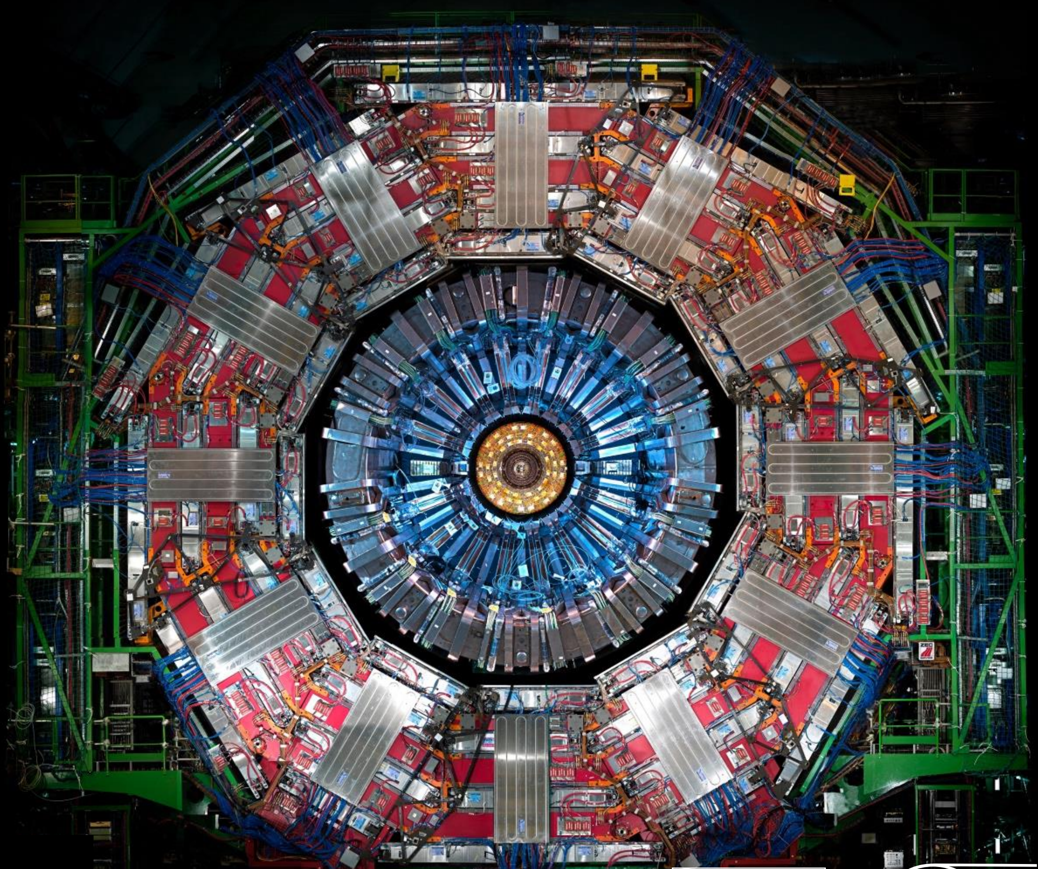
Anomaly Triggers for Exotics Searches



Jannicke Pearkes

Lepton Photon 2025

August 26th, 2025

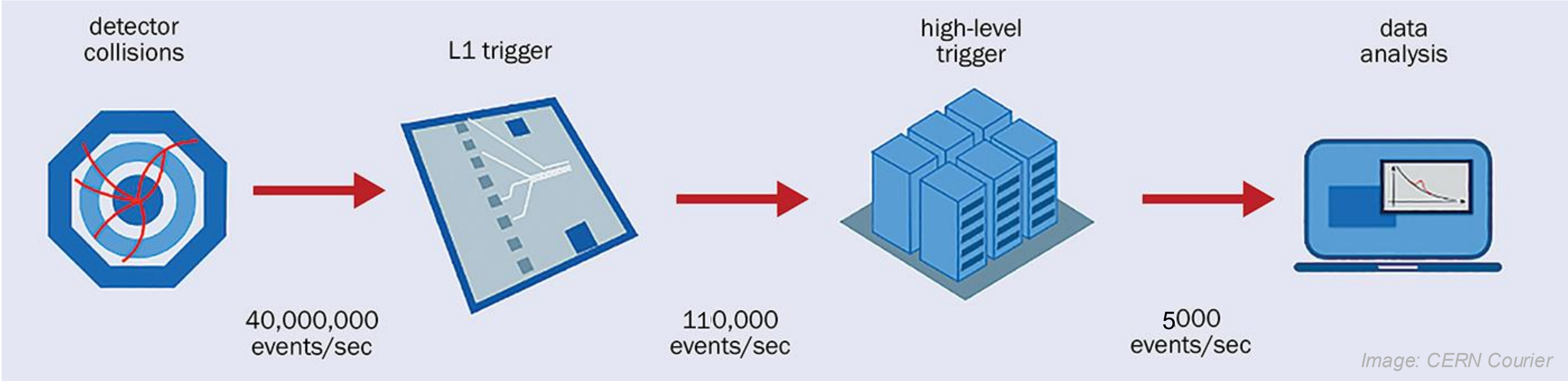
University of Colorado **Boulder**

NexTGen
Next Generation Triggers

CMS

CERN

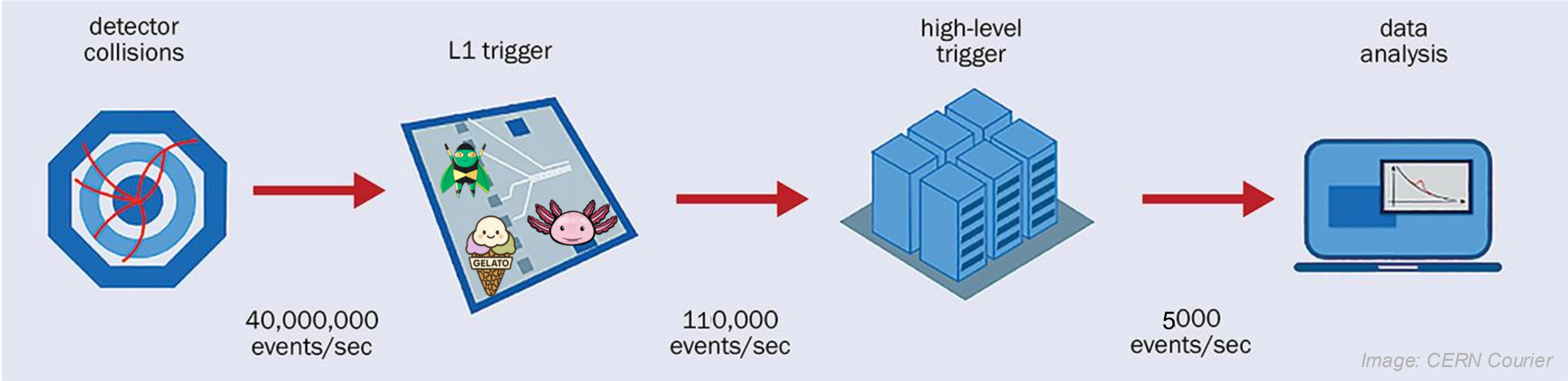
Why anomaly detection at L1?



L1 trigger filters out 99.75% of collision events 

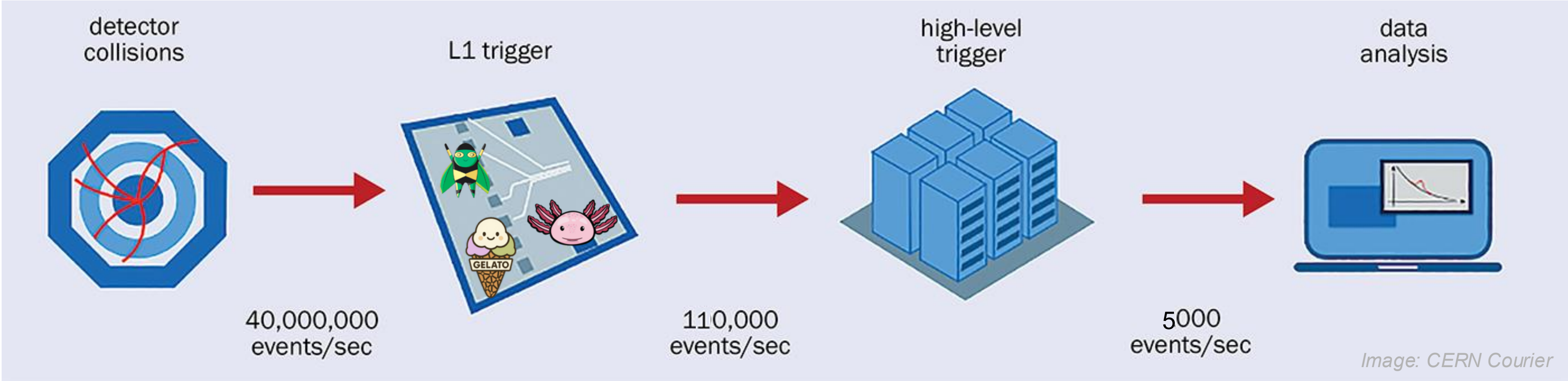
If we don’t identify interesting events in trigger, we lose them forever!

Why anomaly detection at L1?



"What if we are missing new physics because we are looking for the wrong thing?"

Why anomaly detection at L1?



"What if we are missing new physics because we are looking for the wrong thing?"

Anomaly detection algorithms use machine learning to be more model independent.

Anomaly Detection with Autoencoders

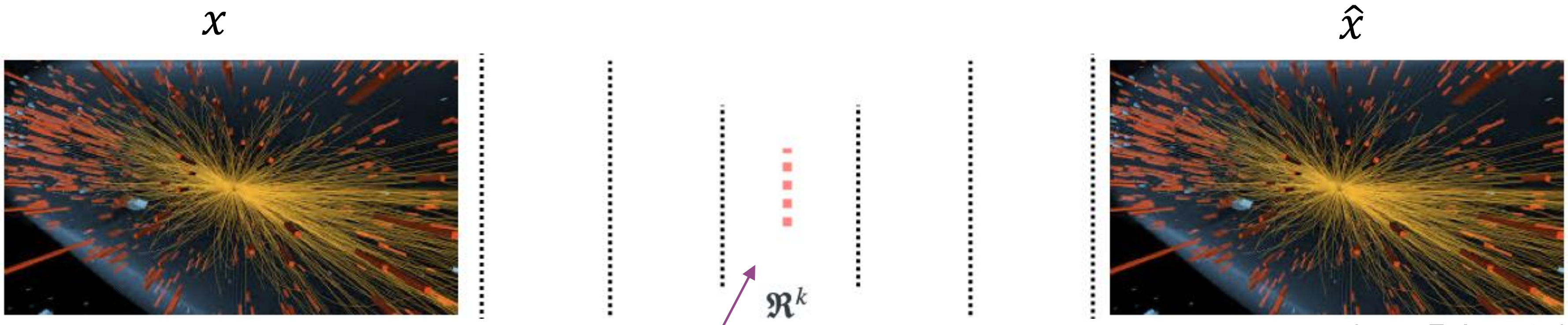


Image T. Aarrestad

Train on randomly sampled *data*

Bottleneck: autoencoder learns to compress high dimensional inputs into low dimensional latent space

Unsupervised learning: $x - \hat{x}$ represents degree of abnormality

Anomaly Detection with Autoencoders

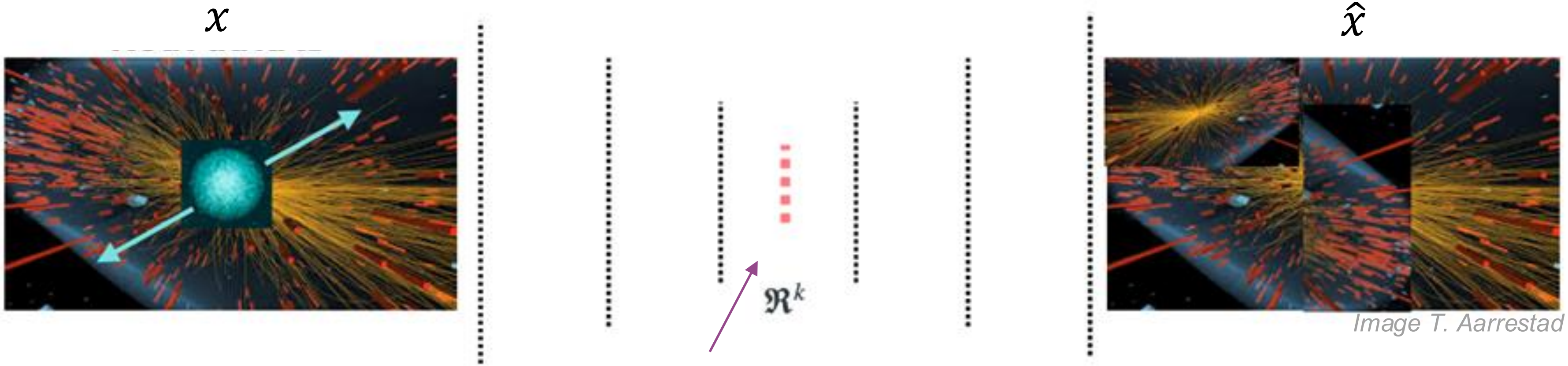


Image T. Aarrestad

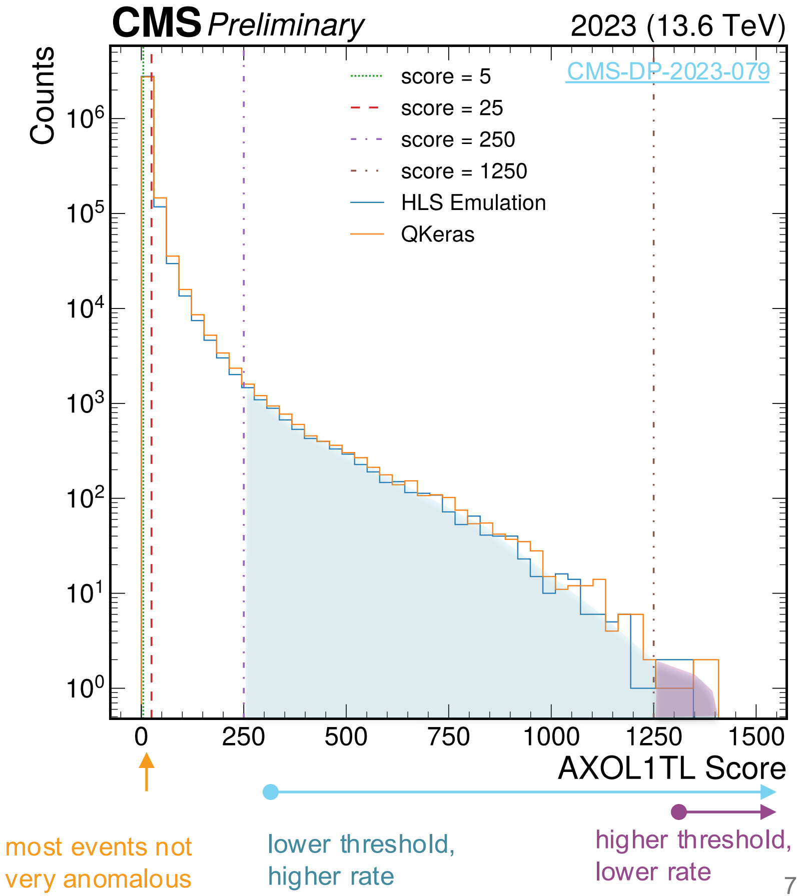
Train on randomly sampled *data*

Bottleneck: autoencoder learns to compress high dimensional inputs into low dimensional latent space

Unsupervised learning: $x - \hat{x}$ represents degree of abnormality

Anomaly Detection with Autoencoders

Different anomaly score thresholds are used to target different trigger rates



Key Assumptions for Anomaly Detection with Autoencoders

- Anomalies are *rare* in training set ✓ Exotics (usually)
- Anomalies are *out-of-distribution* ✓ Exotics (often)
- Anomalies are perceptible in available input objects ✓ Exotics (if the right objects are used)

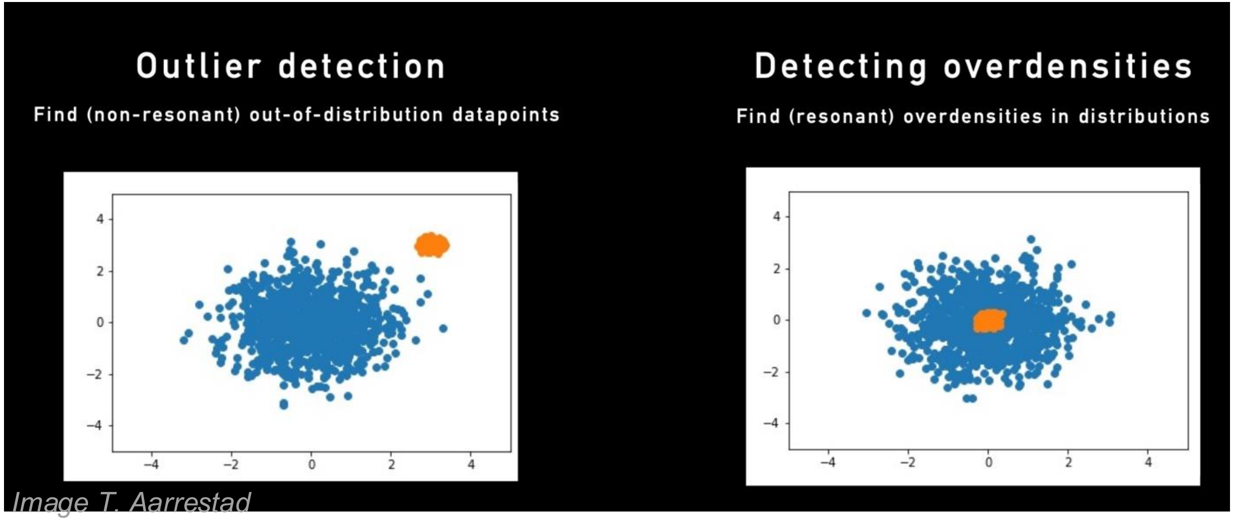


Image T. Aarrestad

Anomaly Detection in the L1 trigger

2024

May 2024

AXOL1TL starts taking data on CMS

Oct 2024

CICADA starts taking data on CMS

Talk by Isobel tomorrow

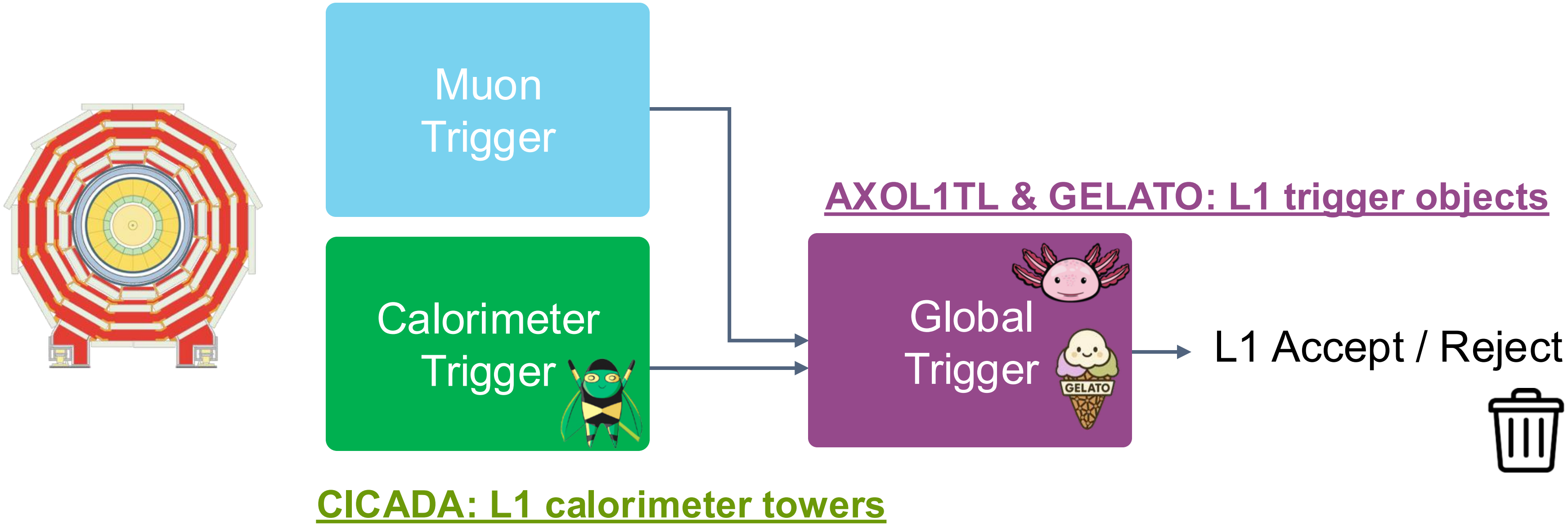
2025

May 2025

GELATO starts taking data on ATLAS

Talk by Kaito tomorrow

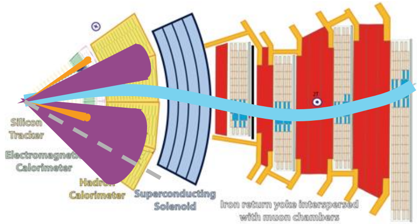
Anomaly Detection in the L1 Trigger



Algorithms must be lightweight enough to fit within the existing L1 trigger system. Latencies of 25 ns-100 ns

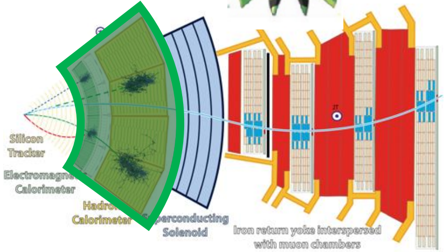
Differences: Inputs

AXOL1TL

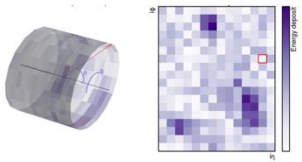


<u>L1 trigger objects (p_T, η, ϕ):</u>
MET, **10** jets, 4 muons,
4 **electrons / photons**
→ *56 variables total*

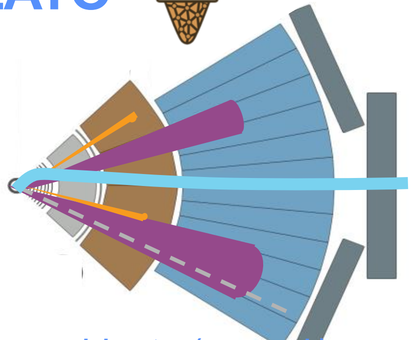
CICADA



<u>L1 calorimeter towers:</u>
E_T deposits corresponding to 14x18 towers in ηxϕ
→ *252 variables total*



GELATO

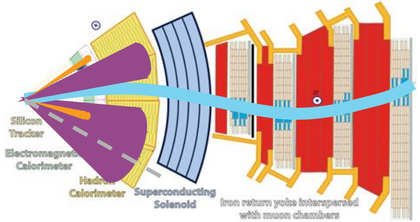


<u>L1 trigger objects (p_T, η, ϕ):</u>
MET, **6** jets, 4 muons,
4 **taus**
→ *44 variables total*

CICADA image by L. Gerlach

Differences: Inputs

AXOL1TL

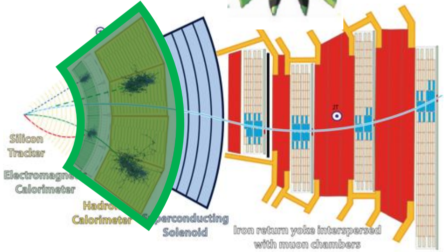


<u>L1 trigger objects (p_T, η, ϕ):</u>
MET, **10** jets, 4 muons,
4 **electrons / photons**
→ *56 variables total*

<u>Higher-level input features:</u>

- Cross-object final states
    - MET+X, Jets+X
- Signatures with muons
    - HNLs, VH

CICADA

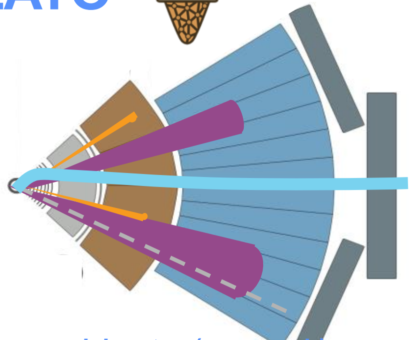


<u>L1 calorimeter towers:</u>
E_T deposits corresponding to 14x18 towers in η×ϕ
→ *252 variables total*

<u>Lower-level input features:</u>

- Interesting jet substructure:
    - Semi-visible jets
    - Emerging jets
- Event level calorimeter deposits
    - SUEPs

GELATO



<u>L1 trigger objects (p_T, η, ϕ):</u>
MET, **6** jets, 4 muons,
4 **taus**
→ *44 variables total*

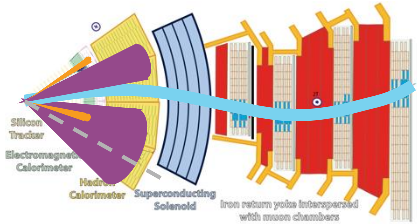
<u>Higher-level input features:</u>

- Same as AXOL1TL
- + signatures with taus

See Roy Cruz's poster here at LP2025!

Differences: Inputs

AXOL1TL

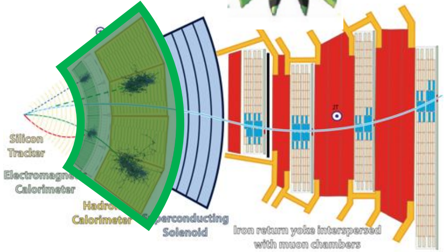


<u>L1 trigger objects (p_T, η, ϕ):</u>
MET, **10** jets, 4 muons,
4 **electrons / photons**
→ *56 variables total*

<u>Higher-level input features:</u>

- HNLs, VH

CICADA

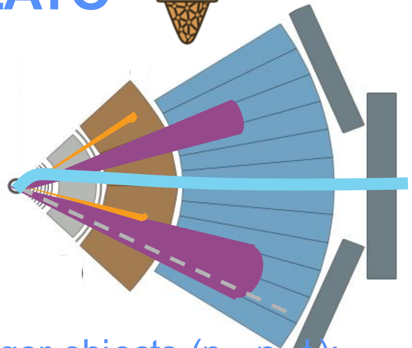


<u>L1 calorimeter towers:</u>
E_T deposits corresponding to 14x18 towers in $\eta \times \phi$
→ *252 variables total*

<u>Lower-level input features:</u>

deposits (SUEPs)

GELATO



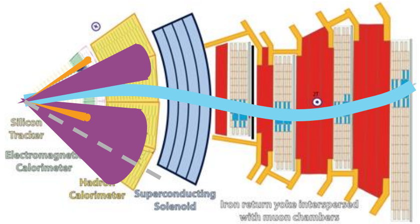
<u>L1 trigger objects (p_T, η, ϕ):</u>
MET, **6** jets, 4 muons,
4 **taus**
→ *44 variables total*

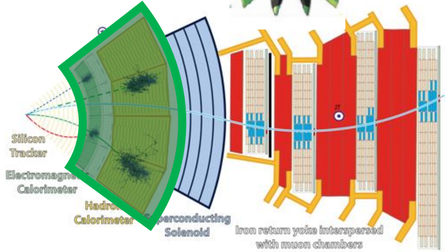
<u>Higher-level input features:</u>

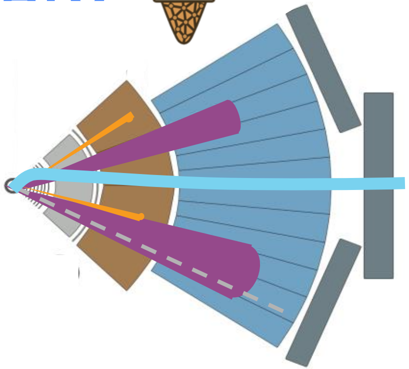
Notes: Sensitivity still to be fully demonstrated by analyses using these triggers. L1 trigger often already very efficient, AD not a replacement for targeted triggers.

See Roy Cruz's poster here at LP2025!

Differences: L1 Strategy

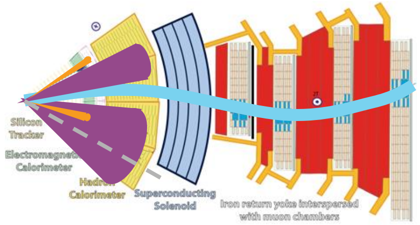





	AXOL1TL	**CICADA**	**GELATO**
Model:	Variational Autoencoder	Autoencoder	Variational Autoencoder
Implementation strategy:	Only encoder deployed 50 ns latency	Knowledge distillation with student model deployed 100 ns latency	Only encoder deployed 25 ns latency
L1 Unique rate:	250 Hz target	250 Hz target	500 Hz target

Differences: HLT Strategy

AXOL1TL

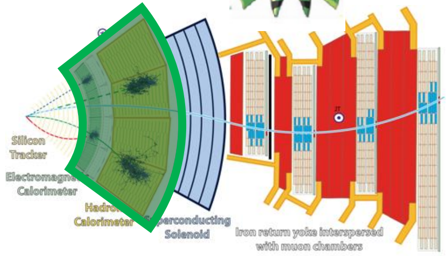


HLT Strategy:

Passthrough trigger at HLT

- Scouting & full reconstruction streams

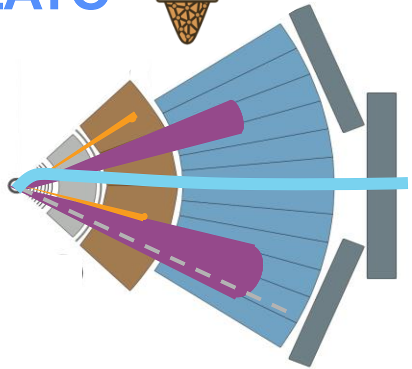
CICADA



Passthrough trigger at HLT

- Scouting & full reconstruction streams

GELATO



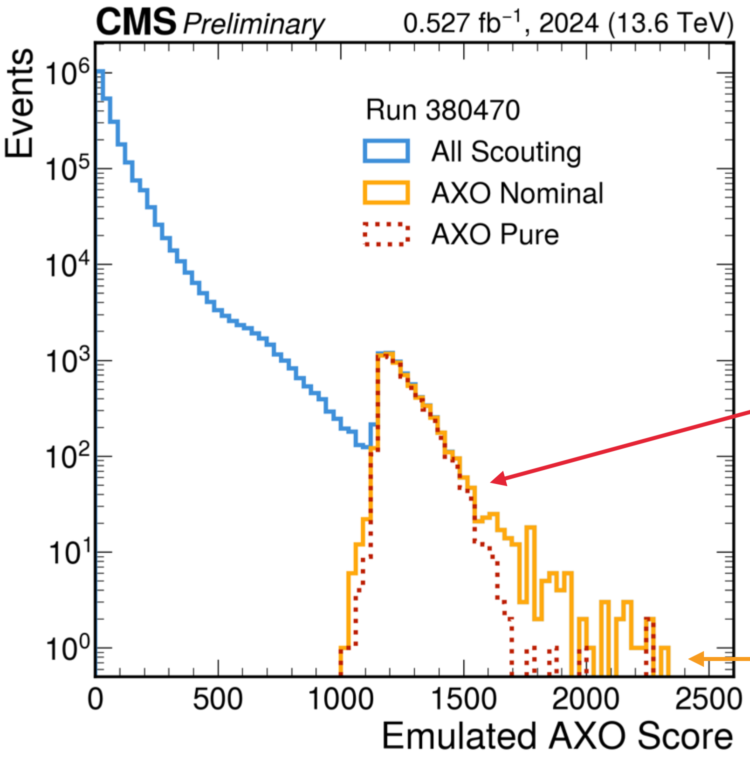
Variational autoencoder at HLT

MET, **6** jets, 3 electrons, 3 photons, 3 muons → *47 variables total*

10 Hz unique rate

AXOL1TL triggered events

CMS-DP-2024-059



Large fraction of unique events recorded that would otherwise be rejected

High anomaly score events, also triggered by existing L1 trigger

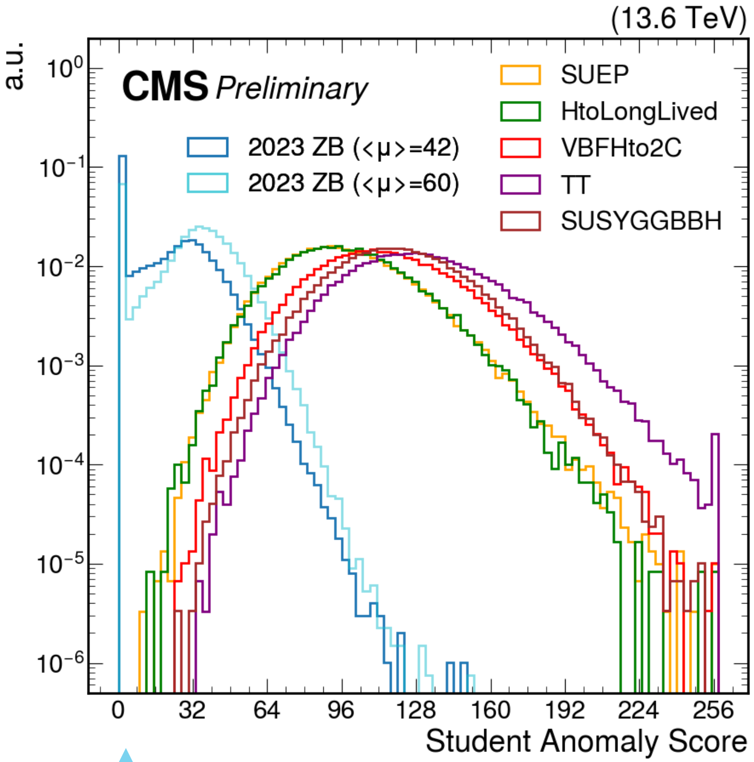
What might we be sensitive to?

CMS-DP-2024-121

Nice separation between Zero Bias data and BSM signatures such as SUEPs and H→SS → 4b (cτ = 900mm)

Important caveat: Domain shift between data (training) and MC (evaluation)

Developing techniques for studying trigger efficiencies in data and evaluating on standard candles.

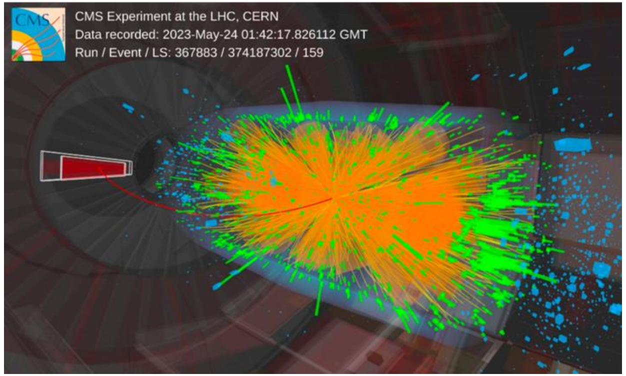


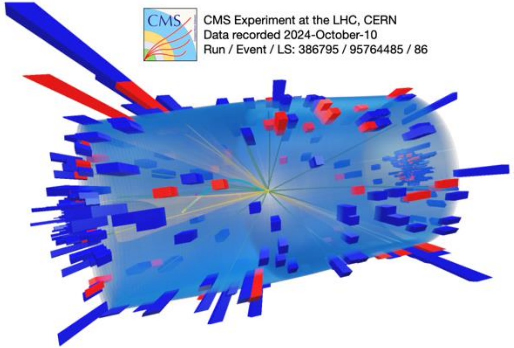
Zero Bias data events have low anomaly score

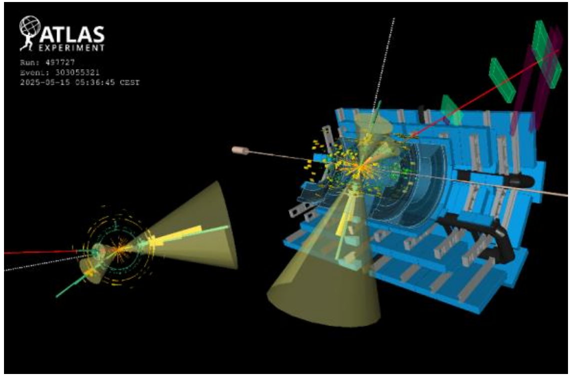
Signal Monte Carlo has high anomaly score

Examples of anomalous events:

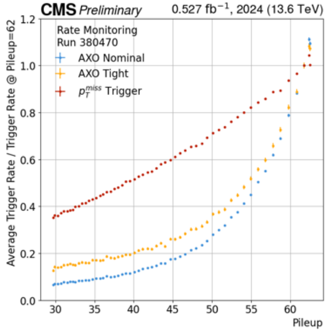
AXOL1TL **CICADA** **GELATO**







High object multiplicity a key feature of autoencoder based triggers



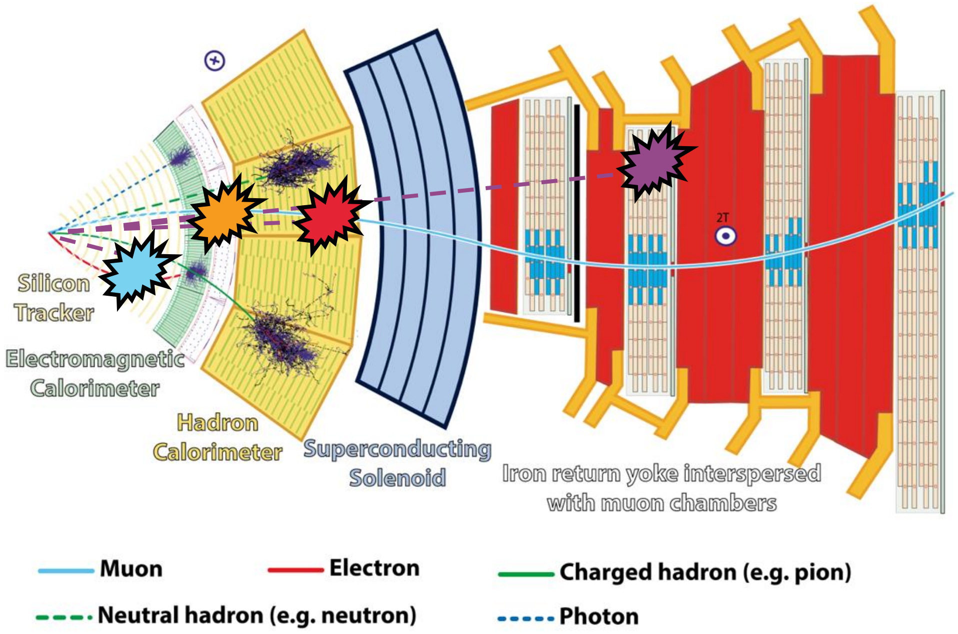
Many options for reducing the dependence on object multiplicity:

- Wasserstein normalized autoencoders
- Pile-up reweighting of anomaly score
- Contrastive learning + data augmentation for robustness

Ideas for increasing AD sensitivity to exotic signatures in the future:

Include new objects as inputs:

- Muon detector showers
- Muon displacement
- HCAL, ECAL timing
- HCAL / ECAL energy ratios
- L1 tracking (CMS HL-LHC)



Beyond variational autoencoders:

- Semi-supervised algorithms?
- Making use of learned embeddings?
- Real-time over-density estimation?

How to do analysis with anomaly detection triggers?

increasing order of novel challenges

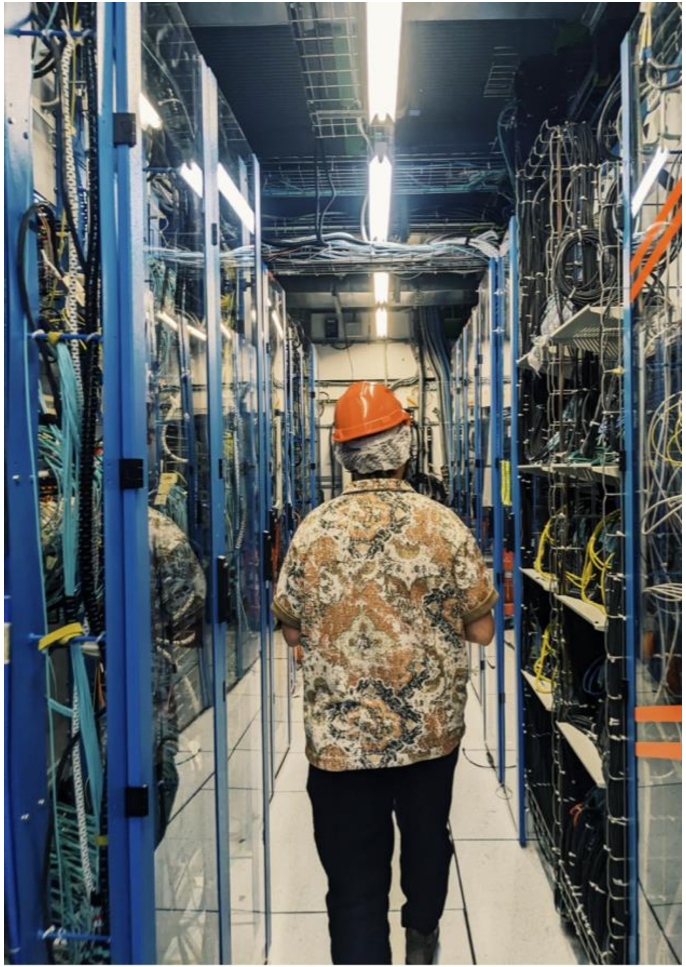
1. Do a bump hunt

2. Repeat an existing targeted search with AD triggered events
   - Potentially useful for searches without dedicated triggers

3. Try a completely new technique
   - E.g. second layer of anomaly detection offline

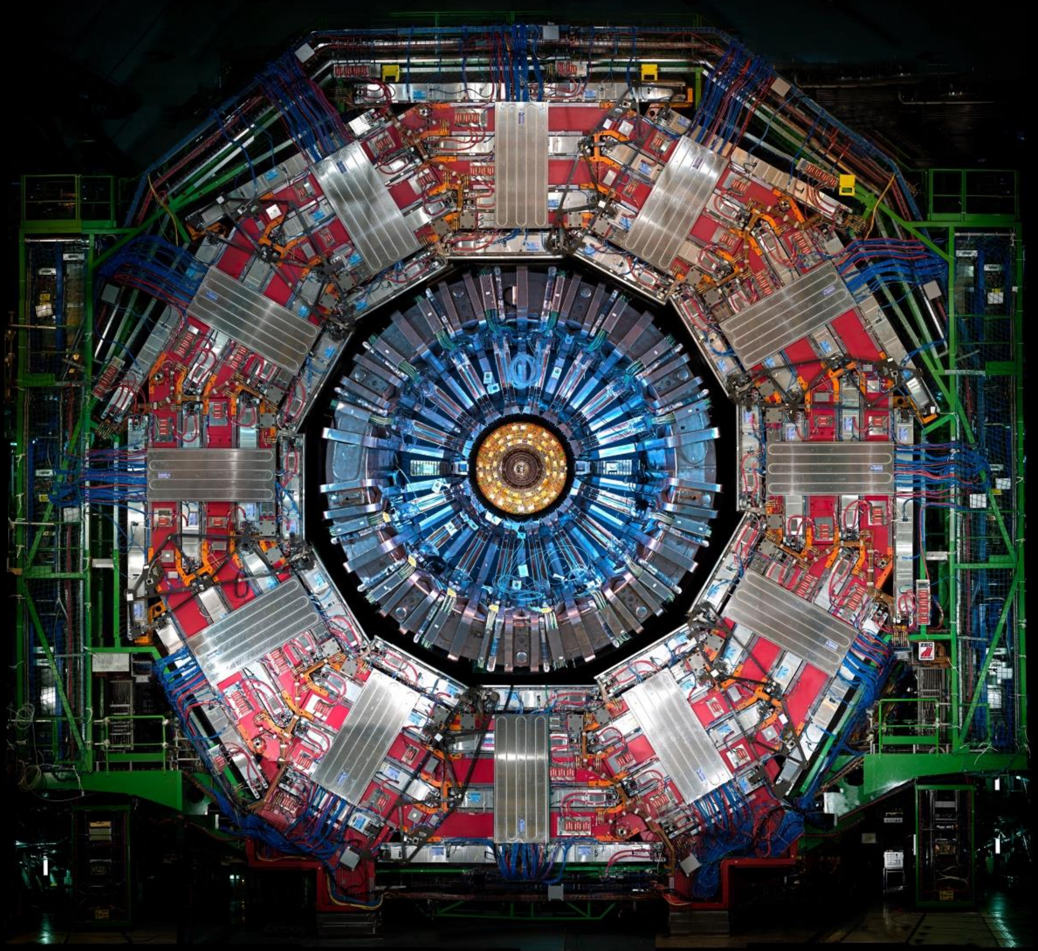
Summary

- Anomaly detection triggers offer a new, model-agnostic approach for triggering on interesting physics
- CMS has been taking data with AXOL1TL and CICADA since 2024. ATLAS has started taking data with GELATO in 2025
- These are the first ML-based anomaly detection triggers deployed at the LHC
- Many extensions in the future are possible
- This is just the very beginning!



Thank you!

Come join the dedicated Triggers AI/ML session (link) tomorrow if you want to hear more!



References:

CMS-DP-2023-079 - AXOL1TL (2023)

CMS-DP-2024-059 - AXOL1TL (2024)

CMS-DP-2023-086 - CICADA (2023)

CMS-DP-2024-121 – CICADA (2024)

ATL-COM-DAQ-2025-039 – GELATO (2025)

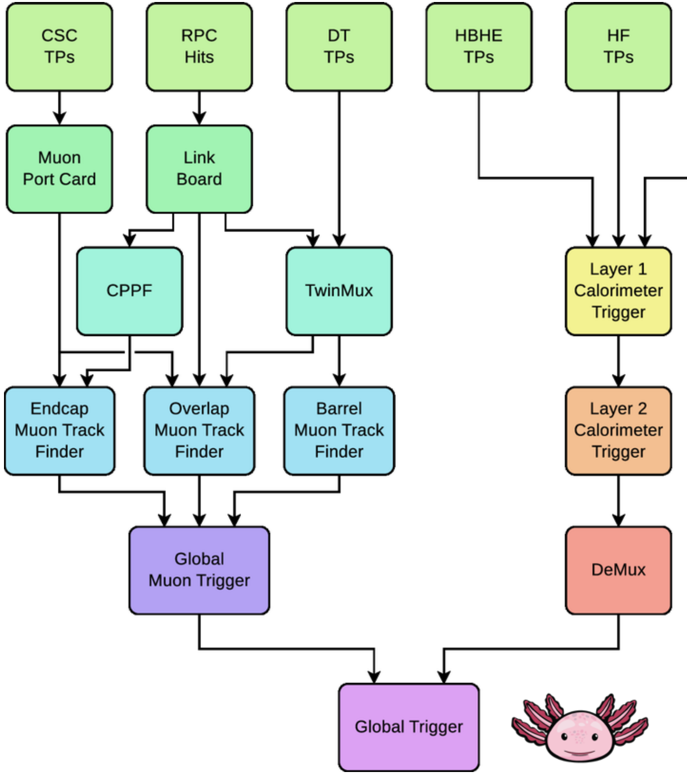
CMS-DP-2024-099 – Muon detector showers

CMS-DP-2023-056 – Muon impact parameters in the barrel

CMS-DP-2024-058 – HCAL depth and timing

CMS-TDR-021 – CMS Phase II L1 Trigger TDR

AXOL1TL & CICADA in the L1 Trigger



<u>CICADA L1 calorimeter towers are inputs:</u>
252 E_T deposits corresponding to 14x18 towers in $\eta x \phi$

<u>AXOL1TL L1 trigger objects are inputs:</u>
MET - (p_T, ϕ)
10 jets - (p_T, η, ϕ)
4 muons - (p_T, η, ϕ)
4 electrons / photons - (p_T, η, ϕ)

CMS AD HLT Strategy 2024

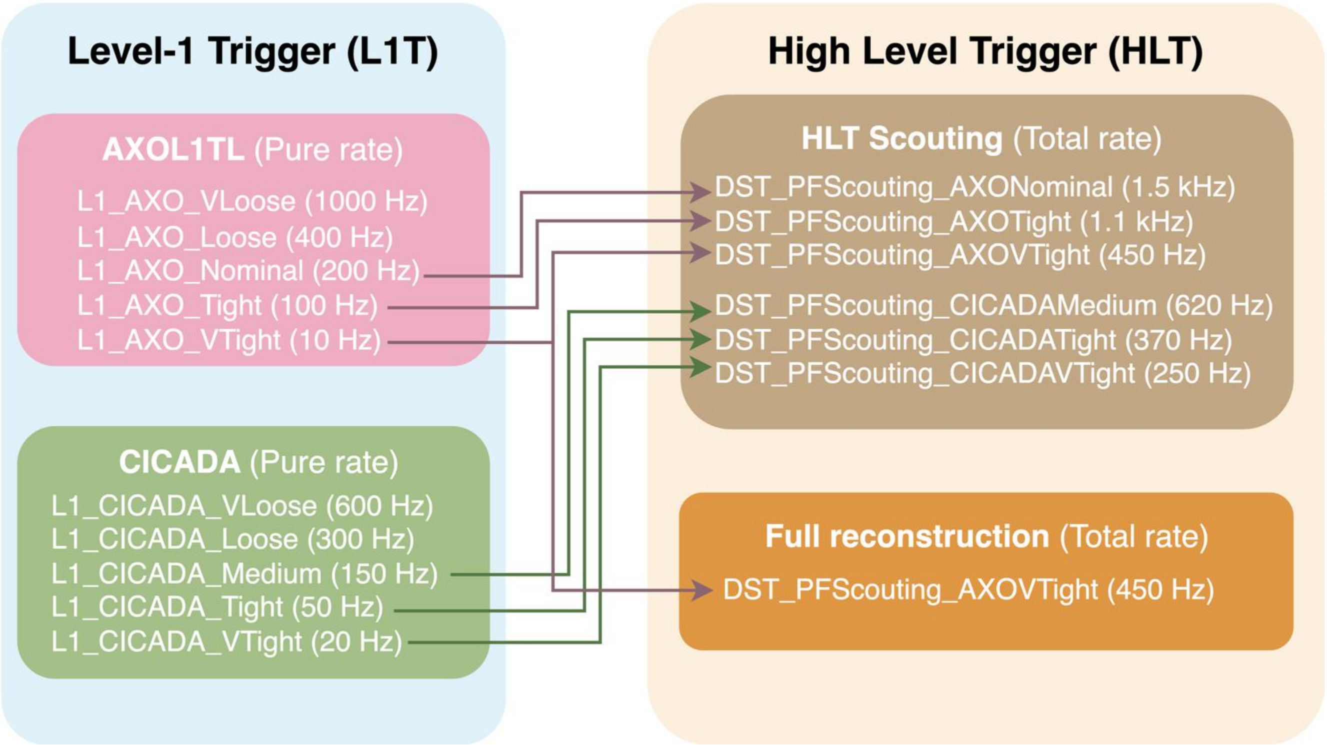
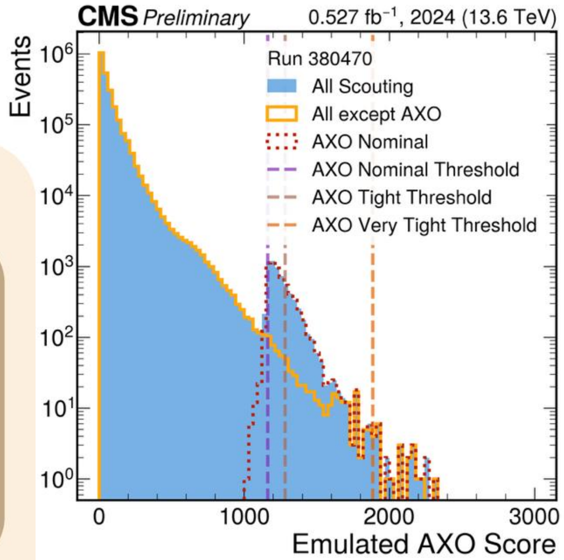


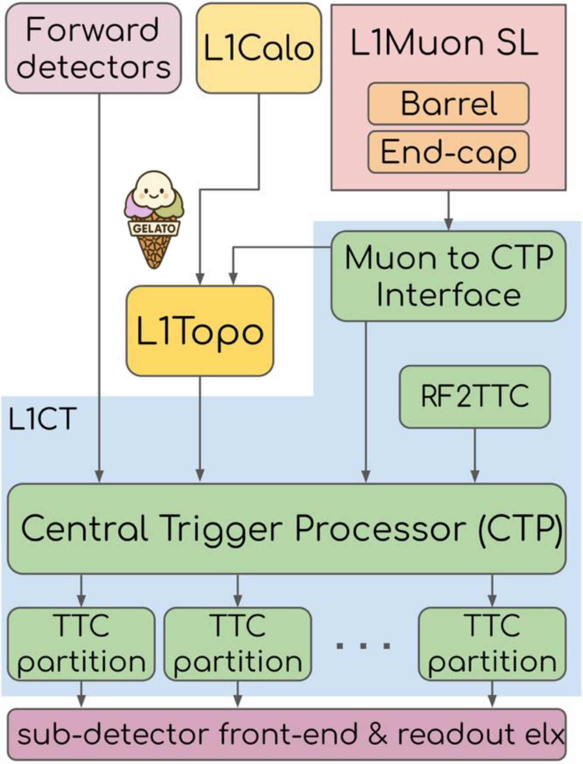
Image: S. Giorgetti



Rates shown are NNv3 target rates

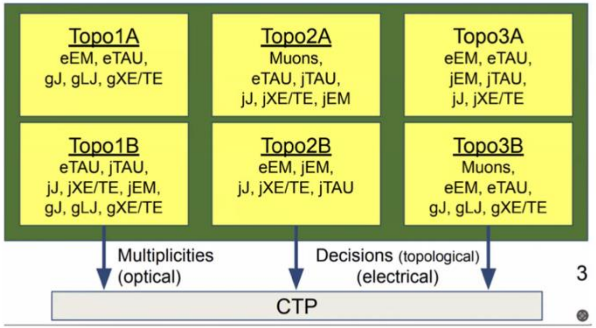
Passthrough trigger:
Anything accepted at L1 is also accepted by the HLT

ATLAS GELATO in the L1 Trigger



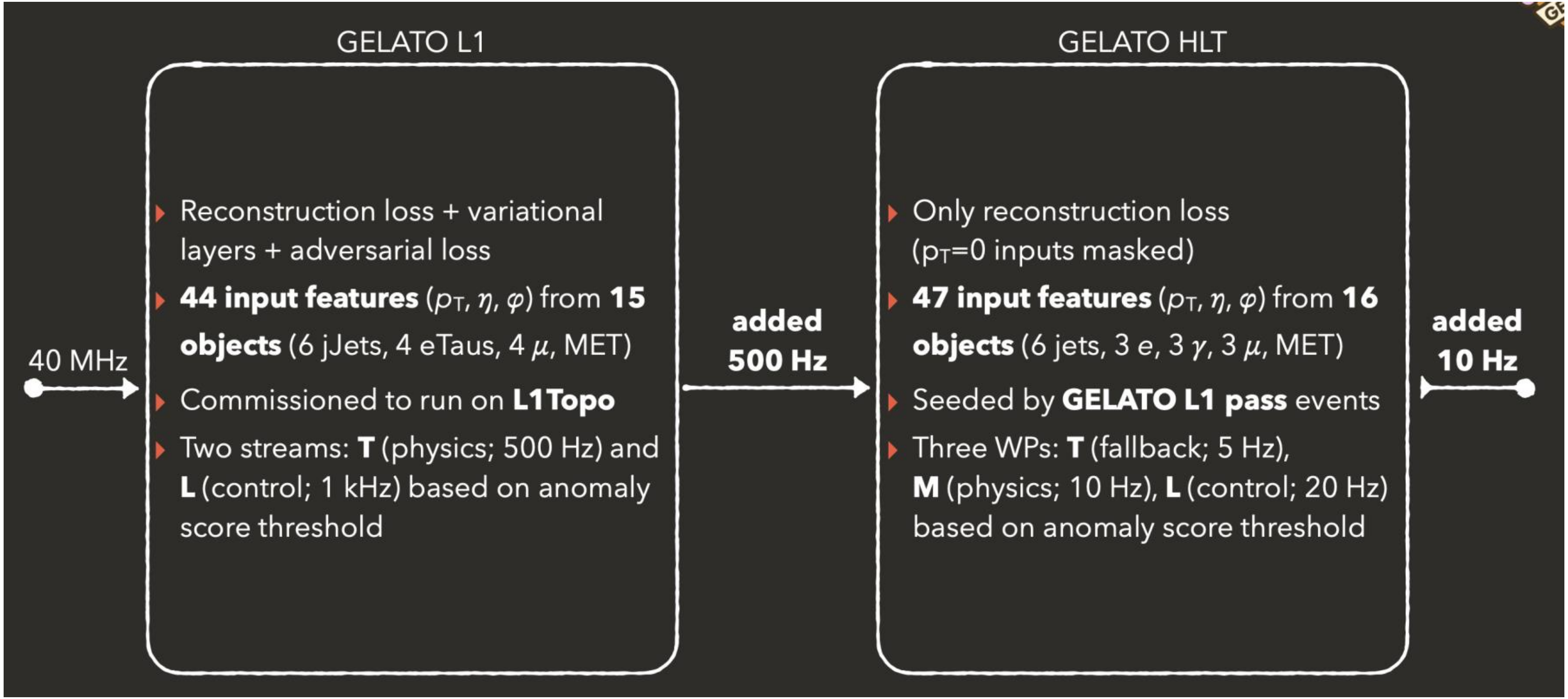
- Each L1Topo board has 2 FPGAs
- Each FPGA receives a subset of physics objects
- Choosing which FPGA to sit on determines which objects we can use for GELATO L1

M. Cohen



ATLAS AD Strategy 2025

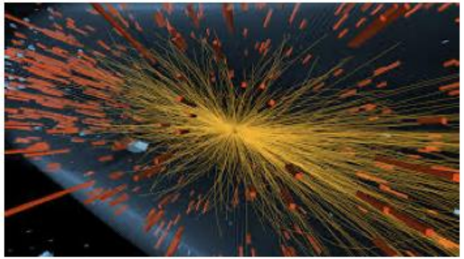
S. Addepalli, ML4Jets

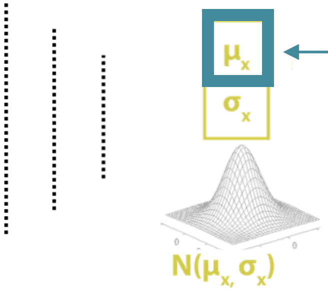


AXOL1TL Implementation

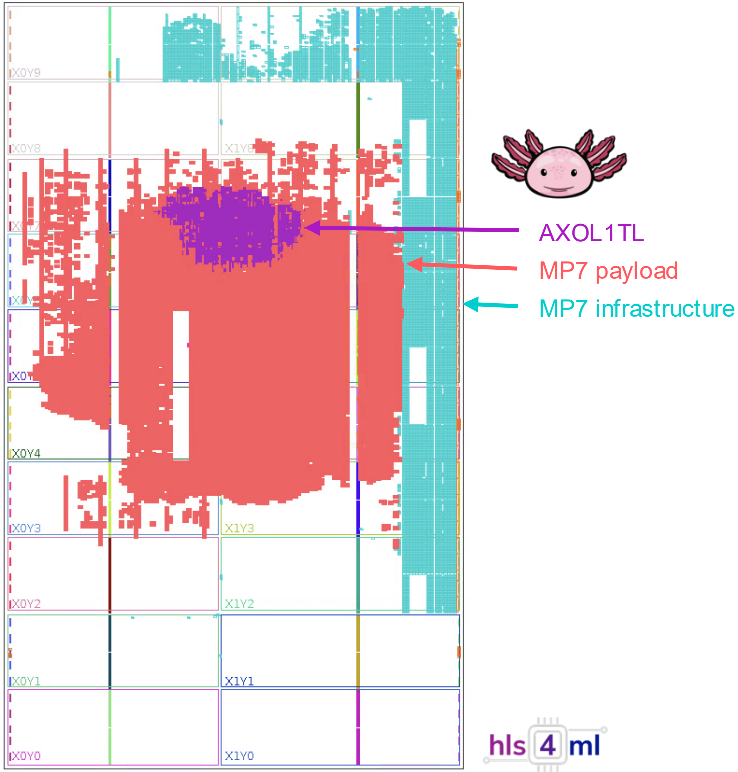
[CMS-DP-2023-079]

- Only deploy encoder half of the network, compute degree of abnormality from latent space directly
- Halves the network size and latency





μ^2 term used for anomaly score



Implemented on Xilinx Virtex-7 FPGA

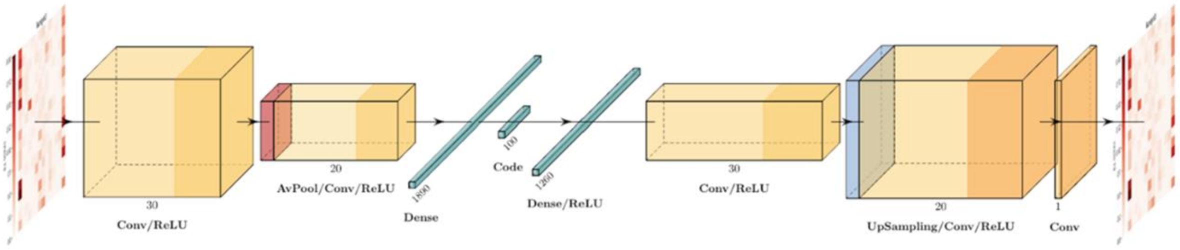
50 ns latency and resource requirements met

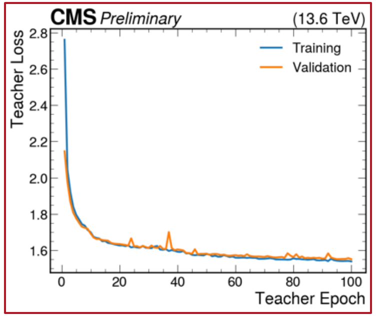
CICADA Implementation

[V. Sharma PPC 2024](#)
[CMS-DP-2024-121](#)

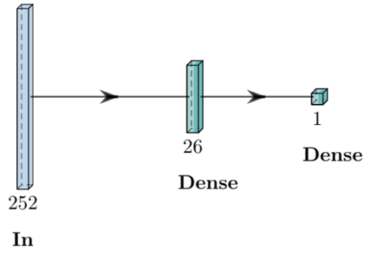
Knowledge distillation: student learns from teacher model

The teacher model

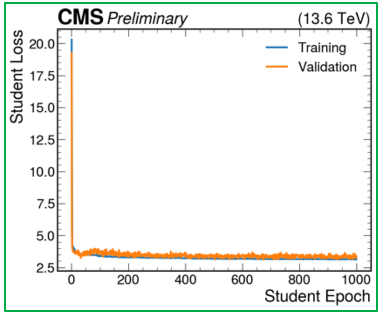




The student model

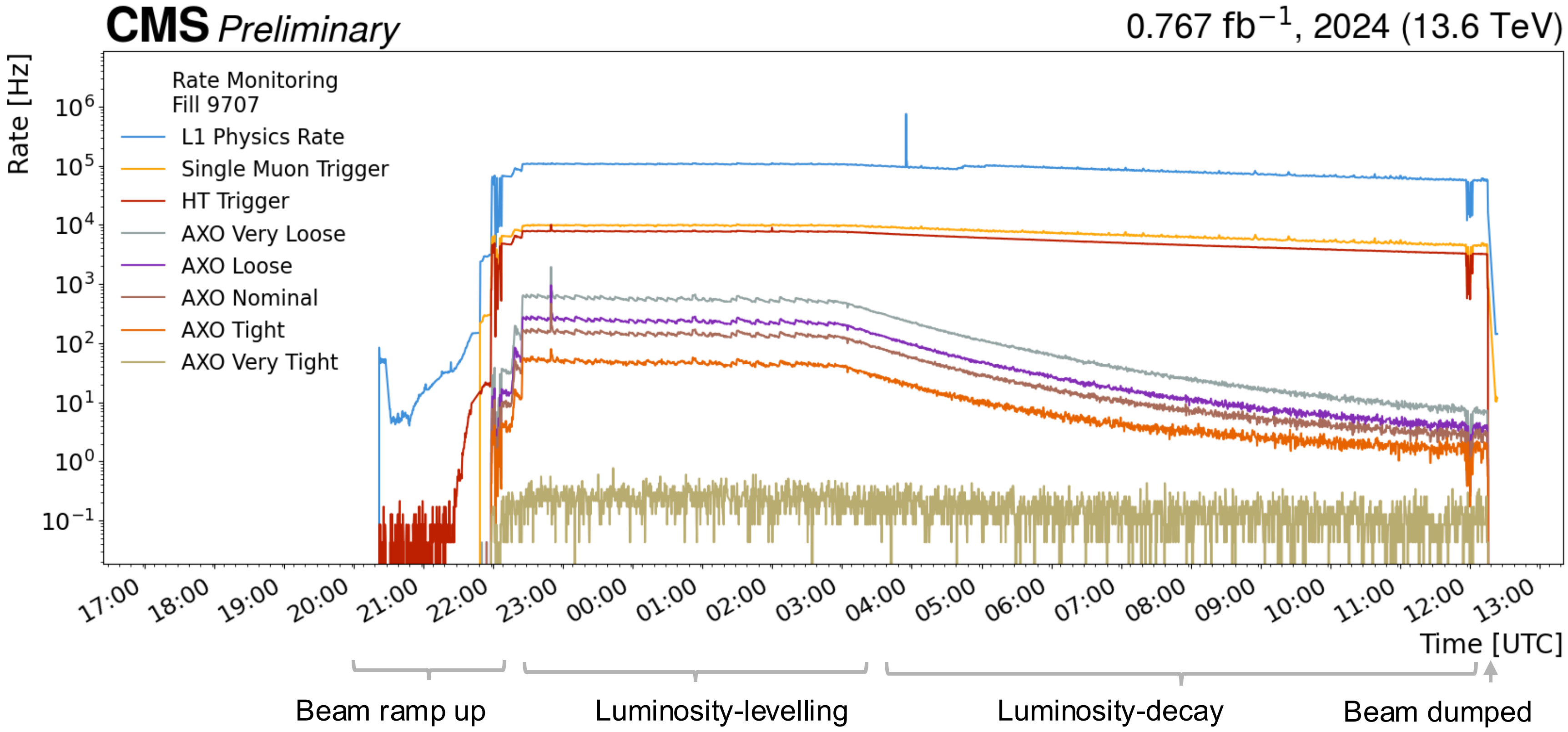


Inference latency ~ 100 ns
on Virtex-7 FPGA



AXOL1TL Rate Stability

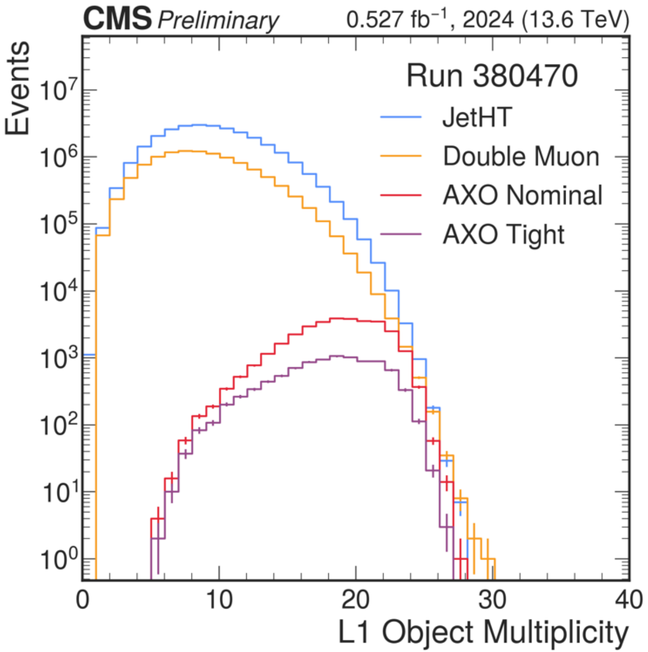
CMS-DP-2024-059

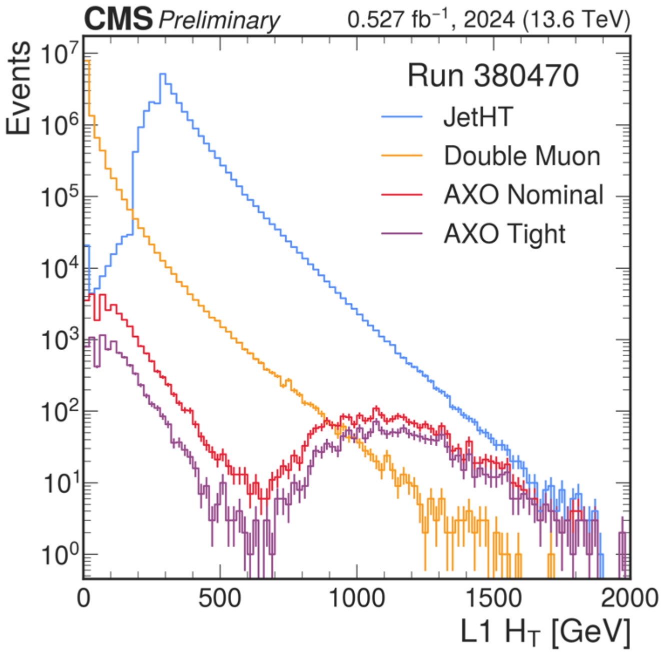


What does AXOL1TL trigger on?

CMS-DP-2024-059

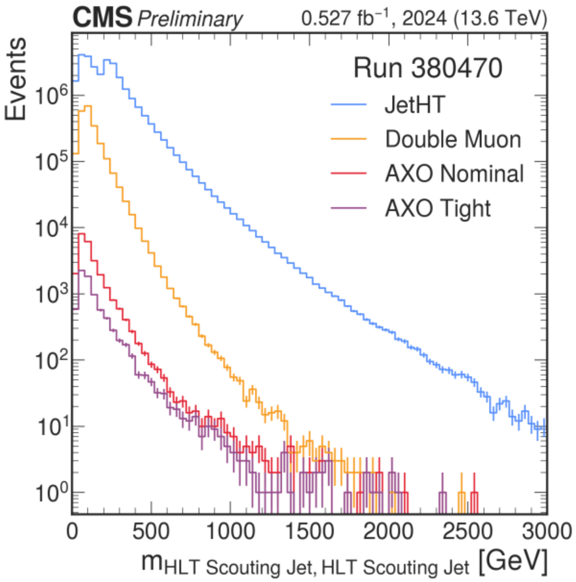
High object multiplicity and total transverse momentum.

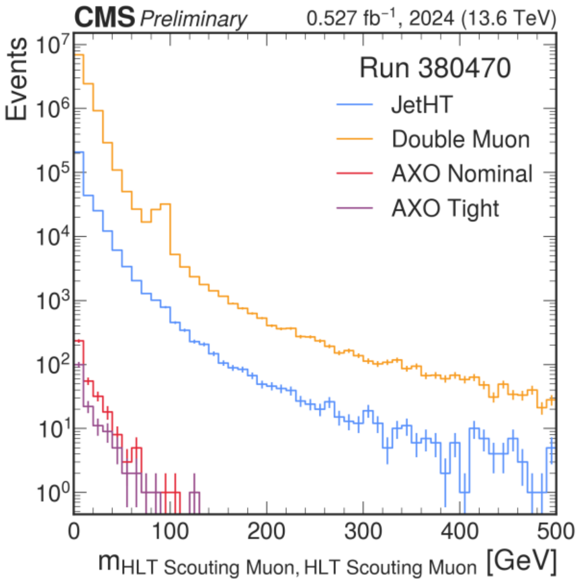


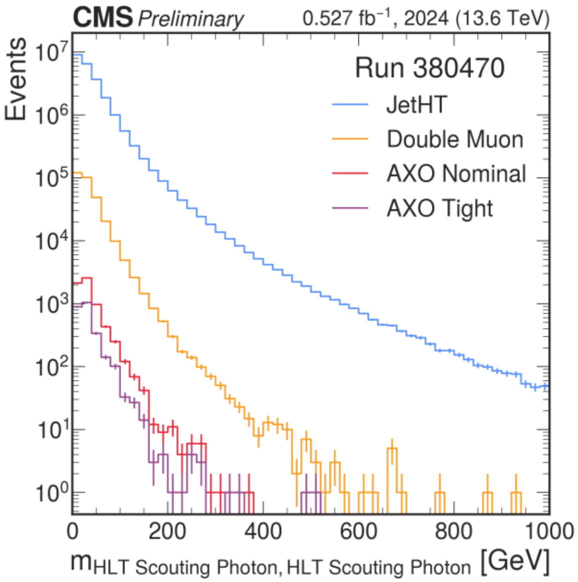


Peak at AXO triggered data:

- Smoothly falling mass distributions shown here in small fraction (<1%) of the 2024 data
- More plots in our DP Note: https://twiki.cern.ch/twiki/bin/view/CMSPublic/AXOL1TL2024







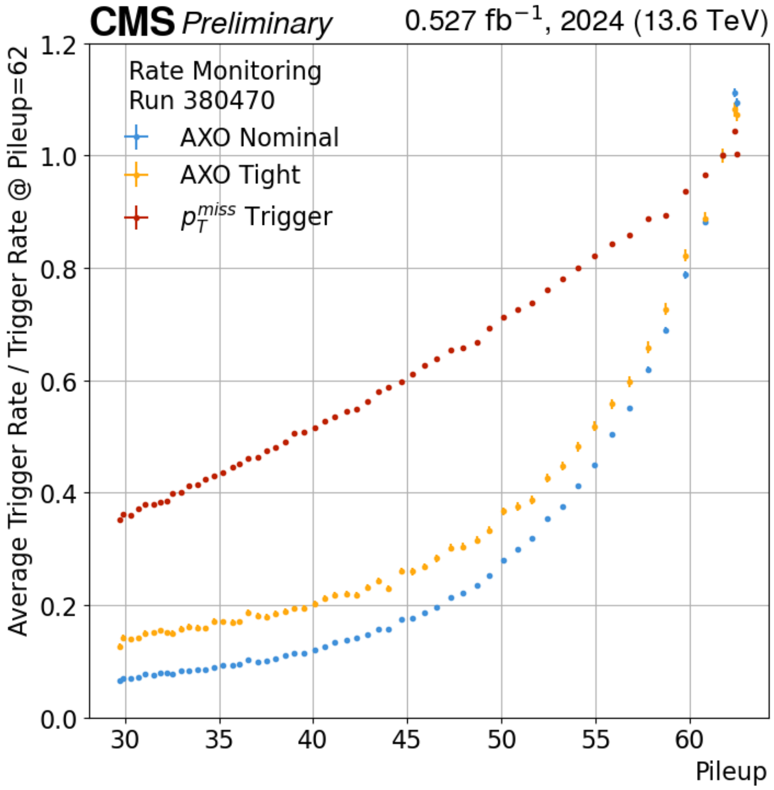
Challenge: Pile-up Dependence

CMS-DP-2024-059

Events with high pileup contain more "information" (more jets, more calorimeter cells) are inherently harder to encode. This contributes to higher rates at high pileup.

Care must be taken during model development and deployment:

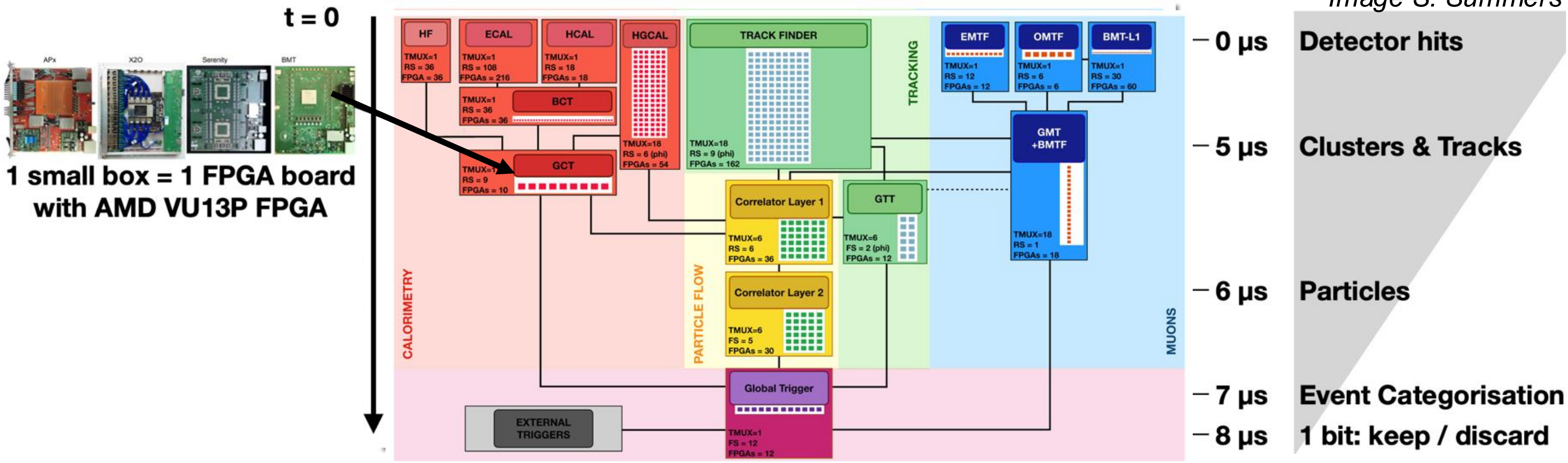
- more robust training procedures
- back-up paths & "emergency off" columns
- conservative rate estimates



HL-LHC Upgrades to the CMS L1 Trigger

- Machine Learning heavily incorporated into upgraded L1 trigger design
- Anticipate **25 billion inferences/s** from ML models

Image S. Summers



- **MLOps challenge**:
  Tracking and monitoring deployed models, and adapting to changing environments

Challenge: Tracking model updates - MLOps

$\int \mathcal{L}\, dt = 47\, fb^{-1}$

$\int \mathcal{L}\, dt = 55\, fb^{-1}$

May 2024

Start of data taking with AXOL1TL

Aug 2024

AXOL1TL **model update**

Oct 2024

CICADA starts taking data

April 2025

AXOL1TL & CICADA **model updates**

- Want to be able to retrain for new detector conditions and update models often
- For analysis, it is essential to store, track and be able to re-emulate all deployed models
- Experiences **now** will be invaluable at HL-LHC where L1 models are expected to contribute **25 billion inferences / second in CMS**