Studying the Primordial Fluid with Deep Learning

Martim Pinto (IST)

Supervised by João A. Gonçalves (LIP)

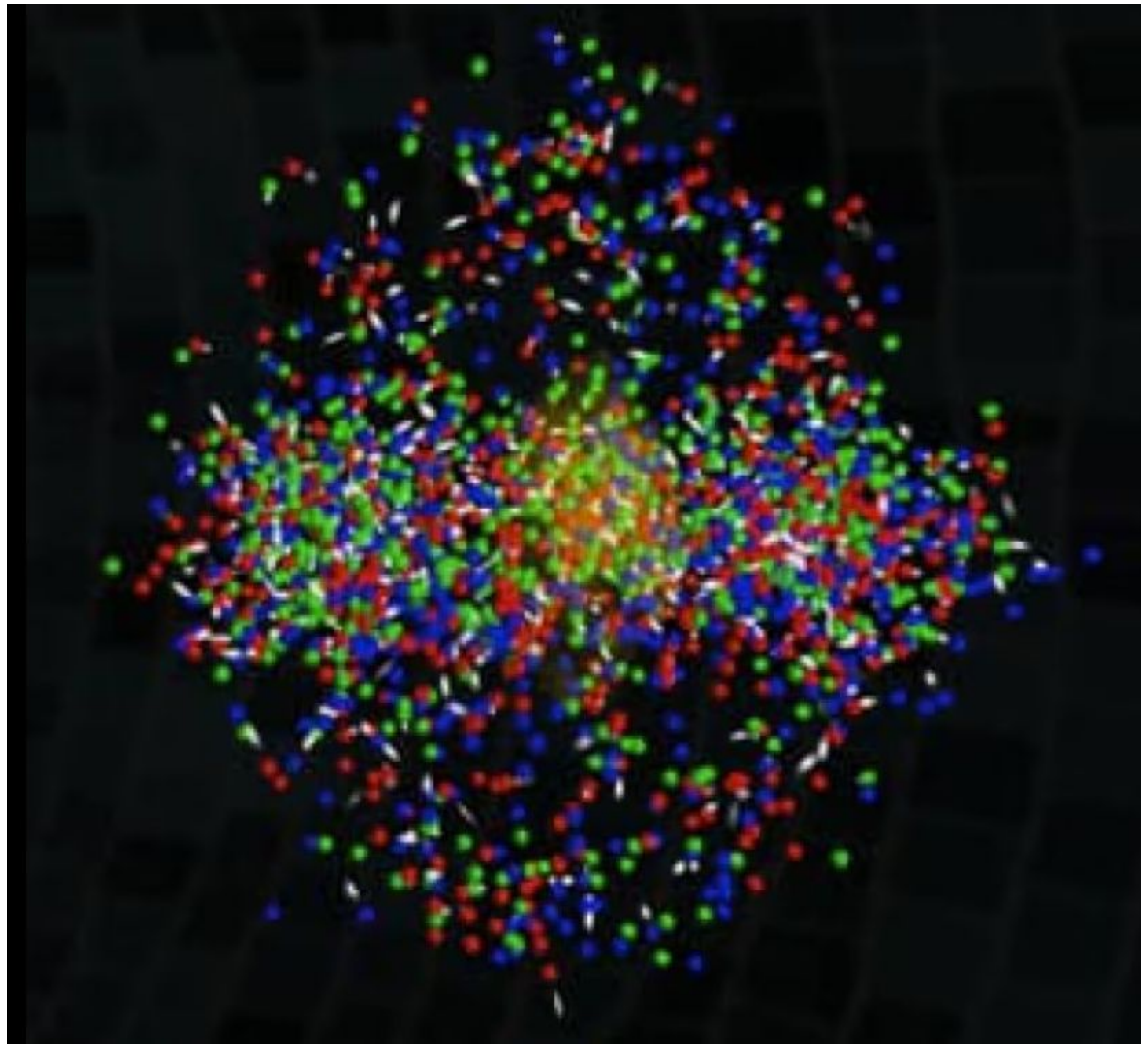
[1] LIP / Laboratório de Instrumentação e Física Experimental de Partículas
[2] IST/ Instituto Superior Técnico

Table of Contents

Introduction –Quark Gluon Plasma

- exotic state of matter that existed in the very early universe in extremely hot and dense conditions, just microseconds after the Big Bang

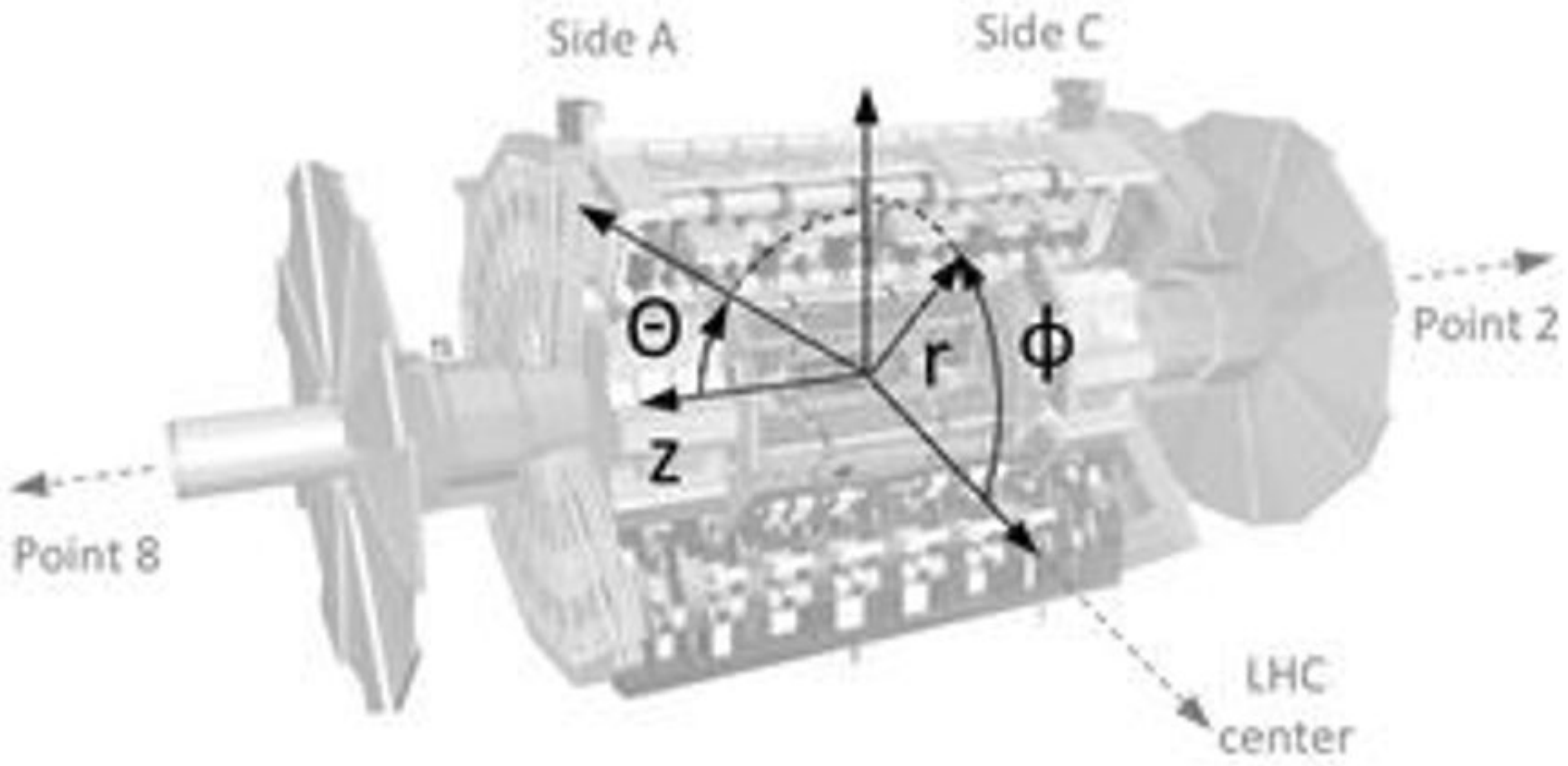
- In this state, **quarks** and **gluons** were not confined inside protons and neutrons. Instead, they moved freely in a kind of "primordial soup" or **plasma**



computer simulation of the QG soup in RHIC (Brookhaven National Laboratory)

Introduction –What is a Jet?

- A stream of aligned particles produced by a high-momentum parton: quarks or gluons (ambiguous)



- Jets are defined by their **size** and **clustering algorithm**, that clusters particles based on the “distances” between each other

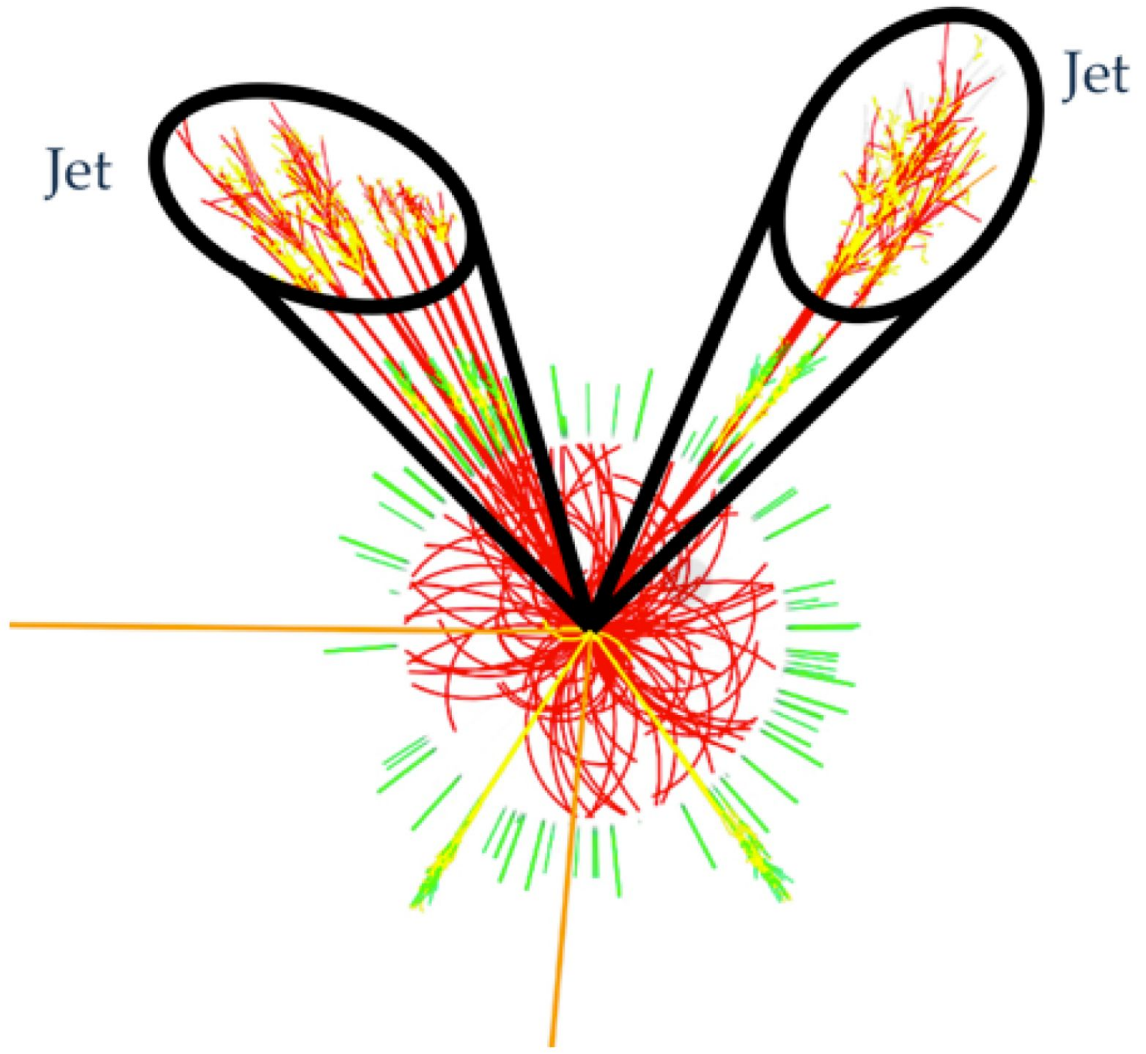
- Jets are suppressed and modified in heavy ion collisions, which serve as powerful probes to the properties of the QGP

Relevant kinematic variables:

- Transverse momentum: p_T
- Rapidity: $y = ½ \cdot \ln (E-p_z)/(E+p_z)$
- Pseudorapidity: $\eta = -\ln \tan ½\theta$
- Azimuthal angle: φ

Introduction –What is a Jet?

- A stream of aligned particles produced by a high-momentum parton: quarks or gluons (ambiguous)



Agafonova, Veronika. "Study of Jet Shape Observables in Au+Au Collisions at S N N = 200 GeV with JEWEL." *Universe*, vol. 5, no. 5, 2019, p. 114, https://doi.org/10.3390/universe5050114. Accessed 1 Sept. 2024.

- Jets are defined by their **size** and **clustering algorithm**, that clusters particles based on the "distances" between each other

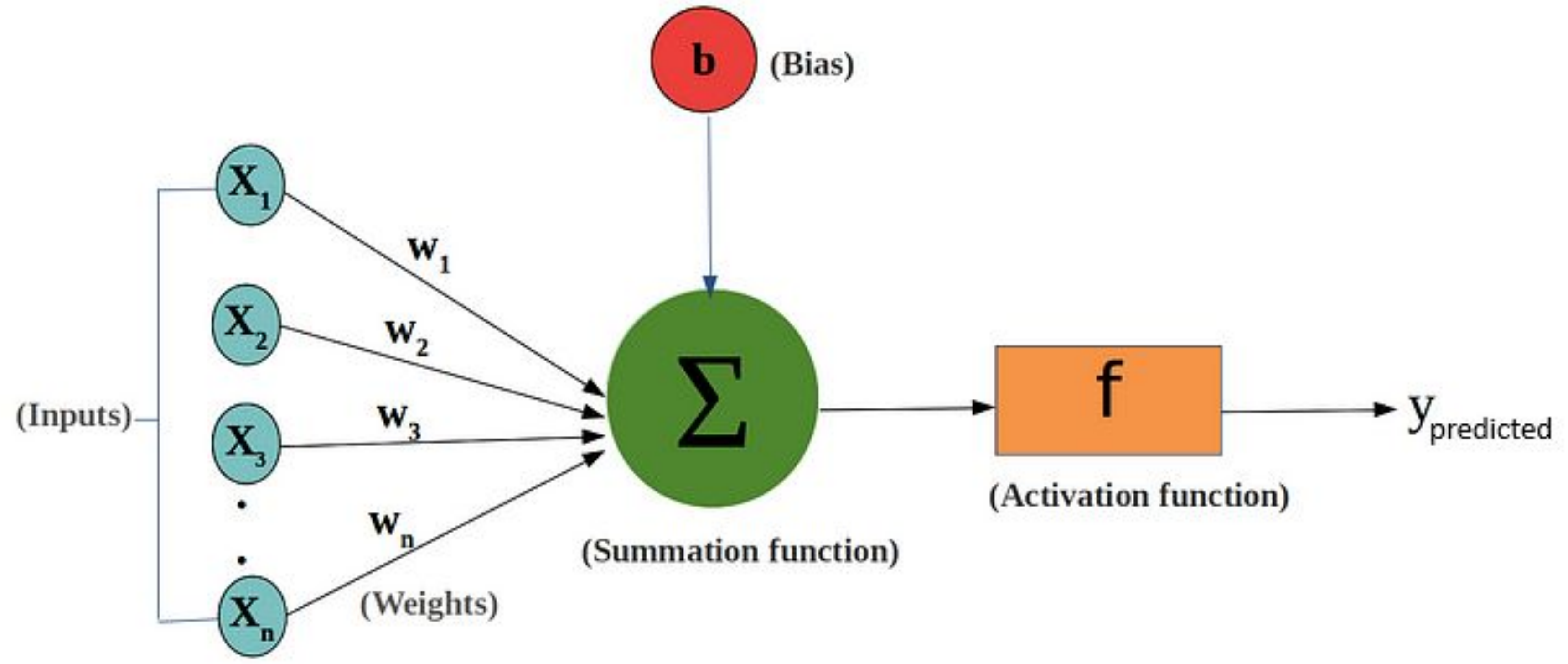
$$\Delta R_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$$

$$d_{ij} = min(p_{T,i}^{2p}, p_{T,i}^{2p}) \frac{\Delta R_{ij}^2}{R^2} \begin{cases} p = 1 & k_T \ algorithm \\ p = 0 & Cambridge/Aachen \\ p = -1 & anti - k_T \ algorithm \end{cases}$$

- Jets are suppressed and modified in heavy ion collisions, which serve as powerful probes to the properties of the QGP

- A **Neural Network** is a computational model that consists of layers of interconnected nodes or "neurons“, which work together to perform complex tasks.

- A **linear transformation** is applied to each layer’s previous outputs (none for inputs) followed by a point-wise **non-linear activation function**.



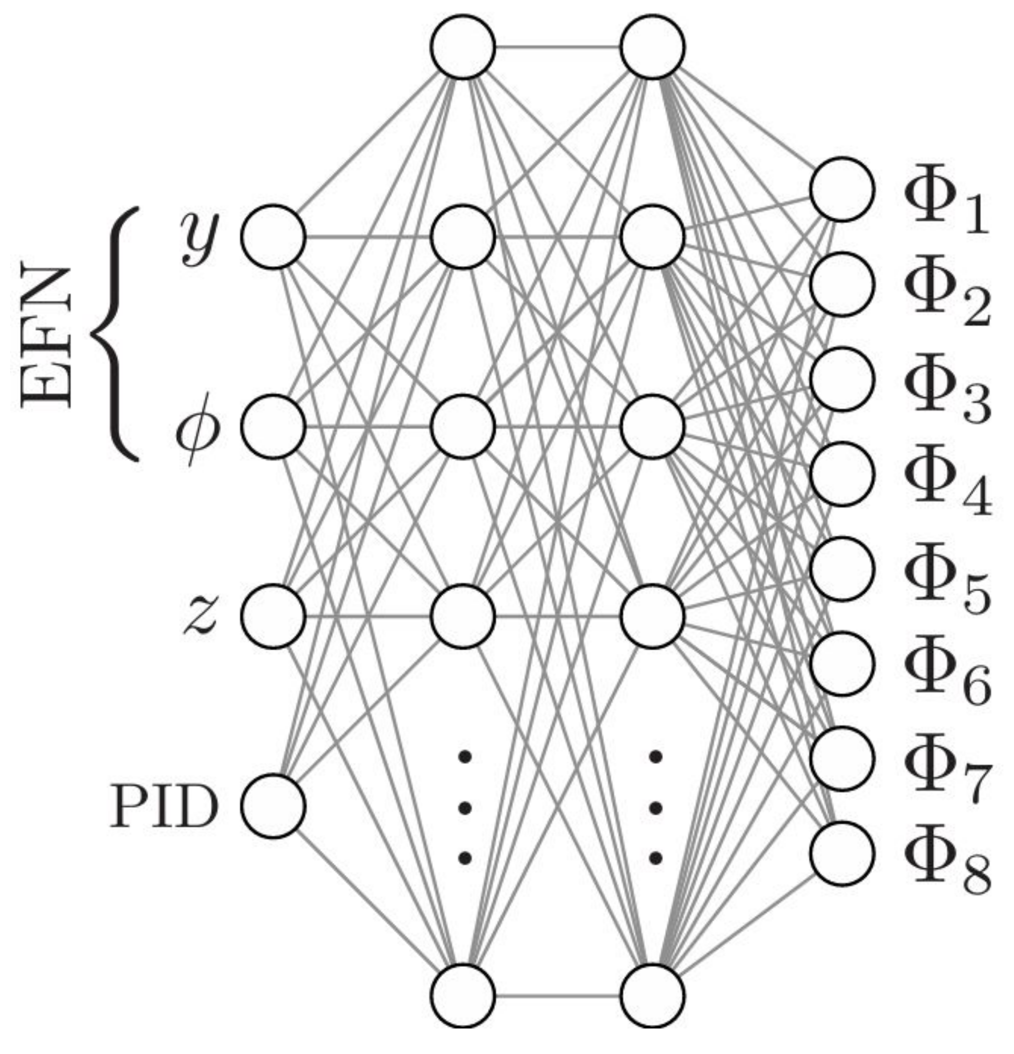
- Used for **extracting relevant features** from raw data or **learning to generalize** from training data to make accurate predictions on unseen (test) data

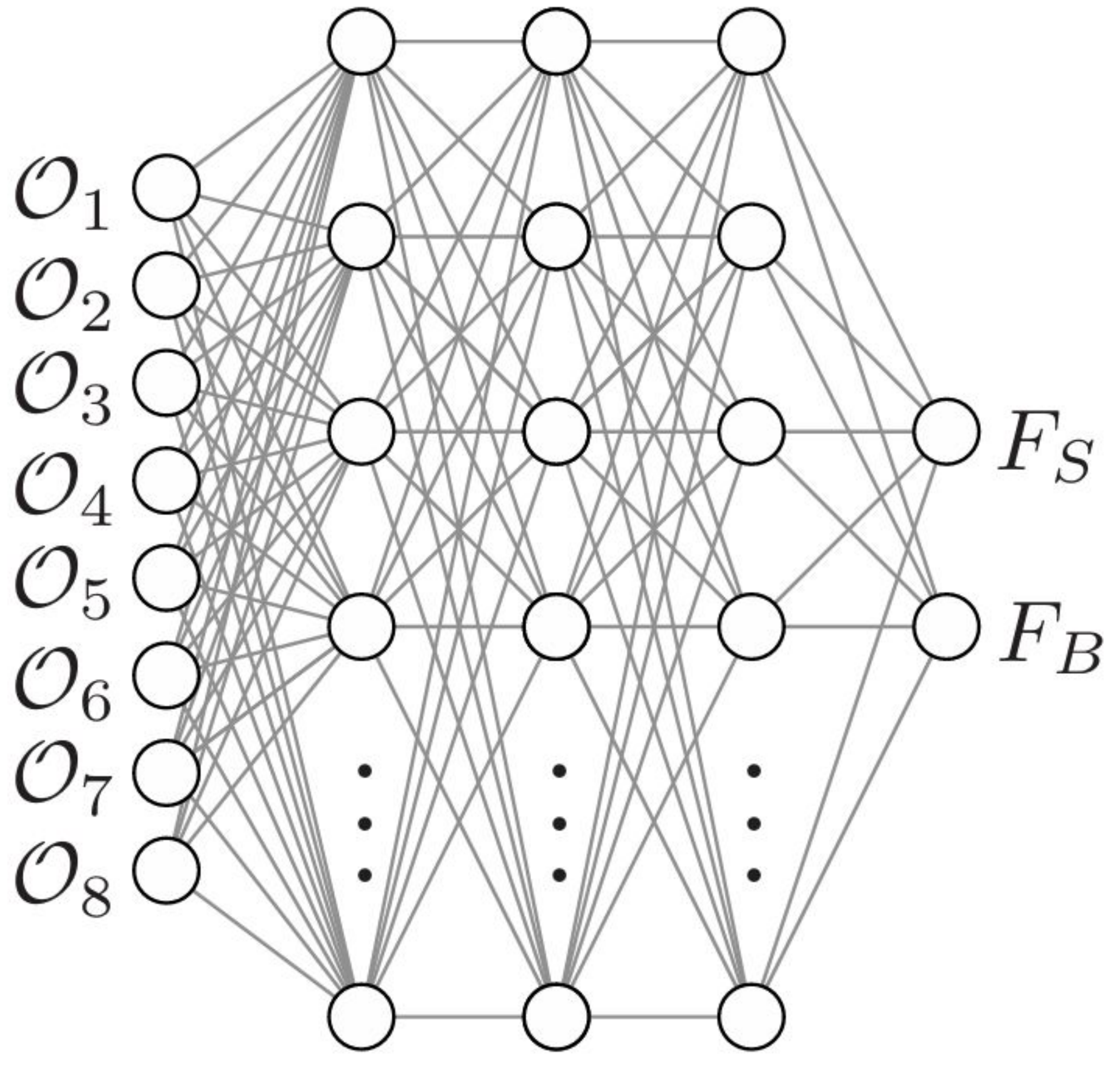
- Goal: to construct an DL architecture capable of learning new observables, relevant to the task p-p vs Pb-Pb jet discrimination

- Idea: Learn new observables, called **latent observables**, through a network **Φ**, by combining the **Φ** outputs for each particle in the jet and feeding them to a classifier network F

- Observables are combined in an **IRC-safe** manner, making the network's predictions physically more reliable and interpretable

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





*Komiske, Patrick T., Eric M. Metodiev, and Jesse Thaler. 2019. "Energy Flow Networks: Deep Sets for Particle Jets." Journal of High Energy Physics 2019

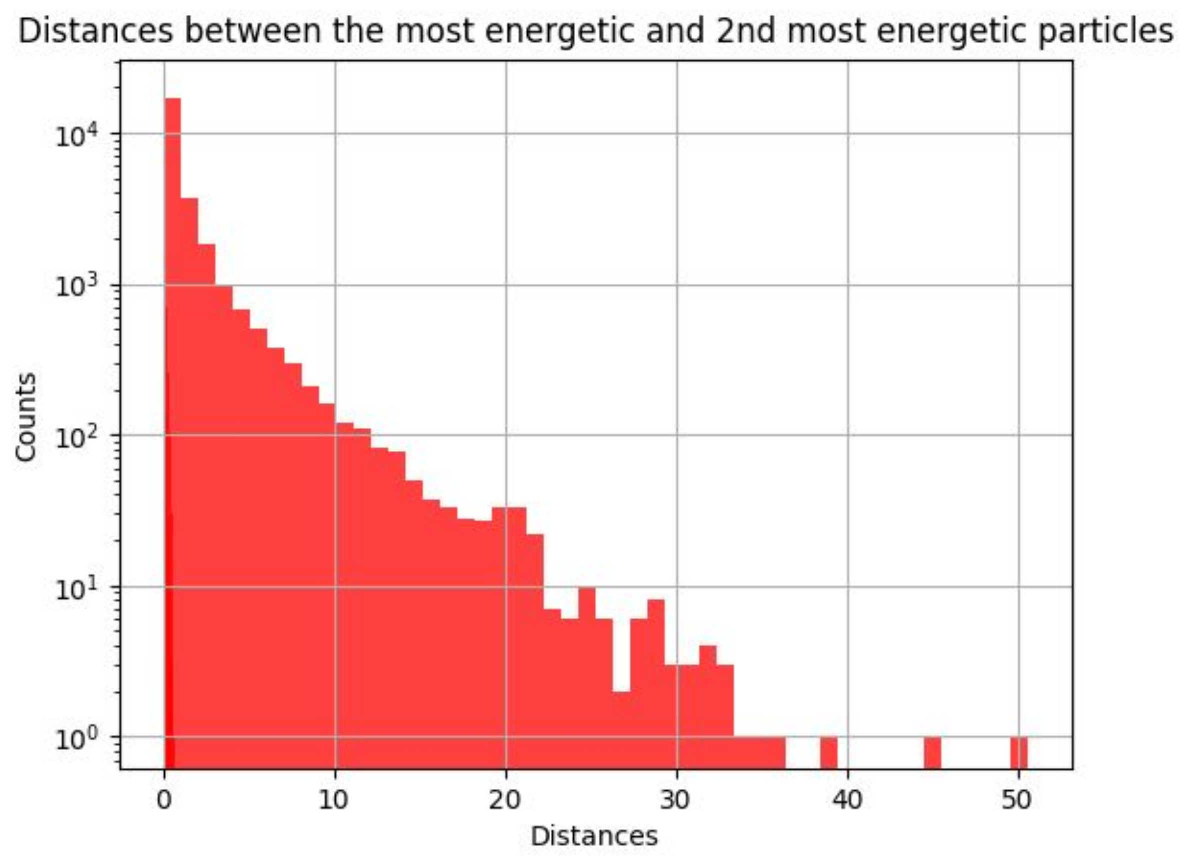
Work Extension

- Directly add particle distances to the EFN observables. Distances will be concatenated with the latent space observables

- For calculating the distances, the most energetic particles are selected, as they best represent the core characteristics of the jet

-Each feature (distance) is Standard Scaled and the kinematic variables (EFN inputs) are normalized w.r.t. energy

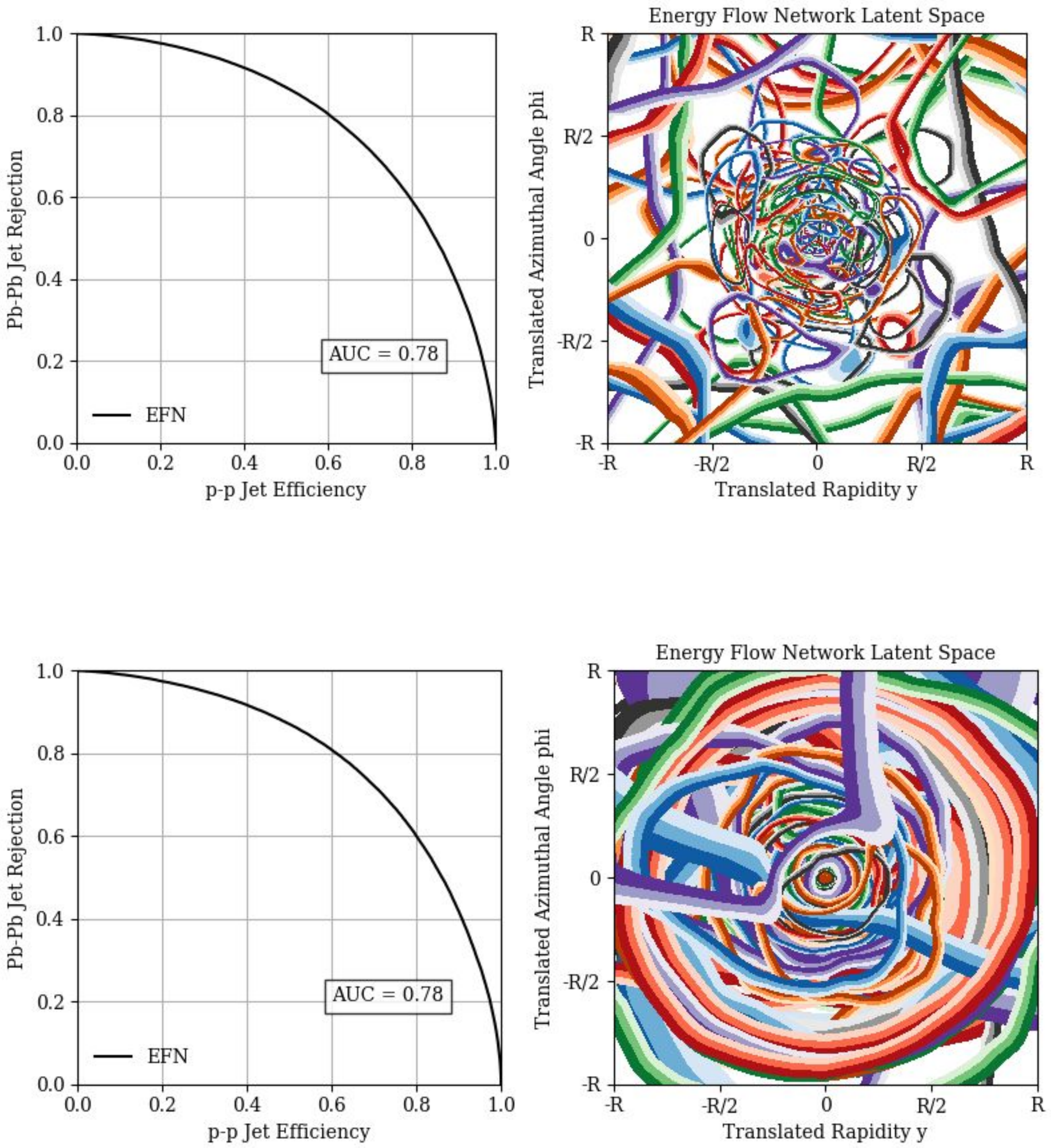
$$p_{T,i} \to \frac{p_{T,i}}{\sum_j p_{T,j}}, \quad y_i \to y_i - \left(\sum_j p_{T,j}\,\hat{p}_j\right)_y, \quad \phi_i \to \phi_i - \left(\sum_j p_{T,j}\,\hat{p}_j\right)_\phi$$



Kt distance histogram (standard scaled)

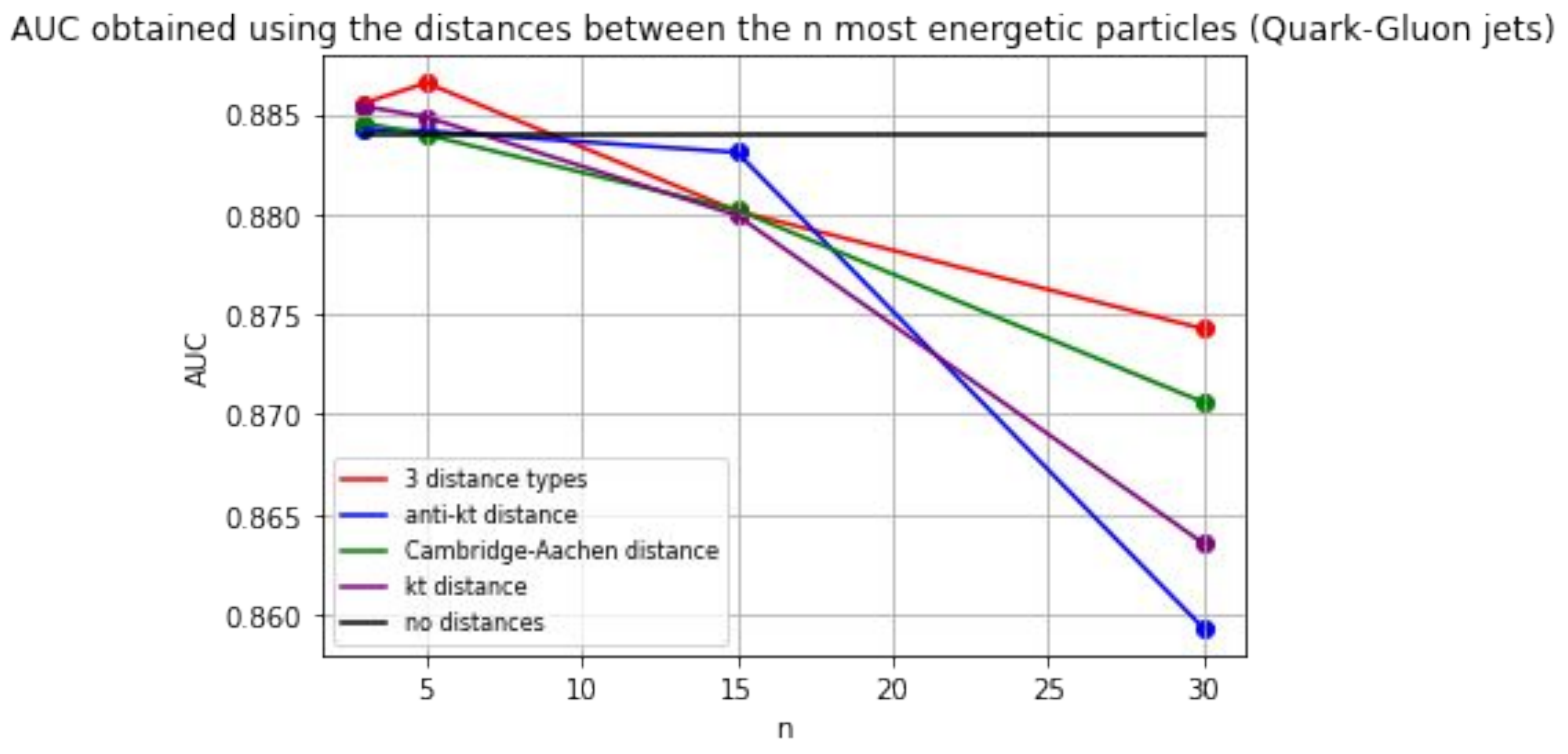
Results and Conclusions

- As we increase the number of used distances, the EFN latent space filters exhibit concentric circular patterns

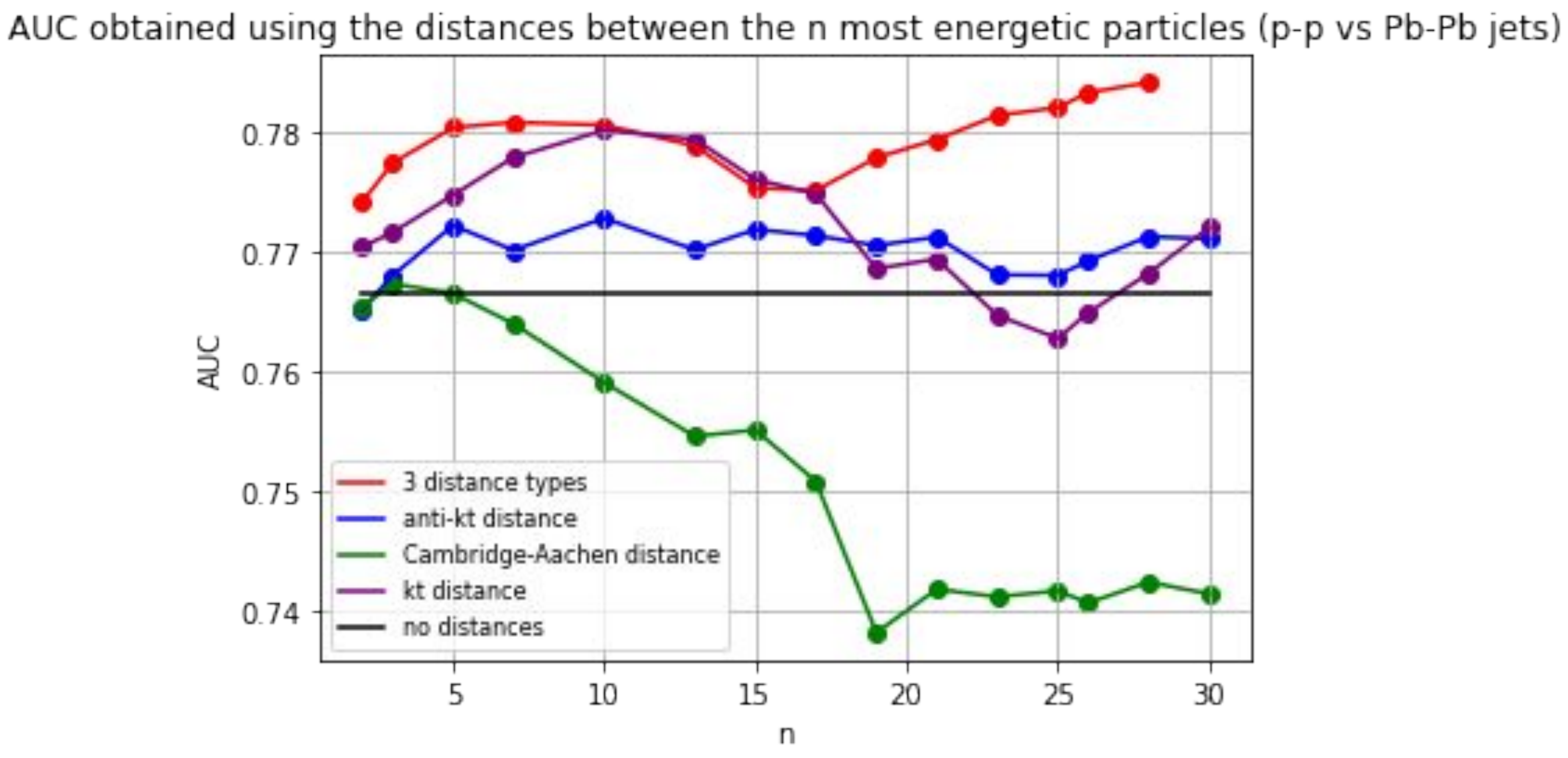


Results and Conclusions

-In most cases, this new model outperforms the base model (no distances)



-Overall, using the 3 types of distances we achieve the best classification accuracy



Hyperparameters

- random seed: 42
- batch size: 1024
- Optimizer: adam (learning rate: 0.0006)
- Phi_sizes = (100,100,128)
- F_sizes = (128,100,100)
- Train_size, val_size, test_size = 1600000, 200000, 200000
- 500 epochs with patience of 30 epochs on the val_loss
- F_dropout: 0.2

Future Work

- Use AutoML (e.g Optuna) to find optimal hyperparameters for the EFNs

- Add relevant observables to the latent space

End of Talk