

# Flavour Tagging with Graph Neural Networks with the ATLAS experiment

Walteri Leinonen on behalf of the ATLAS Collaboration



**Radboud Universiteit**



**EuCAIF conference 2024**

# Physics-informed GNN

Flavour Tagging with Graph Neural Networks with the ATLAS experiment

GN1 and GN2 are ATLAS's state-of-the-art low-level ***b*-tagging** and **boosted Higgs tagging** architectures.

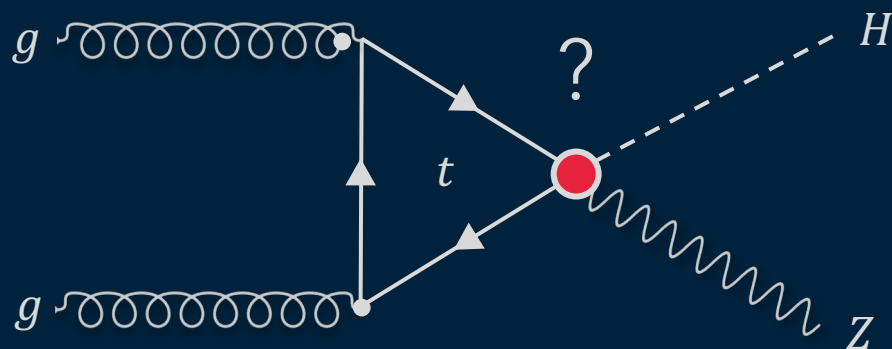
- Physics inspired tagging with **auxiliary tasks** :  $L_{total} = L_{jet} + \alpha L_{vertex} + \beta L_{track}$ 
  - Interpretable model outputs through **vertexing** and **track tagging**
- Massive performance upgrade
  - Many-fold higher background rejection → **Higher tagging efficiency**
- Flexible architecture, suitable for general tagging tasks
  - **Boosted Higgs** →  **$b\bar{b}/c\bar{c}$  tagging** used in **precision measurements**
    - Top tagging folded into the same architecture
  - Unique challenges in producing mass-decorrelated tagger



# Towards the First Time Measurement of $gg \rightarrow ZH$ at the LHC Using Transformer Networks

Geoffrey Gilles, Wouter Verkerke, Marcel Vreeswijk

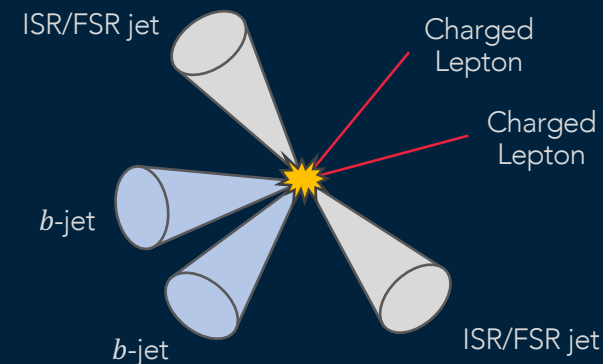
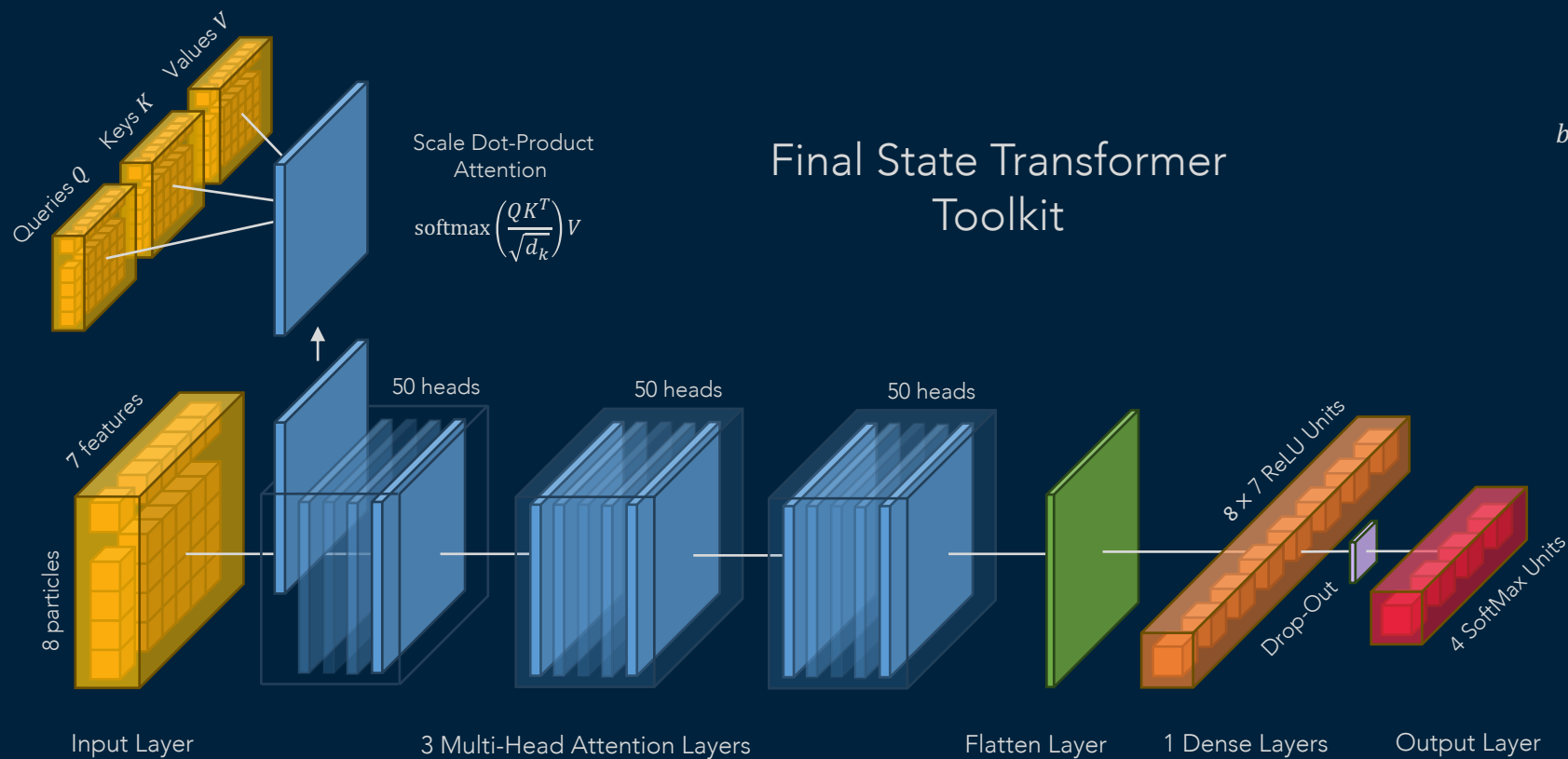
Extract New Physics Contributions from the Challenging  $gg \rightarrow ZH$  process



Simulation Based Inference  
empowered by  
Transformer network

# Towards the First Time Measurement of $gg \rightarrow ZH$ at the LHC Using Transformer Networks

## Learning from Particle Collision Event Final State

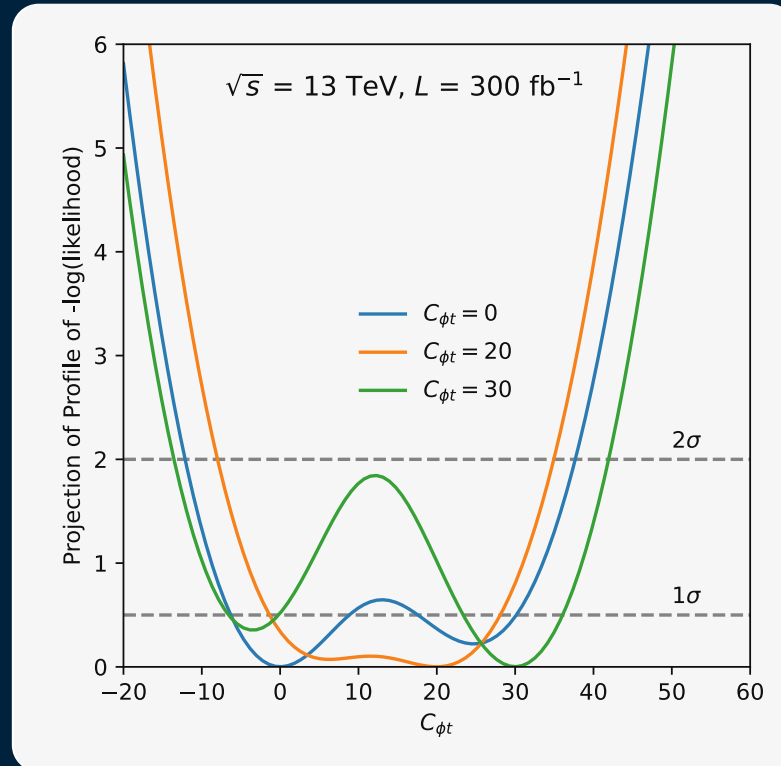


Discriminate signal from background  
 &  
 Assess new physics contributions

## Exploring Simulation Based Inference

Construct likelihood function  
based on Transformer  
output probabilities

$$\begin{aligned} \mathcal{L} = & P_{obs}^{SM} \\ & + c_{\phi t} \cdot P_{obs}^{INT} \\ & + c_{\phi t}^2 \cdot P_{obs}^{QUA} \\ & + \mu_{BKG} \cdot P_{obs}^{BKG} \end{aligned}$$



Capture intricacies of  
new physics interaction  
and detector responses  
in unprecedented details

Towards the First Time Measurement of  
 $gg \rightarrow ZH$  at the LHC Using Transformer Networks

Want to see the full story?

Visit my poster for in-depth findings and analysis!

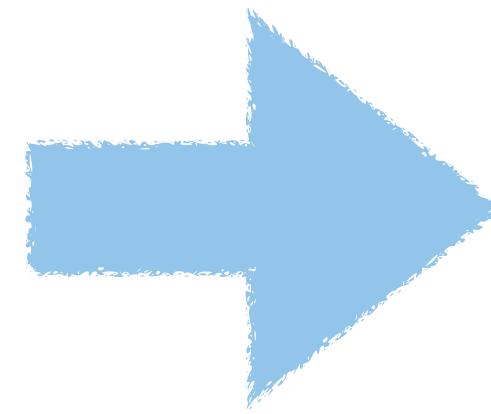
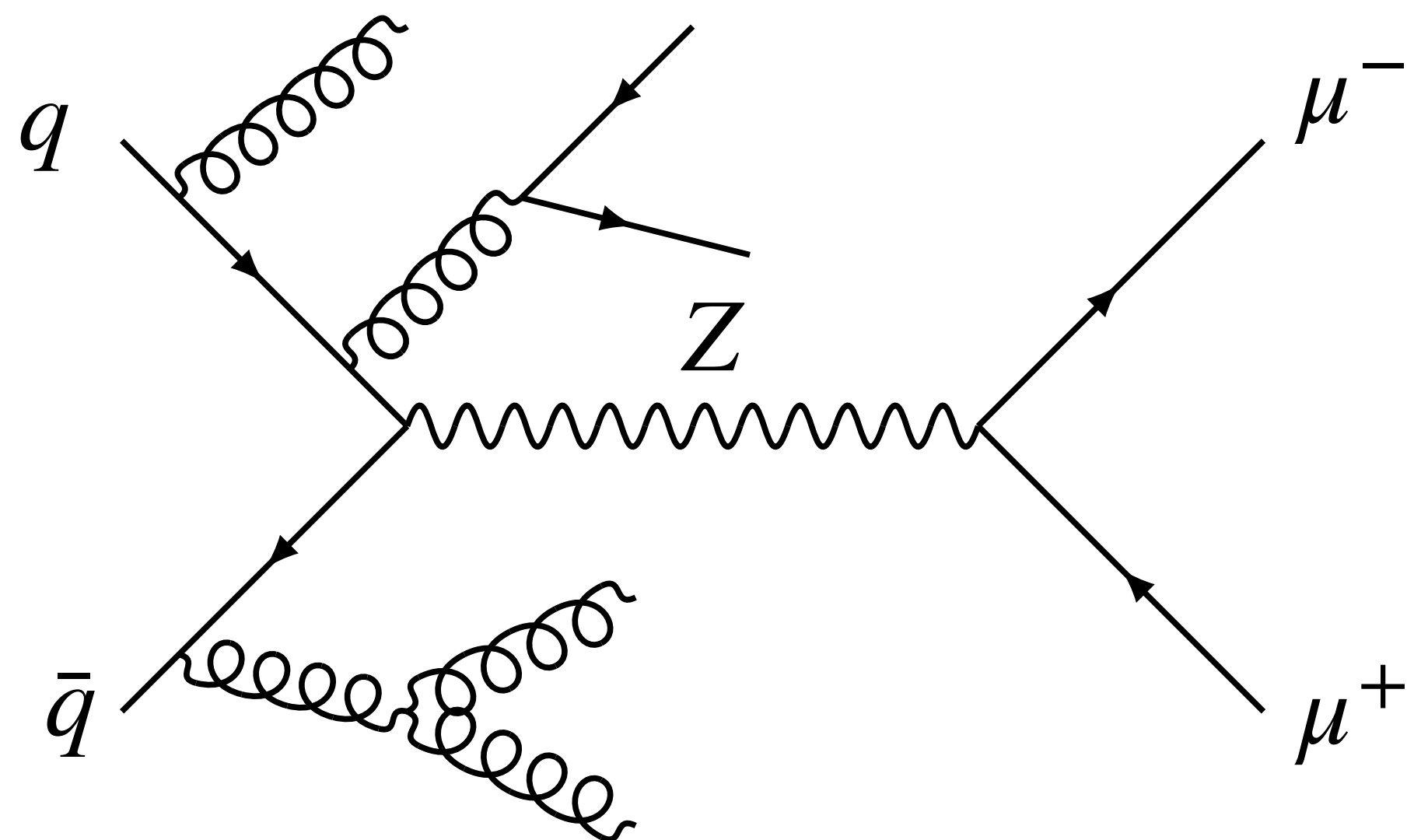


# LHC Event Generation with JetGPT

Anja Butter, Nathan Hütsch, Sofia Palacios Schweitzer, Tilman Plehn, Peter Sorrenson, Jonas Spinner, Nathanael Ediger, Maeve Madigan  
Heidelberg University

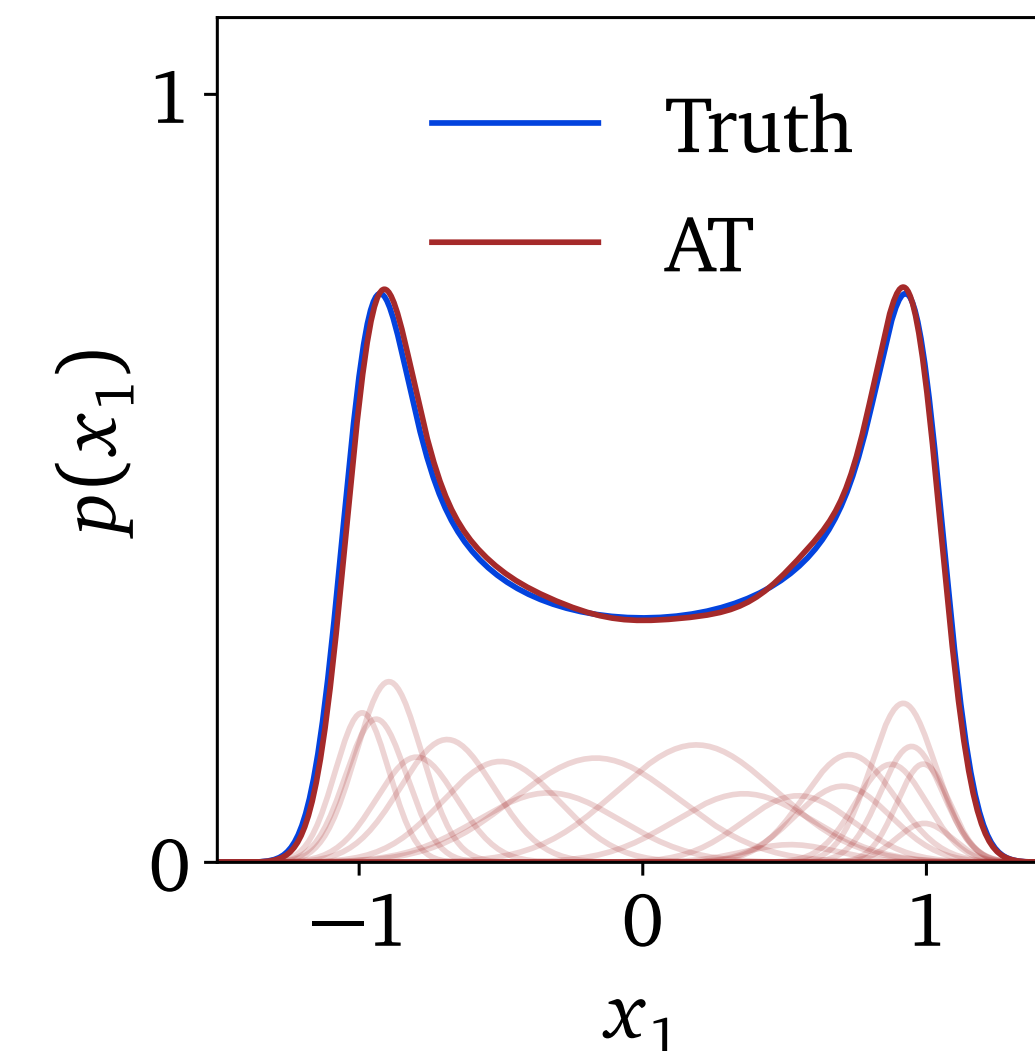
## Physics problem

- **Fast** generation of LHC events
- Learn challenging correlations to **percent-level**
- **Transfer knowledge** from cheap low-multiplicity events to expensive high-multiplicity events



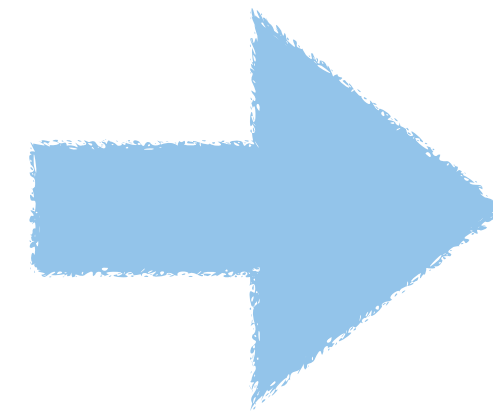
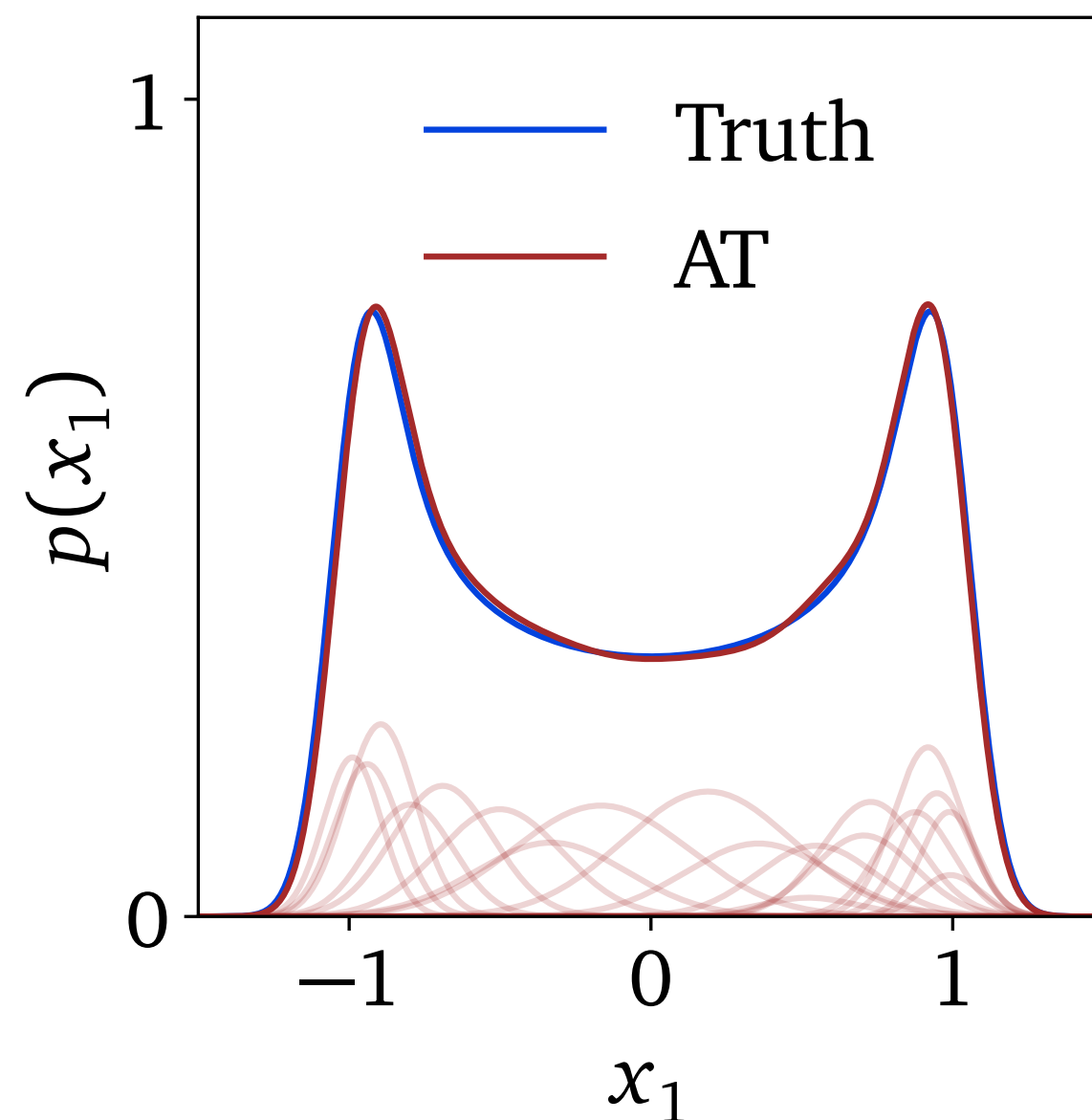
## ML solution

- Autoregressive **transformer**
- **Gaussian Mixture Model** likelihood
- **Neural classifier** to locate and reweight remaining discrepancies



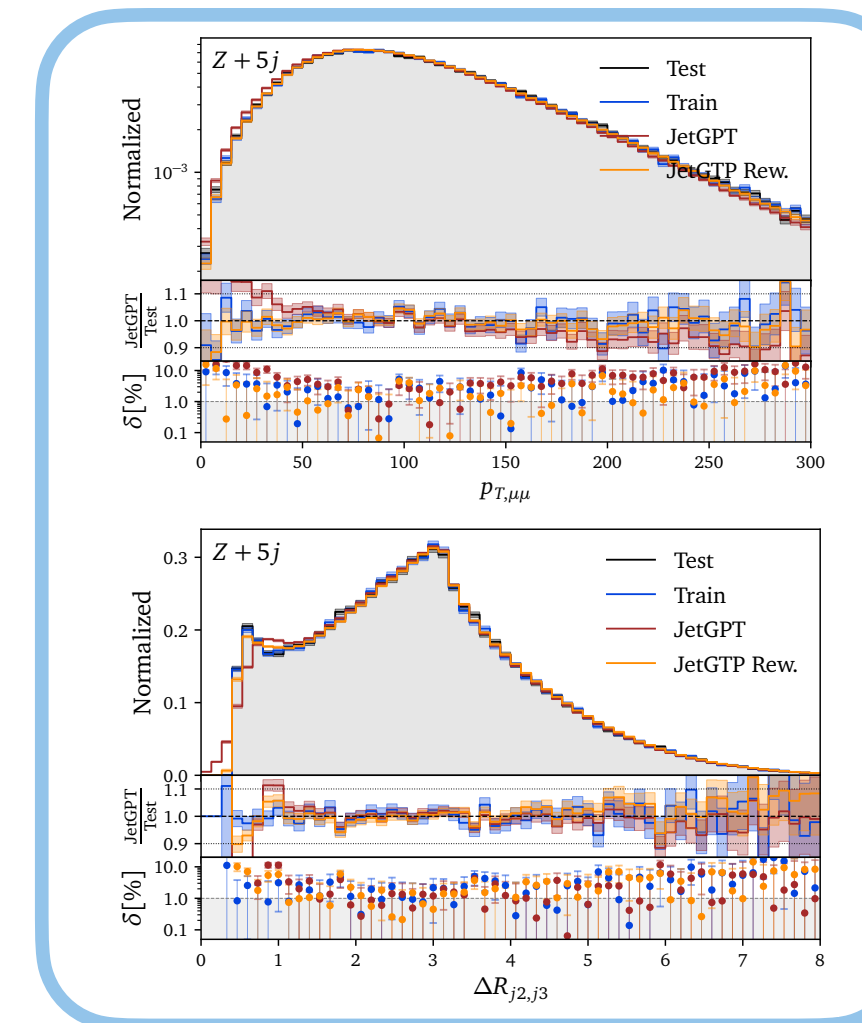
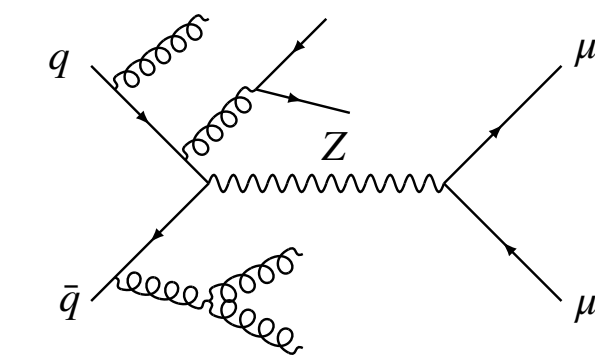
## ML solution

- Autoregressive **transformer**
- **Gaussian Mixture Model** likelihood
- **Neural classifier** to locate and reweight remaining discrepancies



## LHC Event Generation

- **Fast** generation of LHC events
- Learn challenging correlations to **percent-level**
- **Transfer knowledge** from cheap low-multiplicity events to expensive high-multiplicity events



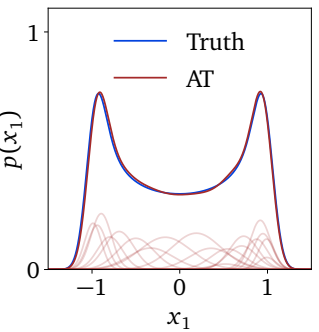
## Results

- **Joint training** on different multiplicities enhances performance and allows knowledge transfer
- **Autoregressive ordering** gives a powerful handle to control which features the model should focus on
- Neural classifiers to **locate and reweight** remaining discrepancies

## Autoregressive Transformer

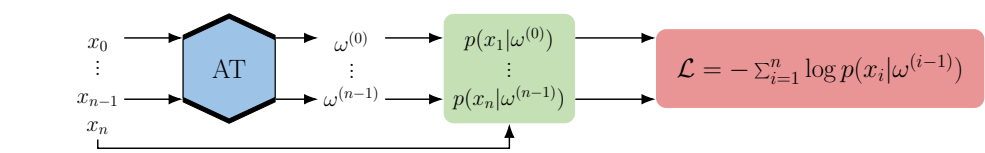
- Autoregressive **Gaussian Mixture Model**

$$p(x_1, x_2, \dots, x_n) = p(x_1)p(x_2|x_1)\dots p(x_n|x_{n-1}) = p(x_1|\omega^{(0)})p(x_2|\omega^{(1)})\dots p(x_n|\omega^{(n-1)})$$

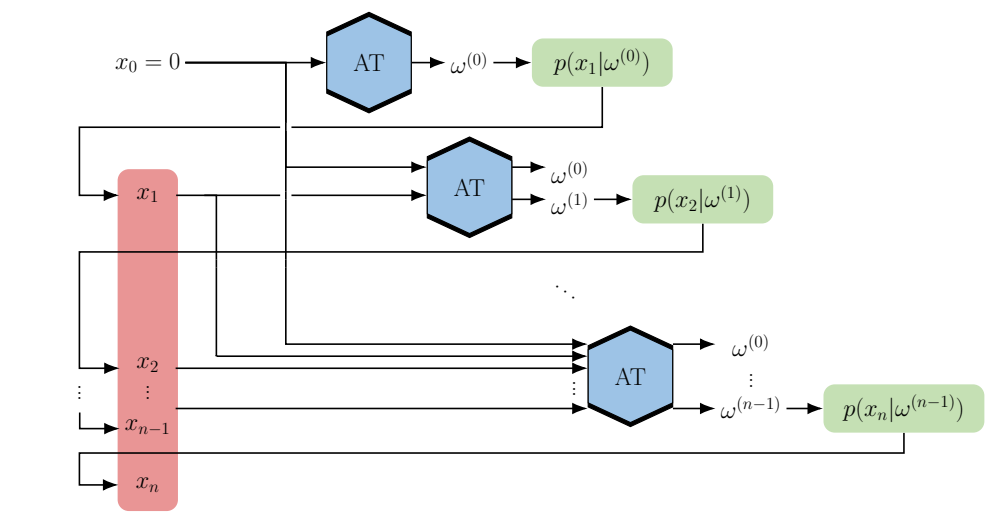


$$p(x_{i+1}|\omega^{(i)}) = \sum_{j=1} w_j^{(i)} \mathcal{N}(x_{i+1}|\mu_j^{(i)}, \sigma_j^{(i)}) \quad \omega^{(i)} = \{w_j^{(i)}, \mu_j^{(i)}, \sigma_j^{(i)}\}$$

- Training: **Parallellised density estimation**

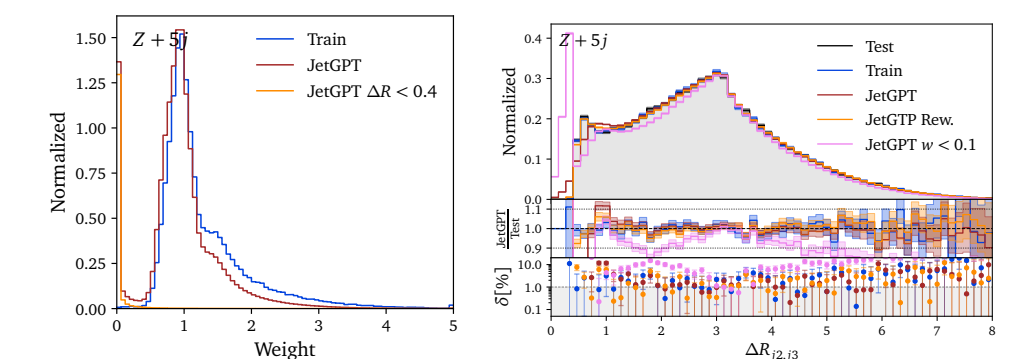


- Generation: **Autoregressive** sampling from one-dimensional distributions



## Classifier Control

- Neural classifiers approximate the **likelihood ratio**
- **Locate** discrepancies: Likelihood ratio as test statistic
- **Reweight** discrepancies: Likelihood ratio as weighting factor







**ATLAS**  
EXPERIMENT

Abstract #126; poster Thur #95

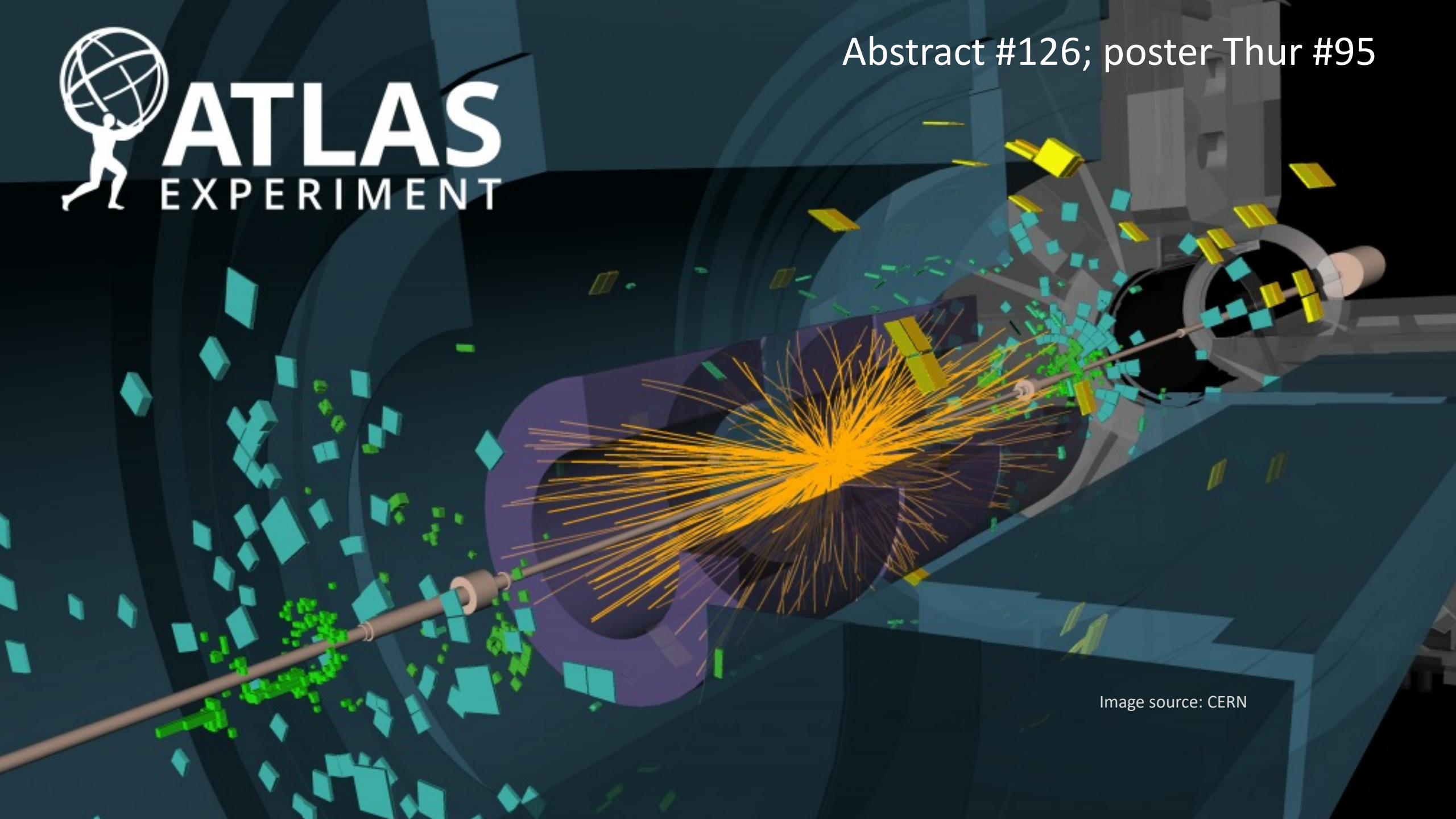


Image source: CERN

Abstract #126; poster Thur #95

Task:  
Reconstruct tracks  
from 3D point cloud



Abstract #126; poster Thur #95

Task:  
Reconstruct tracks  
from 3D point cloud

Approach:  
Transformer-inspired models

Abstract #126; poster Thur #95

# Transformer-inspired ML models for particle track reconstruction

Yue Zhao

High performance machine learning group, SURF, the Netherlands

[yue.zhao@surf.nl](mailto:yue.zhao@surf.nl)