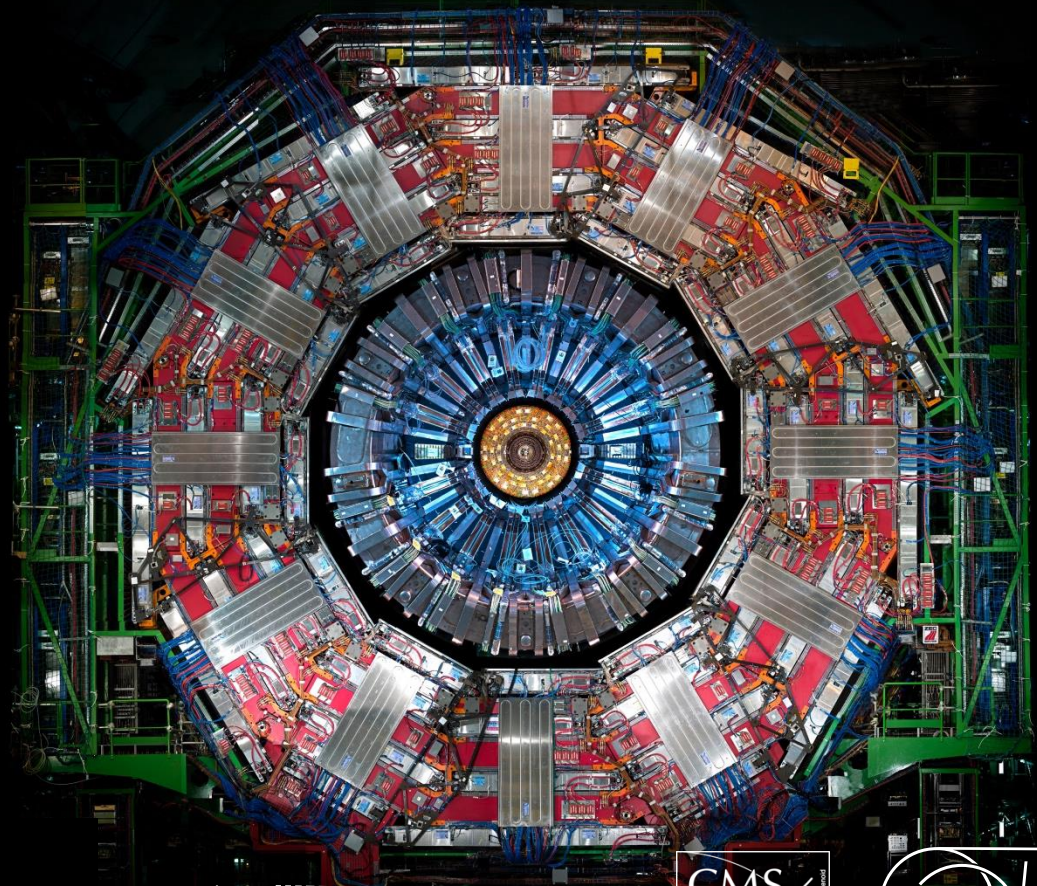


Real-Time Anomaly Detection in the CMS L1 Trigger

Jannicke Pearkes

LLP 2025

June 5th, 2025



University of Colorado **Boulder**

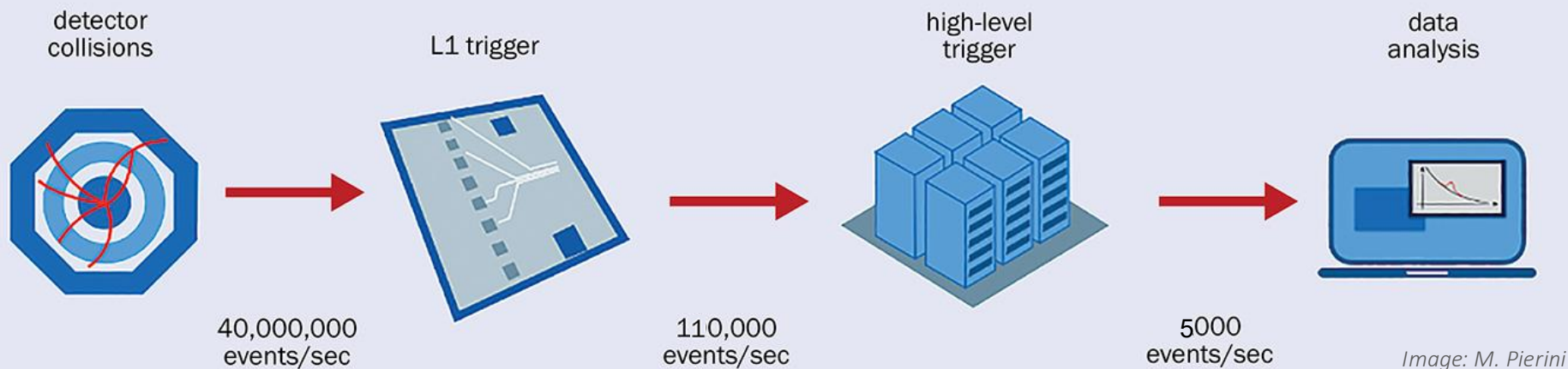


NextGen
Next Generation Triggers



Jannicke Pearkes

Why anomaly detection at L1?

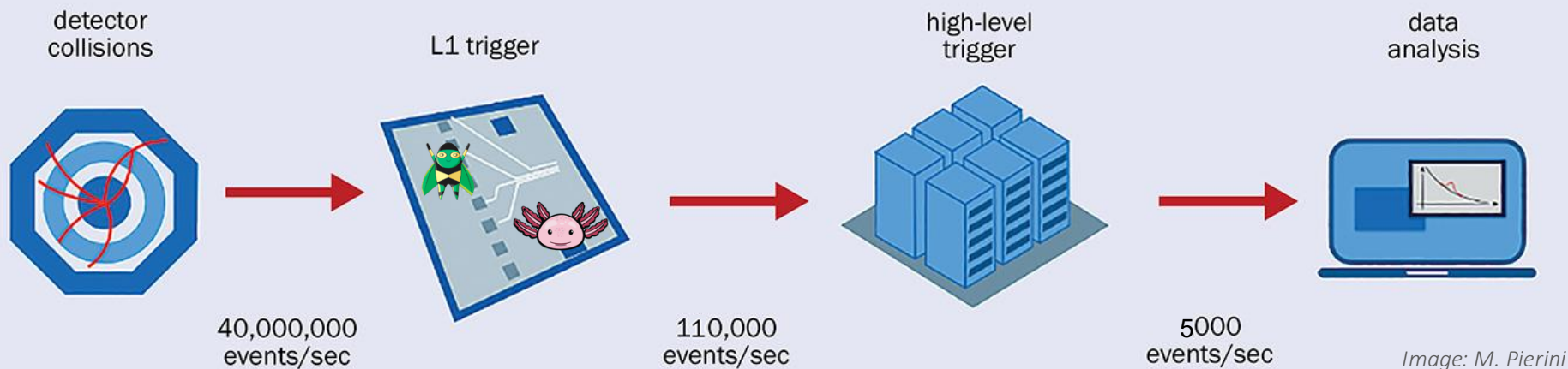


L1 trigger filters of 99.75% of collision events



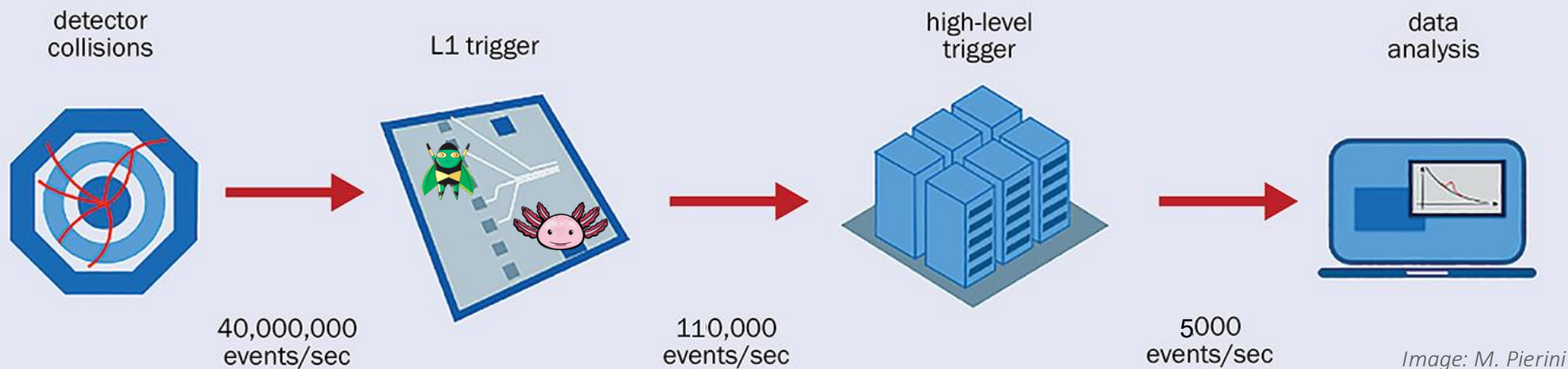
If we don't identify interesting events in trigger, we lose them forever!

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?



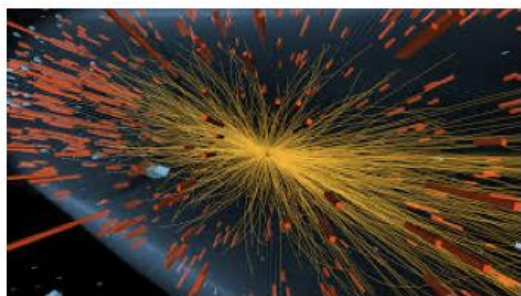
“What if we are missing new physics because we are looking for the wrong thing?”

AXOL1TL and CICADA anomaly detection algorithms use machine learning to be more model independent.

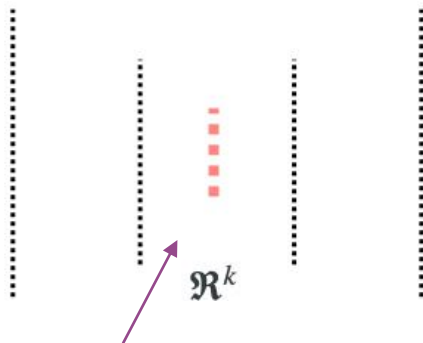


Anomaly Detection with Autoencoders

x



Train on randomly sampled *data*



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

\hat{x}

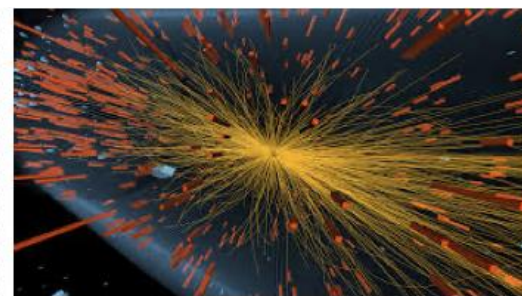
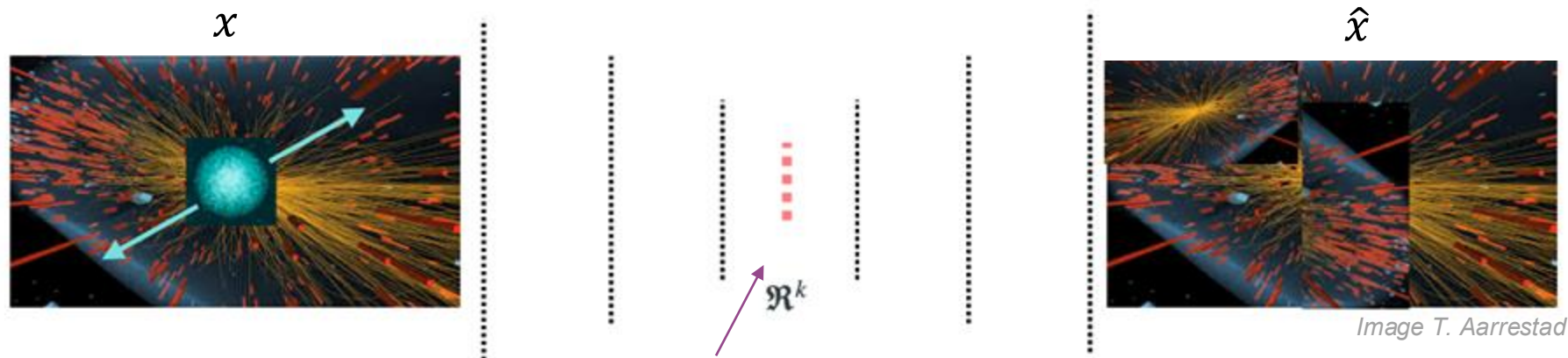


Image T. Aarrestad

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

Anomaly Detection with Autoencoders



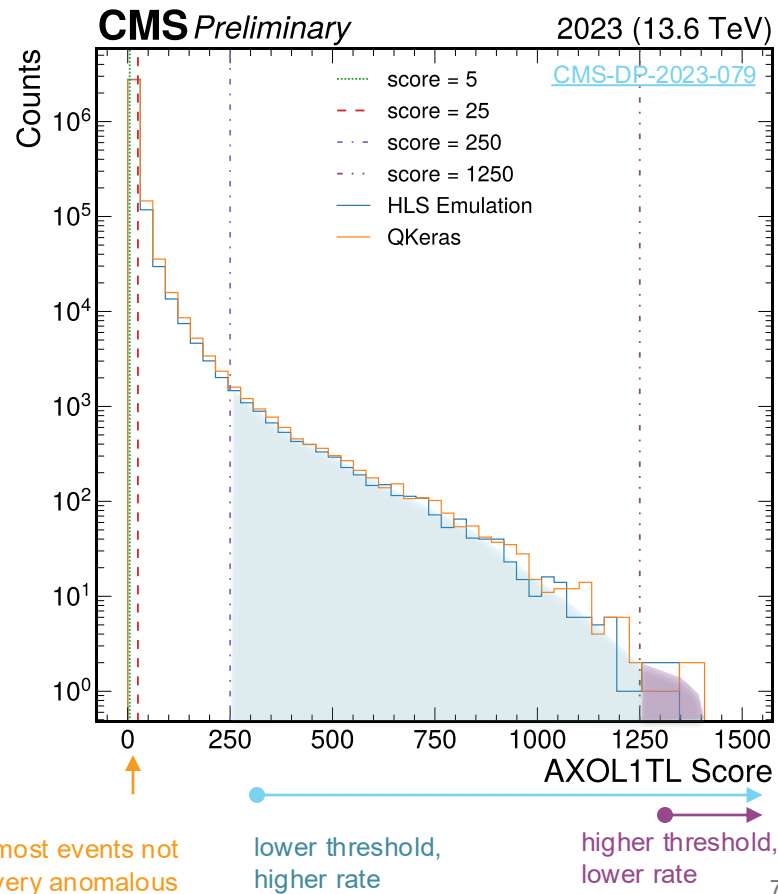
Train on randomly sampled *data*

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

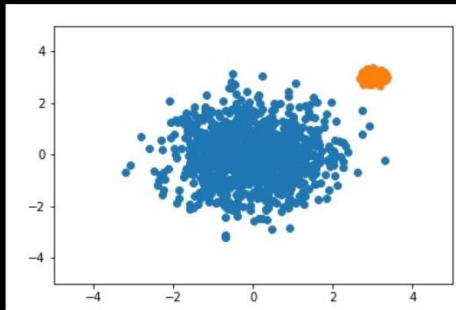


Key Assumptions for Anomaly Detection with Autoencoders

- Anomalies are *rare* in training set ✓ LLPs
- Anomalies are *out-of-distribution* ✓ LLPs (often)
- Anomalies are perceptible in available input objects ~ LLPs
(if you use the right objects)

Outlier detection

Find (non-resonant) out-of-distribution datapoints



Detecting overdensities

Find (resonant) overdensities in distributions

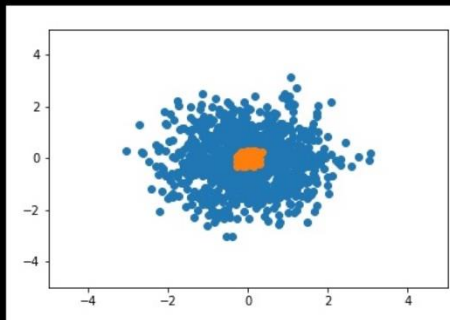
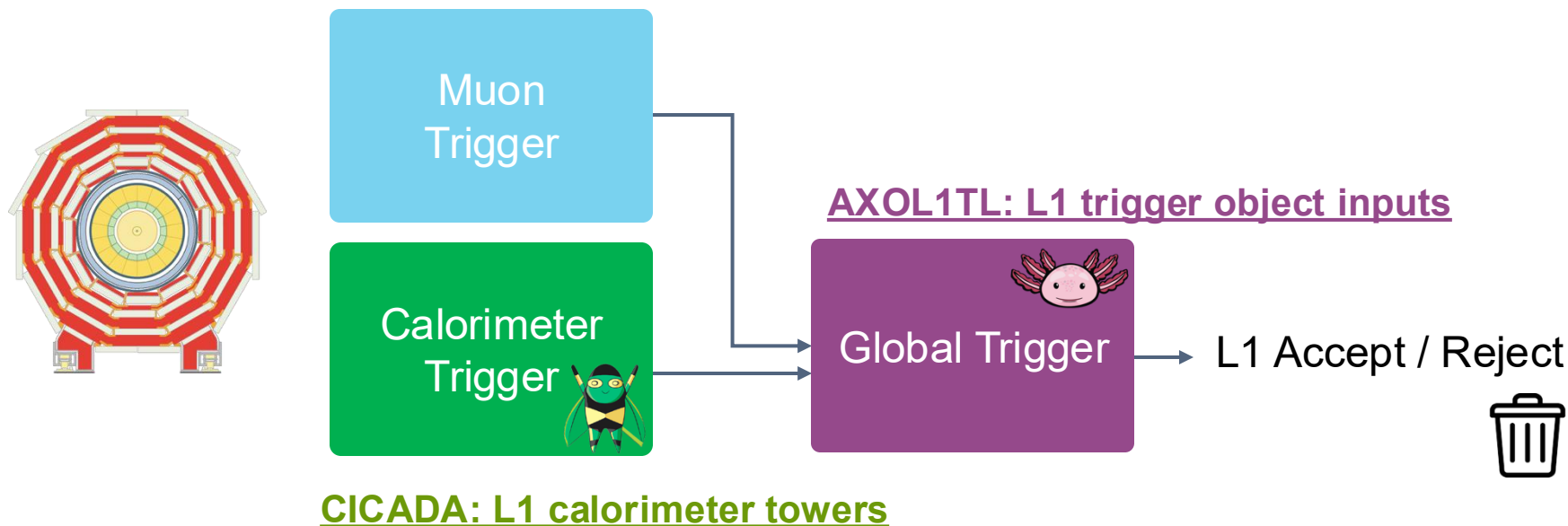


Image T. Arrestad

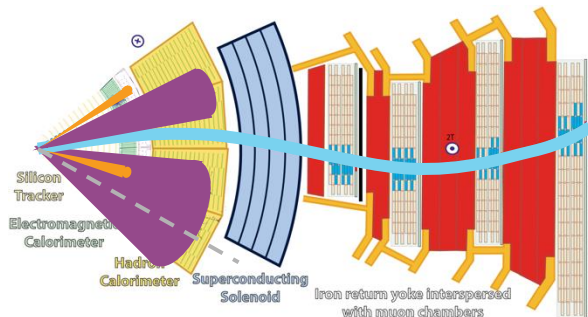
Anomaly Detection in the L1 Trigger



Both algorithms must be lightweight enough to fit within the existing L1 trigger system. Latencies of 50ns-100ns

What are the inputs?

AXOL1TL



L1 trigger objects are inputs:

MET - (p_T , ϕ)

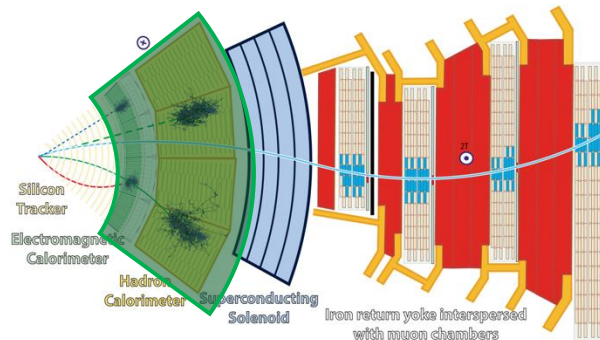
Up to 10 jets - (p_T , η , ϕ)

Up to 4 muons - (p_T , η , ϕ)

Up to 4 electrons / photons - (p_T , η , ϕ)

56 input variables total

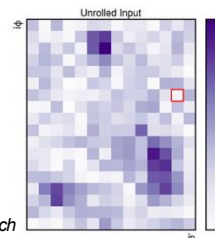
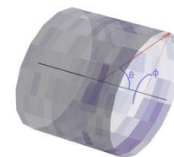
CICADA



L1 calorimeter towers are inputs:

252 E_T deposits corresponding
to 14x18 towers in $\eta \times \phi$

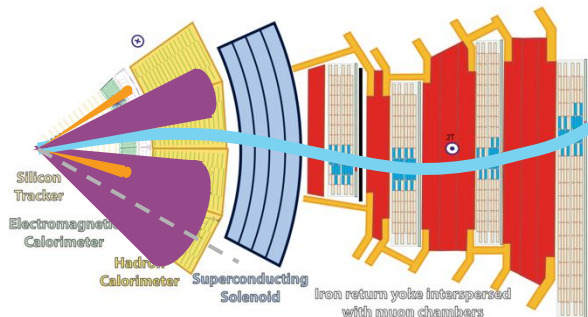
CMS Calorimeter (Schematic)



CICADA image by L. Gerlach

What sorts of LLP signatures could we be sensitive to?

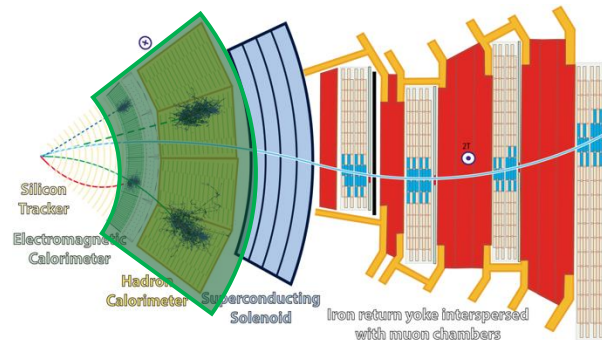
AXOL1TL



Higher-level input features:

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

CICADA

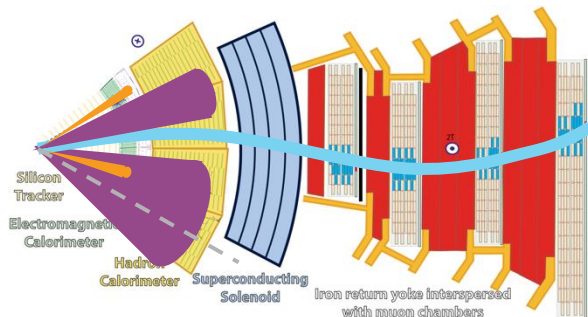


Lower-level input features:

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

What sorts of LLP signatures could we be sensitive to?

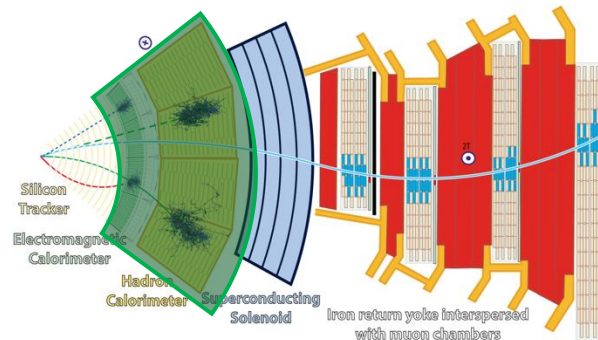
AXOL1TL



Higher-level input features:

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

CICADA



Lower-level input features:

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

Note: L1 trigger often already very efficient, AD not a replacement for targeted triggers

Status of Anomaly Detection Data-taking



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

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

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

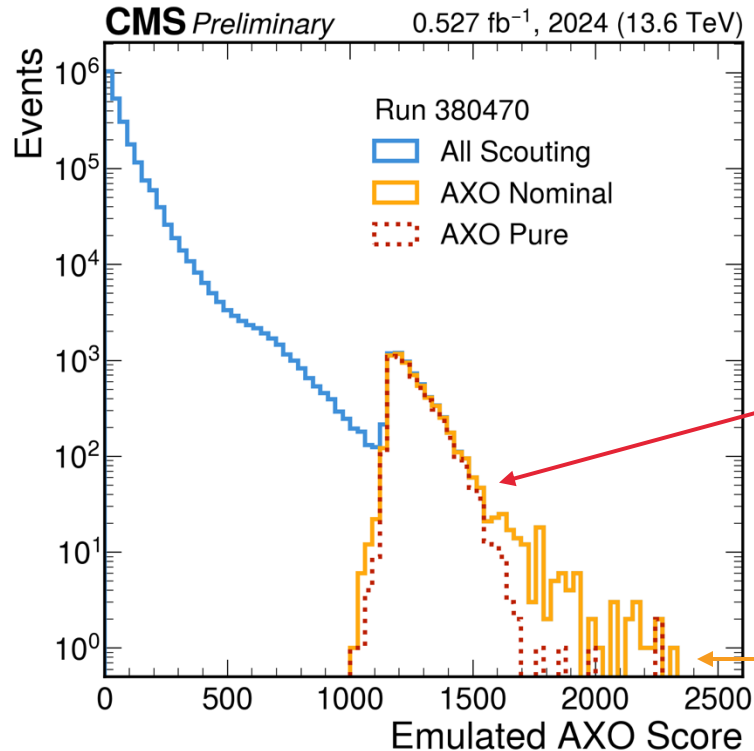


2025?

ATLAS starts taking data with
Anomaly Detection Trigger?

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?

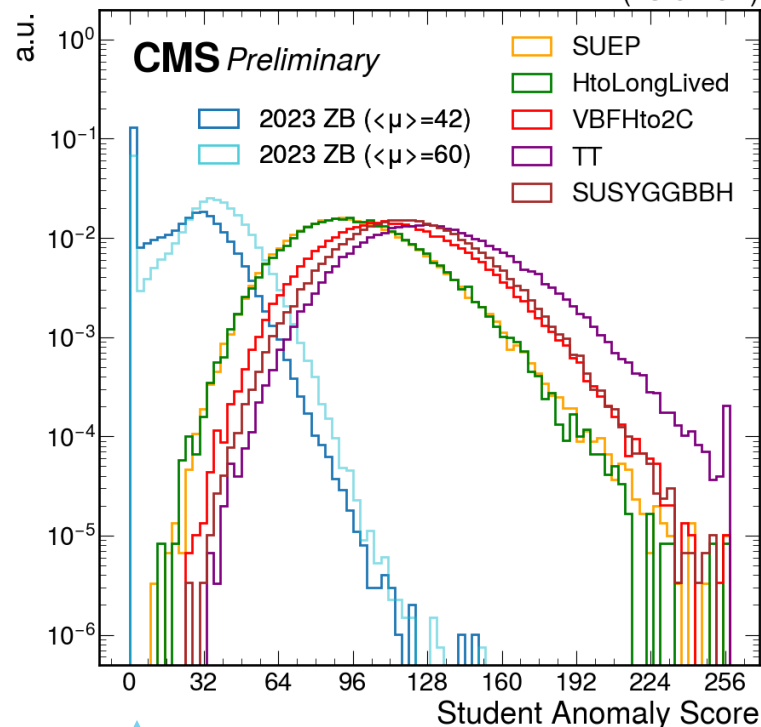
Nice separation between Zero Bias data and BSM signatures such as SUEPs and $H \rightarrow SS \rightarrow 4b$ ($c\tau = 900\text{mm}$)

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

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

CMS-DP-2024-121

(13.6 TeV)

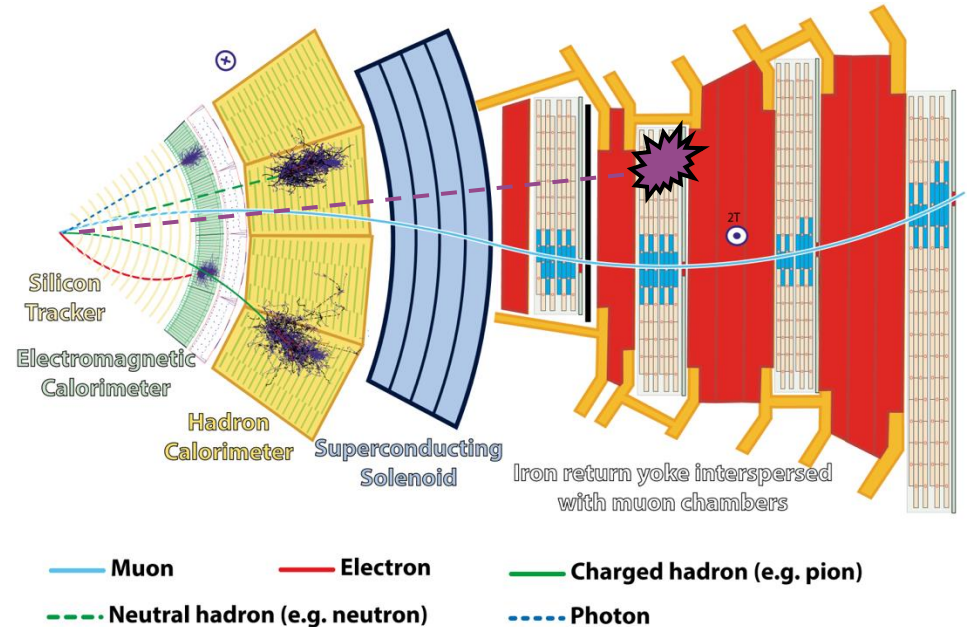


Zero Bias data events
have low anomaly score

Signal Monte Carlo has
high anomaly score

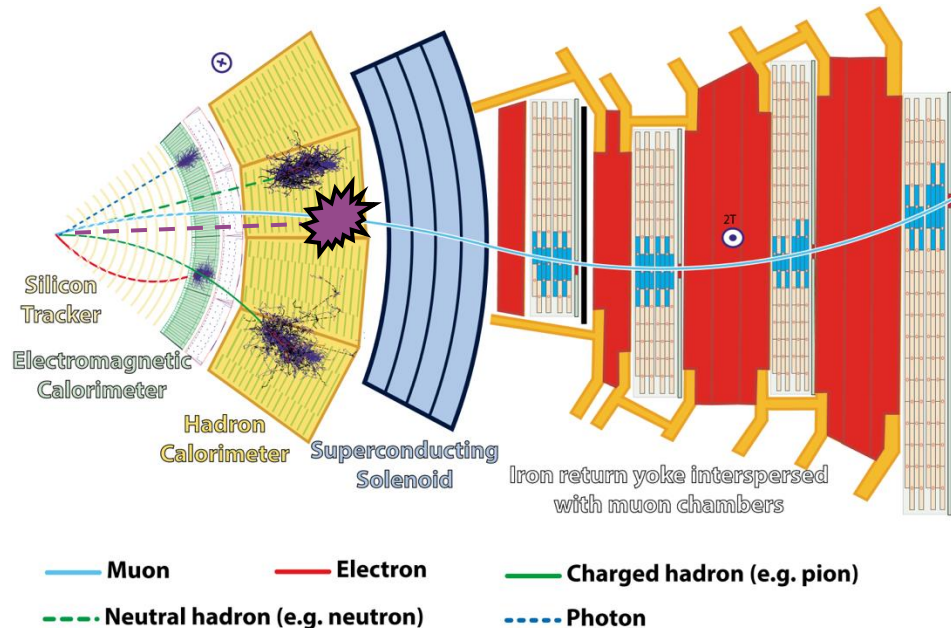
Ideas for increasing AD sensitivity to LLPs:

- (A) Incorporate muon detector shower bits [CMS-DP-2024-099](#)
- (A) Incorporate muon impact parameter bits [CMS-DP-2023-056](#)



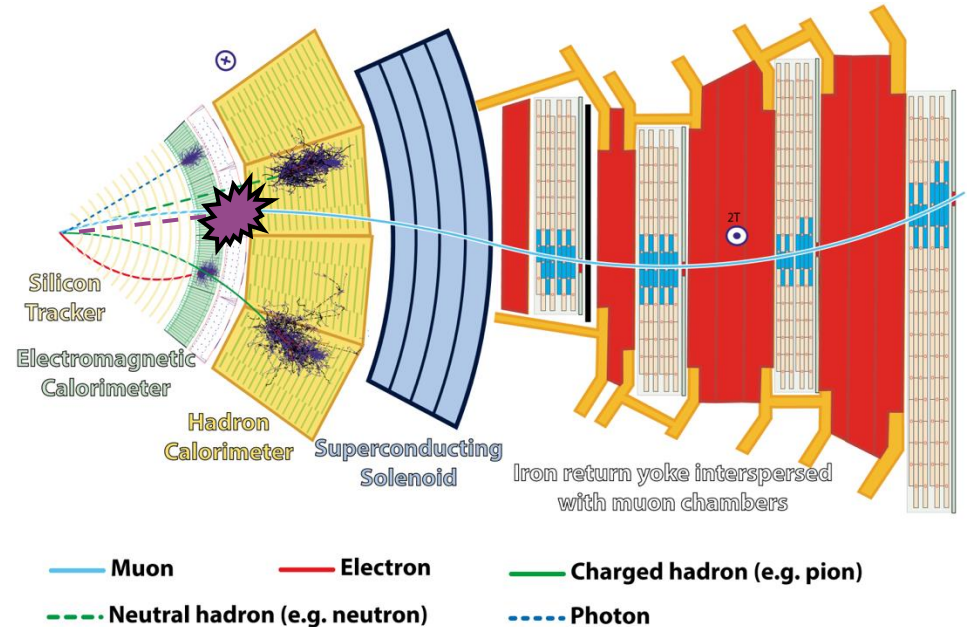
Ideas for increasing AD sensitivity to LLPs:

- (A) Incorporate muon detector shower bits [CMS-DP-2024-099](#)
- (A) Incorporate muon impact parameter bits [CMS-DP-2023-056](#)
- (A/C) Incorporate HCAL depth and timing [CMS-DP-2024-058](#)
- **HL-LHC** HGCal depth and timing



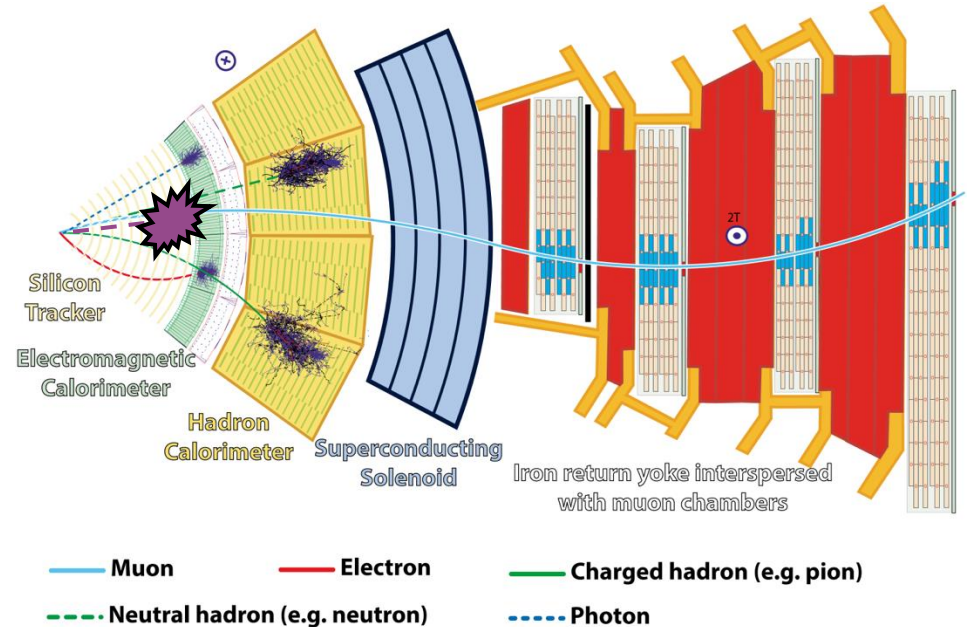
Ideas for increasing AD sensitivity to LLPs:

- (A) Incorporate muon detector shower bits [CMS-DP-2024-099](#)
- (A) Incorporate muon impact parameter bits [CMS-DP-2023-056](#)
- (A/C) Incorporate HCAL depth and timing [CMS-DP-2024-058](#)
- *HL-LHC* HGCALE depth and timing
- (C) Incorporate ratio of $E_{\text{HCAL}} / E_{\text{ECAL}}$
- *HL-LHC* L1 ECAL timing



Ideas for increasing AD sensitivity to LLPs:

- (A) Incorporate muon detector shower bits [CMS-DP-2024-099](#)
- (A) Incorporate muon impact parameter bits [CMS-DP-2023-056](#)
- (A/C) Incorporate HCAL depth and timing [CMS-DP-2024-058](#)
- **HL-LHC** HGCALE depth and timing
- (C) Incorporate ratio of $E_{\text{HCAL}} / E_{\text{ECAL}}$
- **HL-LHC** L1 ECAL timing
- **HL-LHC** Level 1 tracking (hopefully including displaced tracking) [CMS-TDR-021](#)

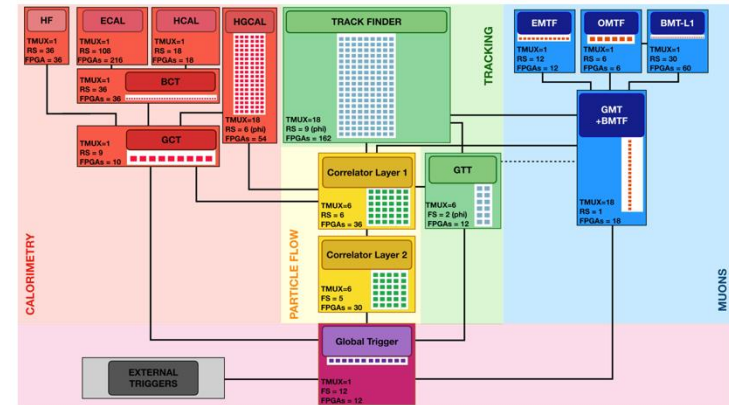


Ideas for increasing AD sensitivity to LLPs:

- (A) Incorporate muon detector shower bits [CMS-DP-2024-099](#)
- (A) Incorporate muon impact parameter bits [CMS-DP-2023-056](#)
- (A/C) Incorporate HCAL depth and timing [CMS-DP-2024-058](#)
- **HL-LHC** HGCal depth and timing
- (C) Incorporate ratio of $E_{\text{HCAL}} / E_{\text{ECAL}}$
- **HL-LHC** L1 ECAL timing
- **HL-LHC** Level 1 tracking (hopefully including displaced tracking) [CMS-TDR-021](#)

Technical constraints:

- Availability of objects
- Stability & modeling of objects
- Latency & size of anomaly detection network

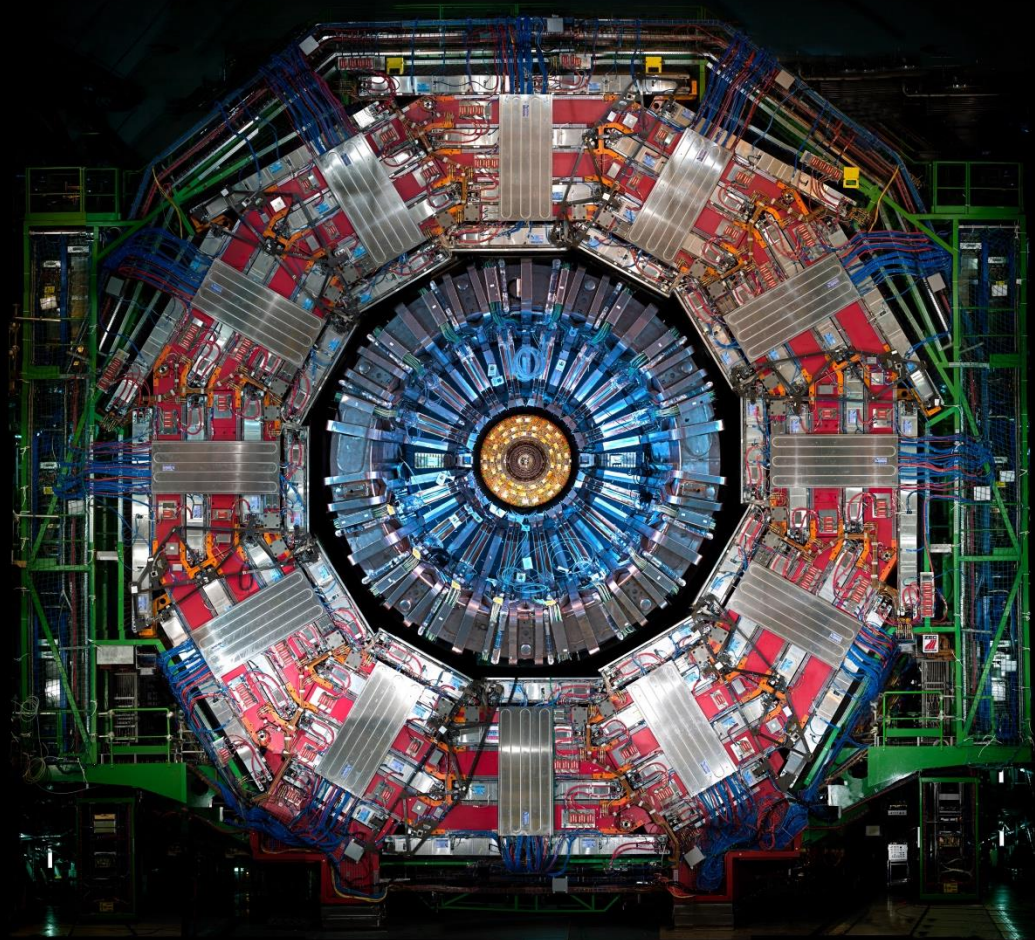


Summary

- 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
- They are the first ML-based anomaly detection triggers deployed at the LHC
- Current algorithms don't use LLP specific inputs, but extensions in the future are possible
- This is just the very beginning!



Thank you!



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)

[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

HLT Strategy 2024

Level-1 Trigger (L1T)

AXOL1TL (Pure rate)

L1_AXO_VLoose (1000 Hz)
L1_AXO_Loose (400 Hz)
L1_AXO_Nominal (200 Hz)
L1_AXO_Tight (100 Hz)
L1_AXO_VTight (10 Hz)

CICADA (Pure rate)

L1_CICADA_VLoose (600 Hz)
L1_CICADA_Loose (300 Hz)
L1_CICADA_Medium (150 Hz)
L1_CICADA_Tight (50 Hz)
L1_CICADA_VTight (20 Hz)

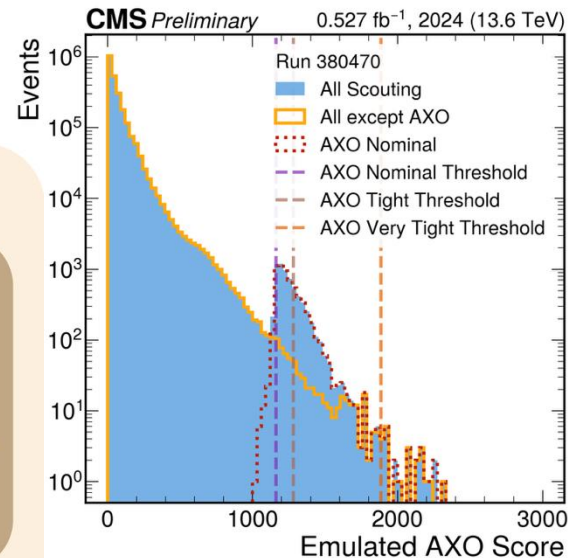
High Level Trigger (HLT)

HLT Scouting (Total rate)

DST_PFScouting_AXONominal (1.5 kHz)
DST_PFScouting_AXOTight (1.1 kHz)
DST_PFScouting_AXOVTight (450 Hz)
DST_PFScouting_CICADAMedium (620 Hz)
DST_PFScouting_CICADATight (370 Hz)
DST_PFScouting_CICADAVTight (250 Hz)

Full reconstruction (Total rate)

DST_PFScouting_AXOVTight (450 Hz)



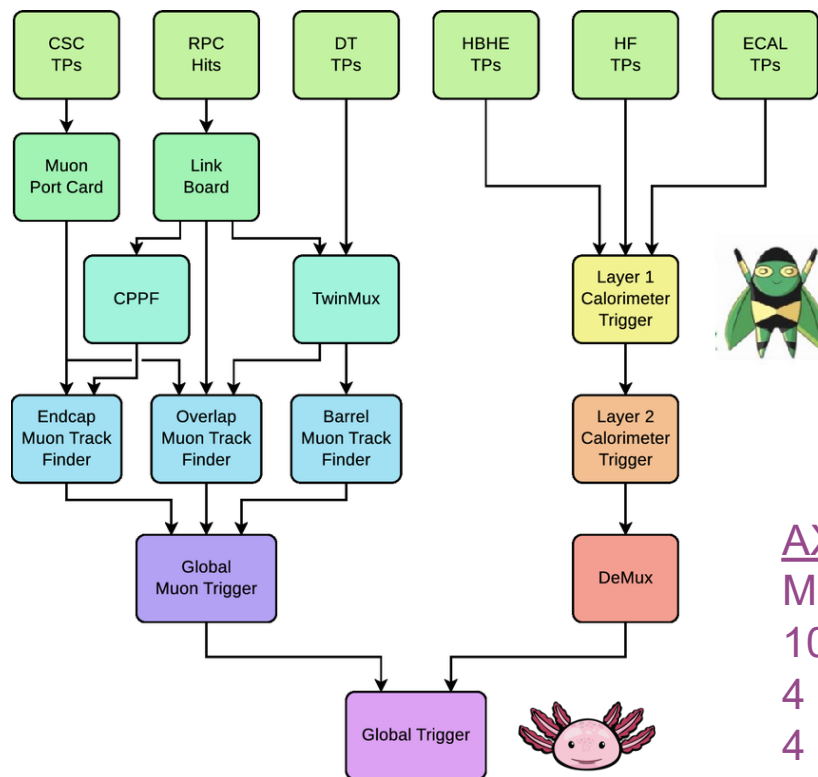
Rates shown are NNv3 target rates

Passthrough trigger:

Anything accepted at L1 is also accepted by the HLT

Image: S. Giorgetti

AXOL1TL & CICADA in the L1 Trigger



CICADA L1 calorimeter towers are inputs:
252 E_T deposits corresponding to 14×18 towers in $\eta \times \phi$

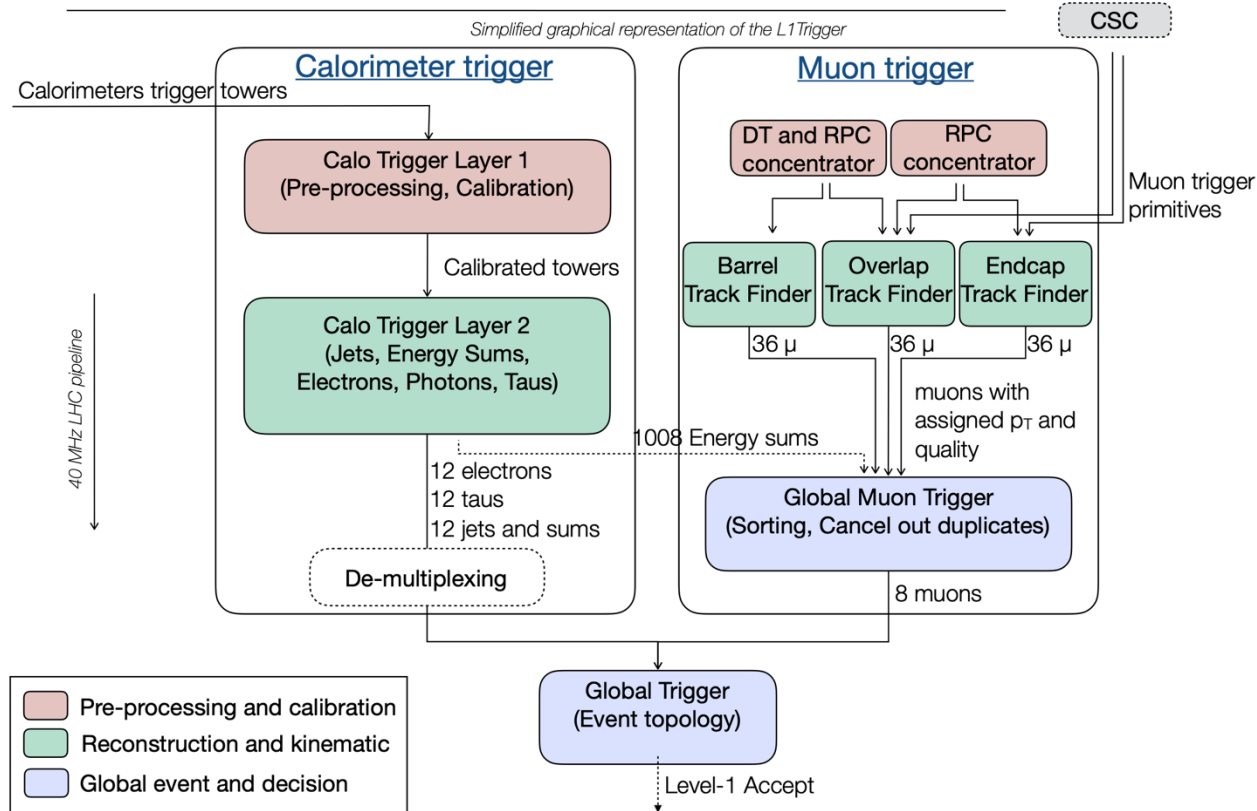


AXOL1TL L1 trigger objects are inputs:
MET - (p_T, ϕ)
10 jets - (p_T, η, ϕ)
4 muons - (p_T, η, ϕ)
4 electrons / photons - (p_T, η, ϕ)



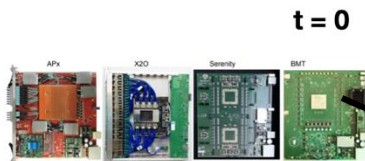
The CMS Level-1 trigger design

[Image: P. Bortignon](#)



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



1 small box = 1 FPGA board
with AMD VU13P FPGA

t = 0

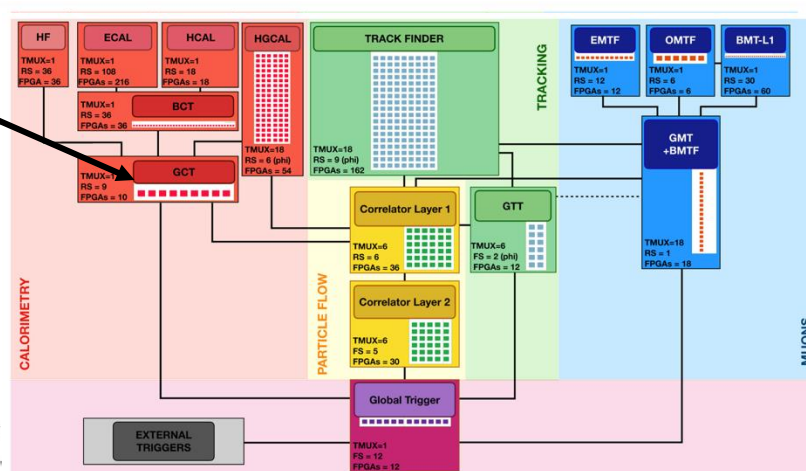


Image S. Summers

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

Challenge: Ultrafast Inference

Image B. Ramhorst

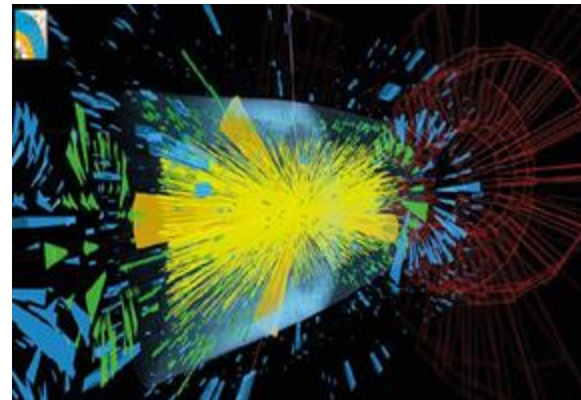
~1-3 seconds

~50ms

~100 ns



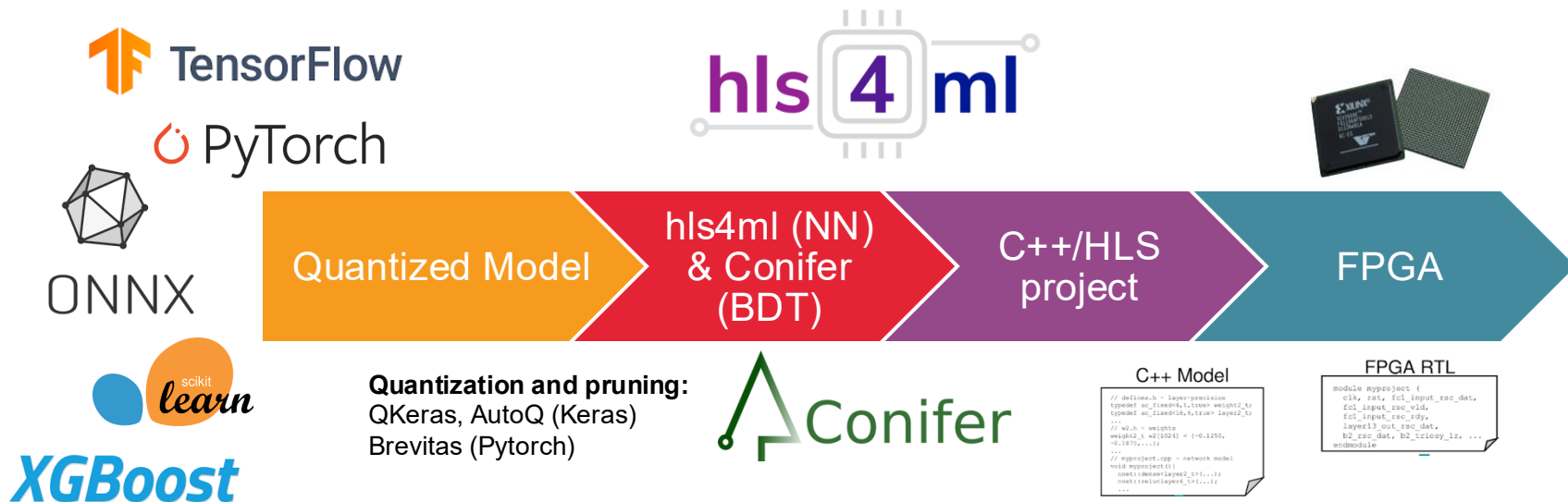
ChatGPT



AXOL1TL and CICADA have to be ultrafast



Deployed Neural Networks on FPGAs

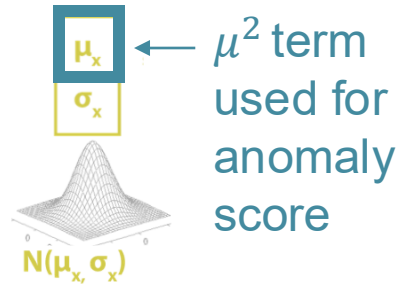
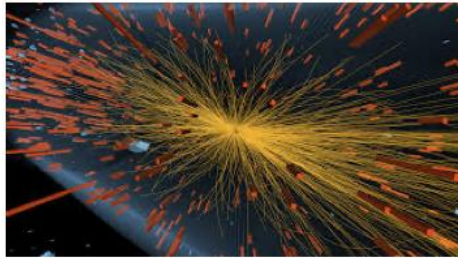


AXOL1TL and CICADA use [hls4ml](#). [hls4ml \(2018\)](#) & [conifer \(2021\)](#) are open source projects primarily developed by LHC community to convert neural networks and boosted decision trees into HLS.

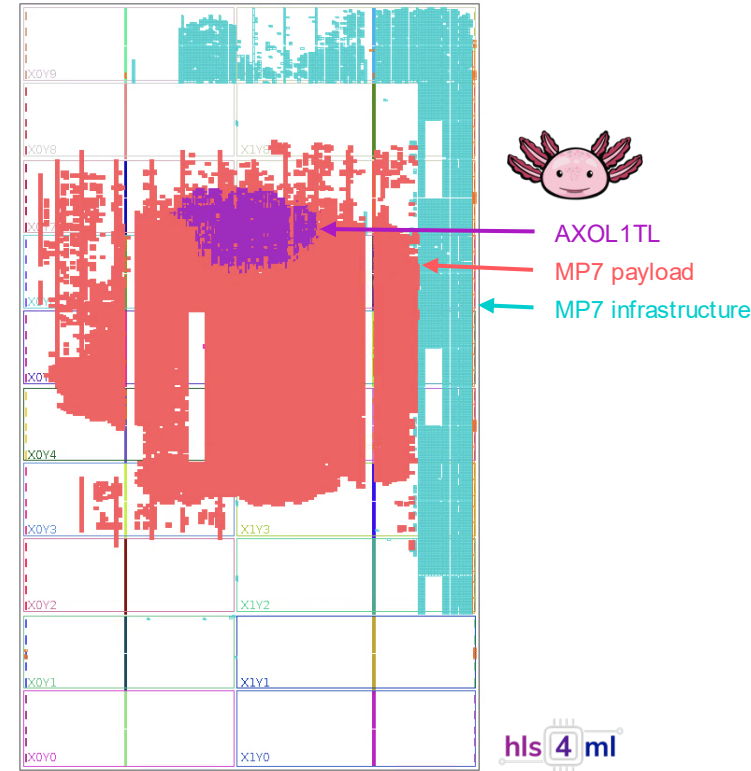
AXOL1TL Implementation

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



Implemented on Xilinx Virtex-7 FPGA
50 ns latency and resource requirements met



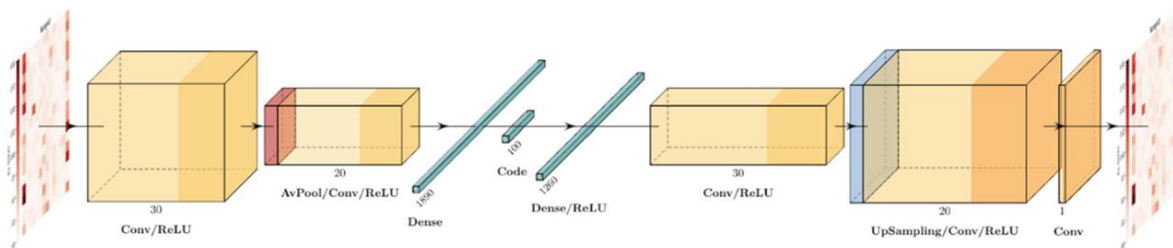
CICADA Implementation



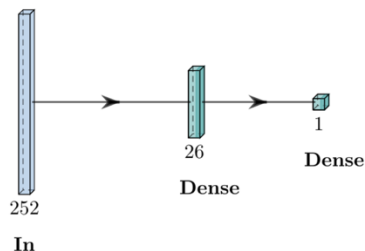
V. Sharma PPC 2024
CMS-DP-2024-121

Knowledge distillation: student learns from teacher model

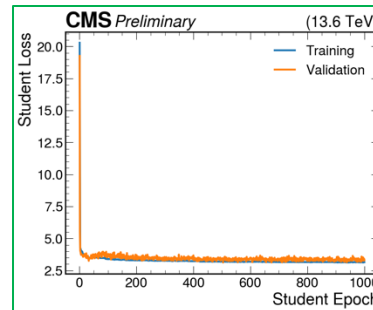
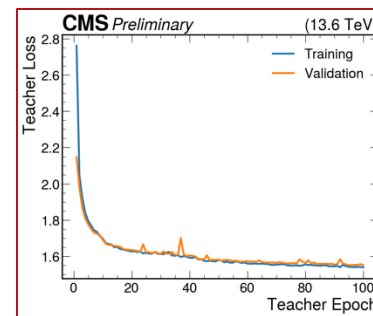
The **teacher** model



The **student** model



Inference latency ~ 100 ns
on Virtex-7 FPGA

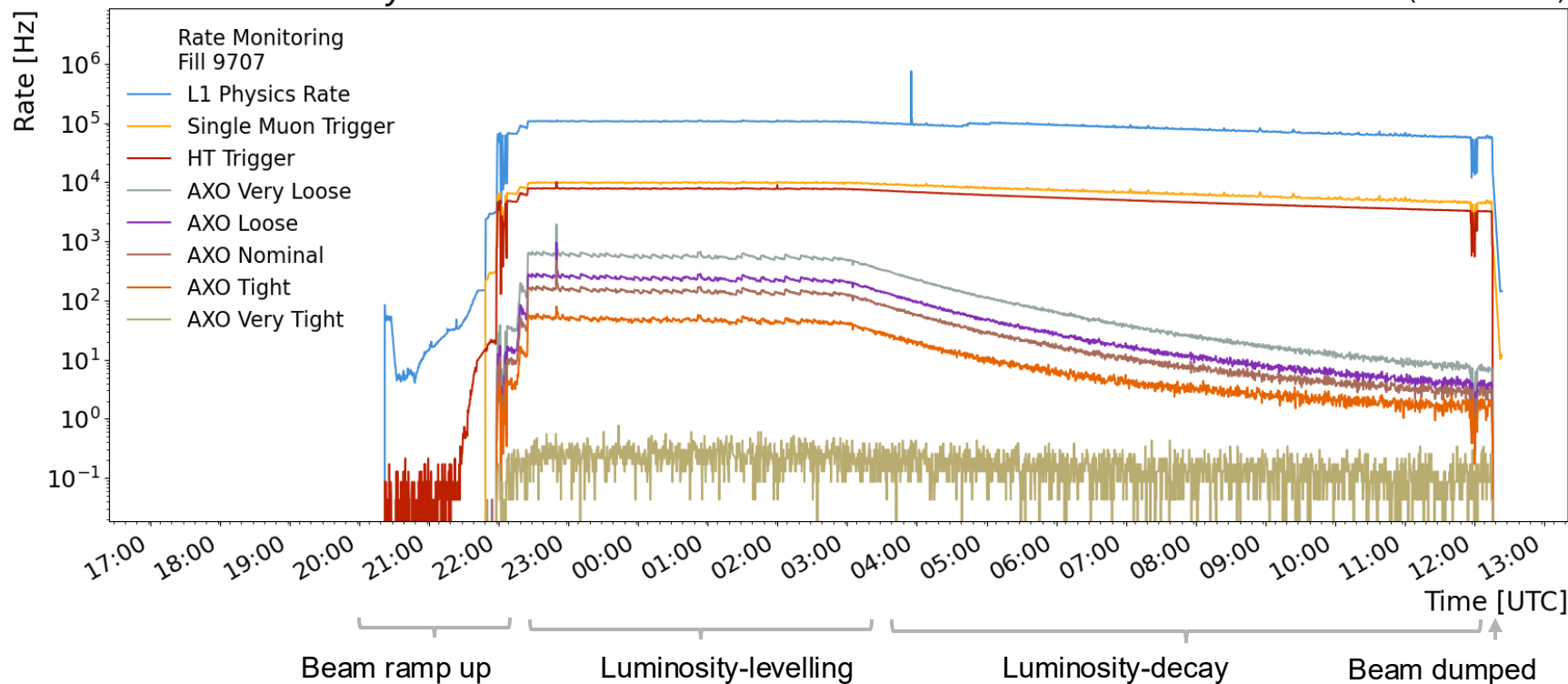


AXOL1TL Rate Stability

[CMS-DP-2024-059](#)

CMS Preliminary

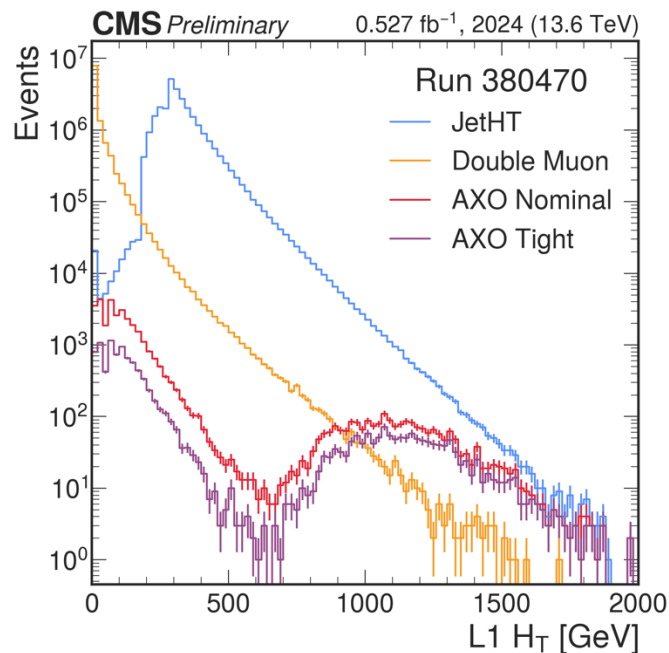
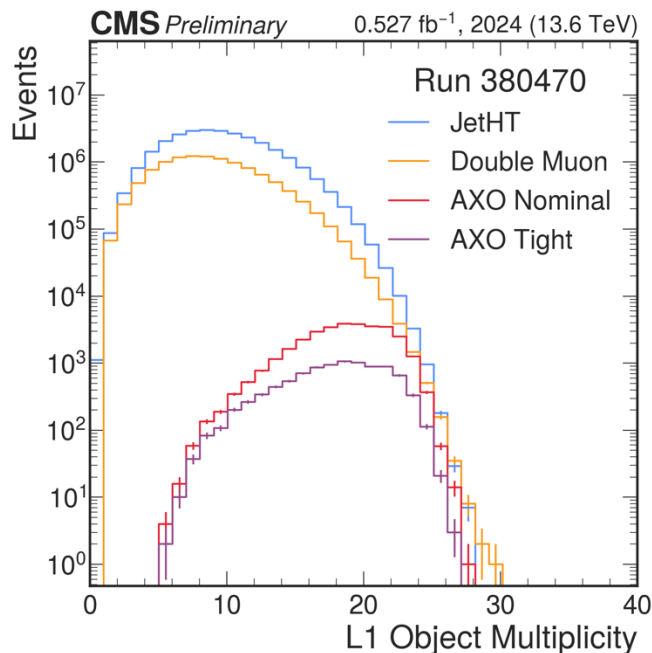
0.767 fb⁻¹, 2024 (13.6 TeV)



What does AXOL1TL trigger on?

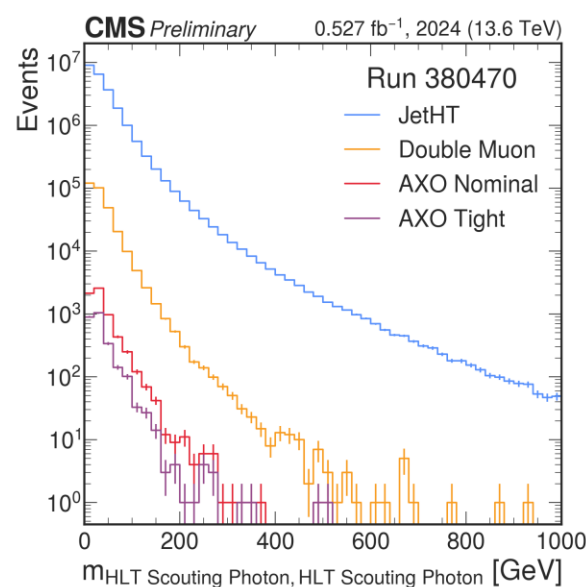
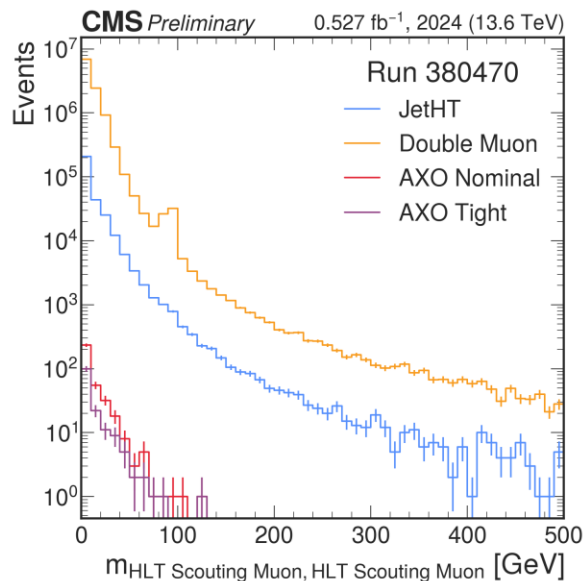
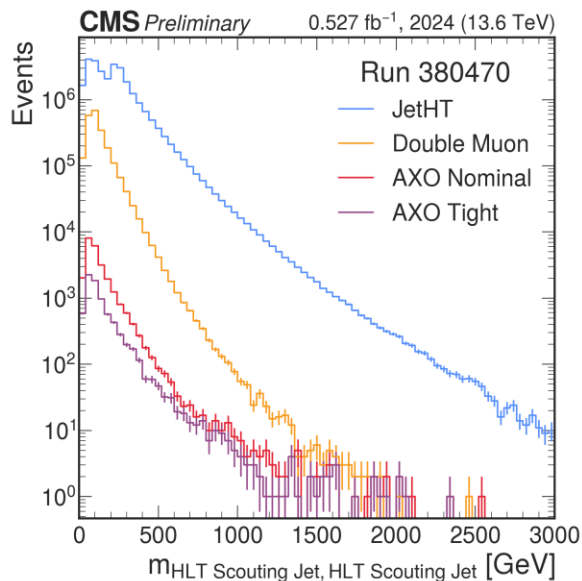
[CMS-DP-2024-059](#)

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>

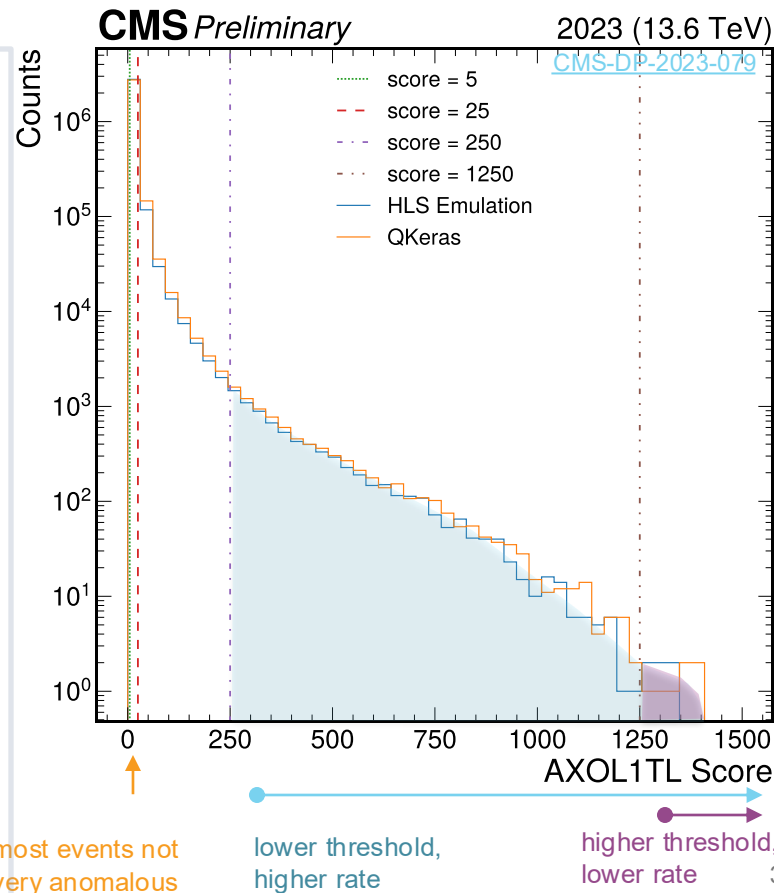
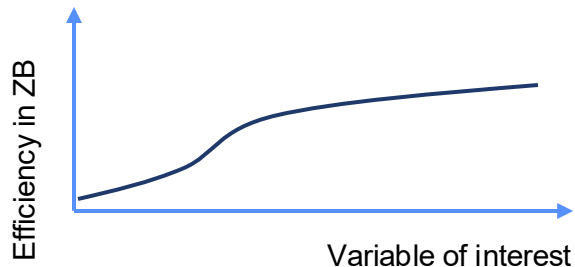


How to think about trigger efficiencies in Zero Bias Data:

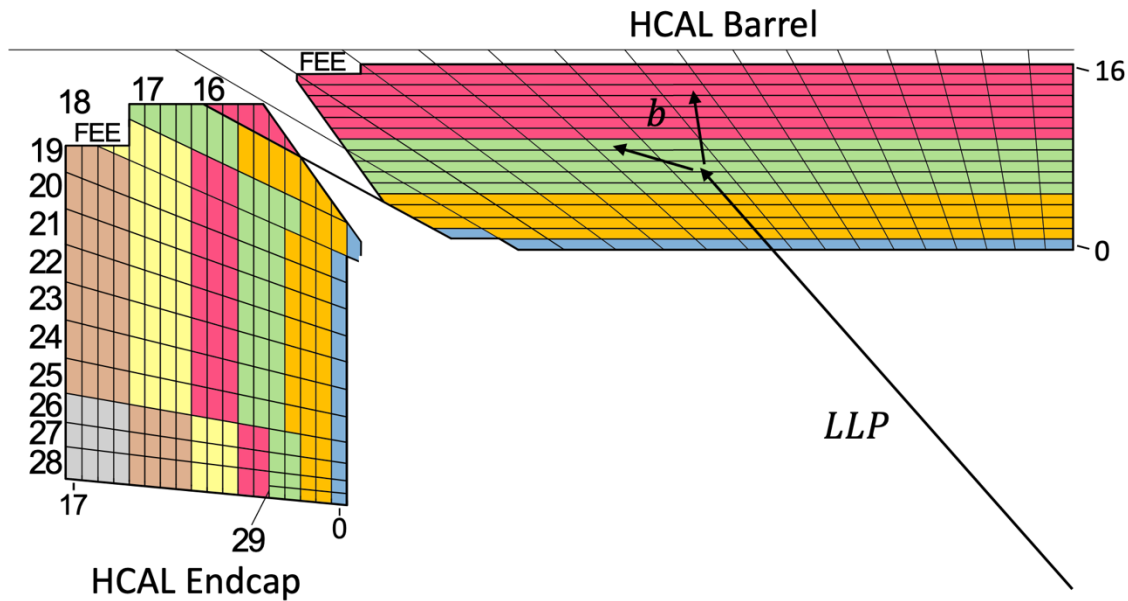
Trigger efficiencies are calculated using the **orthogonal data method** in Zero Bias data (randomly triggered events)

$$\text{Efficiency (\%)} = \frac{\text{ZeroBias \&\& L1_AXO Nominal}}{\text{ZeroBias}}$$

The trigger efficiency in data is a measure of what types of events the AD trigger considers to be "in the tails" of the Zero Bias distribution.

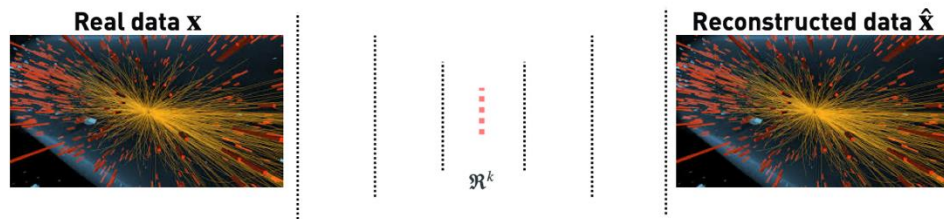


HCal Depth Segmentation [CMS-DP-2024-058](#)



Challenge: Pile-up Dependence

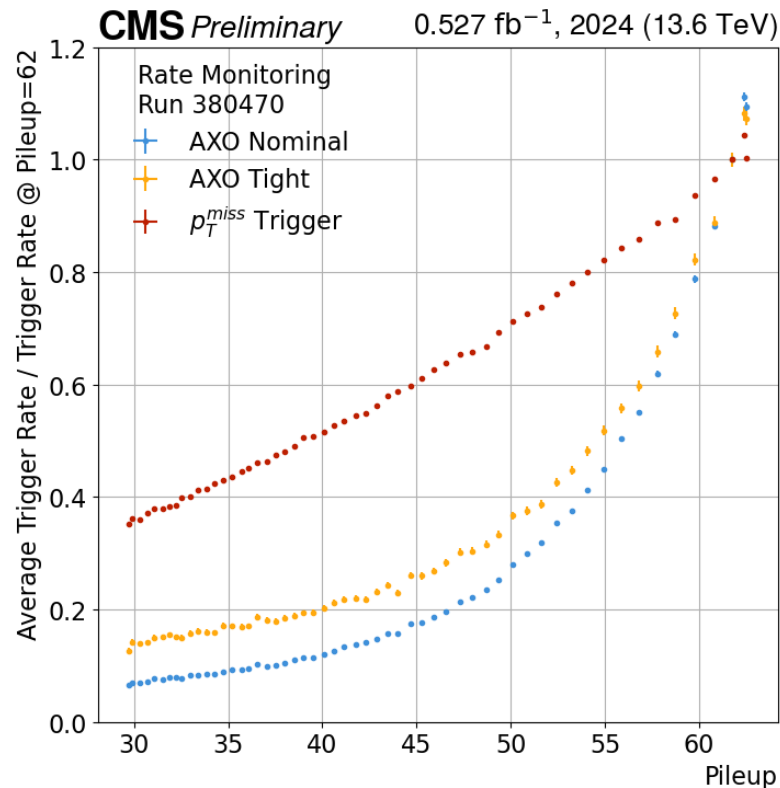
[CMS-DP-2024-059](#)



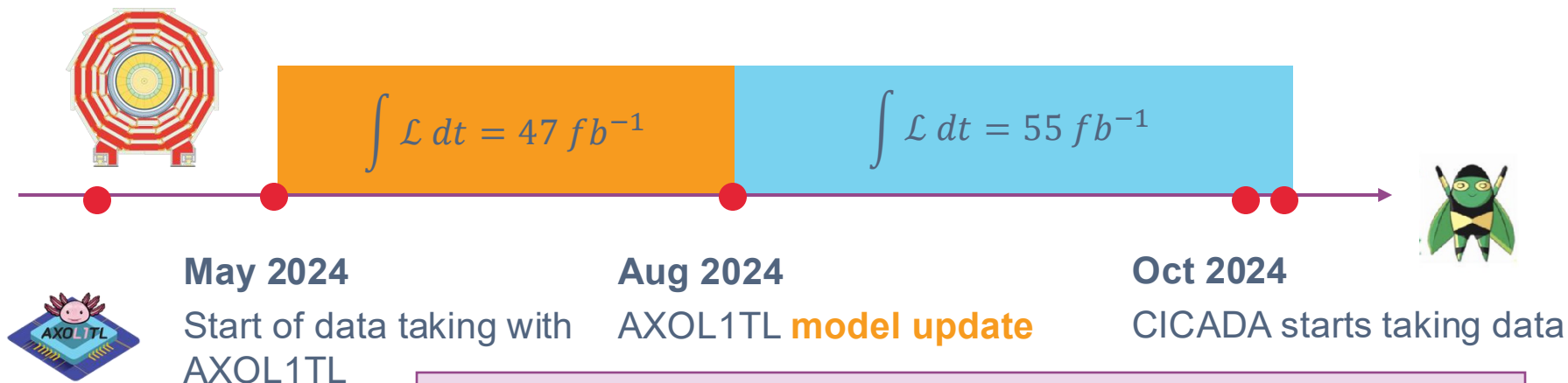
Events with high pileup contain more “information” (more jets, more calorimeter cells) are inherently harder to encode. This contributes to higher rates at high pileup.

Care must be taken during model development and deployment:

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



Challenge: Tracking model updates - MLOps



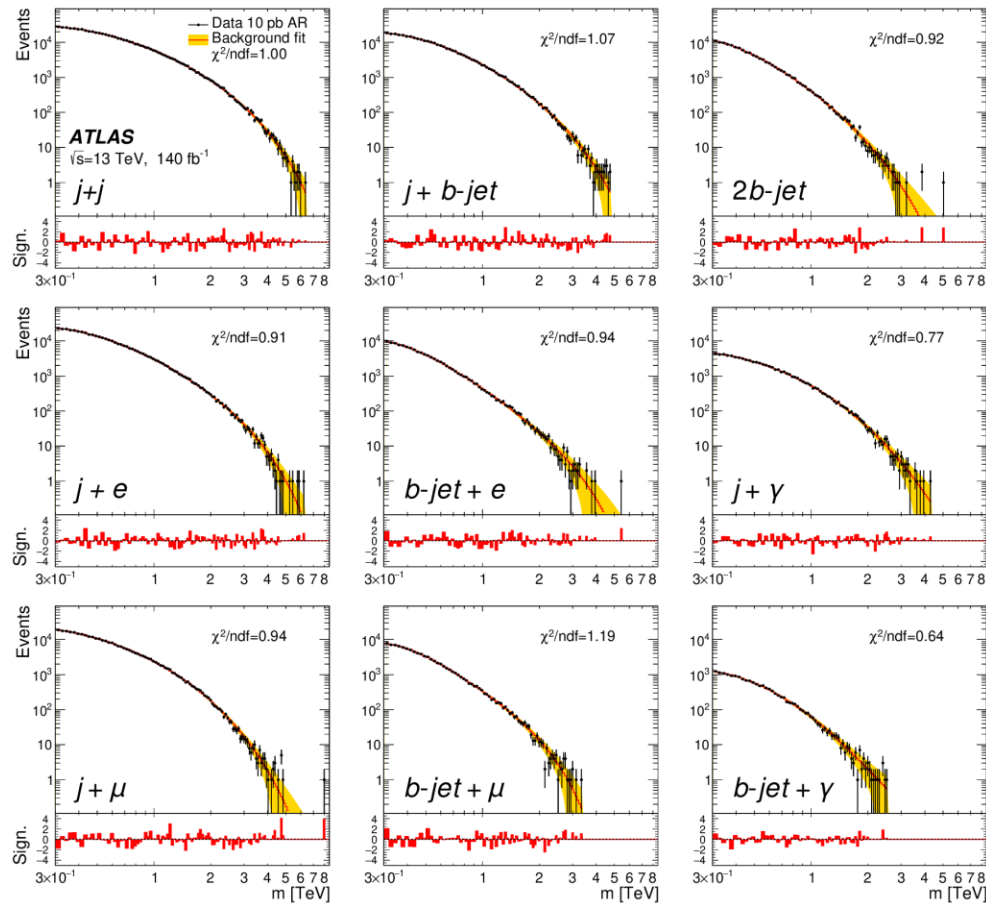
April 2025
AXOL1TL & CICADA
model updates

- 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**

mlflow™

Bump hunt:

- **Start simple:** bump hunt in di-object invariant mass spectra
- Analysis with an anomaly detection triggers has never done before, **lots of technology to develop**
- AXOL1TL is an event-level trigger, can hunt for resonances on **combinations of different objects**, e.g. $\mu + e$, jet + γ , etc.



ATLAS EXOT-2022-07