APPLICATION OF VARIATIONAL AUTOENCODERS TO GENERATE D^0 MESONS KINEMATICS AT THE LHC

Allan Eduardo F. G. Ferreira, César A. Bernardes, Davi L. de Souza
Breno Orzari, Thiago R. F. P. Tomei

INTRODUCTION

With the advancement of generative neural networks, their application in HEP has become increasingly natural. One potential use is data generation, serving as an alternative to Monte Carlo (MC) generators, which often demand significant computational power and may struggle to produce representative data. [1, 2]

We developed a Variational Autoencoder (VAE)-based network to generate D^0 meson kinematics by reconstructing the four-momenta of its daughter particles.

DATASET

- Minimum bias PbPb collisions simulated with Hydjet. [3]
- D^0 mesons decaying into charged pions and kaons are embedded in events using the Pythia8. [4]
- The input file includes the 4-momentum of the D^0 mesons decay products.

VARIATIONAL AUTOENCODERS

The main goal of VAEs is to reconstruct data that closely resembles the input. It consists of three components: [5, 6]

- **Encoder:** reduces the input data's dimensionality and stores it in latent space.
- **Latent space:** represents the data in a given probability distribution;
- **Decoder:** reconstructs the input data from the latent space, returning it to its original dimensionality.

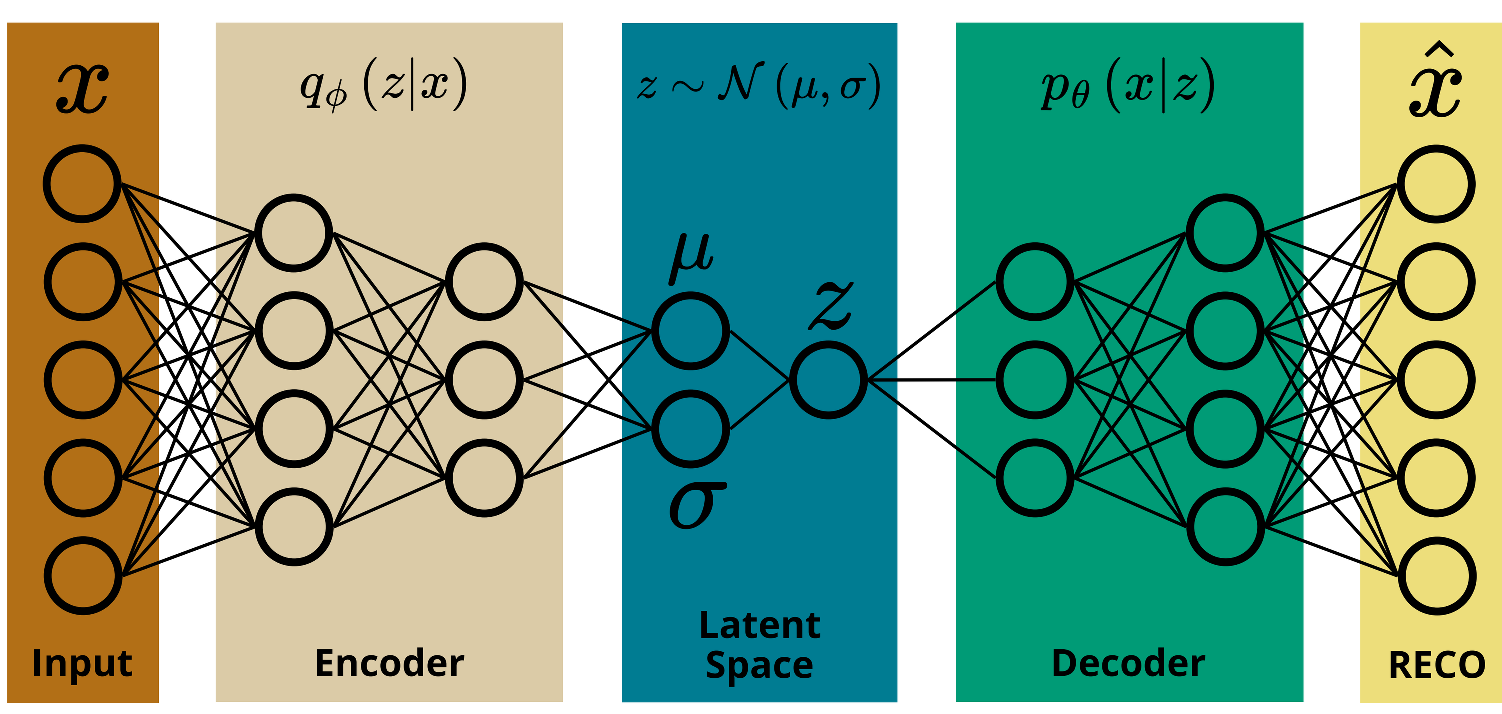

Fig. 1: VAE architecture representation.

To optimize the network, we use a loss function to evaluate the efficiency of data reconstruction, which can be expressed as:

$$\mathcal{L} = \beta L + (1-\beta) D_{KL} + \mathcal{C}$$

Where L is the error term, composed by the mean squared error between the value from the reconstructed variable and its input value, and between the reconstructed D^0 mass and the real one:

$$L = \frac{1}{N}\sum_i^N (x_i - \hat{x}_i)^2 + \frac{1}{N}\sum_i^N (m_i - \hat{m}_i)^2$$

The D_{KL} term is Kullback-Leibler Divergence, which measures how different two distributions are over the same variable:

$$D_{KL} = E_{x\sim q_\phi}\left[\log q_\phi(z|x) - \log p_\theta(z|x)\right]$$
$$= -\frac{1}{2}\left(1 + \log\sigma^2 - \mu^2 - \sigma^2\right)$$

And finally, $\mathcal{C}$ is a condition penalty term, which increases when the conditions gave to the network do not match:

$$\mathcal{C} = \sum_j^{N_\mathcal{C}} w_j \frac{1 - \mathbf{C}_j}{N}$$

Where $\mathbf{C}_j$ is the condition tensor and w_j is the weight term for each condition.

This penalty term defines a Conditional Variational Autoencoder (CVAE) neural network.

RESULTS

Implementing the CVAE network with PyTorch and using Optuna to optimize the network hyperparameters, we are being able to get some daughters 4-momentum distributions:

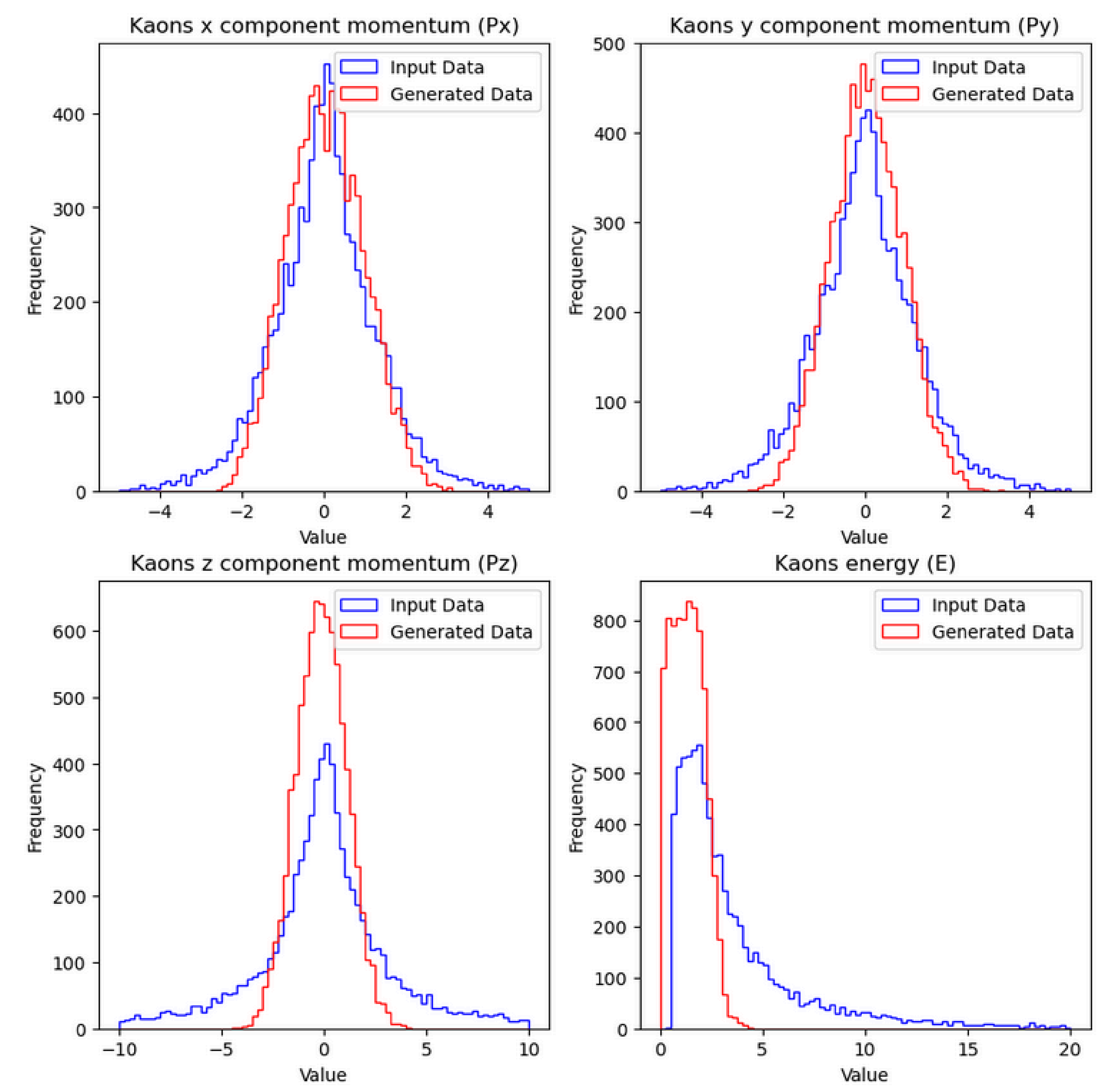

Fig. 2: Reconstructed vs input data kaons' 4-momentum.

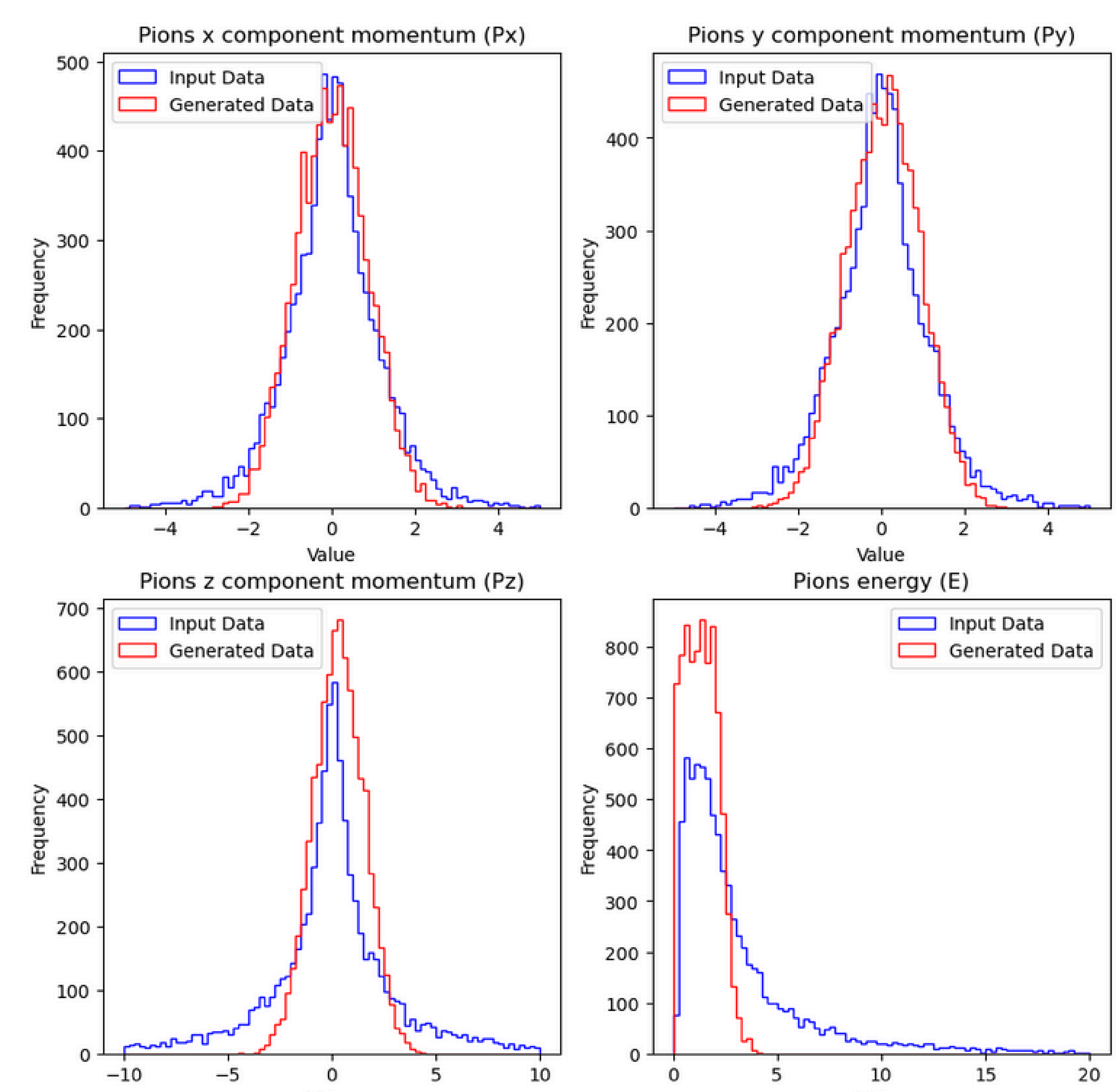

Fig. 3: Reconstructed vs input data pions' 4-momentum.

CONCLUSION

We applied Conditional Variational Autoencoders to generate the kinematics of D^0 mesons. So far, we have achieved better precision in modeling the transverse momentum (relative to the beam direction) of the D^0 meson decay products. Our next steps involve further optimizing the network parameters before proceeding with studies using data from the CMS experiment.

REFERENCES

[1] **Breno Orzari et al.** *LHC hadronic jet generation using convolutional variational autoencoders with normalizing flows.* Machine Learning: Science and Technology, 4(4):045023, October 2023.

[2] **Raghav Kansal et al.** *Evaluating generative models in high energy physics. Phys. Rev. D, 107:076017, Apr 2023.*

[3] **I. P. Lokhtin and A. M. Snigirev.** *A model of jet quenching in ultrarelativistic heavy ion collisions and high-pt hadron spectra at RHIC.* The European Physical Journal C, 45(1):211–217, January 2006.

[4] **Torbjörn Sjöstrand et al.** *An introduction to Pythia 8.2.* Computer Physics Communications, 191:159-177, June 2015.

[5] **Carl Doersch.** *Tutorial on variational autoencoders.* arXiv:2310.13138v2, 2021.

[6] **Diederik P. Kingma and Max Welling.** *An introduction to variational autoencoders.* Foundations and Trends in Machine Learning, 12(4):307–392, 2019