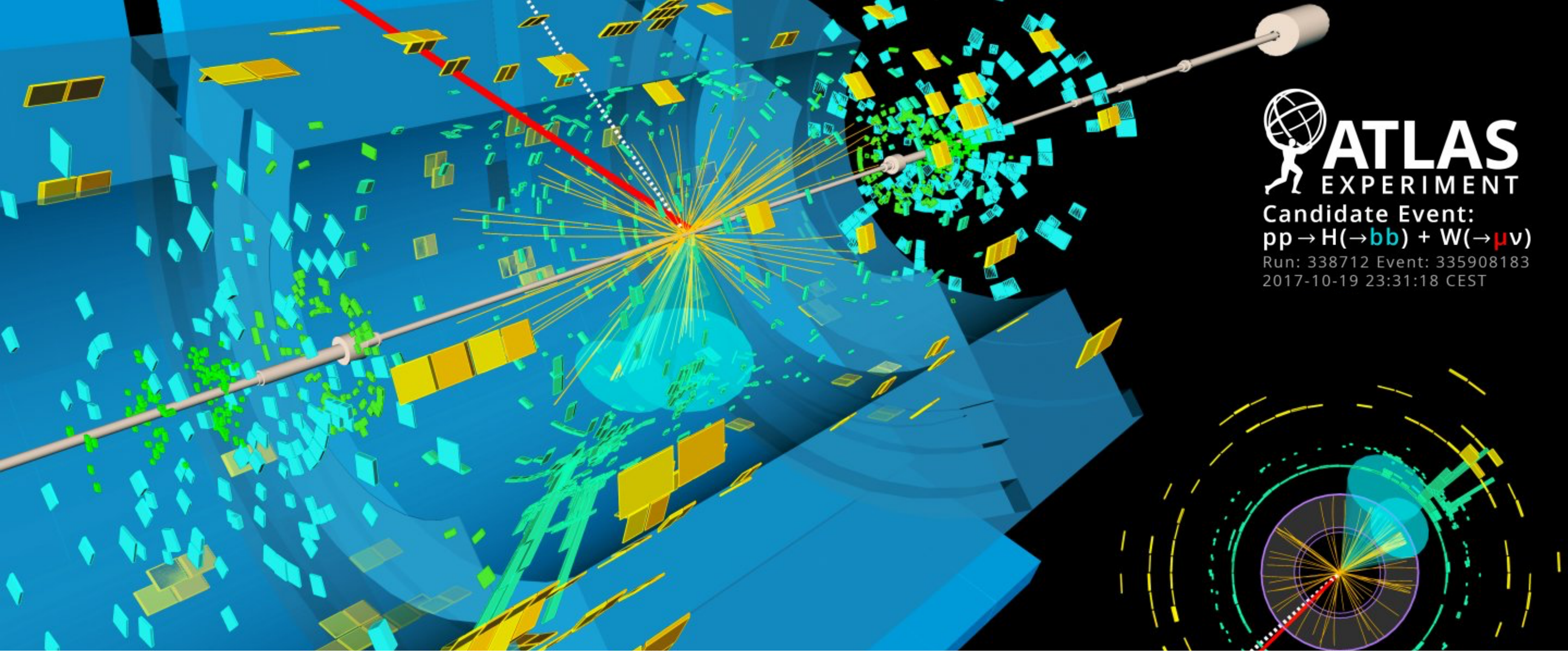


ML in Data Analysis: Use cases

Lecture 2

Sofia Vallecorsa | Ilaria Luise

Thematic CERN School of Computing on Machine Learning

17th October 2024

So WHERE and HOW can we use Deep Learning in HEP?

NB: LLMs are also quickly entering our domain too

Deep Learning in HEP

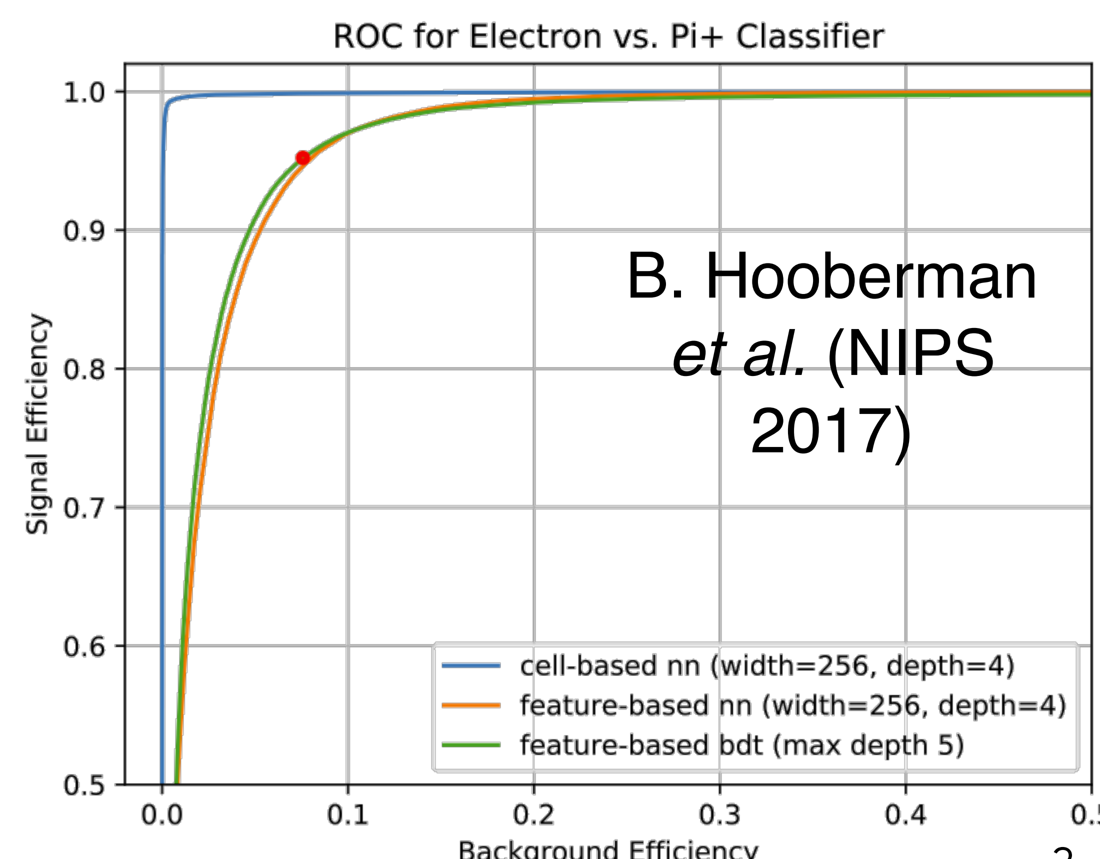
DL can **recognize patterns** in large complicated data sets
- Better performances if applied directly to **raw** data

Re-cast physics problems as “DL problems”
- Interpret detector output as **images** and apply techniques borrowed from **computer vision**
- Interpret physics events as **sentences** and apply **NLP techniques**

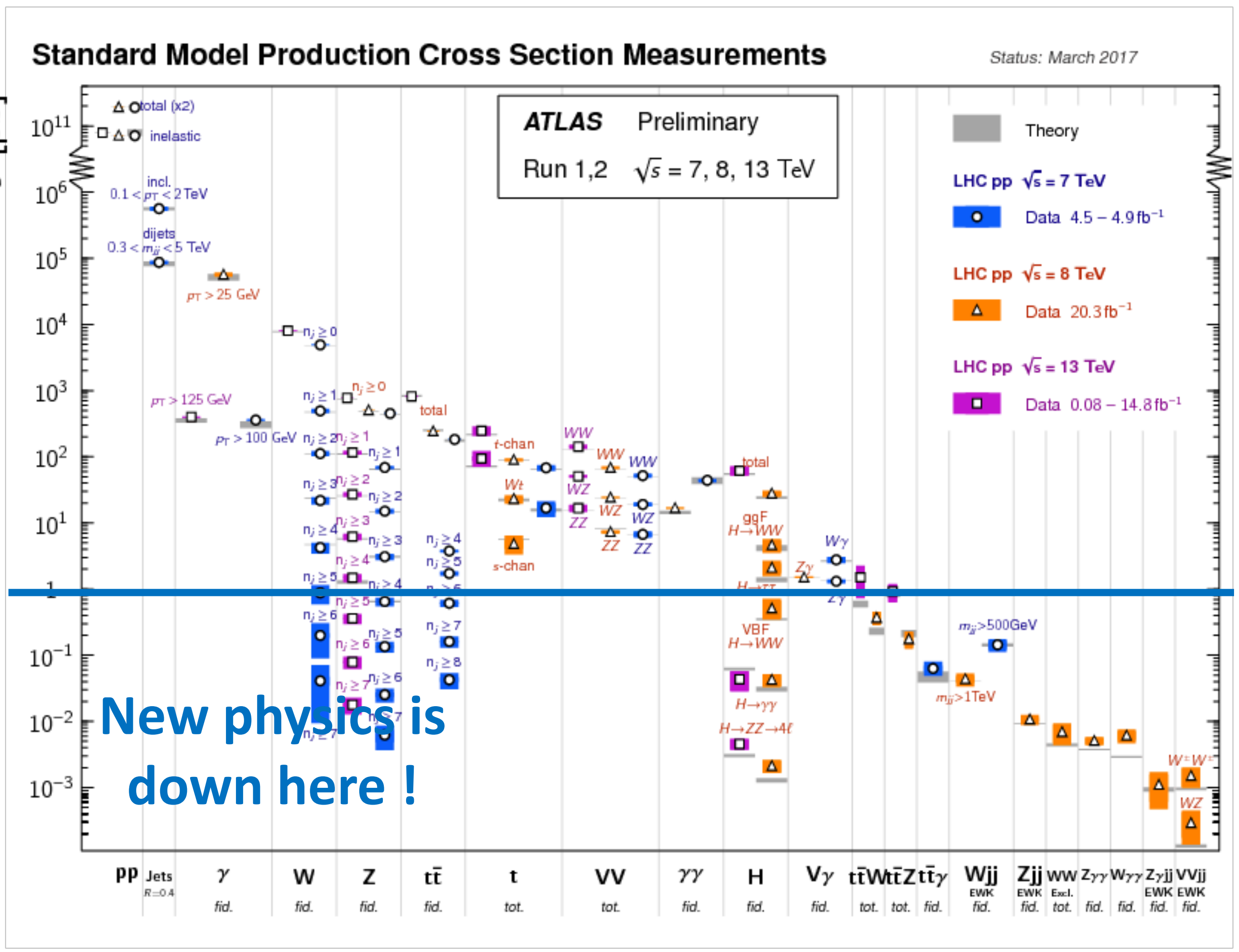
Intense R&D activity

Adapt DL to HEP requirements
- In terms of model **interpretability**
- Results **validation** against classical methods
- Detailed **systematics**

Adopting ”new” computing models
- **Accelerators** and dedicated hardware
- **HPC** integration
- **Cloud** resources
- **Big Data** platforms

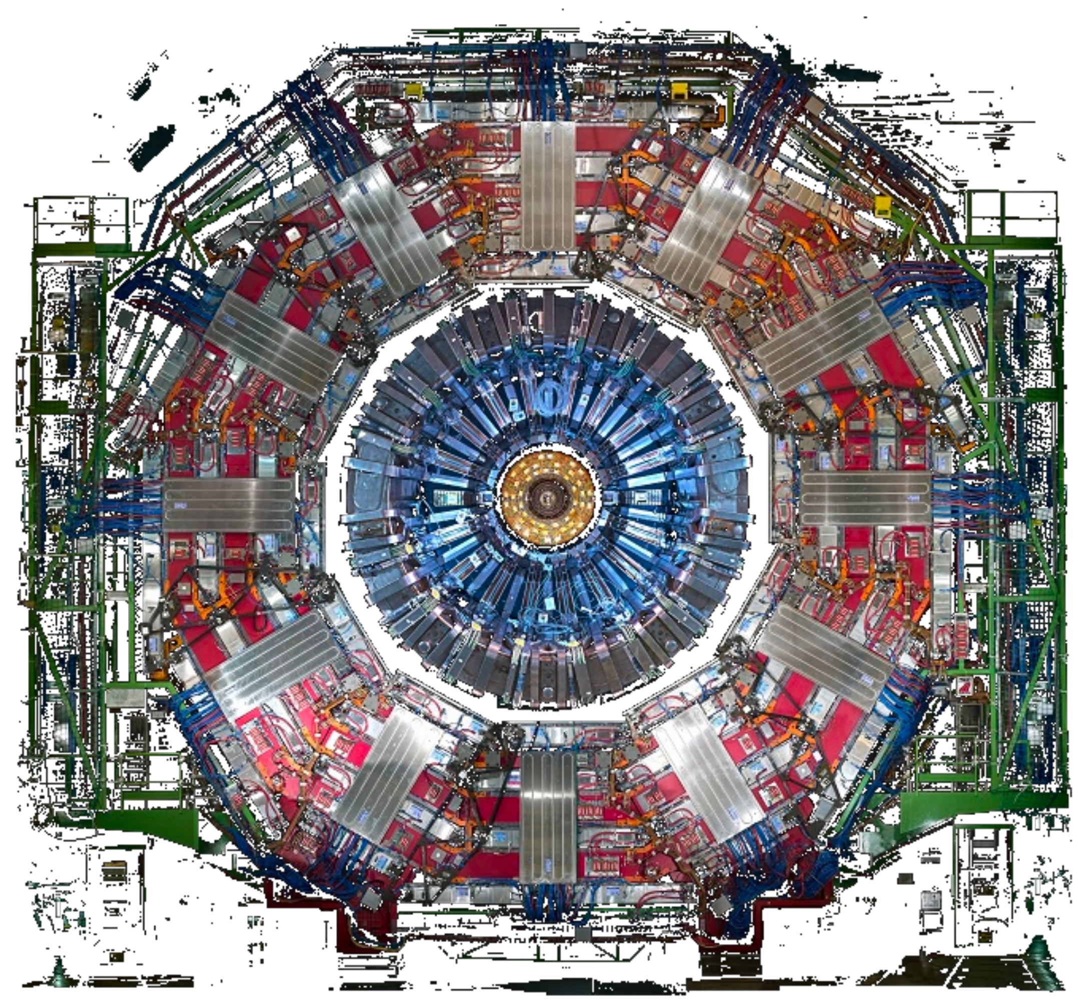


New Physics search as a Big Data problem

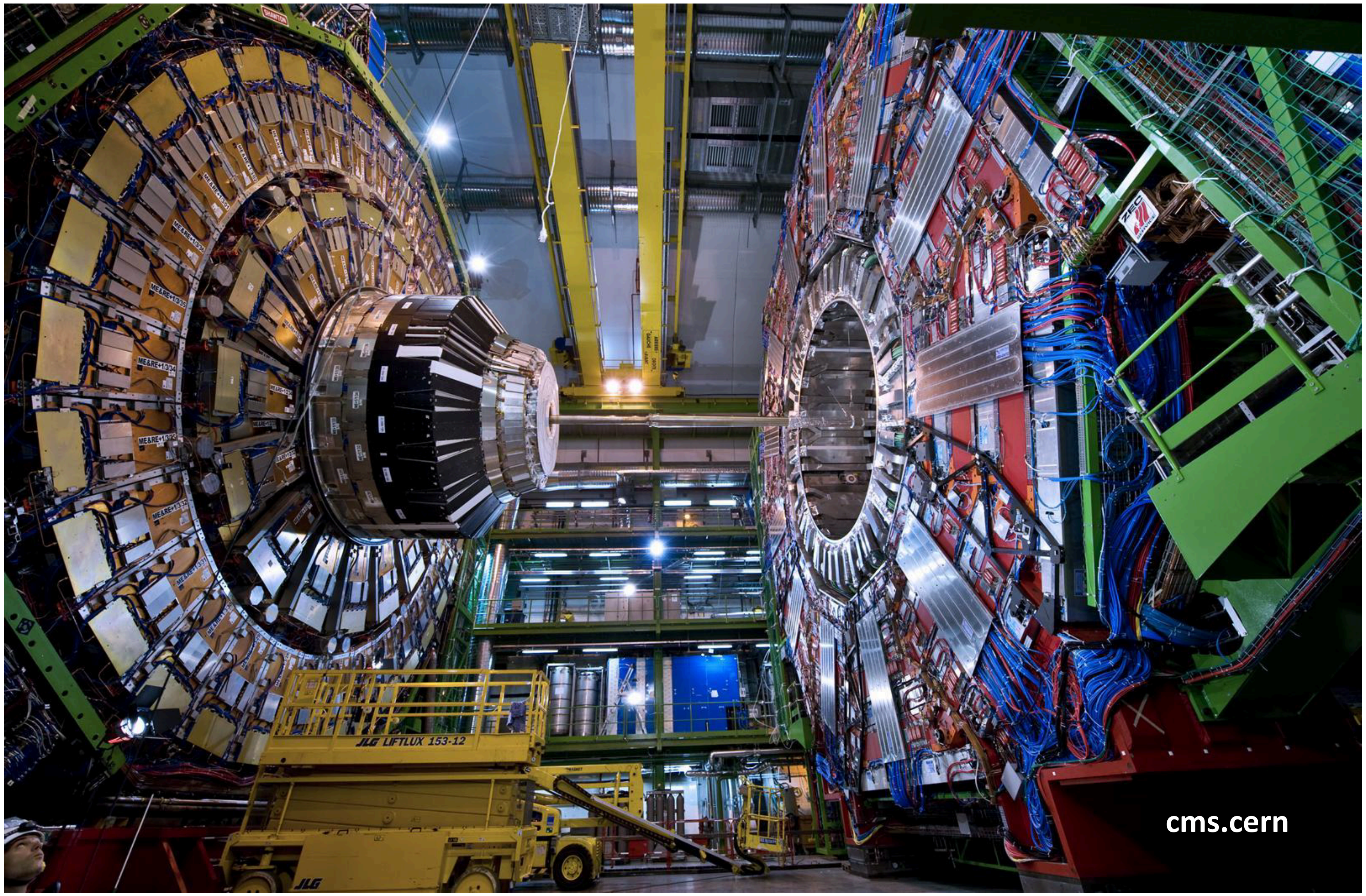


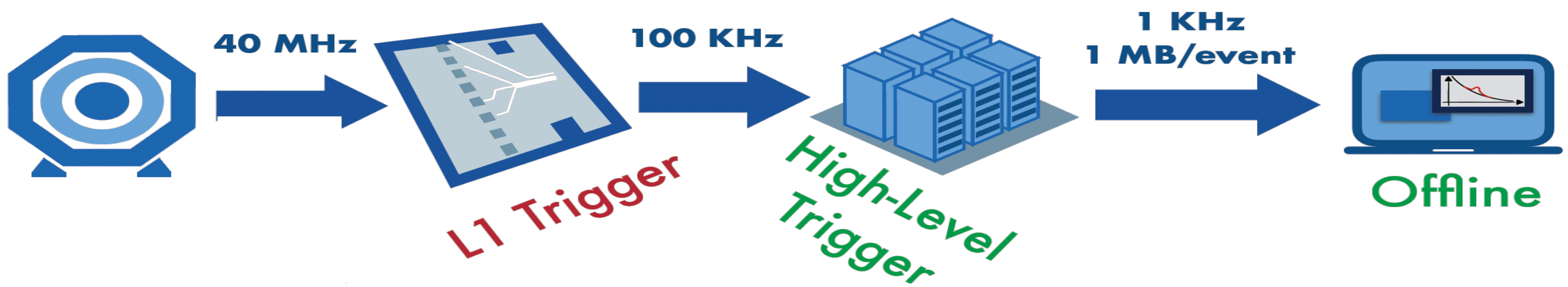
New physics is down here !

> 600 PB of collisions data



LHC data processing





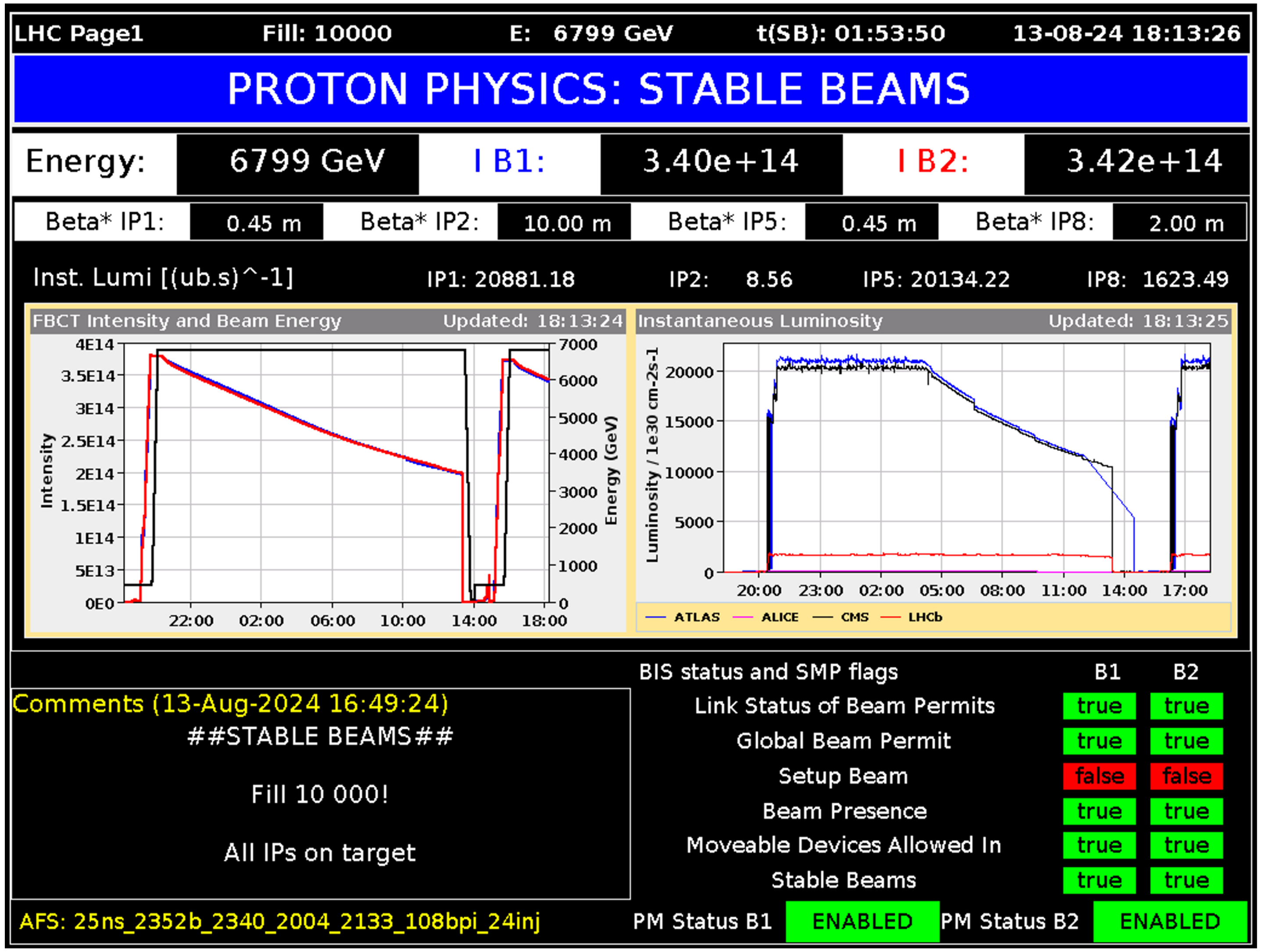
Ilaria Luise, Sofia Vallecorsa CERN - ilaria.luise@cern.ch I sofia.vallecorsa@cern.ch

Table of Content

LHC Run 3 running conditions

Fact sheet:

- Since 2015 we have collisions at 6.5 TeV (6.8 TeV from 2022)
- 25 ns bunch crossing
- Peak collision rate at 30 MHz (2017-2018)
- LHC peak instantaneous luminosity of $2 \times 10^{34} cm^{-2} s^{-1}$ (2023)
- About 50 pileup collisions



https://home.cern/news/news/accelerators/accelerator-report-10-000-lhc-fills

Online Machine/Deep Learning

Constraints on Latency (Need fast inference)

—> can we deploy ML/DL on dedicated ASICs, FPGAs?

Constraints on Model Complexity (need to find ways of compressing / reducing ML/DL size or simplify graph topology)

—> how do we introduce quantisation, compression, distillation, …

Constraints on the quality of data available (Very close to raw data. Physics quantities are known with limited resolution or limited detector information)

—> is this really a limitation for ML/DL ?

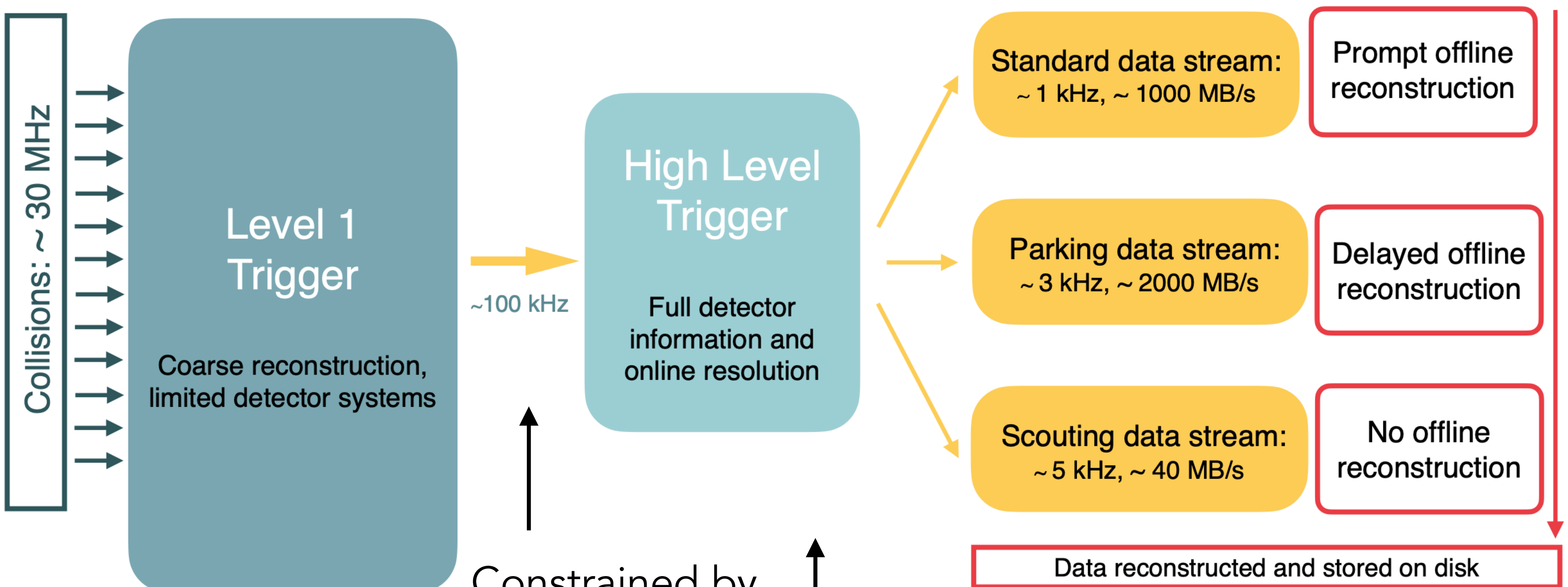
There are a lot of possible applications for ML/DL in the real time detectors operation
Data Quality Monitoring , Adaptive Data Acquisition Systems , Triggers

Sofia Vallecorsa, Ilaria Luise CERN - sofia.vallecorsa@cern.ch | ilaria.luise@cern.ch

Real-time data processing at CMS

- We can process only a **minimal fraction** of collider data
  - Keep only **interesting** events

Data flow for a typical 2018 data-taking scenario



The CMS Collaboration, arXiv:2403.16134 [hep-ex]

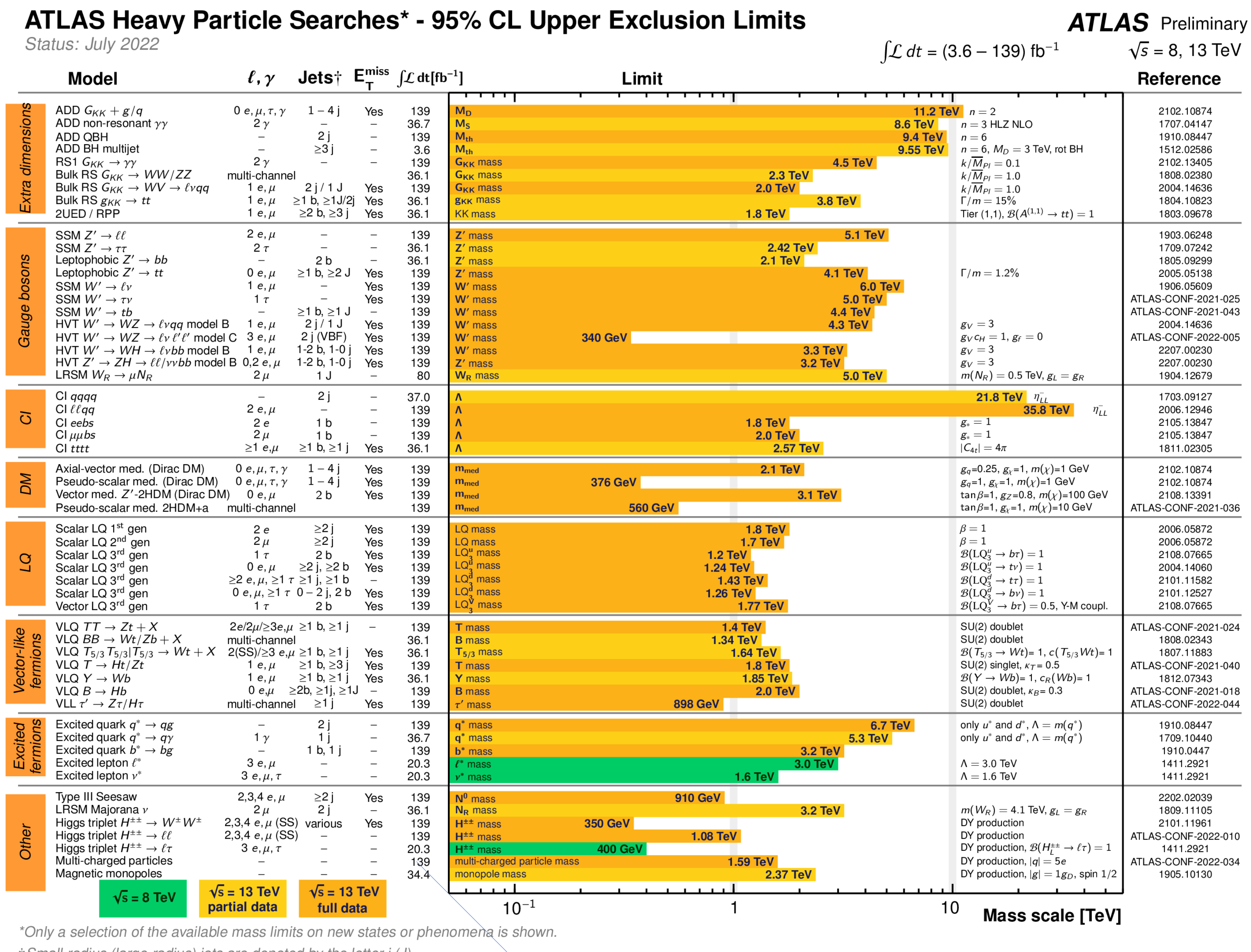
4 μs fixed latency
Information from calorimeters and muons

Constrained by readout and size of raw data to transfer

Constrained by HLT farm "size" (Ex. 600 ms/event in 2018)

NB. DAQ output bandwidth at 20 GB/s for Run 3 is not a bottleneck

Model dependent search results



How to insure we do not miss potential discoveries?

Model agnostic searches

TABLE I. Existing two-body exclusive final state resonance searches at $\sqrt{s} = 8$ TeV. The $\varnothing$ symbol indicates no existing search at the LHC.

	e	μ	τ	γ	j	b	t	W	Z	h
e	$\pm\mp$[4],$\pm\pm$[5]	$\pm\pm$[5, 6] $\pm\mp$[6, 7]	[7]	$\varnothing$	$\varnothing$	$\varnothing$	$\varnothing$	$\varnothing$	$\varnothing$	$\varnothing$
μ		$\pm\mp$[4],$\pm\pm$[5]	[7]	$\varnothing$	$\varnothing$	$\varnothing$	$\varnothing$	$\varnothing$	$\varnothing$	$\varnothing$
τ			[8]	$\varnothing$	$\varnothing$	$\varnothing$	[9]	$\varnothing$	$\varnothing$	$\varnothing$
γ				[10]	[11–13]	$\varnothing$	$\varnothing$	[14]	[14]	$\varnothing$
j					[15]	[16]	[17]	[18]	[18]	$\varnothing$
b						[16]	[19]	$\varnothing$	$\varnothing$	$\varnothing$
t							[20]	[21]	$\varnothing$	$\varnothing$
W								[22–25]	[23, 24, 26, 27]	[28–30]
Z									[23, 25, 31]	[28, 30, 32, 33]
h										[34–37]

From Craig, Draper, Kong, Ng & Whiteson 1610.09392

What if BSM is not like any of the models we have searched for?

There are a lot of BSM scenarios that are not covered by existing searches!

Deep Learning allows us to go model agnostic

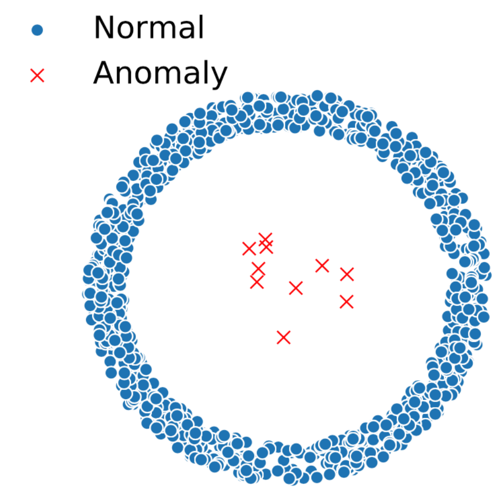
—> Anomaly Detection

											BSM $\to$ SM$_1$ $\times$ SM$_1$				BSM $\to$ SM$_1$ $\times$ SM$_2$				BSM $\to$ complex			
		e	μ	τ	q/g	b	t	γ	Z/W	H	q/g	γ/π^0's	b	$\cdots$	tZ/H	bH		$\cdots$	$\tau qq'$	eqq'	$\mu qq'$	$\cdots$
	e	[37, 38]	[39, 40]	[39]	$\varnothing$	$\varnothing$	$\varnothing$	[41]	[42]	$\varnothing$	$\varnothing$	$\varnothing$	$\varnothing$		$\varnothing$	$\varnothing$	$\varnothing$		$\varnothing$	[43, 44]	$\varnothing$	
	μ		[37, 38]	[39]	$\varnothing$	$\varnothing$	$\varnothing$	[41]	[42]	$\varnothing$	$\varnothing$	$\varnothing$	$\varnothing$		$\varnothing$	$\varnothing$	$\varnothing$		$\varnothing$	$\varnothing$	[43, 44]	
	τ			[45, 46]	$\varnothing$	[47]	$\varnothing$	$\varnothing$	$\varnothing$	$\varnothing$	$\varnothing$	$\varnothing$	$\varnothing$		$\varnothing$	$\varnothing$	$\varnothing$		[48, 49]	$\varnothing$	$\varnothing$	
	q/g				[29, 30, 50, 51]	[52]	$\varnothing$	[53, 54]	[55]	$\varnothing$	$\varnothing$	$\varnothing$	$\varnothing$		$\varnothing$	$\varnothing$	$\varnothing$		$\varnothing$	$\varnothing$	$\varnothing$	
	b					[29, 52, 56]	[57]	[54]	[58]	[59]	$\varnothing$	$\varnothing$	$\varnothing$		[60]	$\varnothing$	$\varnothing$		$\varnothing$	$\varnothing$	$\varnothing$	
	t						[61]	$\varnothing$	[62]	[63]	$\varnothing$	$\varnothing$	$\varnothing$		[64]	[60]	$\varnothing$		$\varnothing$	$\varnothing$	$\varnothing$	
	γ							[65, 66]	[67–69]	[68, 70]	$\varnothing$	$\varnothing$	$\varnothing$		$\varnothing$	$\varnothing$	$\varnothing$		$\varnothing$	$\varnothing$	$\varnothing$	
	Z/W								[71]	[71]	$\varnothing$	$\varnothing$	$\varnothing$		$\varnothing$	$\varnothing$	$\varnothing$		$\varnothing$	$\varnothing$	$\varnothing$	
	H									[72, 73]	[74]	$\varnothing$	$\varnothing$		$\varnothing$	$\varnothing$	$\varnothing$		$\varnothing$	$\varnothing$	$\varnothing$	
BSM $\to$ SM$_1$ $\times$ SM$_1$	q/g										$\varnothing$	$\varnothing$	$\varnothing$		$\varnothing$	$\varnothing$	$\varnothing$		$\varnothing$	$\varnothing$	$\varnothing$	
	γ/π^0's											[75]	$\varnothing$		$\varnothing$	$\varnothing$	$\varnothing$		$\varnothing$	$\varnothing$	$\varnothing$	
	b												[76, 77]		$\varnothing$	$\varnothing$	$\varnothing$		$\varnothing$	$\varnothing$	$\varnothing$	
	$\vdots$																					

From Kim, Kong, Nachman & Whiteson 1907.06659

Two types of anomaly searches

Main assumption: P(anomaly) ≪ P(normal)



Outliers/point anomalies: Detector malfunctions, Background-free search…

- **Pros:**
  - Can be fully unsupervised
  - Suitable for very rare anomalies
- **Cons:**
  - Definition of low p(x) is not invariant but is parametrization dependent

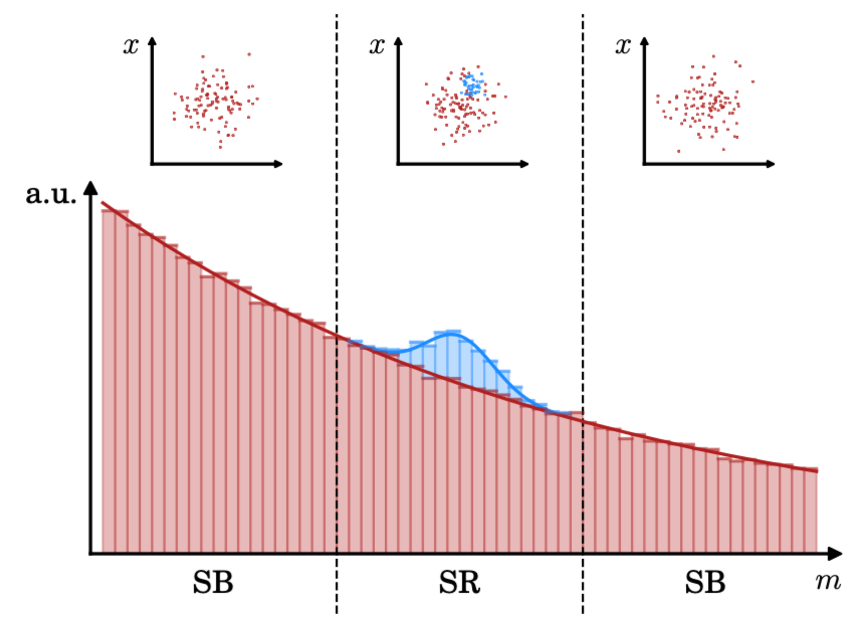
Overdensity detection: Resonant searches …

Pros:

- Re-parametrization invariant

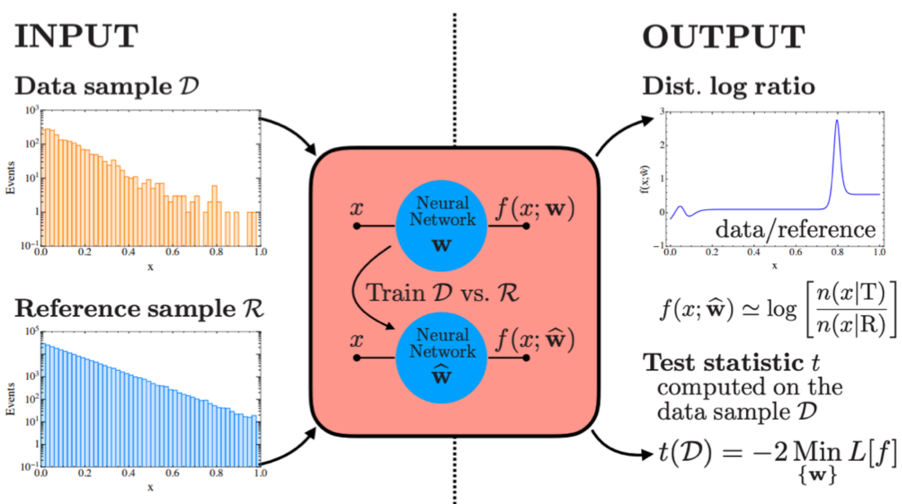
Cons:

- Prior knowledge of bkgd distribution
- Performance suffers when signal is too rare

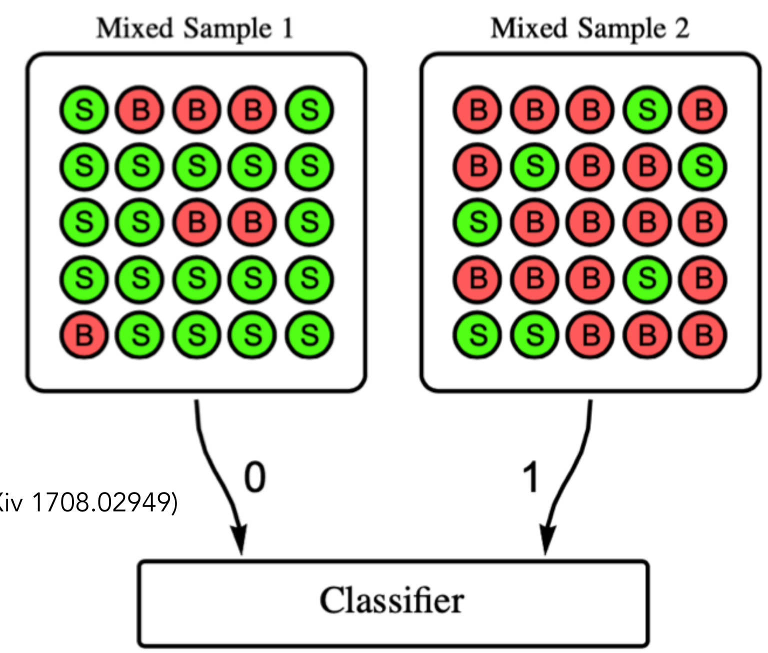


Sofia Vallecorsa, Ilaria Luise CERN - sofia.vallecorsa@cern.ch | ilaria.luise@cern.ch

Different methodologies (supervised)



Learning new physics from a machine - D'Agnolo & Wulzer



Metodiev, Nachman, Thaler (arXiv 1708.02949)

Classifier trained using data and a reference model

- Binned distributions are replaced by NN-approximated smooth functions
- Procedure gives a *p*-value and the likelihood ratio
- Sensitive to both group and low-density anomalies
- Reliant on simulation
- But scales badly with N observables…

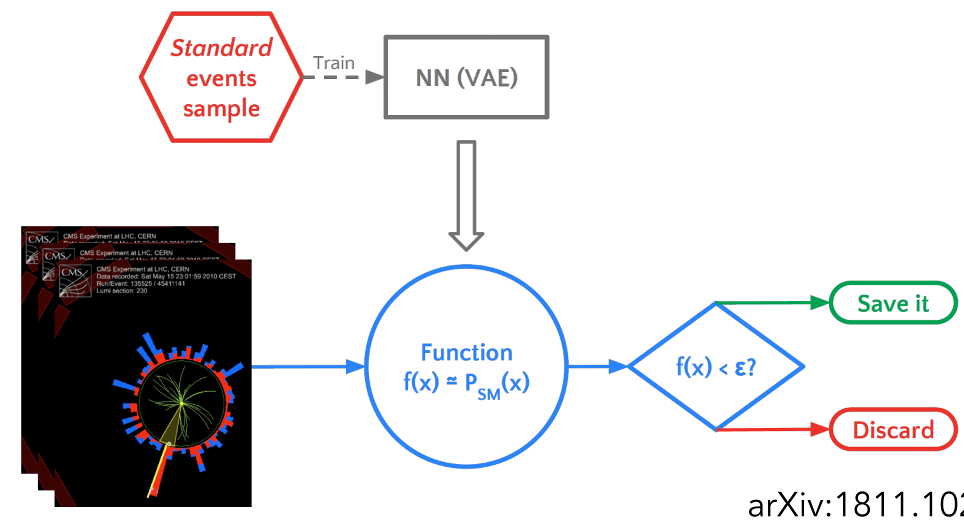
- Two samples, one signal-enriched
- Train a supervised classifier to distinguish between them
- Monotonically related to $L_{S/B}$ for $f_1 > f_2$

$$L_{M_1/M_2} = \frac{f_1 p_S + (1-f_1)p_B}{f_2 p_S + (1-f_2)p_B} = \frac{f_1 L_{S/B} + (1-f_1)}{f_2 L_{S/B} + (1-f_2)}$$

Sofia Vallecorsa, Ilaria Luise CERN - sofia.vallecorsa@cern.ch | ilaria.luise@cern.ch

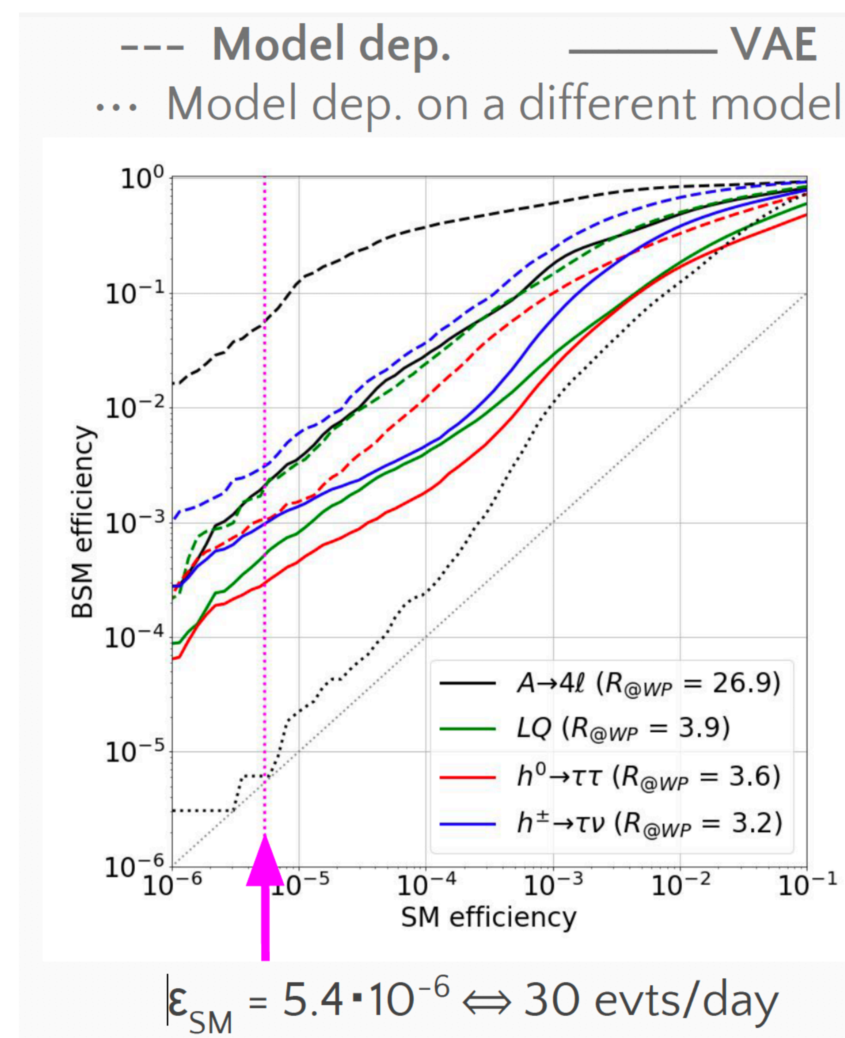
AD with Variational AutoEncoders

First demonstrations as early as 2018 !



arXiv:1811.10276
arxiv: 2005.01598

VAE as **model-independent** new physics selection tool (unsupervised)



Train on known physics

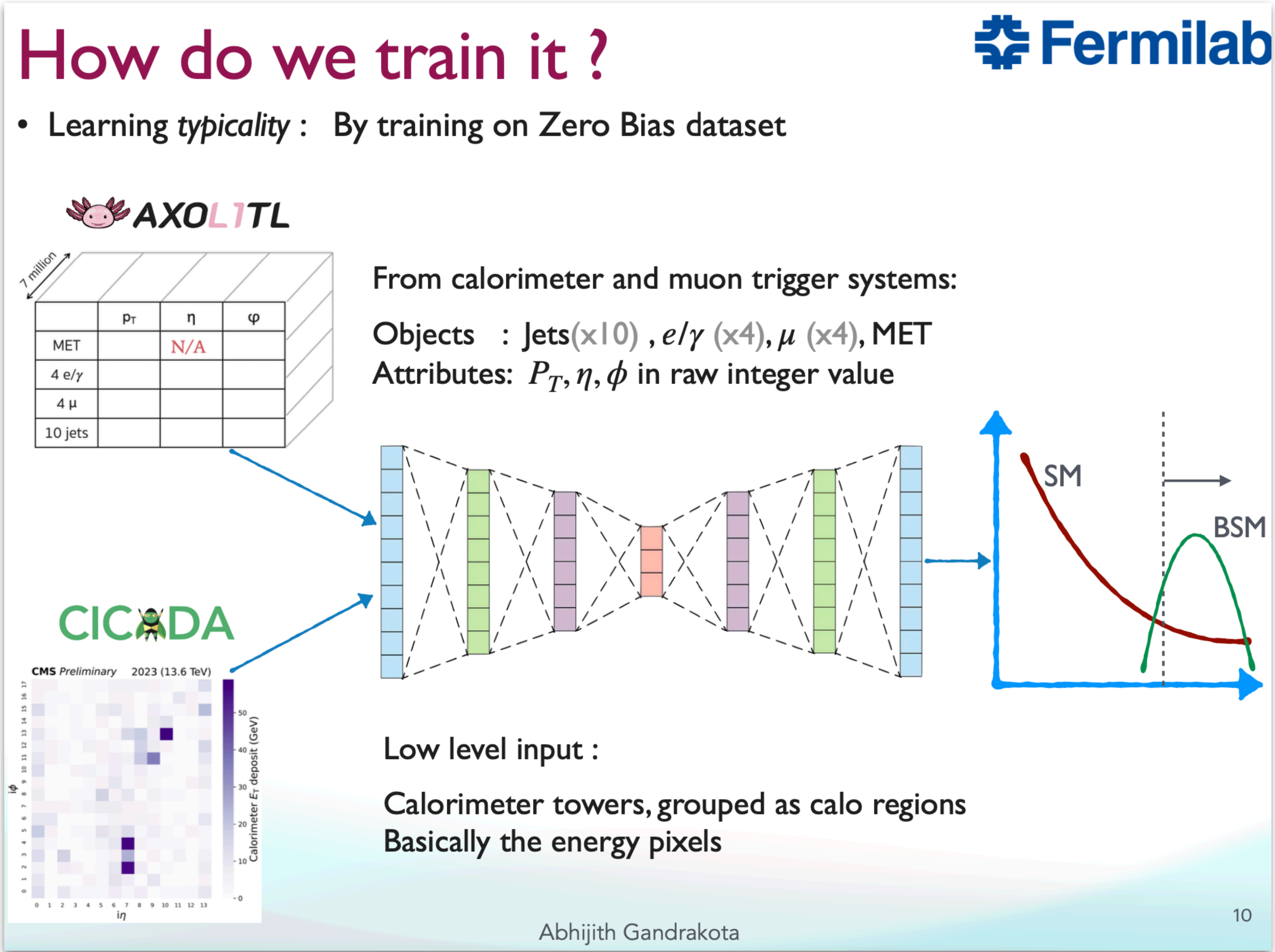
- Monte Carlo
- Real detector data

Minimise input-output difference $\mathcal{L} = ||x - x'||^2$

- Anomalies will exhibit large error!

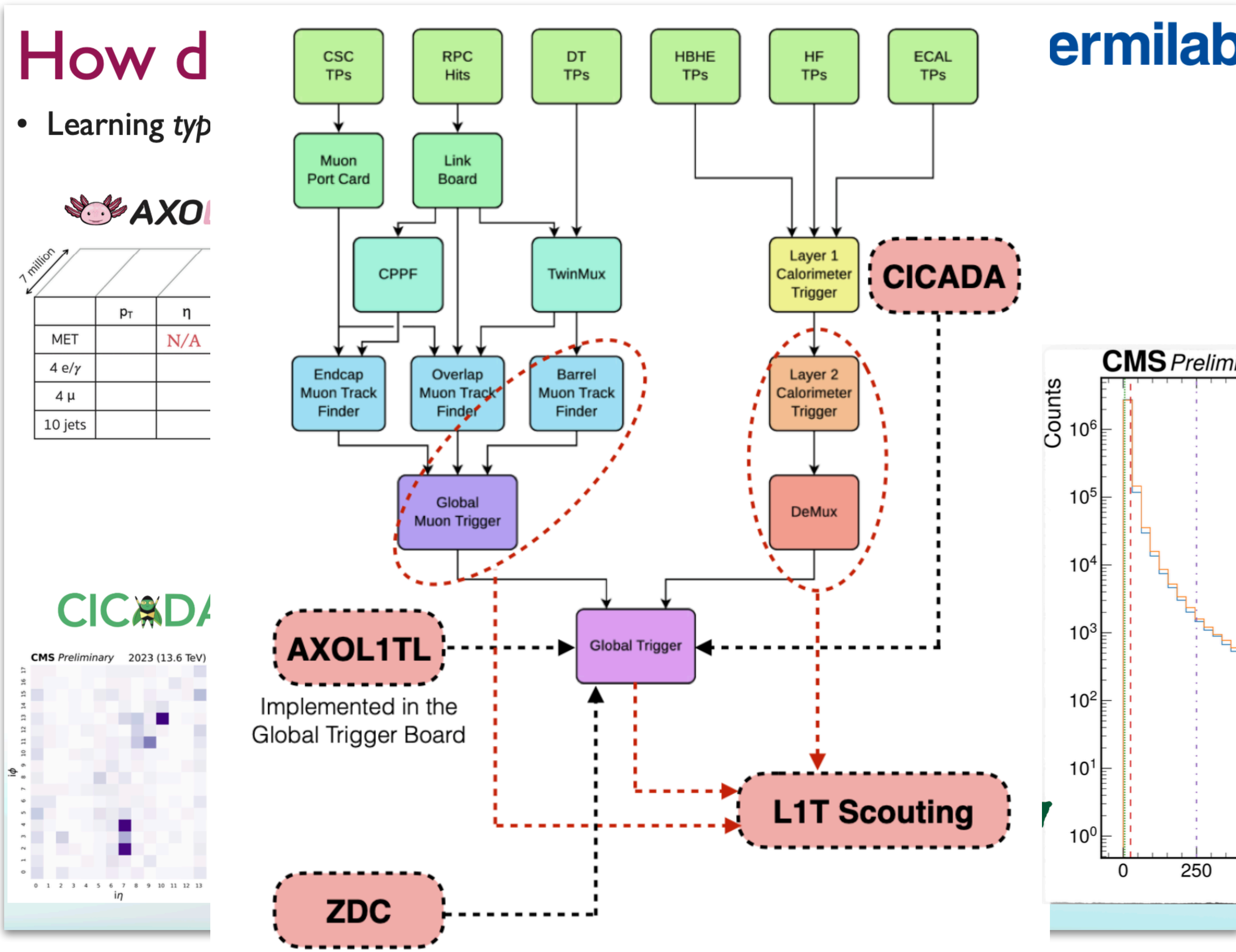
Build an anomaly score

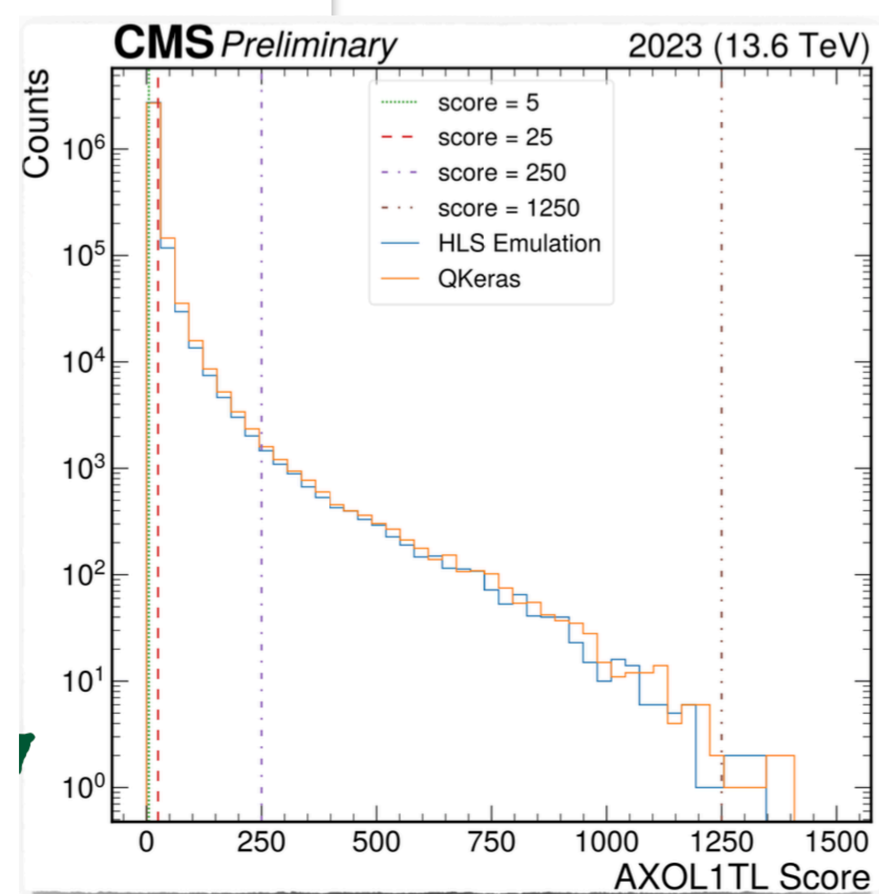
Currently running examples



A. Gandrakota, ICHEP2024

Currently running examples





A. Gandrakota, ICHEP2024

High Level Synthesis for deployment on FPGAs

- AXOL1TL and other models leverage hls4ml for inference deployment on low latency FPGA

- Ingredients for performance:

compression: reduce number of synapses or neurons

quantization: reduces the precision of the calculations (inputs, weights, biases)

parallelization: tune how much to parallelize to make the inference faster/slower versus FPGA resources

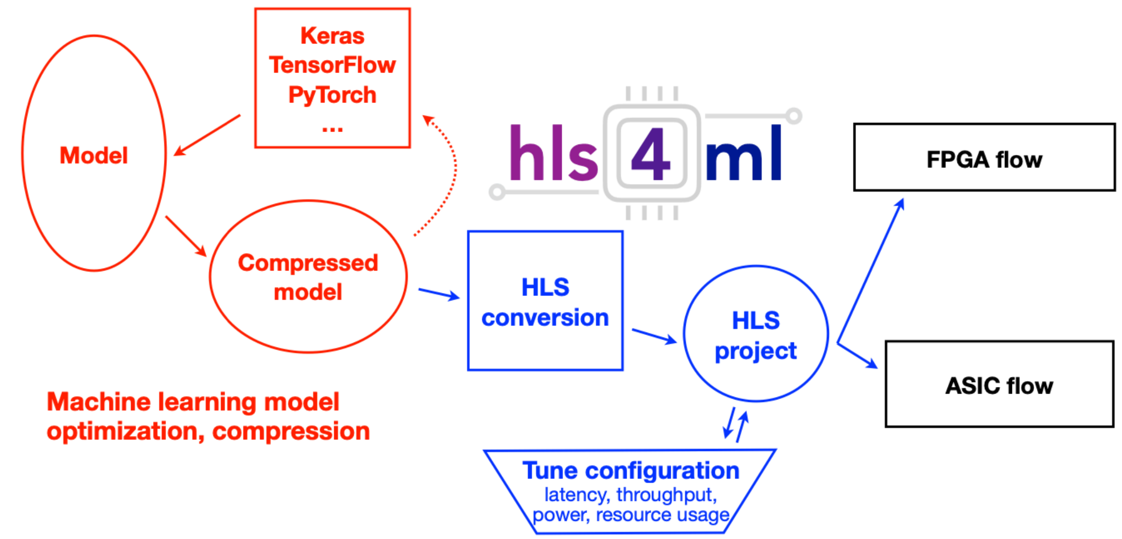
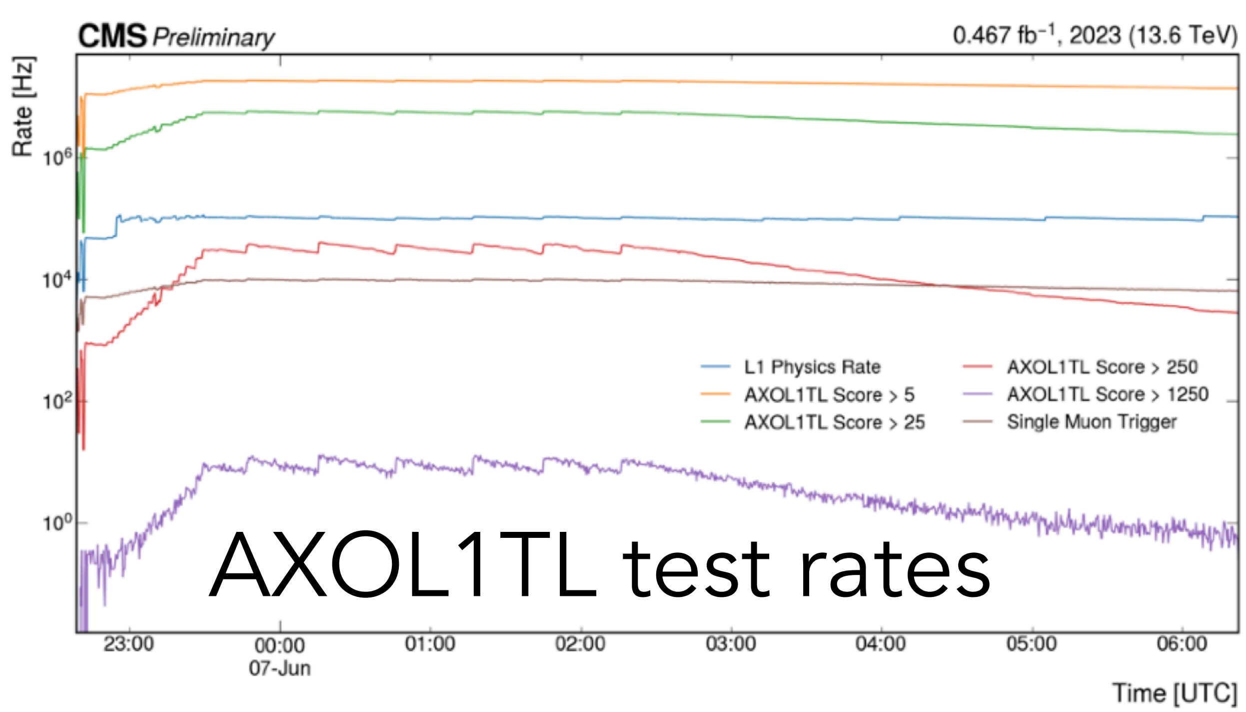


Figure 1: A typical workflow to translate an ML model into an FPGA or ASIC implementation using `hls4ml`. The red boxes (left) describe the model training and compression steps performed within conventional ML software frameworks. The `hls4ml` configuration and conversion steps are shown in the blue boxes (center). The black boxes (right) illustrate possible ways to export and integrate the HLS project into a larger hardware design.



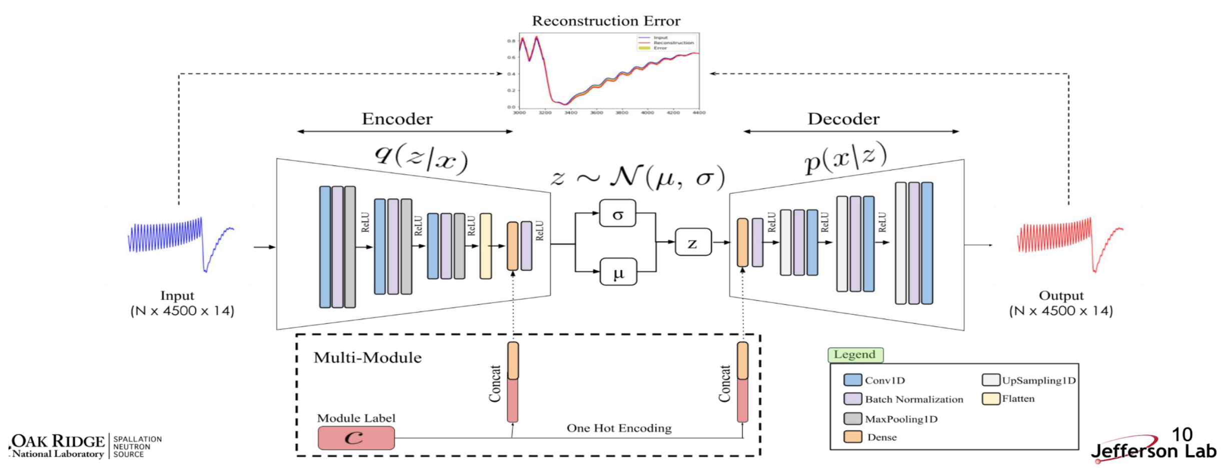
hls4ml

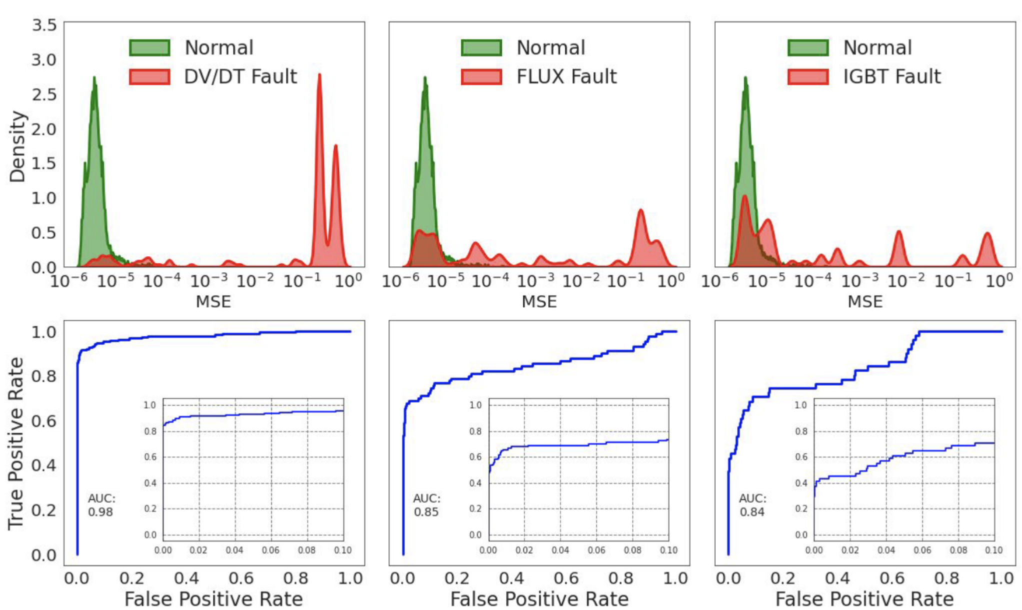
- https://github.com/hls-fpga-machine-learning/hls4ml
- https://fastmachinelearning.org/hls4ml/

Anomaly Detection for hardware

Resilient Variational Autencoder for Unsupervised Anomaly Detection at the SLAC Linac Coherent Light Source, CHEP2023

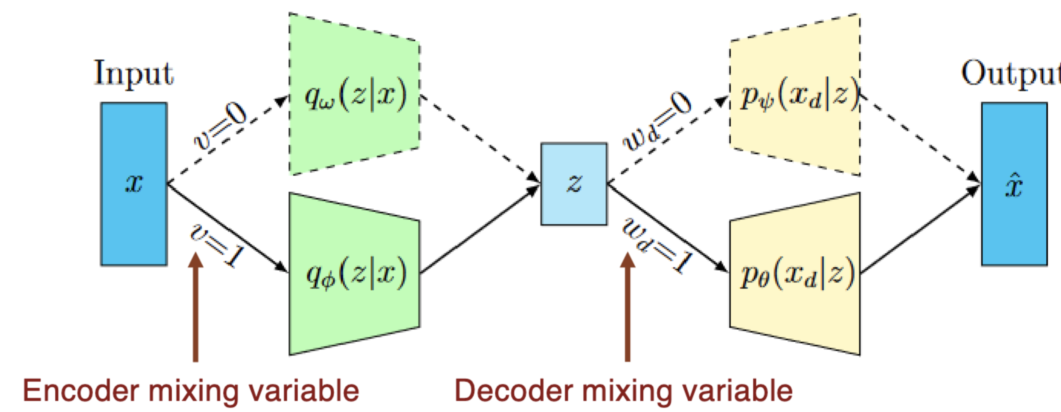
High Voltage Converter Modulators for SNS Linac: a multi-module AE performs better than **dedicated separate** modules



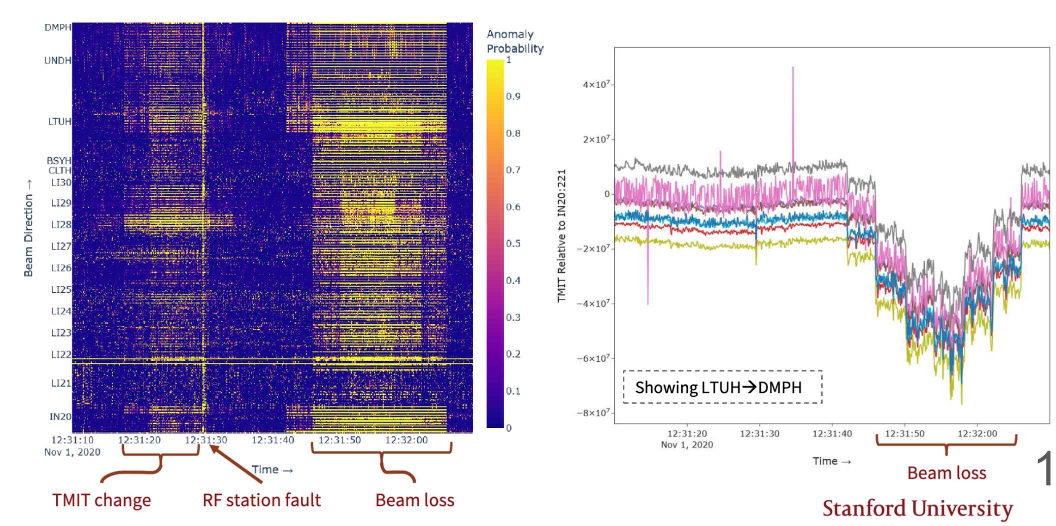


Multi-Module based VAE to predict HVCM faults in the SNS accelerator, CHEP2023

RF cavities of FEL@SLAC : Create "outlier" path through the network and decide according to probabilistic inference



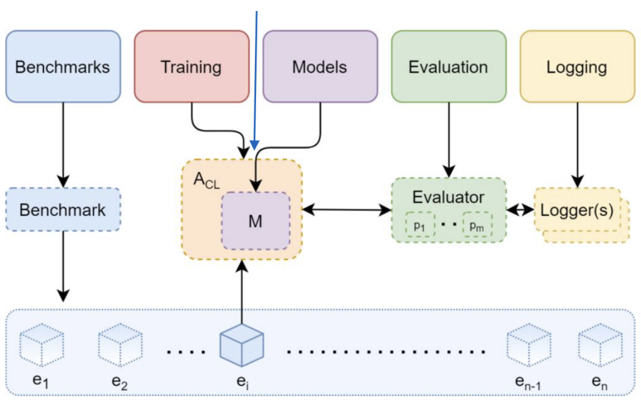
ResVAE used to identify and diagnose LCLS anomalies



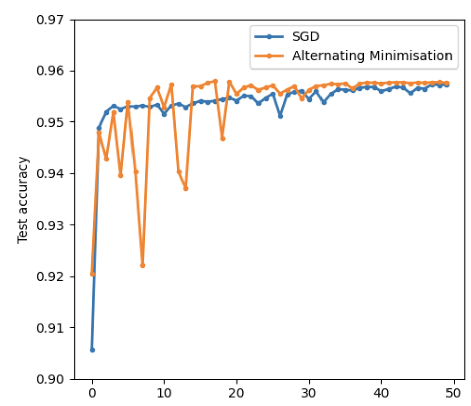
Other online applications

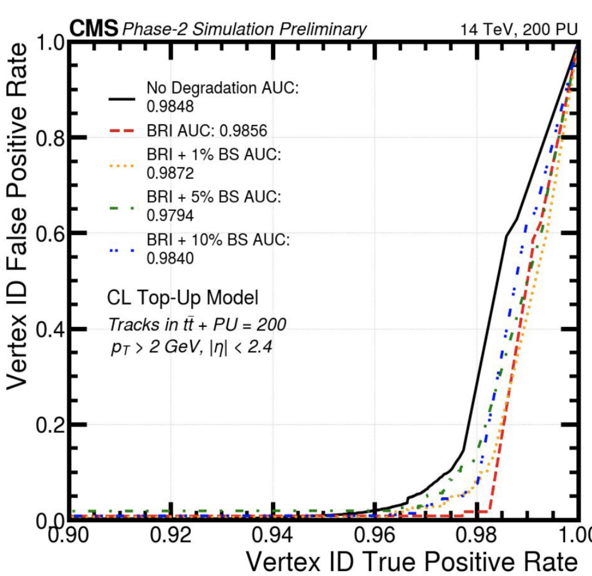
Continual learning

Useful in **online environment and changing conditions**. Avoid retraining. Strong computational constraints. Proposed lightweight alternative to SGD.



Embedded Continual Learning for HEP, CHEP2023



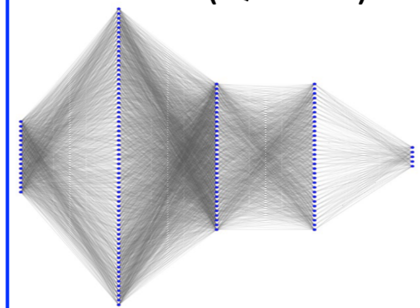


The Deployment of Realtime ML in Changing Environments, CHEP2024

← CMS L2 trigger Vertex finding

Symbolic regression

Baseline (QAT NN)



vs.

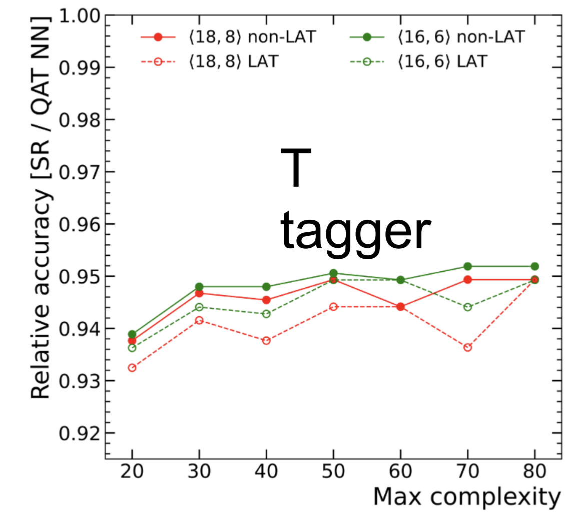
SR (5-line expressions)

Tagger	Expression for the trigonometric model with $c_{\max} = 20$	AUC
g	$\sin(-2C_1^{\beta=1} + 0.31C_1^{\beta=2} + m_{\mathrm{mMDT}} + \mathrm{Multiplicity} - 0.09\mathrm{Multiplicity}^2 - 0.79)$	0.897
q	$-0.33(\sin(m_{\mathrm{mMDT}}) - 1.54)(\sin(-C_1^{\beta=1} + C_1^{\beta=2} + \mathrm{Multiplicity}) - 0.81)\sin(m_{\mathrm{mMDT}}) - 0.81$	0.853
t	$\sin(C_1^{\beta=1} + C_1^{\beta=2} - m_{\mathrm{mMDT}} + 0.22(C_1^{\beta=2} - 0.29)(-C_1^{\beta=2} + C_2^{\beta=1} - \mathrm{Multiplicity}) - 0.68)$	0.920
W	$-0.31(\mathrm{Multiplicity} + (2.09 - \mathrm{Multiplicity})\sin(8.02C_1^{\beta=2} + 0.98)) - 0.5$	0.877
Z	$(\sin(4.84m_{\mathrm{mMDT}}) + 0.59)\sin(m_{\mathrm{mMDT}} + 1.14)\sin(C_1^{\beta=2} + 4.84m_{\mathrm{mMDT}}) - 0.94$	0.866

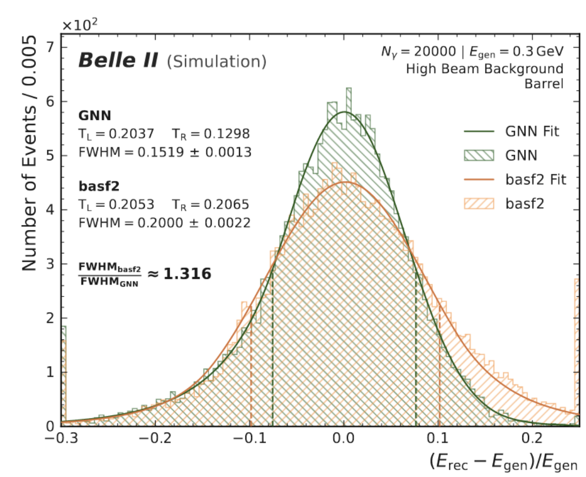
Table 1. Expressions generated by PySR for the trigonometric model with $c_{\max} = 20$. Operator complexity is set to 1 by default. Constants are rounded to two decimal places for readability. Area under the receiver operating characteristic (ROC) curve, or AUC, is reported.

SR models dramatically reduce latency and resources compared to NN

Symbolic Regression on FPGAs for Fast Machine Learning Inference, Thu 11/05



Improved Clustering in the Belle II Electromagnetic Calorimeter with Graph Neural Networks, Thu 11/05



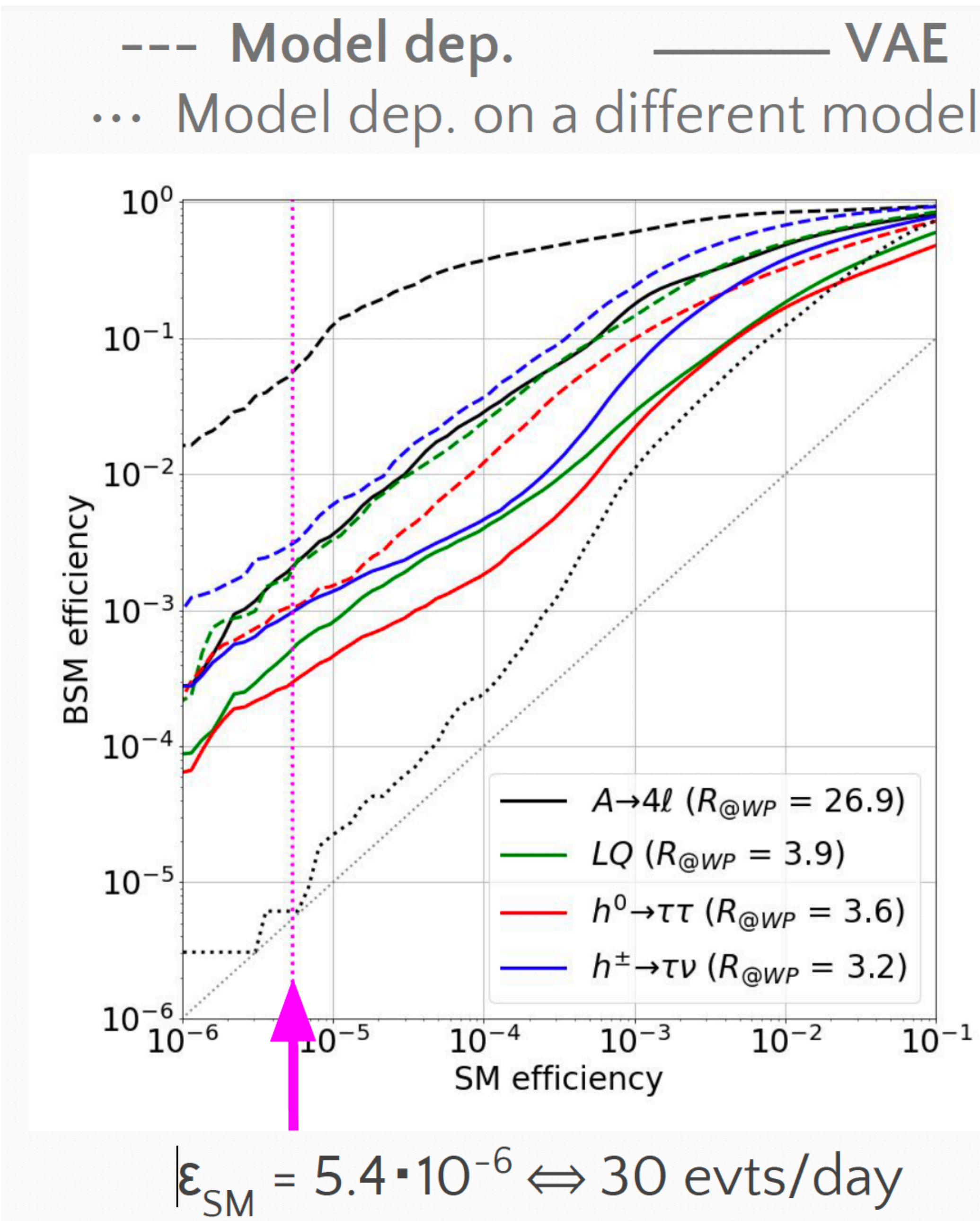
More **GNN**, **Object condensation** and Fuzzy Clustering in Belle II!

Selecting the unknown

A GRAPHIC REPRESENTATION OF DATA (…) UNTHINKABLE COMPLEXITY

William Gibson

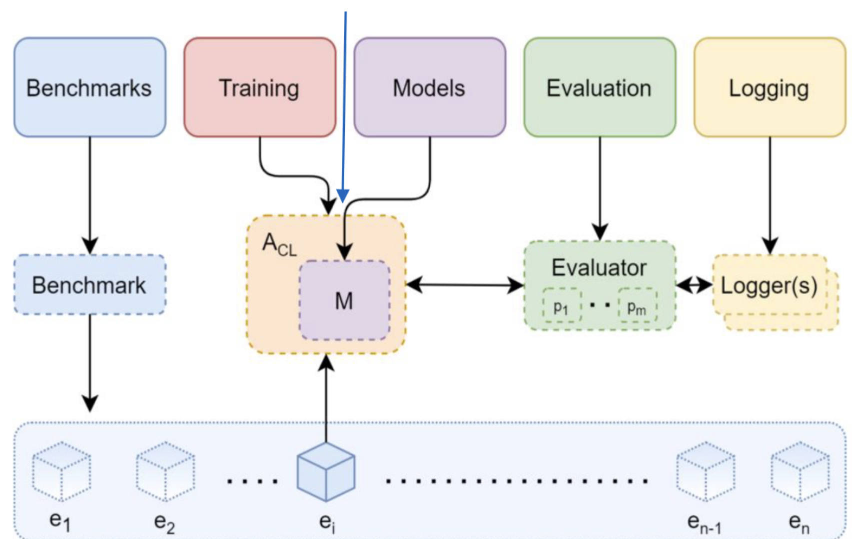
Unsupervised and model independent tools for new physics searches

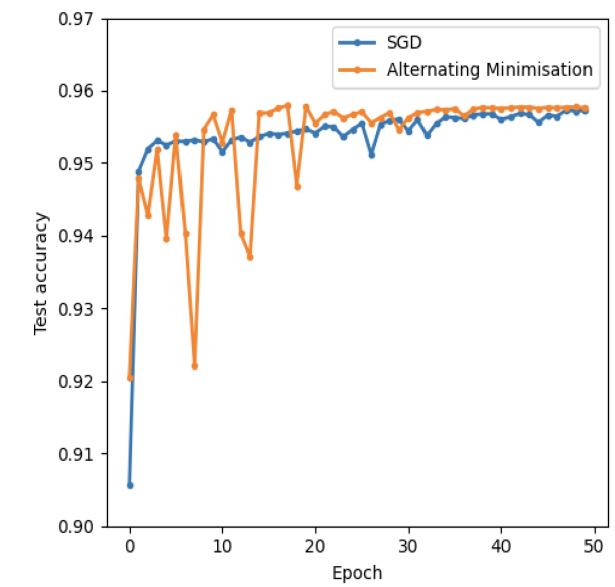


Arxiv:1811.10276. Evolved into:
Knapp, Oliver, et al. "**Adversarially Learned Anomaly Detection on CMS Open Data: re-discovering the top quark.**" *The European Physical Journal Plus* 136.2 (2021): 236.

Continual learning for online environment

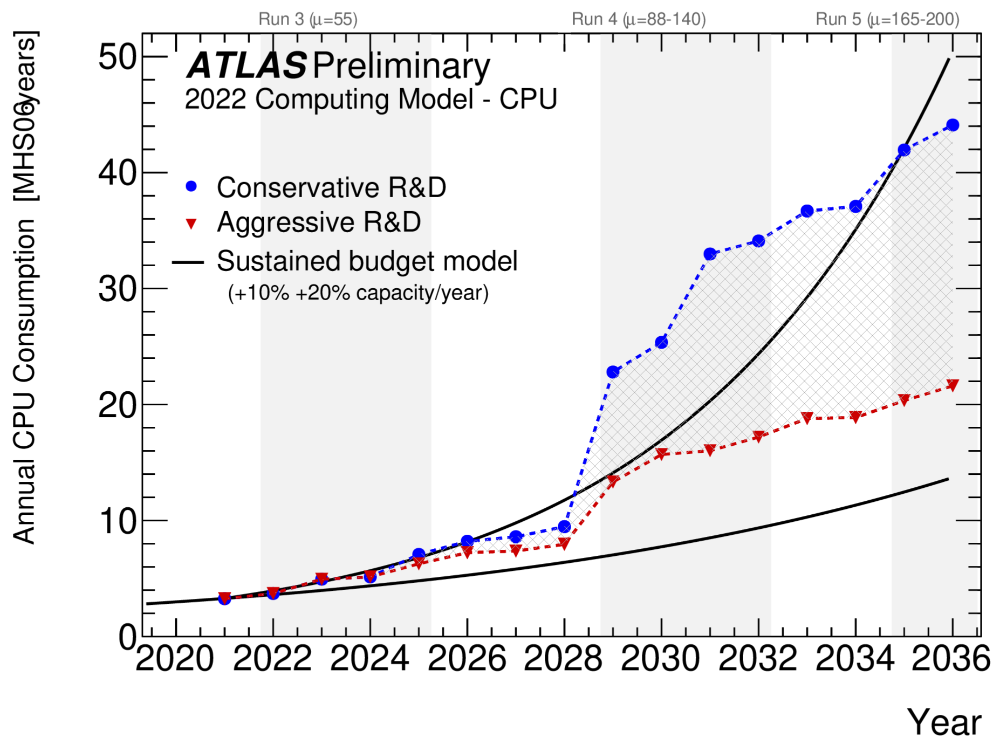
Useful with **changing conditions**. Avoid retraining. Strong computational constraints. Proposed lightweight alternative to SGD.

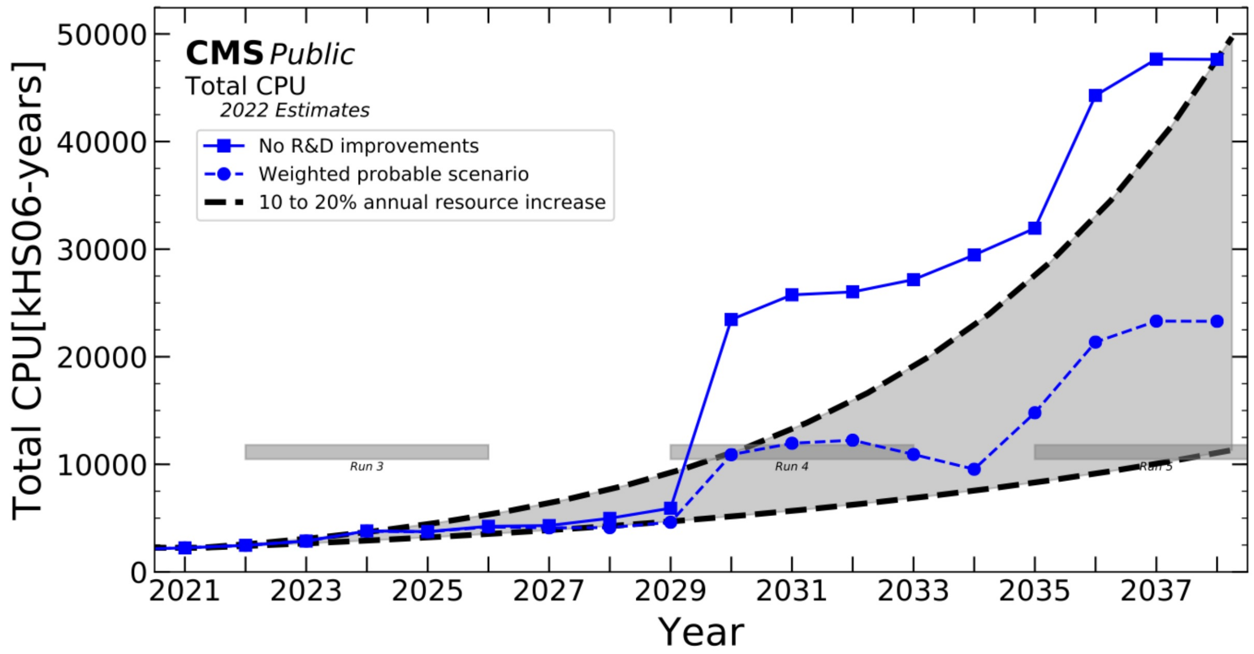


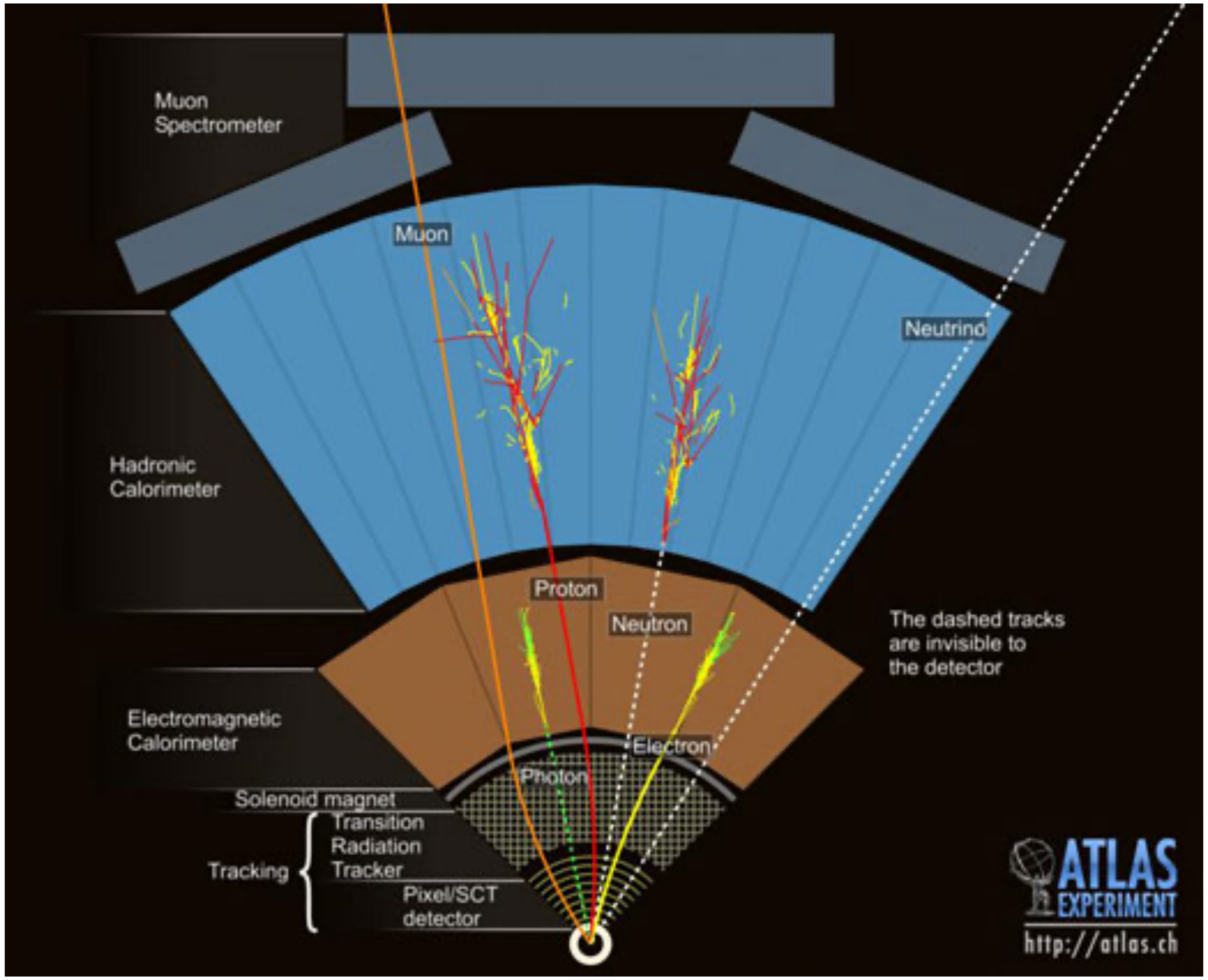


Embedded Continual Learning for HEP, CHEP2023

Offline processing (reconstruction) challenges

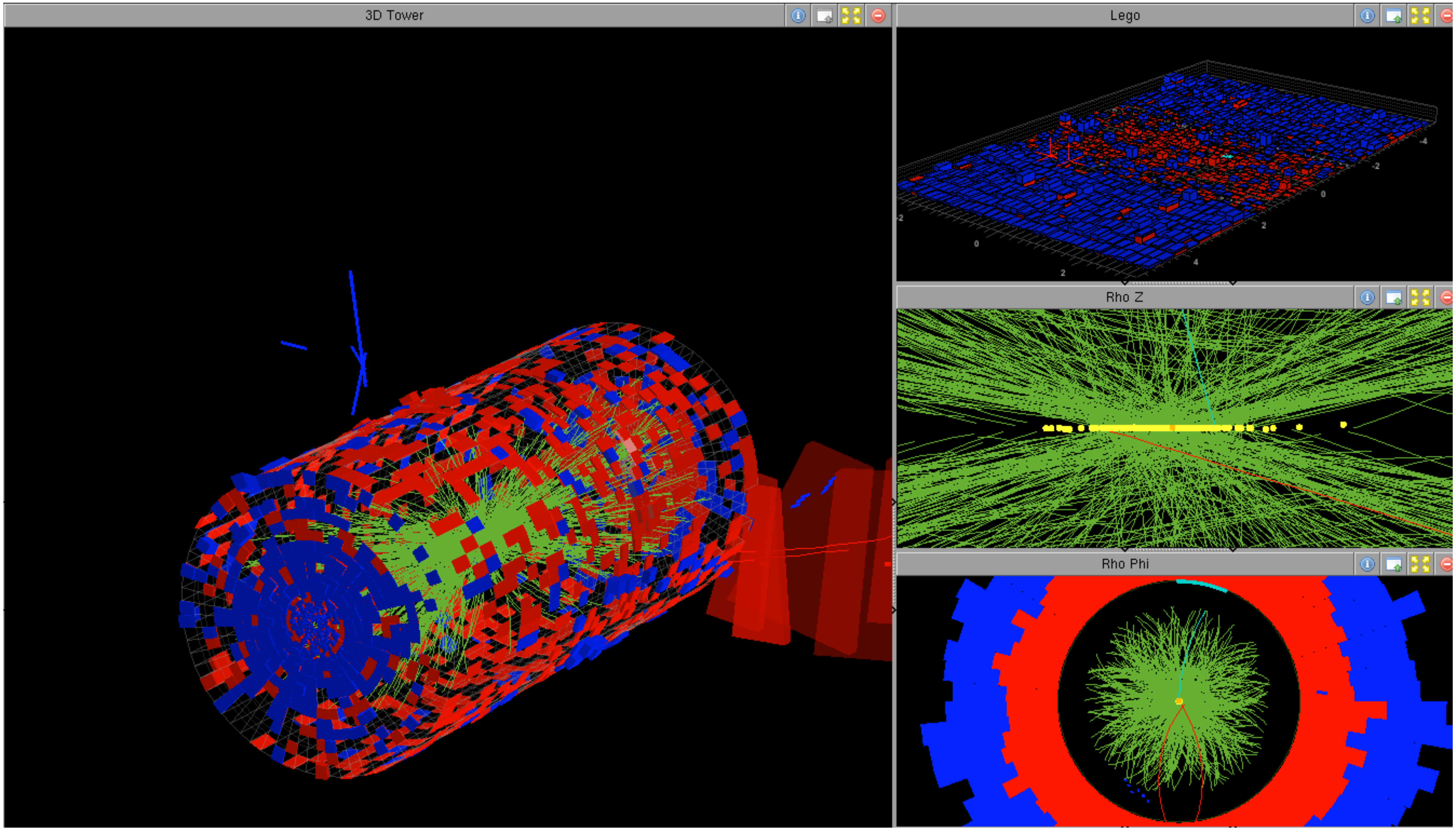






Sofia Vallecorsa, Ilaria Luise CERN - sofia.vallecorsa@cern.ch | ilaria.luise@cern.ch

Tracking



Exa.TrkX

https://exatrkx.github.io/

reports, publications & presentations

Publications

- Performance of a Geometric Deep Learning Pipeline for HL-LHC Particle Tracking (Associated Code) Eur. Phys. J. C **81**, 876 (2021)

Conference Contributions

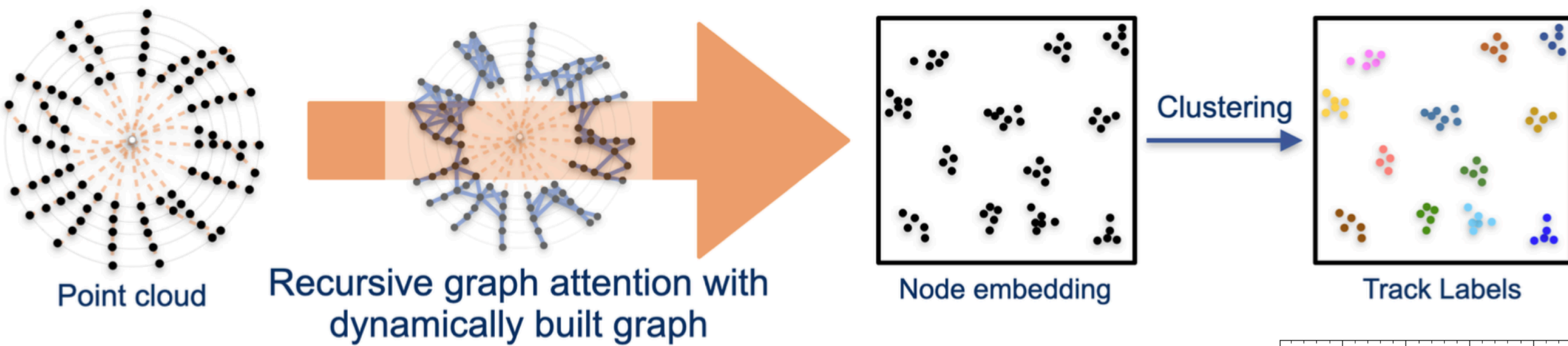
- EggNet: An Evolving Graph-based Graph Attention Network for Particle Track Reconstruction (To be) Presented at CHEP 2024.
- TrackSorter: A Transformer-based sorting algorithm for track finding in High Energy Physics Presented at ACAT 2024 (Associated Code).
- Influencer Loss: End-to-end Geometric Representation Learning for Track Reconstruction Presented at CHEP 2023.
- Physics Performance of the ATLAS GNN4ITk Track Reconstruction Chain Presented at Connecting the Dots 2023 (Associated Code).
- Graph Neural Network-based Tracking as a Service Presented at Connecting the Dots 2023.
- A Language Model for Particle Tracking Presented at Connecting the Dots 2023.
- Hierarchical Graph Neural Networks for Particle Track Reconstruction Presented at ACAT 2022 (Associated Code).
- ATLAS ITk Track Reconstruction with a GNN-based pipeline Presented at Connecting the Dots 2022 (Associated Code).
- Accelerating the Inference Time of Machine Learning-based Track Finding Pipeline Presented at ACAT 2021 (Associated Code).
- Graph Neural Network for Large Radius Tracking Presented at ACAT 2021
- Graph Neural Network for Object Reconstruction in Liquid Argon Time Projection Chambers Presented at CHEP 2021
- Distributed Training of GNNs on HPCs Presented at the 4th Inter-experiment Machine Learning Workshop (Associated Code).
- "Track Seeding and Labelling with Embedded-space Graph Neural Networks". Presented at Connecting the Dots 2020 - (Associated Code).
- "Graph Neural Networks for Particle Reconstruction in High Energy Physics Detectors". Presented at NeurIPS 2019 Workshop "Machine Learning and the Physical Sciences" - (NeurIPS Poster) (Associated Code).

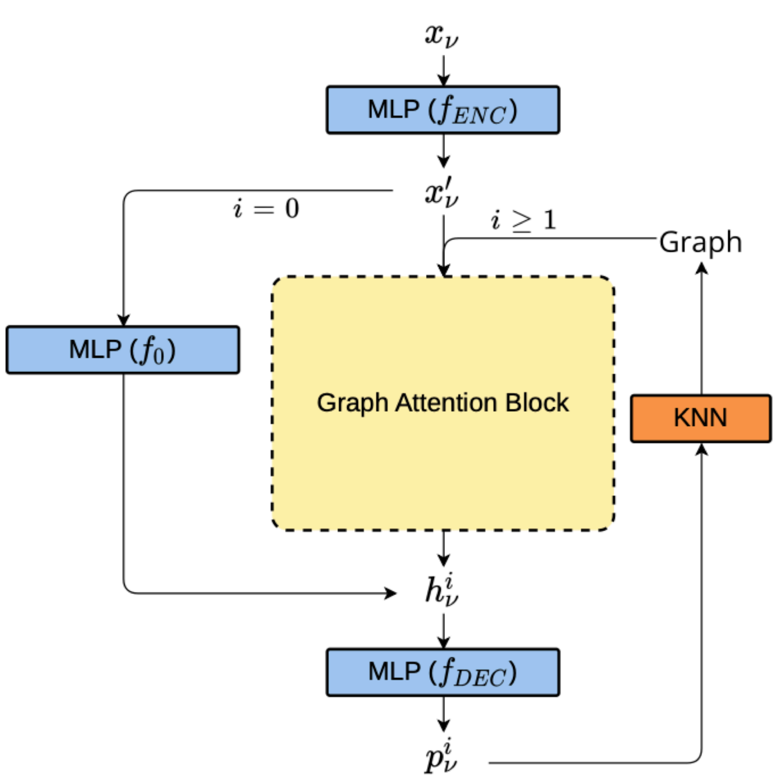
Presentations

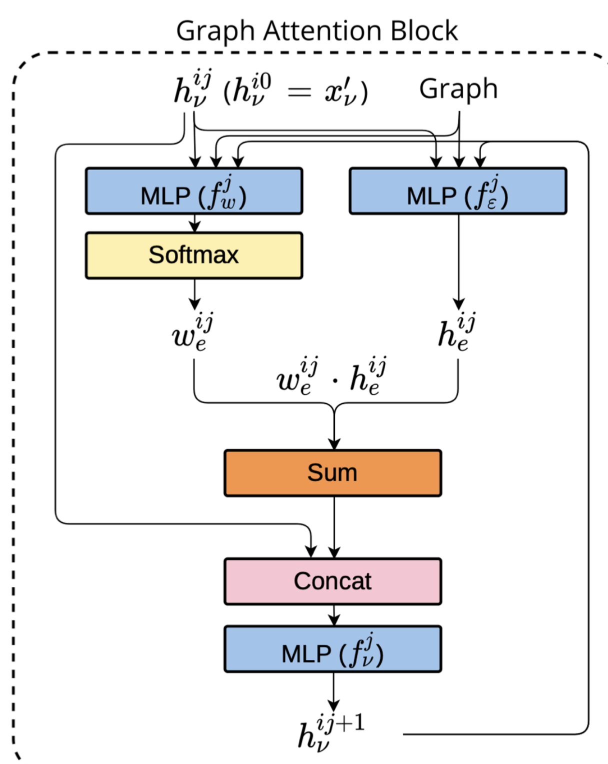
- Full-length tutorial on Tracking with Graph Neural Networks (Sep 2023, Heidelberg) Part 1 - Part 2
- Graph Neural Networks for High Luminosity Track Reconstruction (CERN EP-IT Data science seminar).
- Graph Neural Networks for Reconstruction in DUNE (presented at the Dec 4th CLARIPHY topical meeting).
- Tracking with GNNs (in-depth code walk-through at the 4th Inter-experiment Machine Learning Workshop) (colab notebook) (Associated Code).
- Graph Neural Networks for Particle Tracking (A non-specialist introduction to Exa.TrkX tracking models).

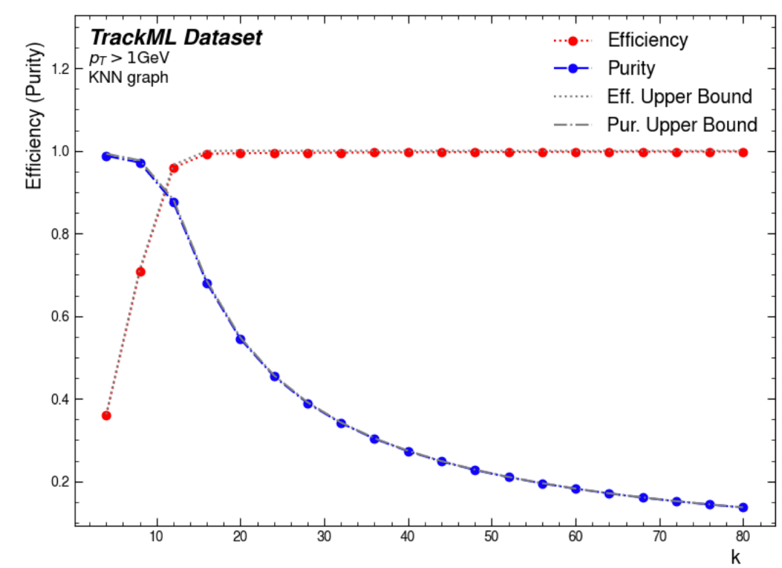
Exa.TrkX: EggNet (a Graph Attention Network)

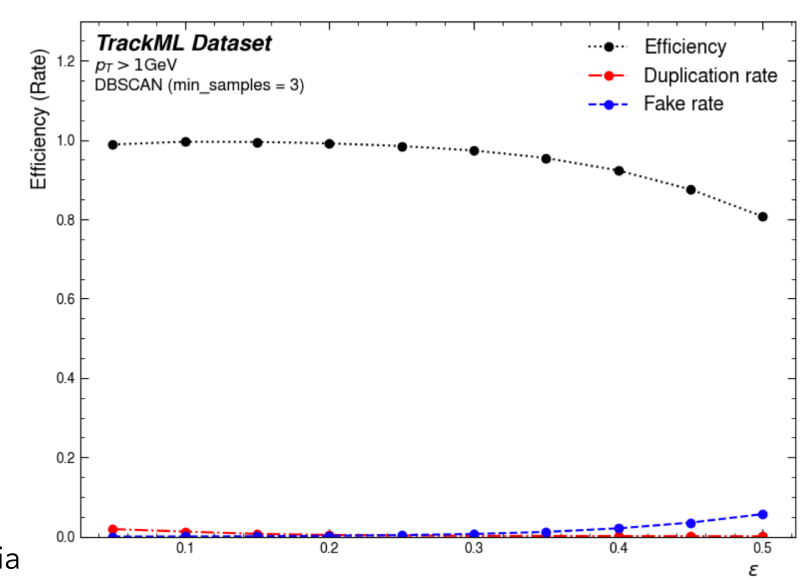
[EggNet: An Evolving Graph-based Graph Attention Network for Particle Track Reconstruction](#) (To be) Presented at [CHEP 2024](#).











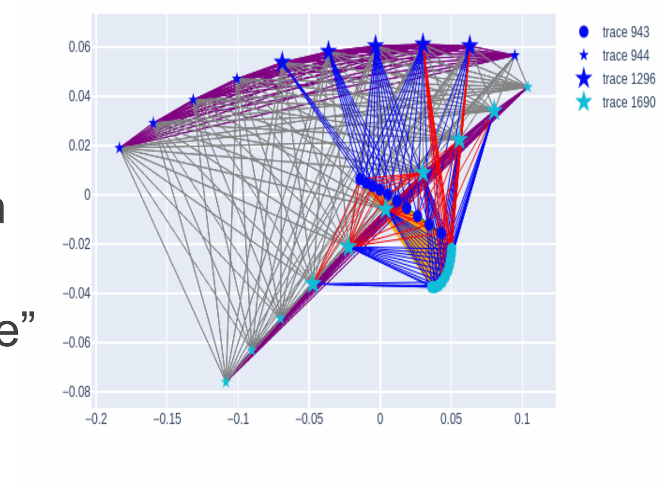
Sofia Vallecorsa, Ilaria Luise CERN - sofia.vallecorsa@cern.ch | ilaria

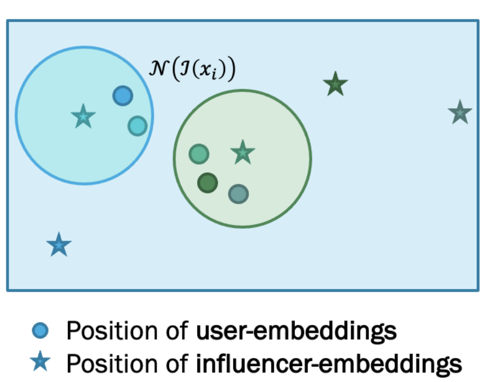
Tracking (highlights from CHEP2023)

Tracking as object condensation

An Object Condensation Pipeline for Charged Particle Tracking

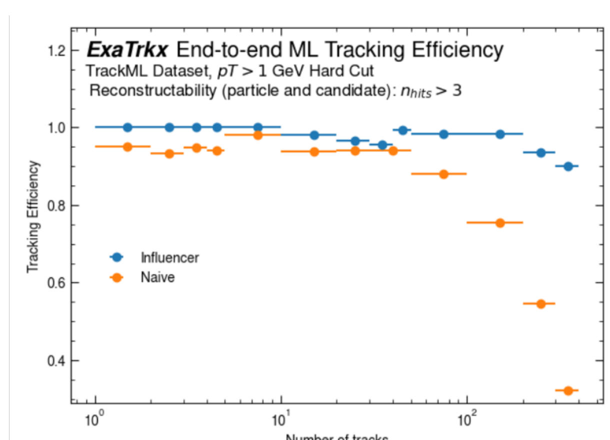
Simultaneously learn embedding similarity space **and** condensation score per hit (a higher score is a more “attractive” point charge in similarity space





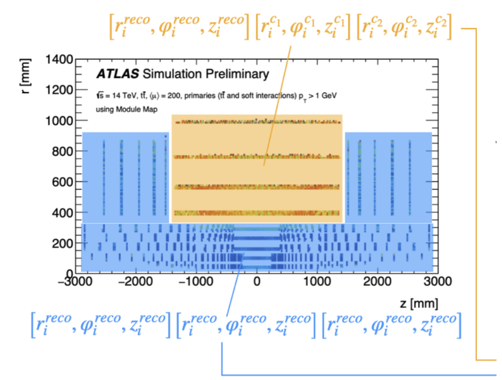
End-to-End Geometric Representation Learning for Track Reconstruction

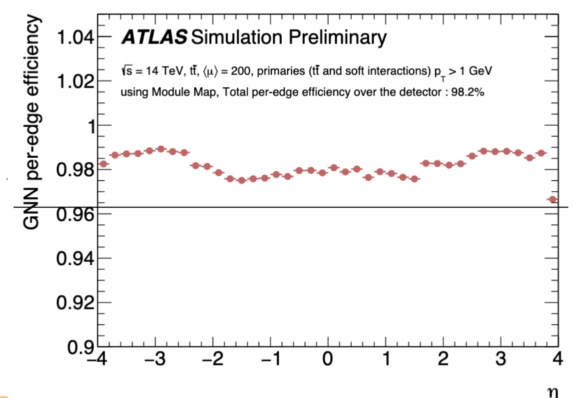
A new twitter inspired loss: **the influencer loss !**



Graph Neural Networks

Novel fully-heterogeneous GNN designs for track reconstruction at the HL-LHC





+ BESIII track reconstruction algorithm based on machine learning

Full ML pipeline for CLAS12 @JLAB

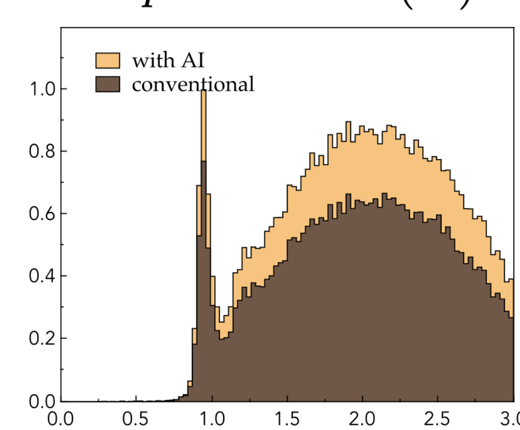
Tracking represents 80% of CLAS12 processing time

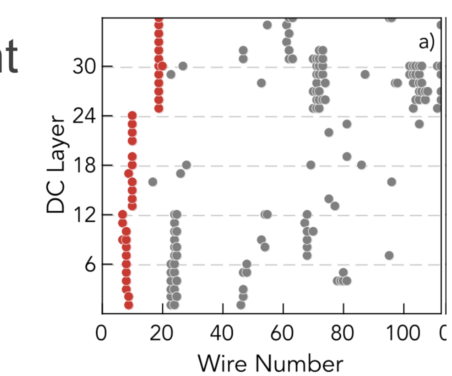
Train a MLP to classify tracks and and AE to account for missing hits

35% improvement

Track Identification for CLAS12 using AI

$$ep \to e'\pi^+\pi^-(X)$$





Develop denoising AE for high luminosity runs

Jets

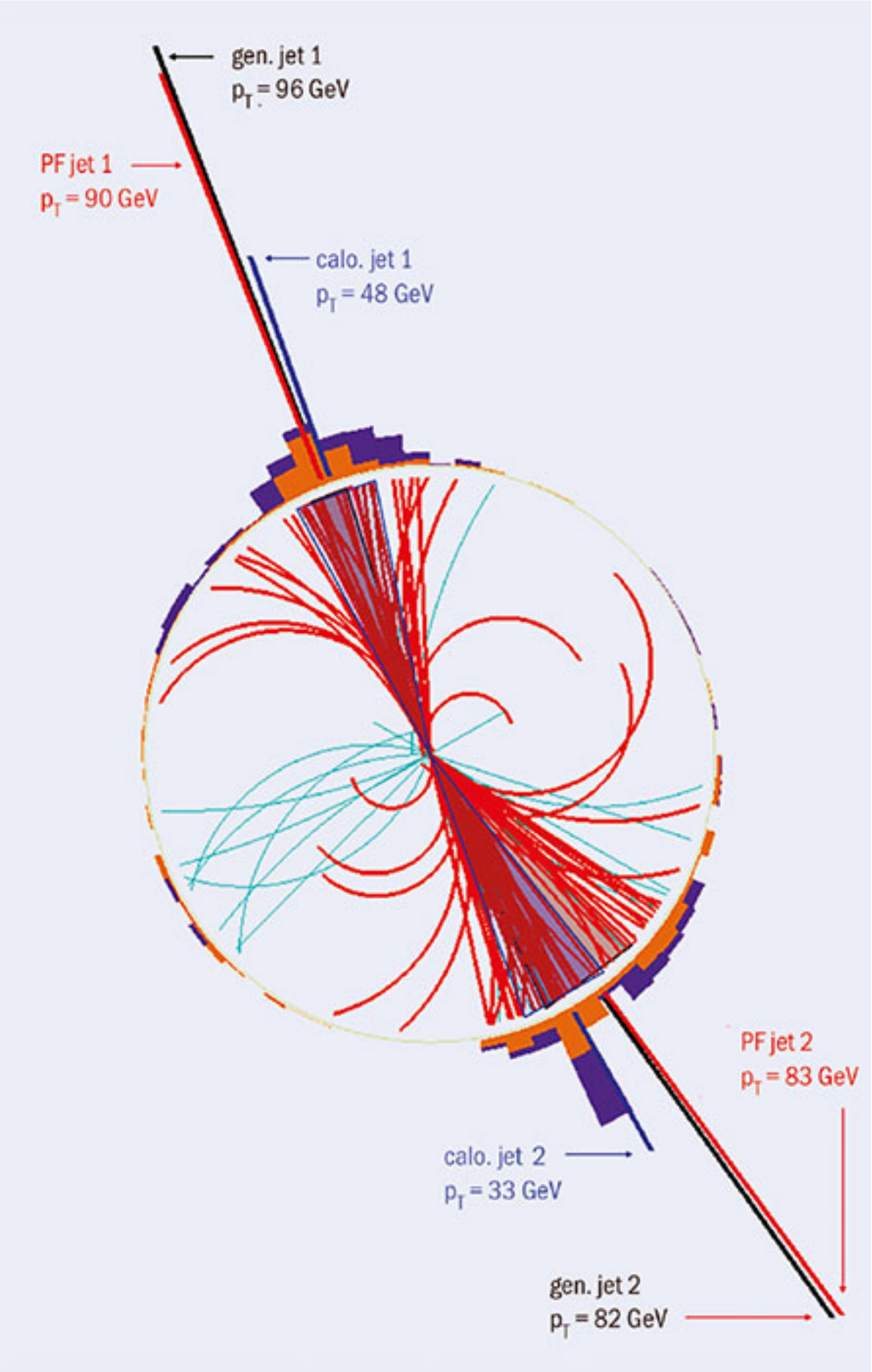
Jets represent a major area of applications for ML.
See ML4Jets https://indico.cern.ch/event/1253794/overview

- **Stable particles defined by MC generators**:"Truth jets".
- **Charged-particle tracks**: "Track Jets". Particularly useful for pile-up mitigation or jet tagging.
- **Calorimeter energy deposits**: "Topo Jets". Requires calorimeter cells clustering and calibration.
- **Combining tracks and energy deposits:** "Particle Flow jets" exploit the best of two very different calorimeters

  Tracks info is limited to charged-particles, while topo-clusters are built from both charged and neutral particles

  Angular resolution of the trackers is "still" better than calorimeters (important for vertex finding). Calorimeter also extend pseudo rapidity coverage.

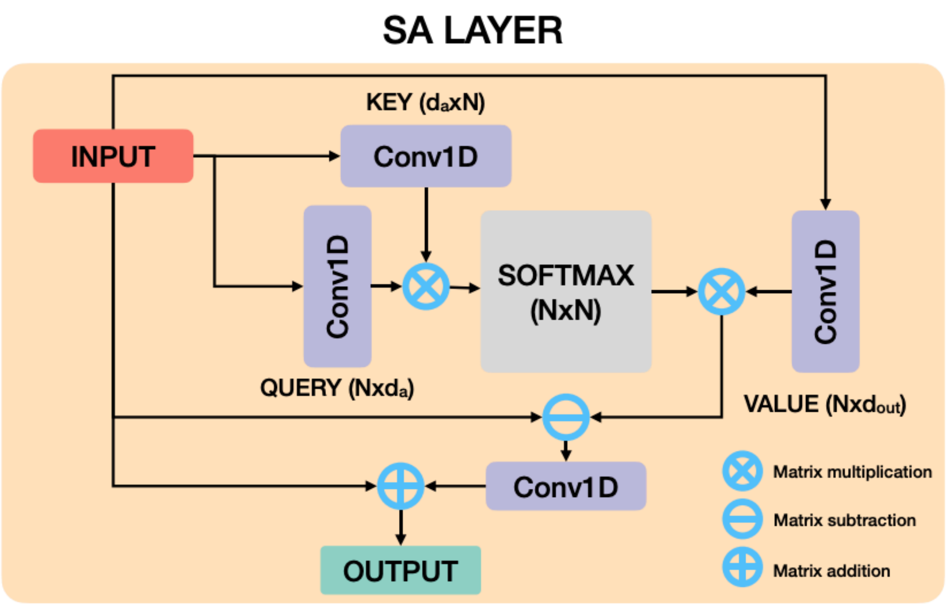
  Inner detector momentum resolution is best for low energy tracks. Energy resolution of the calorimeters is better than the momentum resolution of the inner detector.

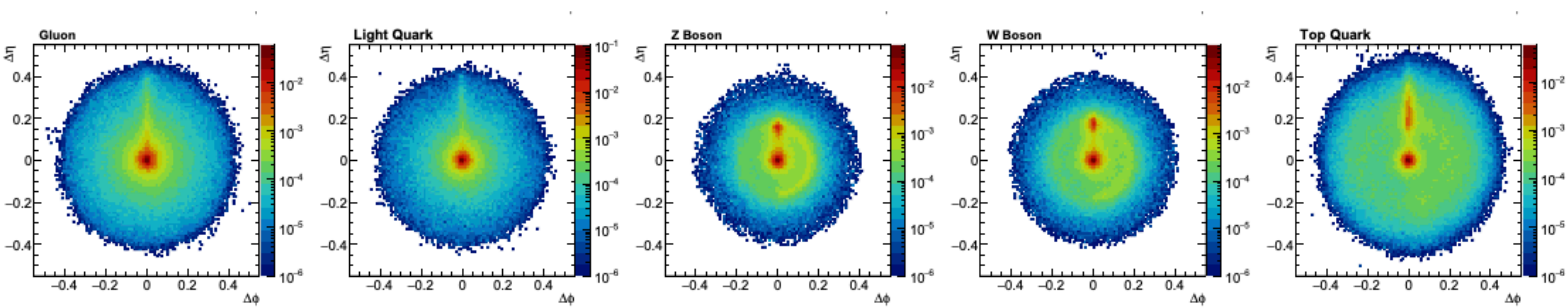


Point Cloud Transformers

Self Attention on point-cloud particle data learns "semantics"

- SA layers extract **different information** for each jet (jet sub-structure)
- **Increased relevance to harder sub-jets** in the case of Z boson, W boson, and top quark initiated jets.
- Light quark and gluon jets have **homogeneous radiation pattern**



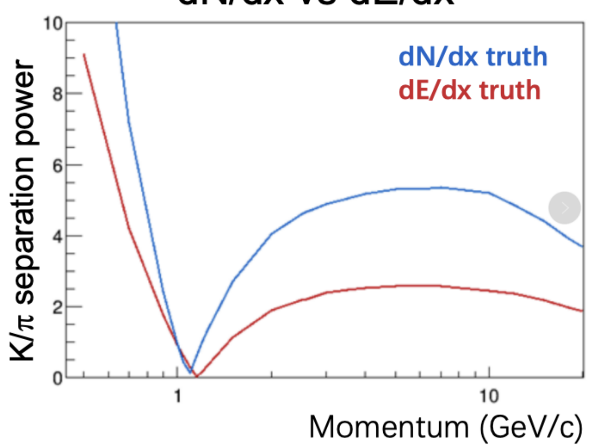


https://iopscience.iop.org/article/10.1088/2632-2153/ac07f6/meta

Reconstruction (Highlights from CHEP2023)

Hybrid RNN/CNN for robust PID based on dN/dx cluster counting in drift chambers

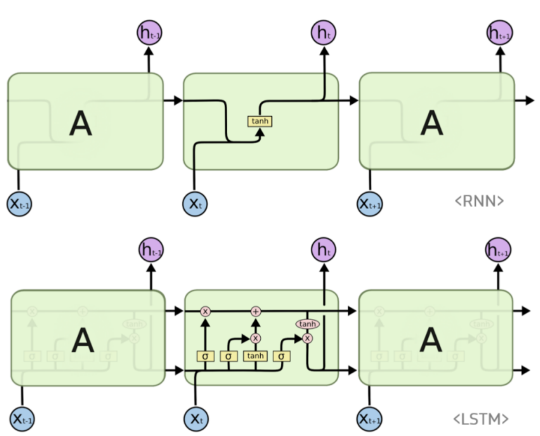
K/π separation power
dN/dx vs dE/dx

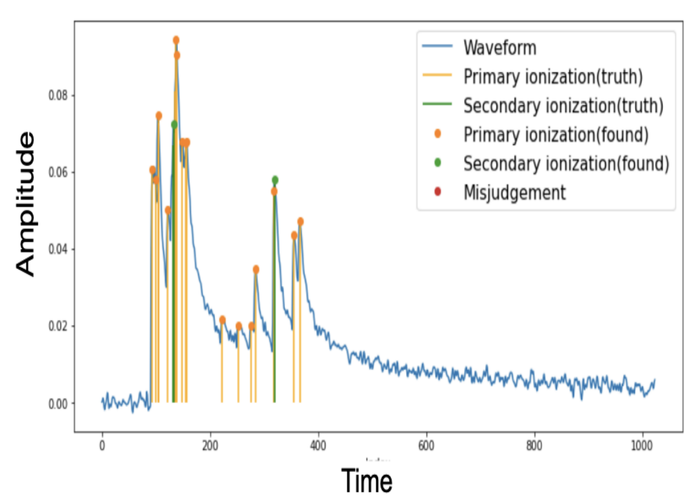


A Deep-Learning Reconstruction Algorithm for Cluster Counting

RNN for peak finding and **CNN for peak clustering** determine the number of clusters per particle trajectory.

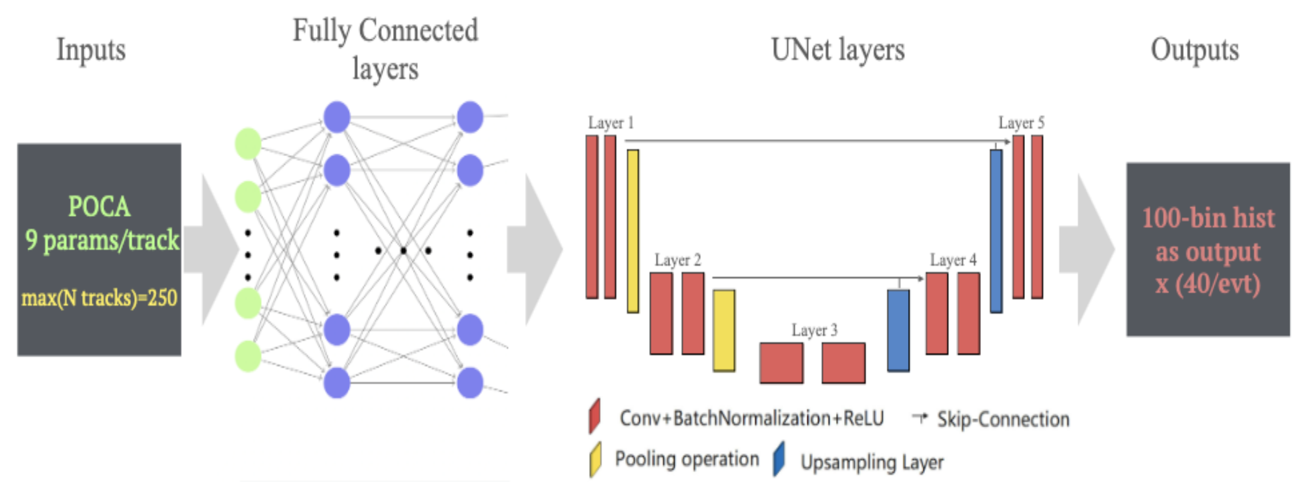
RNN (Recurrent Neural Network)



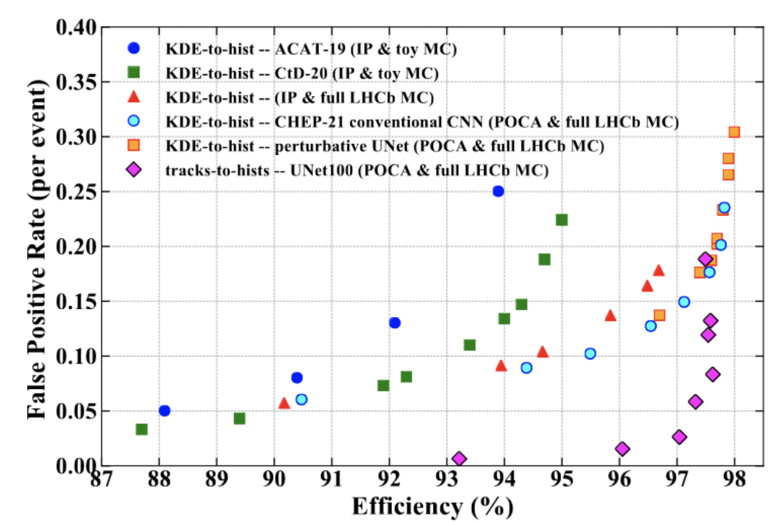


Improved **Primary Vertex** finding @ATLAS and LHCb

Advances in developing deep neural networks for finding primary vertices in proton-proton collisions at the LHC. Tue



UNet based architecture in LHCb

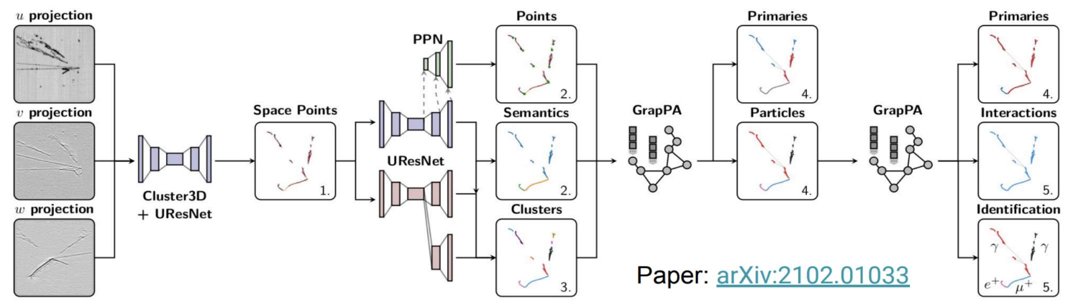


(Beyond Colliders) LAr TPC imaging

DEEP UNDERGROUND NEUTRINO EXPERIMENT

End-to-end reconstruction pipeline integrates various state-of-the art ML architectures.

- **UResNet** for pixel feature extraction, **GrapPA** for superstructure formation
- Used on **ICARUS** sim./data and **DUNE-ND** (high neutrino pileup) sim. **today**
- Check out this ICARUS interactive reconstructed event !



Paper: arXiv:2102.01033

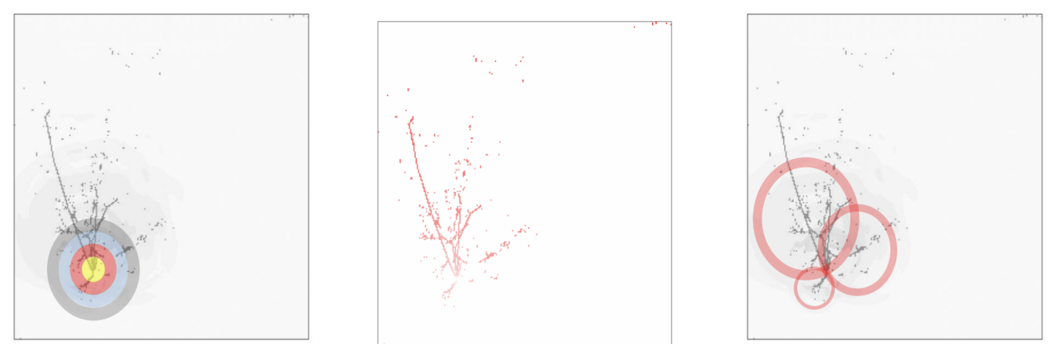
End-to-end, ML-based Reconstruction Chain for Particle Imaging Detectors

Pandora: hybrid Deep Learning + algorithmic pattern recognition outperforms previous binary decision tree algorithm.

In training hits are assigned a class according to distance from true vertex

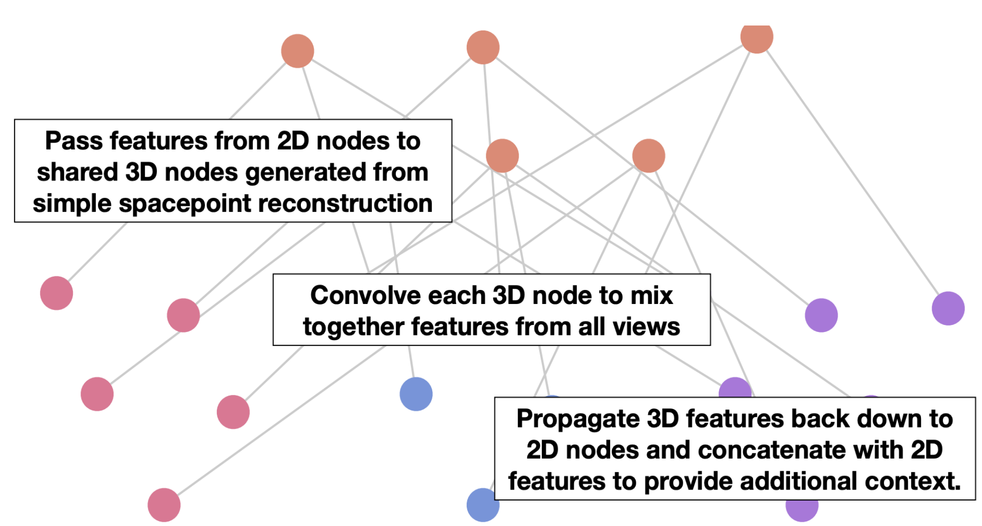
Network trained to learn those distances from input images

Network infers hit distances and resultant heat map isolates candidate vertex



Neutrino interaction vertex-finding in a DUNE far-detector using Pandora deep learning

Nugraph2: 2nd generation of **hit labelling GNN from ExaTrk**
Performance **boosted through application of Nexus Convolutions** to multi-head attention message passing : **70% → 98% accuracy**

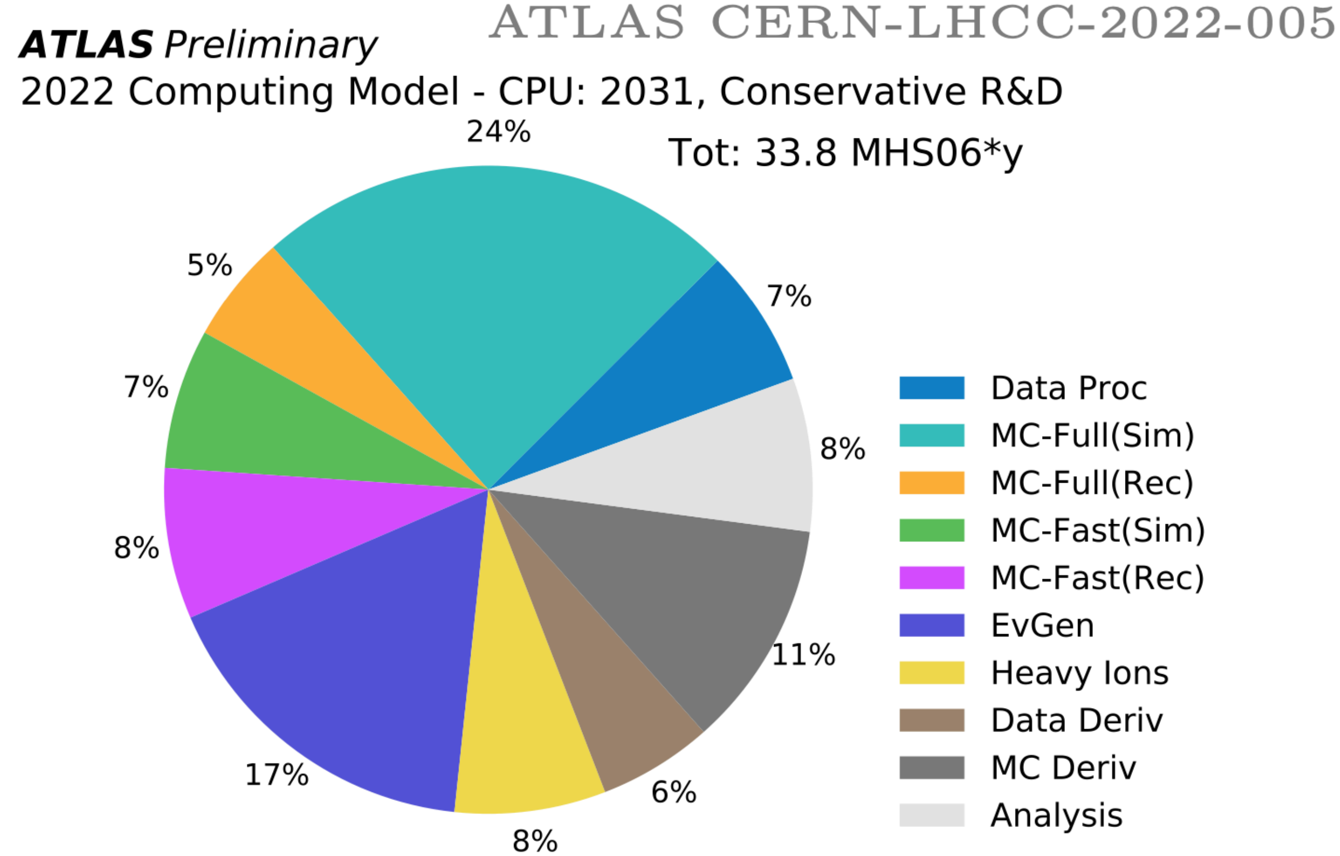


A Graph Neural Network for 3D Reconstruction in Liquid Argon Time Projection Chambers

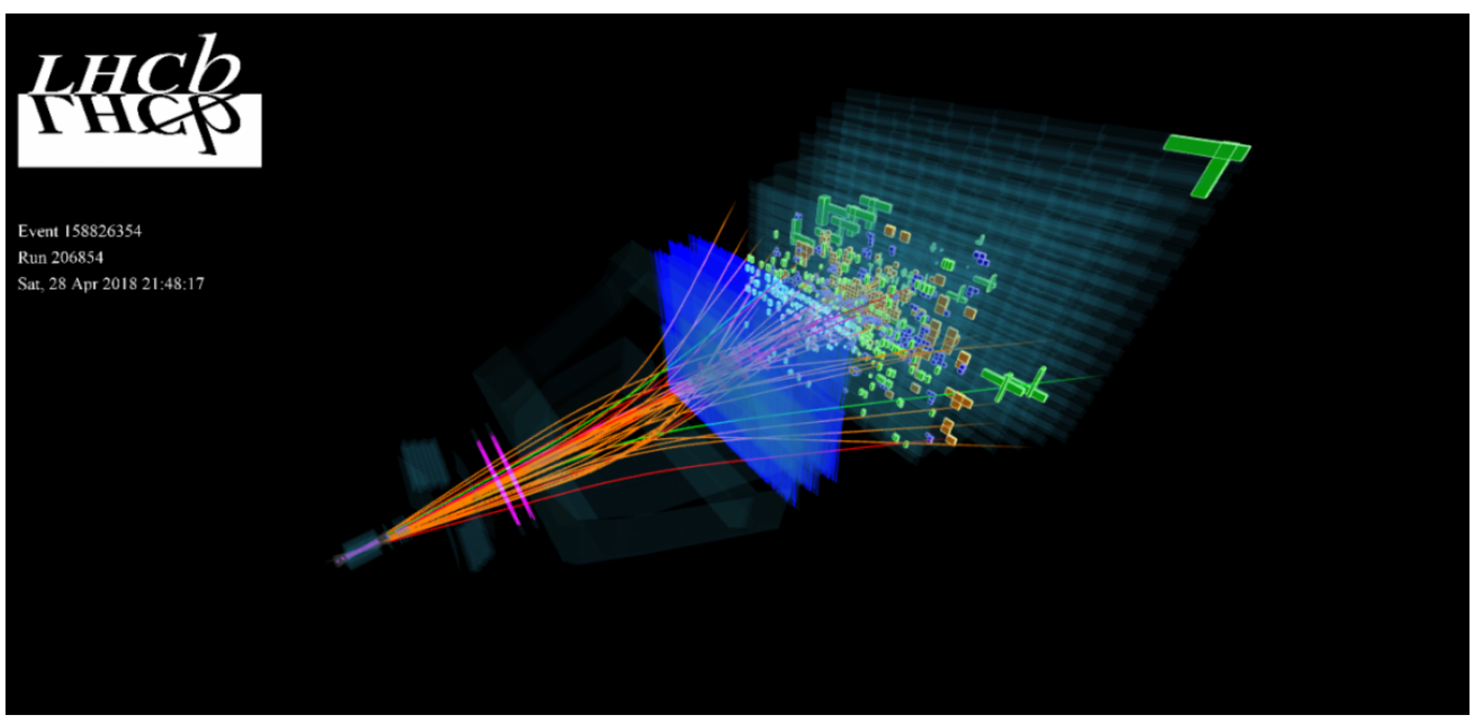
Monte Carlo Simulation

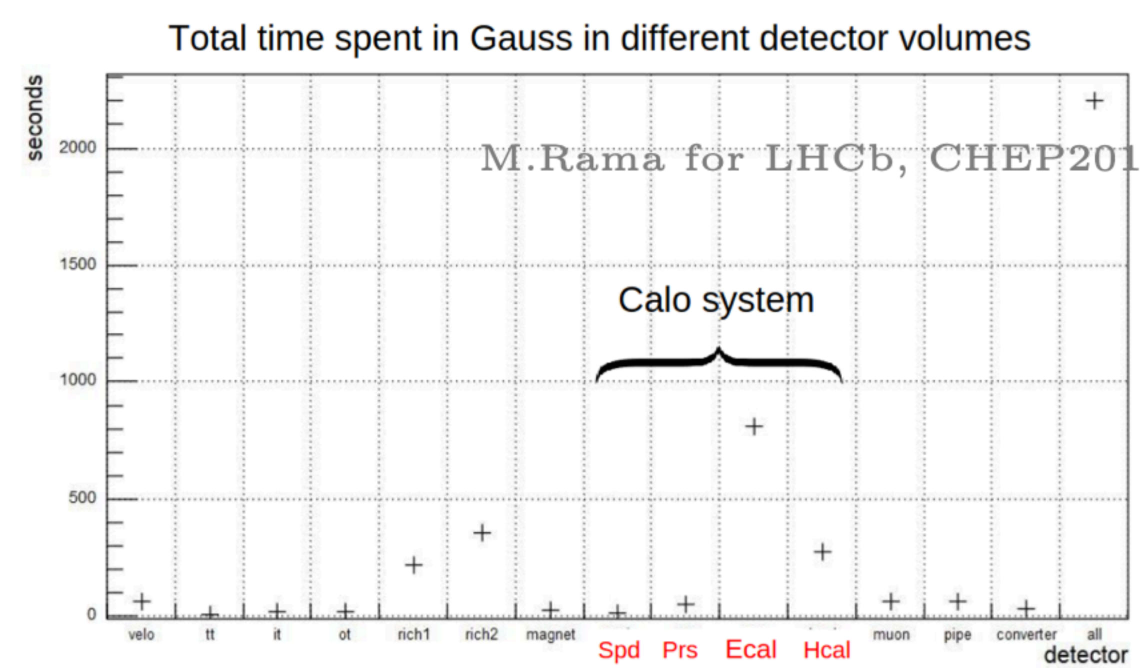
Monte Carlo computational costs

ATLAS CERN-LHCC-2022-005



Calorimeters!

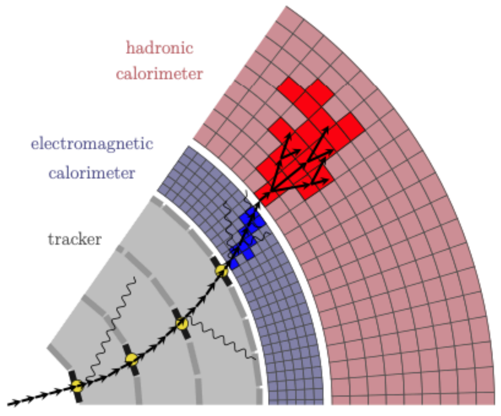




CPU time in calorimeter system: ~ **53%**
CPU time in RICH1+2: ~ **27%**

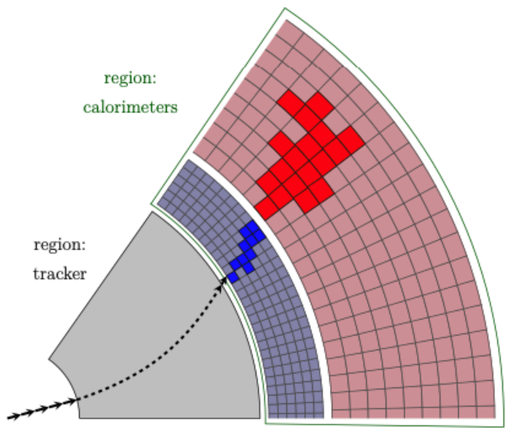
Fast detector simulation

Simulation of particle passage: full vs fast



detailed / "full" simulation
→ GEANT4

parameterisation / "fast" simulation
→ requires input



- detailed detector description (DD4hep⇒GEANT4)
- definitions of particles and processes
- transport in e-m field

- where particles are parametrised
- which particles
- how/what happens

Defining both 'full' and 'fast' simulation within one framework offers great flexibility to seamlessly mix both types.

Given the recent decision to turn towards DDG4 as the full sim framework, the old examples from k4SimGeant4 and native Geant4 examples need to be adapted, but the main principles remain.

Synthetic data generation through DL

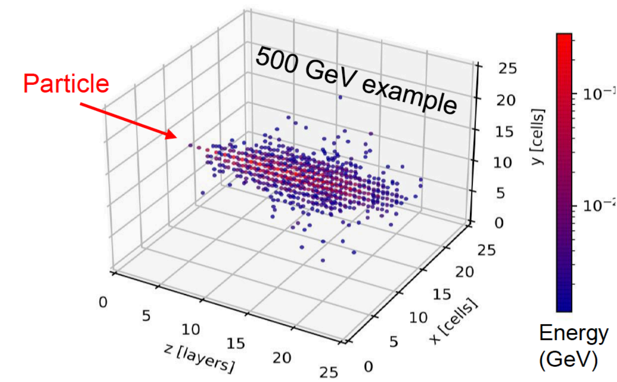
CYBERSPACE. A CONSENSUAL HALLUCINATION EXPERIENCED DAILY BY BILLIONS OF LEGITIMATE OPERATORS

William Gibson

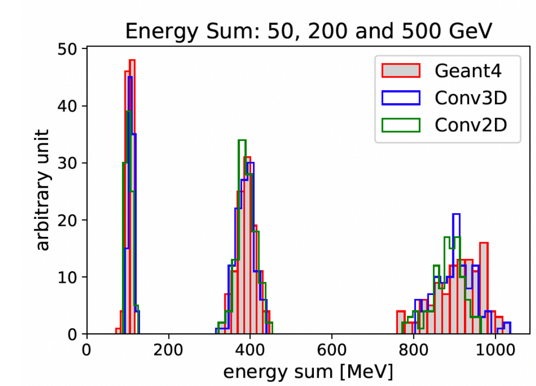
A major task, requiring high accuracy.

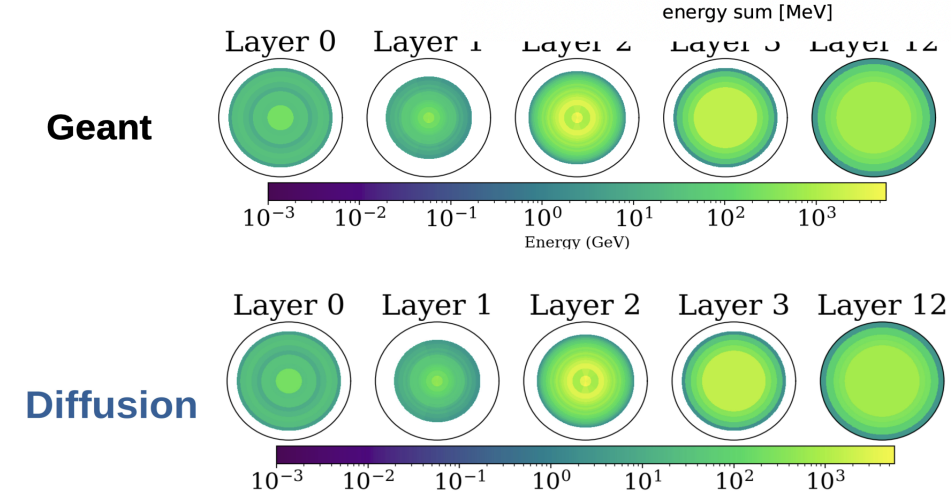
It is computationally expensive **(typically Monte Carlo based)**

Ideal task for state-of the-art generative AI



Rehm, Florian, et al. *arXiv:2105.08960* (2021).





Diffusion models for shower generation, CHEP2023

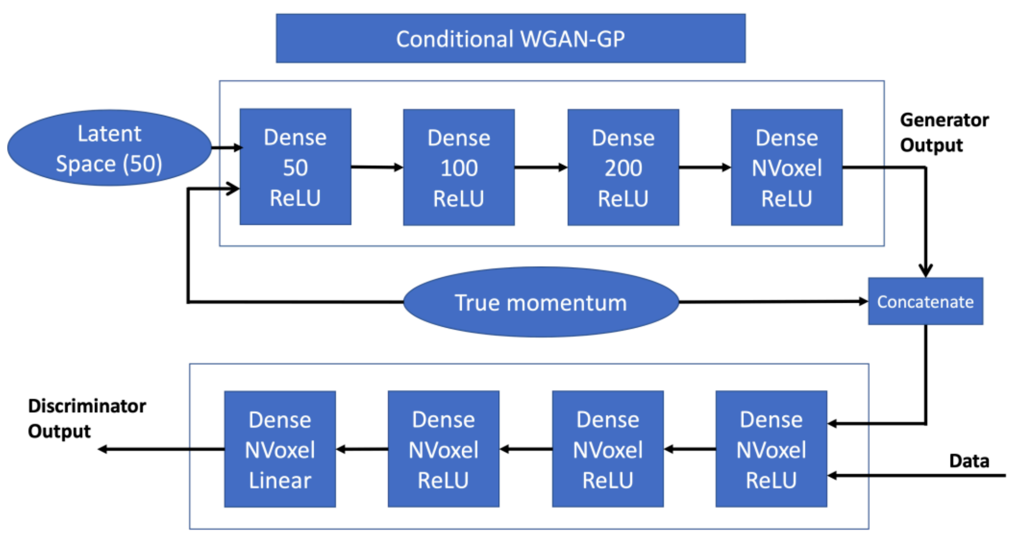
GAN FOR CALORIMETERS

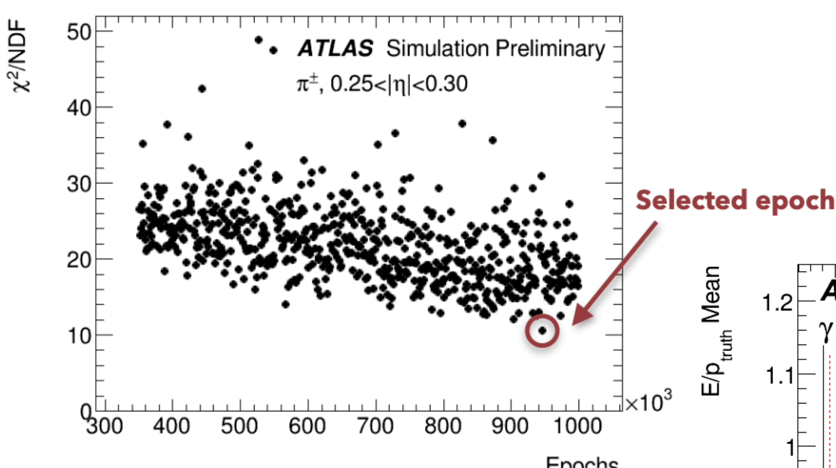
ATLAS

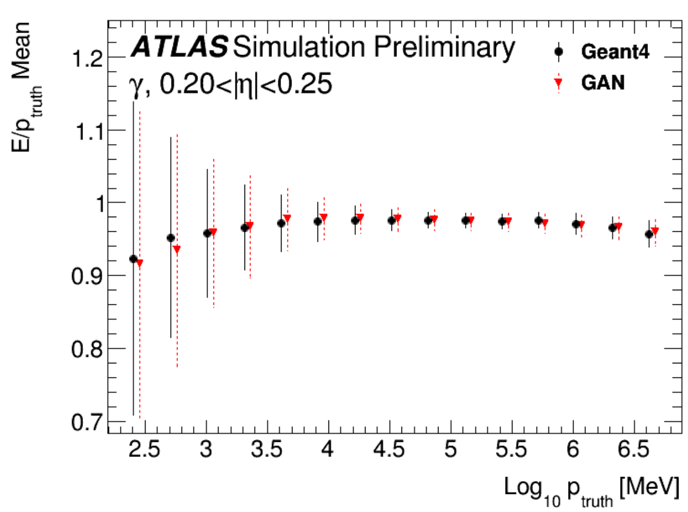
FastCaloGAN: 300 GANs

IN PRODUCTION

CHEP 2023



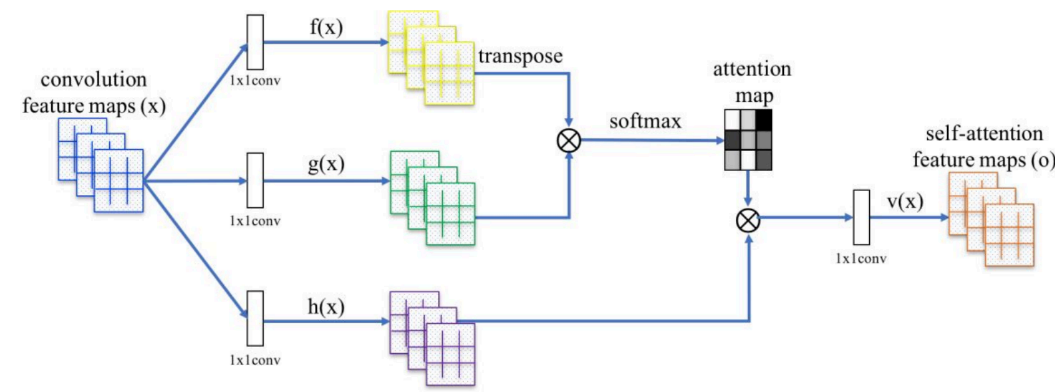


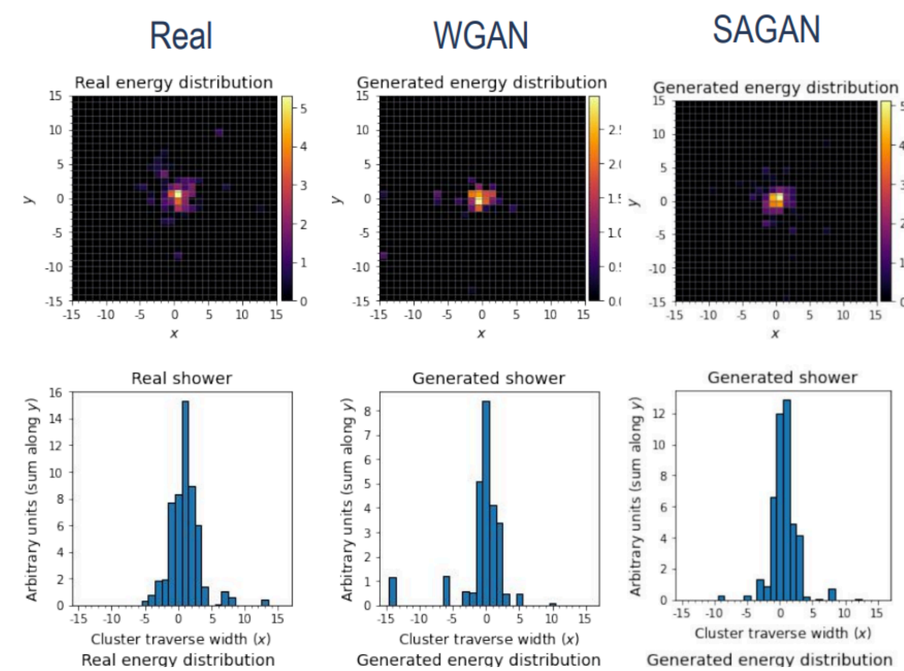


LHCb

Self-Attention GANs

F. Ratnikov, A. Rogachev: https://indico.cern.ch/event/948465/contributions/4324135





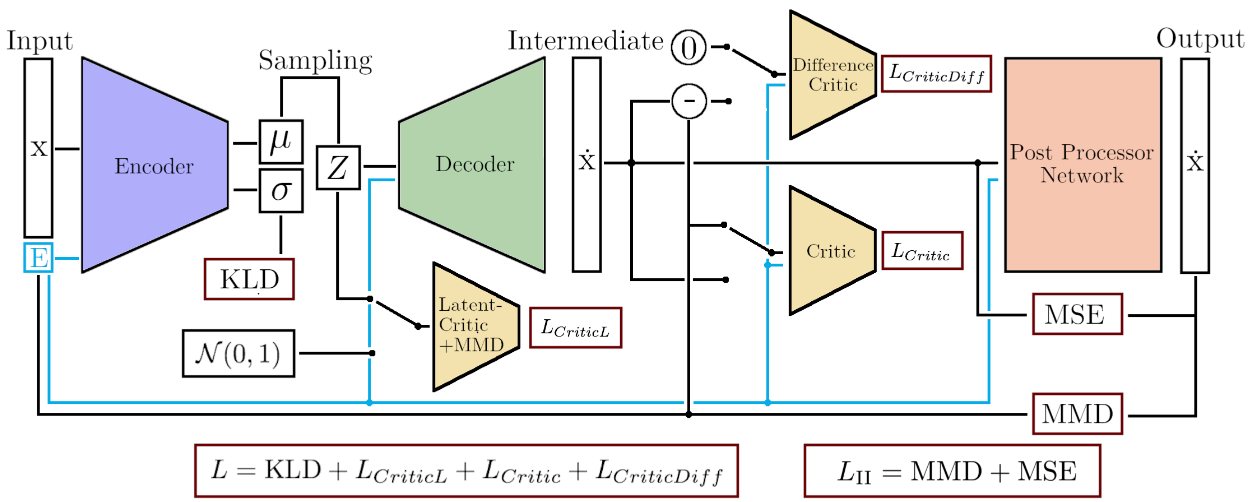
Model	Physics PRD-AUC	Raw Images PRD-AUC
WGAN	0.936	0.971
SAGAN+SN D	0.895	0.901
SAGAN+SN G and D	0.948	0.975

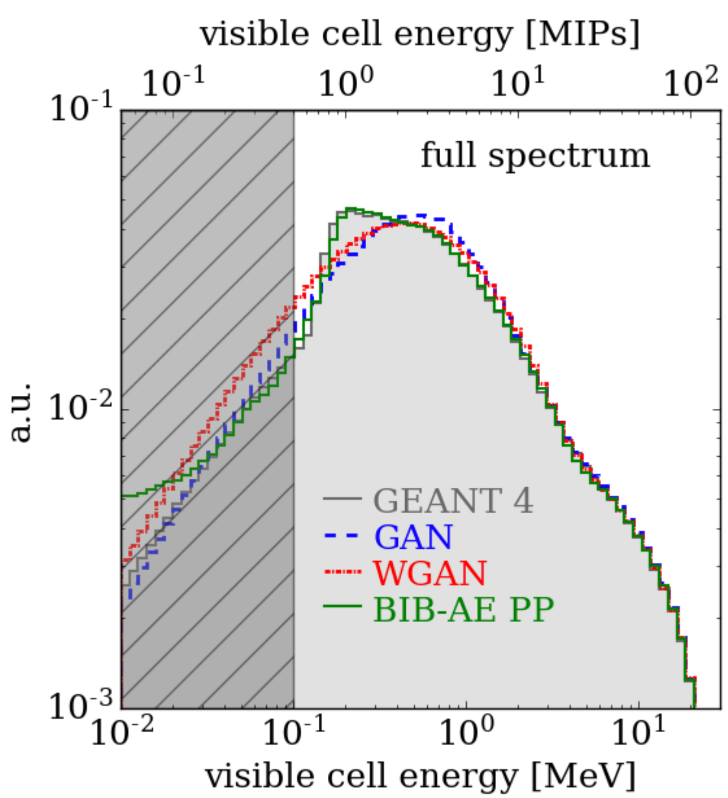
Zhang H. et al. Self-attention generative adversarial networks. – PMLR, 2019 C. 7354-7363.

Increasing complexity

GAN – AutoEncoder hybrid

CMS

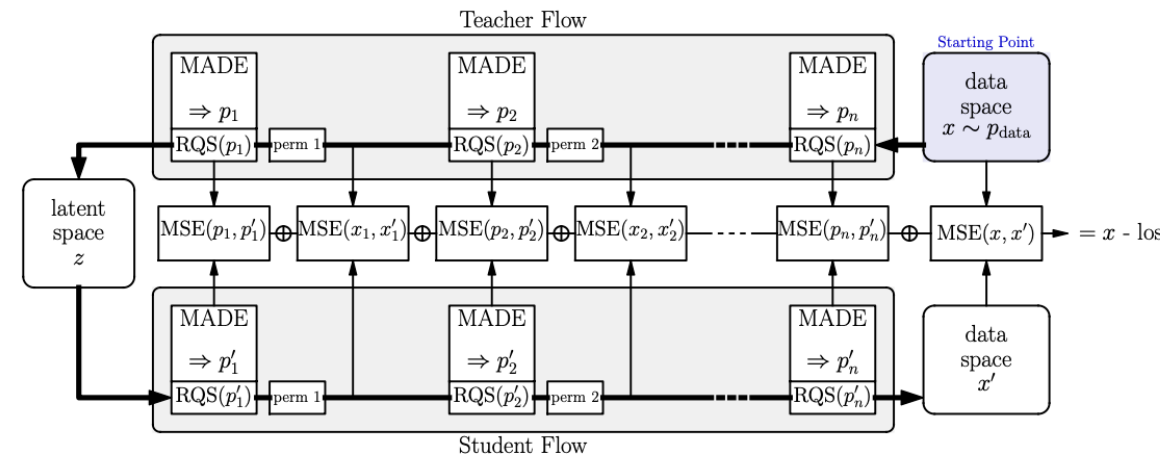


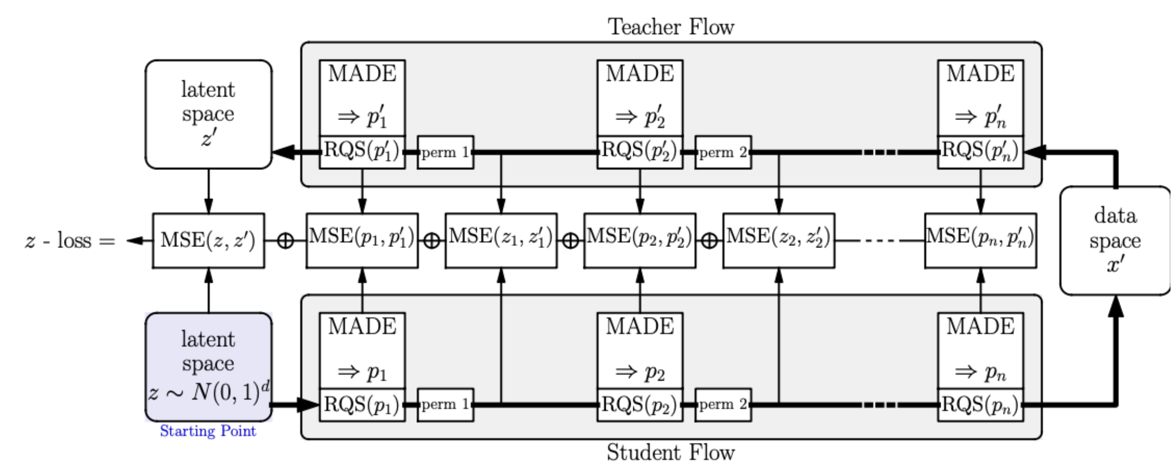


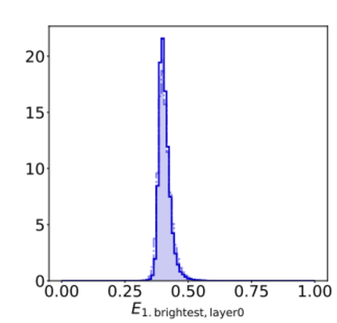
Buhmann, Erik, et al. "Getting high: high fidelity simulation of high granularity calorimeters with high speed." *Computing and Software for Big Science* 5.1 (2021): 1-17.

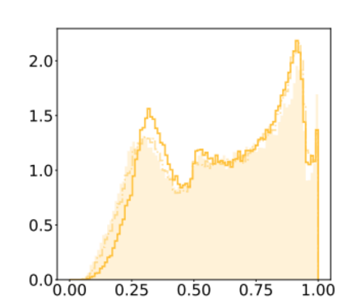
Normalizing Flows

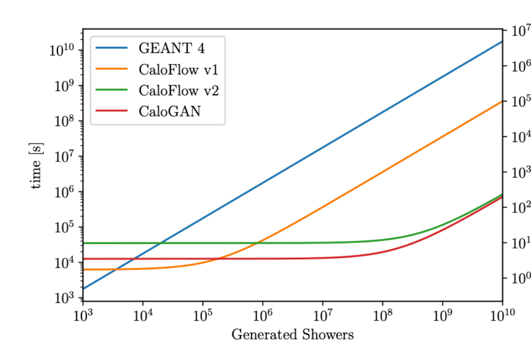
Krause, Claudius, and David Shih. "CaloFlow II: Even Faster and Still Accurate Generation of Calorimeter Showers with Normalizing Flows." *arXiv:2110.11377*











Conditional Diffusion based Transformer

Architecture based on **visual transformers**

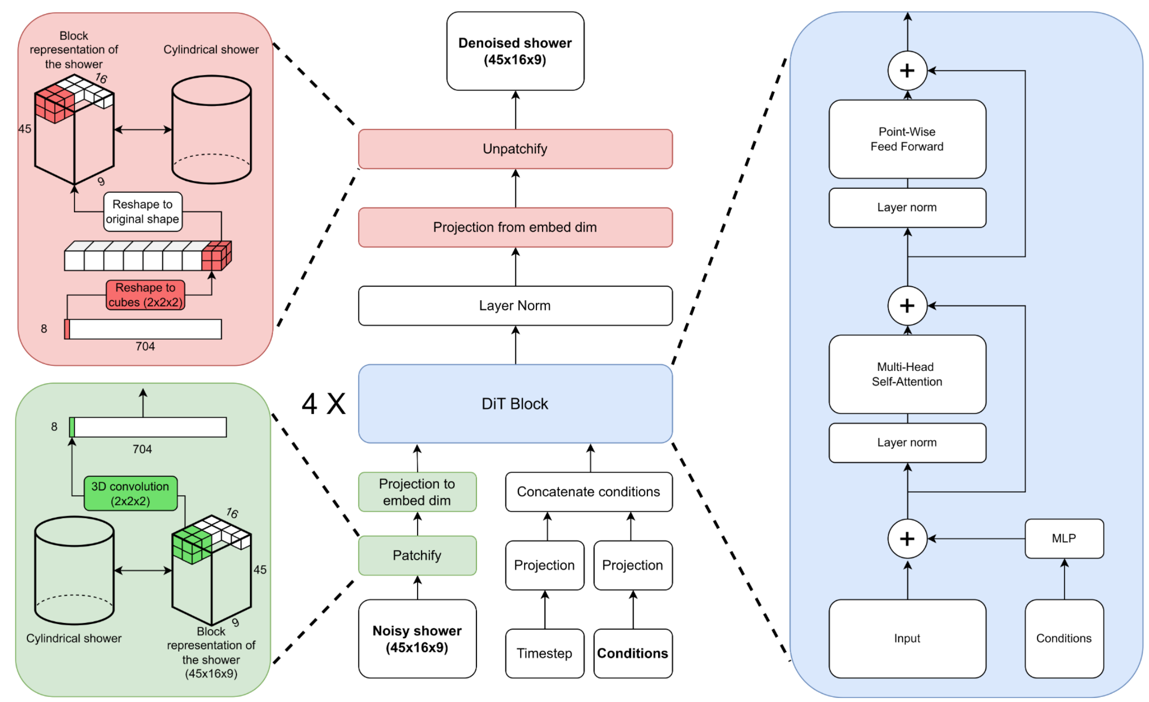
Input condition on Energy, Particle Trajectory, Geometry

Heavy data preprocessing necessary to map calorimeter geometry to image tiles

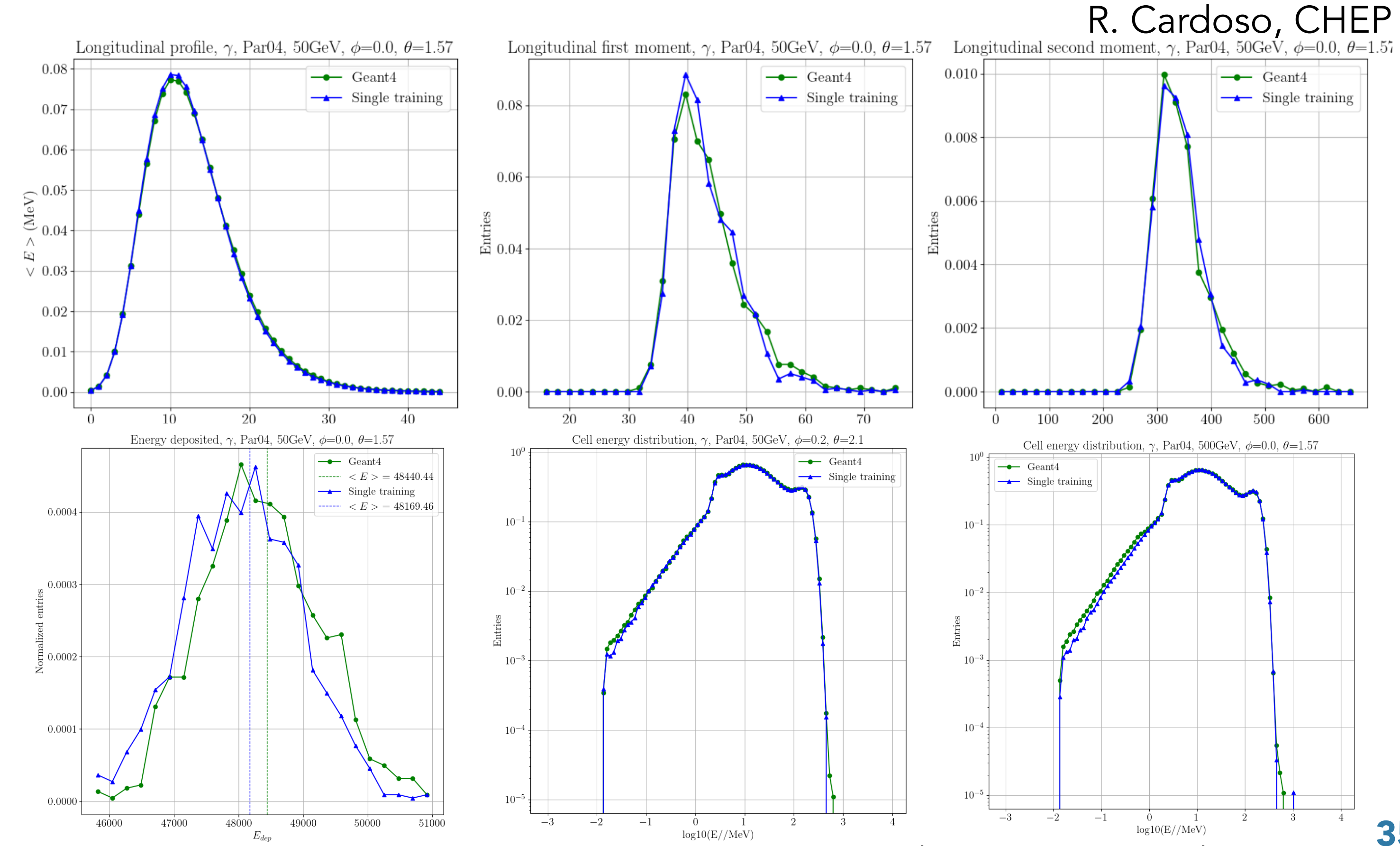
Maybe different data representation could be more convenient?

Results:

- **Good accuracy** throughout all profiles
- **Cell energy shows particular good results** compared to other generative models



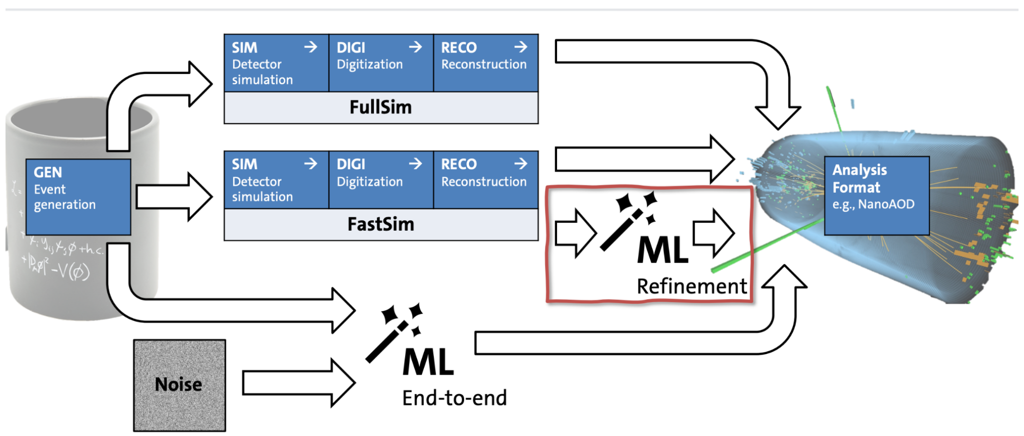
R. Cardoso, CHEP 2032

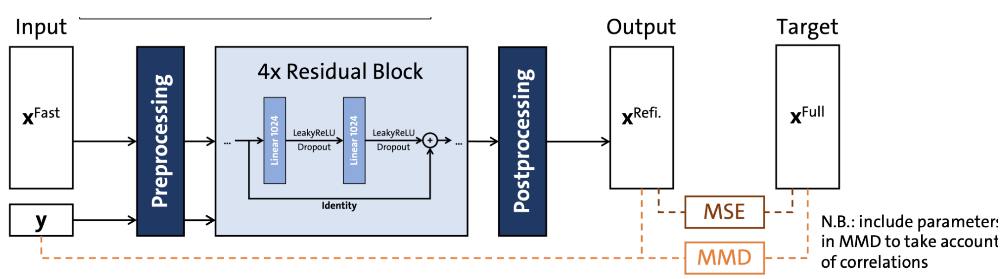


https://arxiv.org/abs/2212.09748

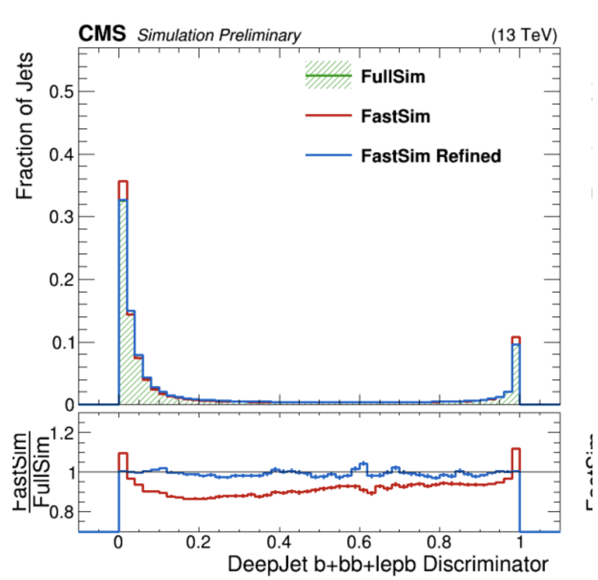
More Simulation

Deep learning to **match fast-sim to fullsim at analysis level Increases fidelity of fastsim**

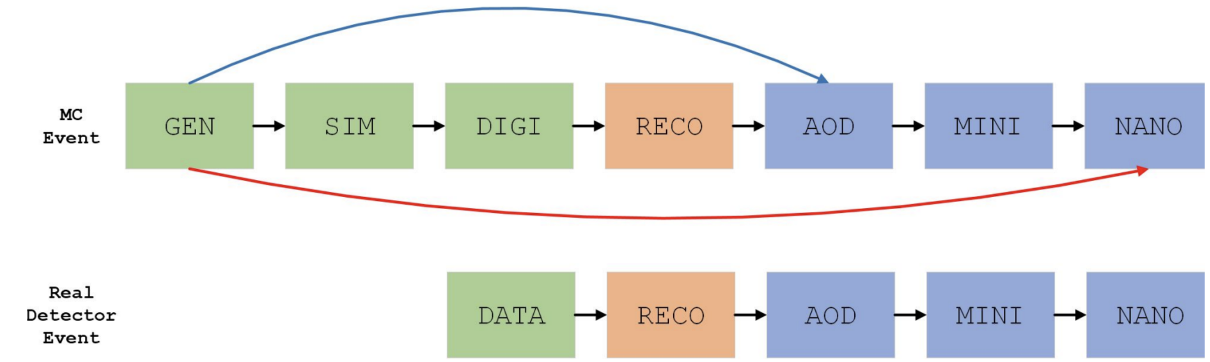


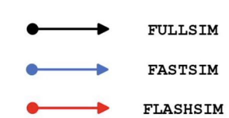


Refining fast simulation using machine learning



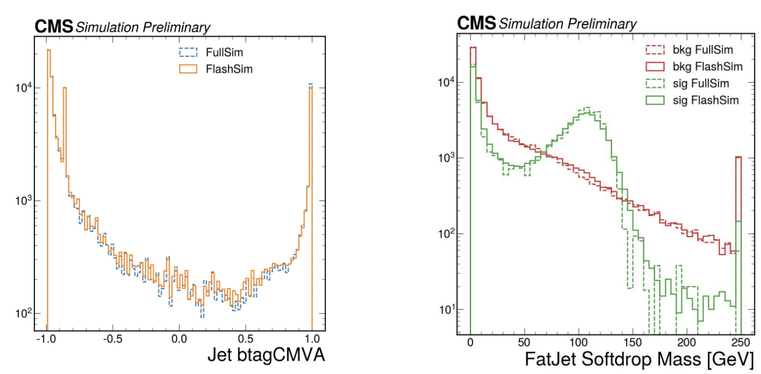
A normalizing-flow based **end-to-end super-fast-sim**, transforming Monte Carlo events directly into high-level analysis objects.





Generator-matched jets

Flashsim: a ML based simulation for analysis datatiers Mon 08/05



More interesting developments in constructing **ML models for event generation (hadronization) or** to have fundamental **data-driven ML representation** for hadronic physics models in Geant4

MLHad: Simulating Hadronization with Machine Learning

Simulation of Hadronic Interactions with Deep Generative Models

Comparing experimental data to theory

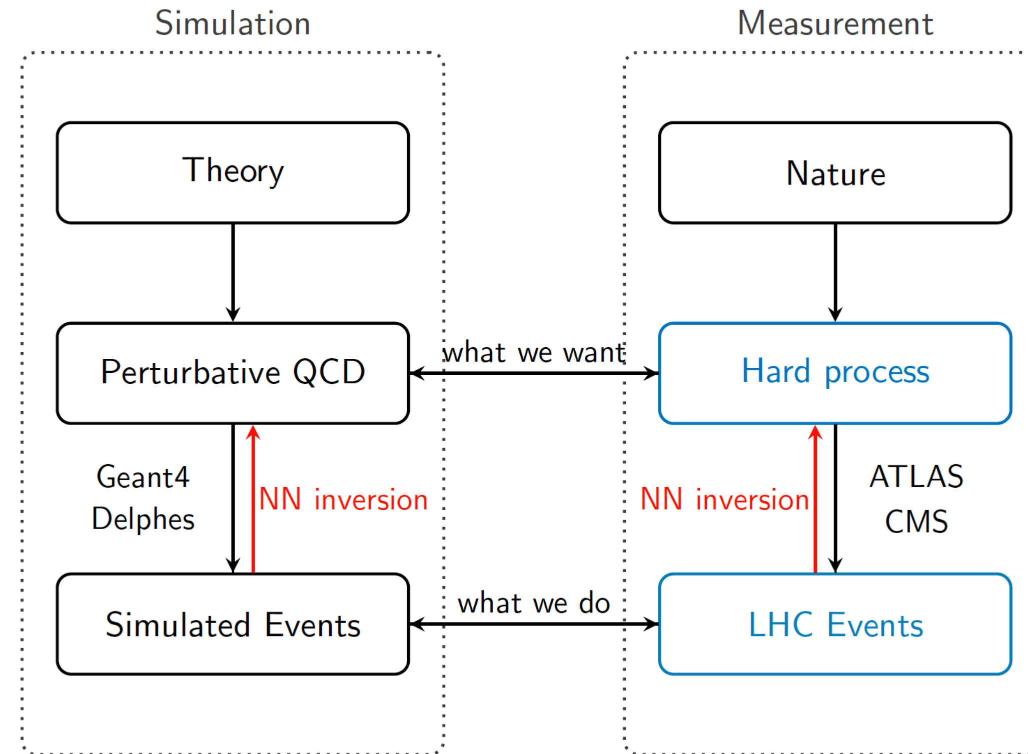
*“How to invert a matrix that should not be inverted”**

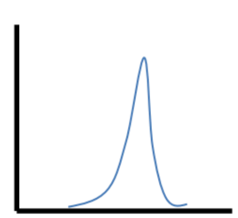
Detectors measure the results of **particle interactions** with matter

But we are interested in the **particle production processes**

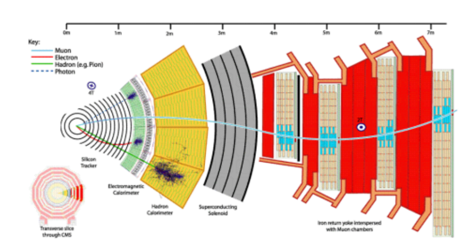
Go **back from experiments to theory**:

- **Disentangle** production process from the experimental setup
- Bayesian problem

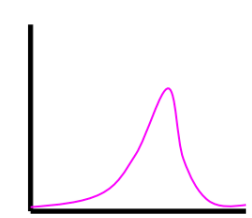




$x \sim p(x)$ = input / true distribution



$p(y|x)$ = Detector smearing



$y \sim p(y)$ = output / observed distribution

- $p(y) = \int p(y|x) p_\theta(x) dx \approx \sum_{x \sim p_\theta(x)} p(y|x)$

*quote from P. Vischia

Inverting the experiment

TIME MOVES IN ONE DIRECTION. MEMORY ANOTHER. WE ARE THAT STRANGE SPECIES THAT CONSTRUCTS ARTEFACTS INTENDED TO COUNTER THE NATURAL FLOW OF FORGETTING

William Gibson

arxiv:1808.04730
arxiv:2006.06685

Inverse problem: given observations **y** determine underlying hidden parameters **x**

Use invertible networks

- Train on the forward process **x** → **y**
- Run backward **y** → **x** to get prediction
- Add latent variable z to compensate information loss during forward process

$$[\mathbf{y}, \mathbf{z}] = f(\mathbf{x})$$

$$p(\mathbf{z})$$

$$\mathbf{x} = f^{-1}(\mathbf{y}, \mathbf{z}) = g(\mathbf{y}, \mathbf{z})$$