Contribution ID: **13** Type: **not specified**

Exploring the Fundamental Limits of Jet Classification

Wednesday, July 23, 2025 2:00 PM (20 minutes)

Jets are ubiquitous observables in collider experiments, composed of complex collections of particles that require classification. Over the past decade, machine learning-based classifiers have significantly enhanced our jet tagging capabilities, with increasingly sophisticated models leading to further improvements. This raises a fundamental question: How close are we to 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 datasets, we estimate the gap between their performance and the theoretical optimum for several jet types. Overall, we find that there remains room for improvement, particularly in cases where the theoretical optimum is higher. Our approach paves the way for investigating what modern classifiers lack in reaching optimal performance. The dataset and software are made publicly available to provide a benchmark task for future developments in jet tagging and other areas of particle physics.

Authors: MÜCK, Alexander (RWTH Aachen); NACHMAN, Benjamin; NISHANK, Gite; REYES-GONZALEZ, Humberto (RWTH Aachen); GEUSKENS, Joep; KRÄMER, Michael (RWTH Aachen University); KOLLER, Sarah (RWTH Aachen); MIKUNI, Vinicius

Presenter: REYES-GONZALEZ, Humberto (RWTH Aachen)

Session Classification: Young Scientists Talks: Session 7