Track 2 Summary

ACAT 2024

Lukas Heinrich, TUM

Track 2: Data Analysis - Algorithms and Tools

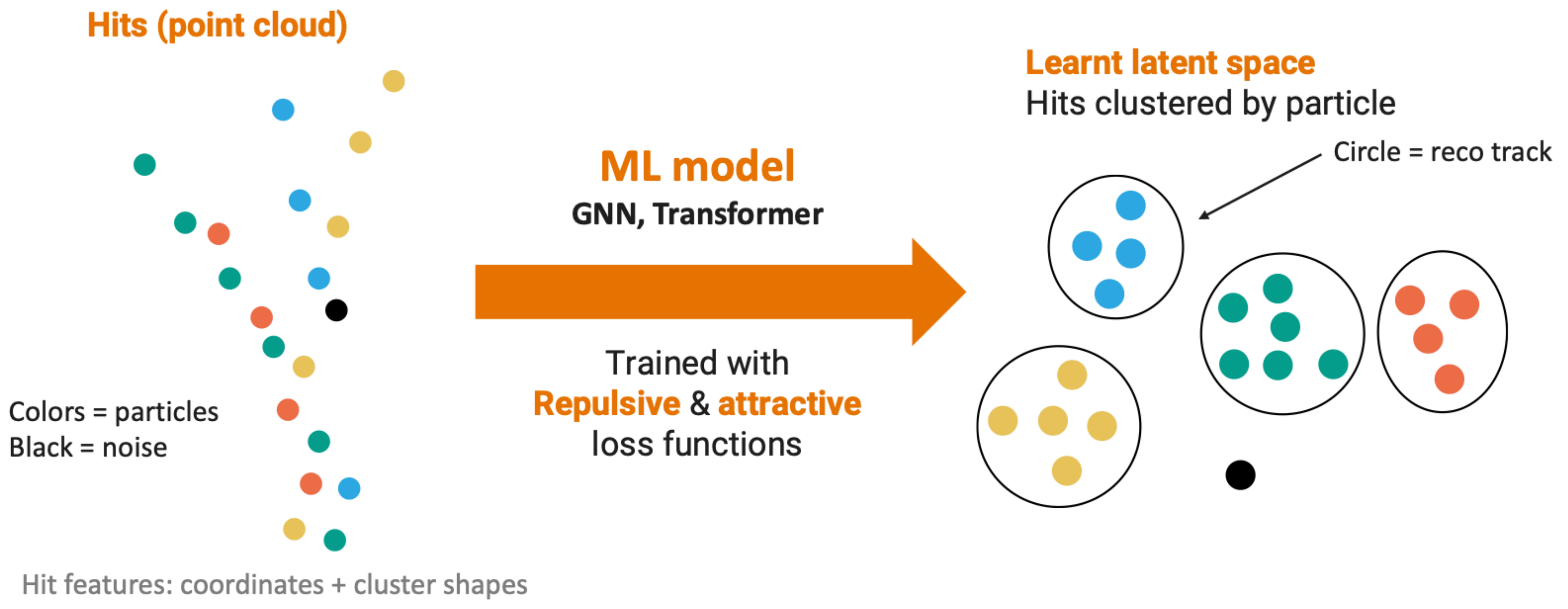
Track to discuss latest developments in reconstruction and analysis tools.

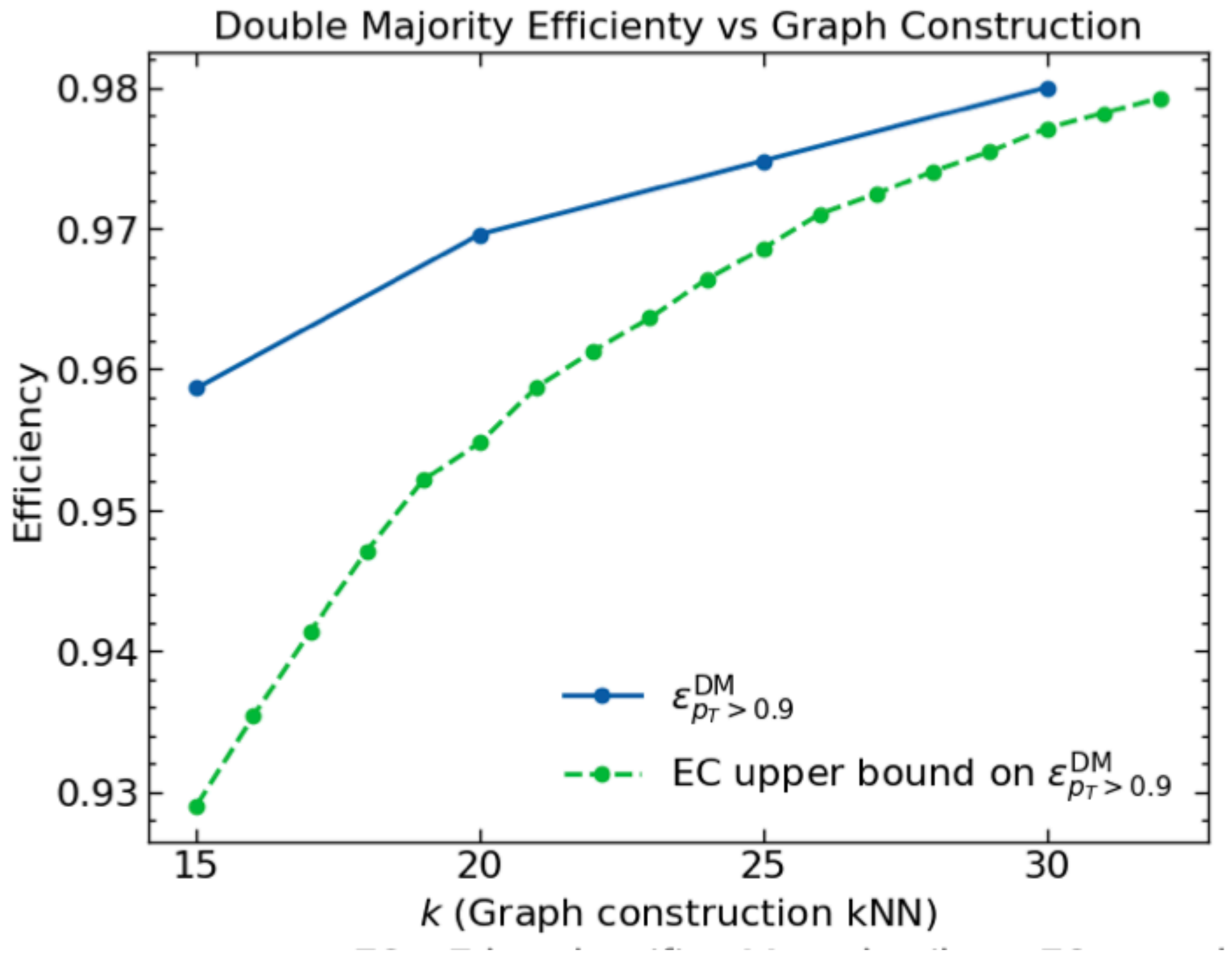
32 Talks. Lots of ML but also great talks that go beyond it incl. online systems, data management, statistics …

Impossible to summarize everything, but short lightning overview & some high-level thoughts

Single Shot Tracking

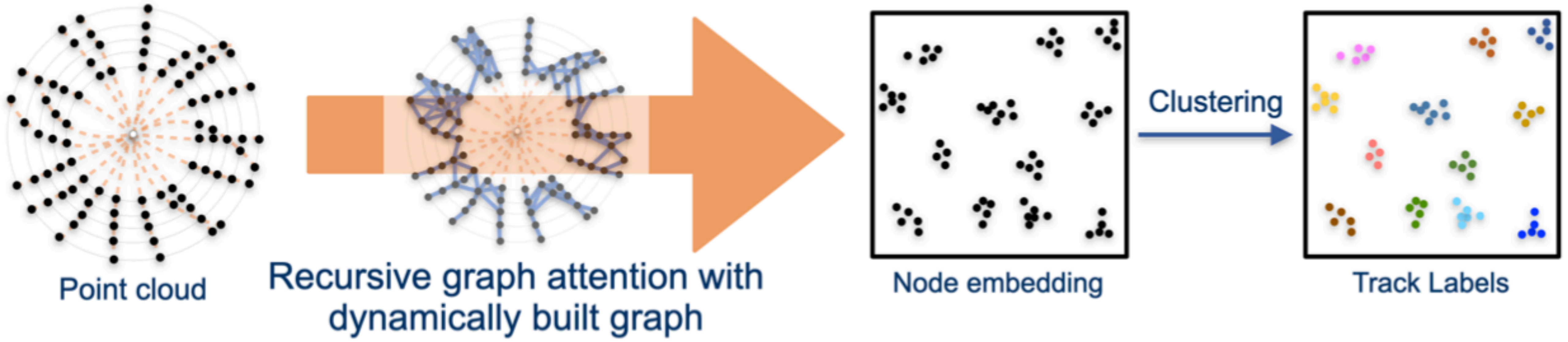
Kilian Lieret

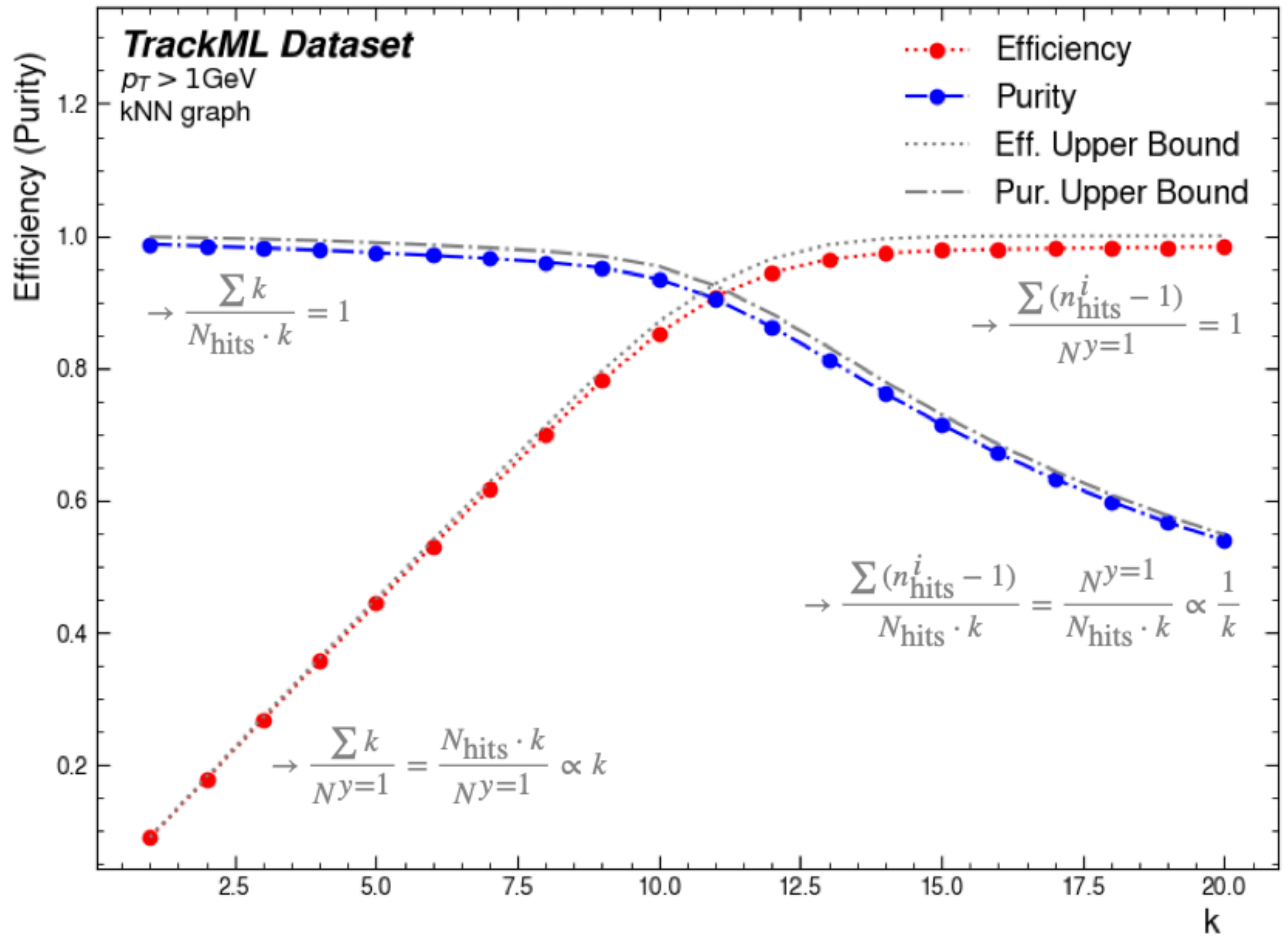




GNN + Object Condensation

Jay Chan

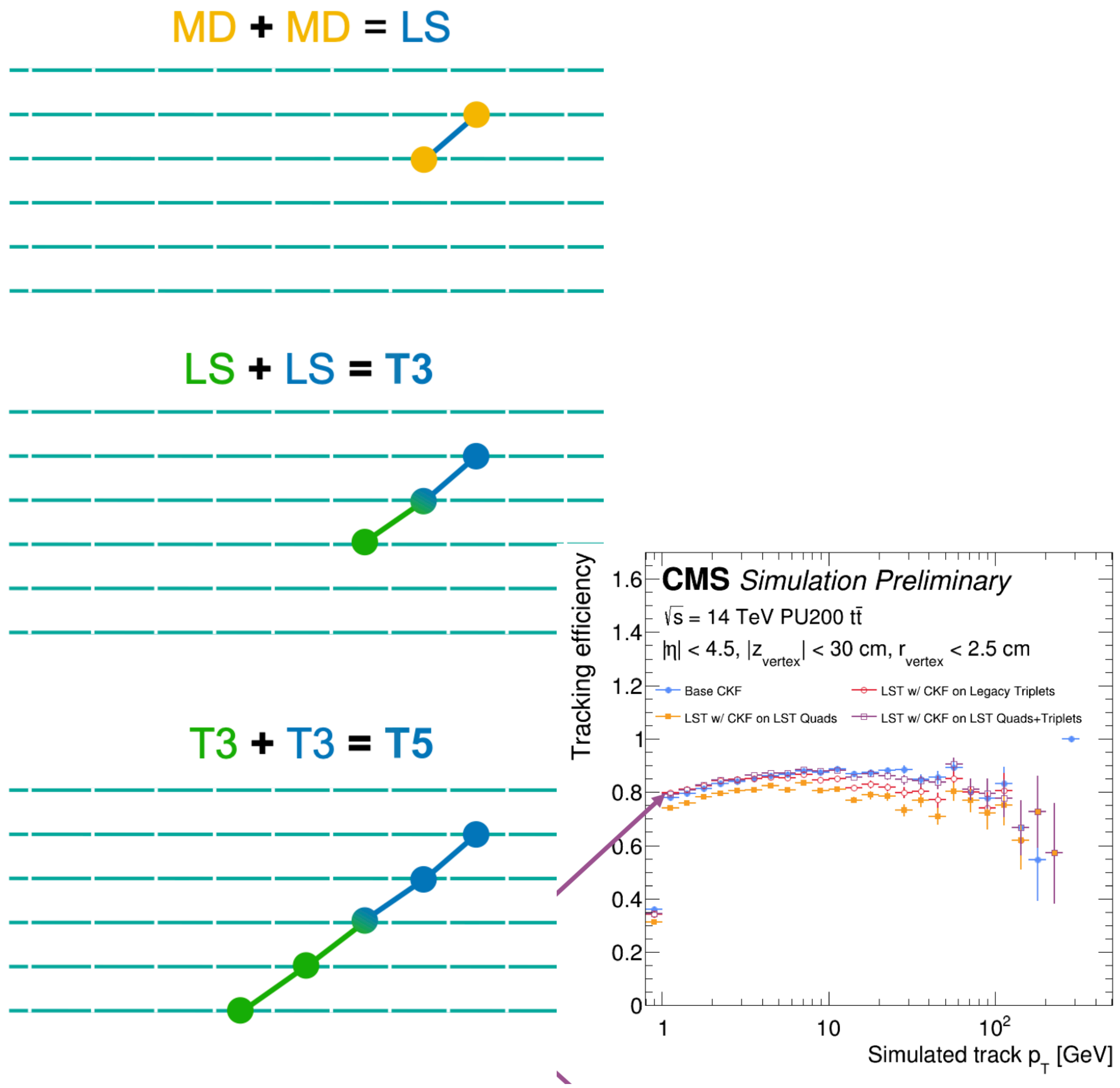




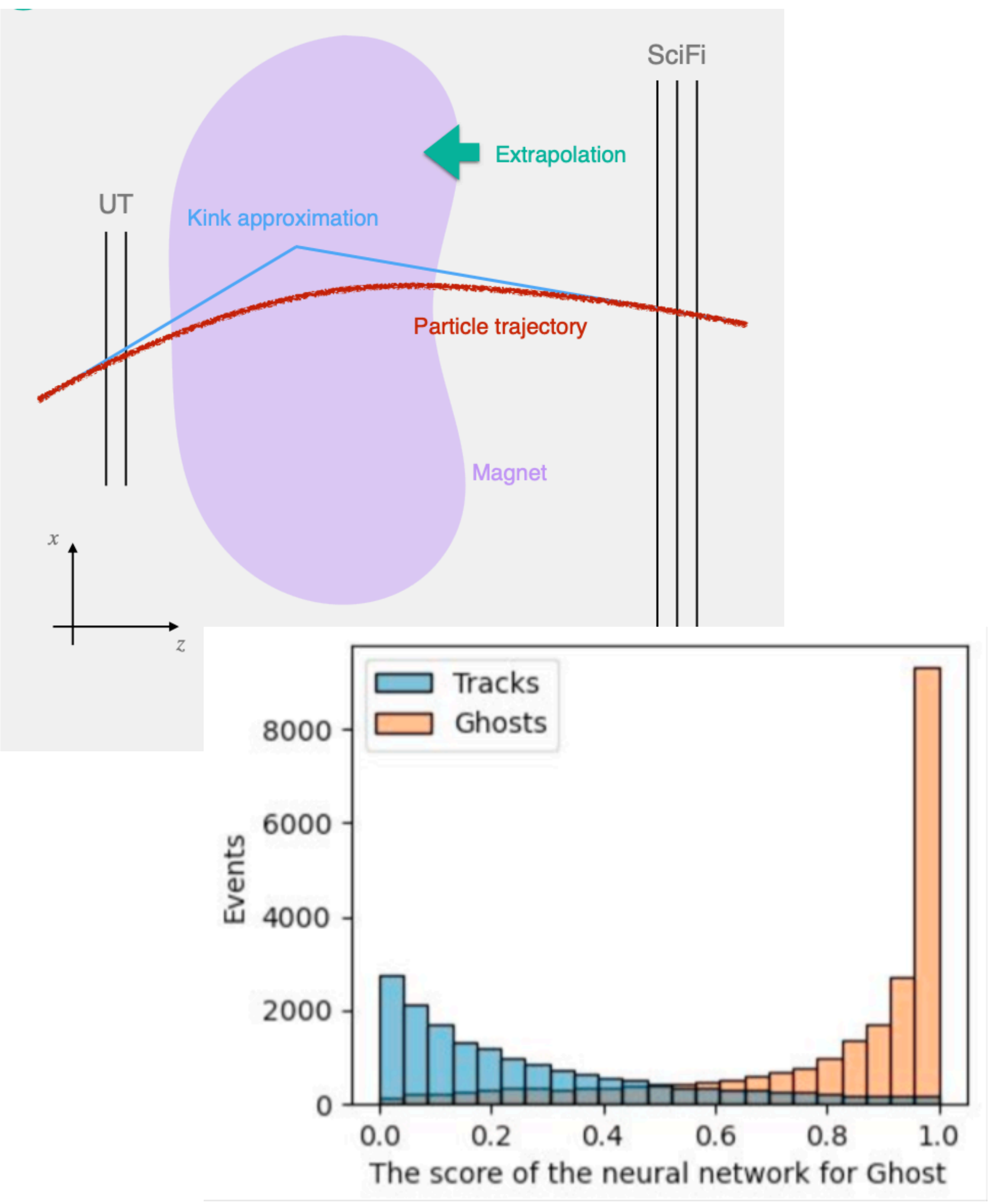
Recursive Graph Attention + DBSCAN

Online Tracking Algorithms

Online Environment is special. Generally not the place for gigantic models, but tailored solution.



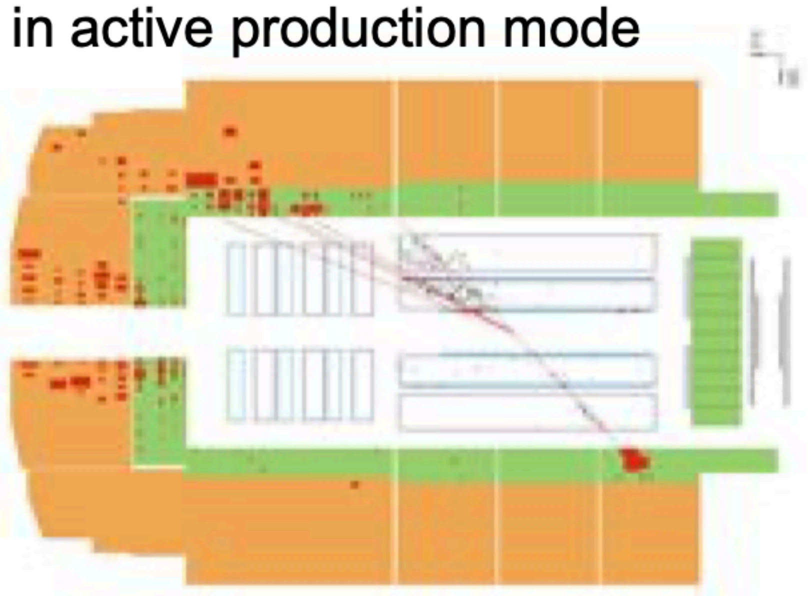
Fast and Parallelizable Tracking for CMS Trigger



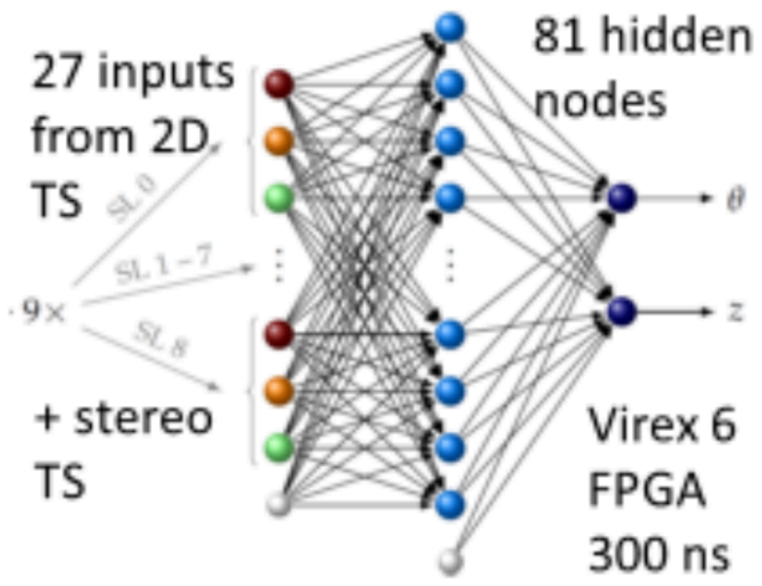
Downstream Tracking in LHCb Trigger

AINHEP 1999

H1 @HERA ep Collider:
First Neural Trigger in HEP
in active production mode



ACAT 2024



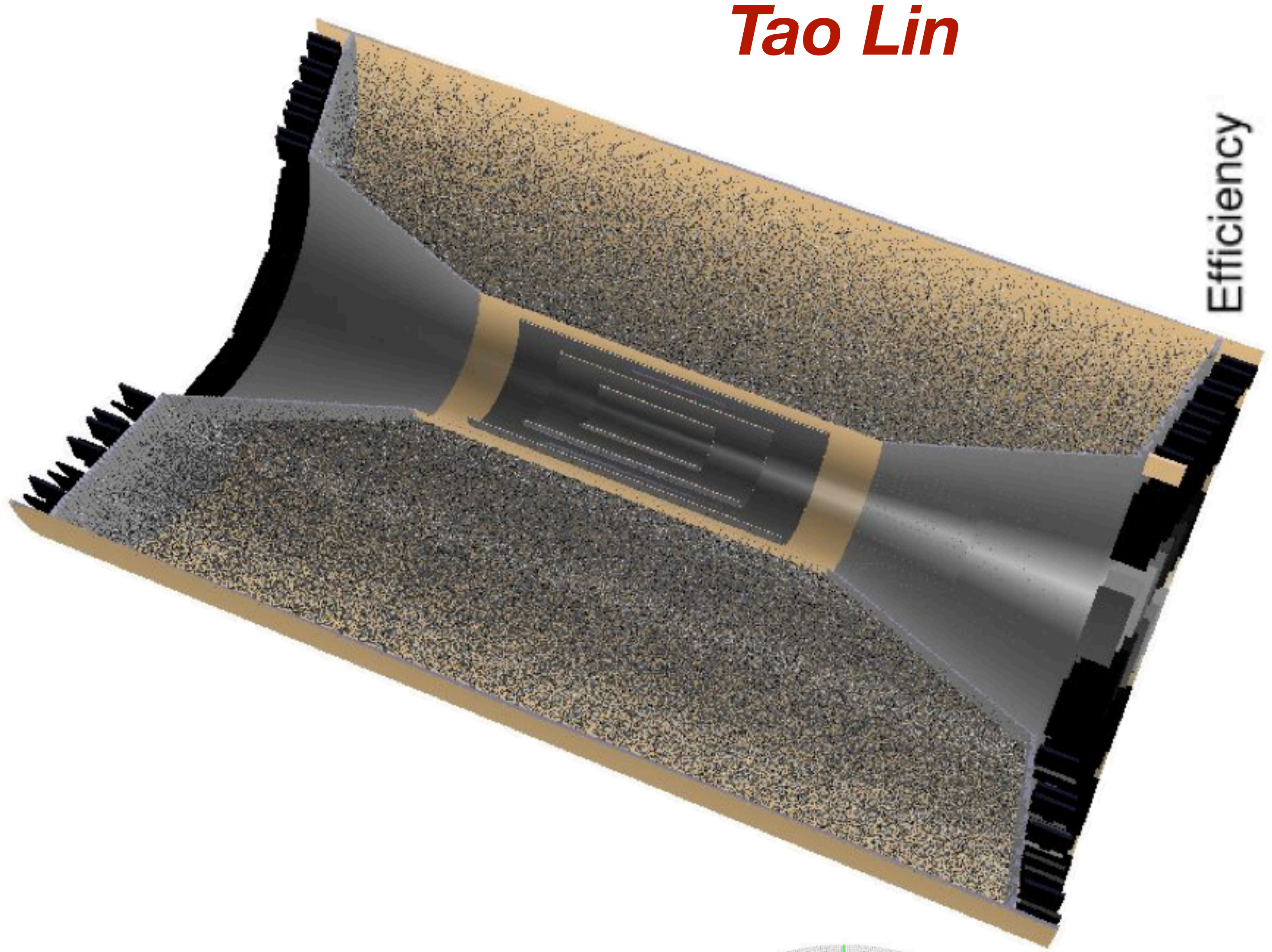
Neural Network Trigger at Belle II

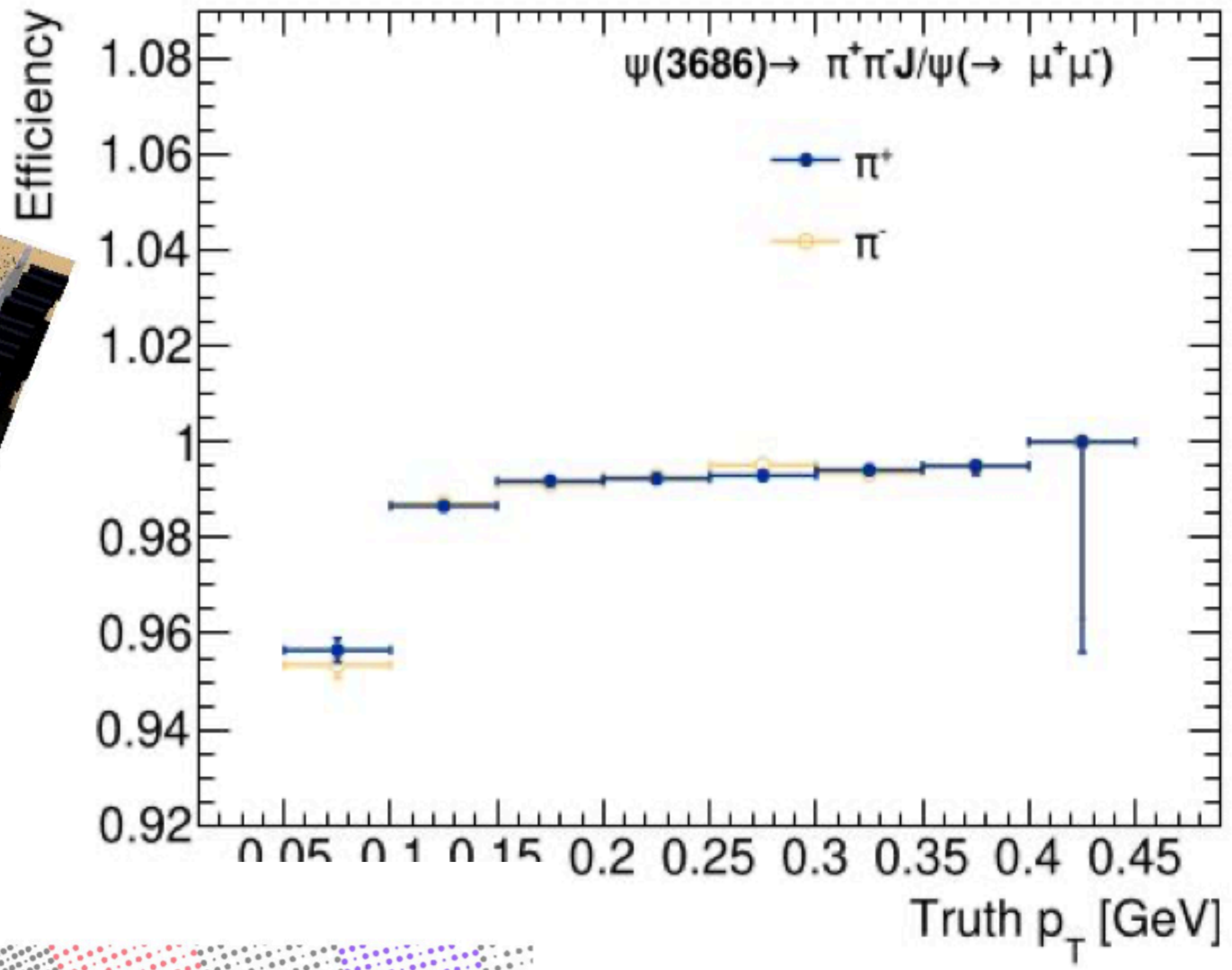
Tracking beyond LHC & Silicon

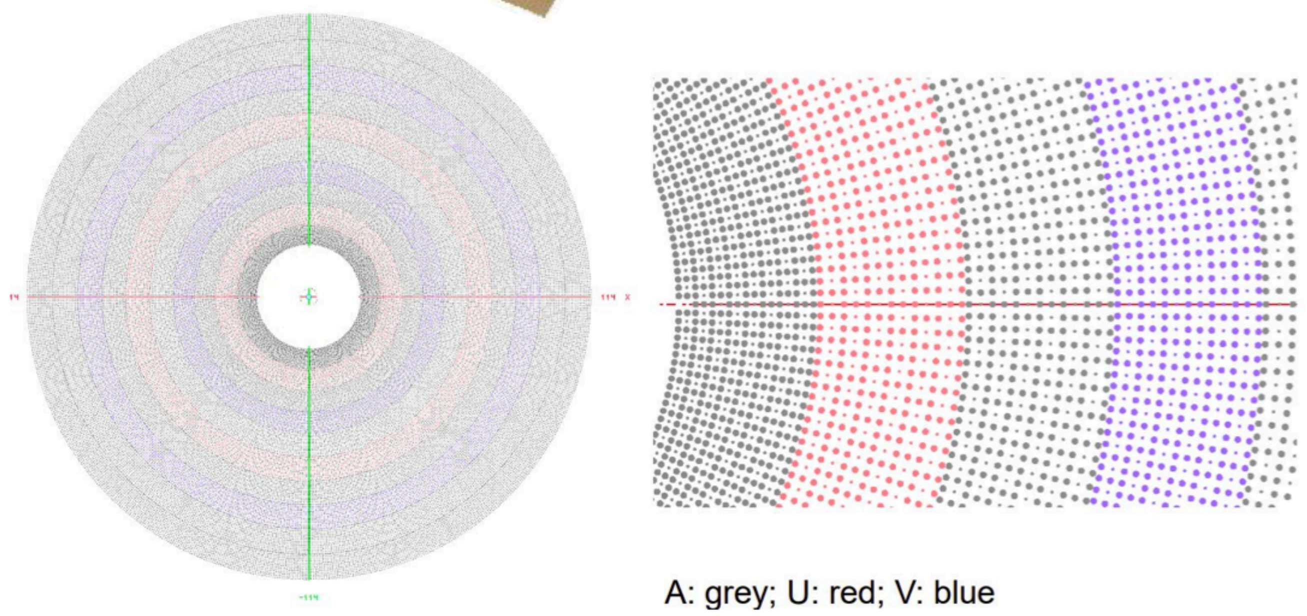
Not everything is the LHC. Unique challenges & Solutions

Tao Lin

More in talk at CHEP2023
and JINST **18** P07026

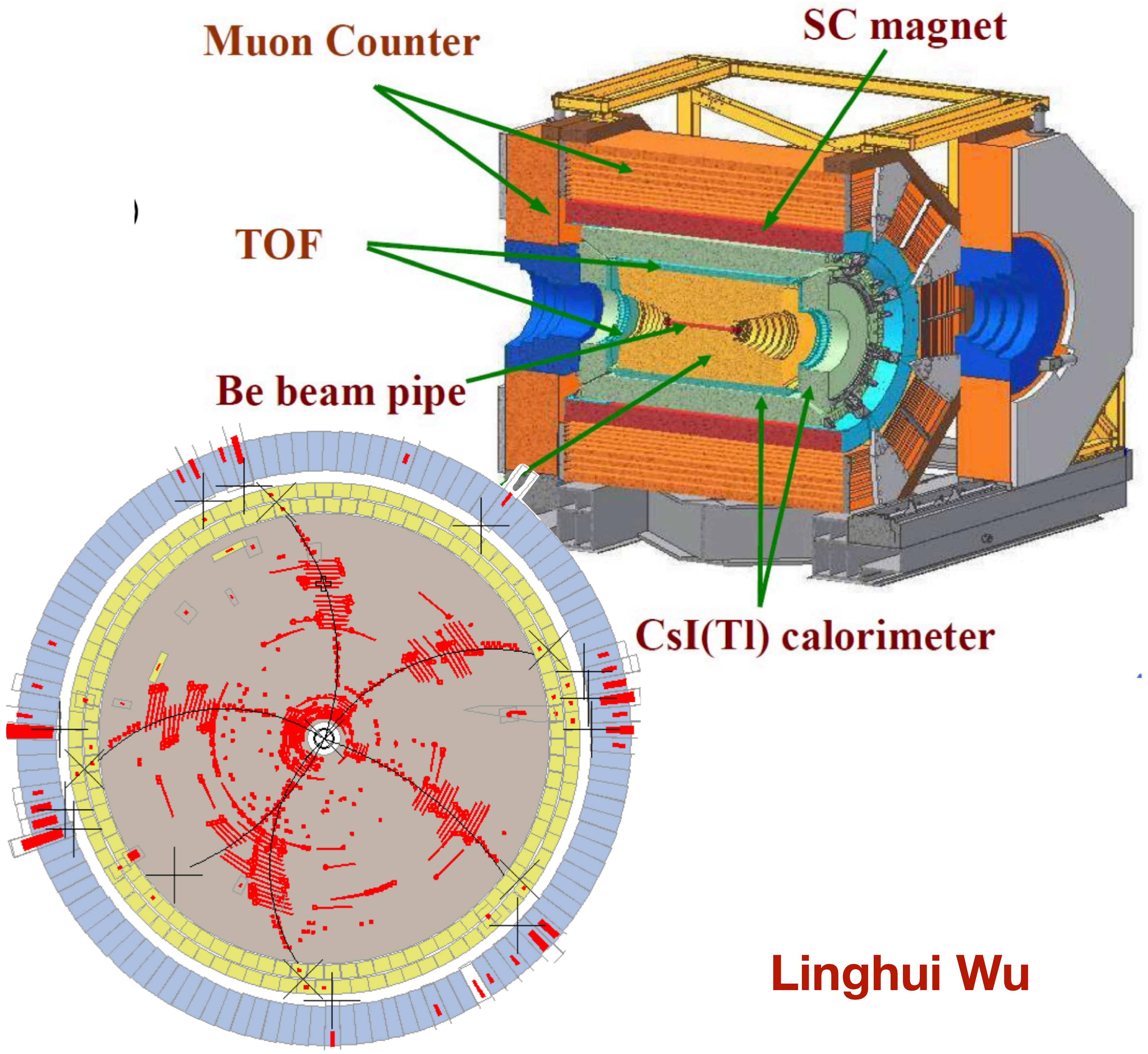






A: grey; U: red; V: blue

ACTS for Gaseous Detectors (
Extending Community Tools & Open Source!

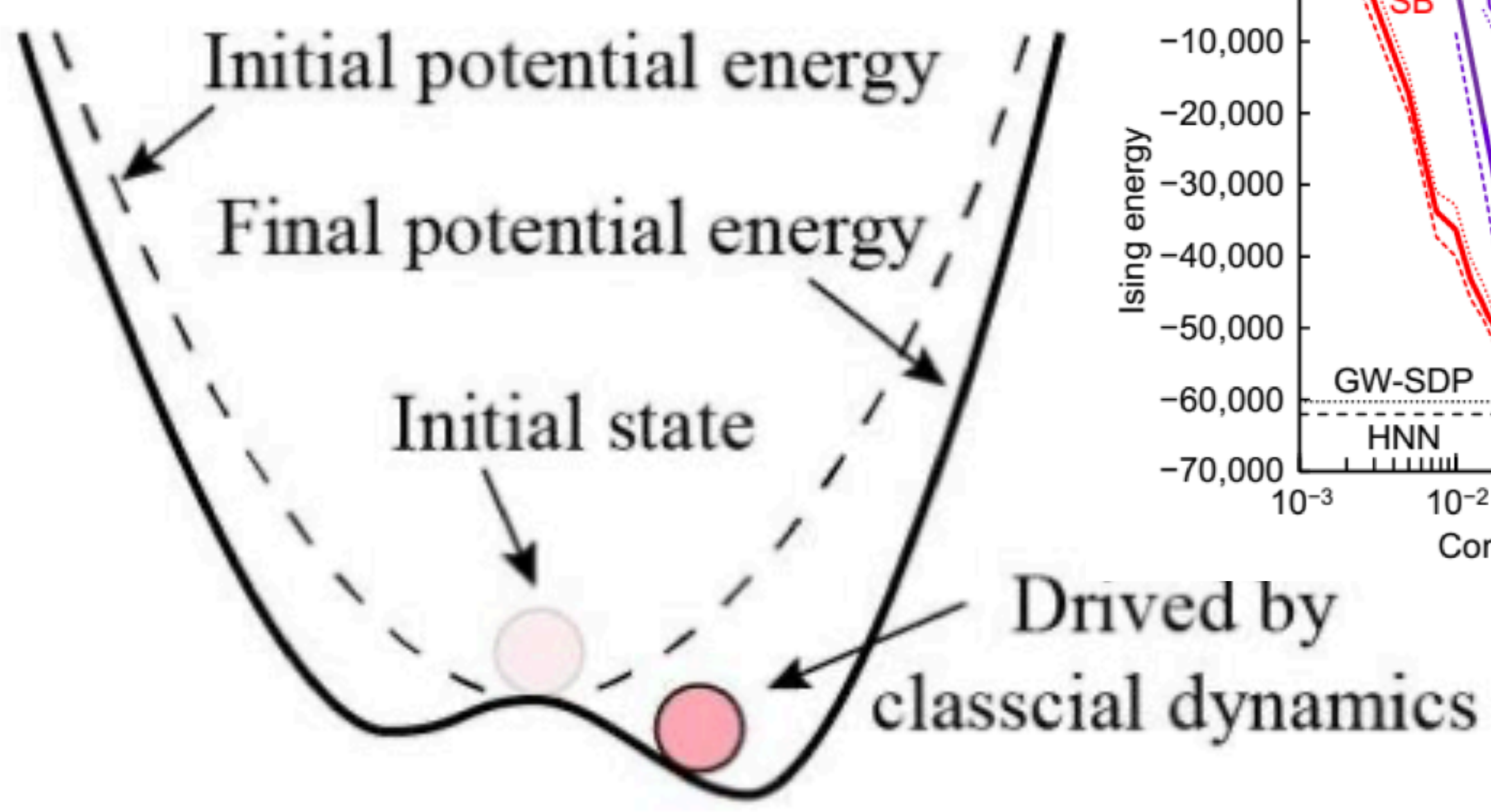


Linghui Wu

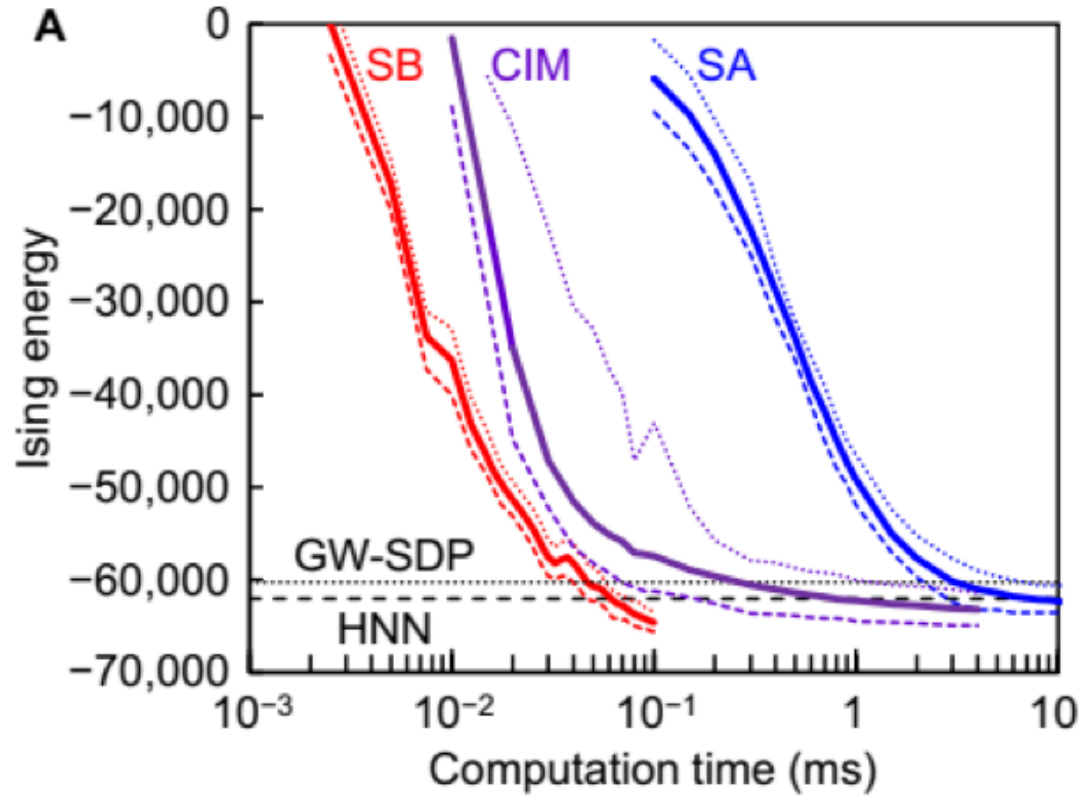
Data Quality for Tracking in MDC (BESIII)

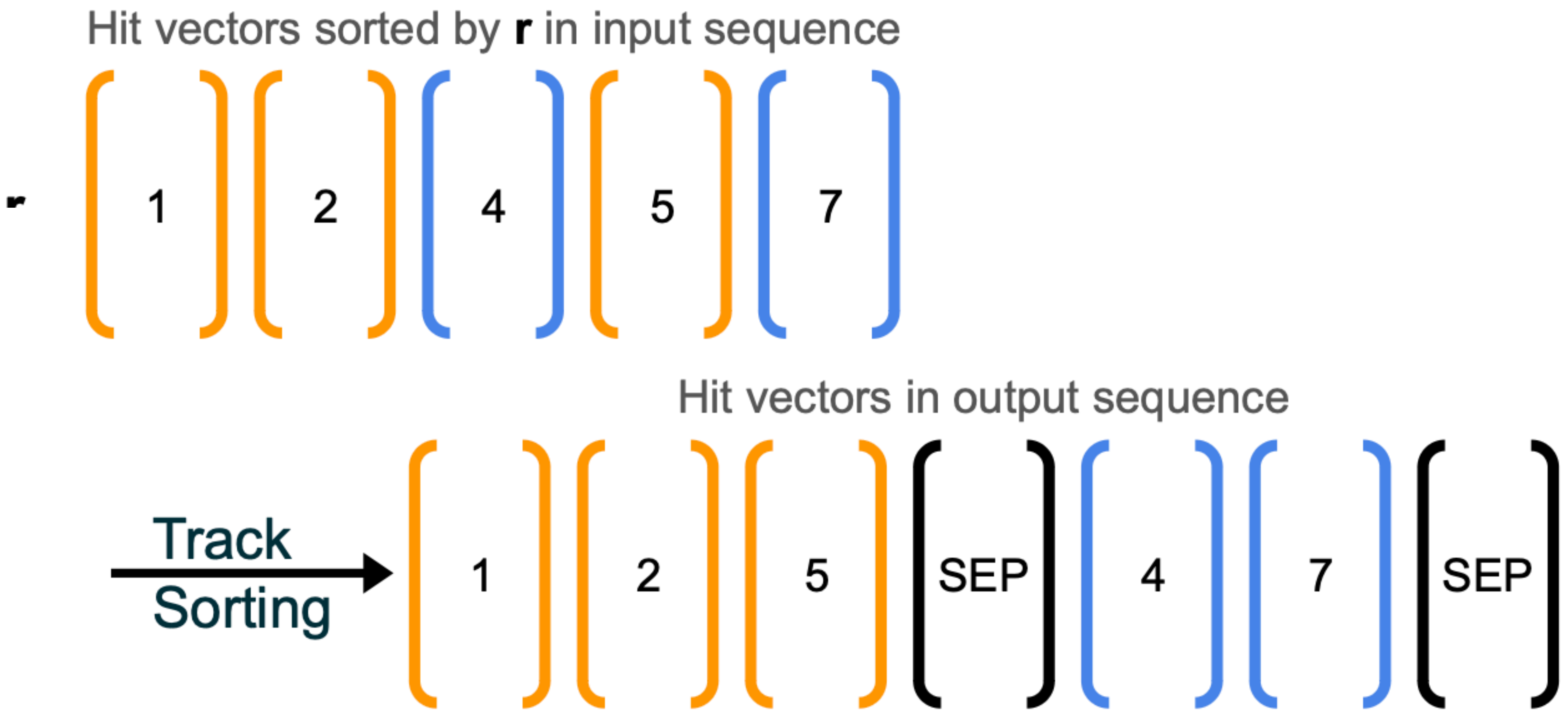
Future Tracking Technologies

Xiangyang Ju



Quantum inspired algorithm





Analogy between HEP and NLP

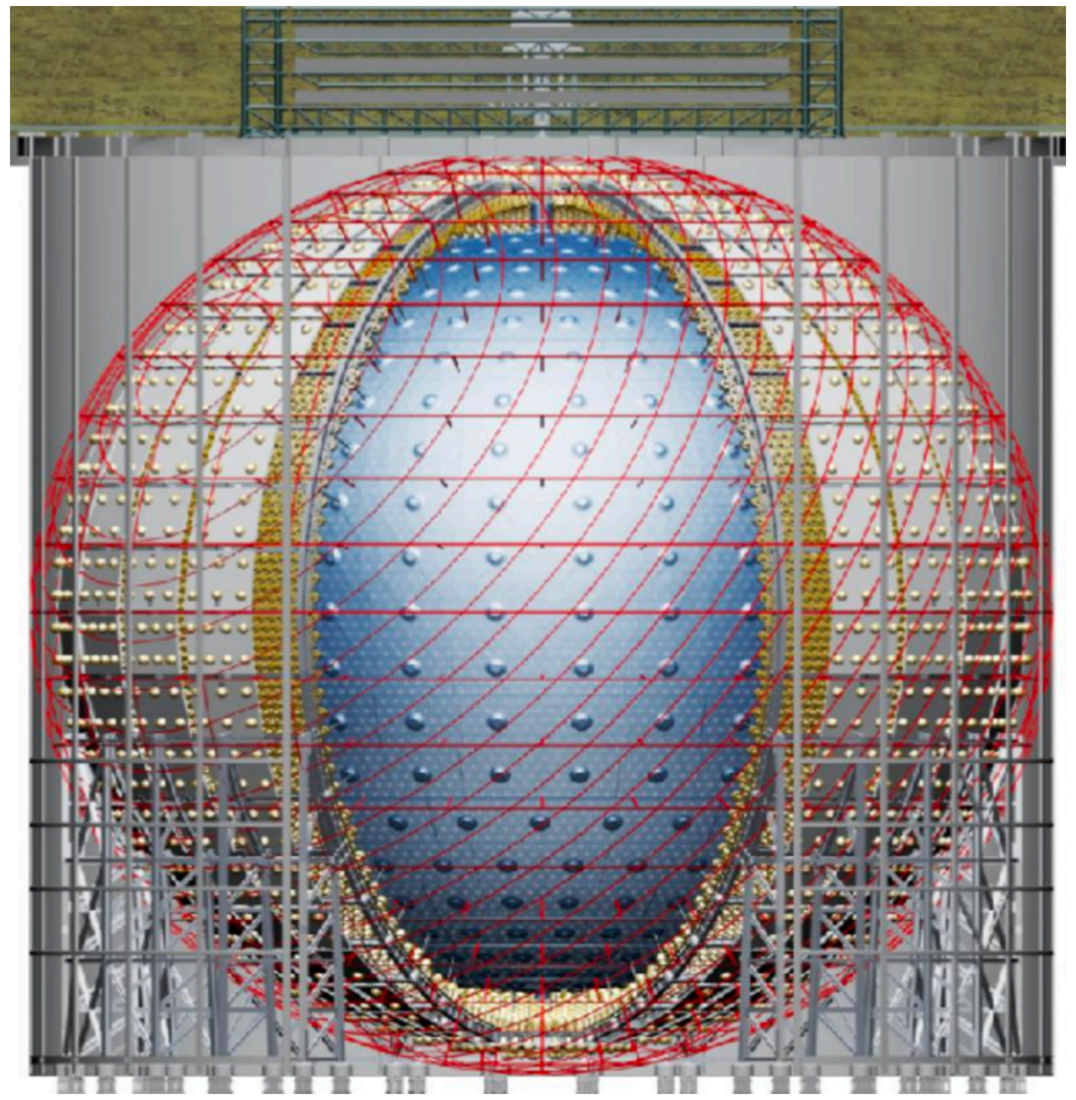
Detector elements	Words
All detector elements	Vocabulary
Trajectories or showers	Sentences
Collision Events	Paragraphs

Hideki Okawa

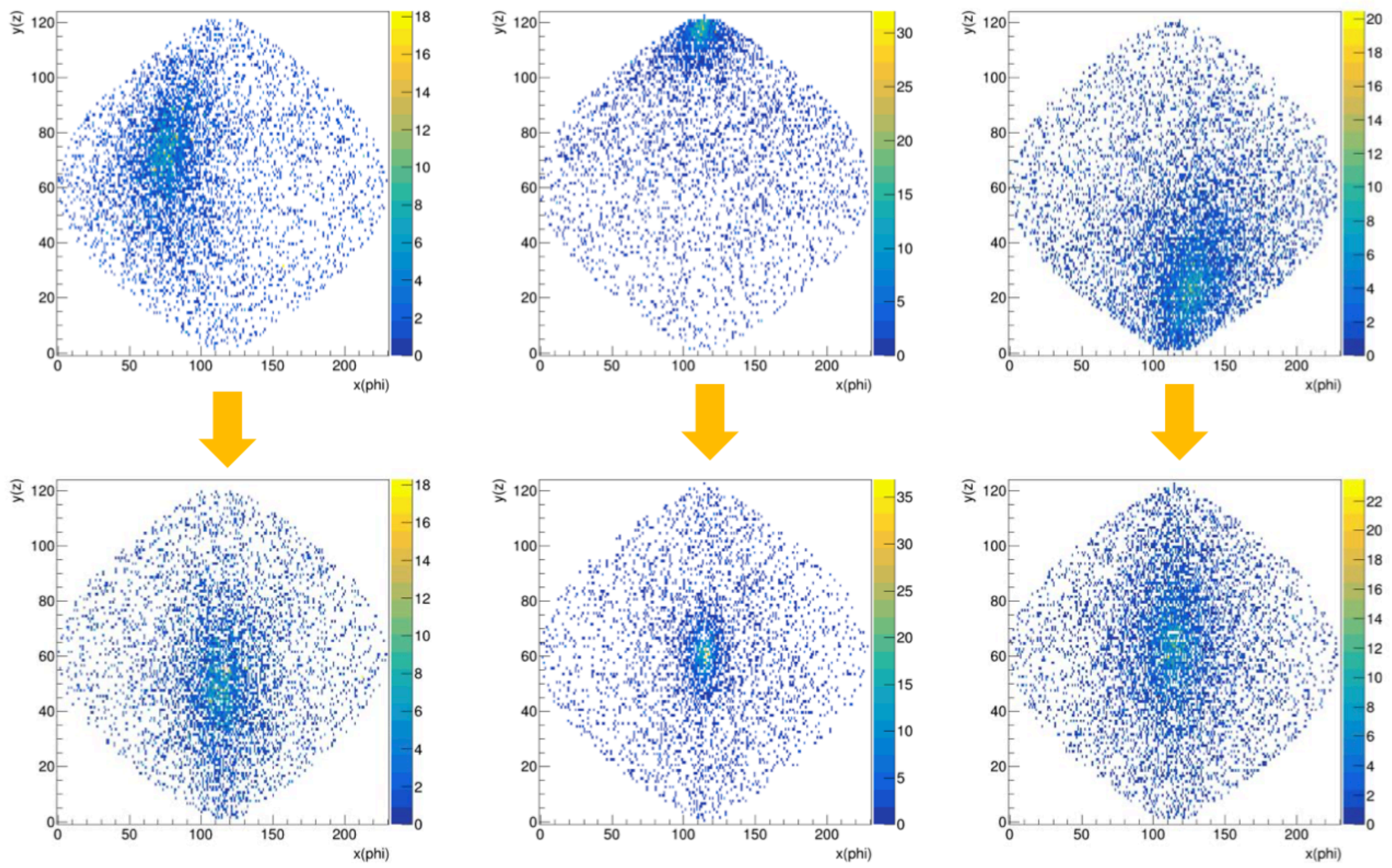
Language models

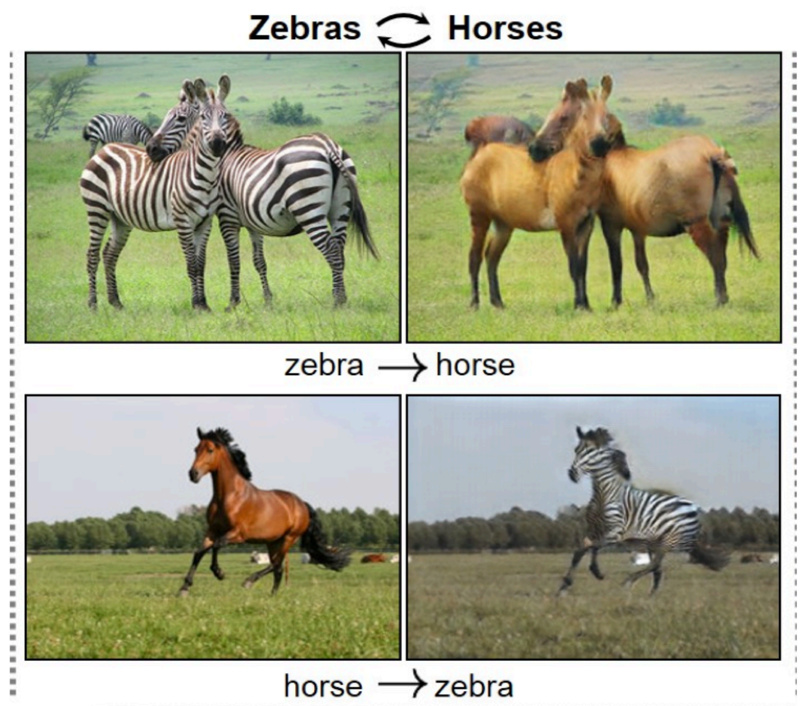
Reconstruction at Neutrinos

Not everything is the LHC. Unique challenges & Solutions



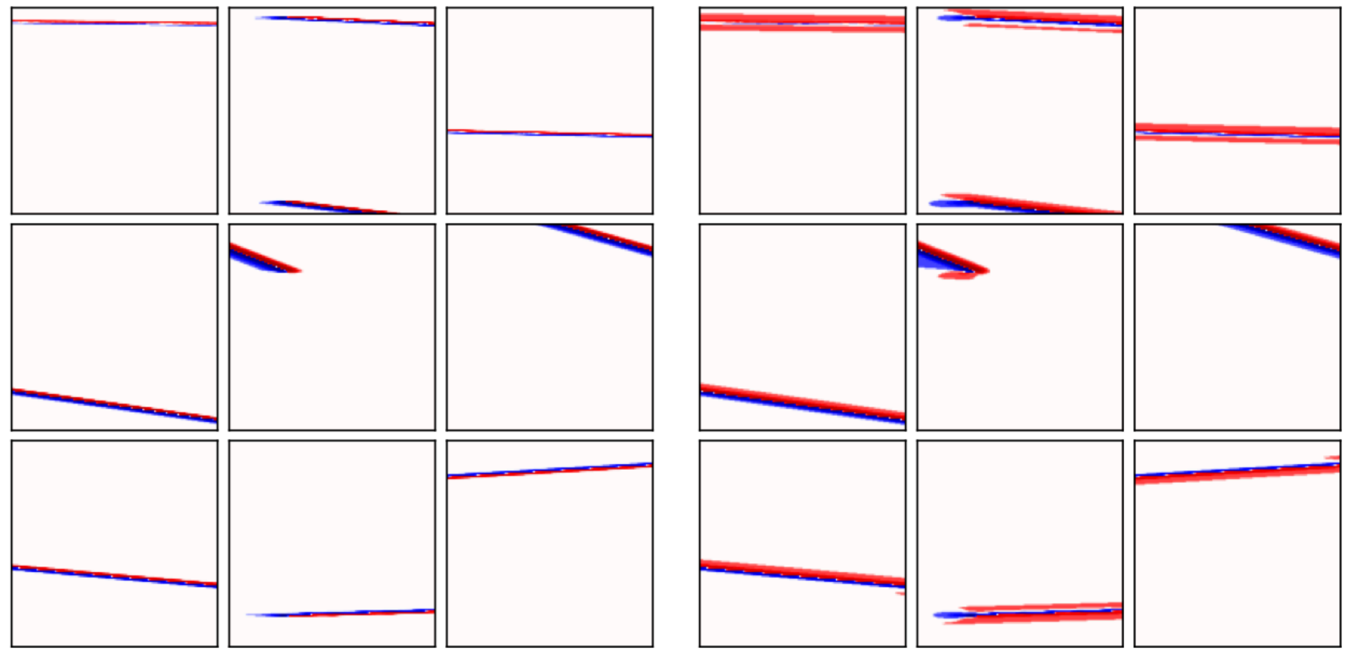
Wenxing Fang



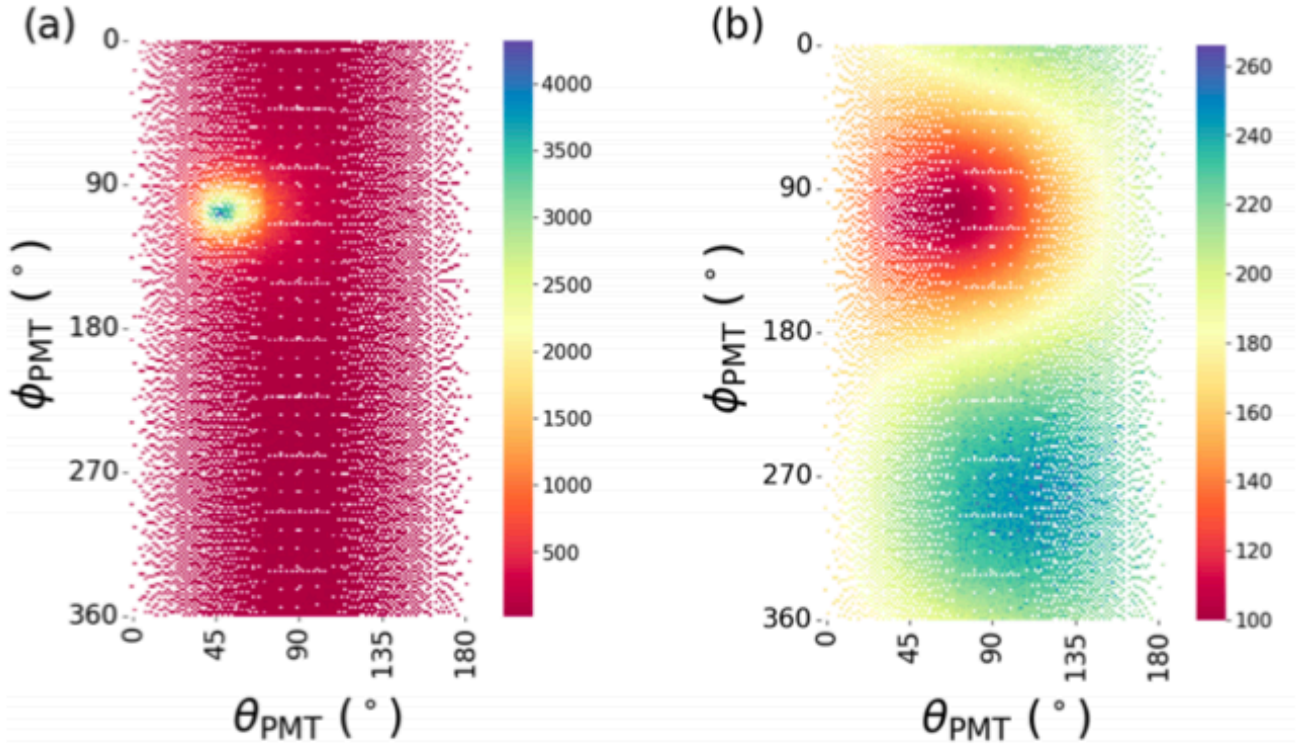


DUNE

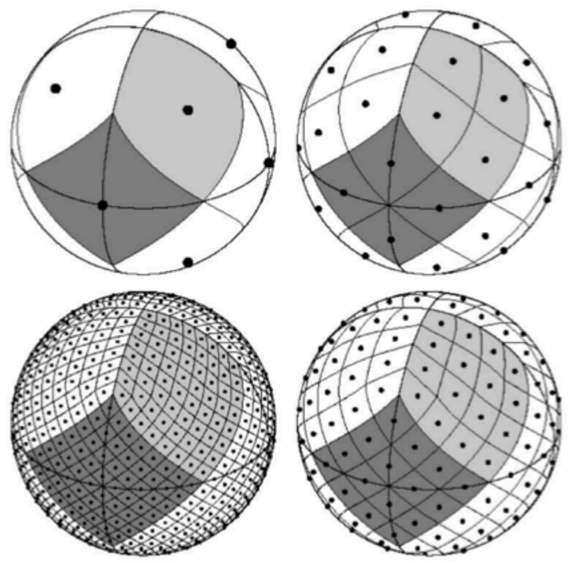
Dmitrii Torbunov



(a) **Domain A** (low-quality simulation) (b) **Domain B** (high-quality simulation)



Wing Yan Ma



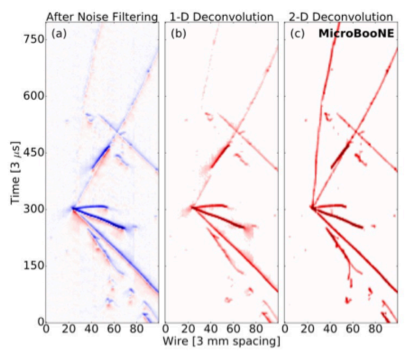
Jiaxi Liu JUNO

TPC simulation
noise filtering
signal processing

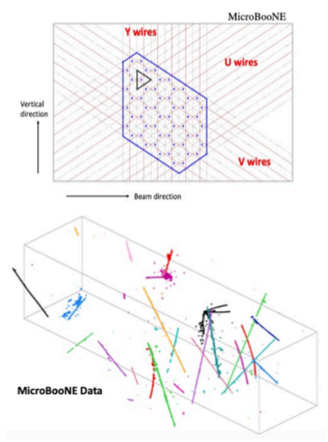
3D imaging
clustering
charge-light matching

3D trajectory & dQ/dx fitting
cosmic muon tagger

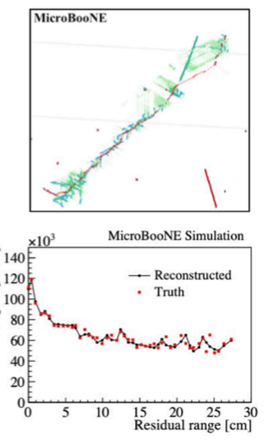
multi-track fitting
DL-3D vertexing
particle identification



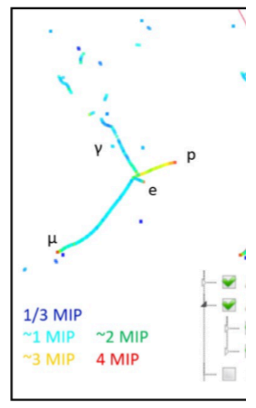
JINST 12 P08003 (2017)
JINST 13 P07006 (2018)
JINST 13 P07007 (2018)
JINST 16 P01036 (2020)



JINST 13 P05032 (2018)
JINST 16 P06043 (2021)

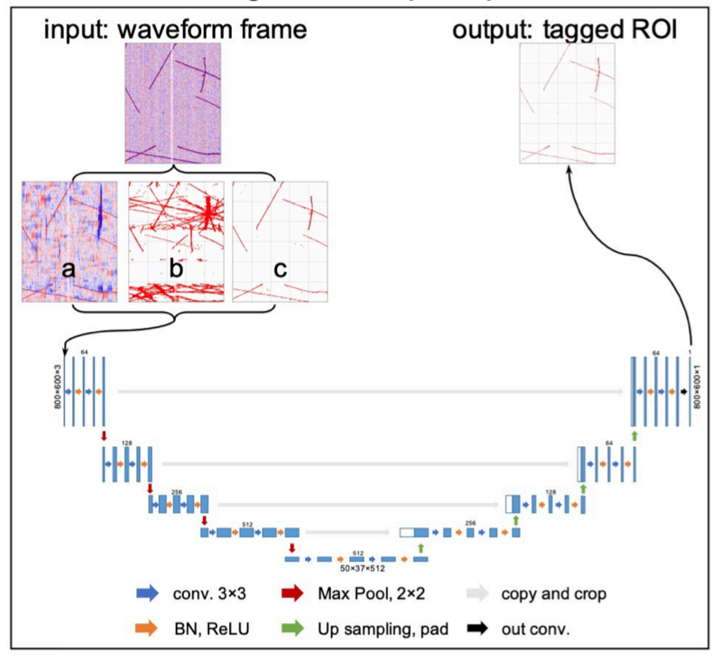


Phys. Rev. Applied 15, 064071 (2021)



JINST 17 P0

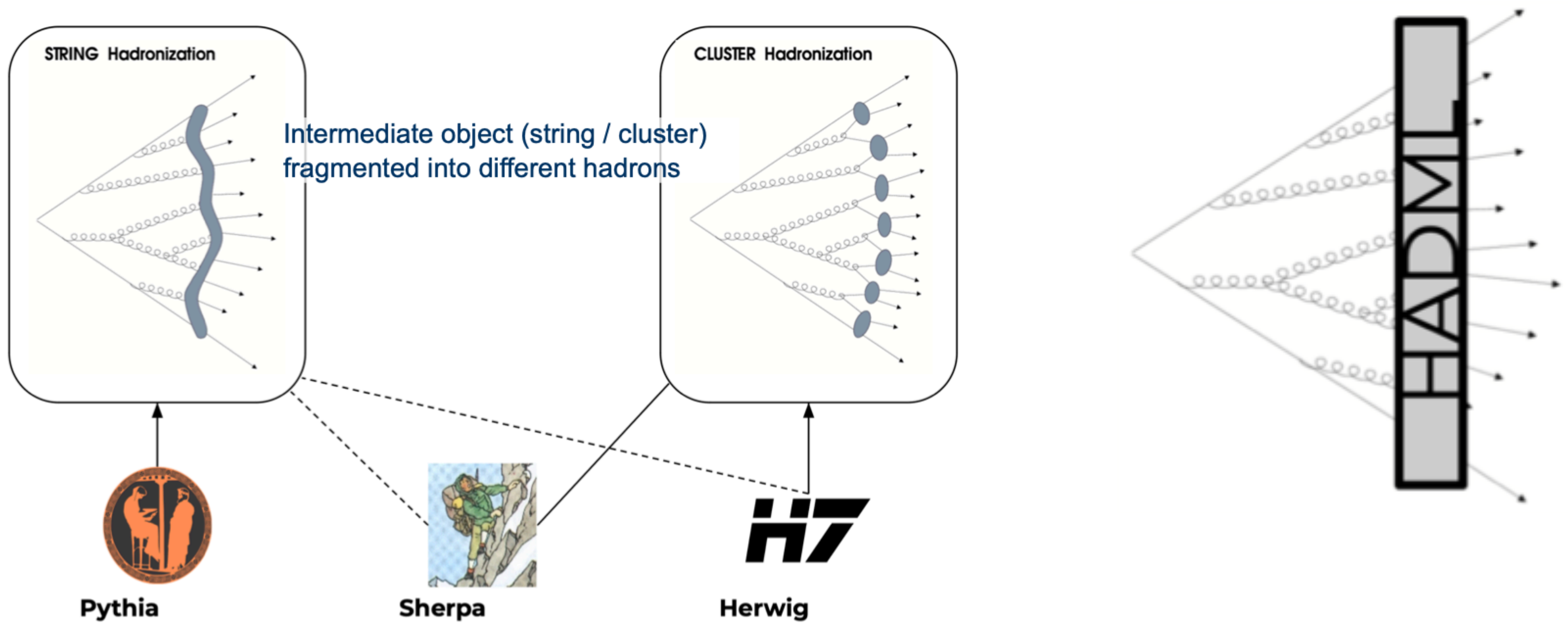
DNN ROI finding with multiple input channel

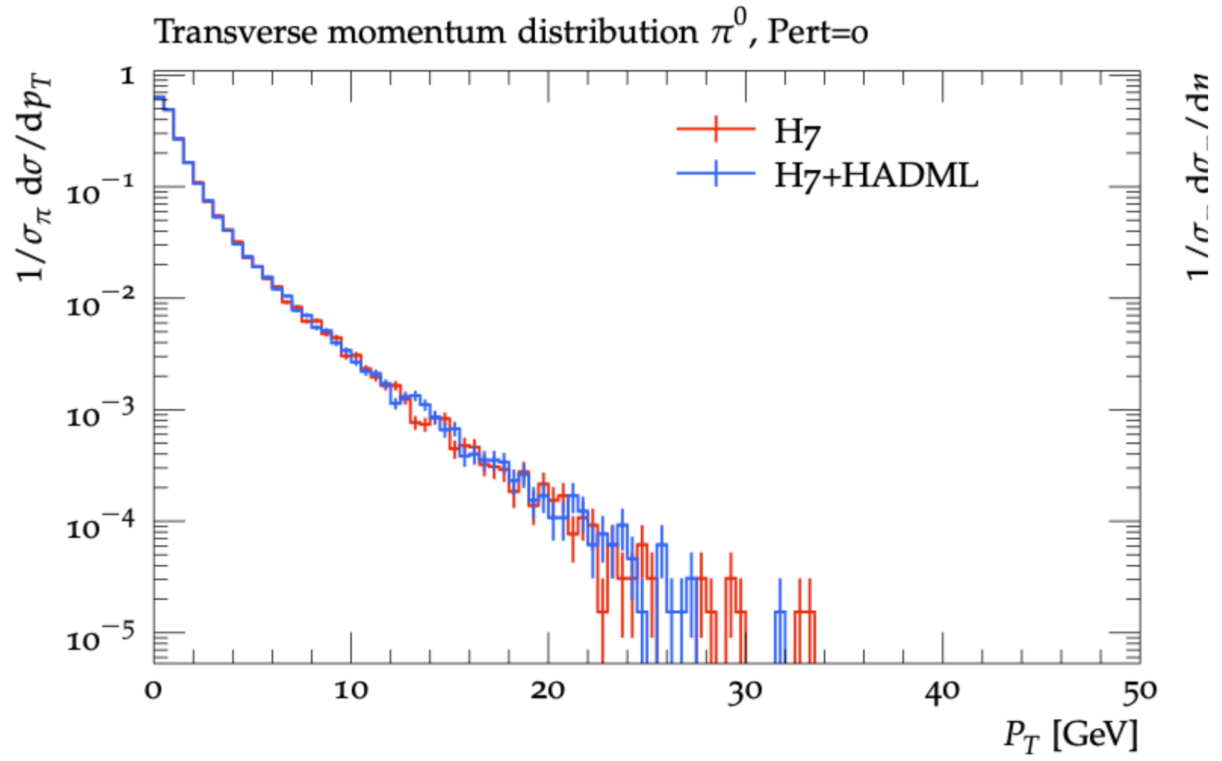


Haiwang Yu

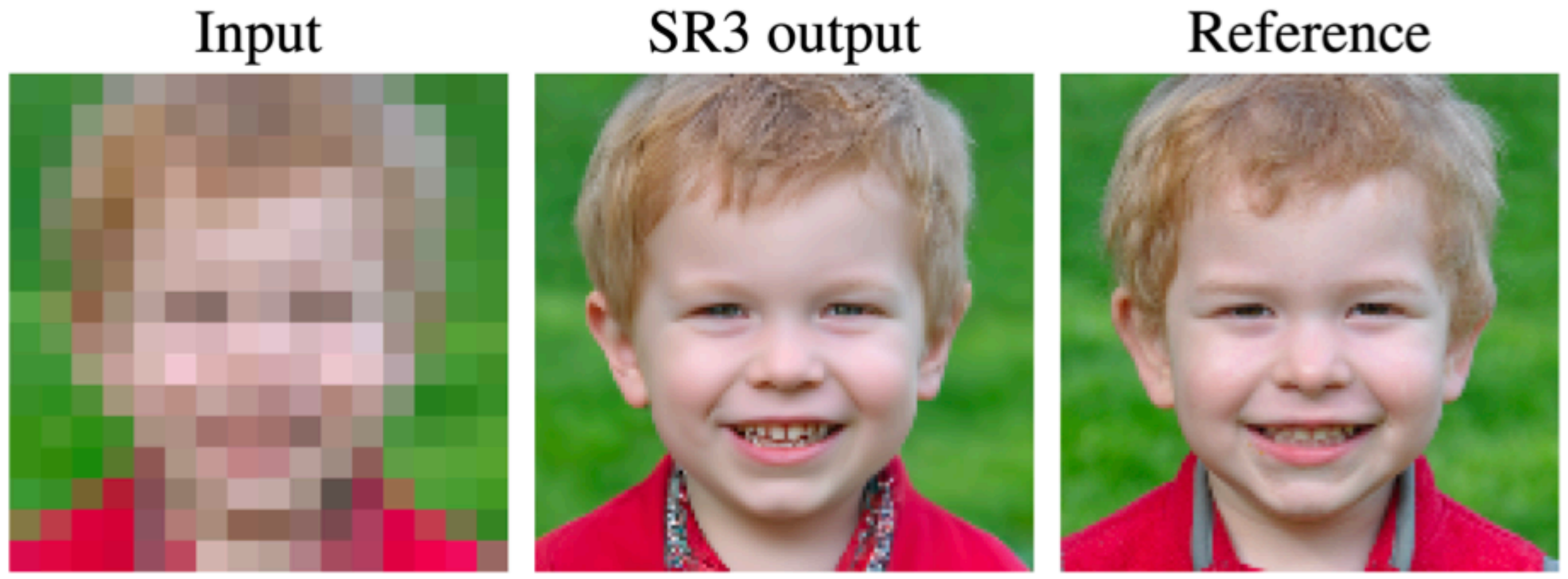
Generative Models

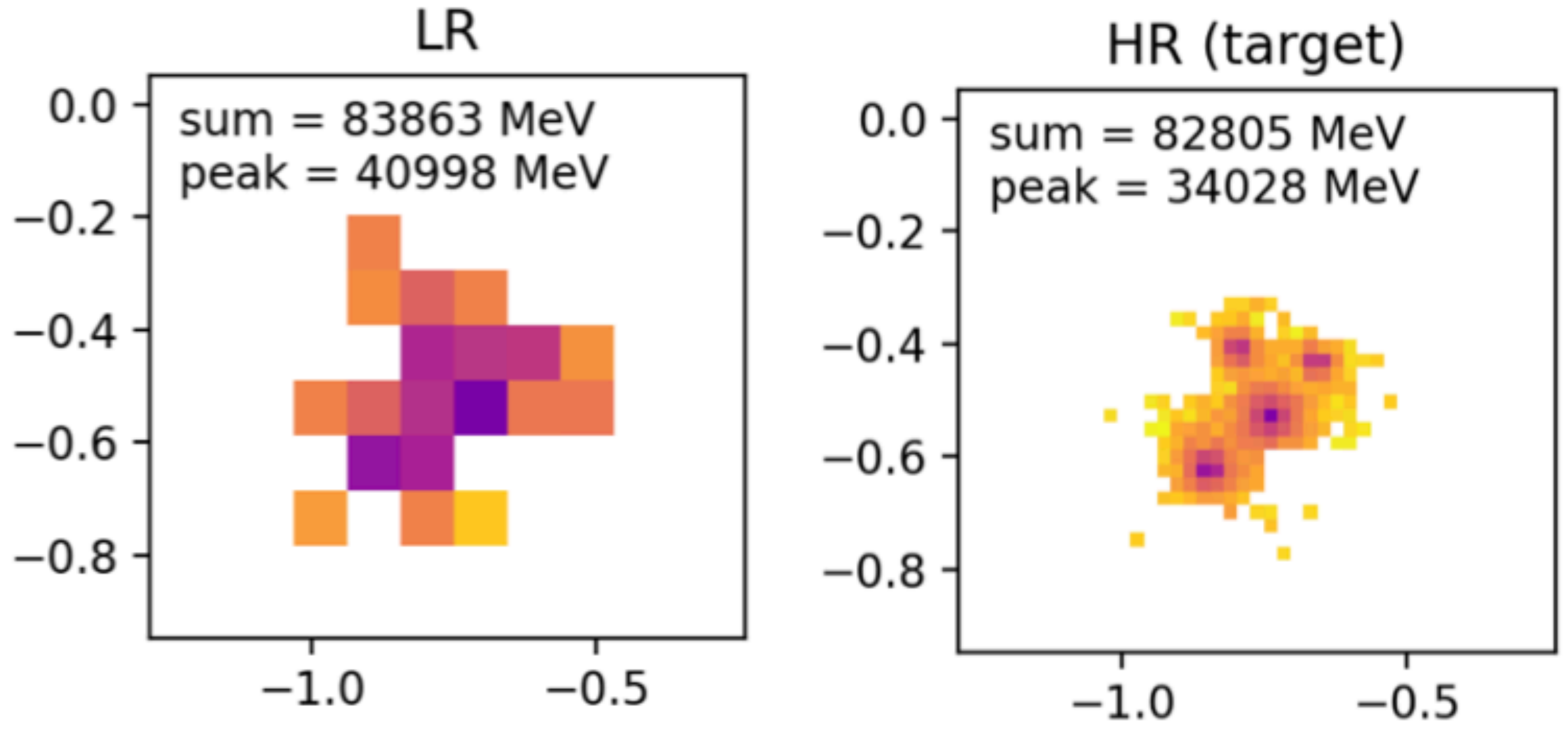
We're not done yet with Generative Models. Both on the Simulation as well as Reconstruction side.





Jay Chan



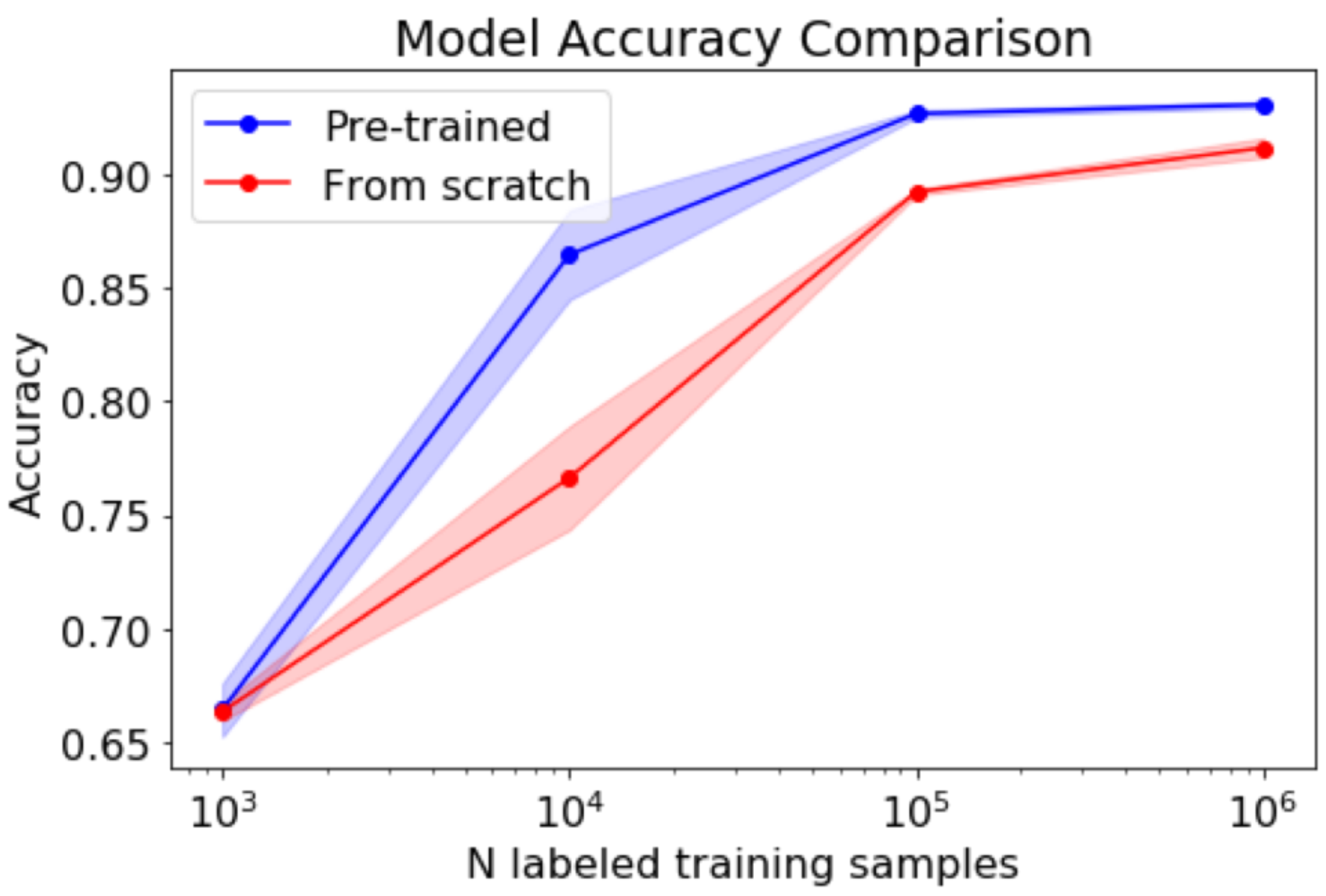


Nilotpal Kakati

Foundation Models

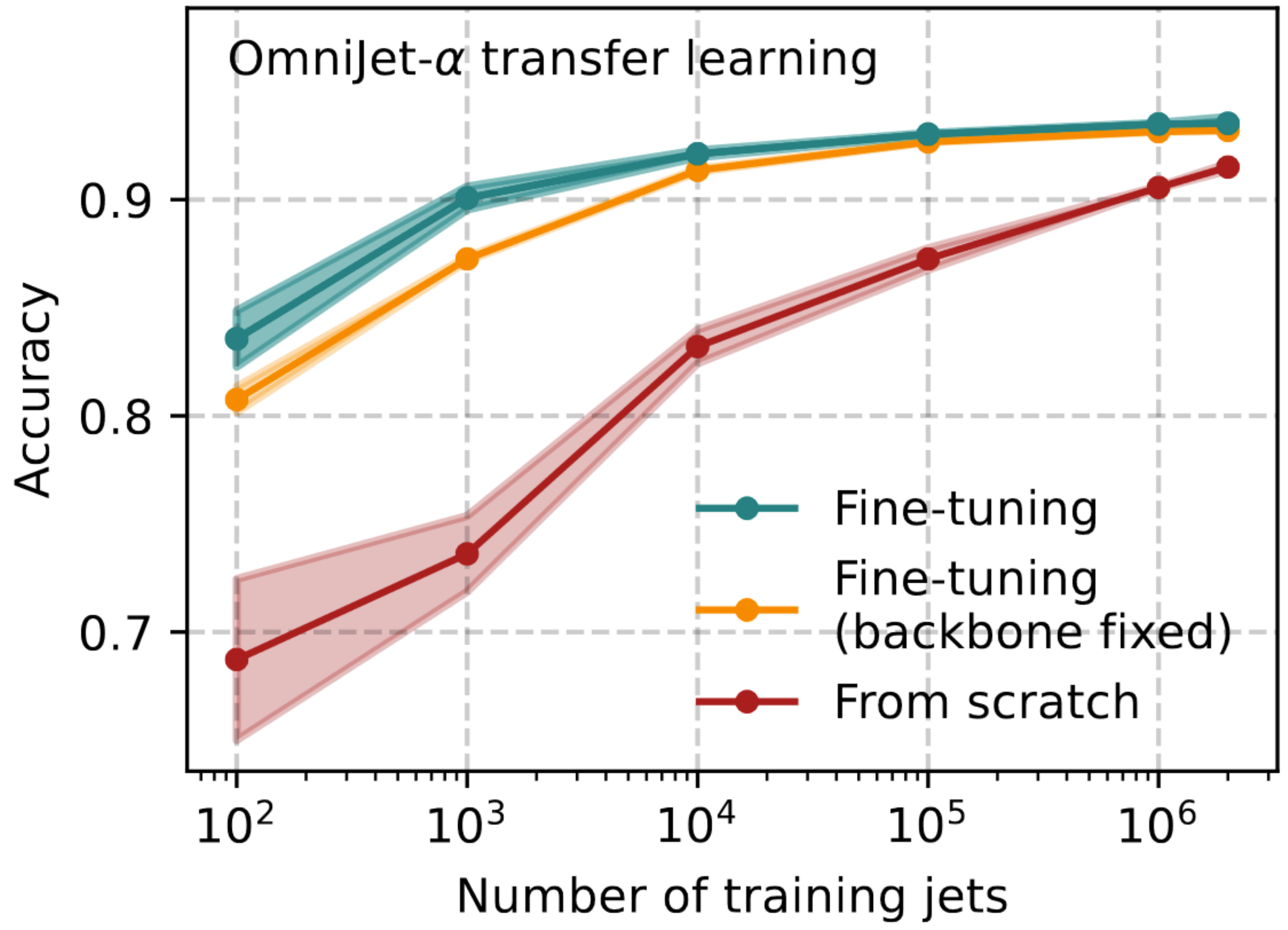
Foundation Models are about data efficiency & finetuning.
Get used to seeing this plot more often in the future

Self-supervised



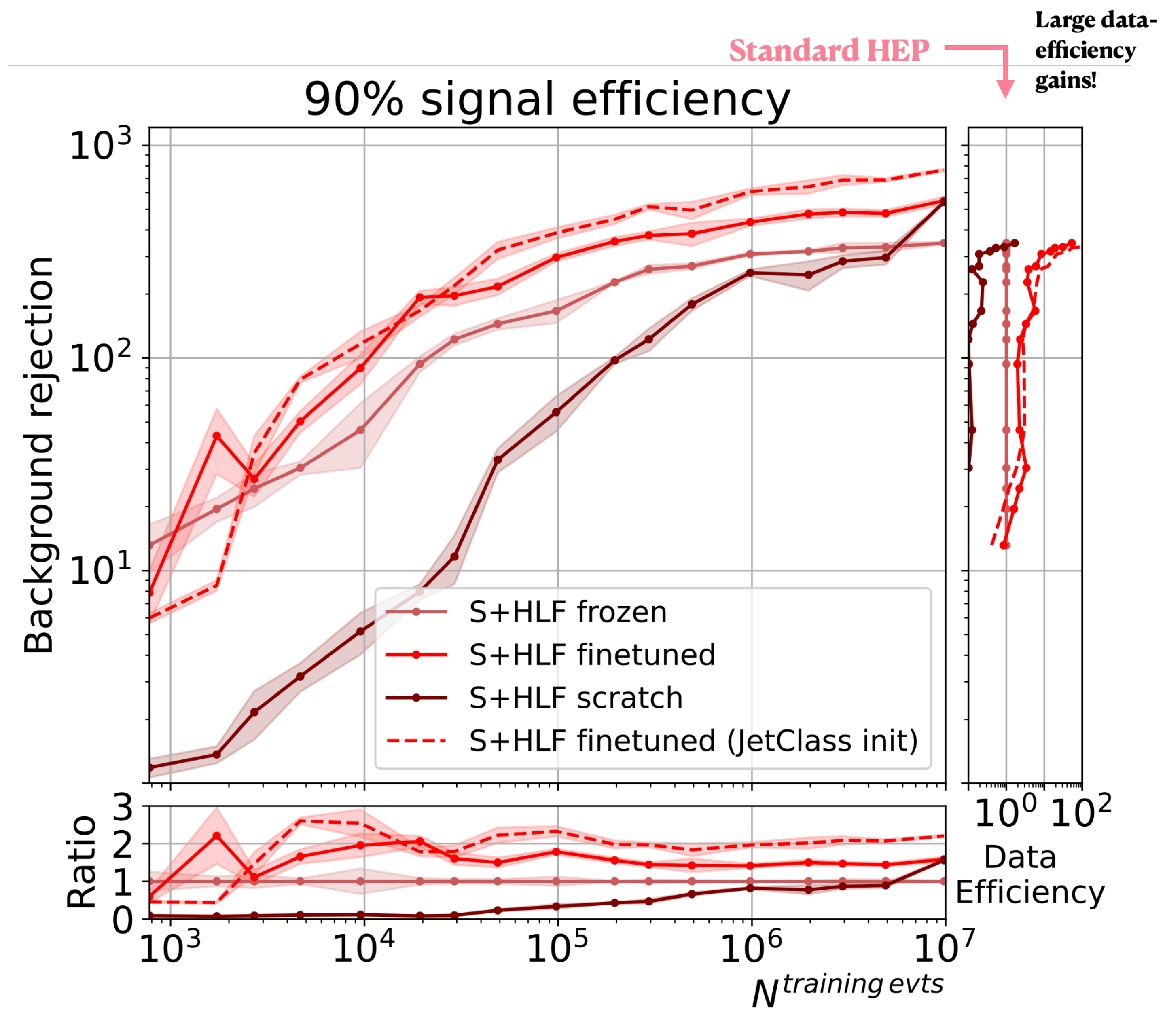
Zihan Zhao

Self-supervised



Anna Hallin

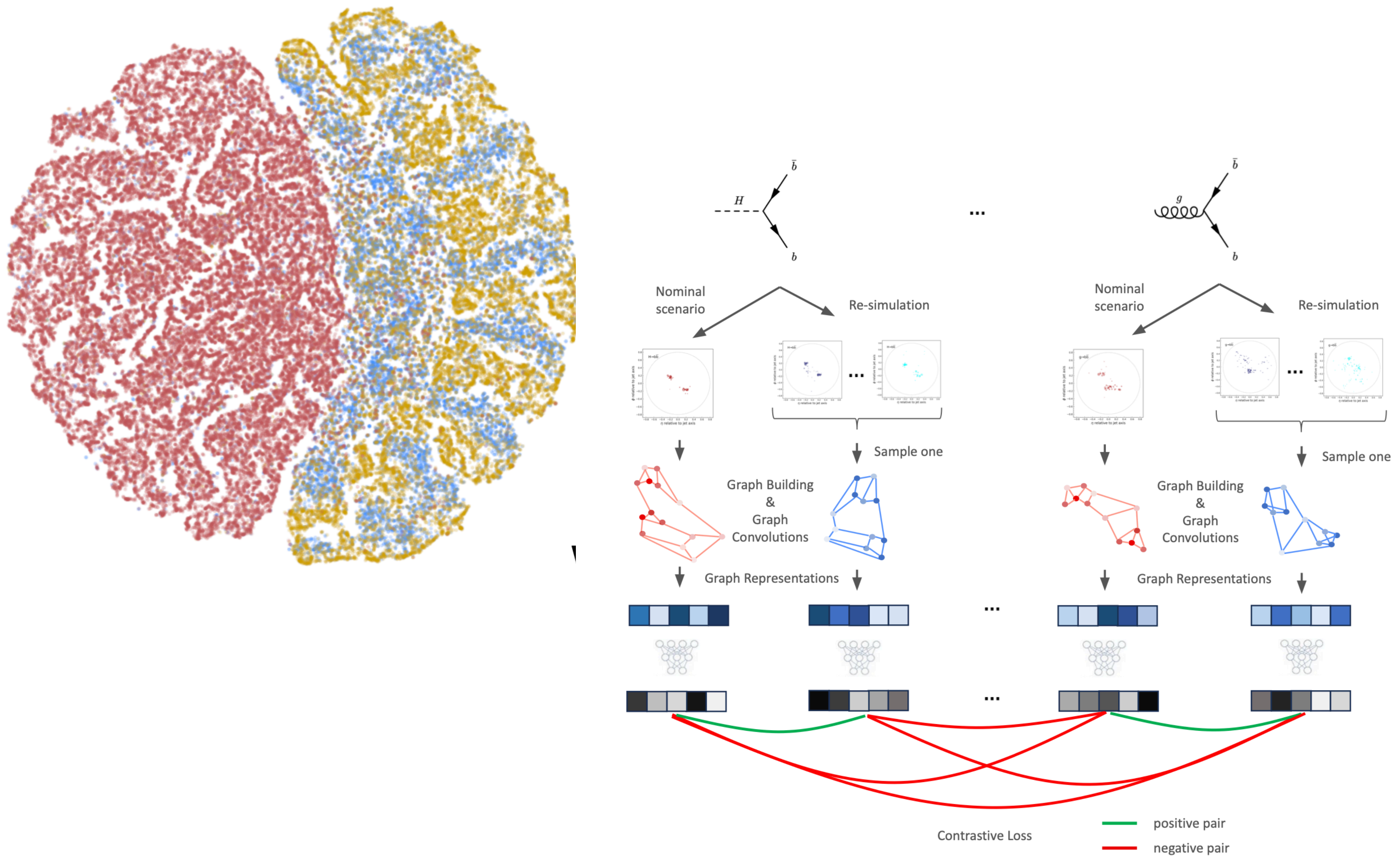
Supervised



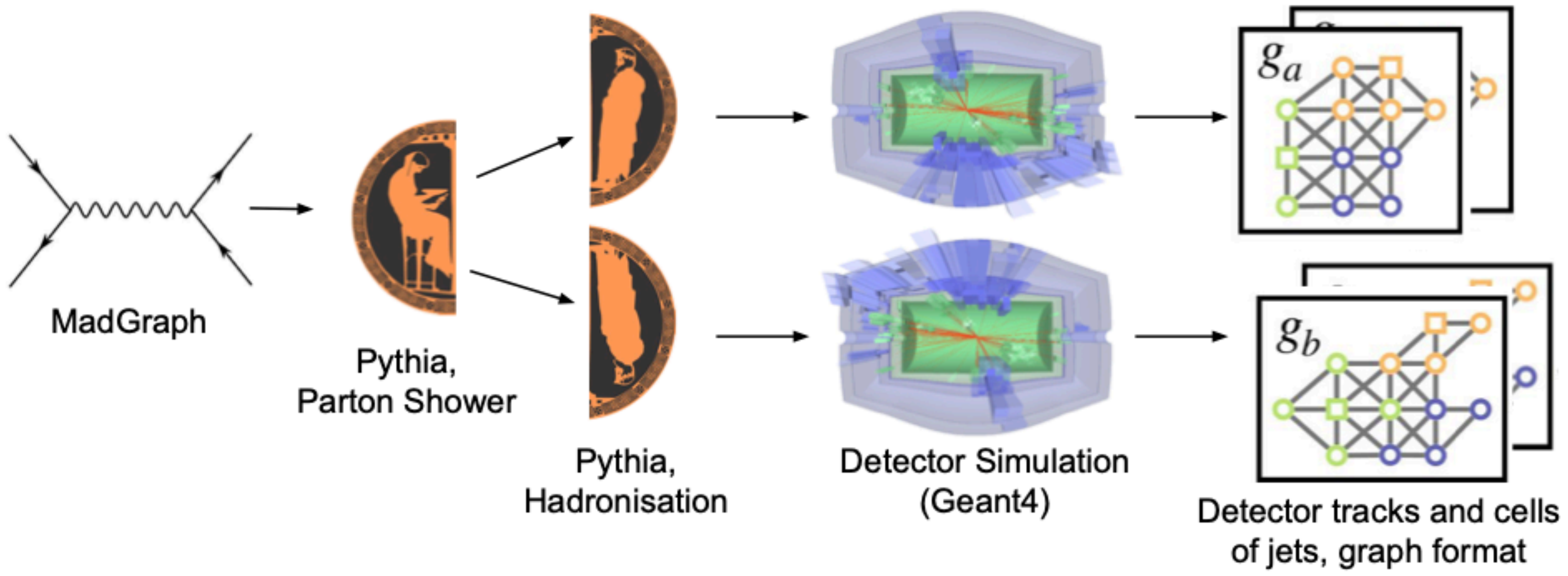
Matthias Vigl

Foundation Models II

Representation Learning & Joint Embedding: Use Simulator to reflect what we actually value & what we consider noise



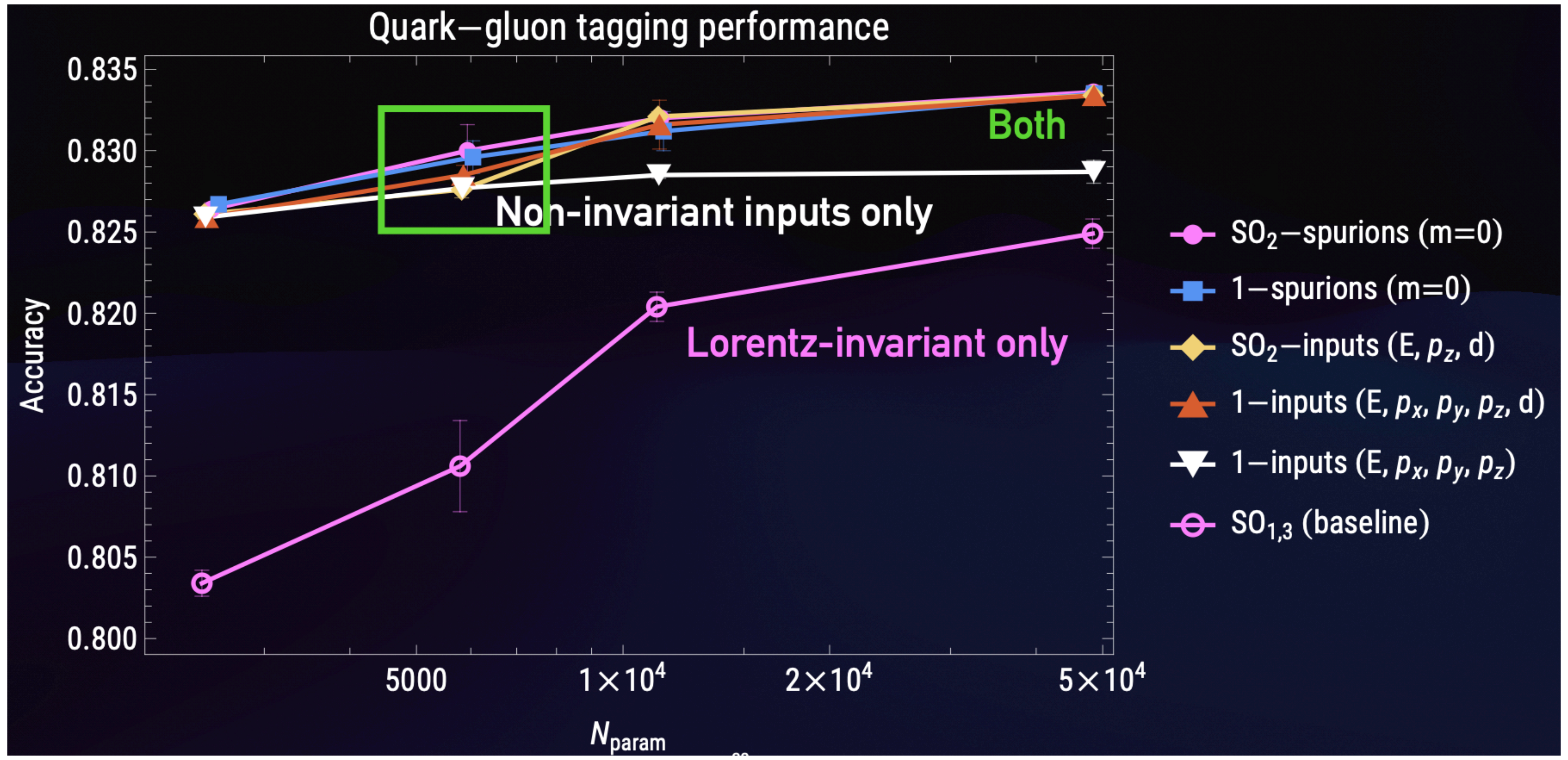
Phil Harris



Patrick Rieck

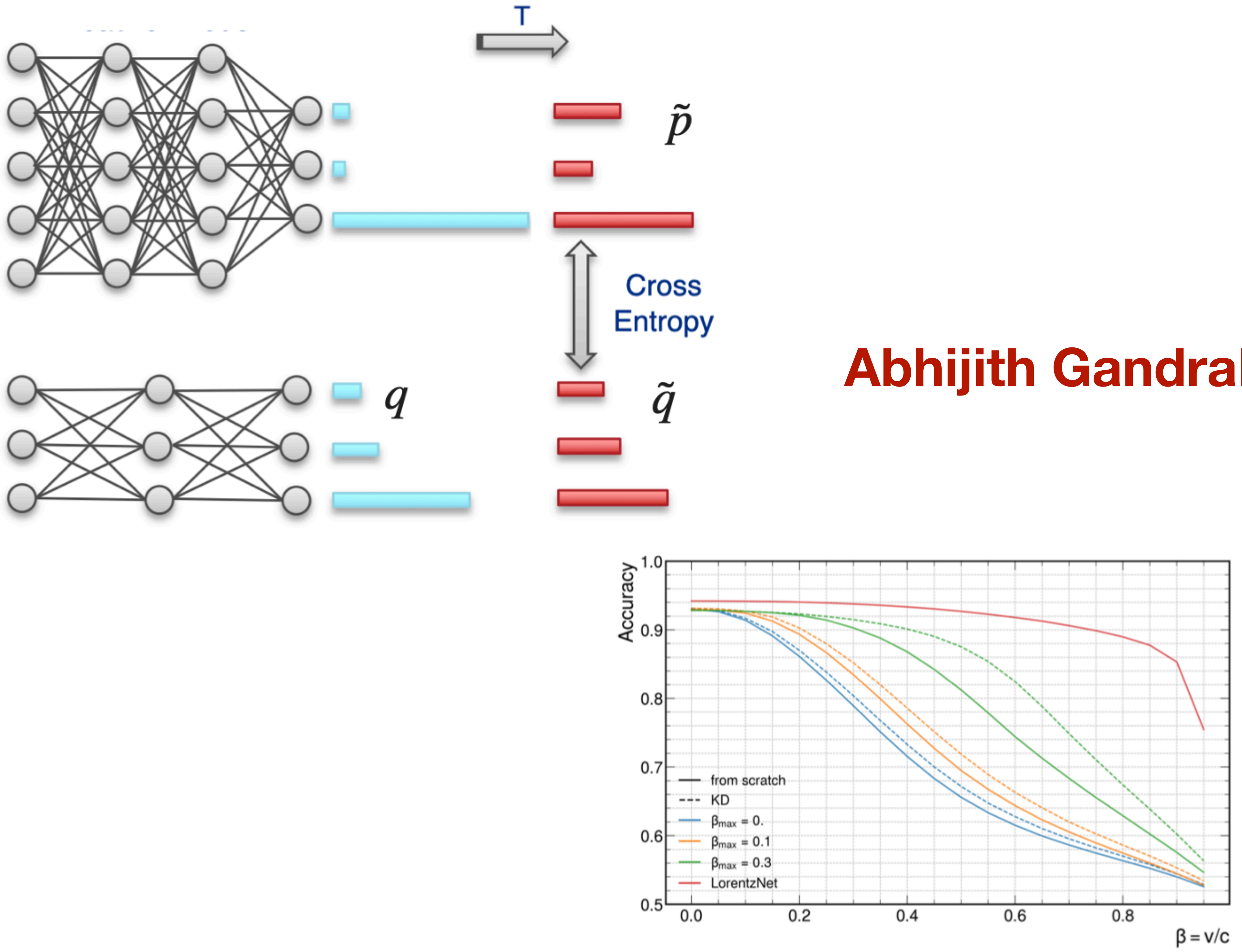
Robustness & Inductive Bias

Inductive Bias is still important, but can also slow down complicate neural networks. Can we resolve this tension?



Alexander Bogatskiy

Inductive Bias Impact



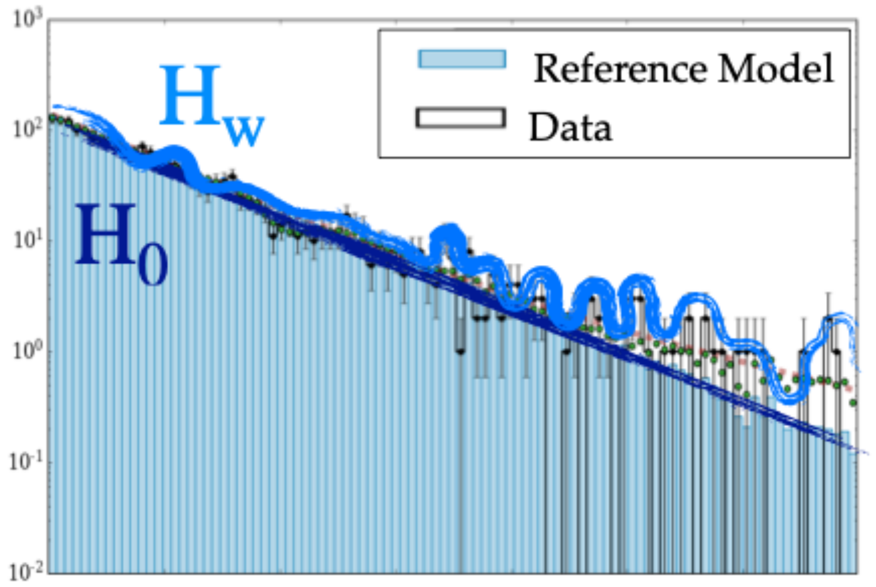
Abhijith Gandrakota

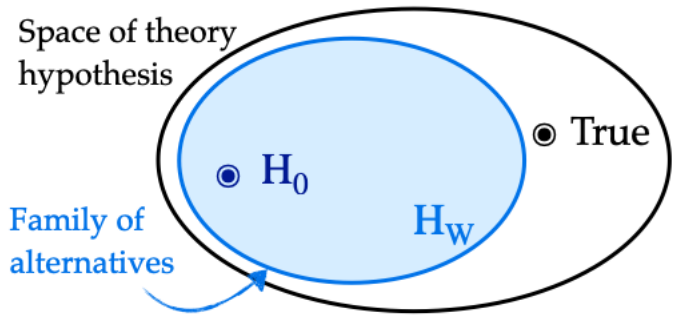
Knowledge Distillation

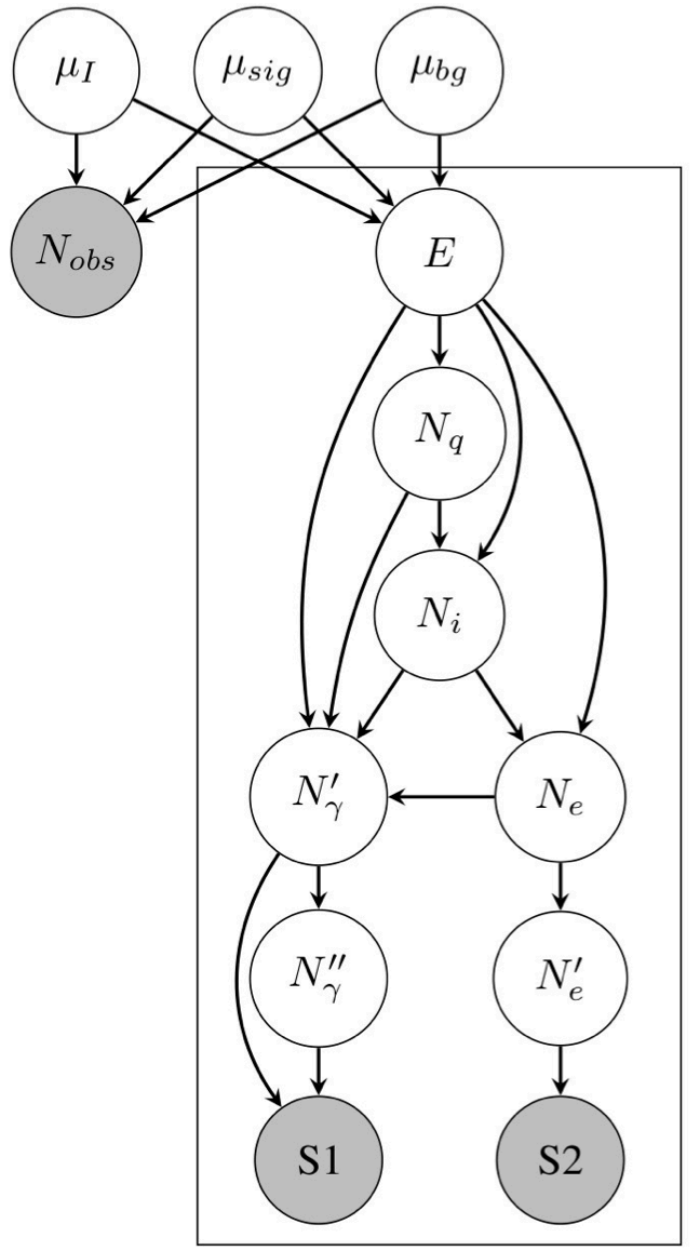
Advanced Inference Methods

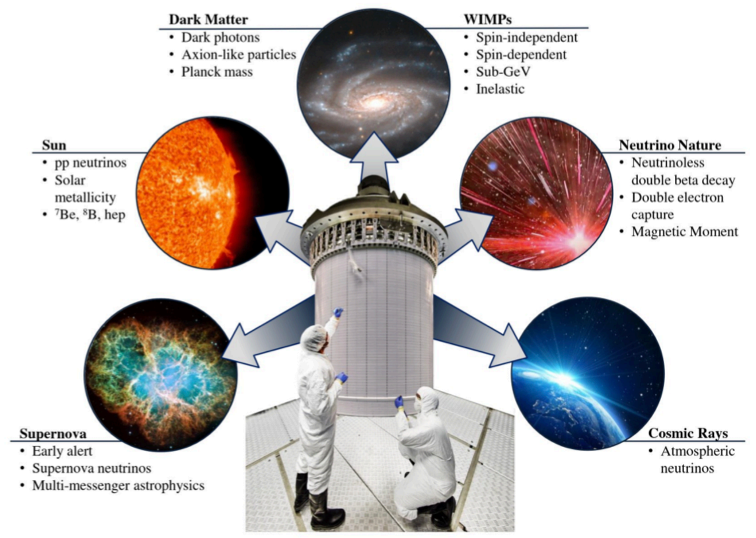
Uptake of new Techniques. Probabilistic & Differentiable Programming, Likelihood Ratio Learning, Bayesian Opt. ...

Gaia Grosso



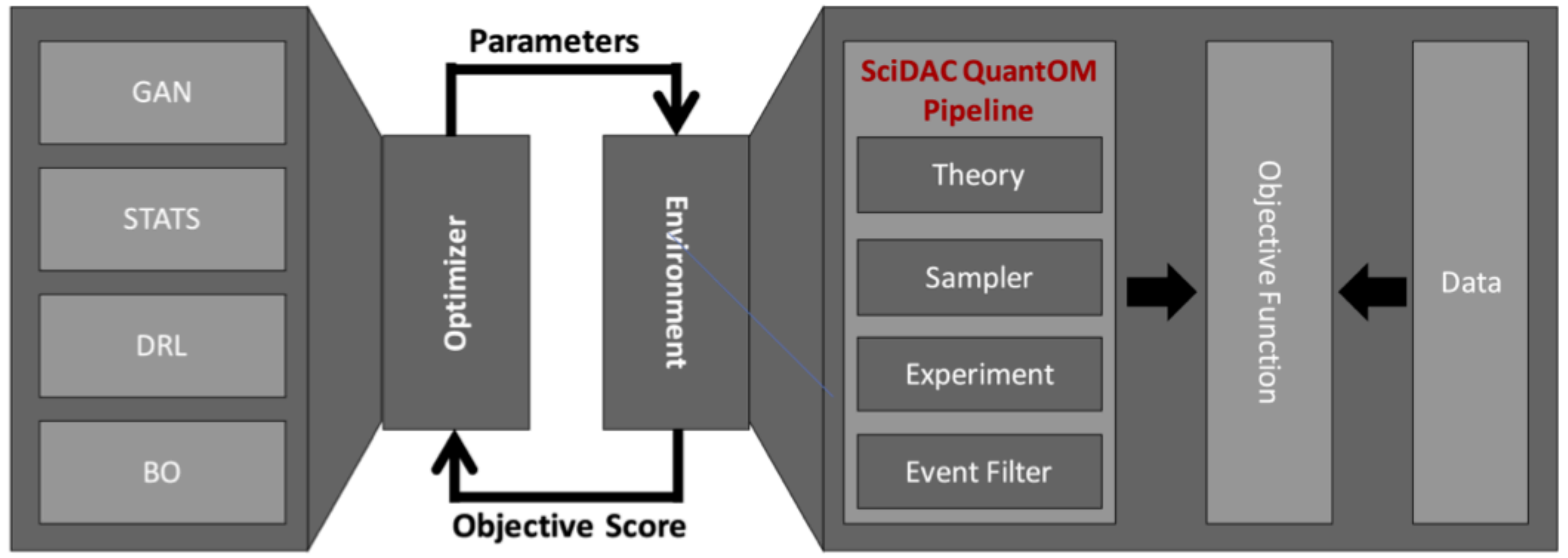


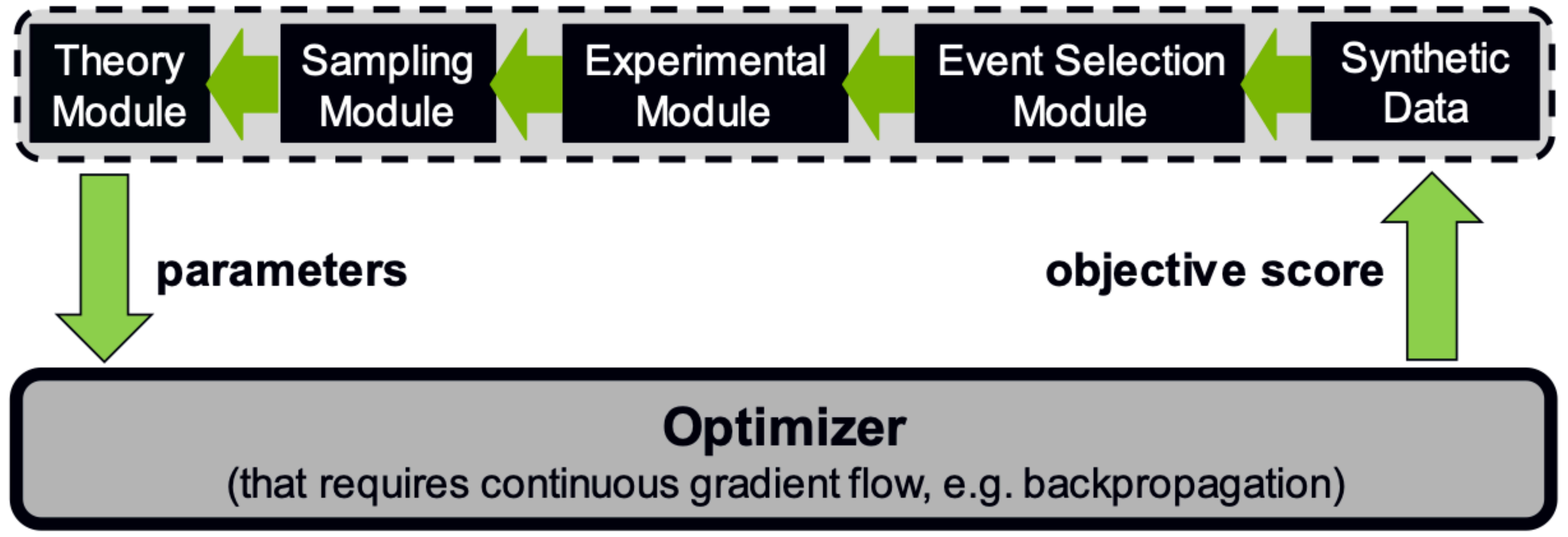


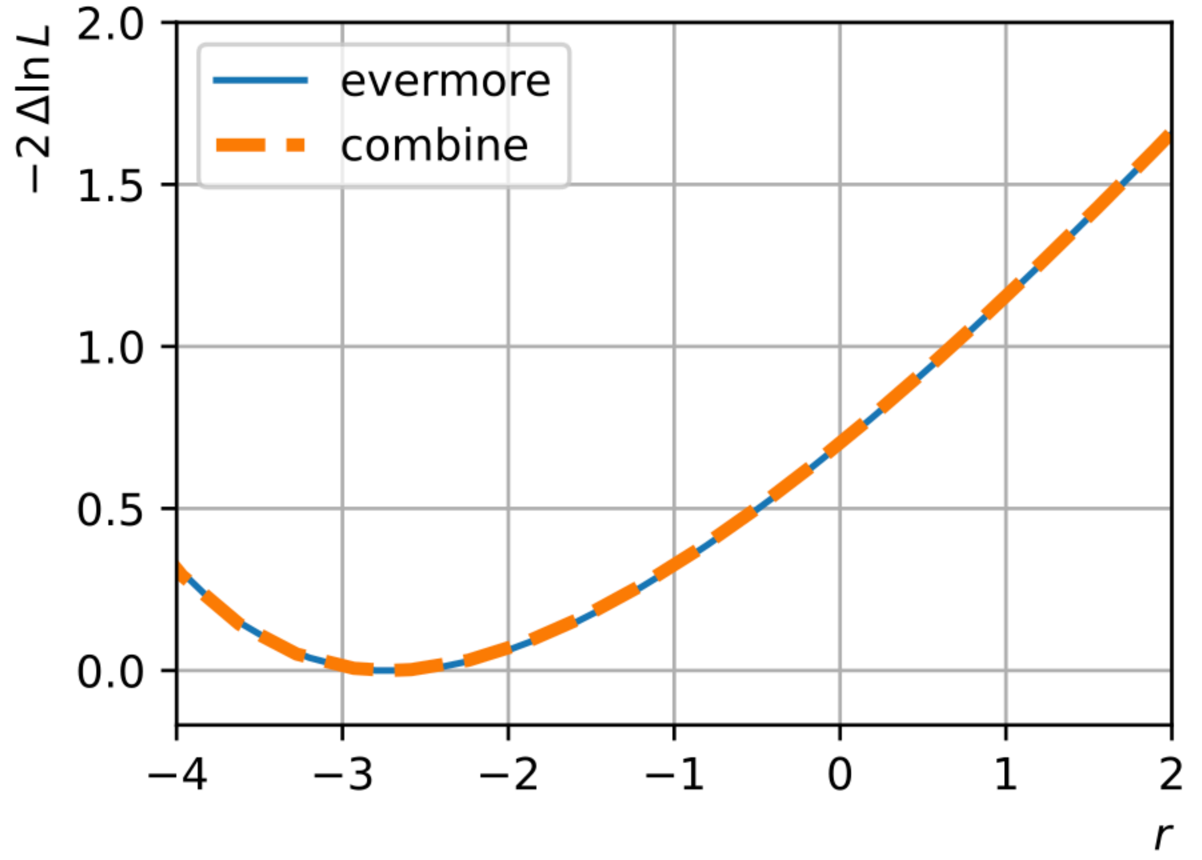


Juehang Qin

Daniel Lersch



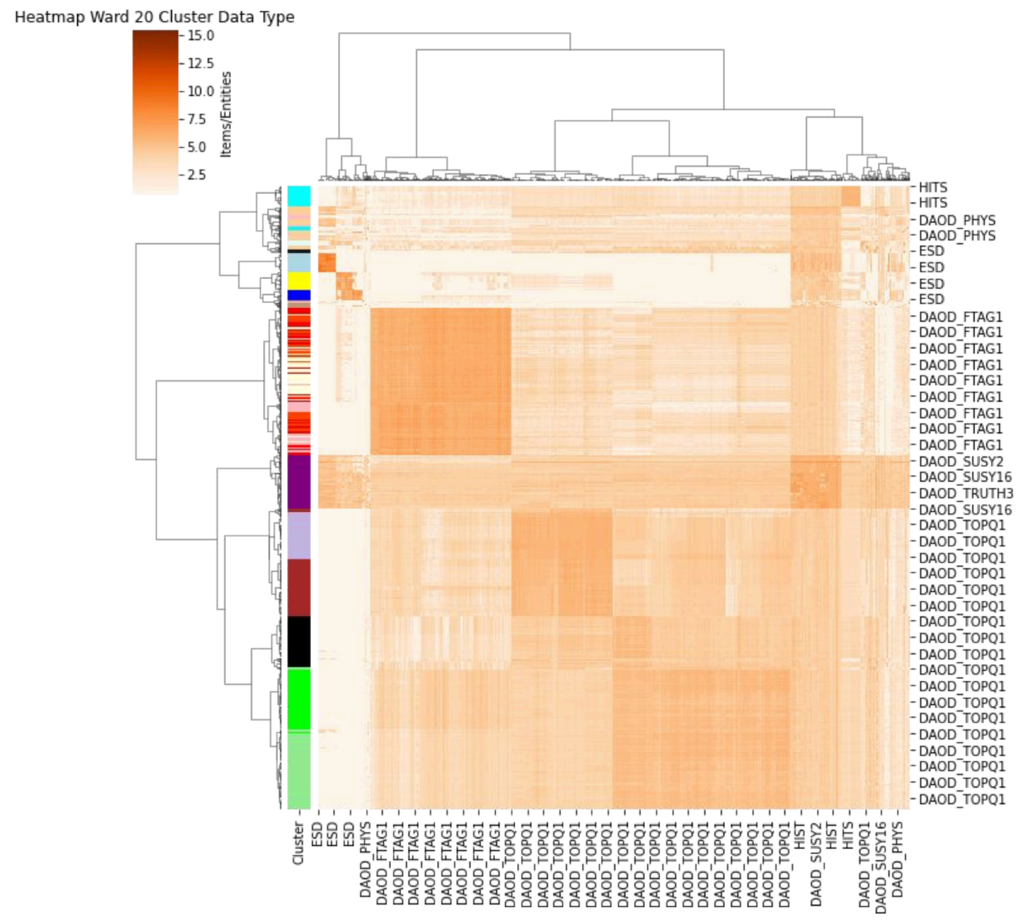




Peter Fackeldey

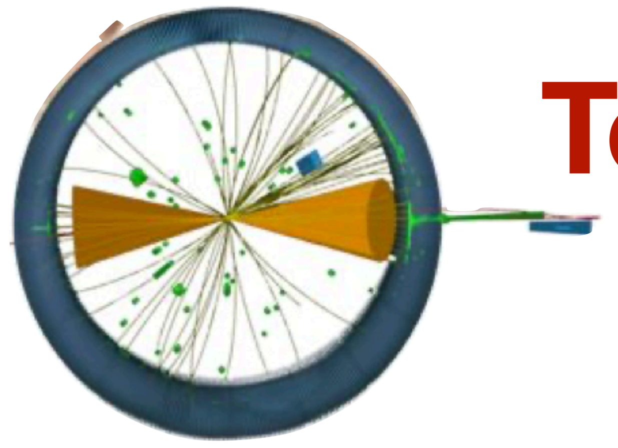
Tooling, Tooling, Tooling

Data Management

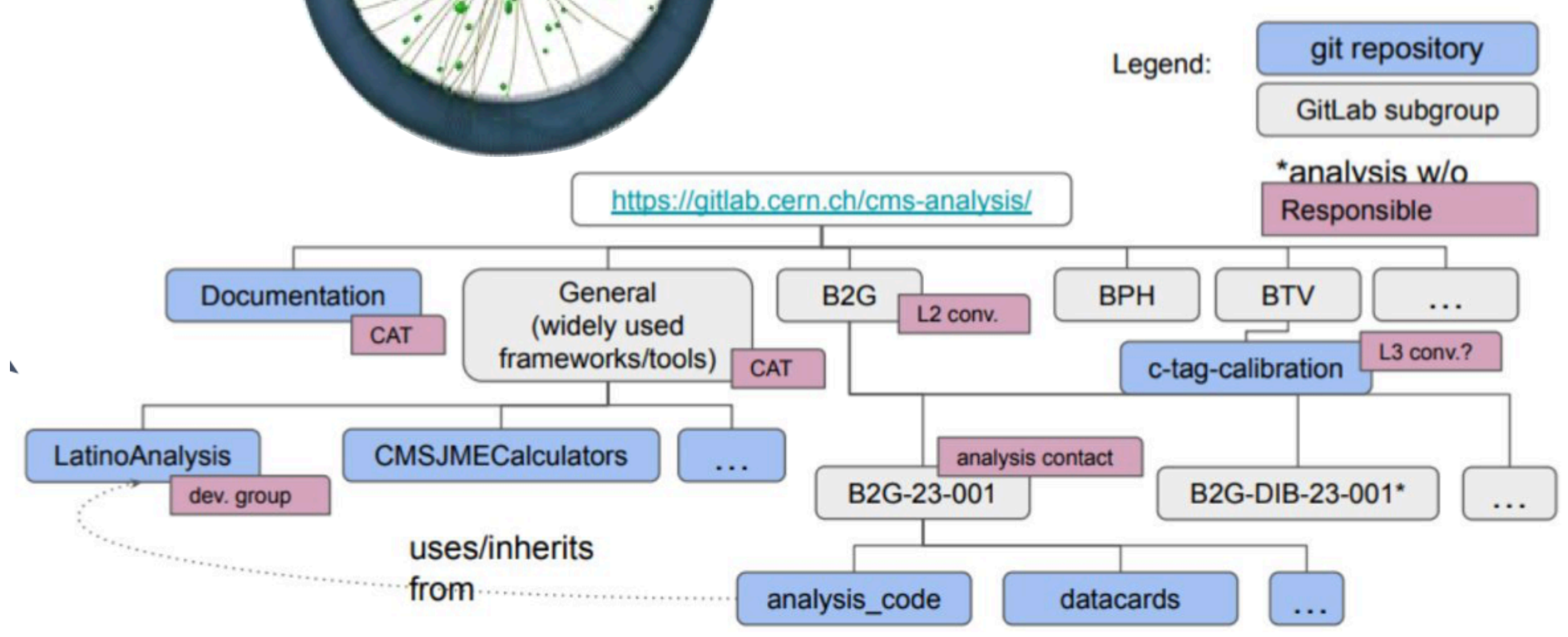


Qiulan Huang

People Management

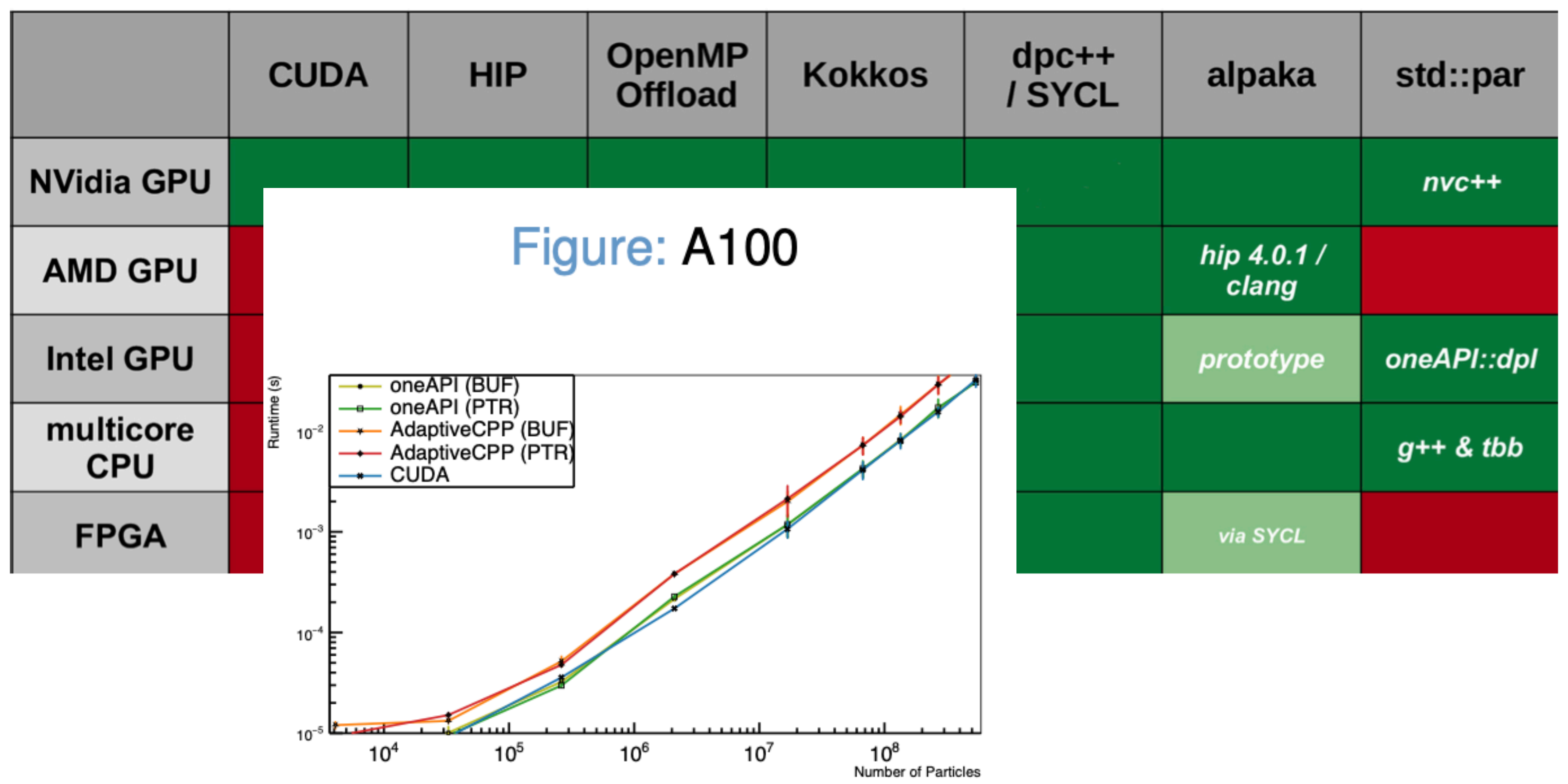


Tomasso Tedesch

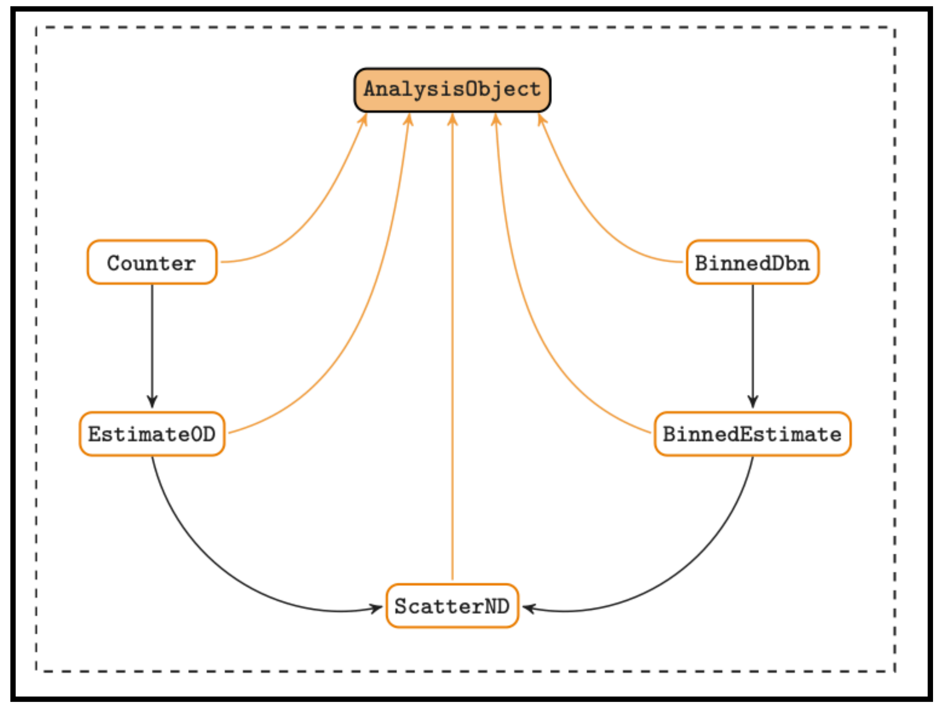


Hardware Managment

Figure: Hardware support of portability layers[1]



Histograms



Christian Gütschow

Declarative Languages

```
topo_clusters = (ds_jz2_exot15
                 .SelectMany(lambda e: e.Jets())
                 .SelectMany(lambda j: j.getConstituents())
                 .Select(lambda tc: tc.pt())
                 .AsAwkwardArray('JetClusterPt')
                 .value()
                )
```

Gordon Watts

High Level Thoughts

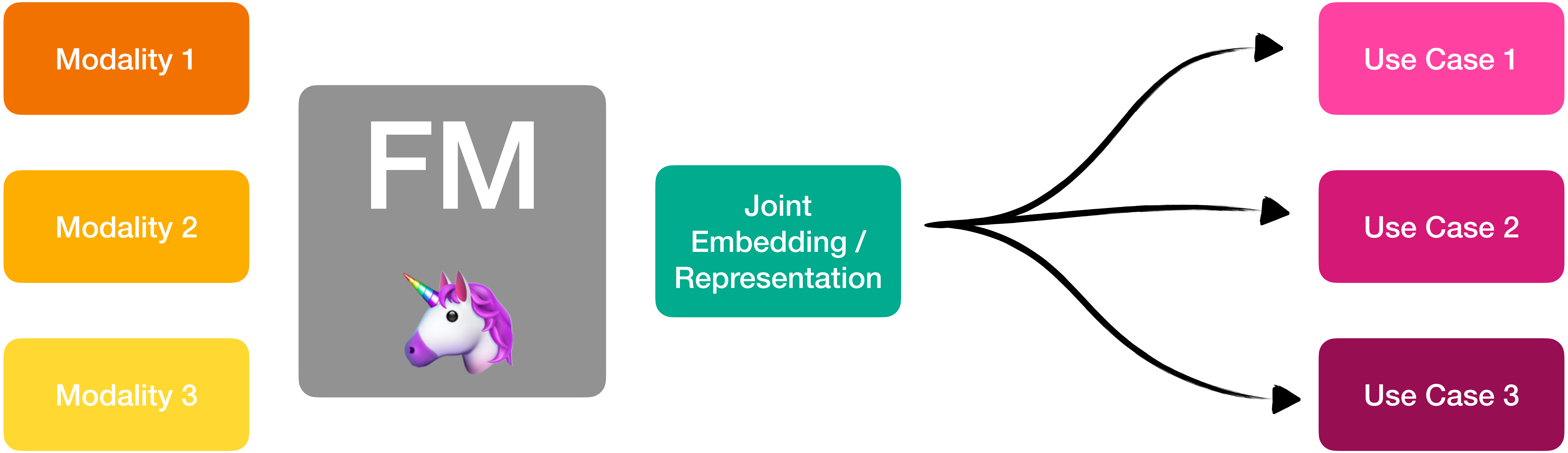
ML is absolutely everywhere. Even in the smallest nooks & crannies. Simple General-Purpose Tool.

The level of sophistication and speed at which ideas from ML research are integrated is impressive.

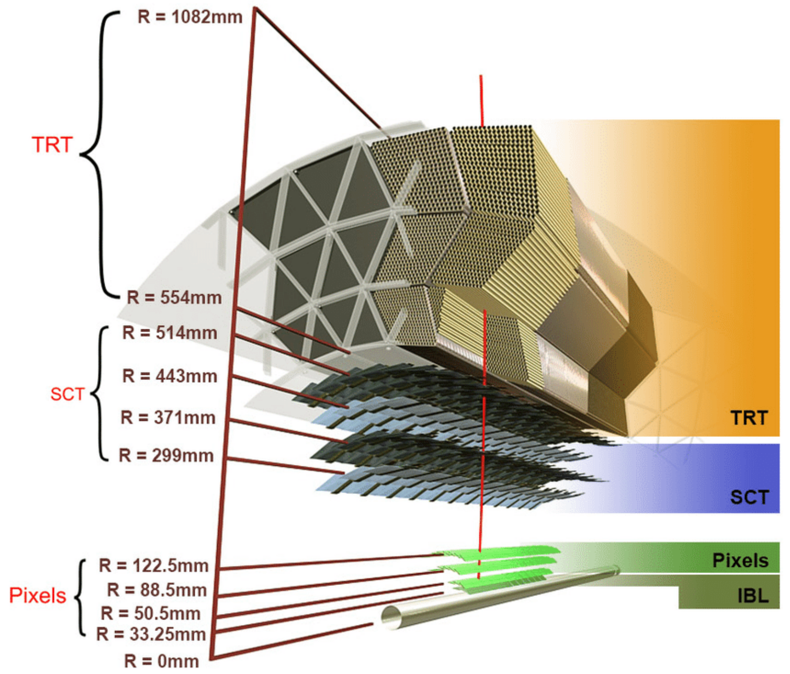
Diverse Physics exposes us to different problems and connection points to ML research.

ML is not everything. At the end of the day the work should make contact with the real world and actually be deployed. Tools and Engineering Matter and should get rewarded.

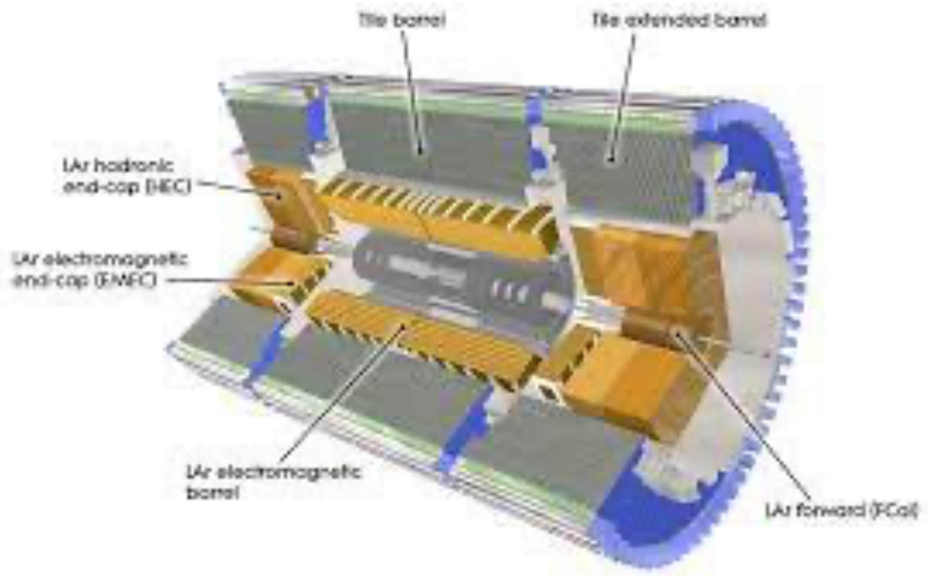
About Foundation Models..


Modality 1
Modality 2
Modality 3
FM
Joint Embedding / Representation
Use Case 1
Use Case 2
Use Case 3

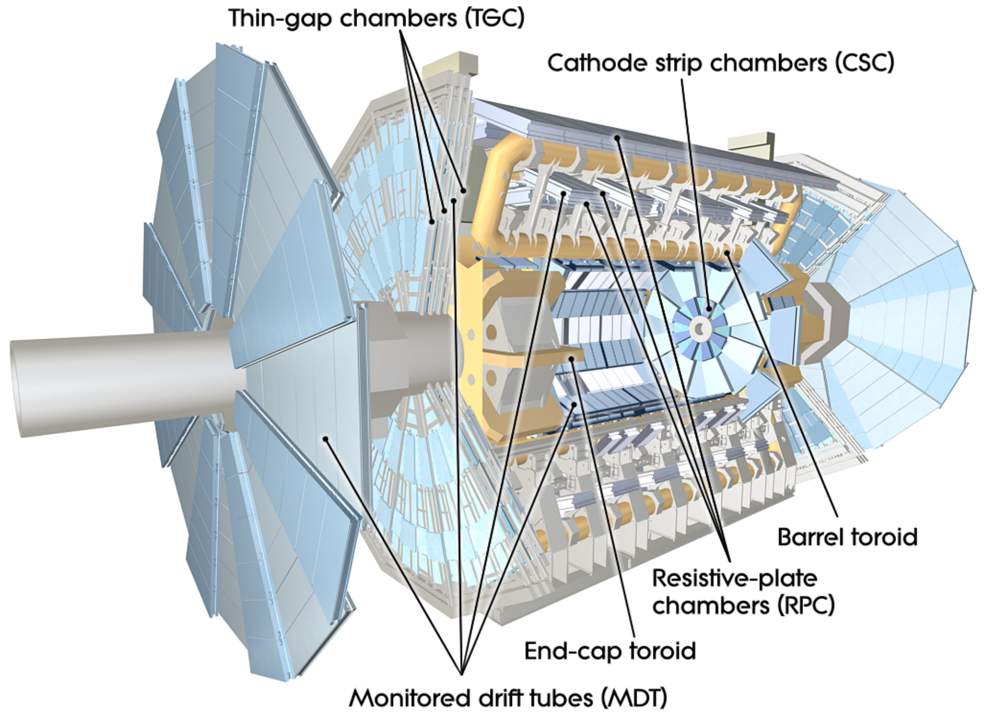
Tracking Data



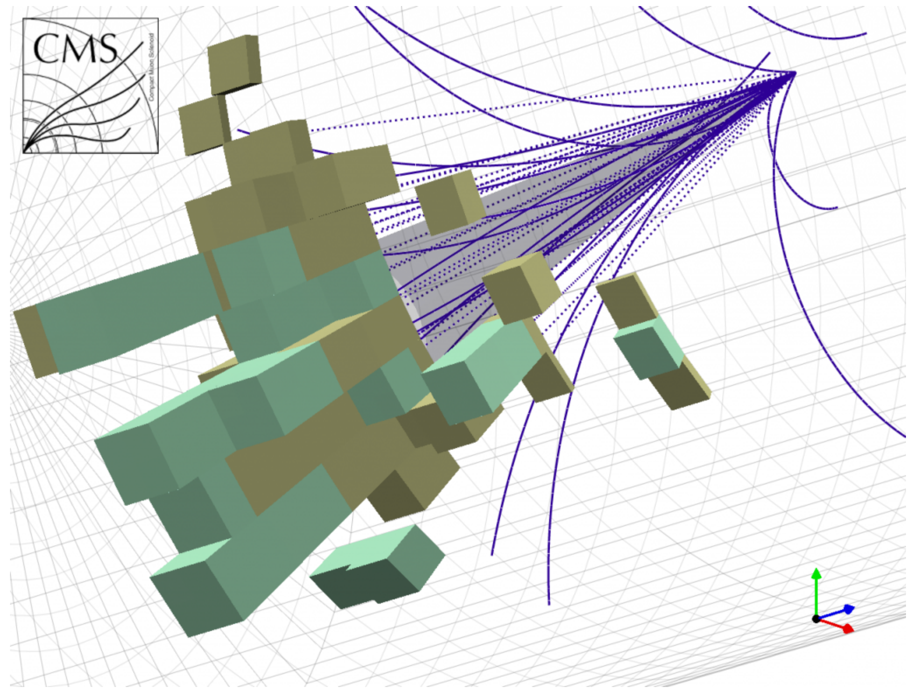
Calorimeter



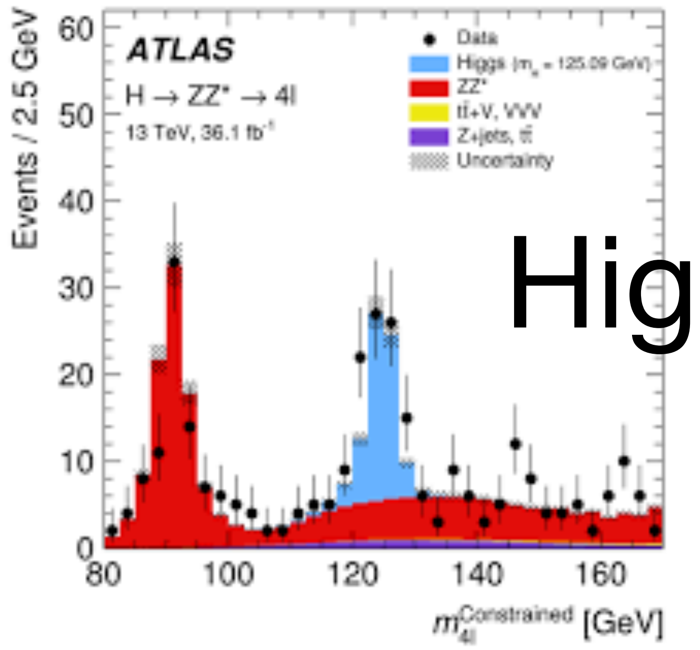
Muon Data



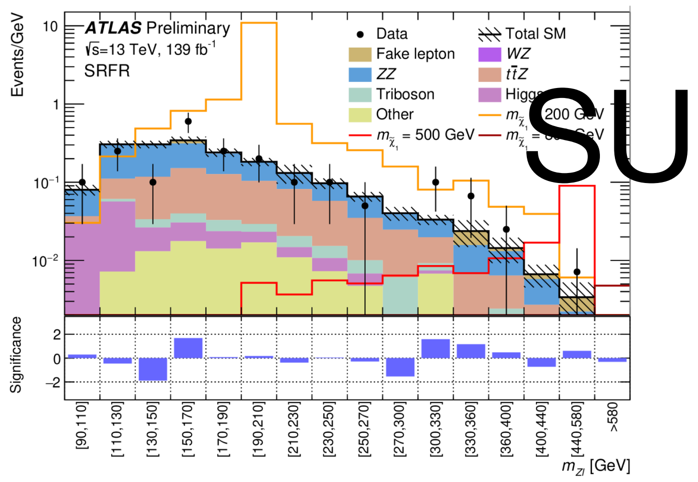
Reco



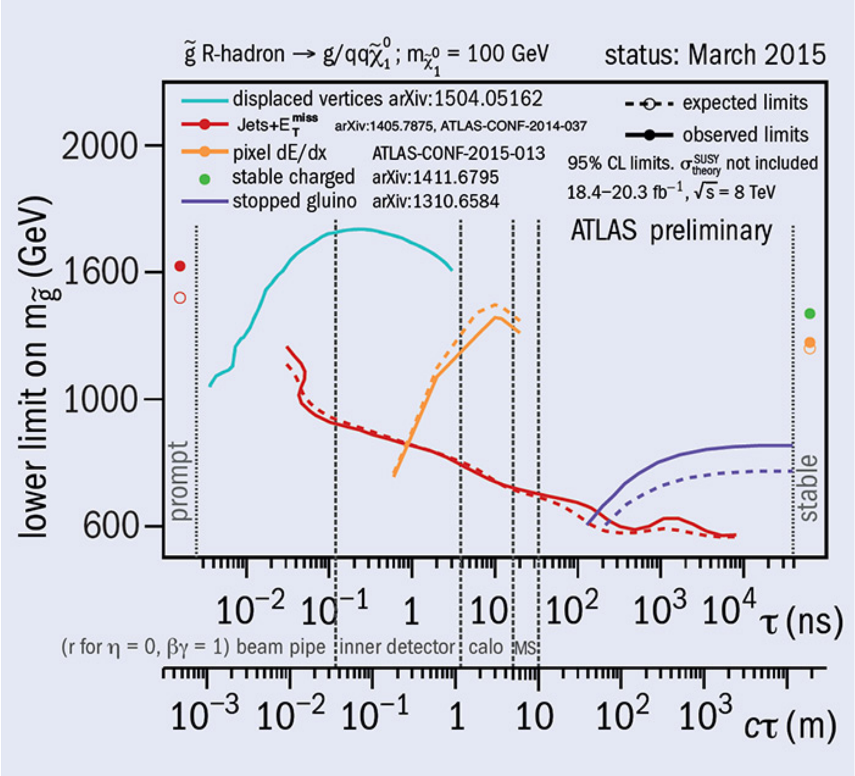
Reconstructed Event



Higgs



SUSY

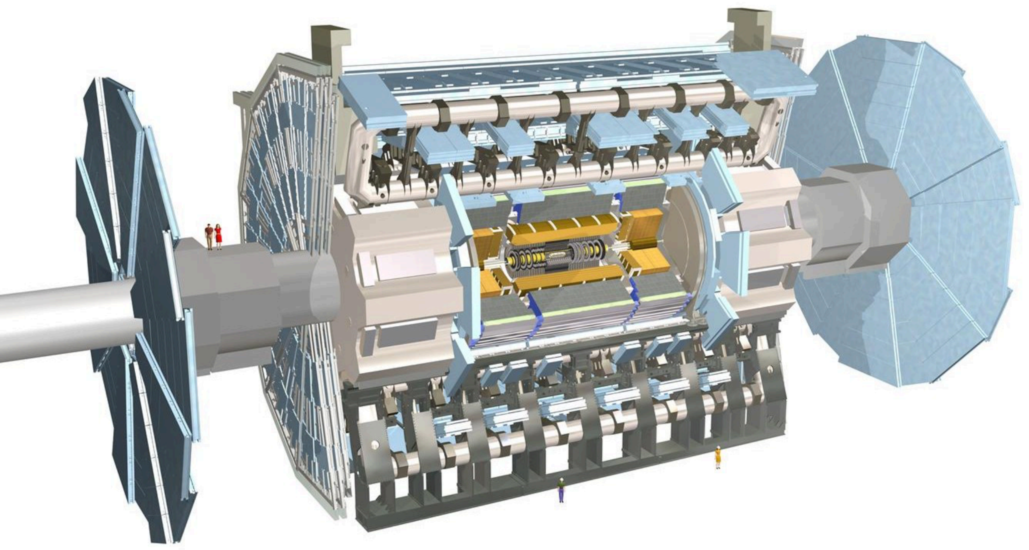


Exotic Particles

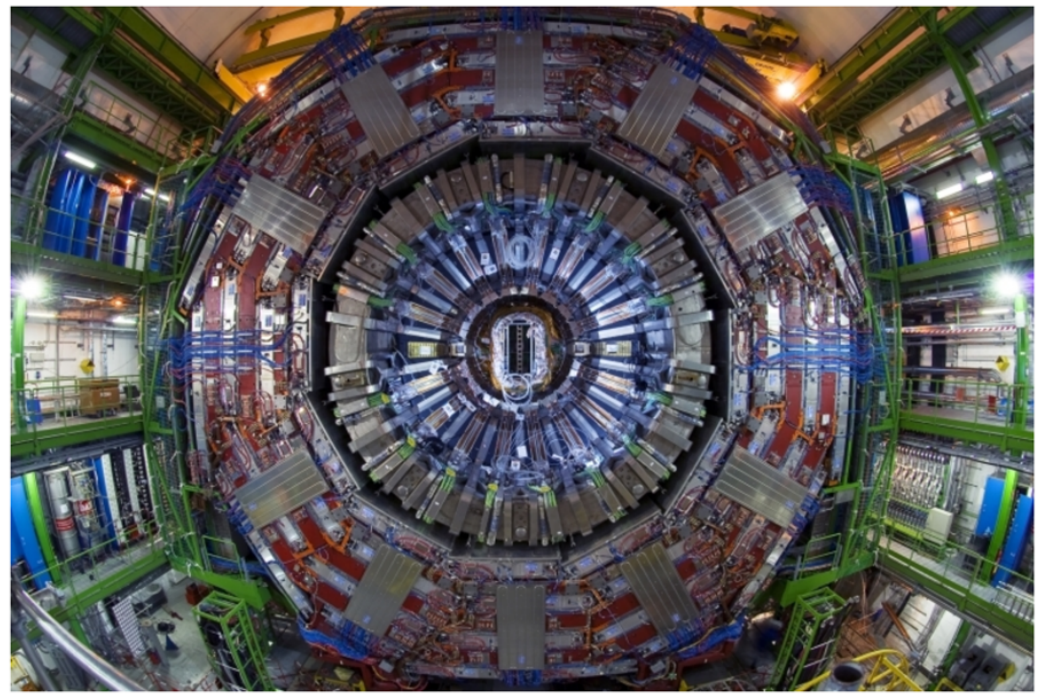
In many ways we've always had a foundation model for general purpose experiments..

Some more detail here: [Slides]

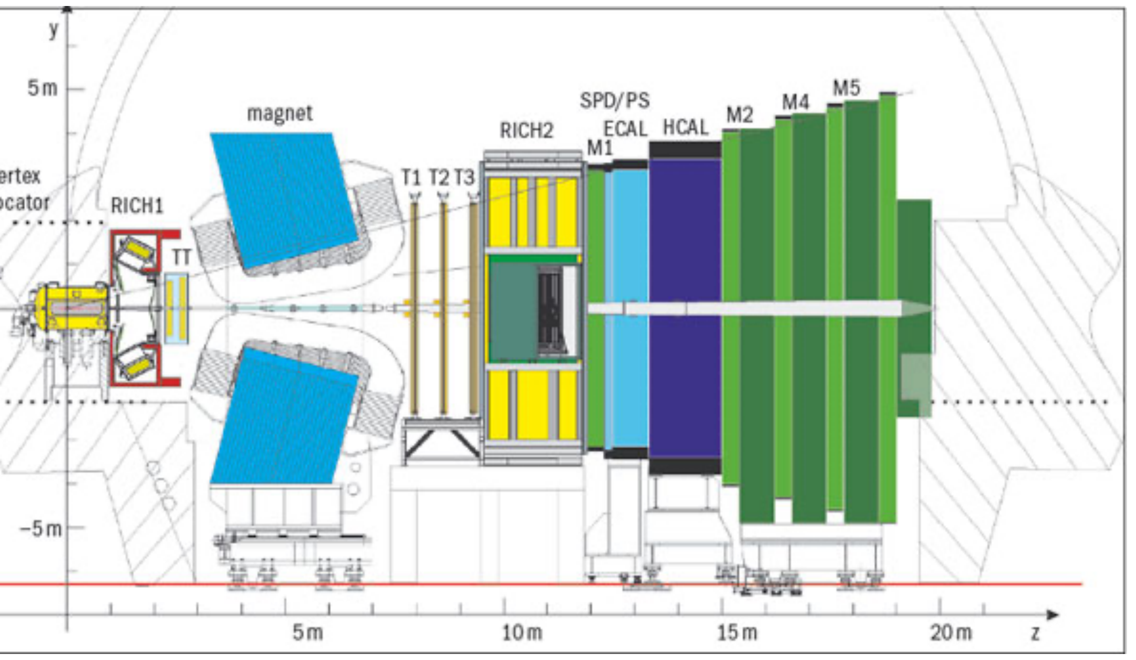
ATLAS



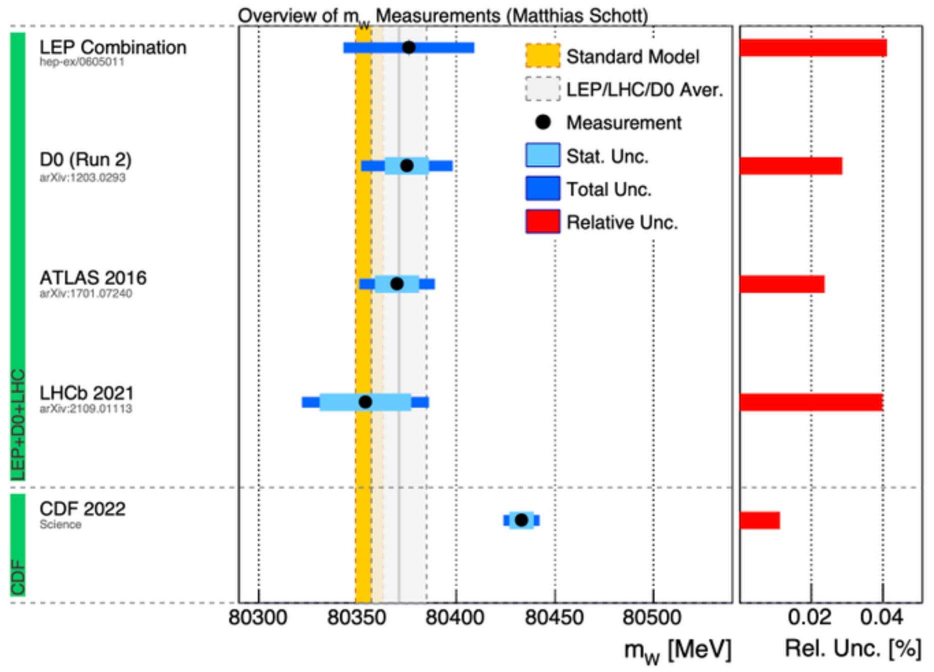
CMS



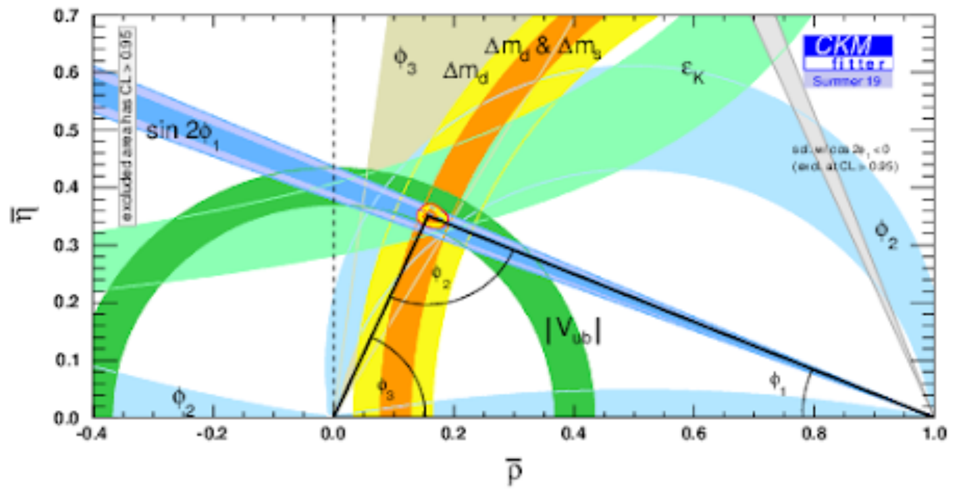
LHCb

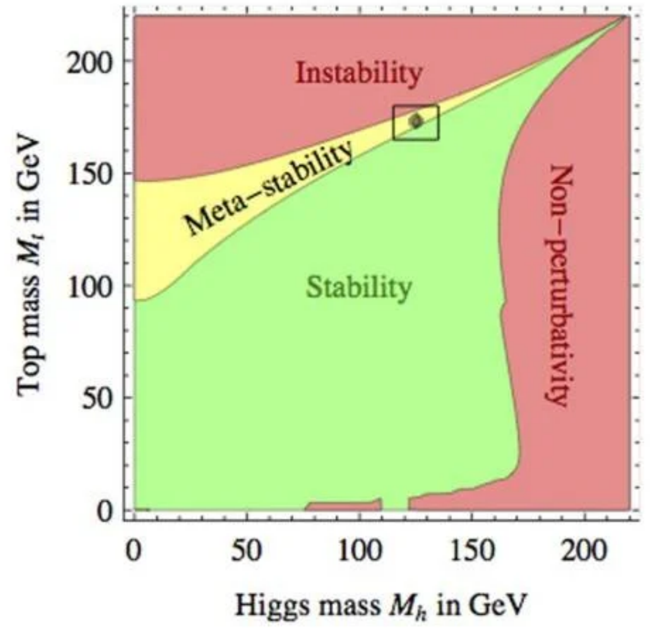


Analysis 🐴



Particle Theory





Theory is the "joint emedding" space in which we compare different experiments..

What do we mean by Foundation Models

Just a common backbone?

Does it have to trained unsupervised?

How portable should it be?

Fixed Inputs vs open-ended data streams?

Does it have to relate to language models ?

Lots of existing questions. We’ll see where it takes us.

What do we mean by Foundation Models

This is <u>not</u> to say we shouldn't work on *new, neural Foundation Models of the current kind !*

The opposite! we see that the way we've always worked is very much aligned with these new technologies

→ a lot of opportunities ahead

→ a bit of care needed to make it mean something and not just the buzzword du jour

Outlook

ACAT always a great opportunity to catch up with latest developments in the field and see a glimpse of the future

Thanks to all the speakers for the great talks and discussion!