



Contribution ID: 8

Type: **Poster**

The fundamental limit of jet tagging

Monday 8 September 2025 11:00 (30 minutes)

Jet tagging, i. e. determining the origin of high-energy hadronic jets, is a key challenge in particle physics. Jets are ubiquitous observables in collider experiments, made of complex collections of particles, that need to be classified. Over the past decade, machine learning-based classifiers have greatly enhanced our jet tagging capabilities, with increasingly sophisticated models driving further improvements. This raises a fundamental question: How far are we from the theoretical limit of jet tagging performance? To explore this, we employ transformer-based generative models to produce realistic synthetic data with a known probability density function. By testing various state-of-the-art taggers on this dataset, we find a significant gap between their performance and the theoretical optimum, signalling a significant room for improvement. Our dataset and software are made public to provide a benchmark task for future developments in jet tagging and other areas of particle physics.

Significance

We explore, for the first time, the jet-tagging capabilities of state-of-the-art taggers on synthetic data with a known optimum. We uncover a big gap between the theoretical optimum (the LLR) and the performance of the taggers. We set a new benchmark for future ML-based jet tagger developments.

References

arXiv:2411.02628

Experiment context, if any

Authors: Dr MÜCK, Alexander (RWTH Aachen University); NACHMAN, Ben (Lawrence Berkeley National Lab. (US)); Dr REYES-GONZÁLEZ, Humberto (RWTH Aachen); KRAMER, Michael; GITE, Nishank Nilesh (Lawrence Berkeley National Lab. (US)); MIKUNI, Vinicius Massami (Lawrence Berkeley National Lab. (US))

Presenter: Dr REYES-GONZÁLEZ, Humberto (RWTH Aachen)

Session Classification: Poster session with coffee break

Track Classification: Track 2: Data Analysis - Algorithms and Tools