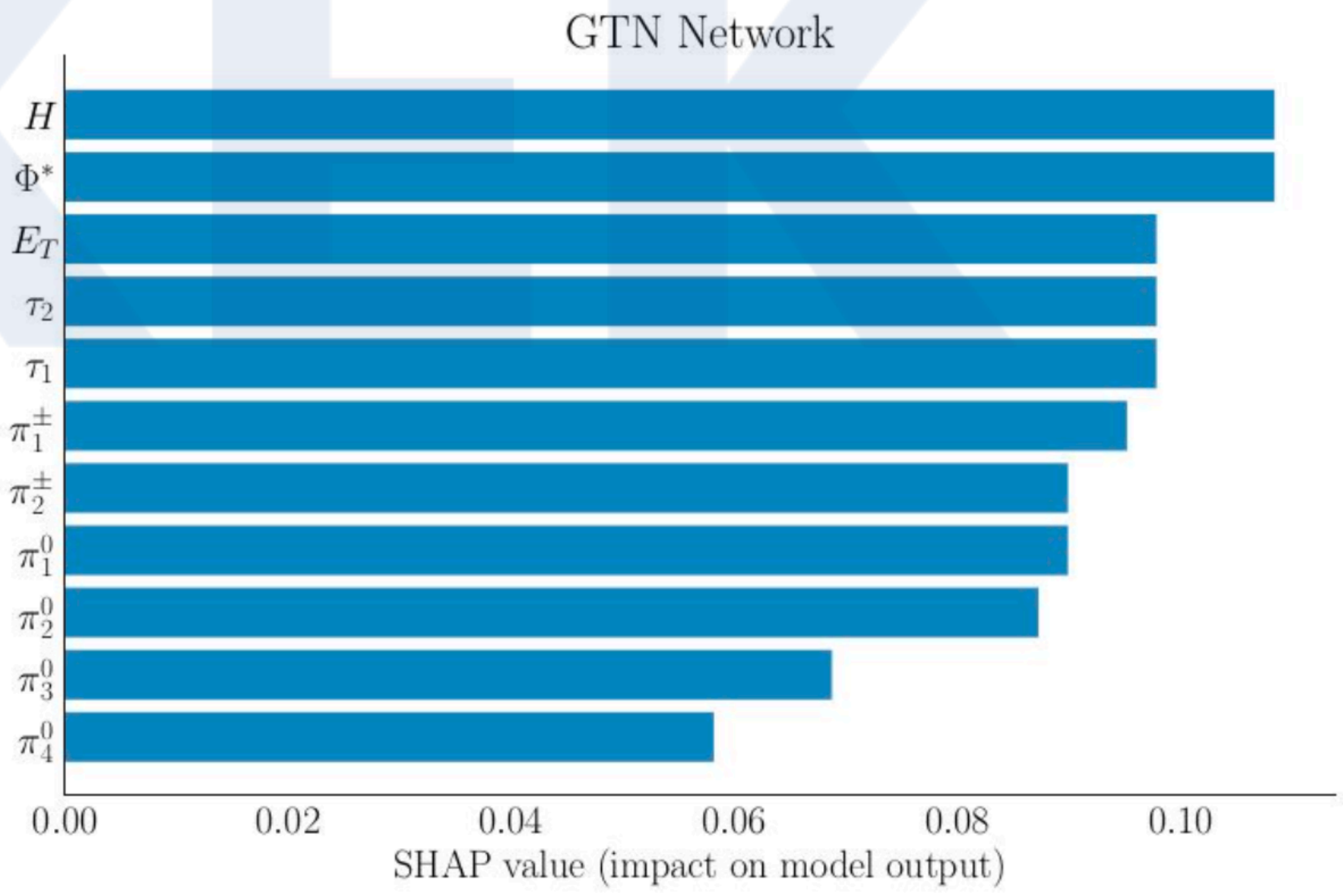


# Results: Interpretation

*Shapley Additive Explanations (SHAP) is a model-agnostic method that assigns each feature a contribution value to explain individual predictions. It ensures fair and consistent attribution by averaging a feature's impact over all possible combinations of features.*

$$\phi_i = \sum_S \frac{|S|!(|N| - |S| - 1)!}{|N|!} [f(S \cup \{i\}) - f(S)]$$



# Summary

- testing CP properties of Higgs is important to confirm SM and could possible be a hint for new physics
- using neural networks improves signal to background yield allowing better measurements
- heterogeneous graphs can describe complex topologies
- GTN allows flexible and powerful framework to analyze such complex topologies and can significantly improve measurements of CP properties at HL-LHC

Thank you

for your attention