PROCEEDINGS OF SCIENCE

End-to-end tau reconstruction and identification using transformers

Nalong-Norman Seeba,[a,*] **Laurits Tani,**[a] **Hardi Vanaveski,**[a] **Joosep Pata**[a] **and Torben Lange**[a]

[a] *National Institute Of Chemical Physics And Biophysics (NICPB), Rävala pst. 10, 10143, Tallinn, Estonia*

E-mail: nalong-norman.seeba@cern.ch, laurits.tani@cern.ch, hardiveski@gmail.com, joosep.pata@cern.ch, torben.lange@cern.ch

Tau leptons play a key role in studying the production of Higgs and electroweak bosons, both within and beyond the Standard Model of particle physics. The precise reconstruction and identification of hadronically decaying tau leptons is essential for present and upcoming high-energy physics experiments. Driven by the advancements in jet tagging, we demonstrate that tau lepton reconstruction can be broken down into tau identification, kinematic reconstruction, and decay mode classification using a multi-task machine learning approach. Based on a new publicly available electron-positron collision dataset with full detector simulation and reconstruction, we demonstrate that standard jet tagging architectures can effectively perform these tasks. Our models achieve momentum resolutions of 2–3%, while the accuracy for reconstructing individual decay modes ranges from 80–95%.

42nd International Conference on High Energy Physics (ICHEP2024)
18-24 July 2024
Prague, Czech Republic

*Speaker

PoS(ICHEP2024)1036

1. Introduction

Tau leptons (τ) have a very short lifetime of 2.9×10^{-13} seconds [1], causing them to decay before interacting with detector material or undergoing radiative processes. One-third of τ decays are leptonic, with the final state particles being electrons or muons, and a neutrino. Since neutrinos do not interact with the detector, the final state resembles regular electrons and muons, for which dedicated reconstruction algorithms exist. In hadronic tau (τ_h) decays, the τ decays into a neutrino, charged hadrons (typically $\pi^{\pm}$ or $K^{\pm}$), and neutral pions (π^0). Since this final state consists of multiple particles, τ_h decays can be challenging to distinguish from jets produced by other high energy processes. Thus, identifying τ_h falls under the category of "jet-tagging", which has seen significant advancements with machine learning (ML), particularly through the use of deep learning techniques like transformers [2]. The effectiveness of these techniques in identifiying τ_h is demonstrated in Ref [3], while Ref [4] showcases their application for τ_h momentum regression and DM classification, using our publicly available `Fuτure` dataset.

2. Dataset & training

The `Fuτure` dataset features Monte Carlo samples of e^+e^- collisions at $\sqrt{s}$ = 380 GeV, including $Z/\gamma^* \to \tau\tau$, $ZH \to Z\tau\tau$ and $Z/\gamma^* \to qq$ processes, with around 2 million events per process. The Compact Linear Collider detector setup [5] was used for simulation and reconstruction. The dataset with additional technical details is available in Ref [6].

Reconstructing τ_h with ML can be framed as a multi-task learning problem, where Φ is a trainable model:

$$\Phi(\text{jet features}, \text{particle features}) \to \{\texttt{isTau}, \mathrm{DM}^{\mathrm{true}}, p_{\mathrm{T}}^{\mathrm{vis,true}}\}$$

Due to their effectiveness in jet related tasks, the ParticleTransformer [7] and LorentzNet [8] architectures are chosen as the models. In order to evaluate the improvement over a basic baseline, a simpler deep sets algorithm [9] (DeepSet) is used as a cross-check for the momentum regression and DM classification tasks, while for the identification a combination of the "hadrons-plus-strips" (HPS) [10] + DeepTau algorithm [11] is used. To learn more about the training, read sections 3 of Ref [3] and Ref [4].

3. Results

All three algorithms achieve τ_h misidentification rates below 0.1% for τ_h identification efficiencies (ϵ_τ) in the range 50-80%, with ParticleTransformer consistently outperforming the other models.

To assess the performance of the three algorithms in τ_h momentum regression, both the scale and resolution of the τ_h momentum distribution are considered. The scale is given by the median ratio of predicted to true visible τ_h momentum, where a median of one indicates accurate prediction, while deviations reflect over- or under-prediction. The resolution is determined by the width of the $p_{\mathrm{T}}^{\mathrm{vis,pred}}/p_{\mathrm{T}}^{\mathrm{vis,true}}$ distribution. Instead of standard deviation, the interquartile range (IQR) is used, as it is less sensitive to outliers. The IQR, defined as the difference between the 25% and 75%

quantiles, normalized by the 50% quantile, measures the width of the central part of the distribution relative to its median. All three algorithms achieve a resolution of 2.5-3% with momentum scale of about 0.5-1% within the true tau momentum.

The precision for DM classification across all three algorithms is broken down by decay mode, with ParticleTransformer outperforming the other models in cases with more final state particles. The confusion matrix for ParticleTransformer not only highlights precision based on each DM but also reveals patterns of misclassification. A precision of 80-95% is achieved overall.

All results are depicted in Fig 1.

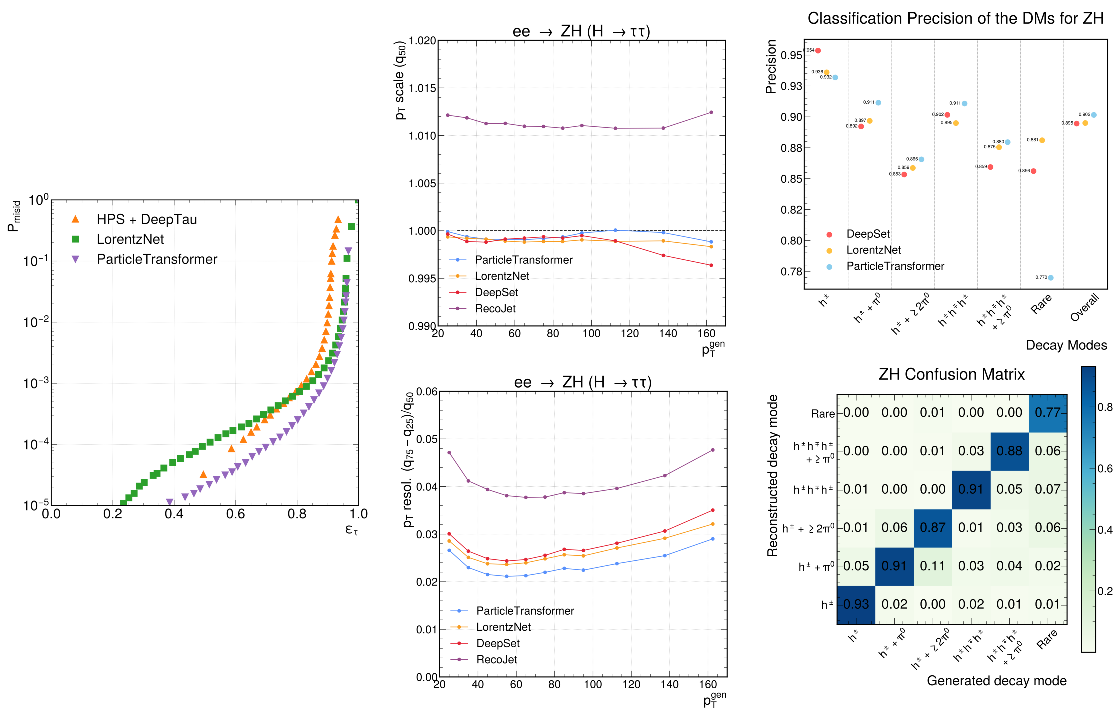

Figure 1: Misidentification rate (**left**) for quark and gluon jets as function of the τ_h identification efficiency. Scale (**upper-middle**) and resolution (**bottom-middle**) of the τ_h momentum response distribution. Precision (**upper-right**) of τ_h DM classification and the DM confusion matrix (**bottom-right**) for ParticleTransformer on the $ZH \rightarrow Z\tau\tau$ dataset.

References

[1] R. L. Workman *et al.*, "Review of Particle Physics," *PTEP*, vol. 2022, p. 083C01, 2022.

[2] S. Mondal and L. Mastrolorenzo, "Machine Learning in High Energy Physics: A review of heavy-flavor jet tagging at the LHC," *arXiv*, Apr. 2024. arXiv: 2404.01071 [hep-ex].

[3] T. Lange, S. Nandan, J. Pata, L. Tani, and C. Veelken, "Tau lepton identification and reconstruction: A new frontier for jet-tagging ML algorithms," *Comput. Phys. Commun.*, vol. 298, p. 109 095, 2024. arXiv: 2307.07747 [hep-ex].

[4] L. Tani, N.-N. Seeba, H. Vanaveski, J. Pata, and T. Lange, *A unified machine learning approach for reconstructing hadronically decaying tau leptons*, 2024. arXiv: 2407.06788 [hep-ex]. [Online]. Available: https://arxiv.org/abs/2407.06788.

[5] L. Linssen, A. Miyamoto, M. Stanitzki, and H. Weerts, "Physics and detectors at CLIC: CLIC conceptual design report," L. Linssen, A. Miyamoto, M. Stanitzki, and H. Weerts, Eds., 2012. arXiv: 1202.5940 [physics.ins-det].

[6] L. Tani, J. Pata, N. Seeba, H. Vanaveski, and T. Lange, *Fuτure - dataset for studies, development, and training of algorithms for reconstructing and identifying hadronically decaying tau leptons*. [Online]. Available: https://doi.org/10.5281/zenodo.12664634.

[7] H. Qu, C. Li, and S. Qian, "Particle transformer for jet tagging," in *International Conference on Machine Learning*, PMLR, 2022, pp. 18 281–18 292.

[8] S. Gong *et al.*, "An efficient lorentz equivariant graph neural network for jet tagging," *Journal of High Energy Physics*, vol. 2022, no. 7, pp. 1–22, 2022.

[9] M. Zaheer, S. Kottur, S. Ravanbakhsh, B. Poczos, R. R. Salakhutdinov, and A. J. Smola, "Deep sets," *Advances in neural information processing systems*, vol. 30, 2017.

[10] V. Khachatryan *et al.*, "Reconstruction and identification of τ lepton decays to hadrons and ν at CMS," *JINST*, vol. 11, no. 01, P01019, 2016. arXiv: 1510.07488 [physics.ins-det].

[11] A. Tumasyan *et al.*, "Identification of hadronic tau lepton decays using a deep neural network," *JINST*, vol. 17, P07023, 2022. arXiv: 2201.08458 [hep-ex].