Measurement of Spin and Quantum Correlations in $Z \to \tau^+\tau^-$ at DELPHI Using Machine Learning

Cen Mo[1], **Chen-Hua Hsu (Ken) [2]**, Colby Joseph Lamore[3], Jingyu Zhang[4], Liang Li[1], Kai-Feng Chen[2], Shih-Chieh Hsu[3],Ting-Hsiang Hsu[2], Wei-Shu Hou[2], Yen-Jie Lee[5],Yi Chen[4], Yu Boxun[6], Yulei Zhang[3],
Electron-Positron Alliance

[1] Shanghai Jiao Tong University
[2] National Taiwan University
[3] University of Washington
[4] Vanderbilt University
[5] Massachusetts Institute of Technology
[6] Peking University

Map

1

What's our Goal

01

2

Why do we need ML

02

3

How ML does each part

03

4

Physics Result

04

$$\langle O \rangle = \langle \Psi | O | \Psi \rangle = \mathrm{Tr}\big(O \underbrace{|\Psi\rangle\langle\Psi|}_{\rho}\big)$$

$$\rho = \frac{1}{4}\left[\mathbb{1} \otimes \mathbb{1} + \sum_i B_i^+ \sigma_i \otimes \mathbb{1} + \sum_j B_j^- \mathbb{1} \otimes \sigma_j + \sum_{ij} C_{ij} \sigma_i \otimes \sigma_j\right]$$

B^+ Polarization B^- Polarization Spin Correlation

	I	n	r	k
I	$\mathbb{1}$	B_n^-	B_r^-	B_k^-
n	B_n^+	C_{nn}	C_{nr}	C_{nk}
r	B_r^+	C_{rn}	C_{rr}	C_{rk}
k	B_k^+	C_{kn}	C_{kr}	C_{kk}

Quantum entanglement *concurrence*

$$\mathscr{C}[\rho] = max(0, \lambda_1 - \lambda_2 - \lambda_3 - \lambda_4)$$

$\mathscr{C} = 0$
separable

$0 < \mathscr{C} \leq 1$
entangled

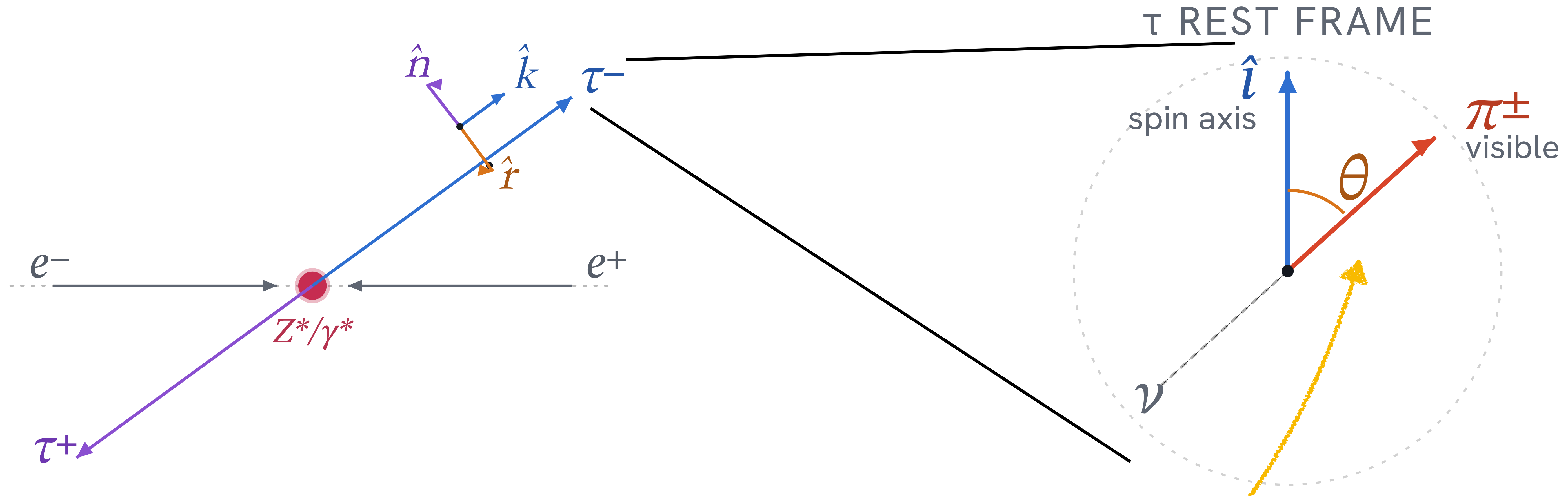
Bell nonlocality *CHSH*

$$B = max_{ij} |C_{ii} \pm C_{jj}| - \sqrt{2}$$

$B \leq 0$
local

$0 < B \leq 2 - \sqrt{2}$
non-local

Goal — Full Spin Density Matrix



event count

analyzing power

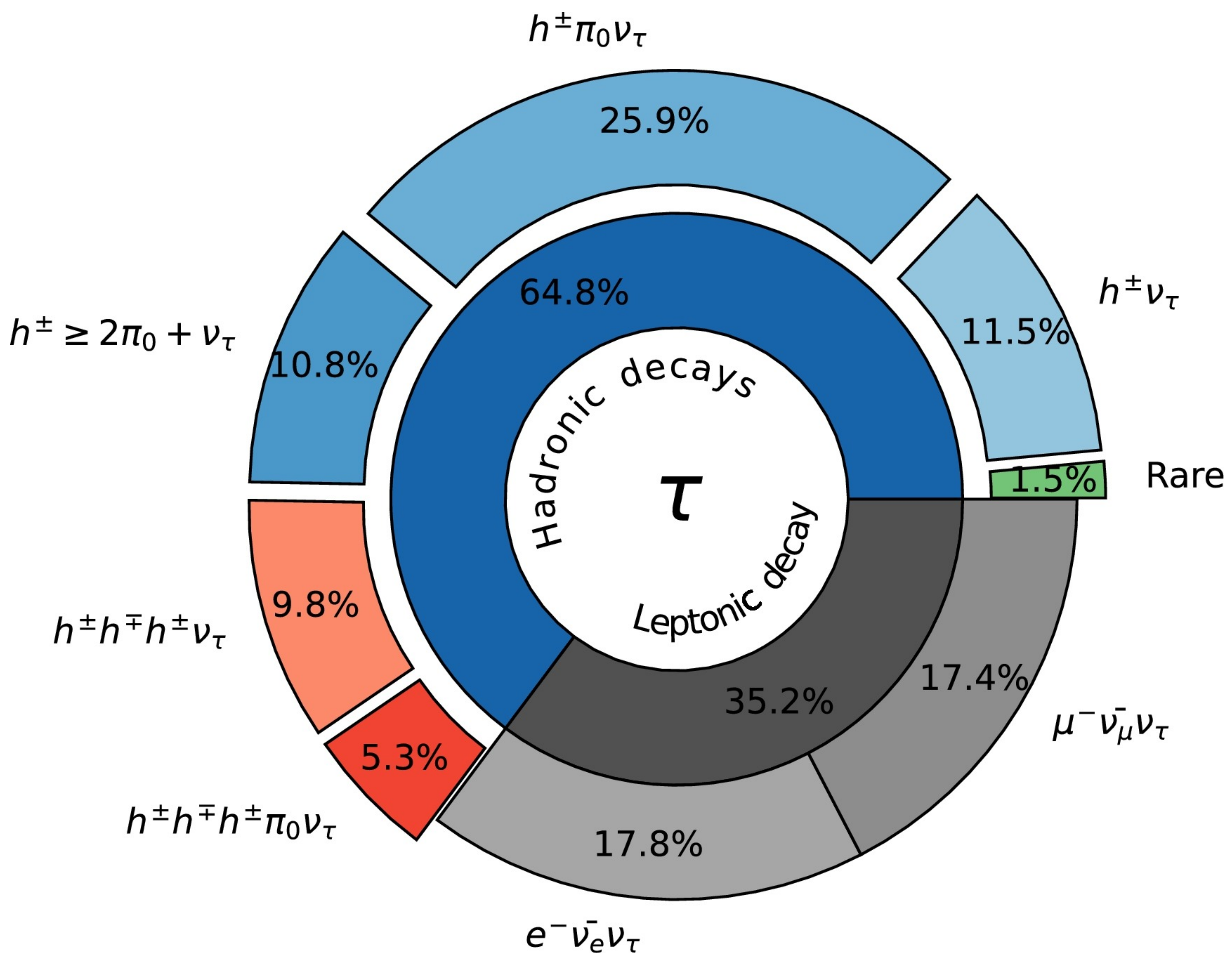
B_i^+ $$\frac{1}{\sigma}\frac{d\sigma}{d\cos\theta_{A,i}} = \frac{1}{2}(1+\kappa_A B_i^+ \cos\theta_{A,i})$$

B_j^- $$\frac{1}{\sigma}\frac{d\sigma}{d\cos\theta_{B,j}} = \frac{1}{2}(1+\kappa_B B_j^- \cos\theta_{B,j})$$

C_{ij} $$\frac{1}{\sigma}\frac{d\sigma}{d\cos\theta_{A,i}\cos\theta_{B,j}} = -\frac{1}{2}(1+\kappa_A\kappa_B C_{ij}\cos\theta_{A,i}\cos\theta_{B,j})\log|\cos\theta_{A,i}\cos\theta_{B,j}|$$

Analyzing Power

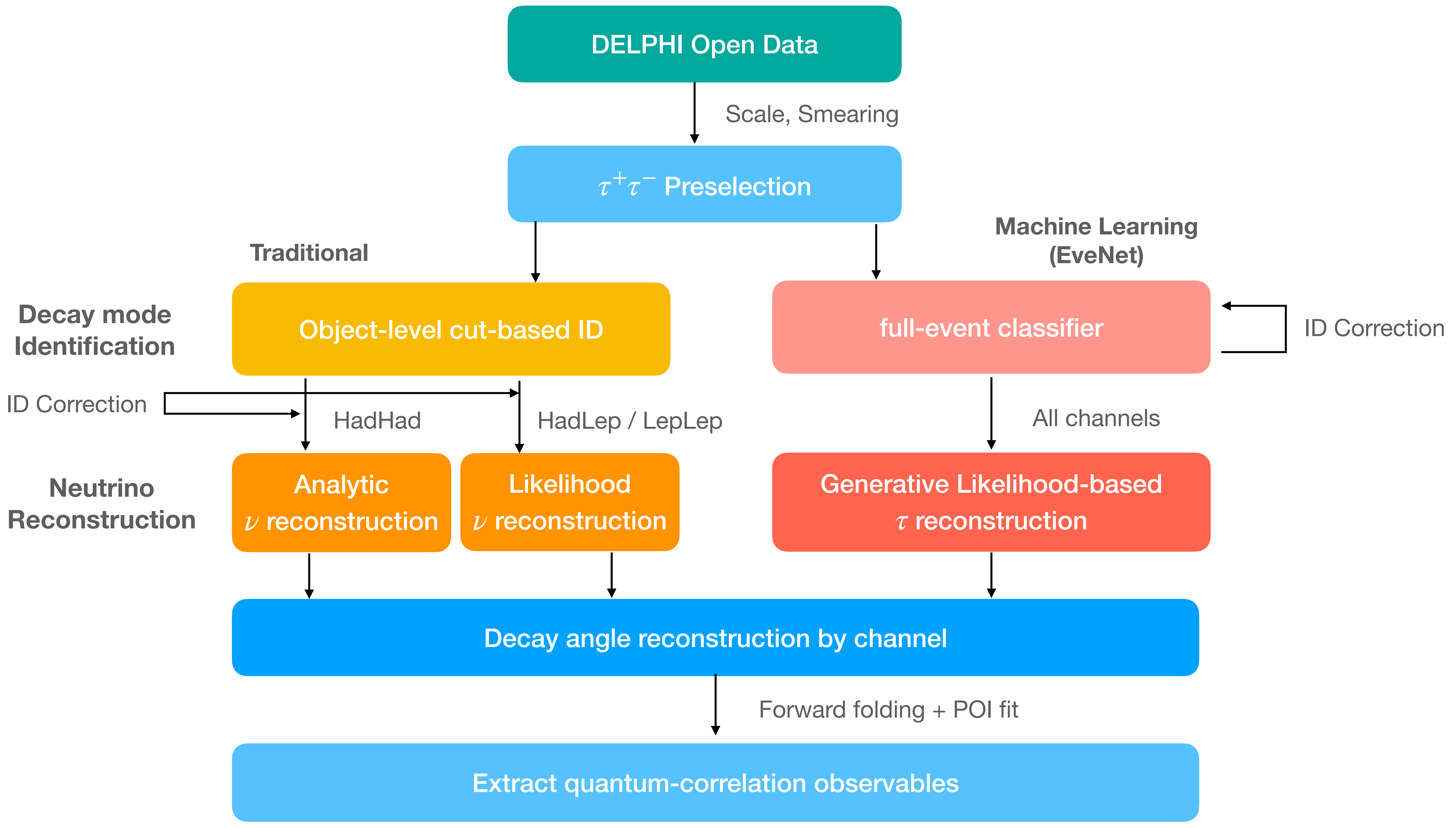
Spin analyzing power (κ) quantifies the **correlation** between the spin direction of a particle and the direction of its decay product A: $-1 \leq \kappa_A \leq 1$, larger $|\kappa_A|$ value → better sensitivity of measurement



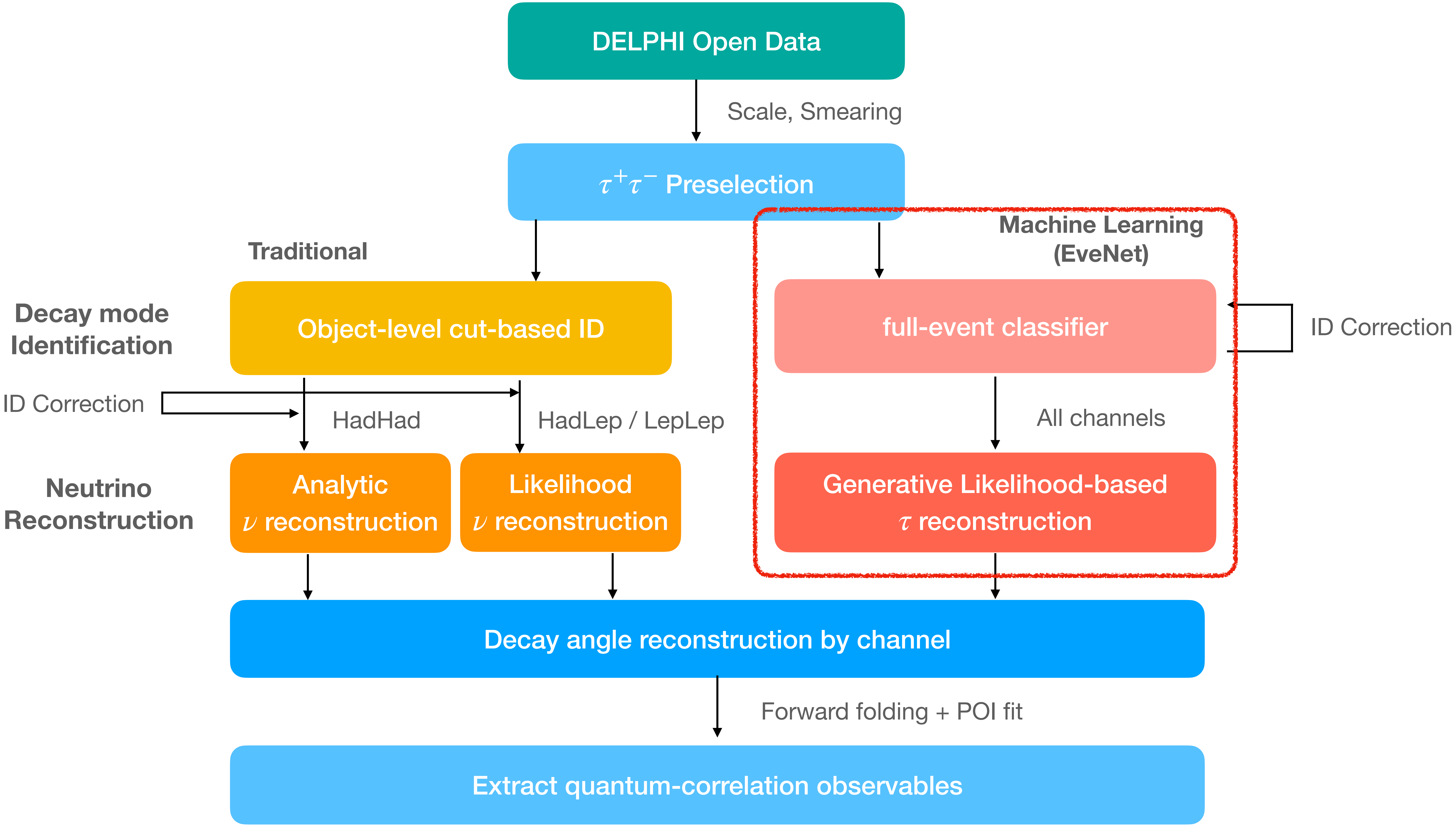
Tau Decay Object	Spin Analyzing Power
πν	κ = 1
ρν	κ ≈ 0.41
ℓνν	κ ≈ −0.33/0.34

Subchannel	Effective analyzing power	Branching fraction
ππ	1 × 1	1.17%
ℓπ	0.33 × 1	3.81%
πρ	1 × 0.41	2.76%
ℓρ	0.33 × 0.41	8.98%
ρρ	0.41 × 0.41	6.50%
ℓℓ	0.33 × 0.33	11.6%

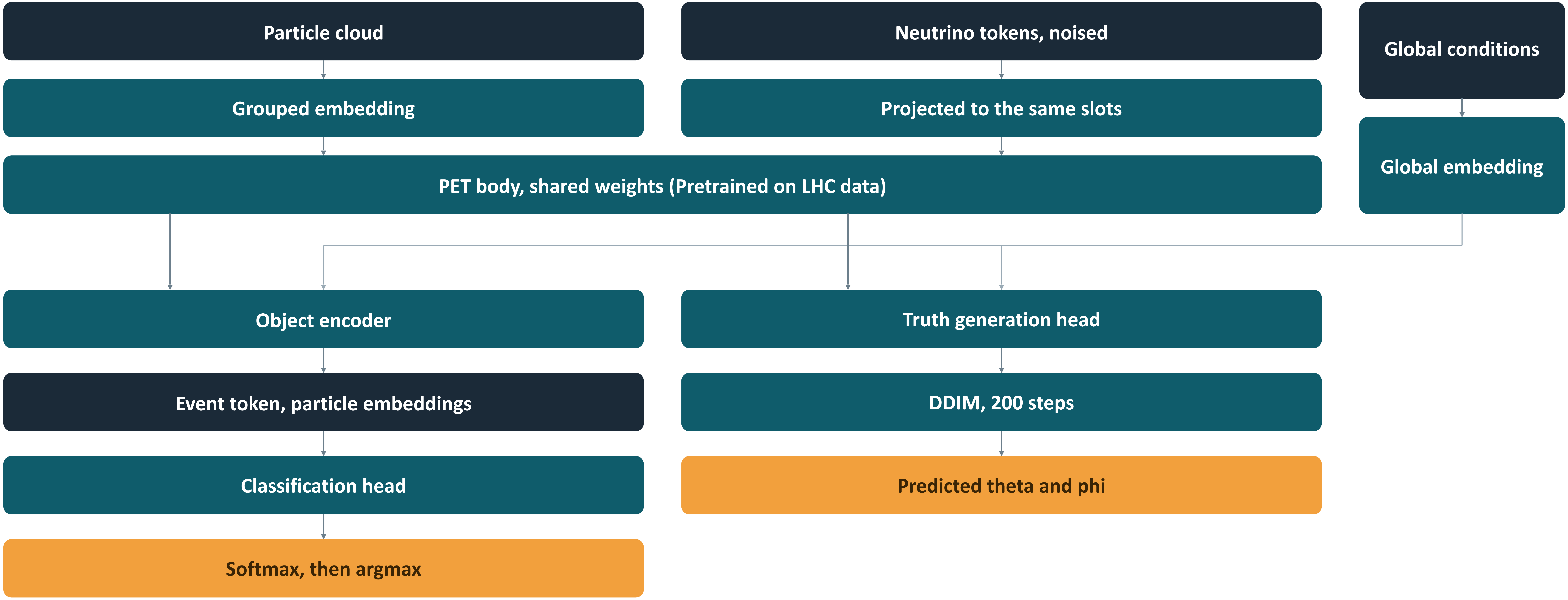
Workflow


DELPHI Open Data
Scale, Smearing
$\tau^+\tau^-$ Preselection
Traditional
Machine Learning (EveNet)
Decay mode Identification
Object-level cut-based ID
full-event classifier
ID Correction
ID Correction
HadHad
HadLep / LepLep
All channels
Neutrino Reconstruction
Analytic ν reconstruction
Likelihood ν reconstruction
Generative Likelihood-based τ reconstruction
Decay angle reconstruction by channel
Forward folding + POI fit
Extract quantum-correlation observables

Workflow


DELPHI Open Data
Scale, Smearing
$\tau^+\tau^-$ Preselection
Traditional
Machine Learning (EveNet)
Decay mode Identification
Object-level cut-based ID
full-event classifier
ID Correction
ID Correction
HadHad
HadLep / LepLep
All channels
Neutrino Reconstruction
Analytic ν reconstruction
Likelihood ν reconstruction
Generative Likelihood-based τ reconstruction
Decay angle reconstruction by channel
Forward folding + POI fit
Extract quantum-correlation observables

Model Structure (EveNet)



[EveNet]

Model Structure (EveNet)

Network Inputs

Particle cloud

HOW THE PARTICLE TENSOR IS BUILT

29 raw branches per particle

Pad to fixed length, keep a mask

Normalize, log for energy and pt

Grouped embedding, 29 into 7 slots

- **Momentum, 4** energy, pt, eta, phi → Slots 0 to 3
- **ObjectTag, 18** pdgId, HPC, HAC and STIC showers → Slots 4 and 5
- **Charge, 7** charge, dR, angle, hemisphere to leads → Slot 6

PET body

Global conditions

HOW THE CONDITION VECTOR IS BUILT

14 values per event

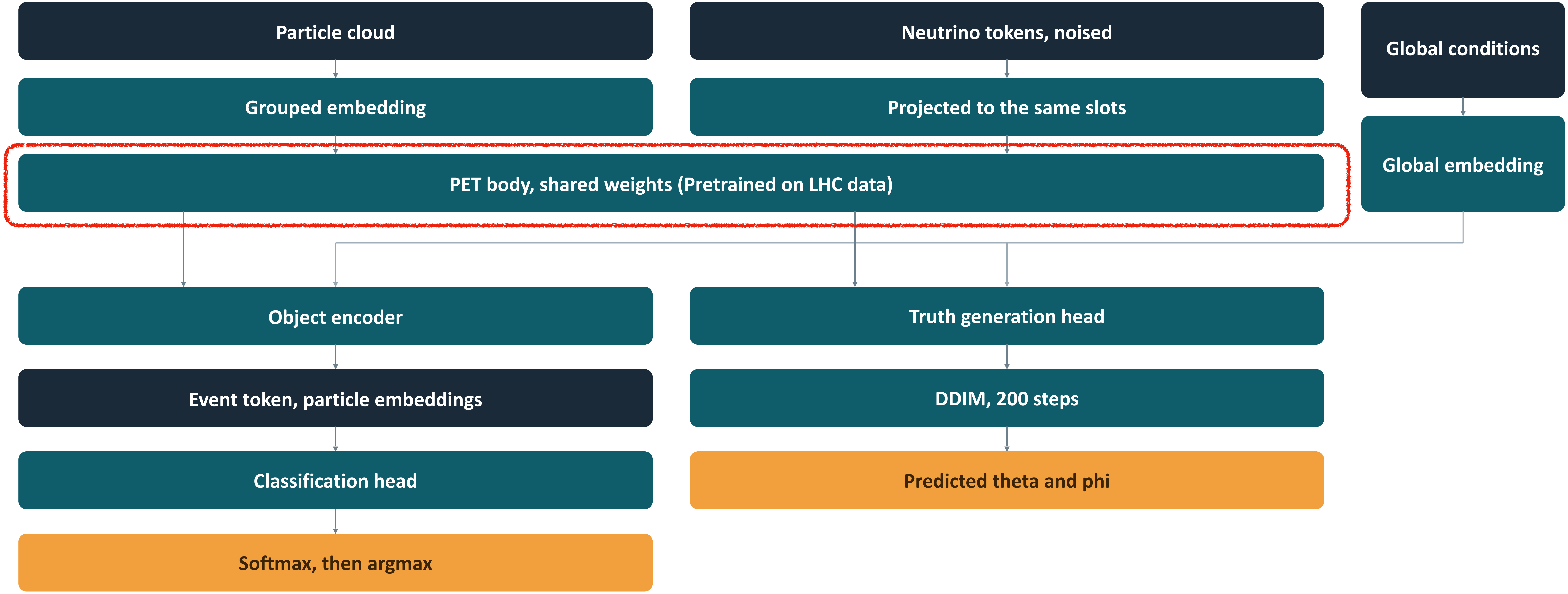
- **Thrust, 4** magnitude and x, y, z
- **Missing momentum, 3** px, py, pz
- **Lead candidates, 6** lead a and lead b momenta
- **Isolation, 1** isolation angle

Normalize each field

Global embedding, 8 GRU blocks

Joined with the particles in the object encoder

Model Structure (EveNet)



PET Body (Shared Body)

7 slots per particle

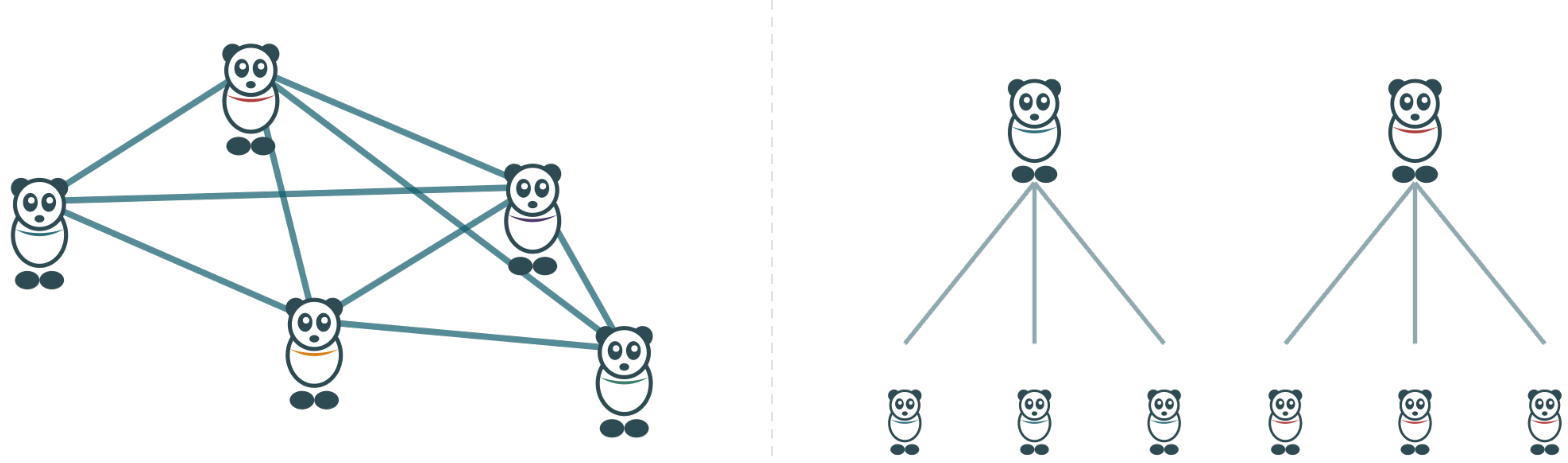
Local embedding

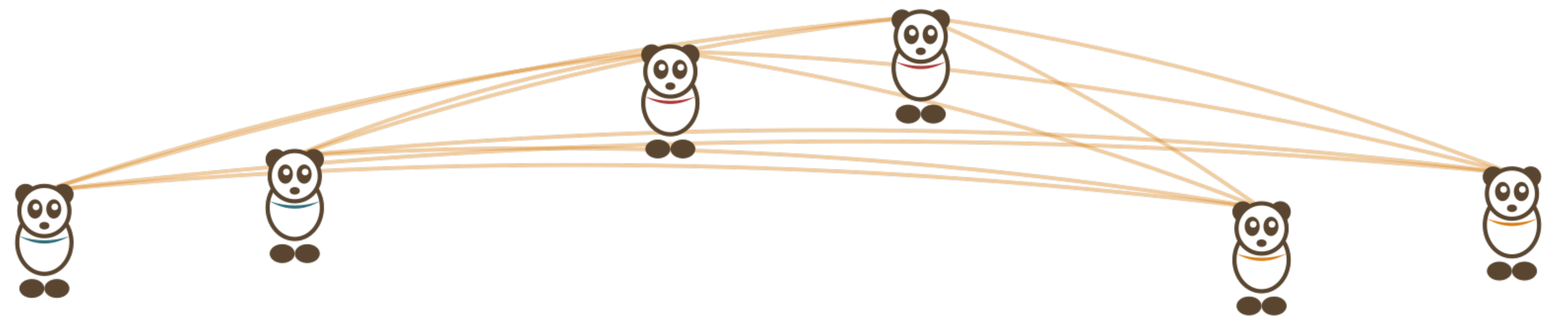
Feature embedding

Add the two branches

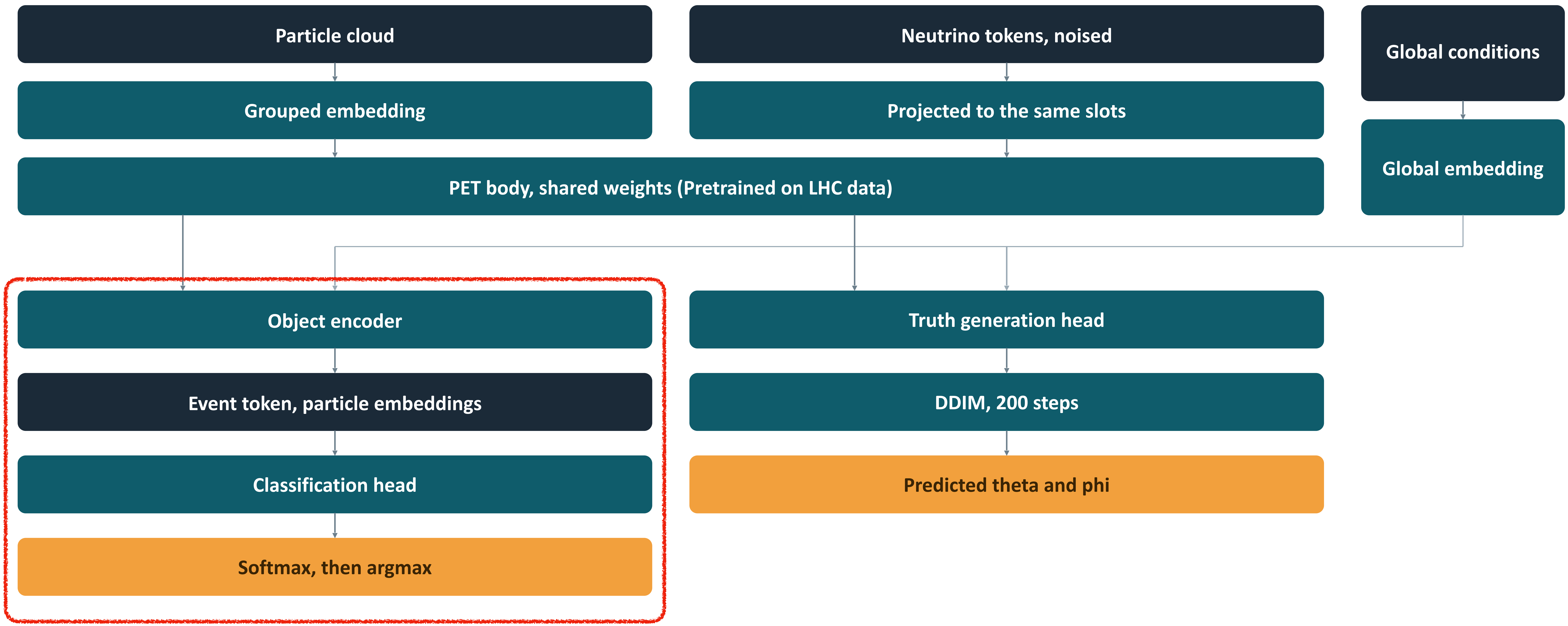
16 transformer layers

256 per particle

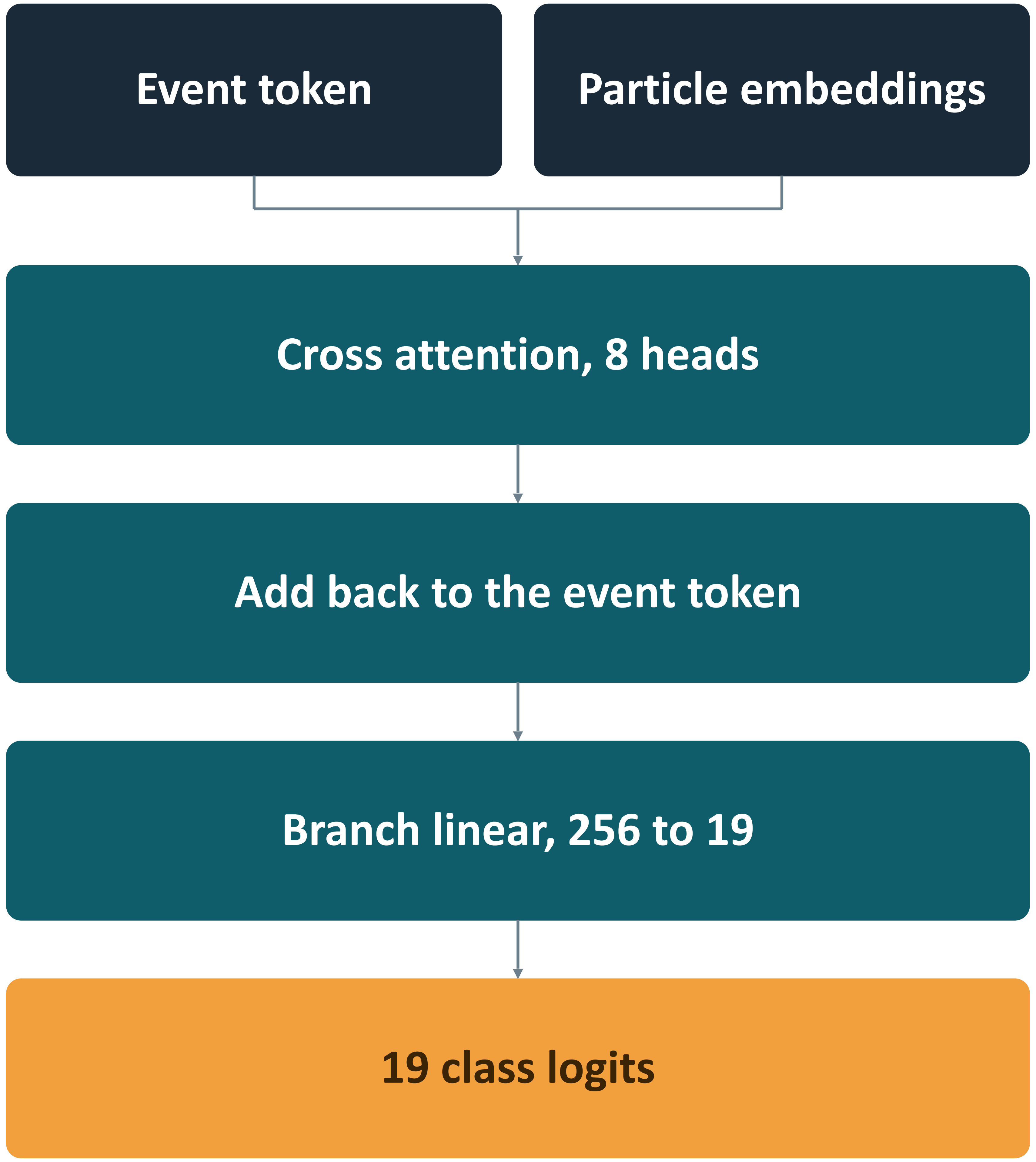




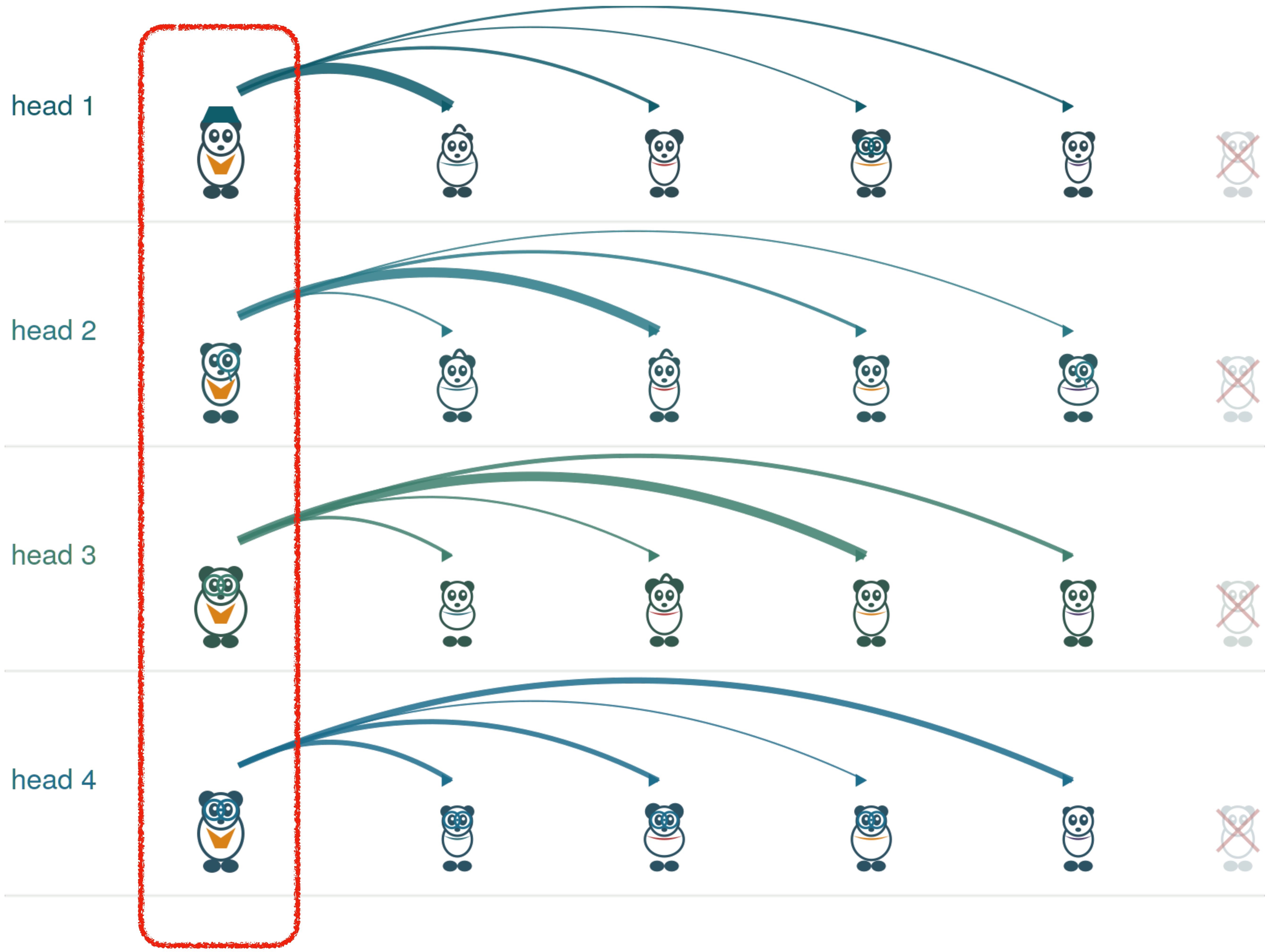
Model Structure (EveNet)



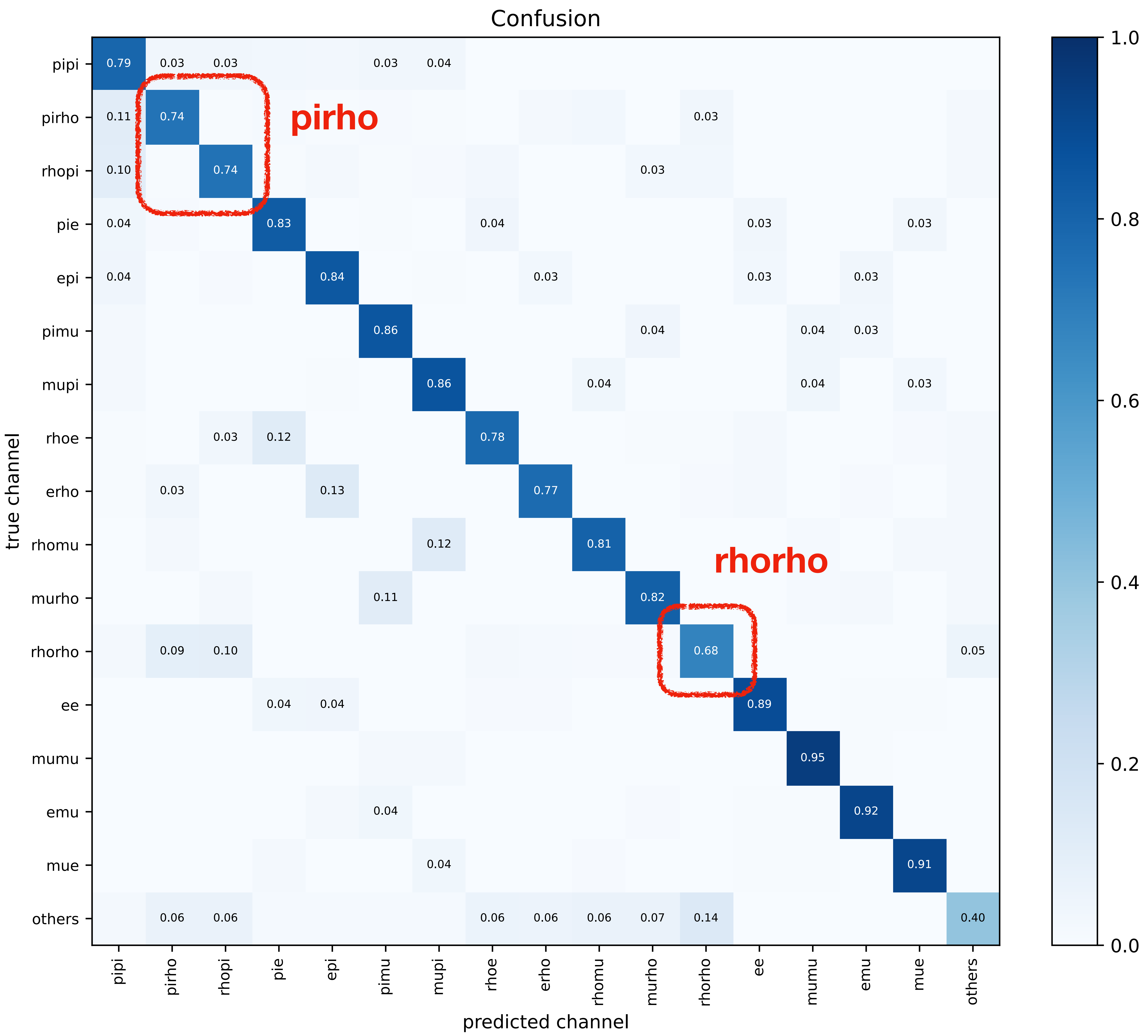
Classification Architecture



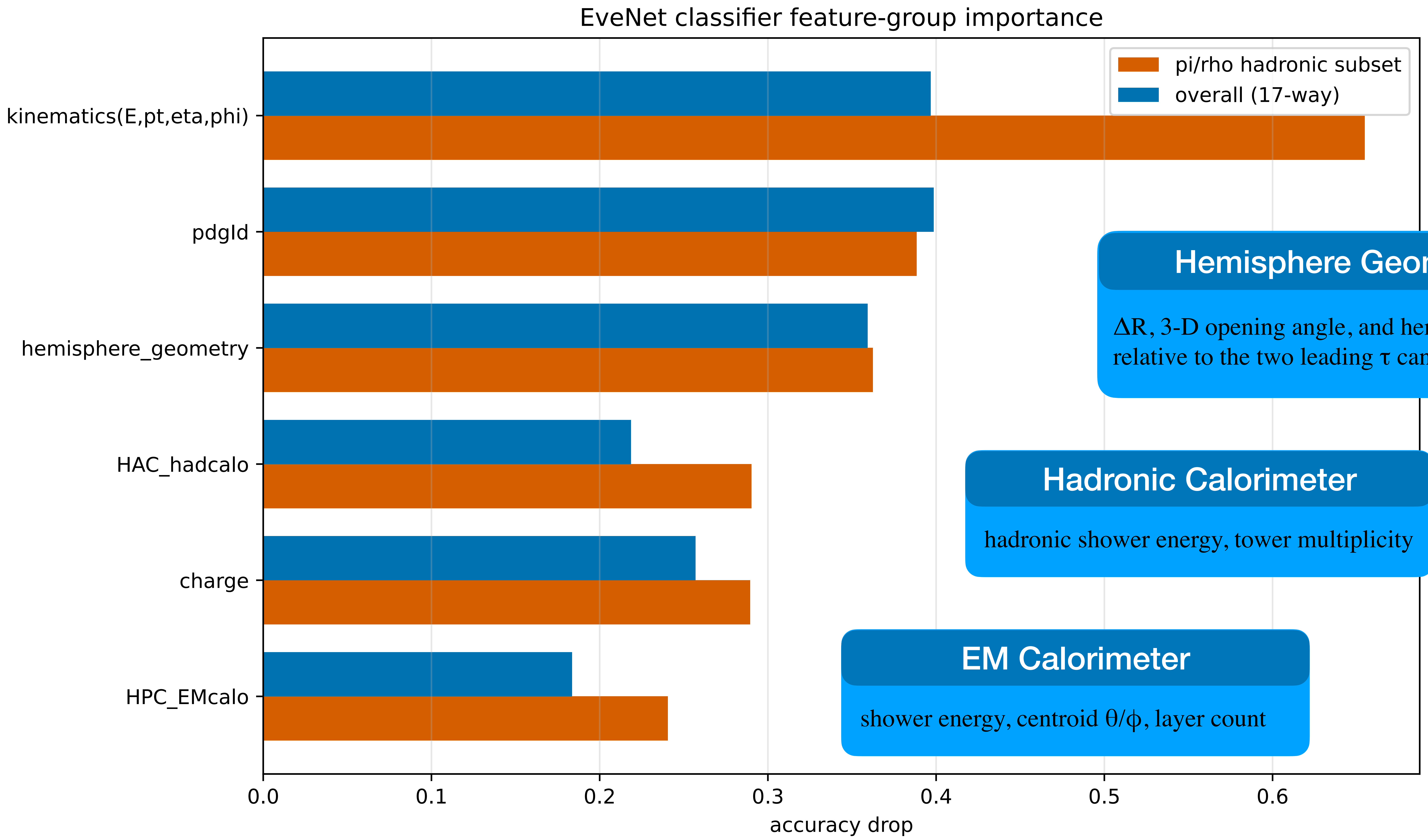
Event token pandas



A First Glance at the Results — Confusion Matrix



Feature Attribution



Hemisphere Geometry

ΔR, 3-D opening angle, and hemisphere flag relative to the two leading τ candidates

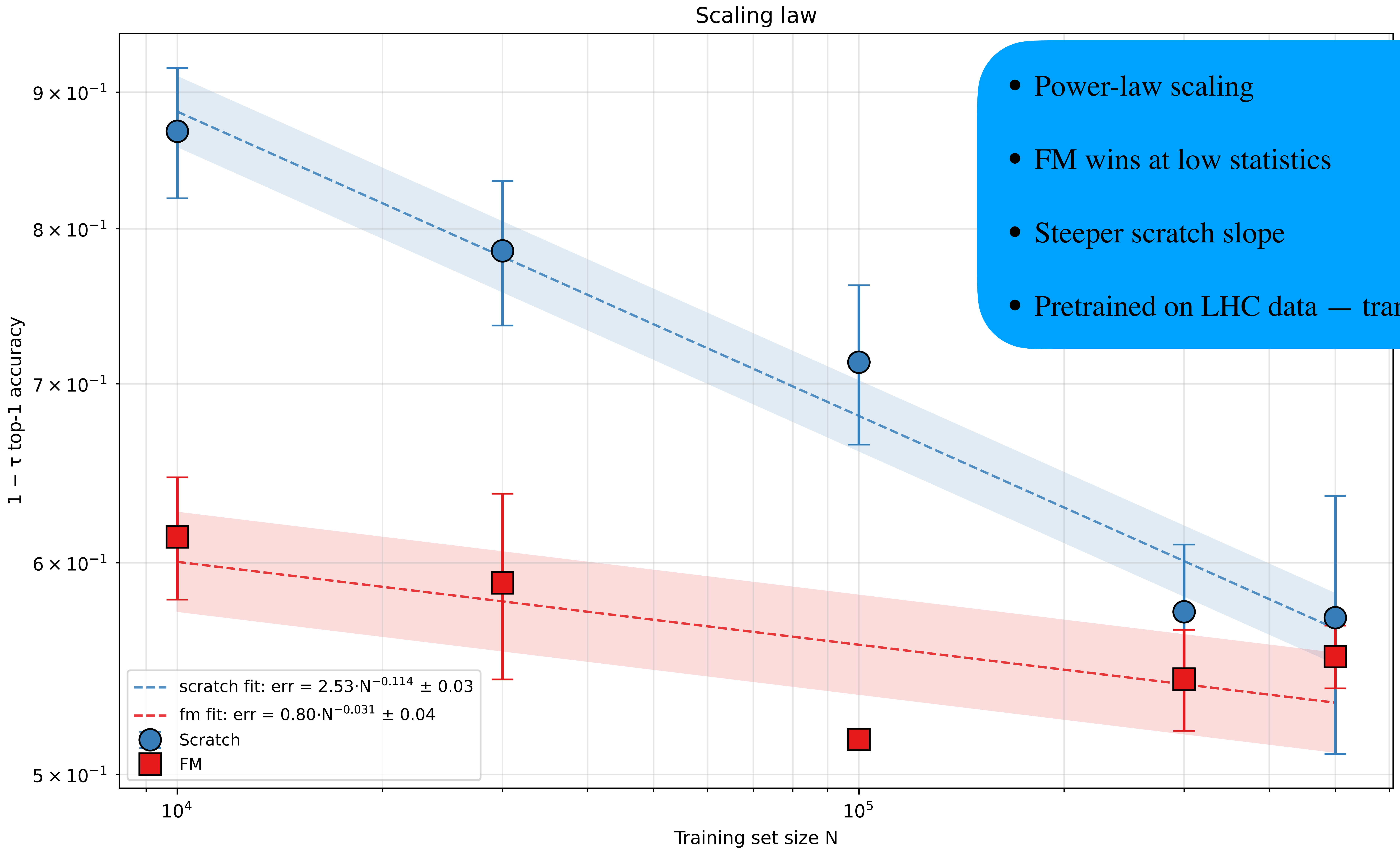
Hadronic Calorimeter

hadronic shower energy, tower multiplicity

EM Calorimeter

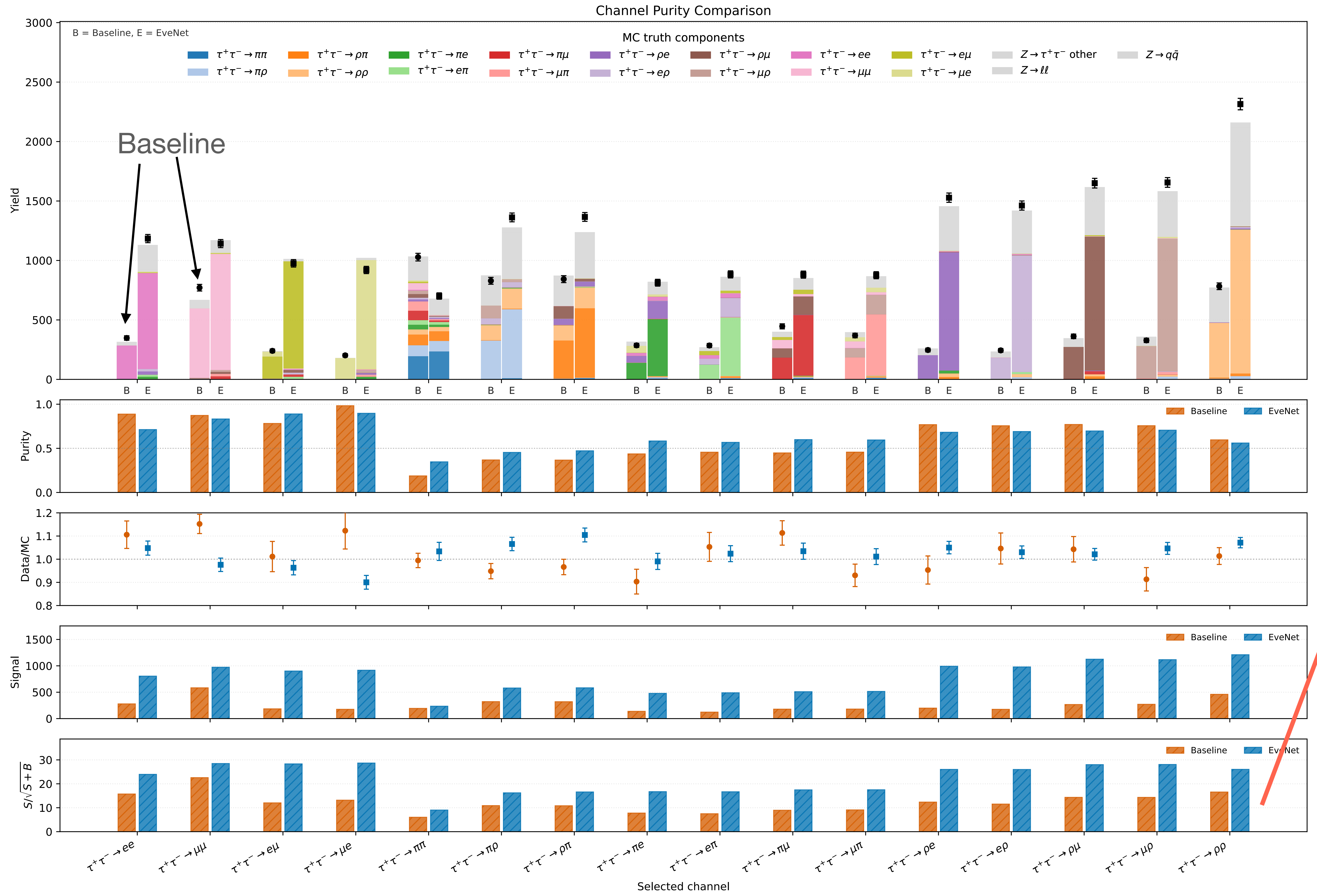
shower energy, centroid θ/ϕ, layer count

Scaling Law on Our Model



- Power-law scaling
- FM wins at low statistics
- Steeper scratch slope
- Pretrained on LHC data — transfer to LEP is real

Advantages of ML — Classification



Key Observations

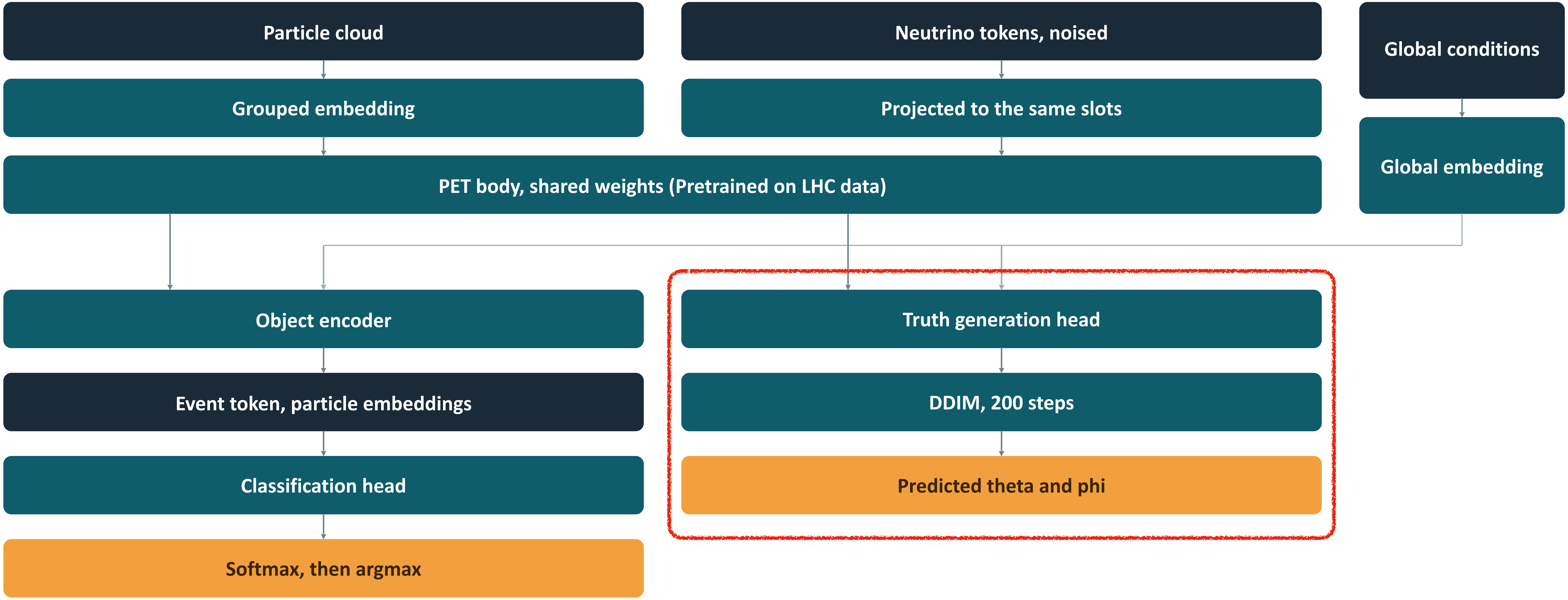
- ℓ channels: acceptance ↑ , purity mildly ↓
- ρ channels: acceptance ↑ , purity mildly ↓
- π channels: purity ↑ , acceptance mildly ↑

Significance Gain

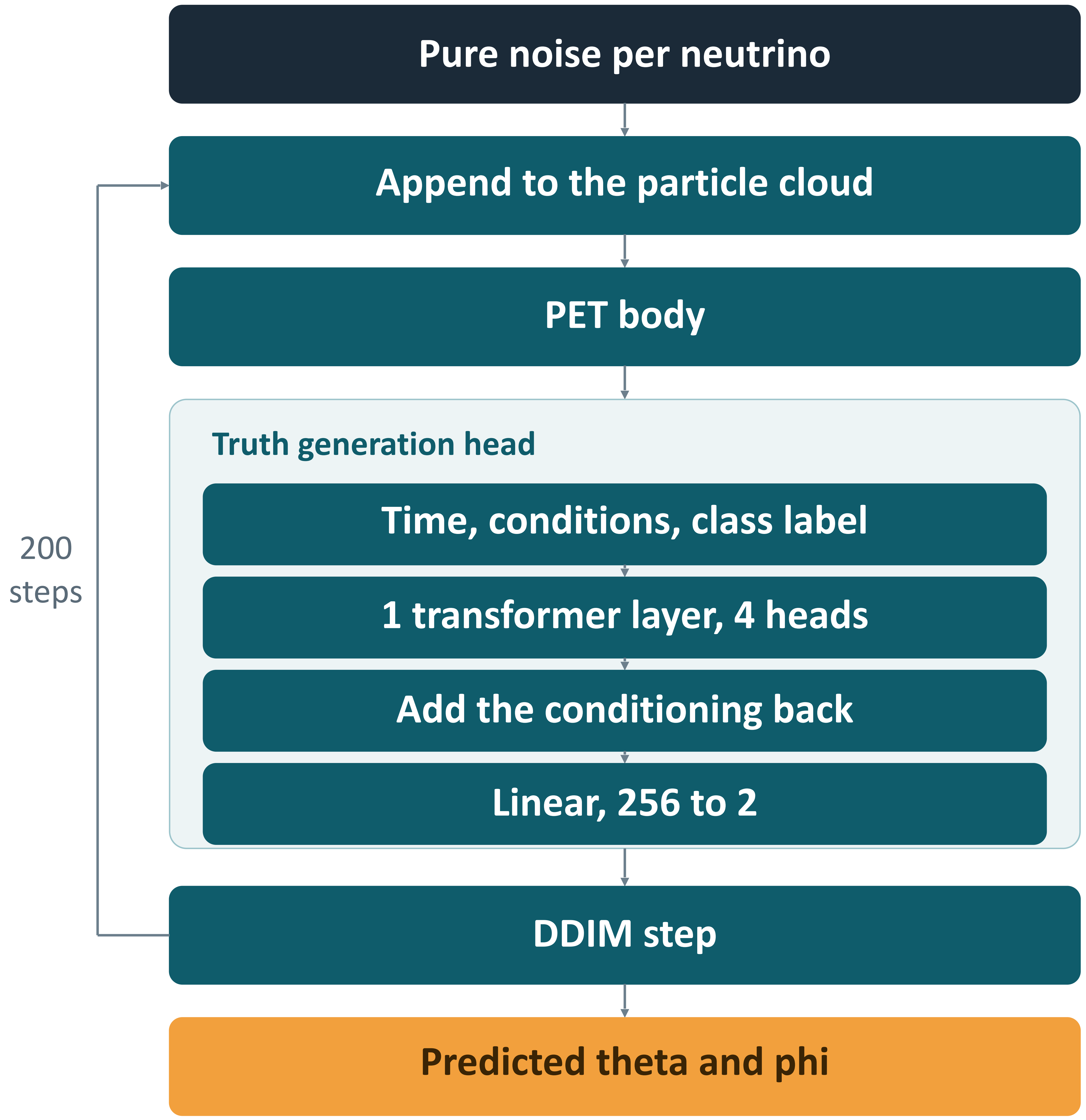
- $\ell\ell$: +67%
- ℓ had: +110%
- $\pi\pi$: +50%
- Other had/had: +58%

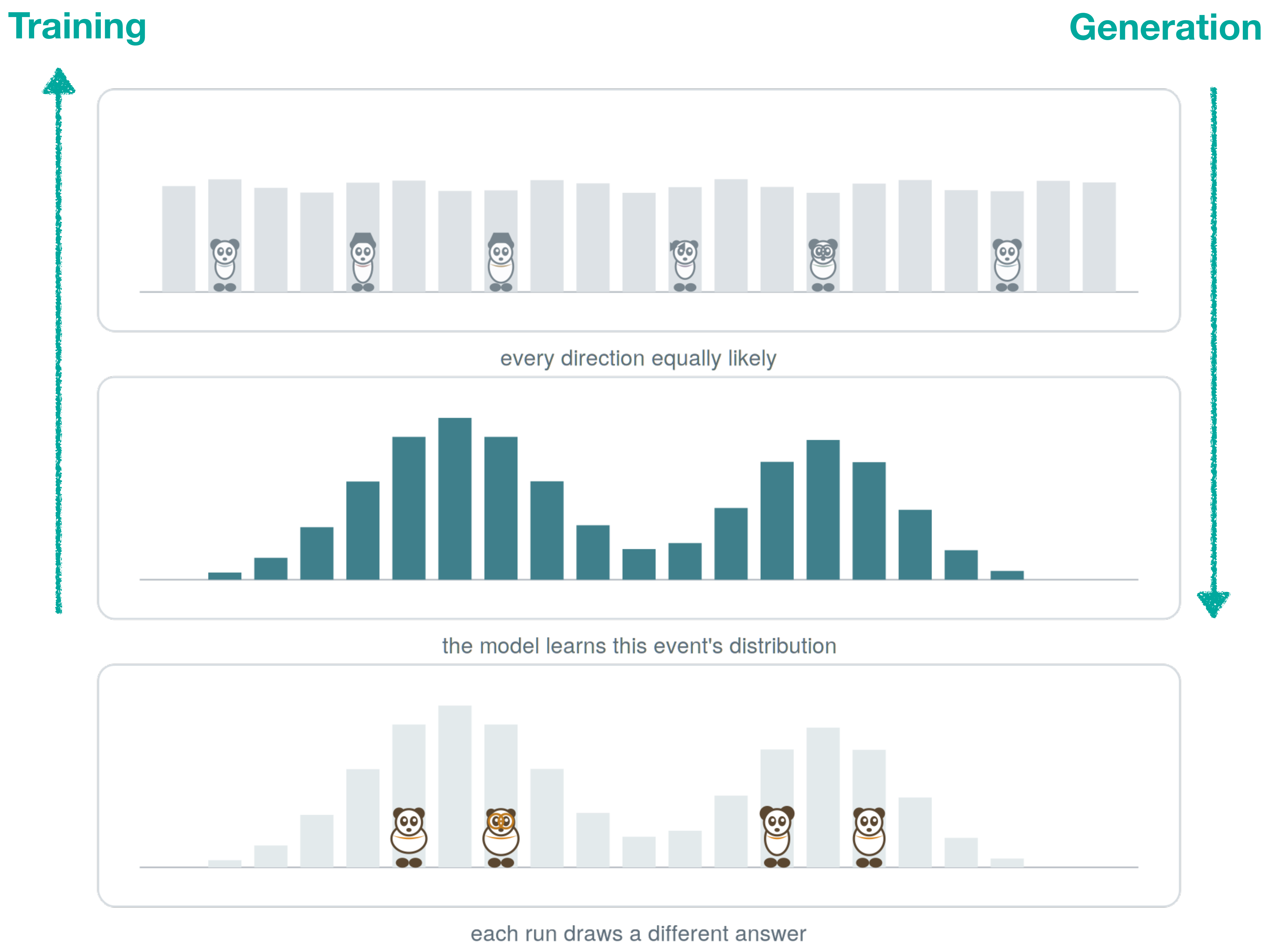
ML significantly increases the signal significance

Model Structure (EveNet)

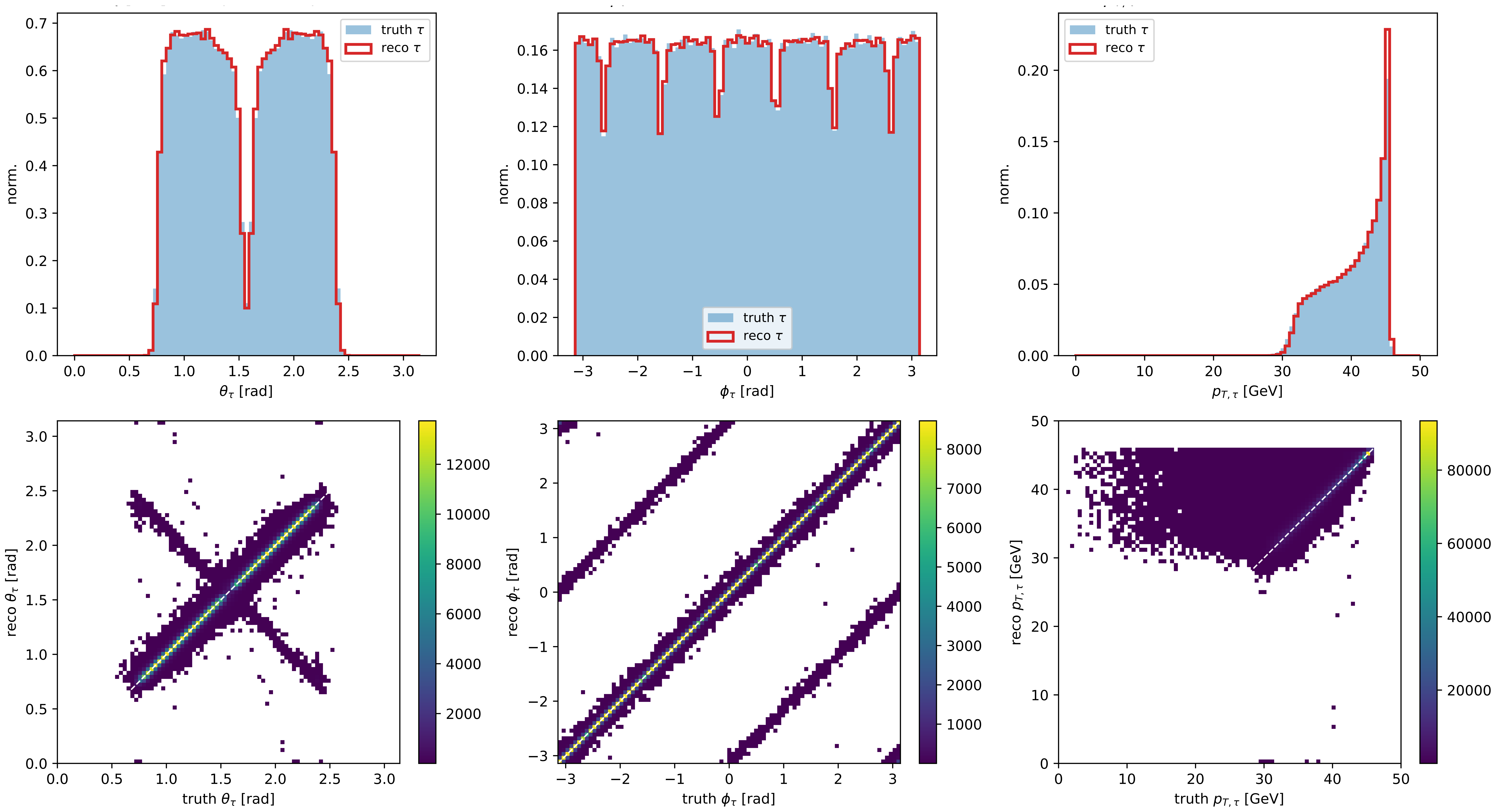


Neutrino Reconstruction Architecture





Neutrino Reconstruction Result



Advantages of ML — Neutrino Reconstruction

ML stays diagonal and robust in response matrix , even where the traditional method fails

Response Matrix: P(reco|truth)
It's diagonal if reco & truth are the same

$\tau^+\tau^- \to \pi^+\pi^-$

$\tau^+\tau^- \to e^+e^-$

Result in Asimov Data

Traditional

Combined $\tau^+\tau^-$ spin density matrix

	l	n	r	k
l	1	+0.00 ±0.13	-0.00 ±0.13	+0.15 ±0.08
n	+0.00 ±0.13	+0.81 ±0.25	+0.00 ±0.63	+0.00 ±0.53
r	+0.00 ±0.13	-0.01 ±0.63	-0.80 ±0.32	-0.00 ±0.51
k	+0.15 ±0.08	-0.01 ±0.53	-0.01 ±0.51	+1.03 ±0.30

BELL NONLOCALITY
0.424± 0.393
1.1σ

CONCURRENCE
0.81 ± 0.264
3.1σ

ML-based

Combined $\tau^+\tau^-$ spin density matrix

	l	n	r	k
l	1	+0.00 ±0.08	-0.00 ±0.08	+0.15 ±0.05
n	+0.00 ±0.08	+0.81 ±0.21	+0.00 ±0.58	+0.00 ±0.35
r	+0.00 ±0.08	-0.01 ±0.58	-0.80 ±0.24	-0.00 ±0.34
k	+0.15 ±0.05	-0.01 ±0.35	-0.01 ±0.34	+1.03 ±0.23

BELL NONLOCALITY
0.424± 0.311
1.4σ

CONCURRENCE
0.81 ± 0.20
4.1σ

Conclusion

1st

OF ITS KIND

Toward a measurement of the full tau-pair spin density matrix with archived DELPHI data

Transfer

LHC Pretrained Backbone

↓

LEP $Z \to \tau^+\tau^-$

Pipeline

Classification

\+

Neutrino Reconstruction

Progress

Higer Purity

Higer Acceptance

More Robust Result

Ongoing

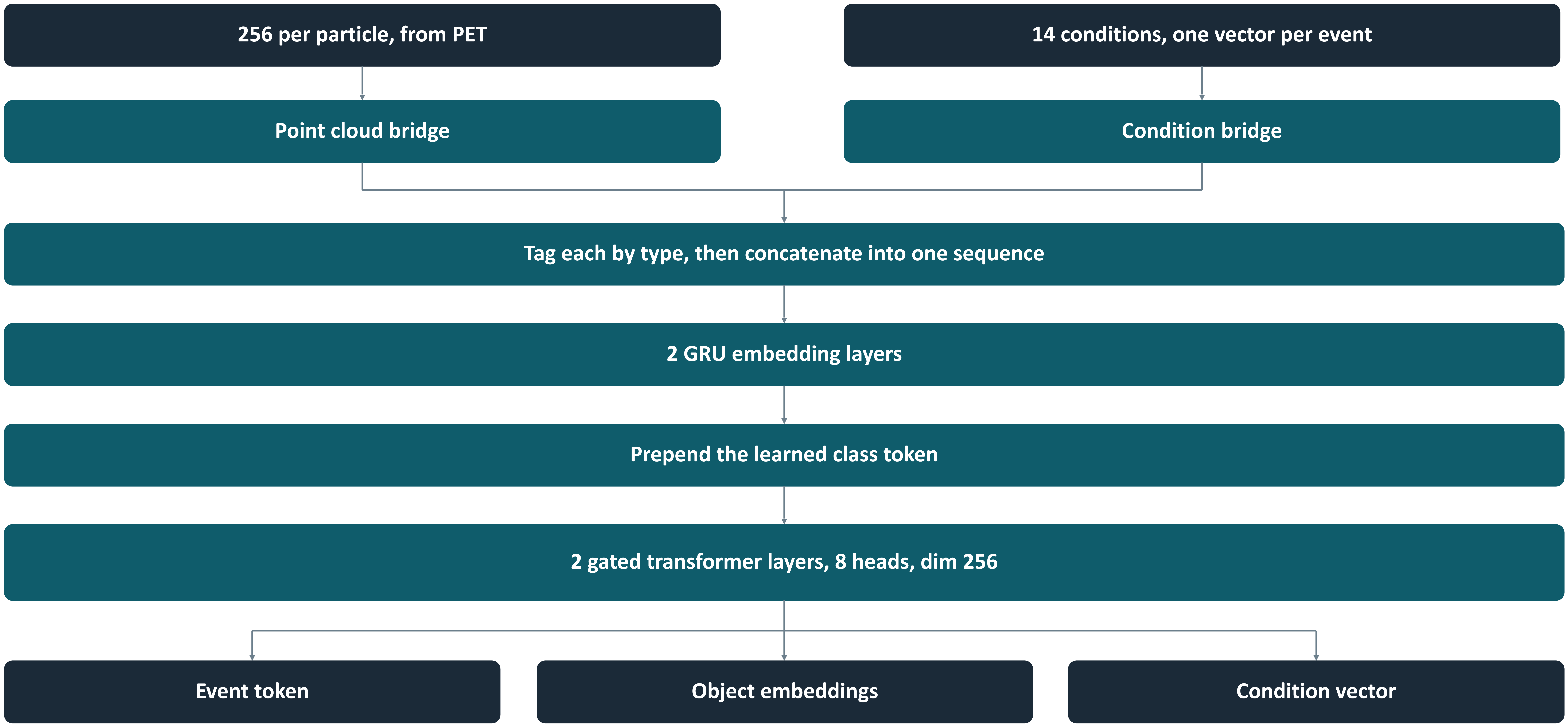
With RL method

from Ting-Hsiang's

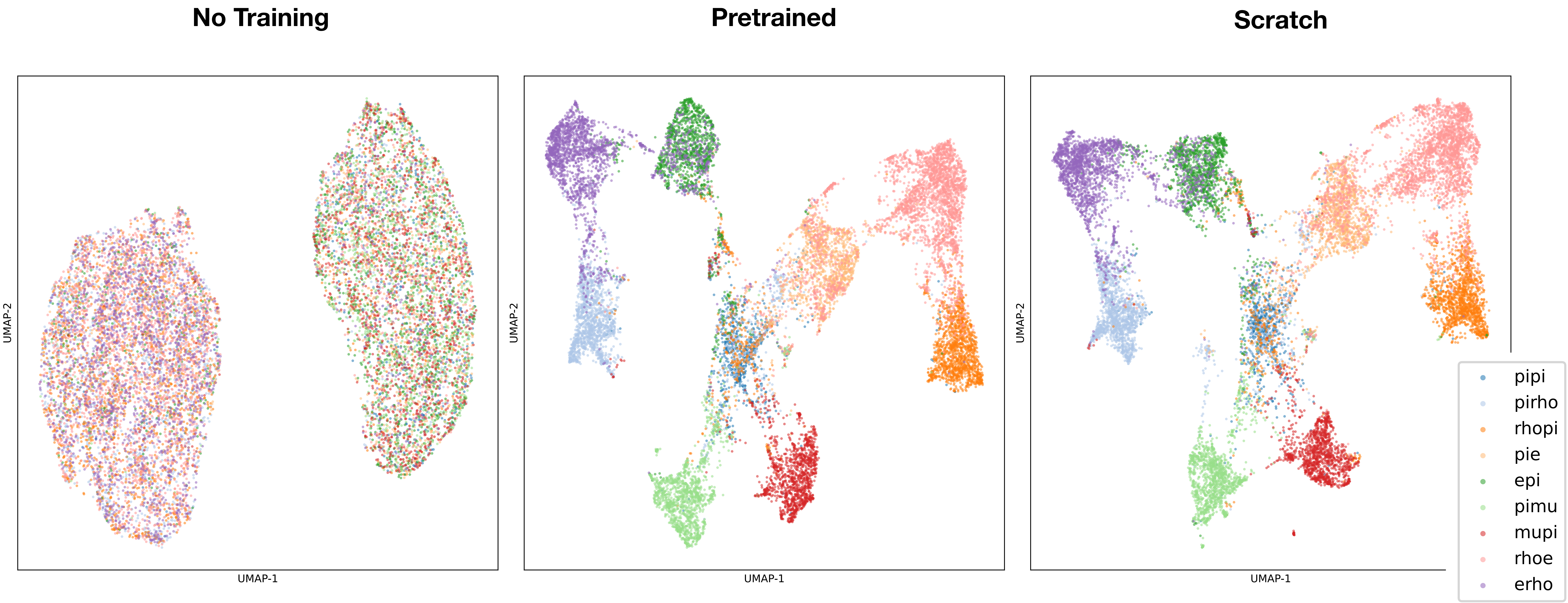
presentation

Back up

Object Encoder



Pretrain vs. Scratch



INPUT

Inputs

- Reconstructed object point cloud
- Kinematics E, p_T, η, φ
- Detector / PID features
- Charge

ENCODER · BACKBONE

EveNet

Point-edge transformers capture local object relations and event-level context.

Pretrained on LHC data
→ transferred to LEP analysis

Event-level $\tau\tau$ classification

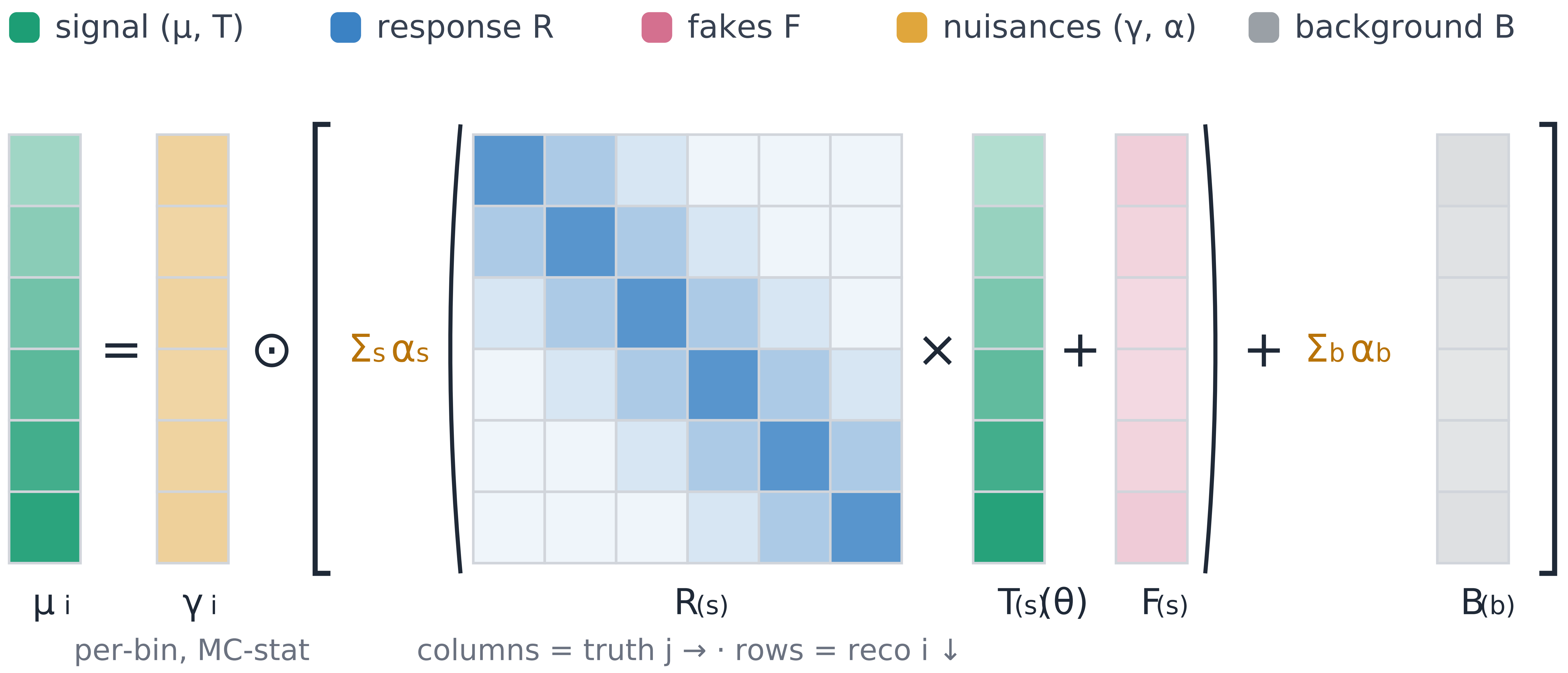
- Uses correlations among objects and event information
- **Main gain:** higher signal acceptance at comparable purity

$\tau\tau$ kinematics reconstruction

- **Diffusion model** reconstructs the angle between τ and its visible decay products
- Post-calibration enforces $\tau^+\tau^-$ system consistency

[EveNet]

Forward-Folding



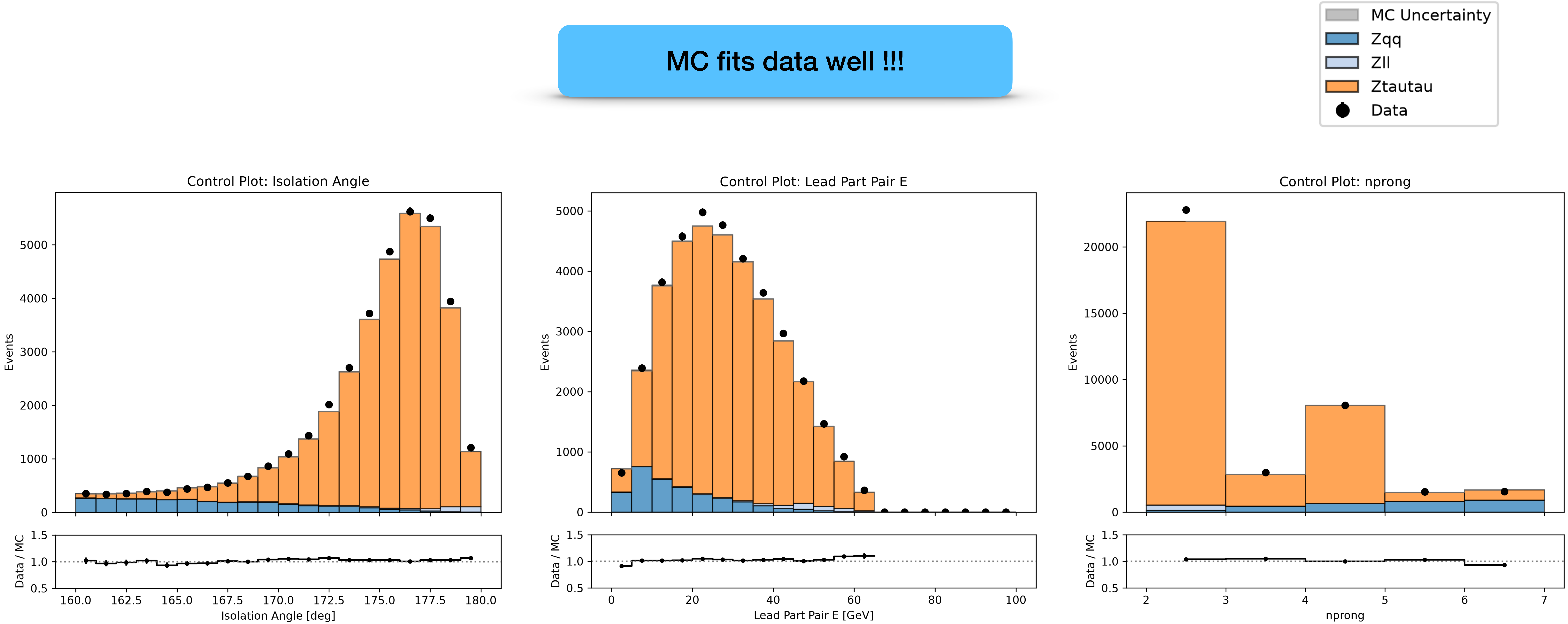
$$-2\ln\mathscr{L}(\theta_{POI},\vec{\alpha},\vec{\gamma}) = \underbrace{2\sum_i\left[\mu_i - n_i + n_i\ln\frac{n_i}{\mu_i}\right]}_{\textbf{Poisson (Baker-Cousins)}} + \underbrace{\sum_k \alpha_k^2}_{\textbf{Gauss on }\alpha} + \underbrace{\sum_i 2\beta_i\left(\gamma_i - 1 - \ln\gamma_i\right)}_{\textbf{Gamma on }\gamma}$$

$$\hat{\theta}_{POI} = \arg\min_{\theta_{POI}}\left[\min_{\vec{\alpha},\vec{\gamma}}\left(-2\ln\mathscr{L}\right)\right]$$

$$\Delta(-2\ln\mathscr{L}) = 1 \qquad \Rightarrow \pm 1\sigma$$

$$\Delta(-2\ln\mathscr{L}) = \chi^2_{0.95,1} \approx 3.84 \qquad \Rightarrow\ 95\,\%\ \textbf{CL}$$

MC fits data well !!!



Baseline Neutrino Reconstruction

Hadron - Analytic method

Only one massless missing ν_τ from τ decay.

With the constraint from m_τ , missing energy, we could solve ν_τ analytically.

Lepton - MMC method

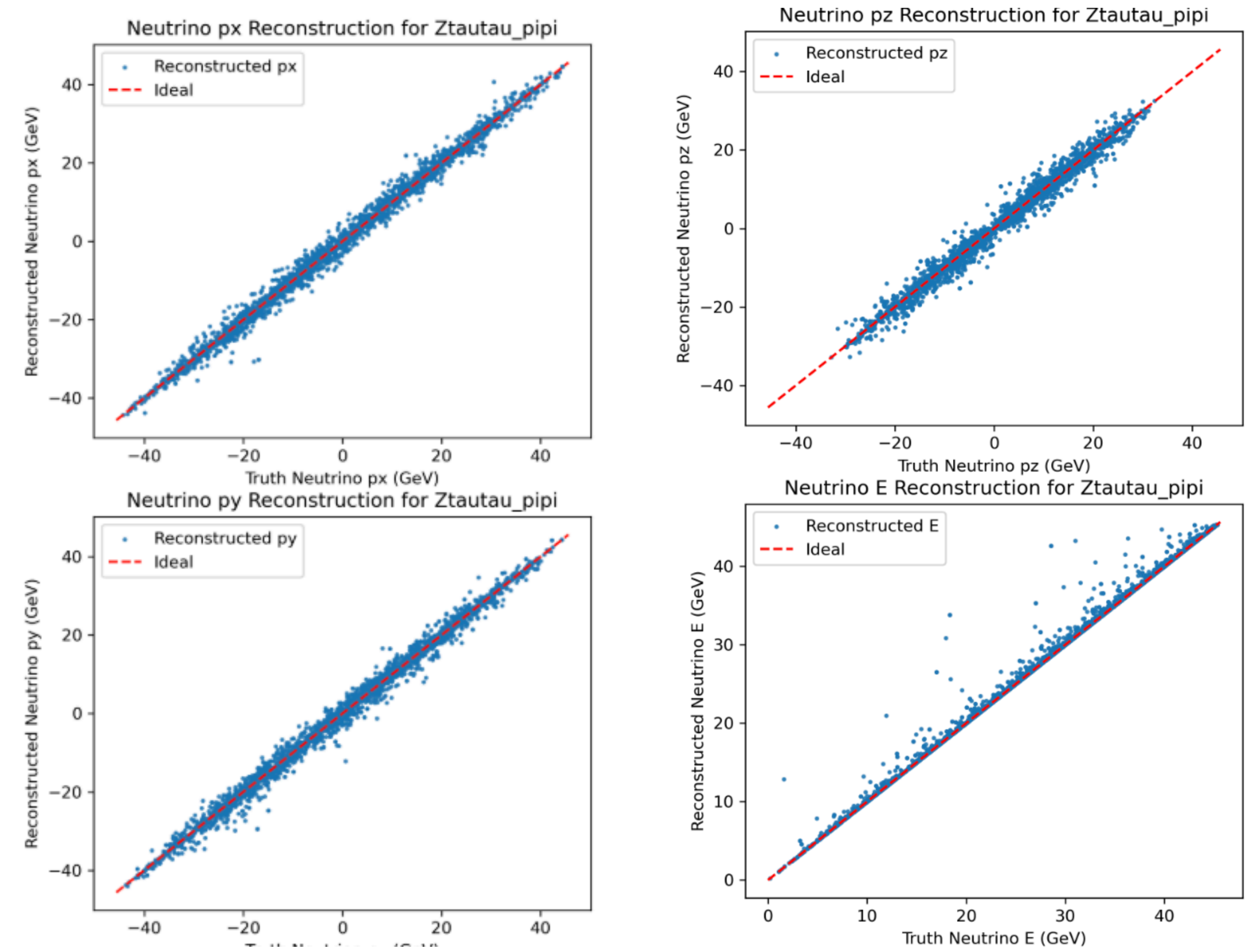
Assume back-to-back $\tau\tau$ pair with $E = \dfrac{\sqrt{s}}{2}$, scan $\theta_{\nu_1}, \phi_{\nu_1}, \theta_{\nu_2}, \phi_{\nu_2}$ to maximize the likelihood

$$\begin{aligned}\mathscr{L}_{kinematics} &= P(\Delta R \,|\, p_\tau)\\ &= w \cdot \mathcal{N}(\Delta R \,|\, \mu, \sigma) + (1-w)L(\Delta R \,|\, u, c)\end{aligned}$$

$$\mathscr{L}_{phase\ space} = m_\nu^2 * (m_\tau^2 - m_\nu^2)^2 \Theta((m_\tau - m_\mu)^2 - m_\nu^2)\Theta(m_\nu)$$

$$\mathscr{L} = \mathscr{L}^{\tau_1}_{phase\ space}\mathscr{L}^{\tau_2}_{phase\ space}\mathscr{L}^{\tau_1}_{kinematics}\mathscr{L}^{\tau_2}_{kinematics}$$

$\pi\pi$



ee