



EuCAIFCon 2024

EUROPEAN AI FOR
FUNDAMENTAL PHYSICS
CONFERENCE

Amsterdam, 30 April - 3 May

Anomaly detection search for BSM physics in ATLAS experiment at LHC

Francesco Ciotto^{1,2}, Francesco Conventi^{2,3}, Elvira Rossi^{1,3}

¹Università degli studi di Napoli Federico II

²Università degli studi di Napoli Parthenope

³INFN Napoli

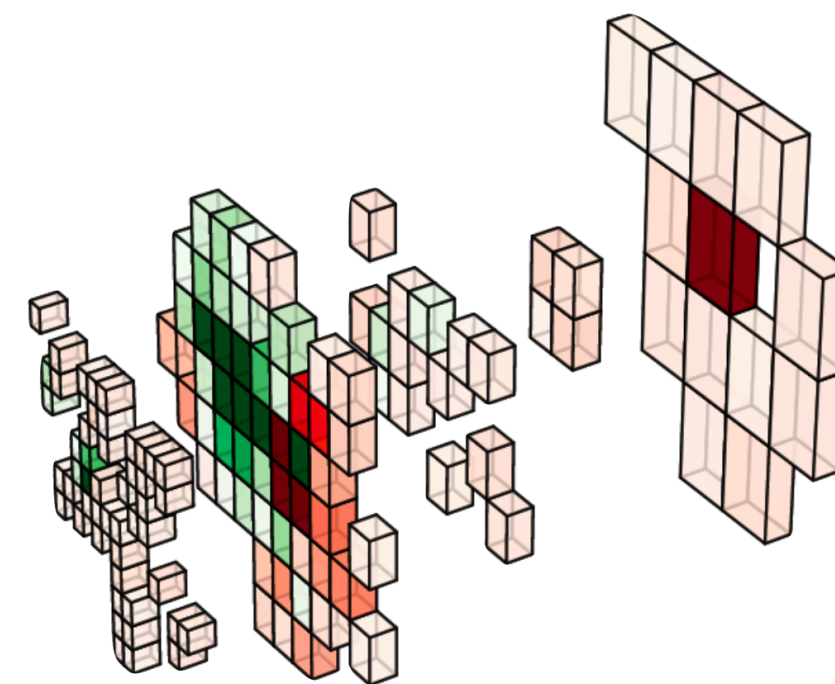
WHERE IS LHC GOING?

Finding anomalies in Jets

- No Physics Beyond Standard Model (BSM) has been observed at the LHC (yet!)
- The currently most used search paradigm is using model-dependent approaches
 - ↳ What if these models have blind spots for unconventional new physics signatures?
 - ↳ If there's new Physics in the current LHC data we can't miss it!
- Anomaly detection can find deviations in Standard Model events, without any signal dependency

Use jets as tools!

- Searches in full hadronic final states
 - ↳ Investigate its substructure by studying jet constituents
 - ↳ What we use: p_T, η, ϕ of each constituent

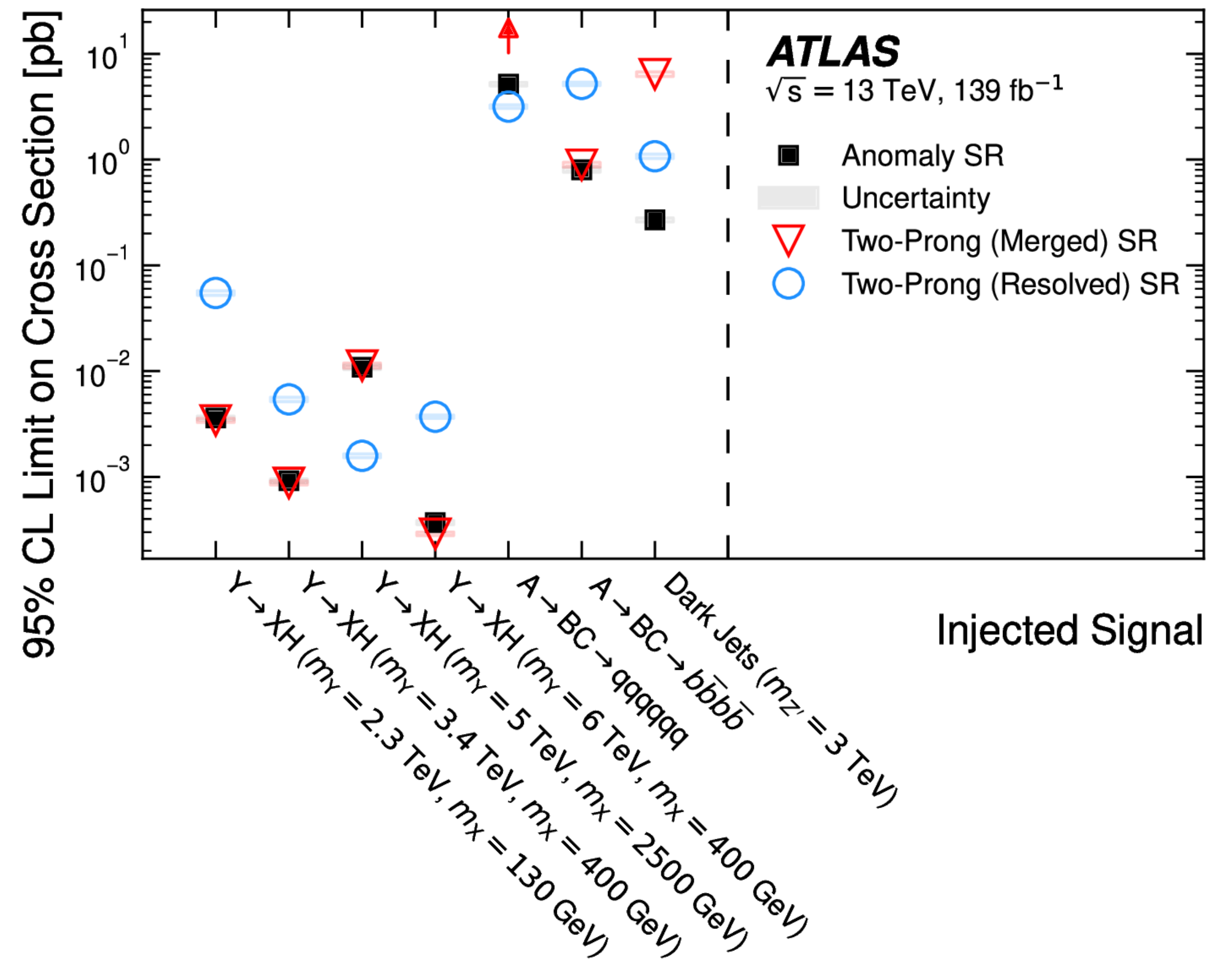


$Y \rightarrow XH$

The First Use of Unsupervised Learning on ATLAS Data

- The $Y \rightarrow XH$ analysis searches for heavy resonances decaying into a Higgs boson and new particle X in a fully hadronic final state
- Developed an unsupervised Variational Recurrent Neural Network to define an "Anomaly Signal Region"
 - ↳ Recurrent neural network that updates a VAE latent space at each time step, accommodating variable-length input sequences

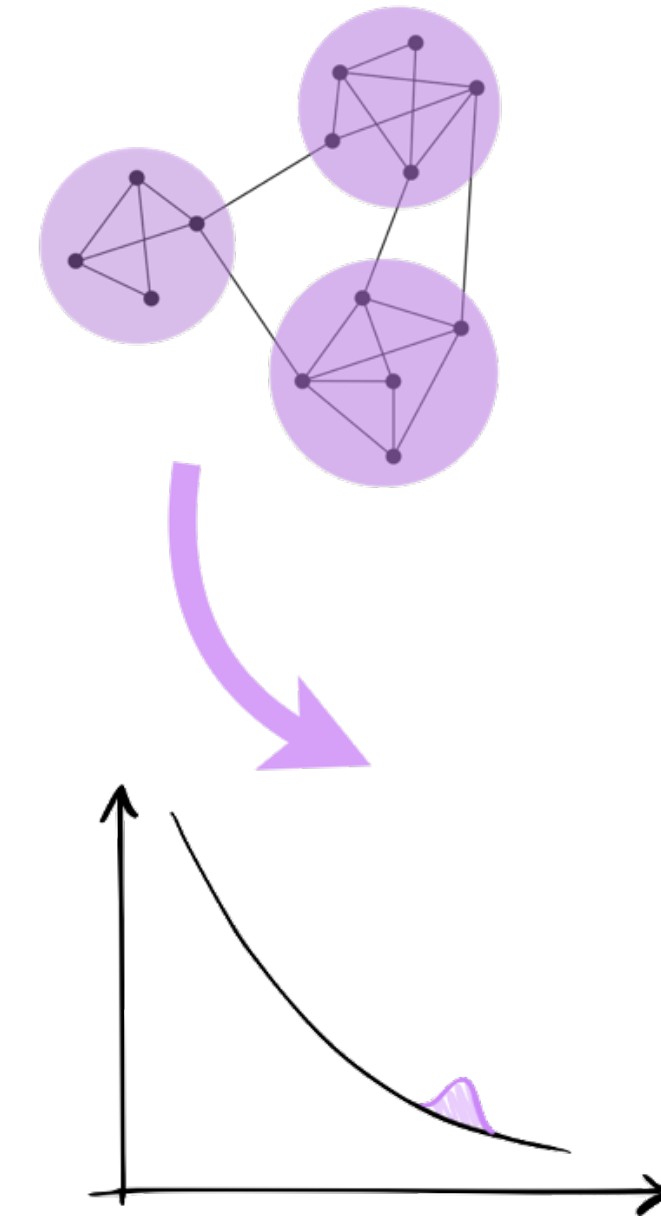
Can we improve the "classical" approach?



DIGGING DEEPLY WITH GRAPHS

A developing approach

- Graph-structured data are ubiquitous across science, engineering, and many other domains
 - ↳ Used to describe and analyze relations and interactions
 - ↳ Can encapsulate object or event information
- Our strategy: to represent jets as graphs and then apply machine learning to build an anomaly detection algorithm
 - ↳ Developing both **object and event level graphs** to detect anomalies
- Preliminary results on LHC Olympics dataset



Poster 76

Poster 76: Anomaly detection search for BSM physics in ATLAS experiment at LHC

Authors: Francesco Cirotto^{1,3}, Francesco Conventi^{1,3}, Evira Rossi^{1,3}

Finding anomalies in HEP
Anomaly Detection (AD) uses unsupervised Machine Learning architectures to identify outliers in a set of "standard" objects. In High Energy Physics, this means the identification of features of detector data inconsistent with the expected background.

Using jets for Anomaly Detection
Many Beyond Standard Model theories predict new massive resonances which can decay hadronically, leading to final states involving jets. For massive particles, their decay products become collimated, or "boosted", in the direction of the progenitor particle. For massive particles that are sufficiently boosted, it is advantageous to reconstruct their hadronic decay products as a single large-radius (R) jet.

Anomaly Detection in ATLAS in fully hadronic final states
The first application with AD technique in ATLAS is a search for a heavy resonance Υ decaying into a Standard Model Higgs boson H and a new particle X in a fully hadronic final state ($\Upsilon \rightarrow XH$) using the full Run-2 dataset collected by ATLAS from 2015 to 2018, corresponding to an integrated luminosity of 36.1 fb⁻¹.

Towards jets representation as Graphs
Graph Neural Networks (GNN) have proved to be innovative machine learning techniques useful to perform anomaly detection in the search for resonances in fully hadronic final states. In this paper, we propose a novel representation of jets as graphs, where nodes represent particles and edges represent interactions between them.

Preliminary results
Social architecture on R&D LHC Olympics 2020 dataset (50 generated SM multi-jet background and 100k signal) $\Upsilon \rightarrow XH$ signal search, where X and Υ are reconstructed as jets.

OmniJet- α : the first cross-task foundation model for particle physics

EuCAIFCon 2024, Amsterdam, May 1st 2024



Joschka Birk



Anna Hallin



Gregor Kasieczka



Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

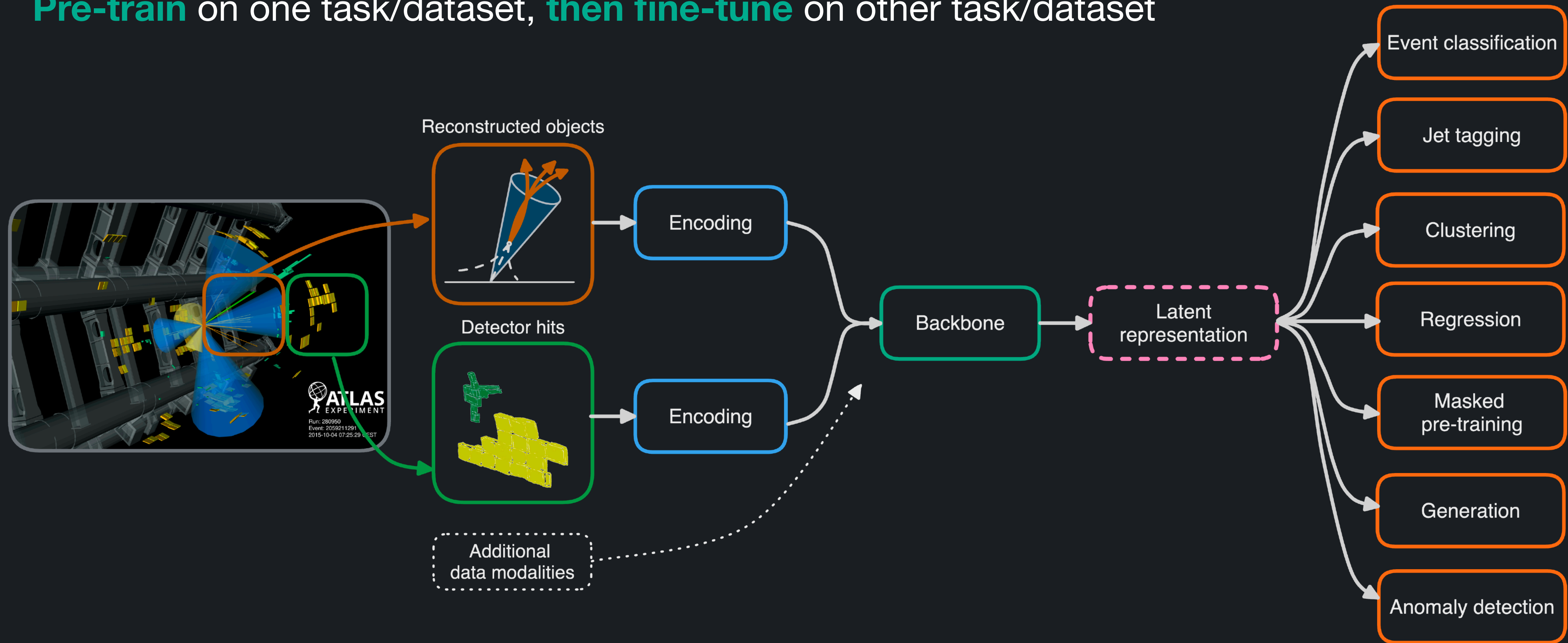
CLUSTER OF EXCELLENCE
QUANTUM UNIVERSE

joschka.birk@uni-hamburg.de

[arXiv:2403.05618](https://arxiv.org/abs/2403.05618)

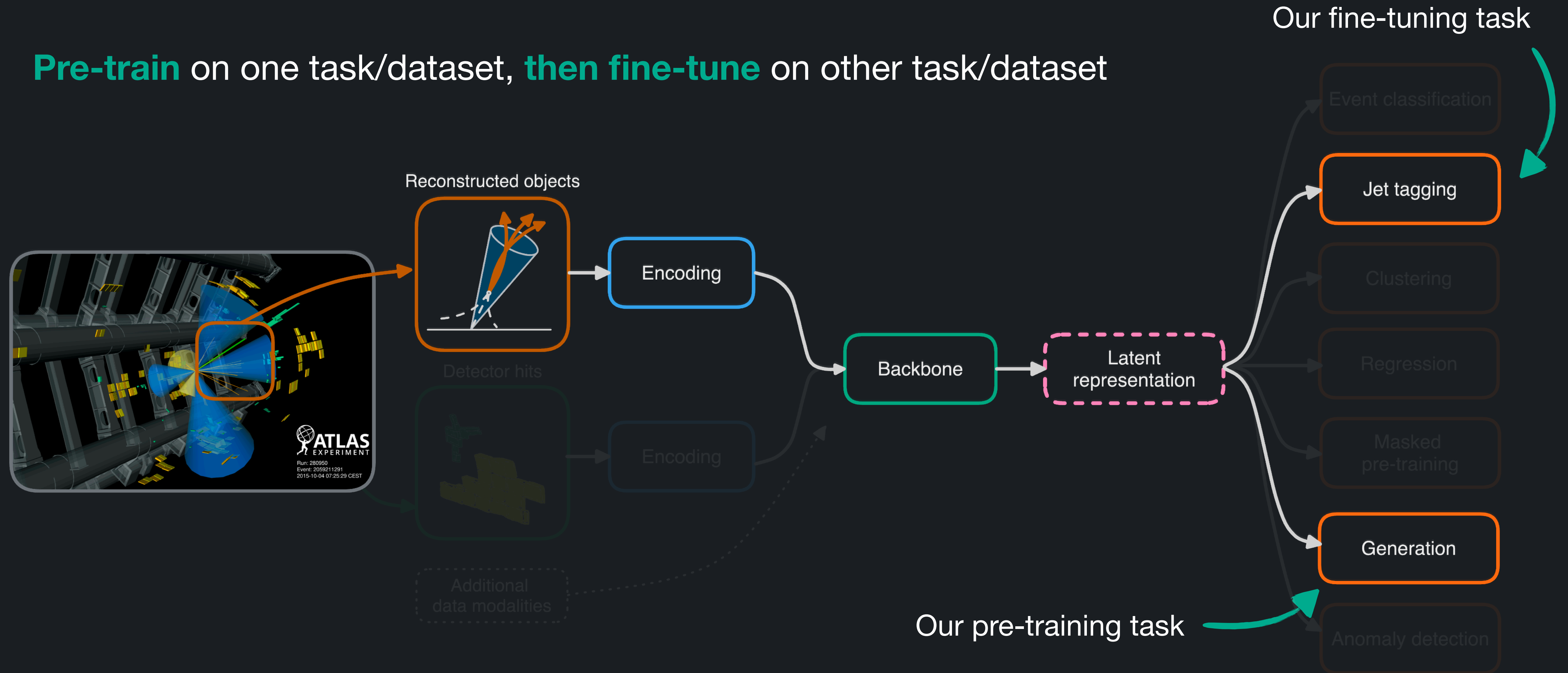
Foundation models for HEP

Pre-train on one task/dataset, **then fine-tune** on other task/dataset



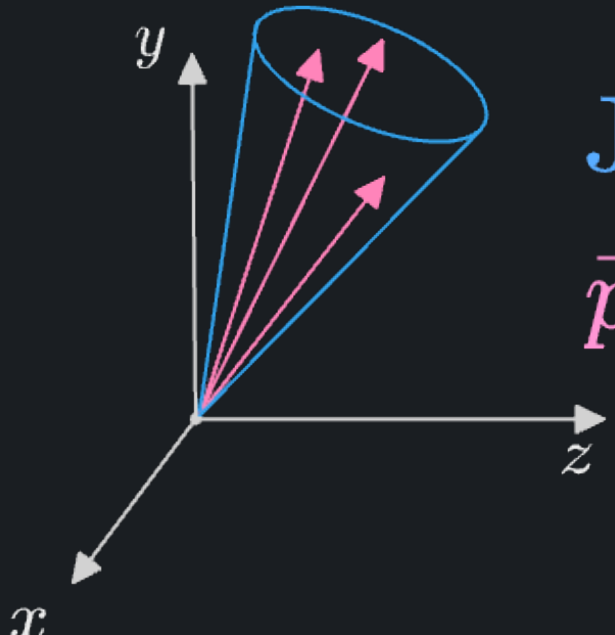
Foundation models for HEP

Pre-train on one task/dataset, **then fine-tune** on other task/dataset



Our approach

Jet constituents with **continuous features**



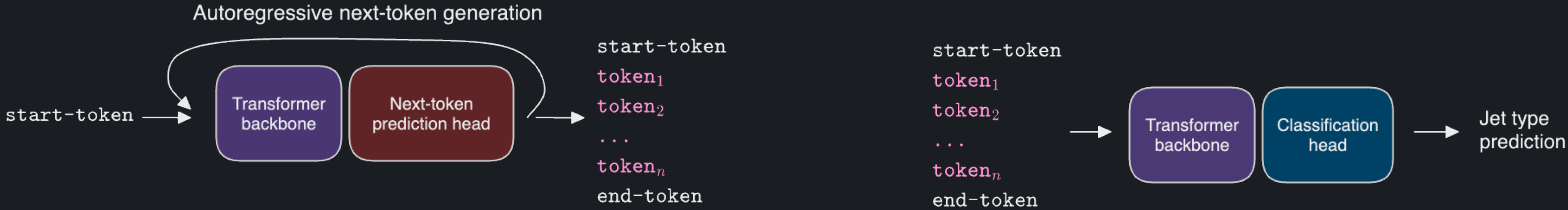
$$\text{Jet} = \{\vec{p}_1, \vec{p}_2, \dots, \vec{p}_n\}$$
$$\vec{p}_i = (p_T, \eta^{\text{rel}}, \phi^{\text{rel}})$$

Constituents are **tokenized with a VQ-VAE**
(using the approach presented by Sam Klein earlier)

$$\text{Jet} = \{\text{start-token}, \text{token}_1, \dots, \text{token}_n, \text{end-token}\}$$
$$\text{token}_i = \text{integer value} \in [1, \dots, 8192]$$

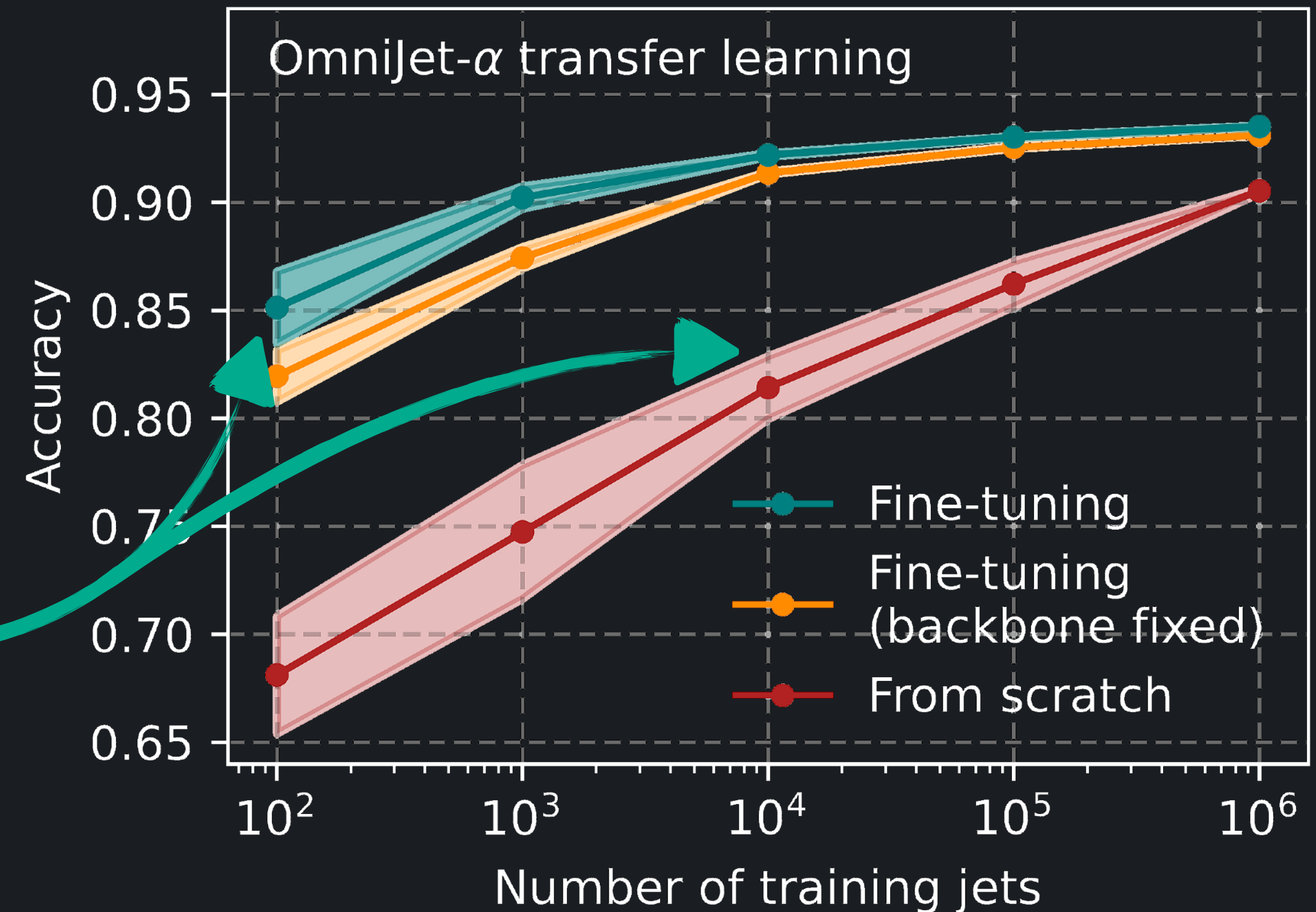
Unsupervised pre-training of transformer backbone
on generative task (next-token prediction)

Fine-tuning to classification task:
Swap model head and copy over the
weights from the pre-trained backbone



Does generative pre-training help for classification?

- Classification: $t \rightarrow bqq'$ vs. q/g jets
- Generative pre-training with both jet types
- **Pre-trained / fine-tuned model (●●)** reaches same performance as from scratch training with **100-1000x less training samples**



→ **Generative pre-training is a promising target for unsupervised pre-training in HEP**

Poster number: **54**

Applying hierarchical autoregressive neural networks for three-dimensional Ising model

Poster no. 74

Mateusz Winiarski,
Piotr Białas, Piotr Korcyl, Tomasz Stebel, Dawid Zapolski

Faculty of Physics, Astronomy and Applied Computer Science,
Jagiellonian University in Kraków, Poland

EuCAIFCon 2024
May 1st, 2024

1



3D Ising model

Autoregressive neural network:

$$q_{\theta}(s) = \prod_{i=1}^N q_{\theta}(s_i | s_1, s_2, \dots, s_{i-1})$$

Hierarchical structure:

$$p(s) = p(B(s))p(I(s)|B(s))$$

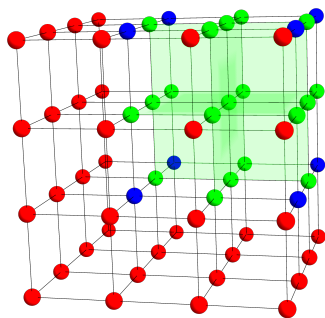


Figure: Scheme of hierarchical decomposition of $4 \times 4 \times 4$ cube

Neural Importance Sampling:

$$Z = \frac{1}{N} \sum_{i=1}^N \frac{e^{-\beta E(s_i)}}{q_{\theta}(s_i)}$$

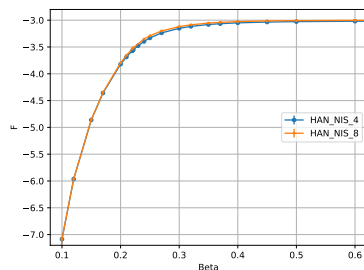


Figure: Free energy of β for HAN neural network