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 Cirotto[1,2], Francesco Conventi[2,3], Elvira Rossi[1,3]

[1]Università degli studi di Napoli Federico II
[2]Università degli studi di Napoli Parthenope
[3]INFN Napoli

WHERE IS LHC GOING?

Finding anomalies in Jets

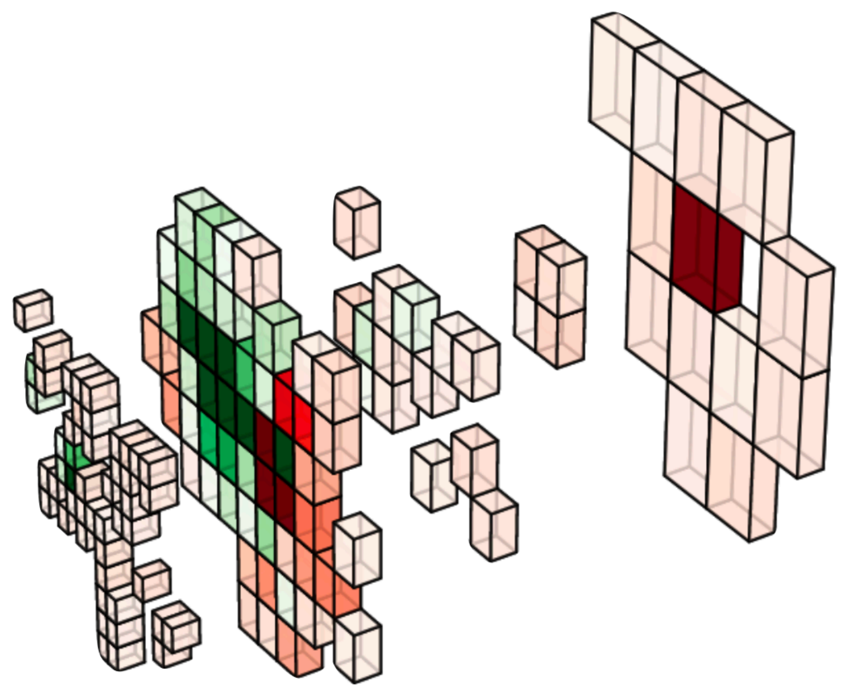
- No Physics Beyond Standard Model (BSM) has been observed at the LHC (yet!)

- 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



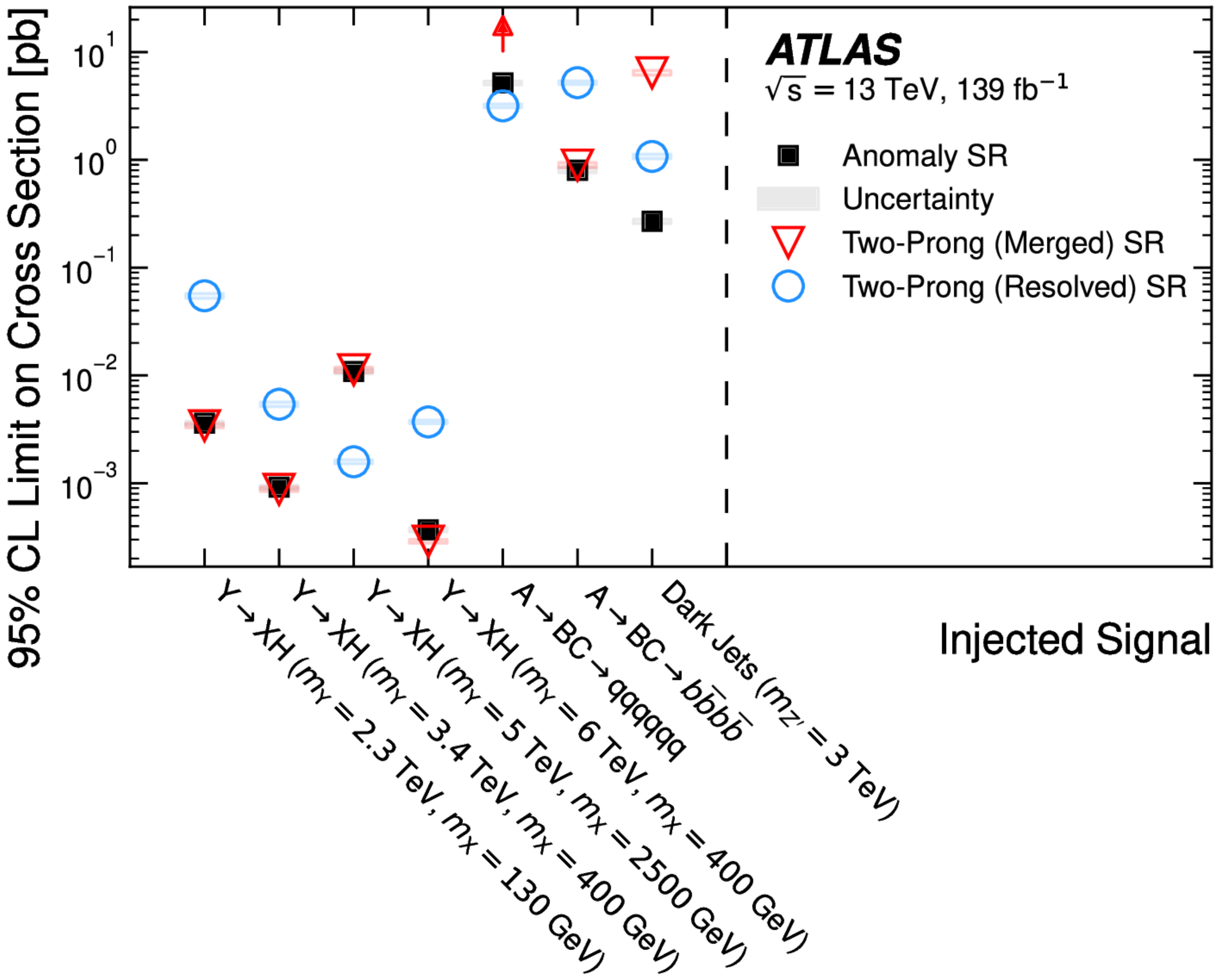
$Y \to XH$

The First Use of Unsupervised Learning on ATLAS Data

- The Y → 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

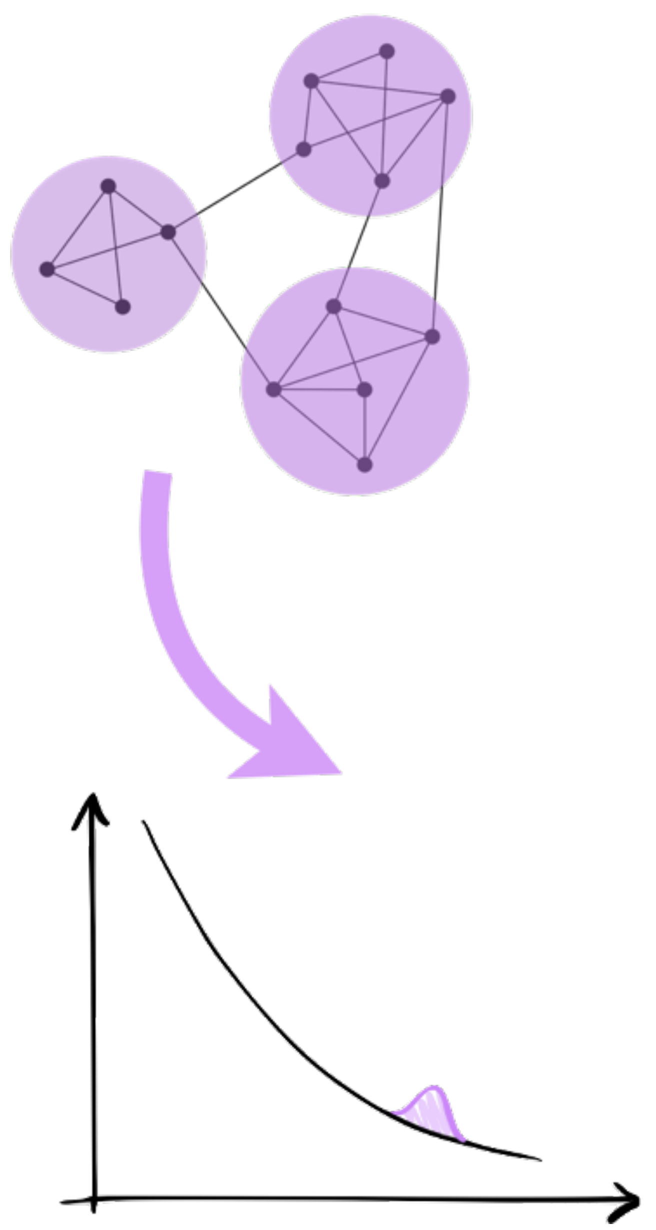
Can we improve the "classical" approach?



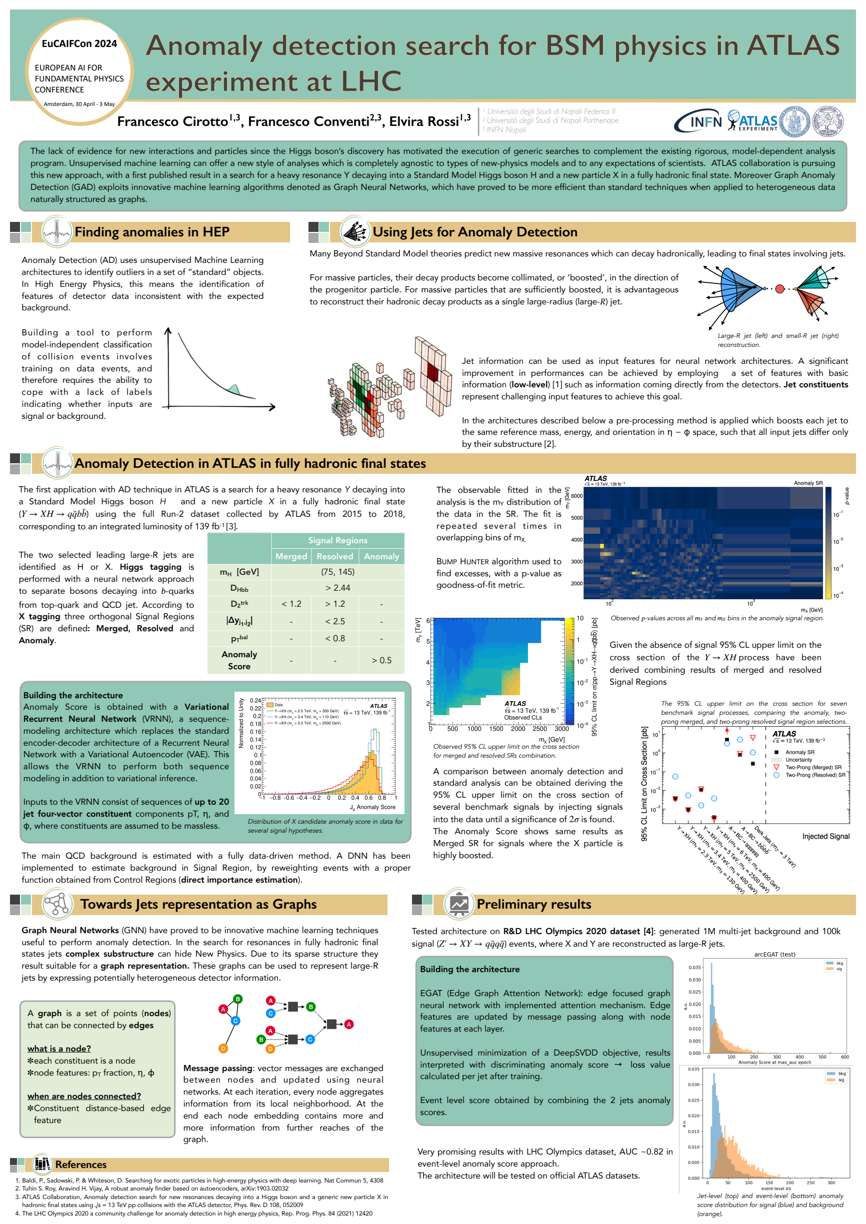
DIGGING DEEPLY WITH GRAPHS

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



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

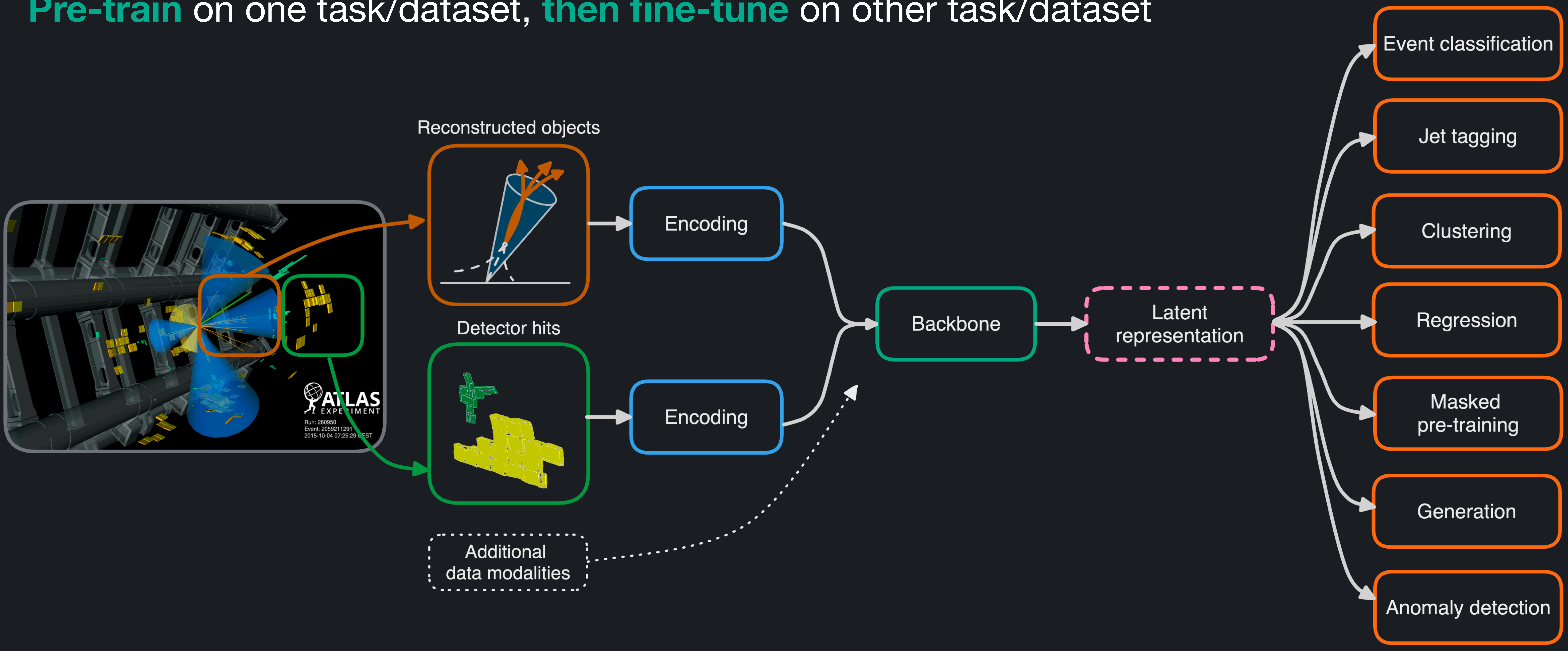
CLUSTER OF EXCELLENCE
QUANTUM UNIVERSE

joschka.birk@uni-hamburg.de

arXiv:2403.05618

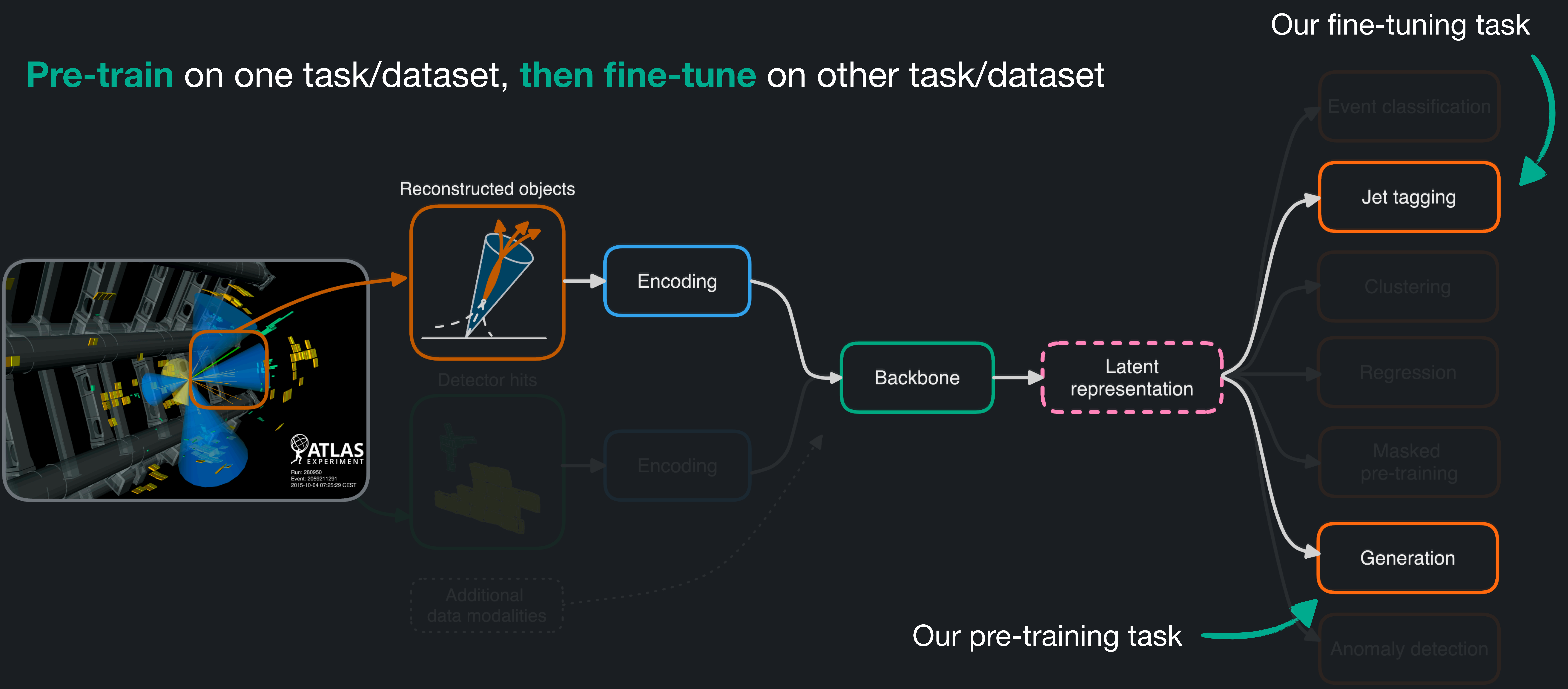
Foundation models for HEP

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



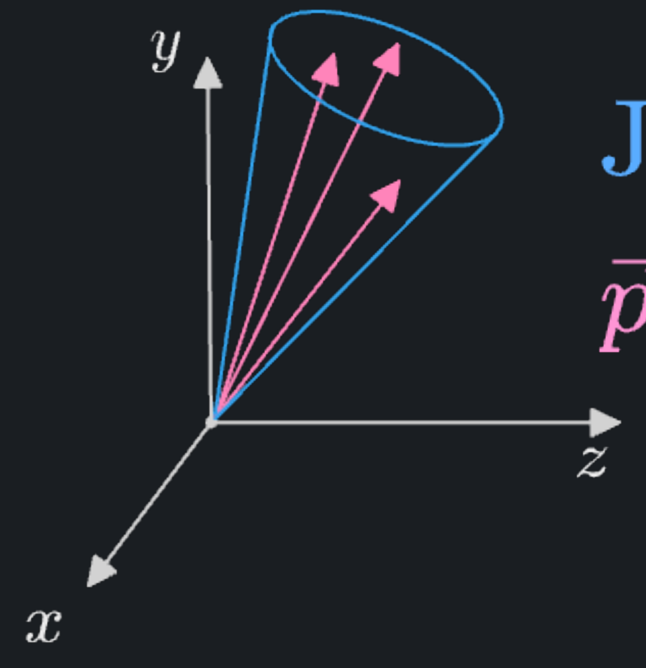
Foundation models for HEP

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



Our approach

Jet constituents with **continuous features**

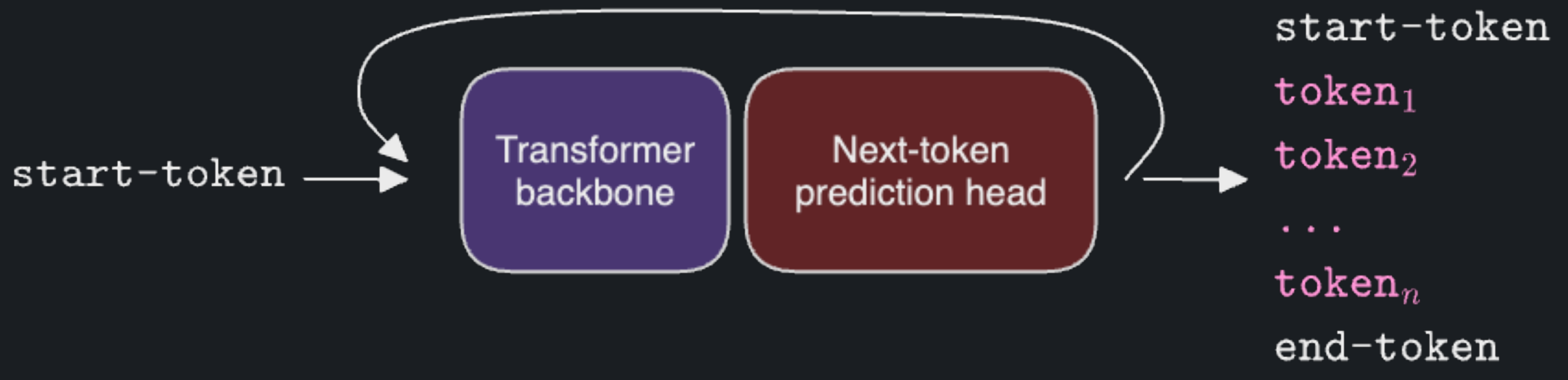


$$\text{Jet} = \{\vec{p}_1, \vec{p}_2, \ldots, \vec{p}_n\}$$
$$\vec{p}_i = (p_\mathrm{T}, \eta^{\mathrm{rel}}, \phi^{\mathrm{rel}})$$

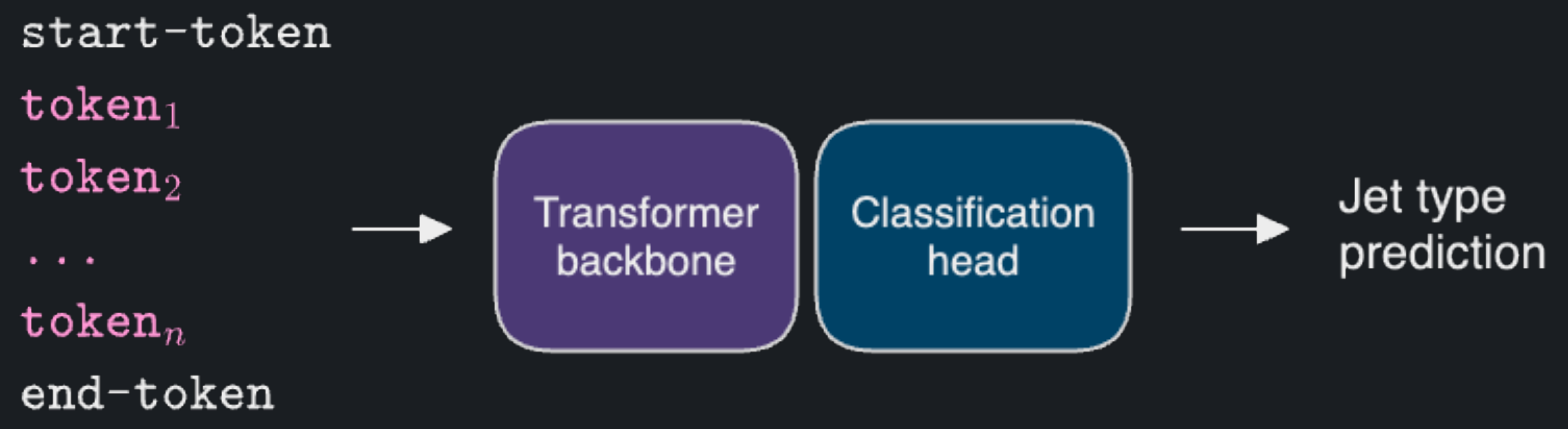
Constituents are **tokenized with a VQ-VAE**
(using the approach presented by Sam Klein earlier)

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

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

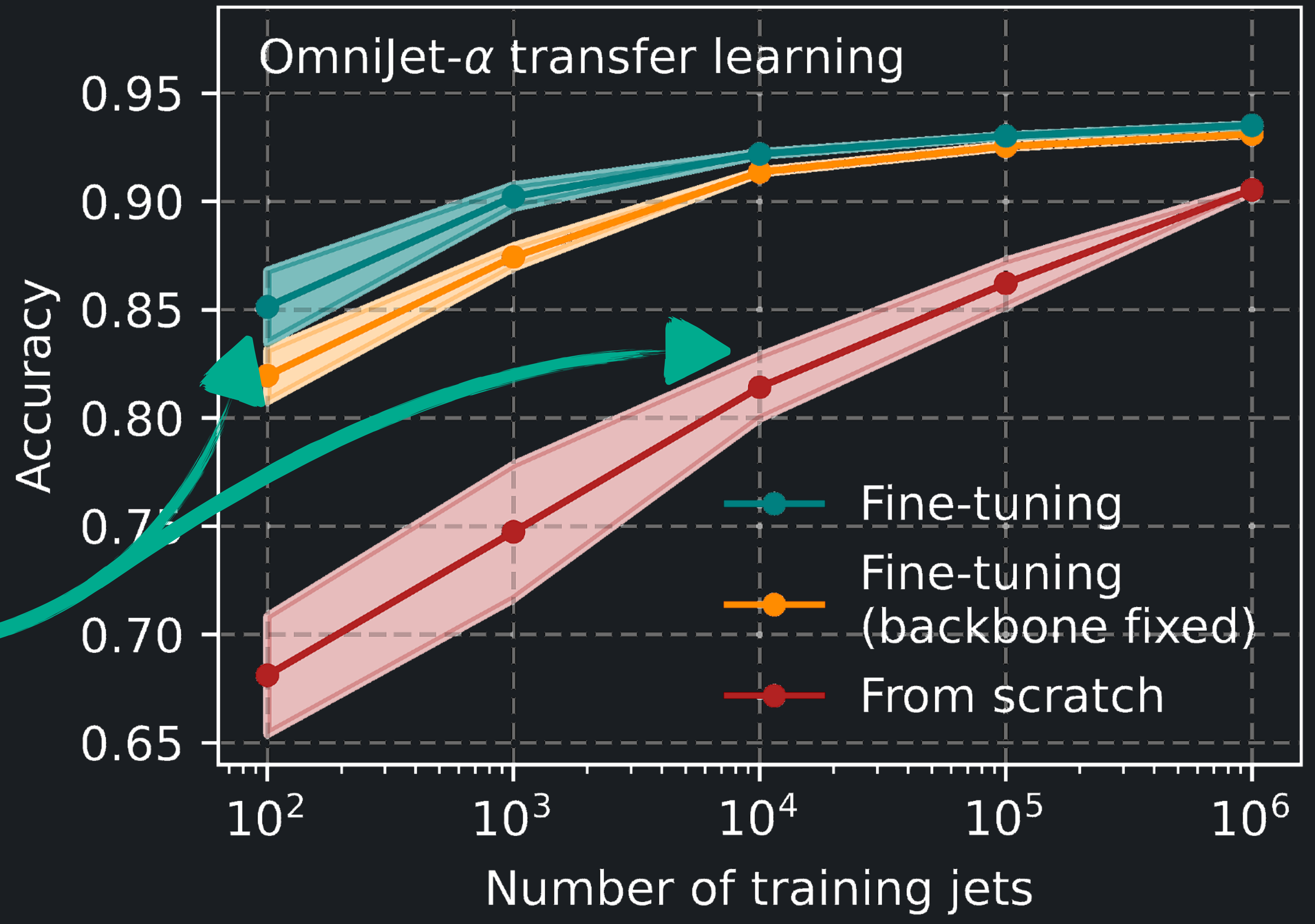


Fine-tuning to classification task:
Swap model head and copy over the weights from the pre-trained backbone



Does generative pre-training help for classification?

- Classification: $t \to bqq'$ 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

3D Ising model

Autoregressive neural network:

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

Hierarchical structure:

$$p(\mathrm{s}) = p(\mathrm{B}(\mathrm{s}))p(\mathrm{I}(\mathrm{s})|\mathrm{B}(\mathrm{s}))$$

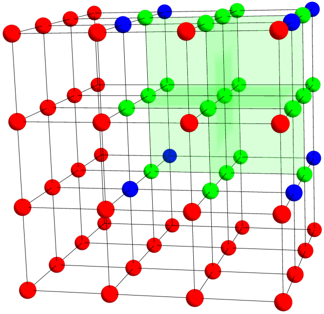


Figure: Scheme of hierarchical decomposition of $4 \times 4 \times 4$ cube

Neural Importance Sampling:

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

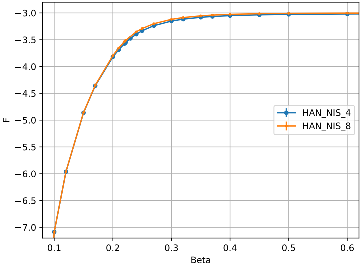


Figure: Free energy of β for HAN neural network

NATIONAL SCIENCE CENTRE POLAND

Project is supported by National Science Centre, grant no. 2021/43/D/ST2/03375.