



Contribution ID: 79

Type: **Oral**

## Generic representations of jets at detector-level with self-supervised learning

*Wednesday 13 March 2024 15:50 (20 minutes)*

Supervised learning has been used successfully for jet classification and to predict a range of jet properties, such as mass and energy. Each model learns to encode jet features, resulting in a representation that is tailored to its specific task. But could the common elements underlying such tasks be combined in a single foundation model to extract features generically? To address this question, we explore self-supervised learning (SSL), inspired by its applications in the domains of computer vision and natural language processing. Besides offering a simpler and more resource-effective route when learning multiple tasks, SSL can be trained on unlabeled data, e.g. large sets of collision data. We demonstrate that a jet representation obtained through SSL can be readily fine-tuned for downstream tasks of jet kinematics prediction and jet classification. Compared to existing studies in this direction, we use a realistic full-coverage calorimeter simulation, leading to results that more faithfully reflect the prospects at real collider experiments.

### Significance

Going beyond previous work, we present novel approaches to the training of our foundation model, with the aim of leveraging large unlabeled datasets (opening the door to novel data-driven analysis techniques) and learning transferable jet representations that are invariant to detector properties.

### References

- Presentation at the ML4Jets Workshop, Hamburg, 2023: <https://indico.cern.ch/event/1253794/contributions/5588641/>
- Paper presenting the collider detector simulation used in this work: “Configurable calorimeter simulation for AI applications”

### Experiment context, if any

**Primary authors:** KOBYLIANSKII, Dmitrii (Weizmann Institute of Science (IL)); GROSS, Eilam (Weizmann Institute of Science (IL)); DREYER, Etienne (Weizmann Institute of Science (IL)); MERZ, Garrett; CRANMER, Kyle Stuart (University of Wisconsin Madison (US)); SOYBELMAN, Nathalie (Weizmann Institute of Science (IL)); KAKATI, Nilotpal (Weizmann Institute of Science (IL)); RIECK, Patrick (New York University (US))

**Presenter:** RIECK, Patrick (New York University (US))

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

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