



Contribution ID: 54

Type: **Flashtalk with Poster**

Omnijet- \boxtimes : The first cross-task foundation model for particle physics

Wednesday, 1 May 2024 16:52 (3 minutes)

Foundation models are multi-dataset and multi-task machine learning methods that once pre-trained can be fine-tuned for a large variety of downstream applications. The successful development of such general-purpose models for physics data would be a major breakthrough as they could improve the achievable physics performance while at the same time drastically reduce the required amount of training time and data.

We report significant progress on this challenge on several fronts. First, a comprehensive set of evaluation methods is introduced to judge the quality of an encoding from physics data into a representation suitable for the autoregressive generation of particle jets with transformer architectures (the common backbone of foundation models). These measures motivate the choice of a higher-fidelity tokenization compared to previous works.

Finally, we demonstrate transfer learning between an unsupervised problem (jet generation) and a classic supervised task (jet tagging) with our new Omnijet- \boxtimes model. This is the first successful transfer between two different and actively studied classes of tasks and constitutes a major step in the building of foundation models for particle physics.

Primary authors: Dr HALLIN, Anna (University of Hamburg); Prof. KASIECZKA, Gregor (University of Hamburg); BIRK, Joschka (University of Hamburg)

Presenter: BIRK, Joschka (University of Hamburg)

Session Classification: 4.3 Physics-informed AI, Foundation models and related techniques

Track Classification: Session B