Track Reconstruction Using Transformers

Nadezhda Dobreva

nadezhda.dobreva@ru.nl

ICIS, Radboud University

MSc student in Data Science and Artificial Intelligence

Team

- **The project collaborates:**
    - Radboud University (Nadezhda Dobreva)
    - Nikhef (Sascha Caron, Zef Wolffs, Uraz Odyurt)
    - SURF (Yue Zhao)
- **Via IMAPP, ELLIS and Radboud AI**

The Problem

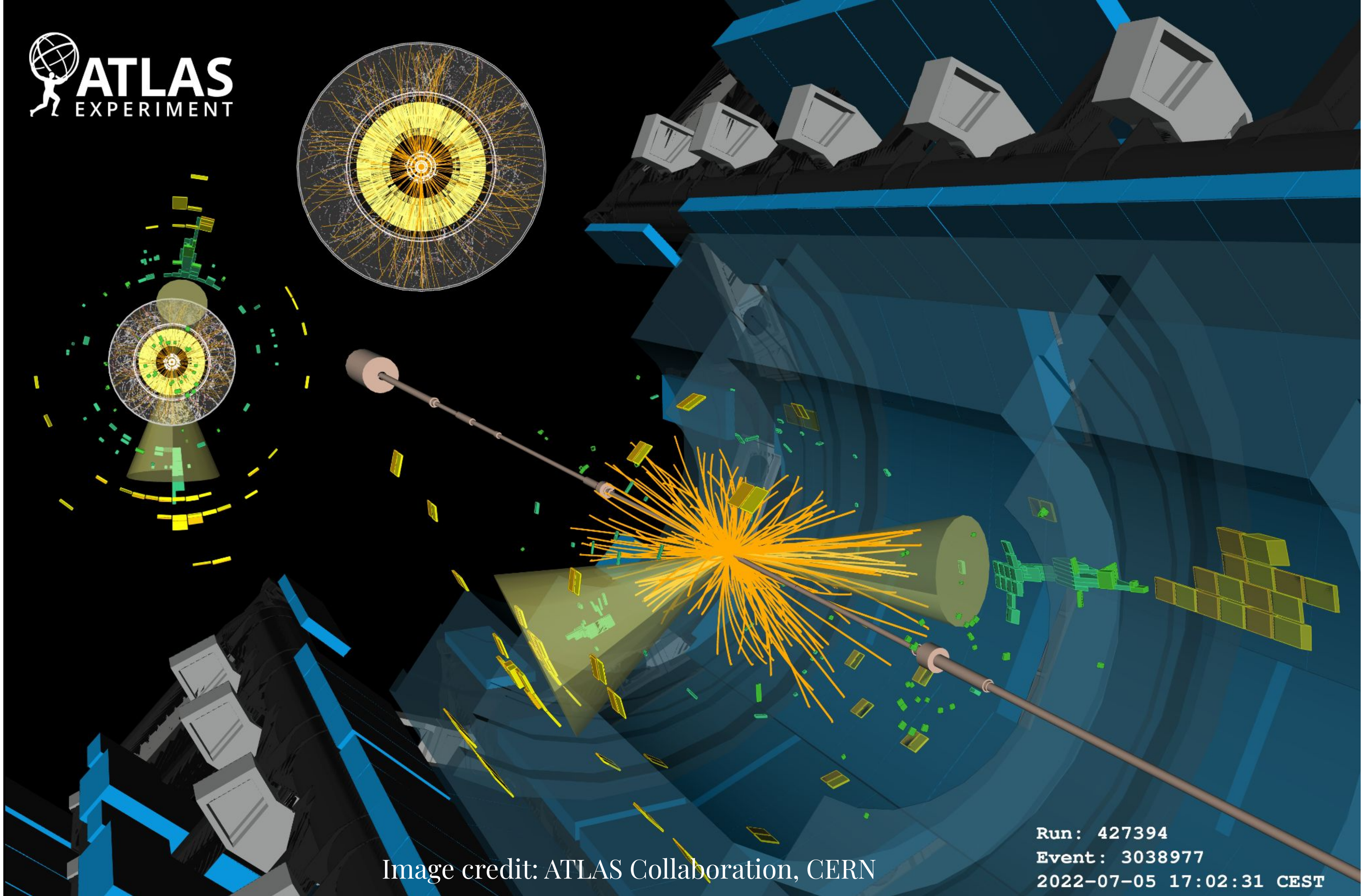


Image credit: ATLAS Collaboration, CERN

Scalability Issue

- **Kalman filters: scaling quadratically, inherently sequential [1, 2]**
- **High Luminosity LHC: detector occupancy scales up [3]**

[1] Lantz, Steven, et al. "Speeding up particle track reconstruction using a parallel Kalman filter algorithm." Journal of Instrumentation 15.09 (2020): P09030.

[2] Tsaris, Aristeidis, et al. "The HEP. TrkX project: deep learning for particle tracking." Journal of Physics: Conference Series. Vol. 1085. IOP Publishing, 2018.

[3] Apollinari, Giorgio, Lucio Rossi, and Oliver Brüning. High luminosity LHC project description. No. CERN-ACC-2014-0321. 2014.

Alternatives

- **An active field of research**
  - TrackML challenge [4]
  - GNN tracking - graph-based solutions [5]
  - Other deep learning approaches

[4] https://www.kaggle.com/competitions/trackml-particle-identification/

[5] Caillou, Sylvain, et al. ATLAS ITk Track Reconstruction with a GNN-based pipeline. No. ATL-ITK-PROC-2022-006. ATL-COM-ITK-2022-057, 2022.

The Transformer

What is a Transformer?

- Deep learning architecture
- Success in NLP

“The butterfly
hits
is beautiful.” → transformer → “Пеперудата
tracks
е красива.”

Transformer Architecture [6]

- Encoder
- Decoder
- Multi-head attention mechanism

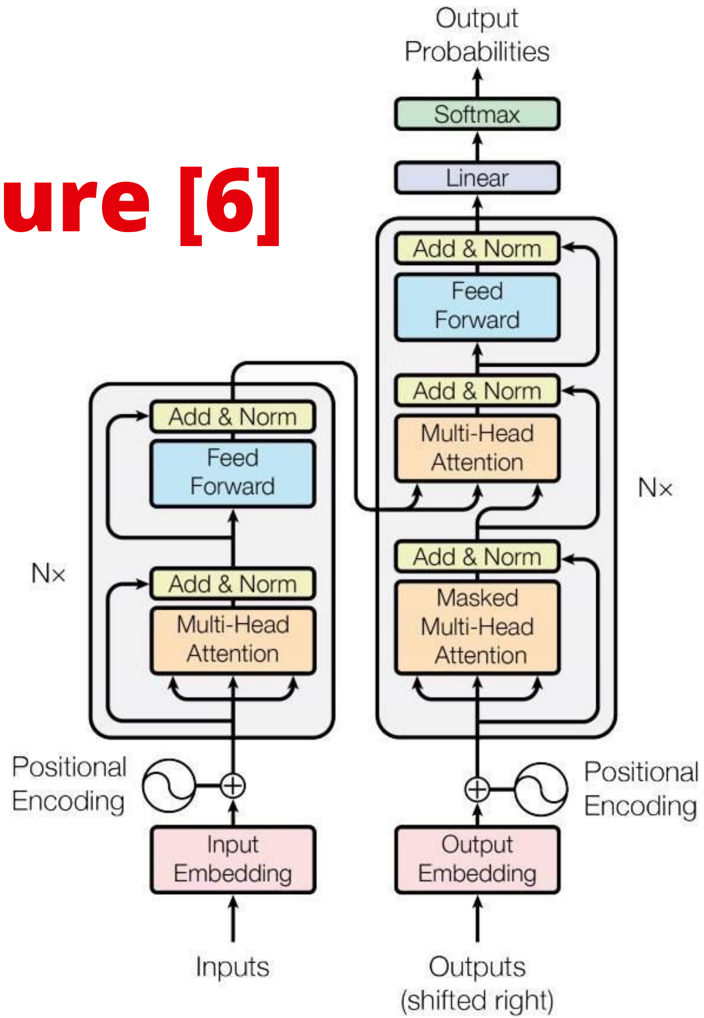


Image credit: [6]

[6] Vaswani, Ashish, et al. "Attention is all you need." Advances in neural information processing systems, 30 (2017).

Motivation

- Runs in parallel
- Transformer pipeline can be trained end-to-end
- Can handle variable length input
- Equivariant to input order
- Good at capturing complex non-linear dynamics in data
- Good at sequential data

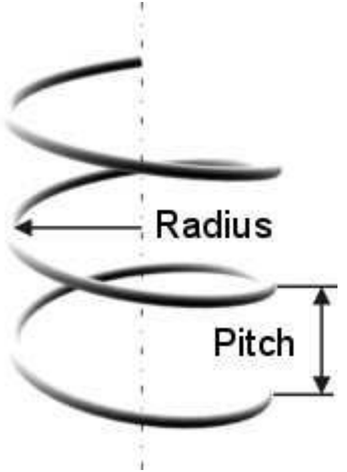
The Data

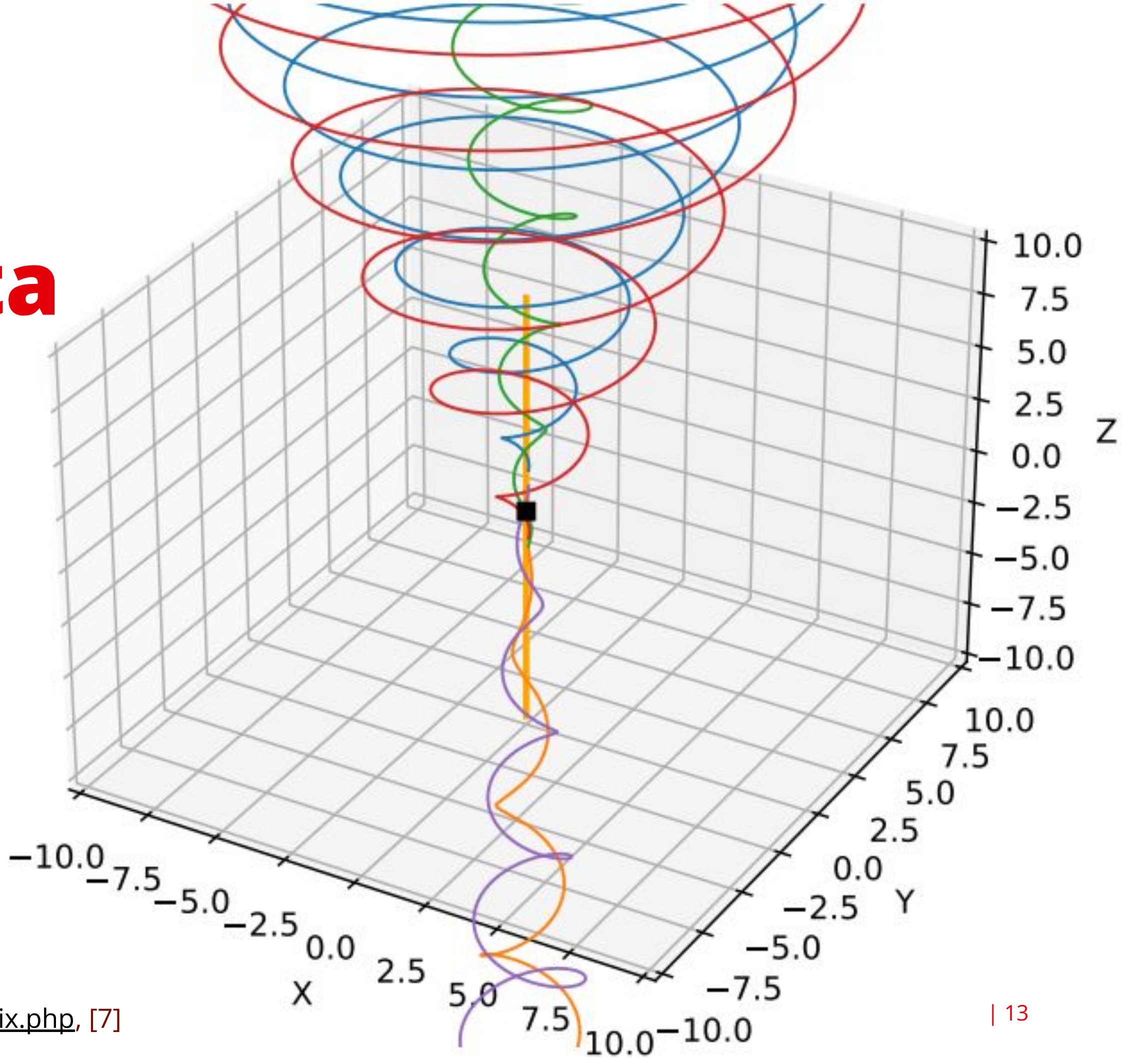
Complexity-Reduced Datasets

- REDuced VIrtual Detector (REDVID) for simulation of events [7, 8]
- Iterative increase of complexity

[7] Reduced Simulations for High-Energy Physics, a Middle Ground for Data-Driven Physics Research, Uraz Odyurt, Stephen Nicholas Swatman, Ana-Lucia Varbanescu, Sascha Caron, 2023
[8] https://virtualdetector.com/redvid/

Example
3D Helical Data



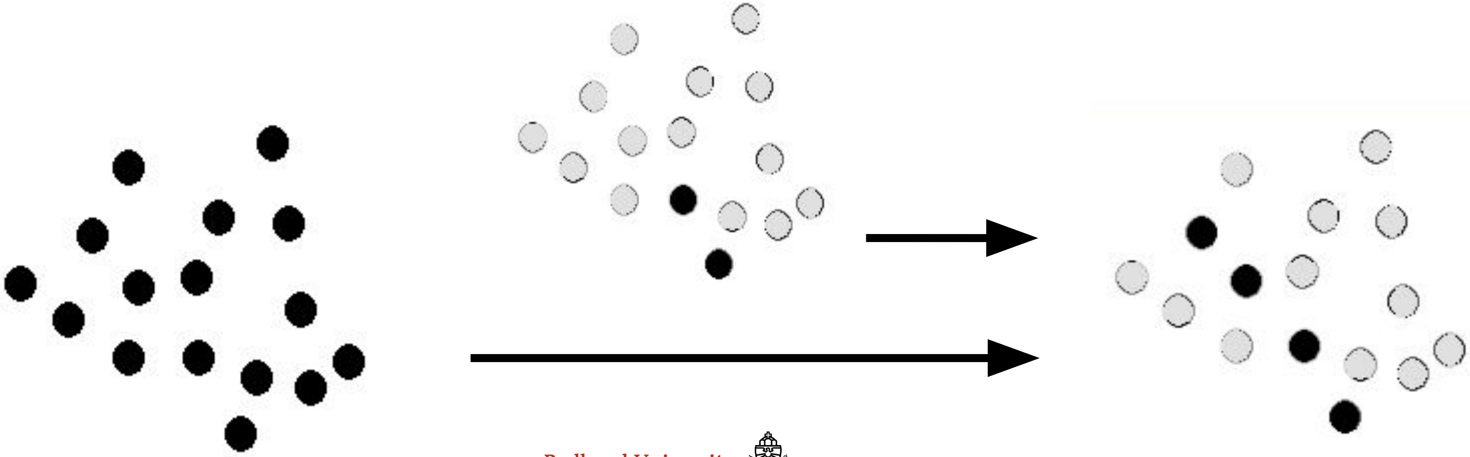


Images: https://www.to-calculate.com/geometry/evolute-helix.php, [7]

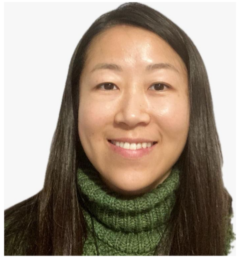
Proposed Approaches

Encoder-Decoder Model

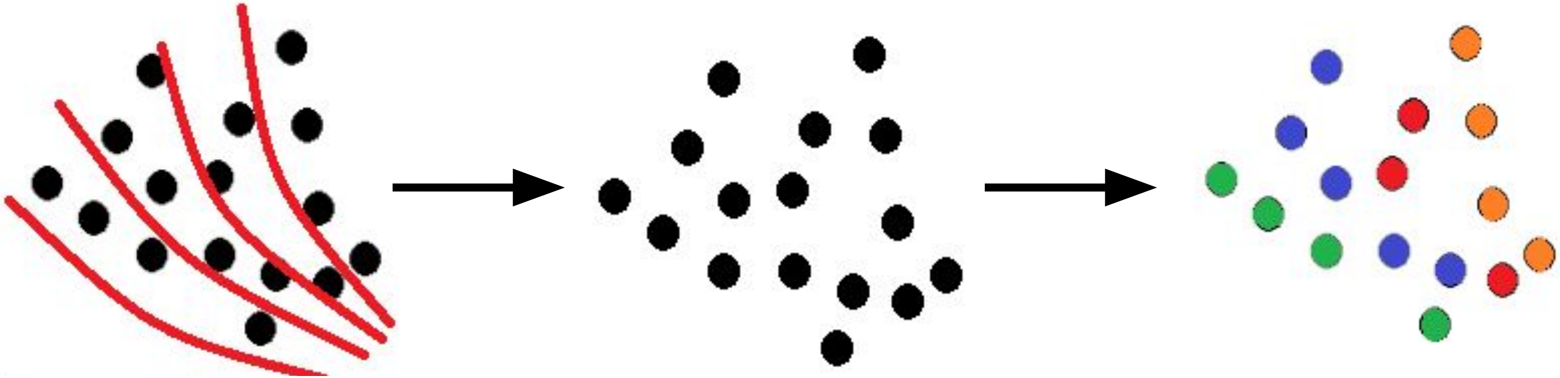
- **Encoder: Encodes event hits**
- **Decoder: Predicts next hit in track. Autoregressively builds the full track, starting from a given seed**



Encoder-only Classifier

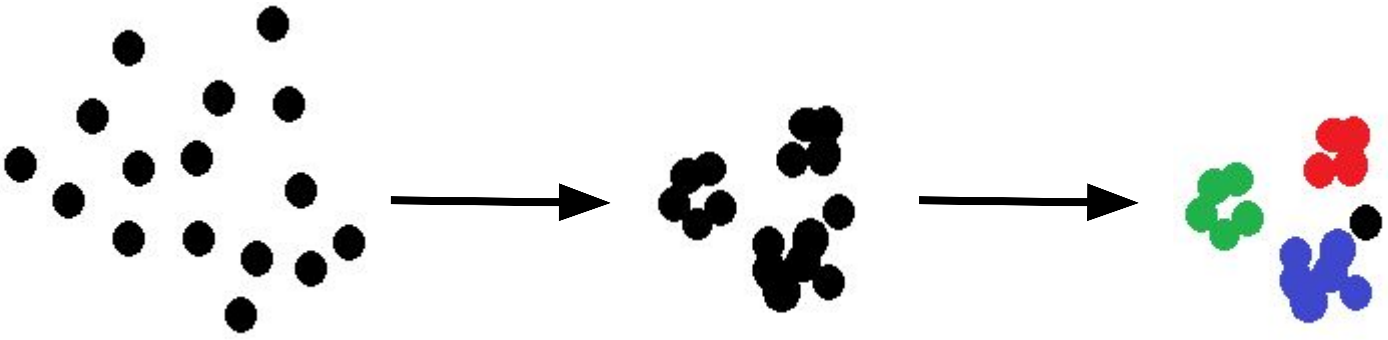


- Track defining parameters placed in equally sized bins (i.e. classes)
- Transformer predicts the class of each hit



Encoder-only Regressor

- Used for regressing track-defining parameters
- Clustering hits based on regressed parameters



Preliminary Results

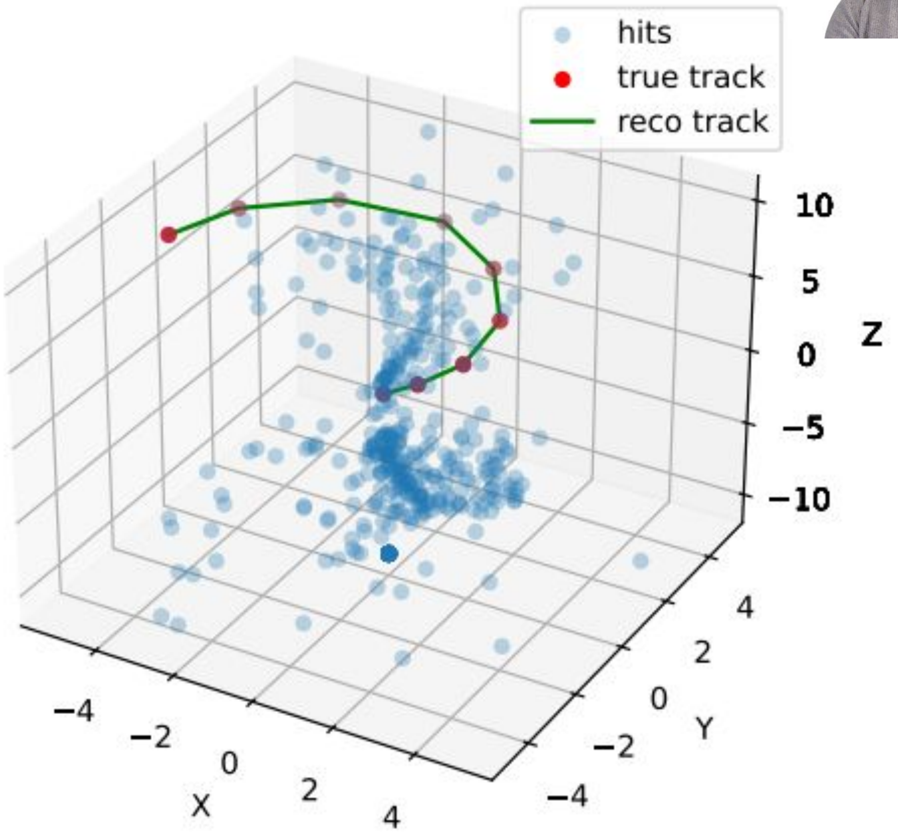
Performance Evaluation

- **TrackML score**
  - A reconstructed track is scored as “correct” if more than 50% of its hits originate from the same truth particle [9]
  - Used for all state-of-the-art models
- **Reported results on 3D helical noisy data with 10 - 50 tracks per event**

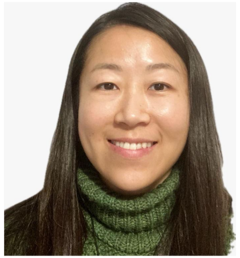
[9] Kiehn, Moritz, et al. "The TrackML high-energy physics tracking challenge on Kaggle." EPJ Web of Conferences. Vol. 214. EDP Sciences, 2019.

Encoder-Decoder Model: 85%

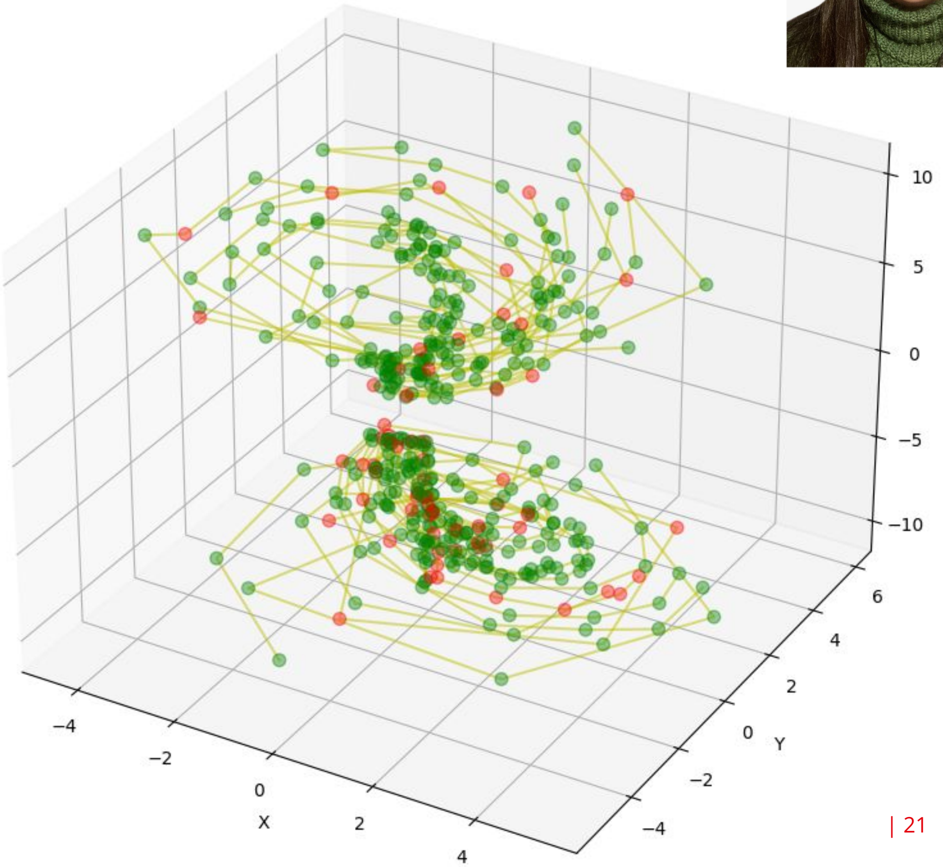
- Example event
- Green: the reconstructed track
- Red: the hits from the true track



Encoder-only Classifier: 98%

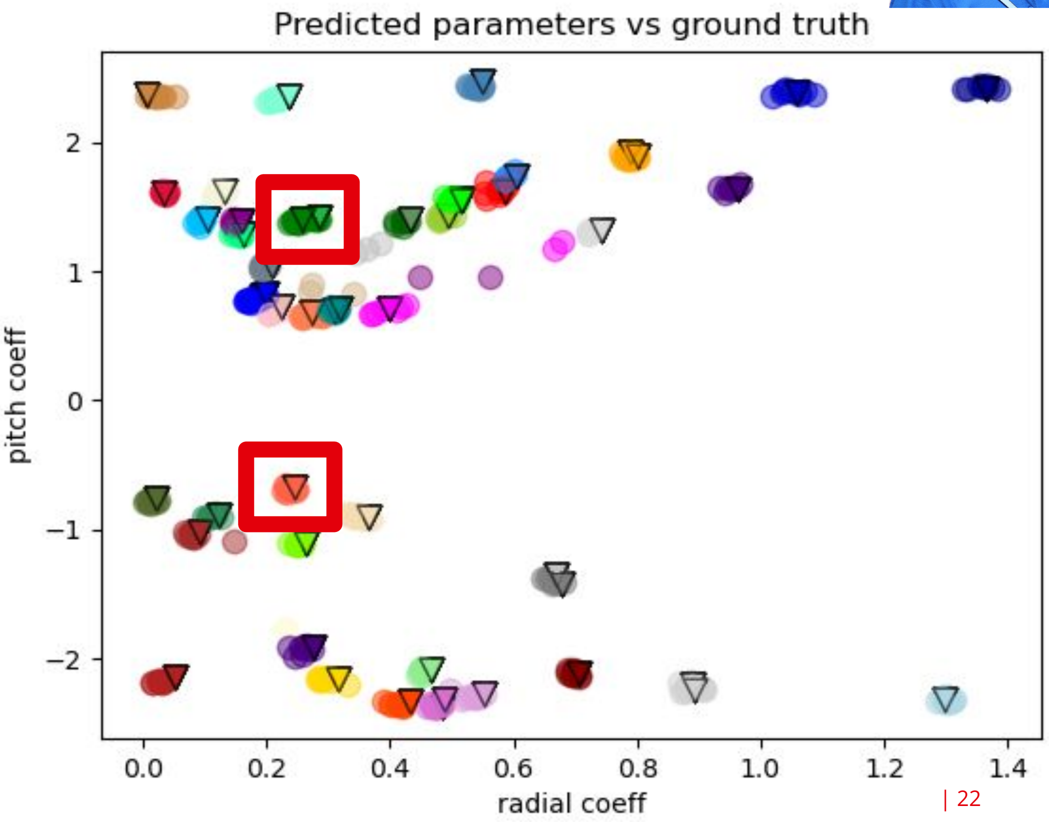


- Example event
- Green: correctly classified hits
- Red: incorrectly classified hits



Encoder-only Regressor: 87%

- Example event
- Circles: regressed track parameters
- Triangles: actual track parameters



Conclusion

Questions?
Thank you.