ARTIFICIAL INTELIGENCE & MACHINE LEARNING IN HIGH ENERGY PHYSICS

ANDRE SZNAJDER
UERJ

Outline

- A Short History of ML in HEP
- AI / ML / DL
- Deep Learning
- HEP Paradigm
- Applications of ML in HEP

History of ML in HEP

The use of ML techniques is revolutionizing how we interpret data samples, greatly increasing the discovery potential of present and future experiments.

ELSEVIER
Computer Physics Communications
Volume 49, Issue 3, June 1988, Pages 429-448

Neural networks and cellular automata in experimental high energy physics

B. Denby

Show more

Add to Mendeley Share Cite

https://doi.org/10.1016/0010-4655(88)90004-5

ELSEVIER
Computer Physics Communications
Volume 70, Issue 1, May 1992, Pages 167-182

Pattern recognition in high energy physics with artificial neural networks — JETNET 2.0

Leif Lönnblad, Carsten Peterson, Thorsteinn Rögnvalsson

Show more

Add to Mendeley Share Cite

https://doi.org/10.1016/0010-4655(92)90099-K Get rights and content

ELSEVIER
Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment
Volume 543, Issues 2–3, 11 May 2005, Pages 577-584

Boosted decision trees as an alternative to artificial neural networks for particle identification

Byron P. Roe [a], Hai-Jun Yang [a], Ji Zhu [b], Yang Liu [c], Ion Stancu [c], Gordon McGregor [d]

Show more

Add to Mendeley Share Cite

https://doi.org/10.1016/j.nima.2004.12.018

nature
Explore content About the journal Publish with us Subscribe
nature > review articles > article
Review | Published: 01 August 2018

Machine learning at the energy and intensity frontiers of particle physics

Alexander Radovic, Mike Williams, David Rousseau, Michael Kagan, Daniele Bonacorsi, Alexander Himmel, Adam Aurisano, Kazuhiro Terao & Taritree Wongjirad

Nature 560, 41–48 (2018) | Cite this article

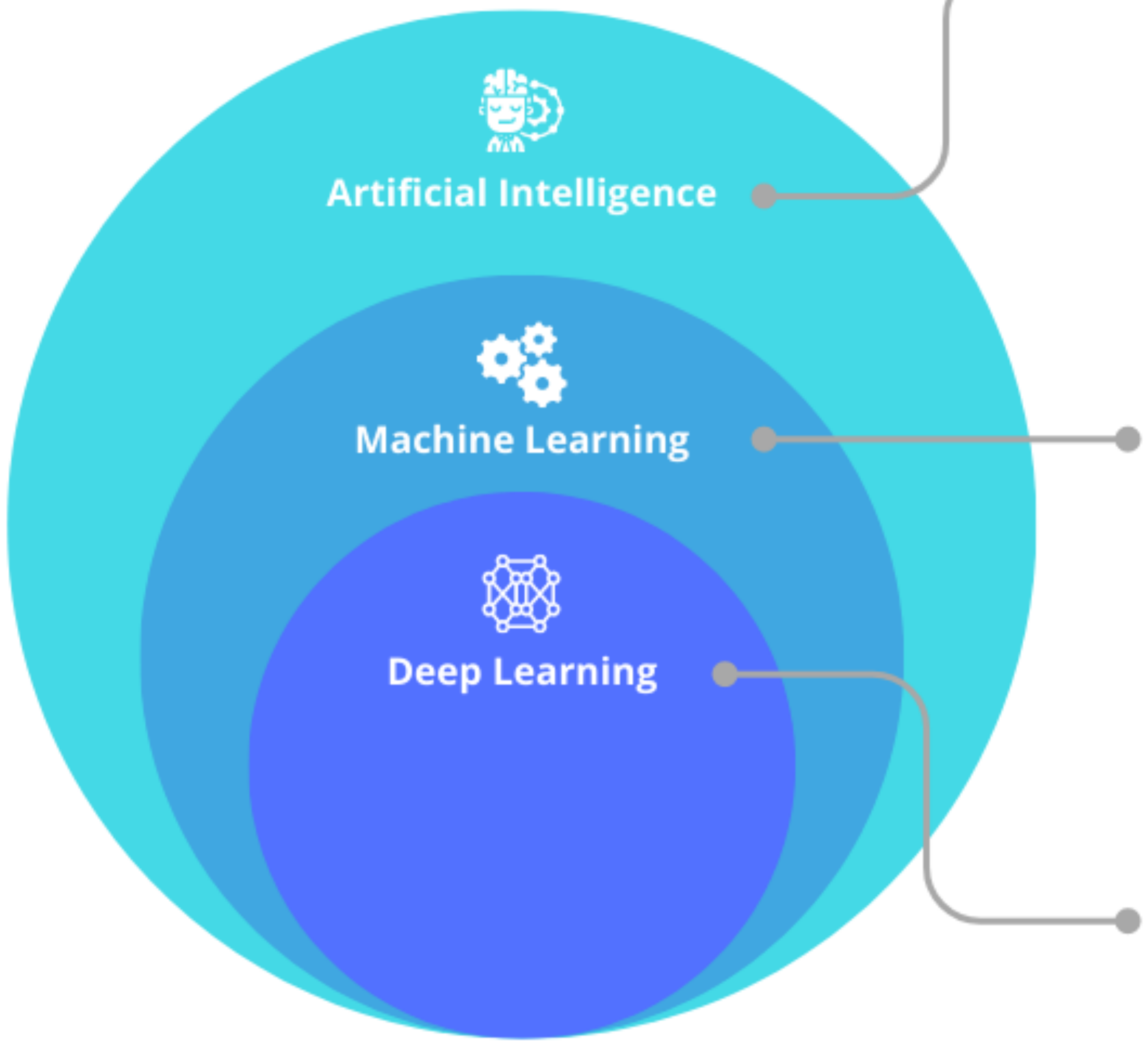
nature communications
Explore content About the journal Publish with us
nature > nature communications > articles > article
Article | Published: 02 July 2014

Searching for exotic particles in high-energy physics with deep learning

P. Baldi, P. Sadowski & D. Whiteson

Nature Communications 5, Article number: 4308 (2014) | Cite this article

Artificial Intelligence / Machine Learning / Deep Learning



ARTIFICIAL INTELLIGENCE (AI)

Algorithms which enables machines to mimic human behaviour

MACHINE LEARNING (ML)

Algorithms which enables machines to learn from data without being explicitly programmed to perform a task

DEEP LEARNING (DL)

Subset of ML which make the computation of multi-layer neural network feasible

Artificial Intelligence

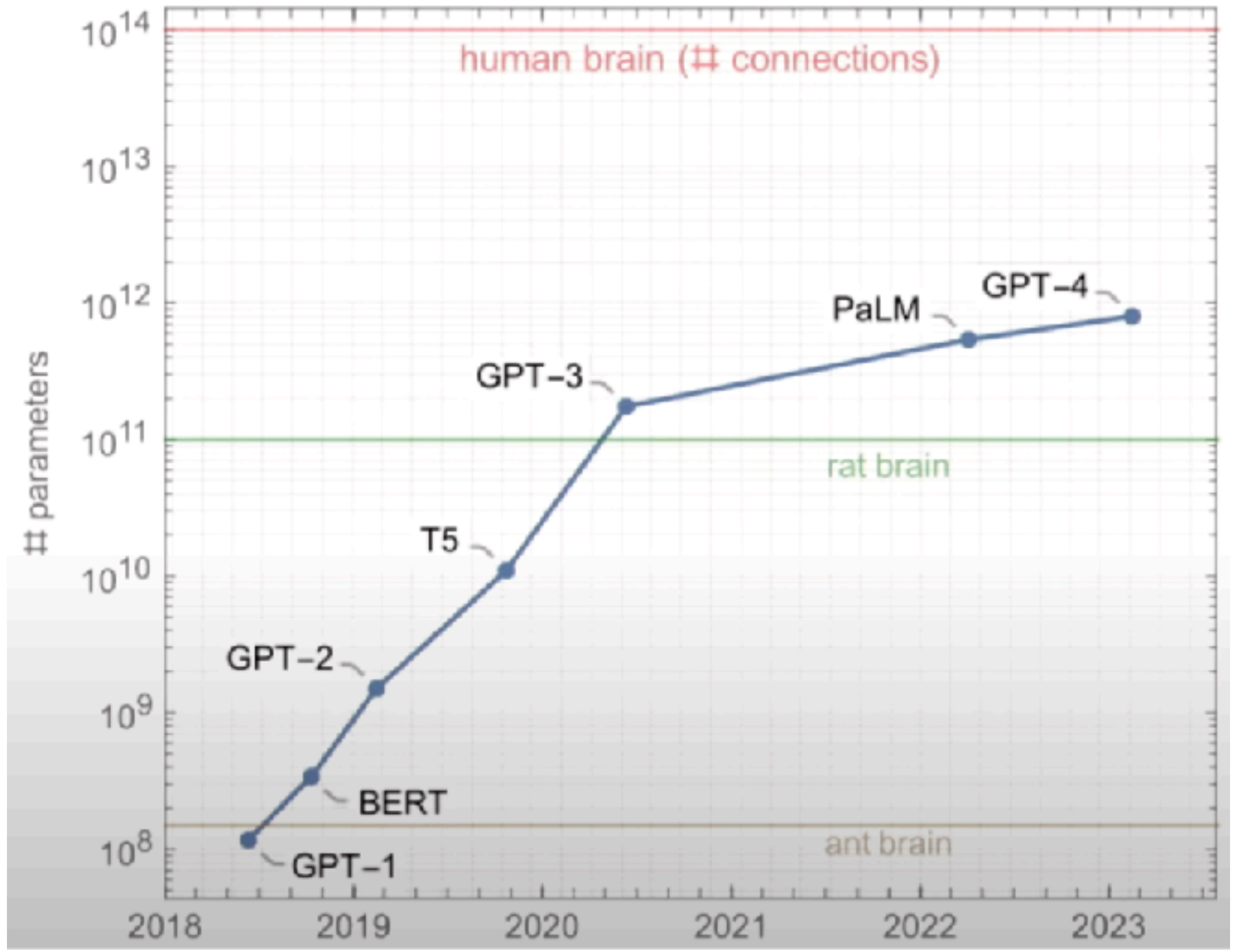
A Path Towards Autonomous Machine Intelligence
Version 0.9.2, 2022-06-27

Yann LeCun
Courant Institute of Mathematical Sciences, New York University yann@cs.nyu.edu
Meta - Fundamental AI Research yann@fb.com

June 27, 2022

Animals and humans exhibit learning abilities and understanding far beyond the capabilities of current Artificial Intelligence (AI) systems.

- A teenager who has never sat behind a steering wheel can learn to drive in about 20 hours
- By contrast, the best autonomous driving systems today need billions of pieces of labeled training data and millions of reinforcement learning trials in virtual environments.



LeCun proposes that one of the most important challenges in AI today is devising learning paradigms and architectures that would allow machines to learn world models in a self-supervised fashion and then use those models to predict, reason, and plan

Deep Learning

Deep neural networks exploits the compositional character of nature. The network layers learns a hierarchical representation of the data with multiple levels of abstraction.

nature

Explore content ⌄ About the journal ⌄ Publish with us

nature > review articles > article

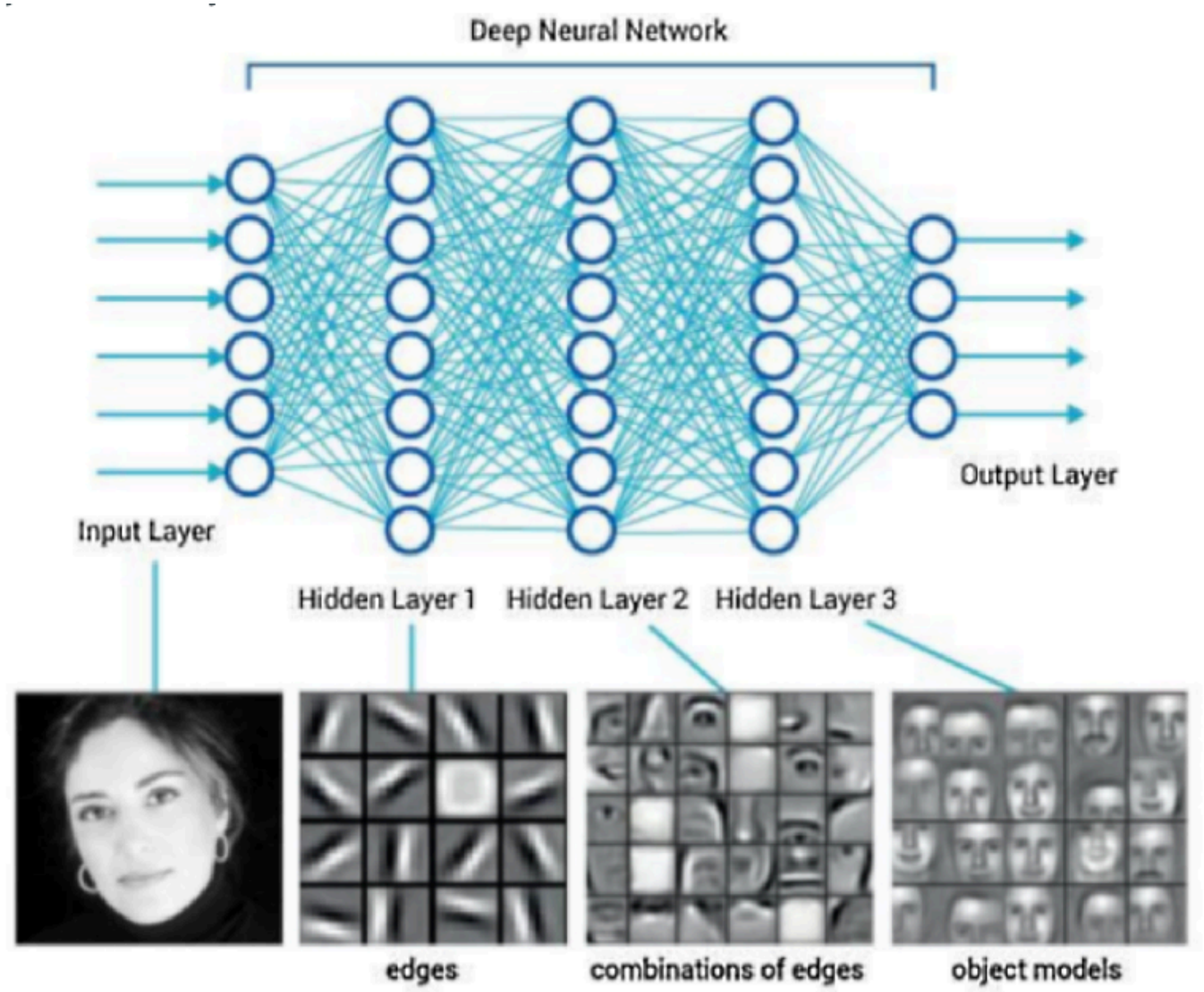
Review Article | Published: 27 May 2015

Deep learning

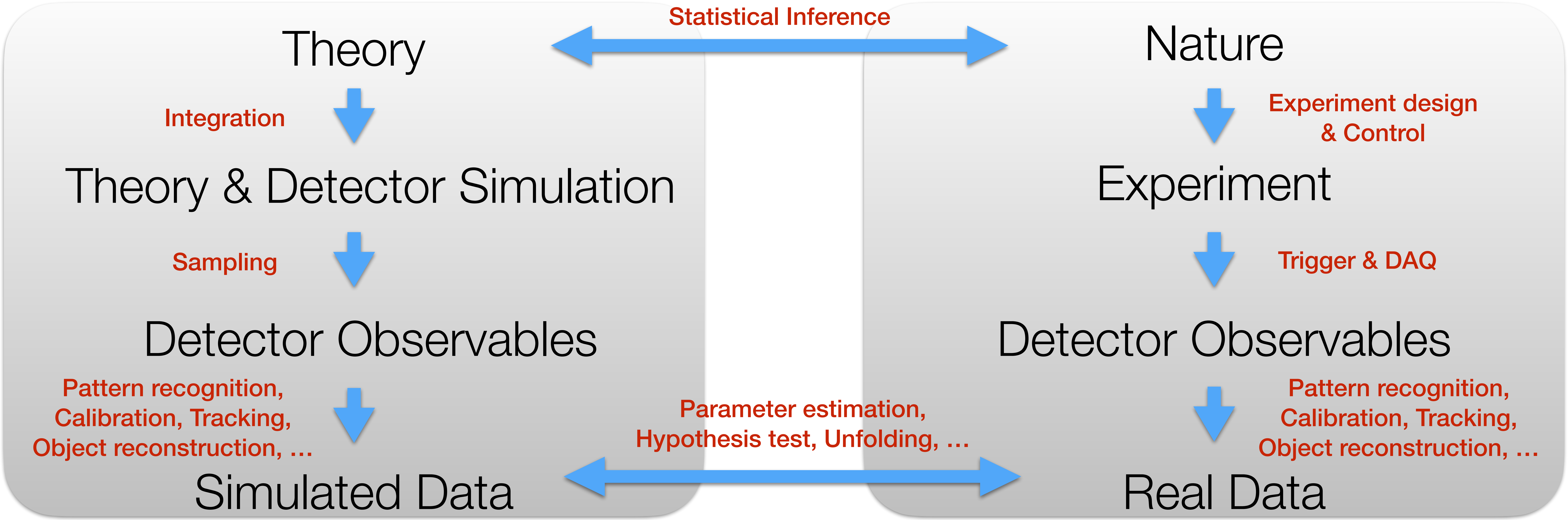
Yann LeCun, Yoshua Bengio & Geoffrey Hinton

Nature **521**, 436–444 (2015) | Cite this article

983k Accesses | **37k** Citations | **1438** Altmetric | Metrics

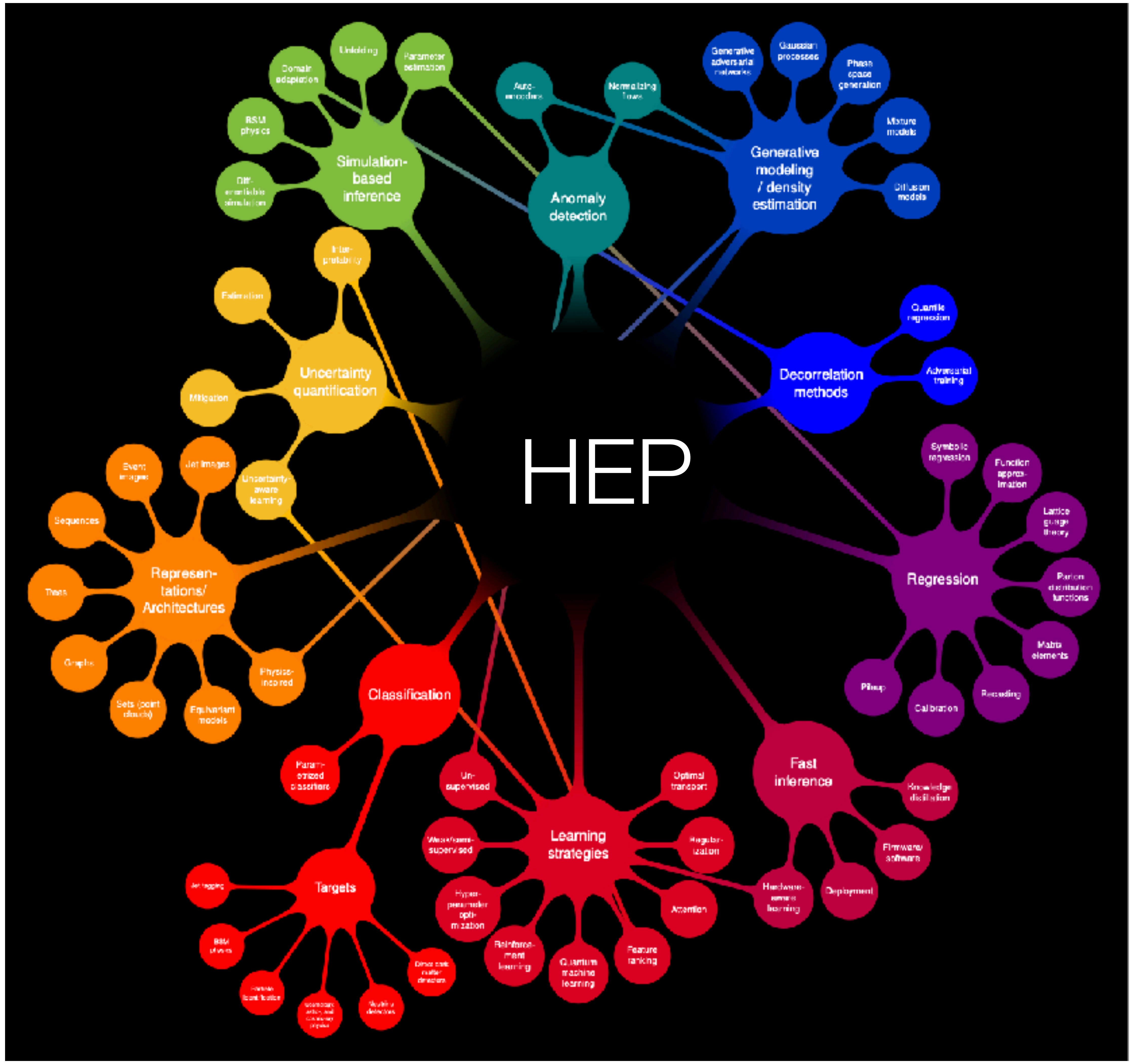


HEP Paradigm



ML can play a role in every step of the HEP paradigm !

ML in HEP

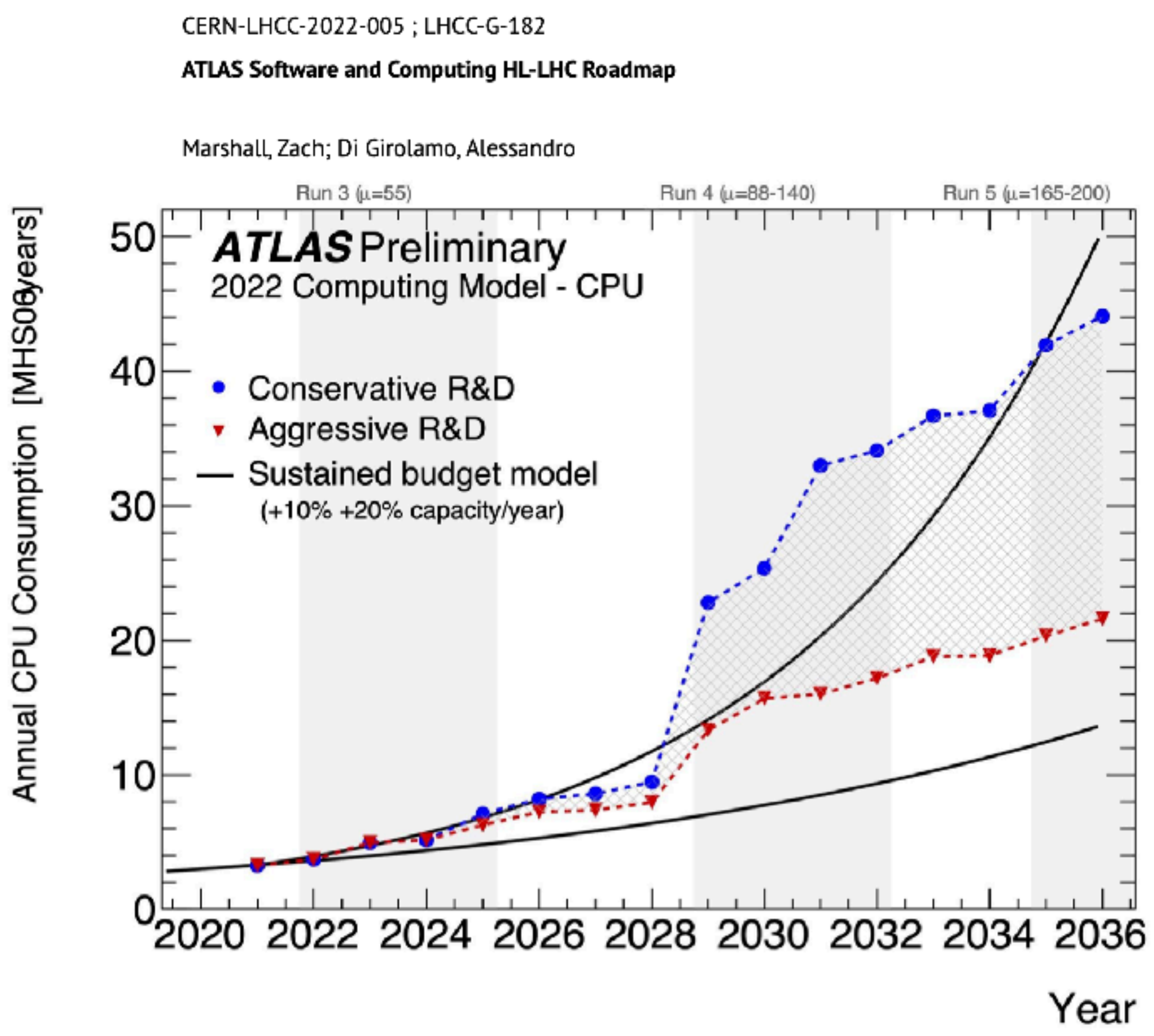


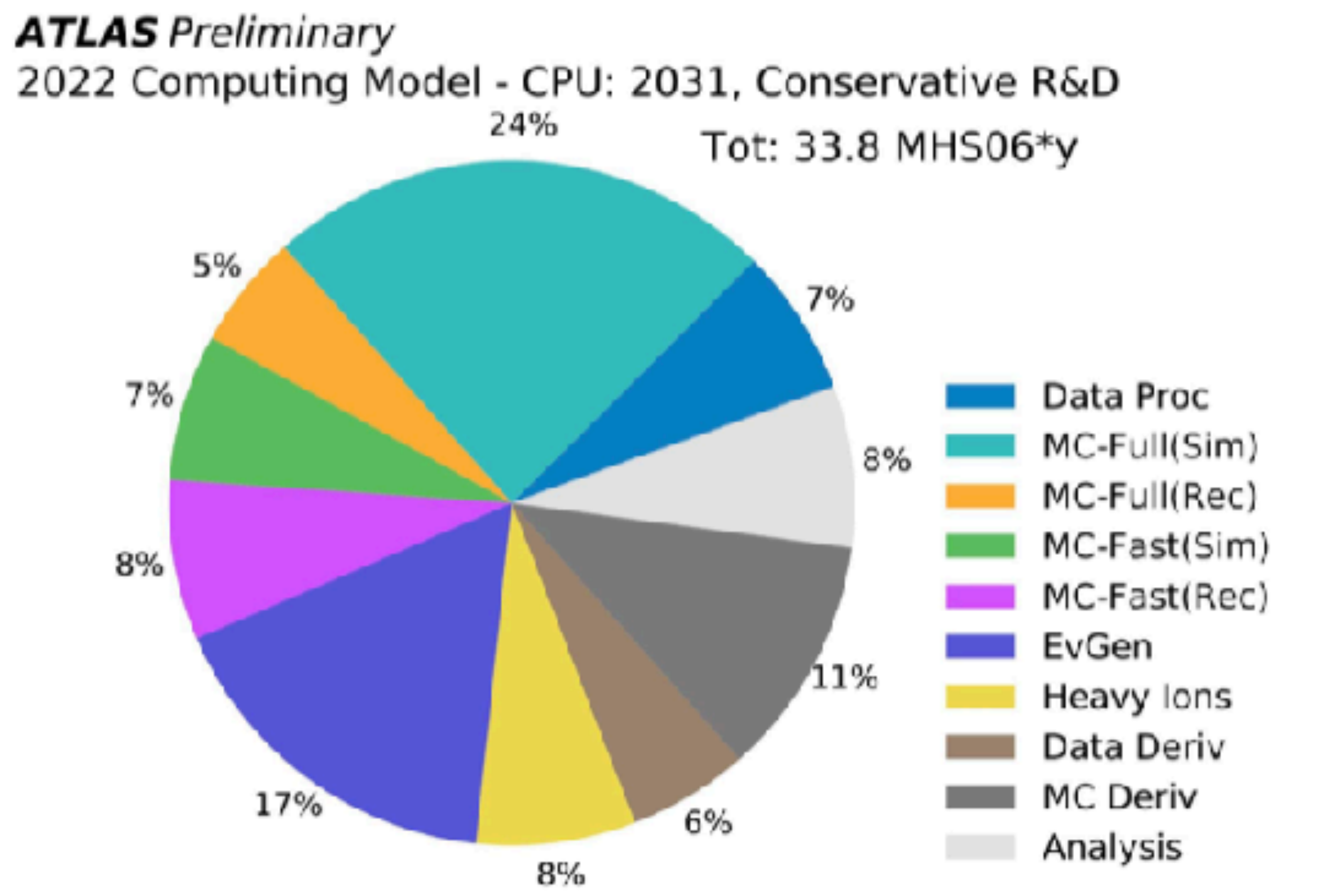
Typical HEP problems:

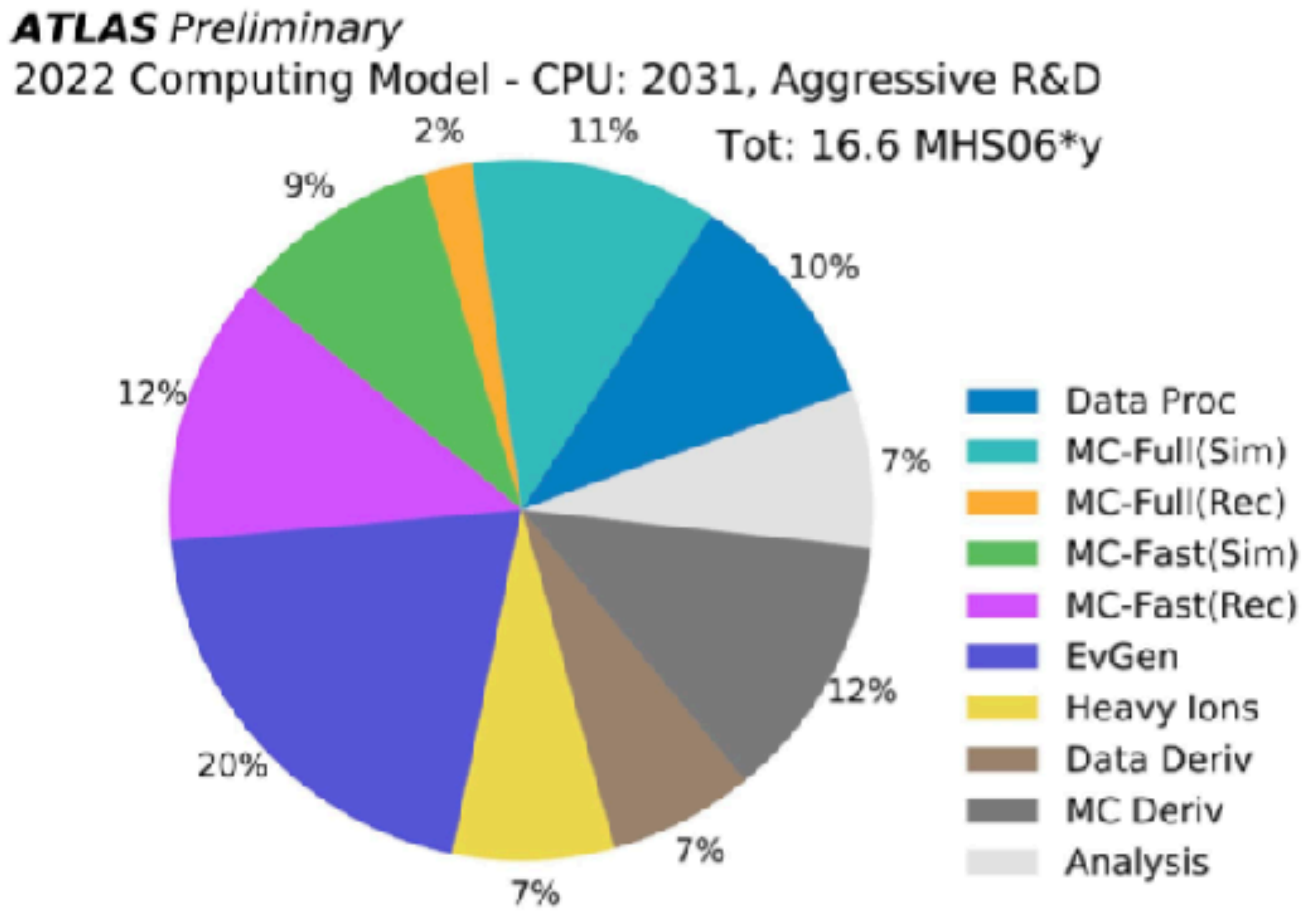
- Simulation
- Classification
- Regression
- Triggering/Filtering
- Unfolding

ML in Simulation

Simulation is a key driver of CPU needs for the HL-LHC







ML can be used reduce CPU simulation needs

ML in Simulation

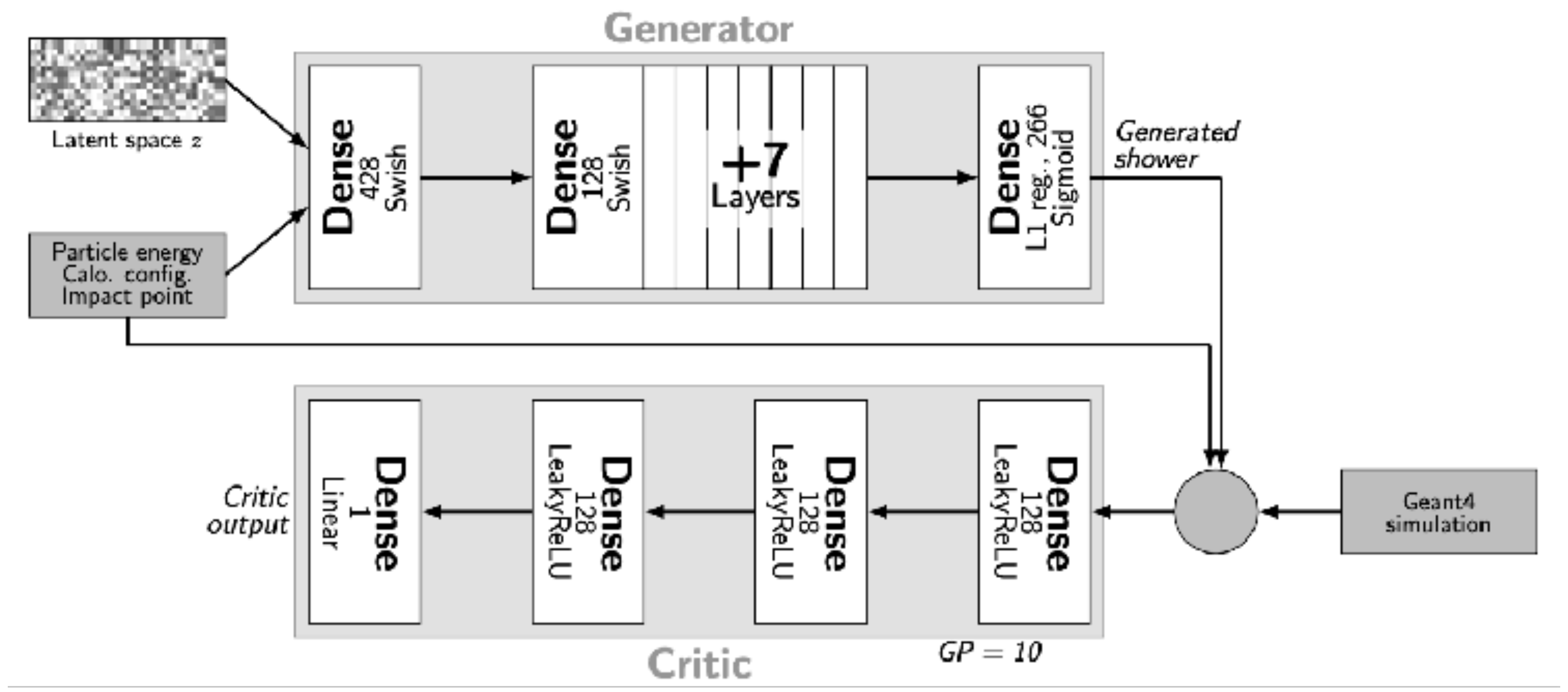
Variational autoencoders (VAE) and generative adversarial networks (GAN) are capable of quickly simulating electromagnetic showers with correct total energies and stochasticity

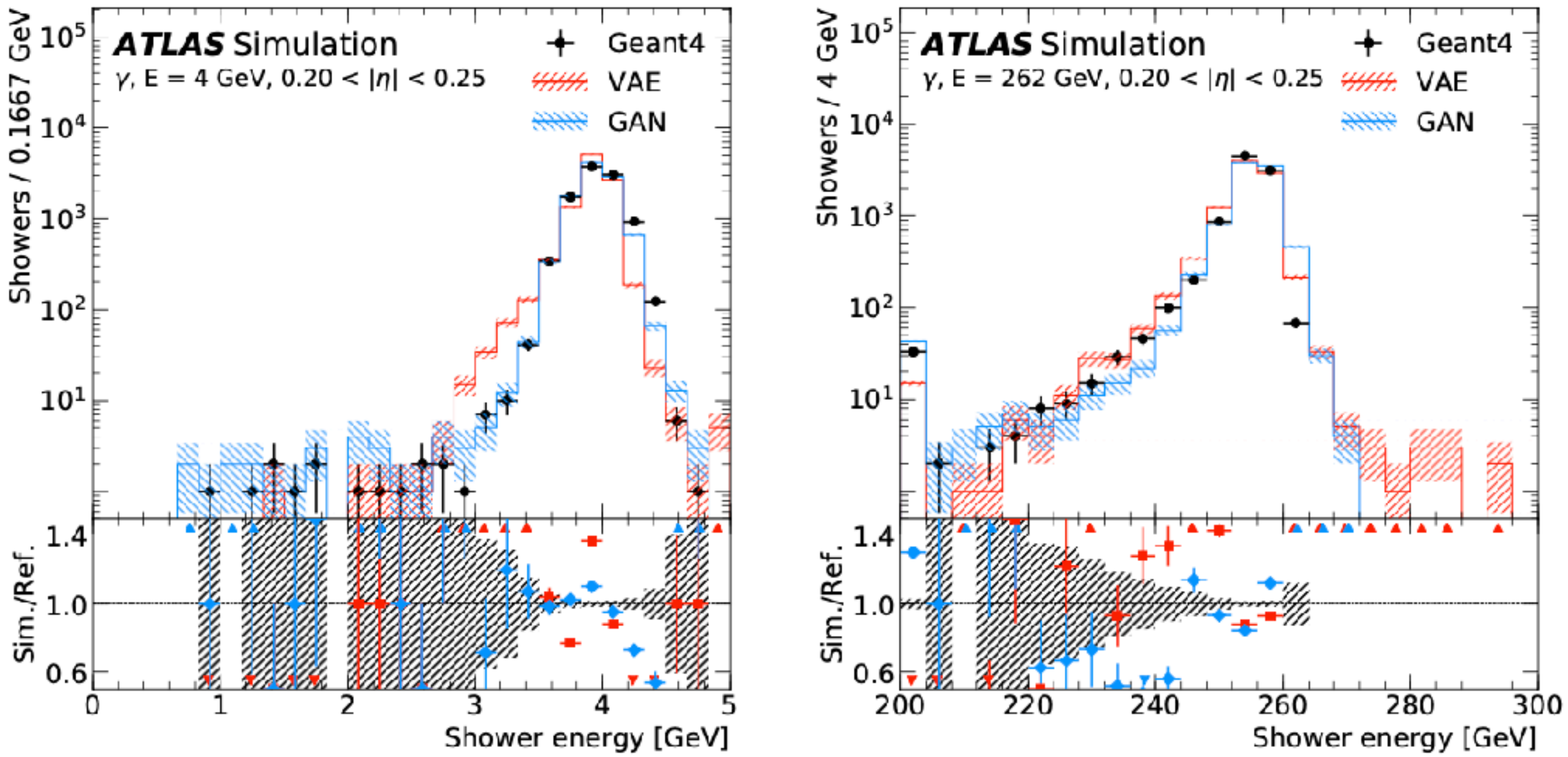
Computing and Software for Big Science (2024) 8:7
https://doi.org/10.1007/s41781-023-00106-9

RESEARCH

Deep Generative Models for Fast Photon Shower Simulation in ATLAS

The ATLAS Collaboration*
CERN, Geneva, Switzerland

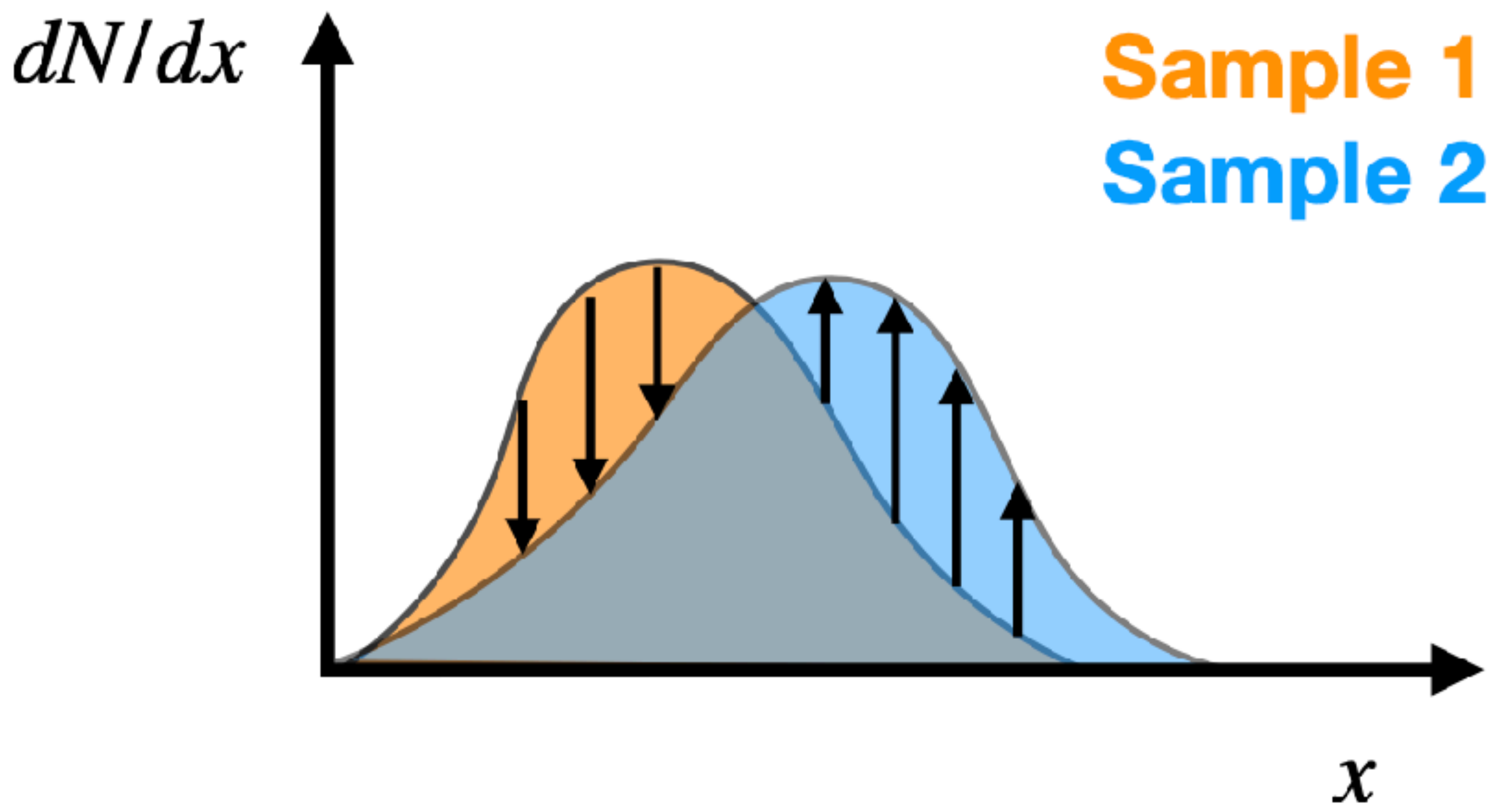




ML in Simulation

Neural reweight of simulated samples to different model parameters or different models, avoiding need of simulating the detector response multiple times. Consider a neural network classifier NN(x) and two distributions.

$$dN/dx \rightarrow w(x) \times dN/dx$$



CMS-PAS-MLG-24-001

Reweighting of simulated events using machine learning techniques in CMS

Likelihood Ratio Trick:

$$w(x) = \frac{p_1(x)}{p_2(x)} \simeq \frac{NN(x)}{1 - NN(x)}$$

Reweight MC for evaluation of systematic uncertainties or higher order correction

ML in DAQ

A real-time autoencoder-based anomaly detection system using semi-supervised machine learning has been developed for the online Data Quality Monitoring system of the electromagnetic calorimeter of the CMS detector at the CERN LHC. A CAE network architecture is build exploiting ECAL data processed as 2D image.

arXiv > physics > arXiv:2407.20278

Search... Help | Advar

Physics > Instrumentation and Detectors

(Submitted on 25 Jul 2024)

Anomaly Detection Based on Machine Learning for the CMS Electromagnetic Calorimeter Online Data Quality Monitoring

Abhirami Harilal, Kyungmin Park, Manfred Paulini (On behalf of the CMS Collaboration)

Table 1. Summary of FDR using 99% anomaly detection threshold for the ECAL barrel fake anomaly scenarios.

	FDR for 99% anomaly detection		
	Missing Supermodule	Zero Occup. Tower	Hot Tower
AE no correction	3.6%	51%	2.8%
AE after spatial correction	3.1%	49%	2.9%
AE after spatial and time corrections	0.13%	4.1%	< 0.01%

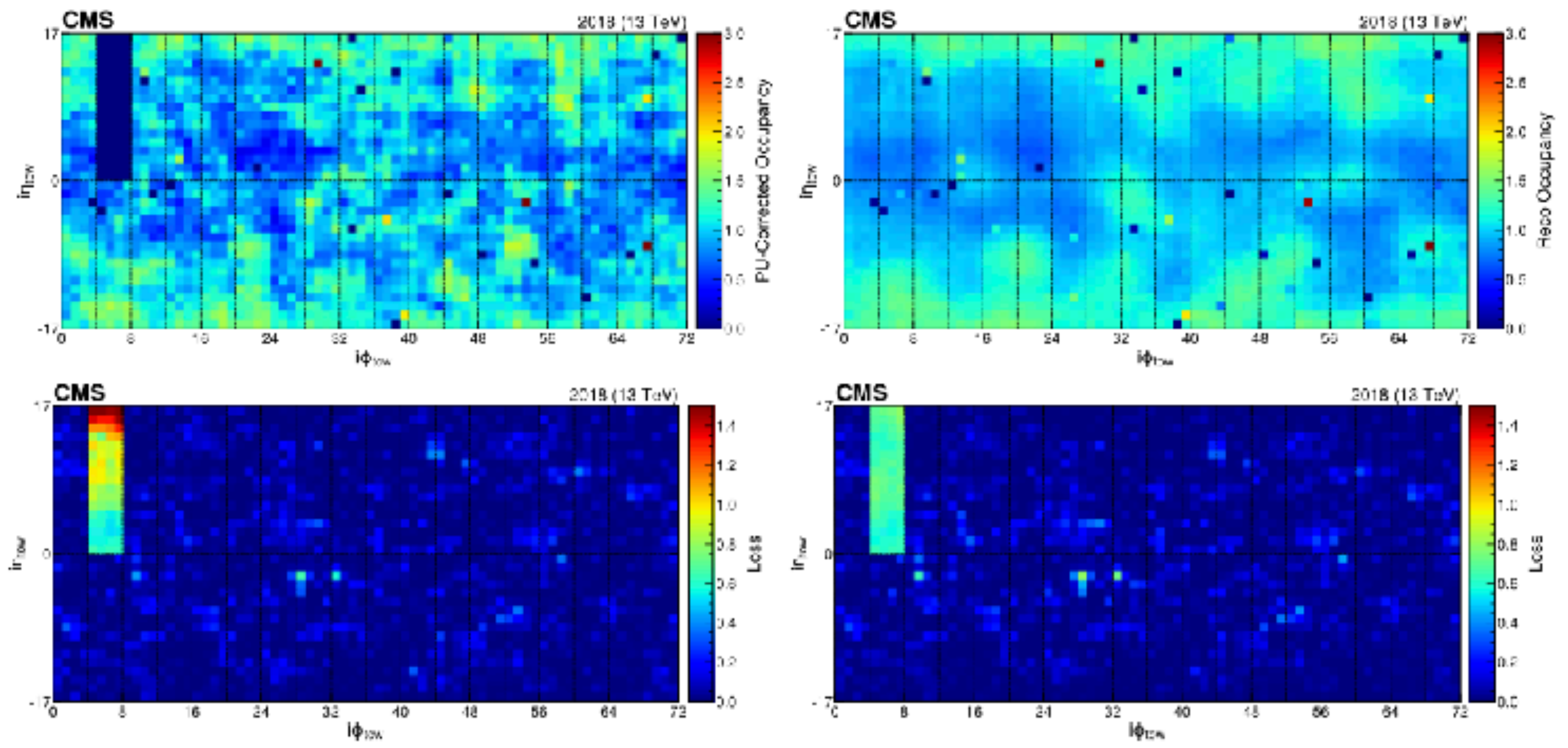


Figure 4. Top-left: Occupancy map with a missing supermodule in the barrel. Top-right: AE-reconstructed occupancy map. Bottom-left: Loss map showing the missing supermodule, indicating higher loss at high $|\eta|$ owing to differences in the detector response. Bottom-right: Loss map after the spatial correction.

The AE-based anomaly detection system labeled MLDQM has been deployed in the CMS ECAL online DQM workflow for the barrel starting in LHC Run 3 in 2022 and for the endcaps in 2023

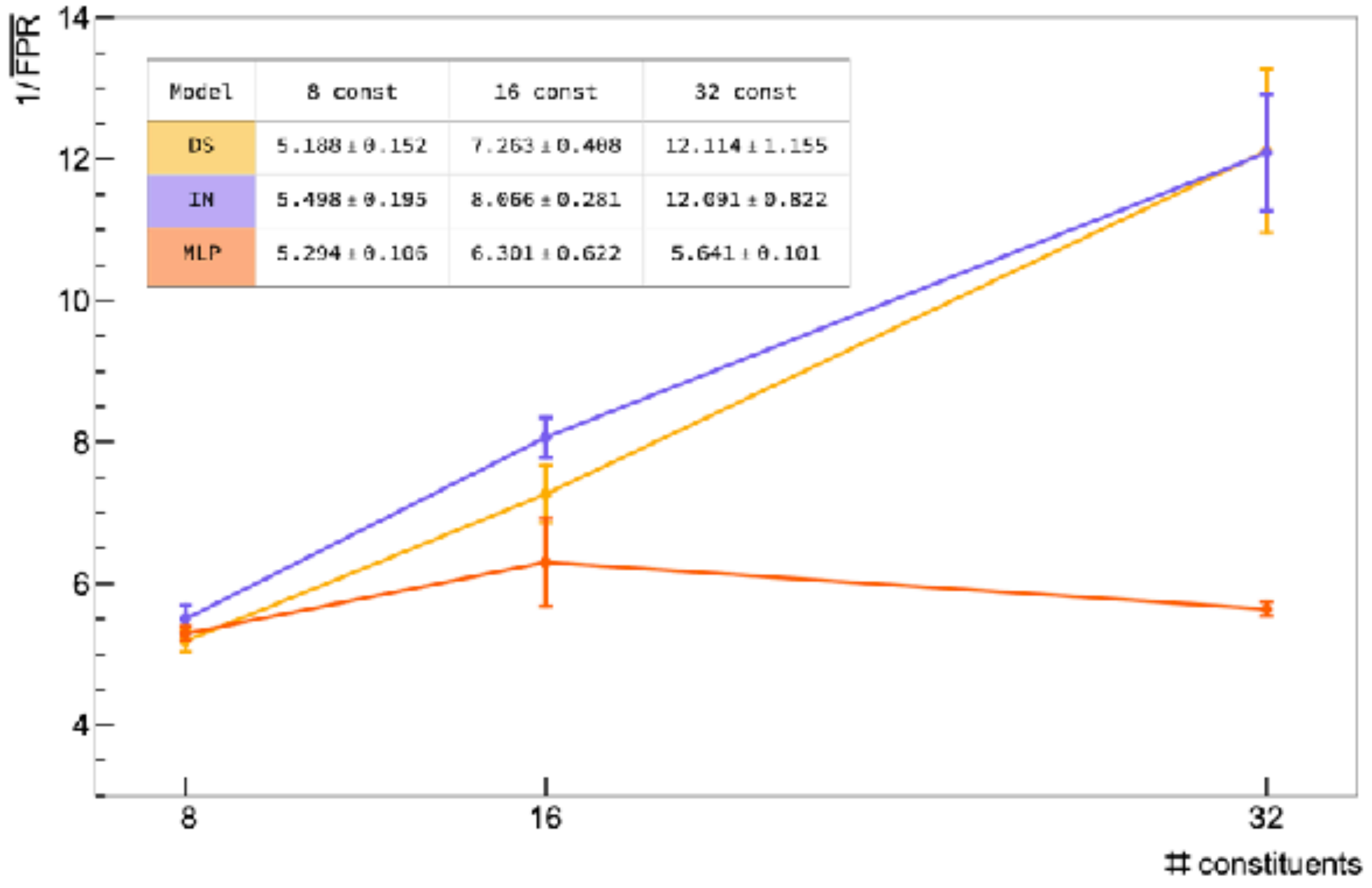
ML in L1 Trigger

Ultrafast jet classification on FPGAs for the HL-LHC

Patrick Odagiu, Zhiqiang Que, Javier Duarte, Johannes Haller, Gregor Kasieczka, Artur Lobanov, Vladimir Loncar, Wayne Luk, Jennifer Ngadiuba, Maurizio Pierini, Philipp Rincke, Arpita Seksaria, Sioni Summers, Andre Sznajder, Alexander Tapper, Thea K. Aarrestad

Three machine learning models are used to perform jet origin classification. These models are optimized for deployment on a field-programmable gate array device. In this context, we demonstrate how latency and resource consumption scale with the input size and choice of algorithm. Moreover, the models proposed here are designed to work on the type of data and under the foreseen conditions at the CERN LHC during its high-luminosity phase. Through quantization-aware training and efficient synthetization for a specific field programmable gate array, we show that $O(100)$ ns inference of complex architectures such as Deep Sets and Interaction Networks is feasible at a relatively low computational resource cost.

Comments: 13 pages, 3 figures, 3 tables. Mach. Learn.: Sci. Technol (2024)
Subjects: **High Energy Physics - Experiment (hep-ex)**; Machine Learning (cs.LG); Instrumentation and Detectors (physics.ins-det)
Report number: FERMILAB-PUB-24-0030-CMS-CSAID-PPD
Cite as: arXiv:2402.01876 **[hep-ex]**
(or arXiv:2402.01876v2 **[hep-ex]** for this version)
https://doi.org/10.48550/arXiv.2402.01876
Related DOI: https://doi.org/10.1088/2632-2153/ad5f10

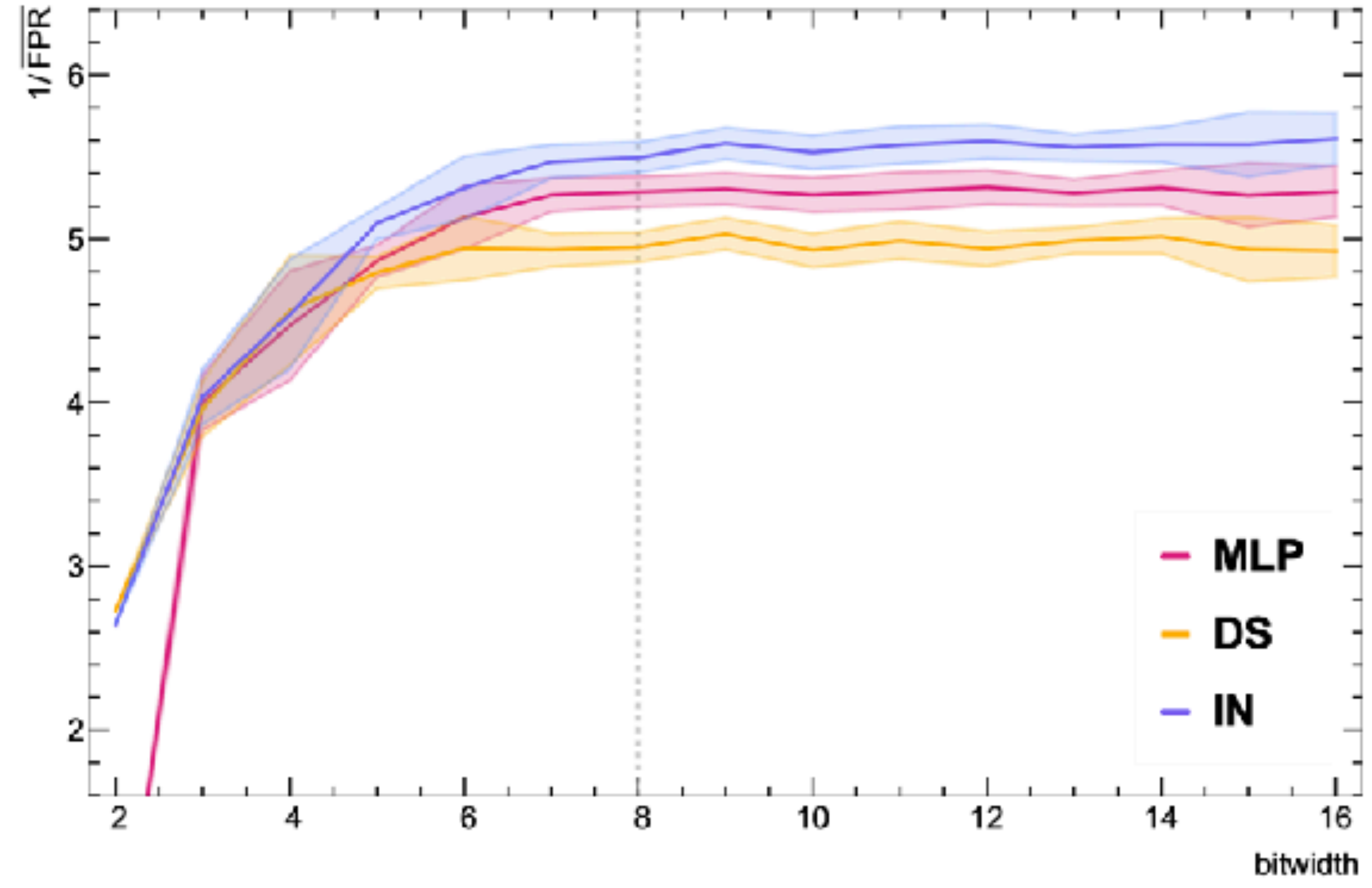


Model	8 const	16 const	32 const
DS	5.188 ± 0.152	7.263 ± 0.408	12.114 ± 1.155
IN	5.498 ± 0.195	8.066 ± 0.281	12.091 ± 0.822
MLP	5.294 ± 0.106	6.301 ± 0.622	5.641 ± 0.101

FPGA: Xilinx Virtex UltraScale+ VU13P

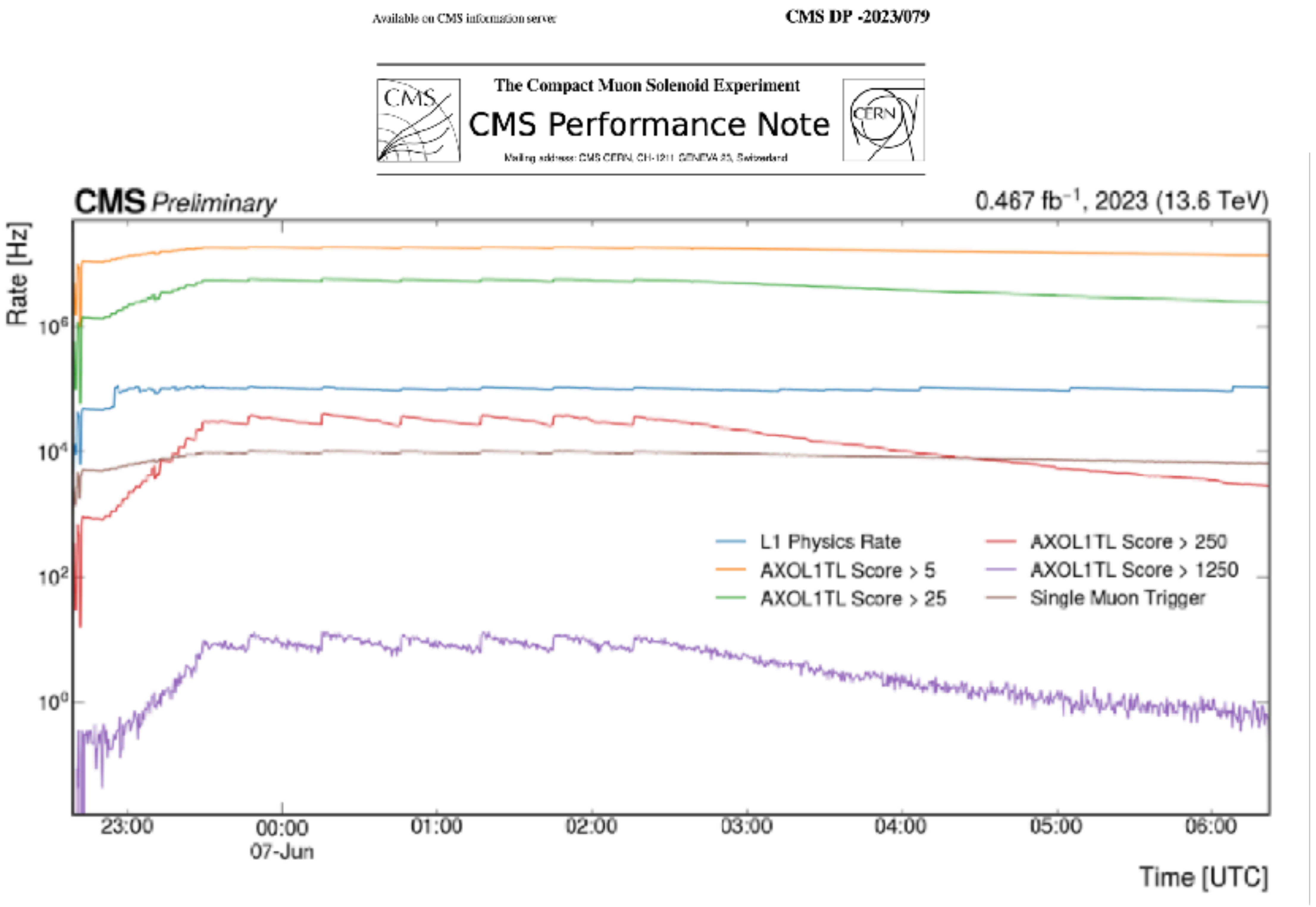
Architecture	Constituents	RF	Latency [ns] (cc)	II [ns] (cc)	DSP	LUT	FF	BRAM18
MLP	8	1	105 (21)	5 (1)	262 (2.1%)	155,080 (9.0%)	25,714 (0.7%)	4 (0.1%)
	16	1	100 (20)	5 (1)	226 (1.8%)	146,515 (8.5%)	31,426 (0.9%)	4 (0.1%)
	32[a]	1	105 (21)	5 (1)	262 (2.1%)	155,080 (7.2%)	25,714 (0.7%)	4 (0.1%)
DS	8	2	95 (19)	15 (3)	626 (5.1%)	386,294 (22.3%)	121,424 (3.5%)	4 (0.1%)
	16	4	115 (23)	15 (3)	555 (4.5%)	747,374 (43.2%)	238,798 (6.9%)	4 (0.1%)
	32[a]	8	130 (26)	10 (2)	434 (3.5%)	903,284 (52.3%)	358,754 (10.4%)	4 (0.1%)
IN	8	2	160 (32)	15 (3)	2,191 (17.8%)	472,140 (27.3%)	191,802 (5.5%)	12 (0.2%)
	16	4	180 (36)	15 (3)	5,362 (43.6%)	1,387,923 (80.3%)	594,039 (17.2%)	52 (1.9%)
	32[a]	8	205 (41)	15 (3)	2,120 (17.3%)	1,162,104 (67.3%)	761,061 (22.0%)	132 (2.5%)

[a] Pruning to a sparsity of 50% is applied to the 32-constituent IN model such that it can fit within the resource constraints of the FPGA. For consistency, the same pruning sparsity is applied to the 32-constituent MLP and DS models.



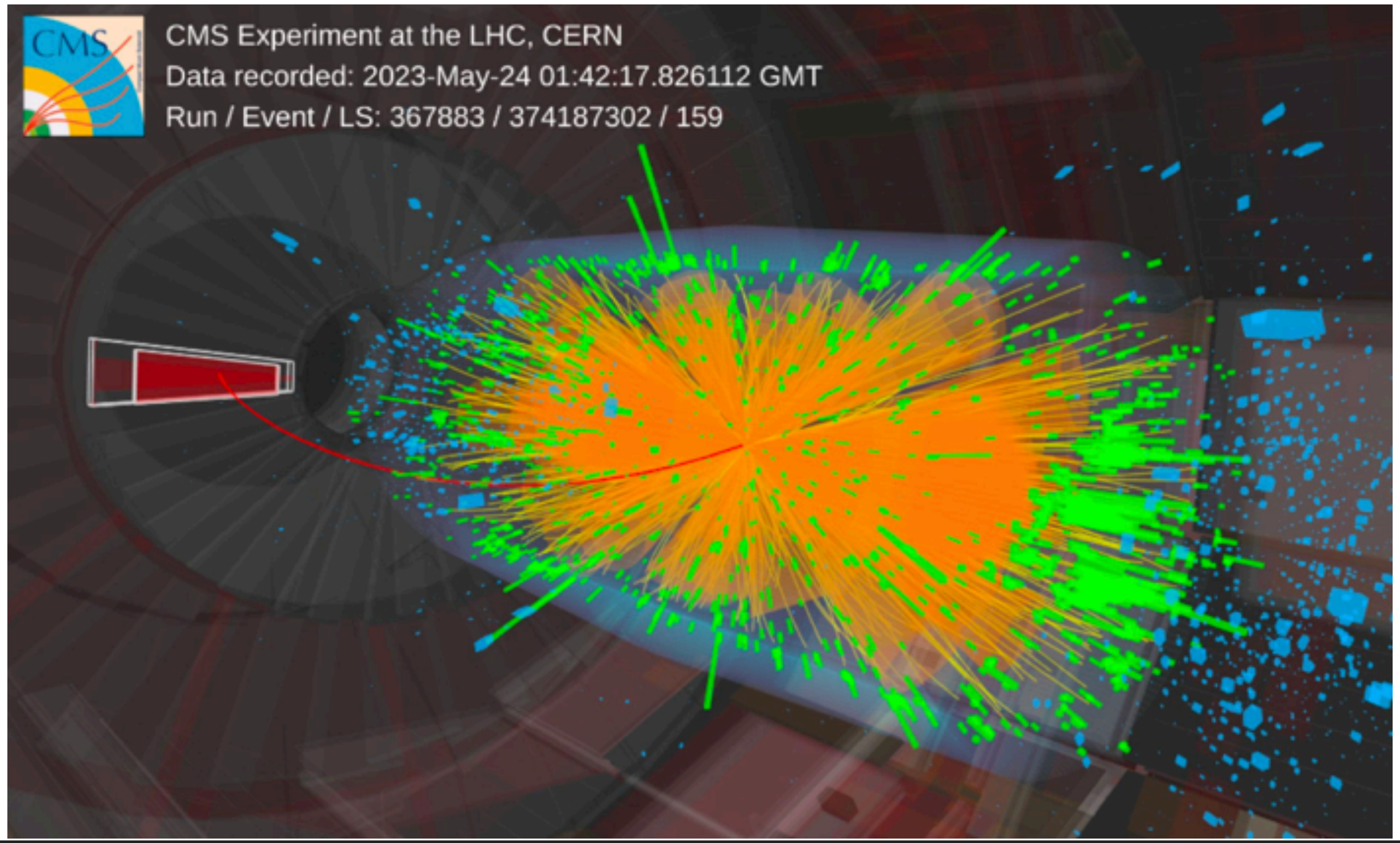
ML in L1 Trigger

AXOL1TL anomaly detection algorithm for the level-1 trigger based on a variational autoencoder implemented on a FPGA. Model is trained on ZeroBias data and used to detect data outliers.



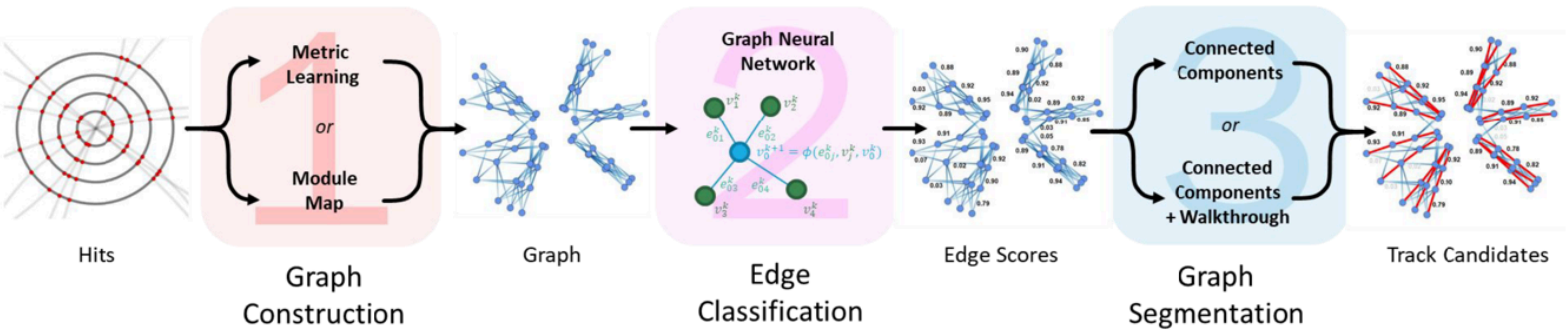
Vivado latency and resource utilization report for Anomaly Detection trigger on Xilinx Virtex-7 FPGA

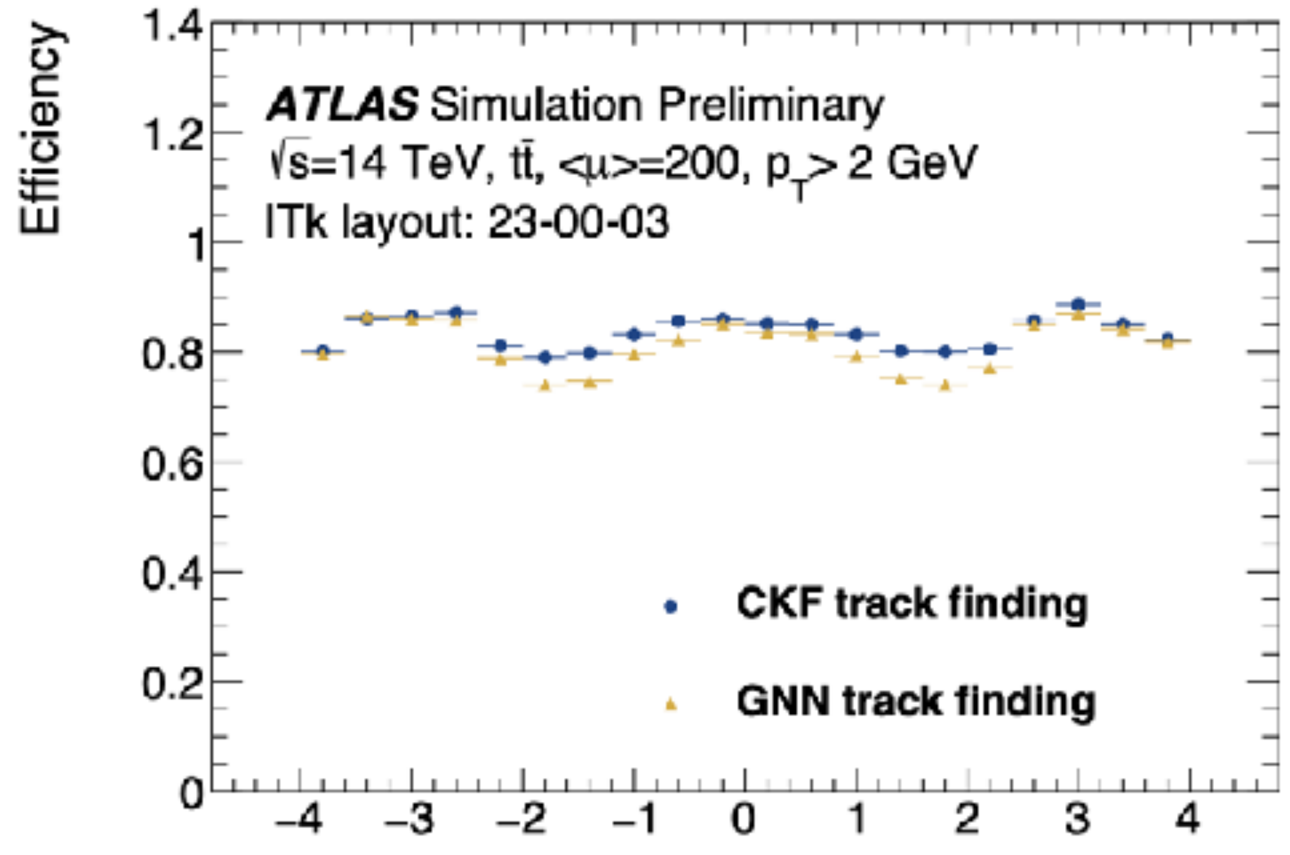
	Latency	LUTs	FFs	DSPs	BRAMs
AXOL1TL	2 ticks 50 ns	2.1%	~0	0	0

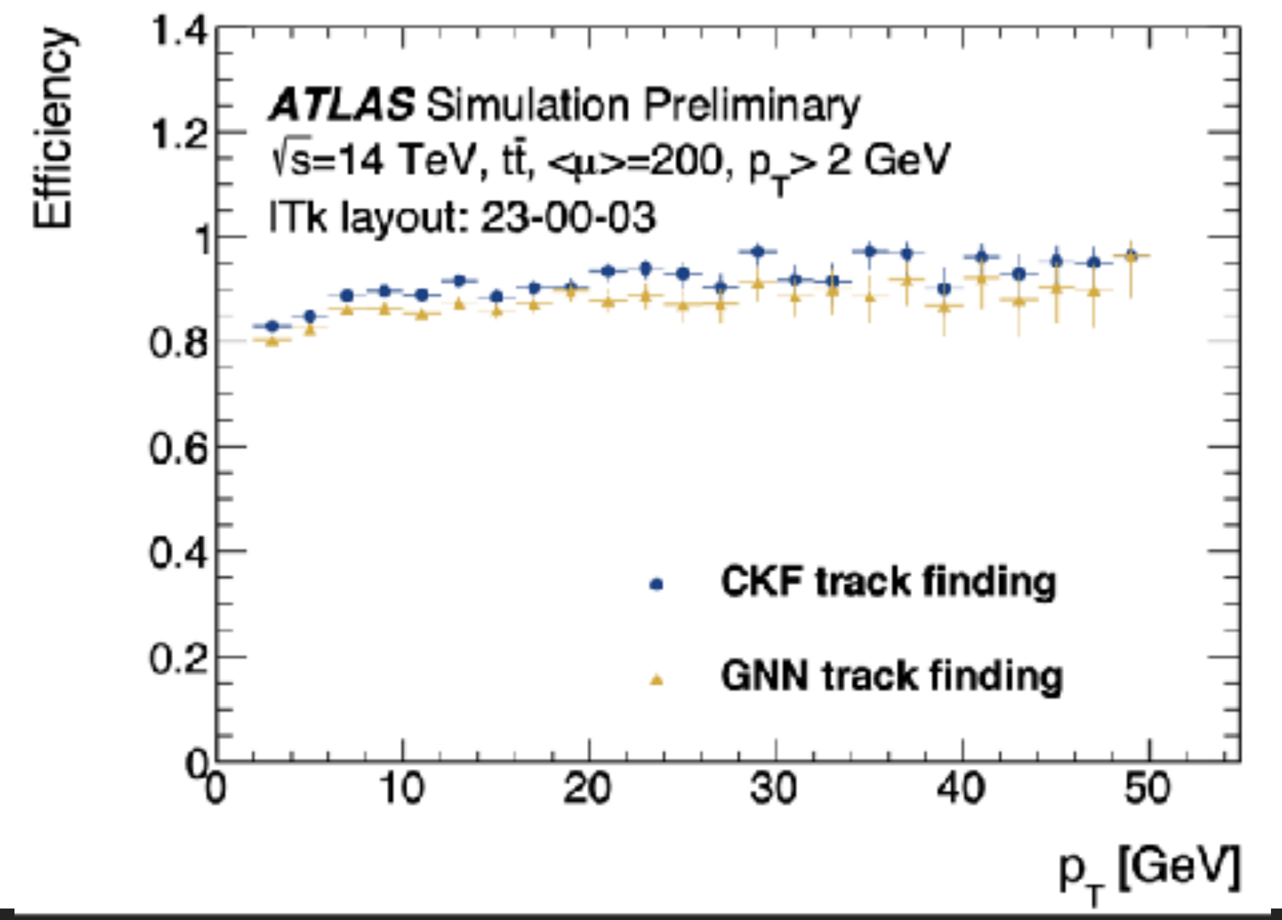


ML in Track Reconstruction

The goal of this Graph Neural Network (GNN) based pattern reconstruction is to identify the subsets of space points in the data that correspond to individual charged particles







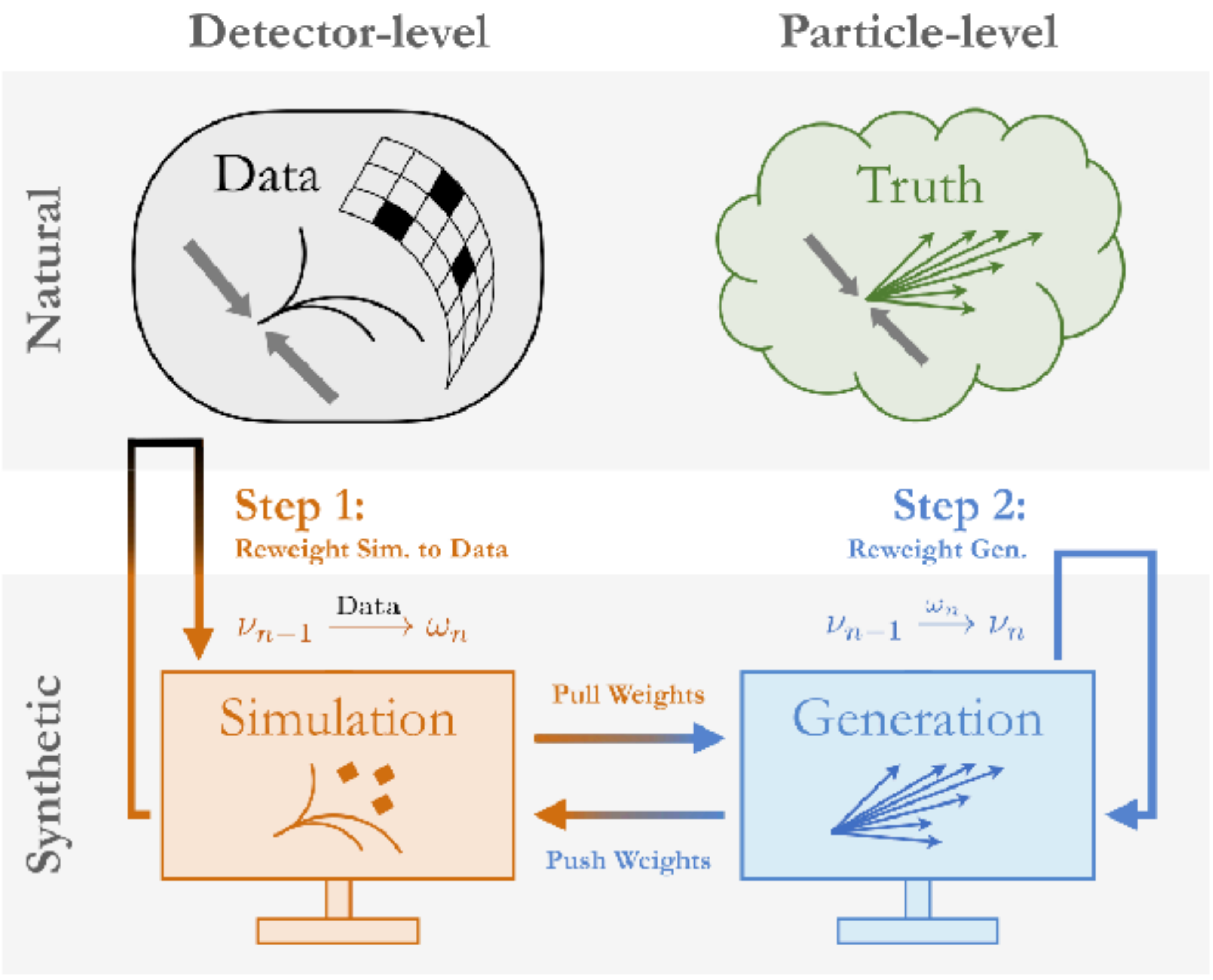
https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PLOTS/IDTR-2023-06/

ML in Data Unfolding

OMNIFOLD, an unfolding method that iteratively reweights a dataset using machine learning. The unbinned approach works for arbitrarily high-dimensional data, naturally incorporating information from the full phase space

OmniFold: A Method to Simultaneously Unfold All Observables

Anders Andreassen, Patrick T. Komiske, Eric M. Metodiev, Benjamin Nachman, and Jesse Thaler
Phys. Rev. Lett. **124**, 182001 – Published 7 May 2020



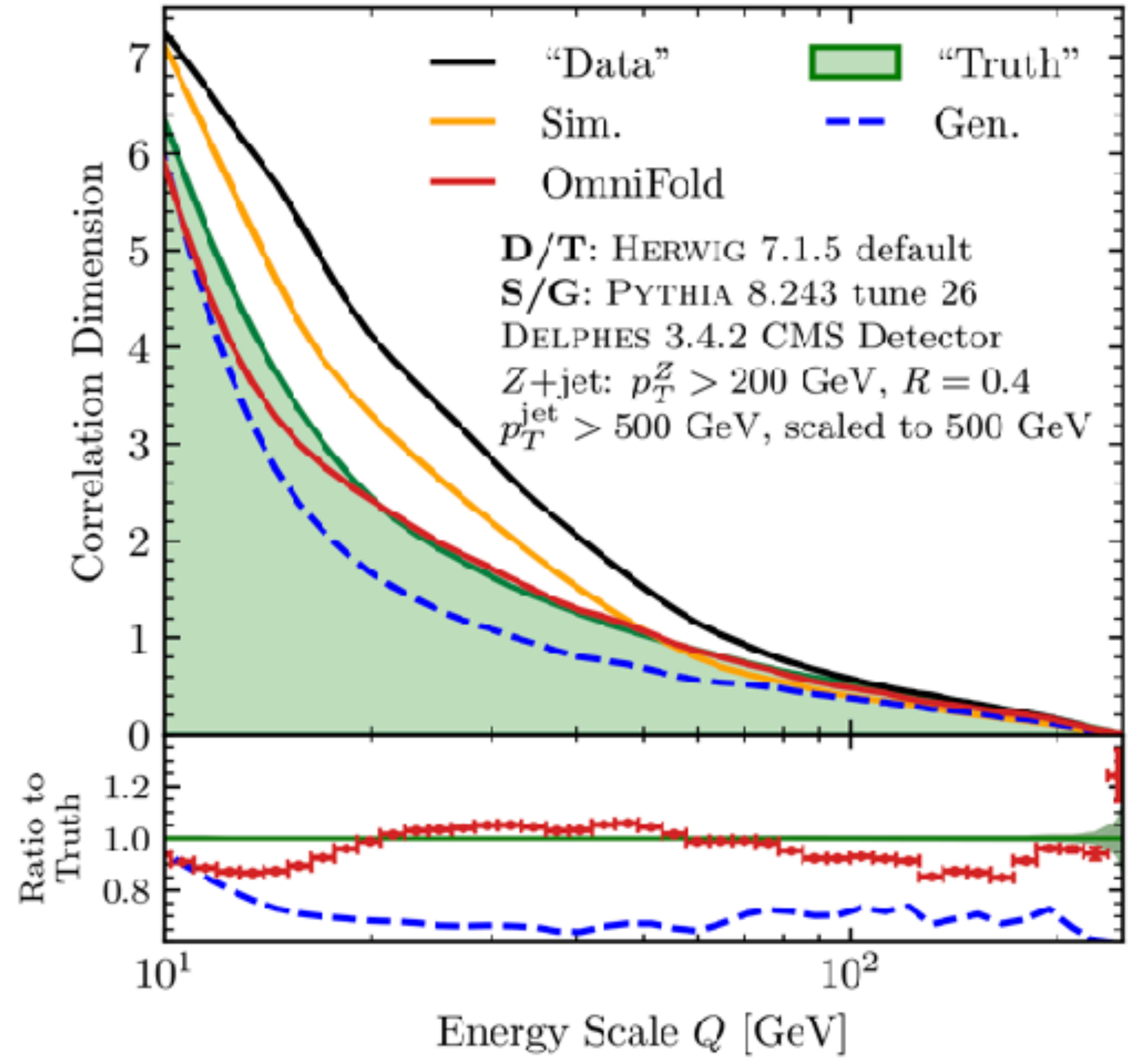


FIG. 3. The correlation dimension of the space of jets, unfolded with OMNIFOLD. The unfolded results closely match the truth-level dimension over most of the energy range, tending toward the prior in the more difficult phase space region at low Q. Unfolding a complicated statistic such as the correlation dimension is challenging with standard methods.

Foundation Models in HEP

Foundation models are multi-dataset and multi-task machine learning models that once pre-trained can be fine-tuned for a large variety of downstream applications

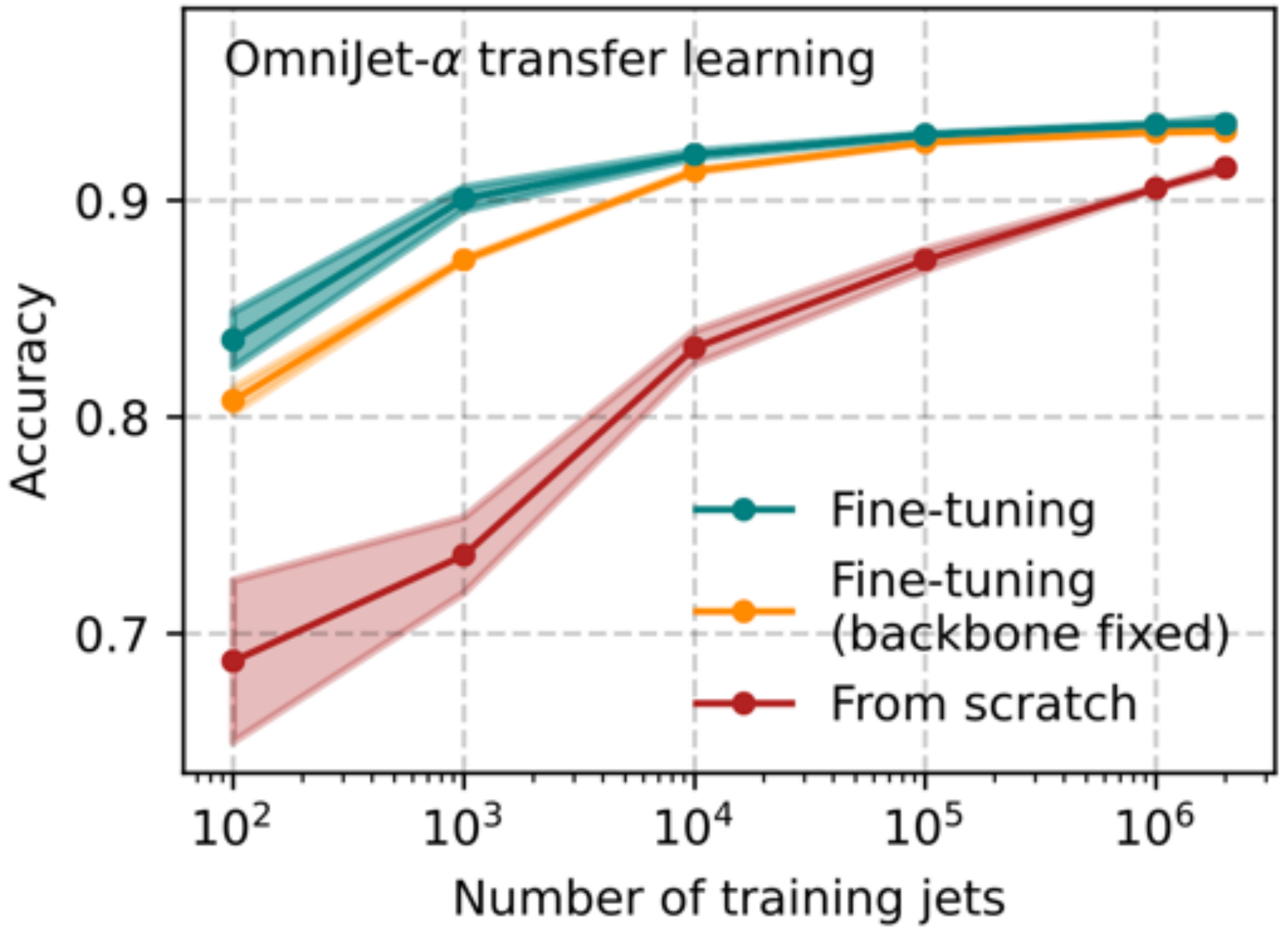
arXiv > hep-ph > arXiv:2403.05618

High Energy Physics – Phenomenology

[Submitted on 8 Mar 2024]

OmniJet-α: The first cross-task foundation model for particle physics

Joschka Birk, Anna Hallin, Gregor Kasieczka



Performance of pre-trained and non-pre-trained models for the task of t → bqq′ vs q/g jet classification.

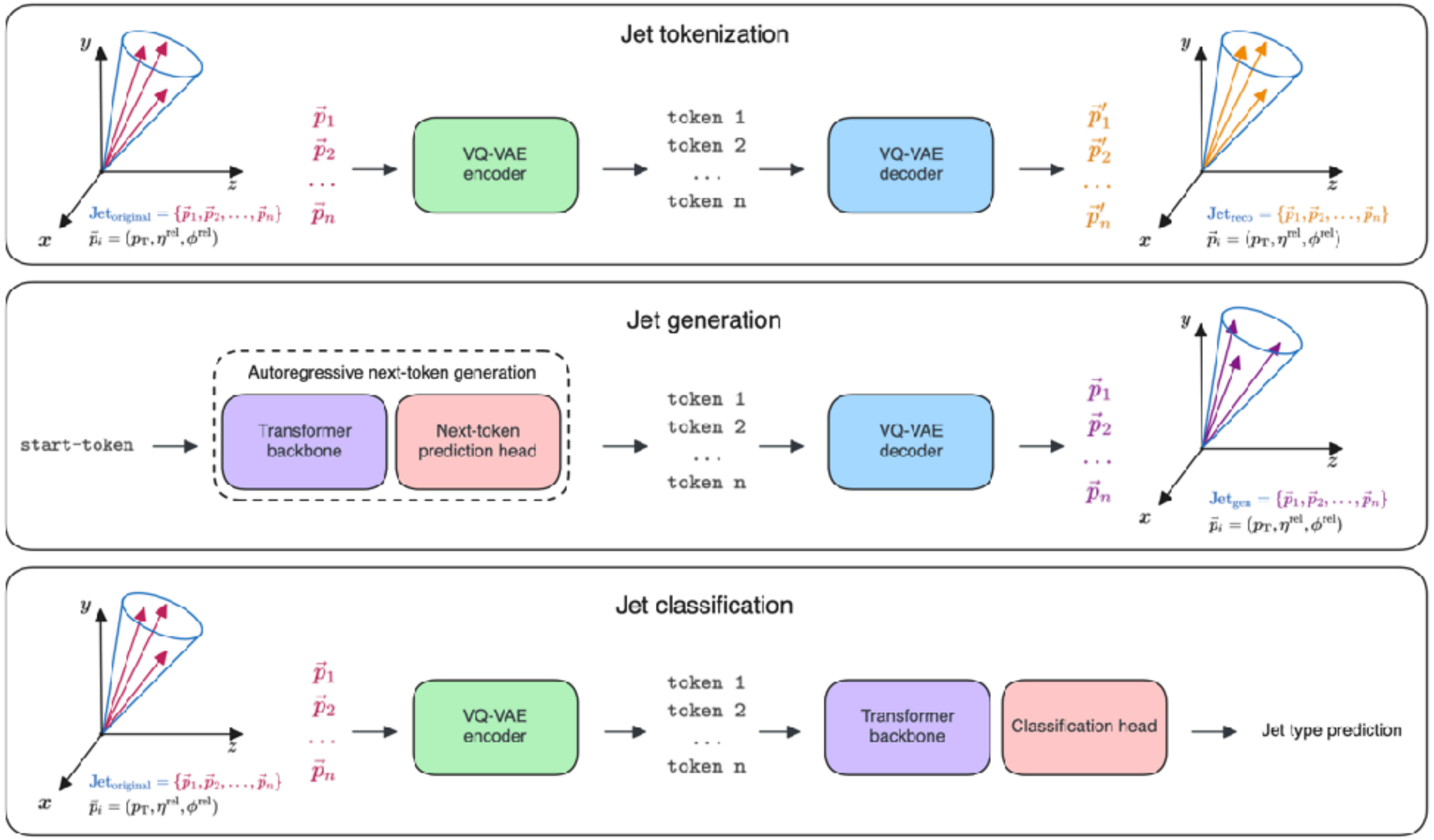


Figure 1: Schematics of the different steps (tokenization, generation, classification) in the OmniJet-α model.