IRIS-HEP Fellows Program Project Proposal

Tagging low momentum taus in CMS

Applicant: Kyrylo Filonenko
Mentor: Valeriia Lukashenko

1 Context of the project

The tau lepton (τ) is the least studied lepton in flavor physics, especially in B-meson decays [1]. It may provide a window into New Physics, as indicated by anomalies observed in measurements of $R(D)$ and $R(D^*)$.

The main challenge is the reconstruction of the tau. Typically, taus are reconstructed through their leptonic or hadronic decays. However, the efficiency of standard hadronic decay reconstruction algorithms drops significantly below p_T of 20 GeV. The taus of interest in this study have $p_T < 10$ GeV, requiring specialized reconstruction methods [1].

The goal of this project is to develop a dedicated tau reconstruction (tagging) algorithm [3] for the CMS Run 3 dataset. This will be based on and extend the low-momentum 3-prong tau tagger developed for CMS Run 2, utilizing the ABCNet model [2] — a graph neural network enhanced with attention mechanisms for improved performance.

As a benchmark, we focus on the decay $B^0 \to K^{*0}\tau^+\tau^-$, which has a relatively higher expected Standard Model branching ratio and thus, is the most promising candidates for observation.

The project will build upon the existing low-p_T tau tagging algorithm by adapting it for the conditions of Run 3, which include higher pileup and increased complexity of events [1].

2 Proposal

The goal is to become familiar with modern machine learning (ML) techniques, particularly those relevant to particle physics applications. The initial focus will be on retraining the existing low-momentum tau tagger based on ABCNet [2] for the CMS Run 3 data-taking period.

Internship steps:

- Retrain the ABCNet-based tau tagger [2] using Run 3 data, with the $B^0 \to K^{*0}\tau^+\tau^-$ decay as the primary benchmark process.
- Explore and adapt possible alternative models, such as the Particle Transformer [3], currently utilized for jet tagging, or any other advanced architectures that may improve the performance in the low-p_T regime.

- Acquire practical skills in modern ML frameworks, focusing on Python and libraries such as TensorFlow [4].
- Develop robust training, validation, and evaluation workflows for ML models.
- Eventually integrate the optimized tau tagger into the centralized CMS software framework for broader usage within the collaboration.

3 Timeline

- **Week 1**: Familiarization with ABCNet architecture and CMS low-p_T tau reconstruction.
- **Weeks 2-3**: Setting up training and validation workflow.
- **Weeks 4-5**: Training and validating ABCNet for Run 3.
- **Week 6-10**: Exploring alternative models (particle transformer) and comparing to the ABCNet baseline.
- **Week 11**: Defining benchmark, writing documentation, integrating with CMS software.
- **Week 12**: Writing report.

References

[1] CMS Collaboration, "Performance of the low-pT tau identification algorithm," CMS DP-2020/039. [Link]

[2] V. Mikuni and F. Canelli, "ABCNet: an attention-based method for particle tagging," Eur. Phys. J. Plus (2020) 135:463. [Link]

[3] H. Qu and L. Gouskos, "Particle Transformer for jet tagging," [GitHub] and [arXiv:2202.03772]

[4] TensorFlow: Open-source Machine Learning Framework, [Official Website]