UC San Diego

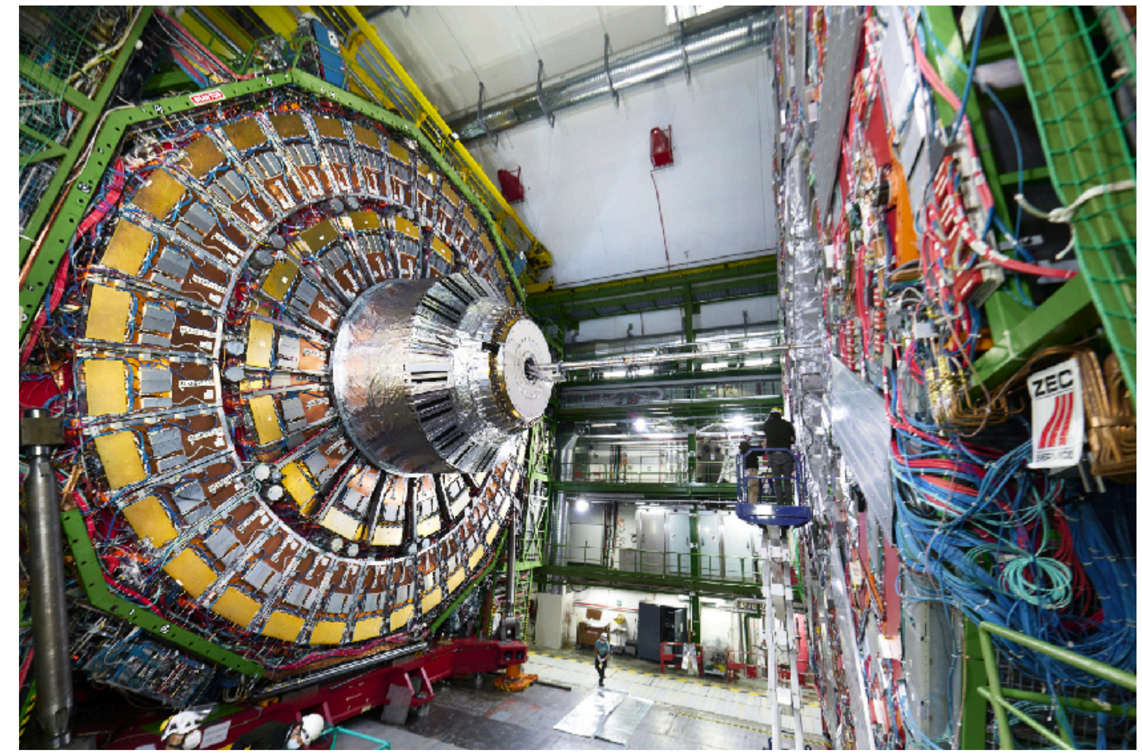
Anomaly Detection in the CMS L1 Trigger

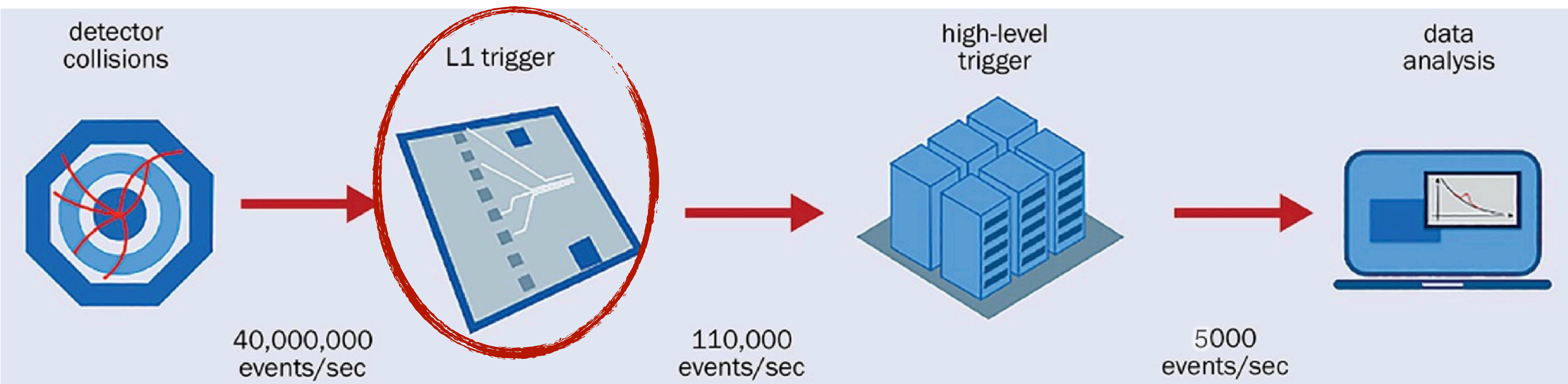
Melissa Quinnan* on behalf of the CMS Collaboration

*UC San Diego

CMS Level-1 Trigger

- L1 trigger rejects 99.75% of LHC events
  - Constrained by low latency of 50 ns and low resource utilization on FPGAs
- What if we are missing new physics because we did not design the right trigger?

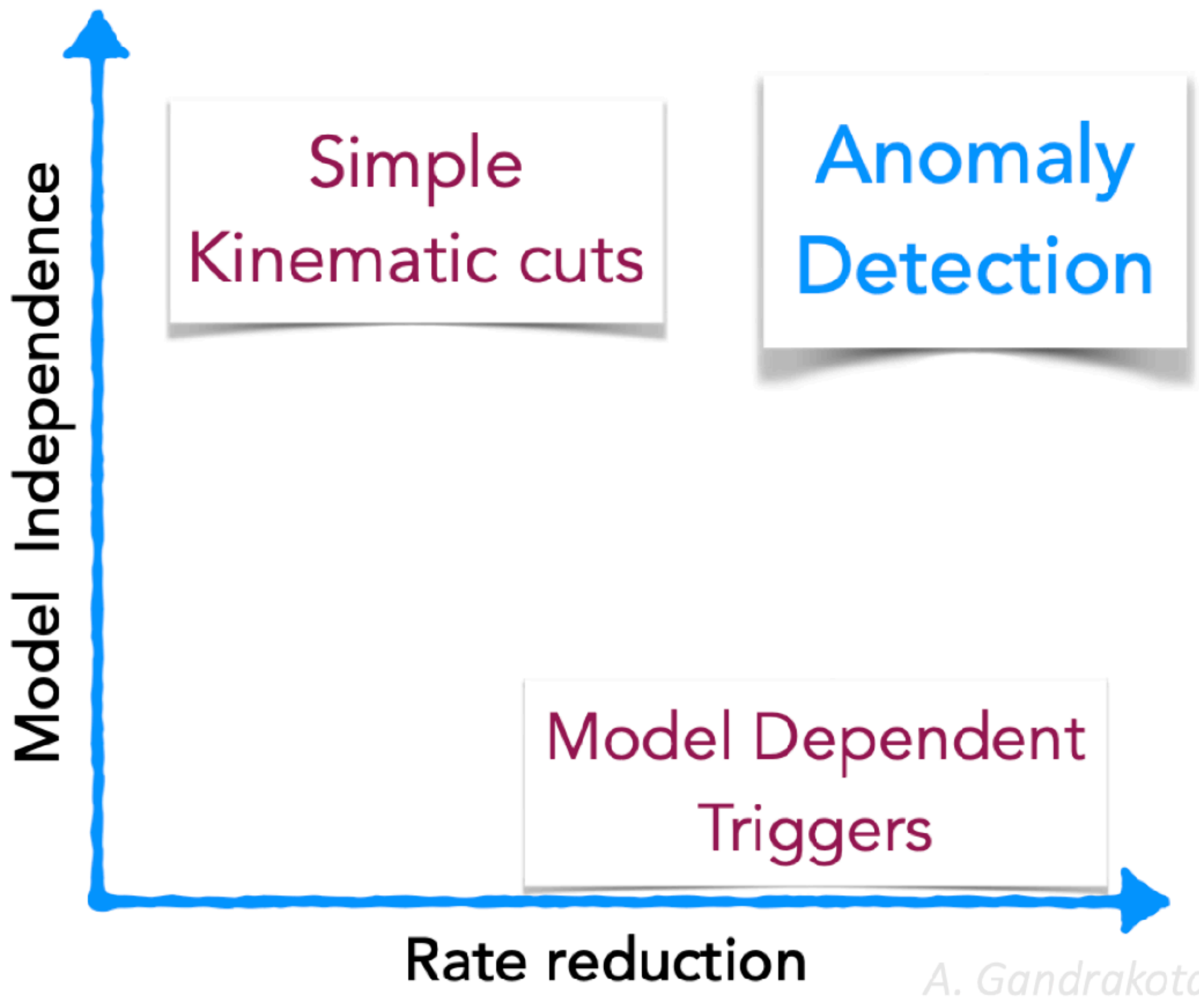




Anomaly Detection in the Trigger

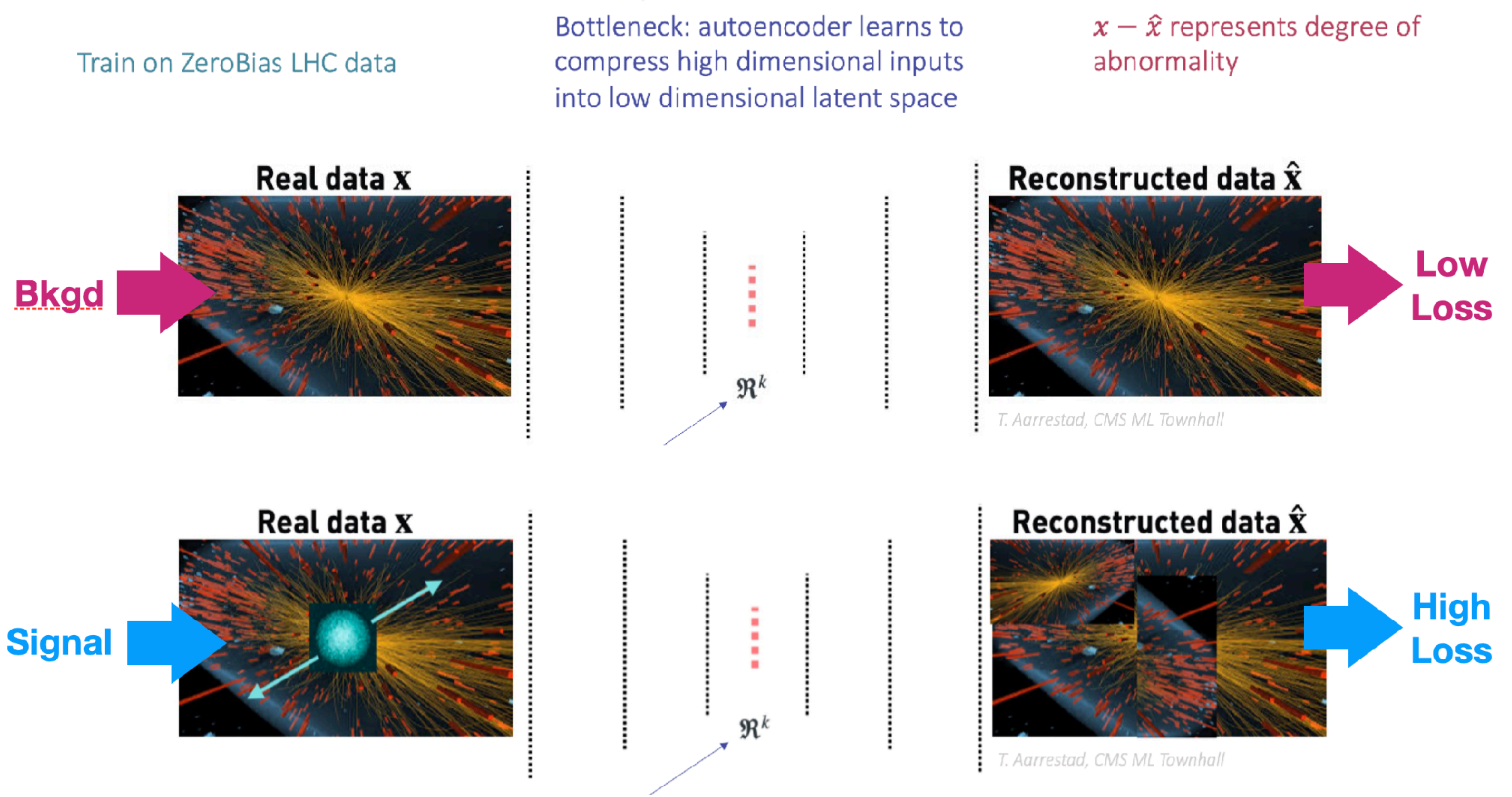
→What if we are rejecting new physics with our existing trigger algorithms?

- Signal agnostic - Applicable to signatures for which we have not had the foresight to design specific triggers
- High sensitivity – Can improve signal efficiency for signatures that are limited by L1 trigger bandwidth



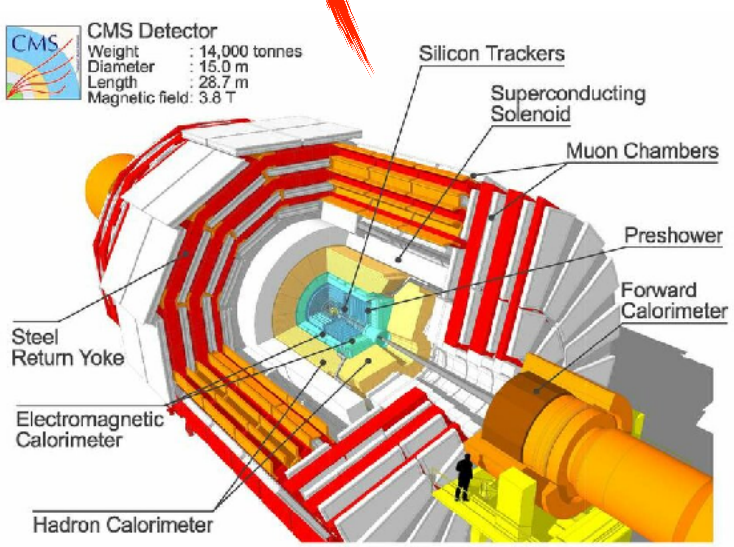
Autoencoders for Anomaly Detection

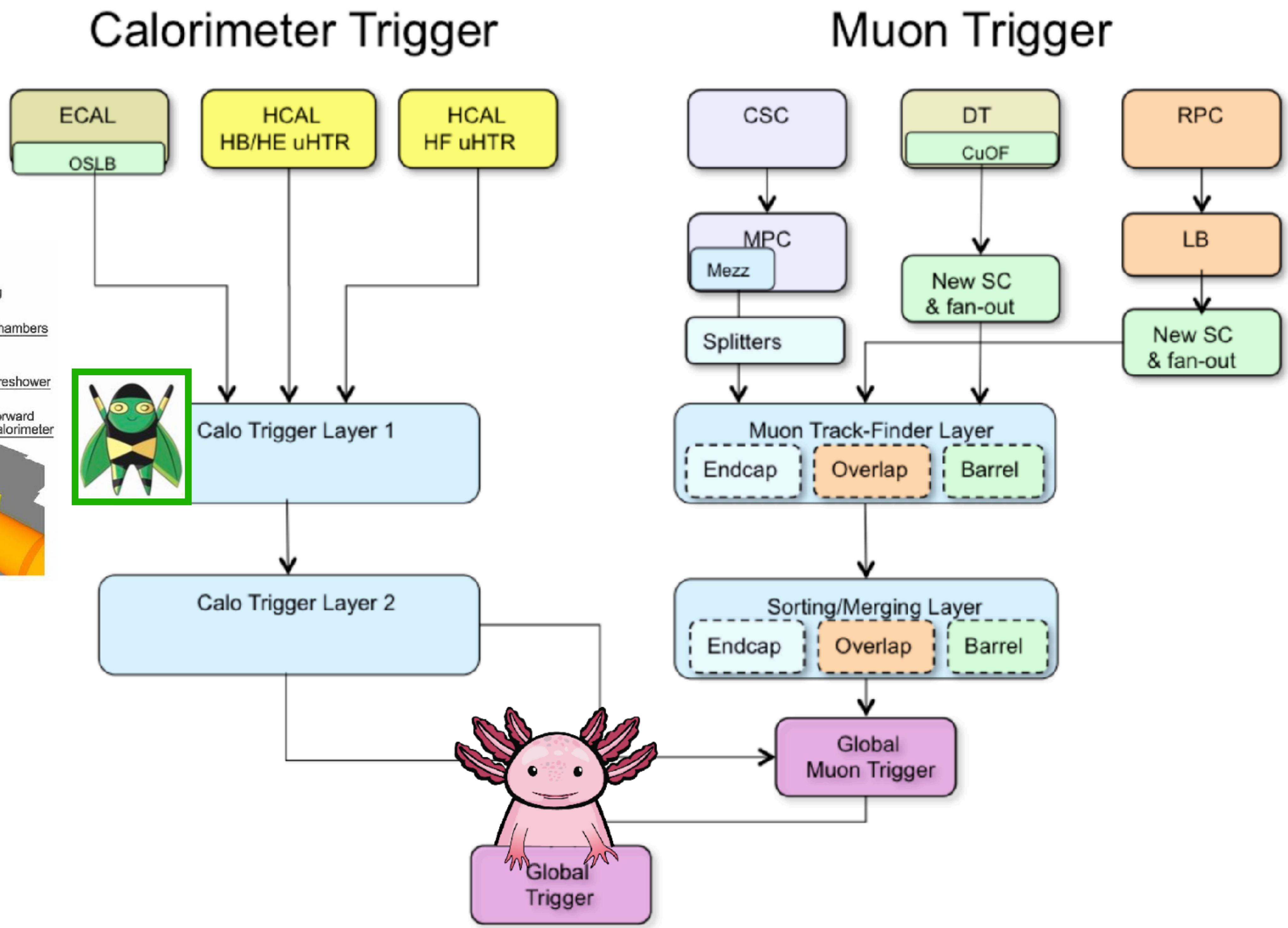
- Unsupervised ML learns efficient encodings of backgrounds



Current L1 Trigger Structure

- AXOL1TL+CICADA: autoencoder-based anomaly detection in L1 Trigger

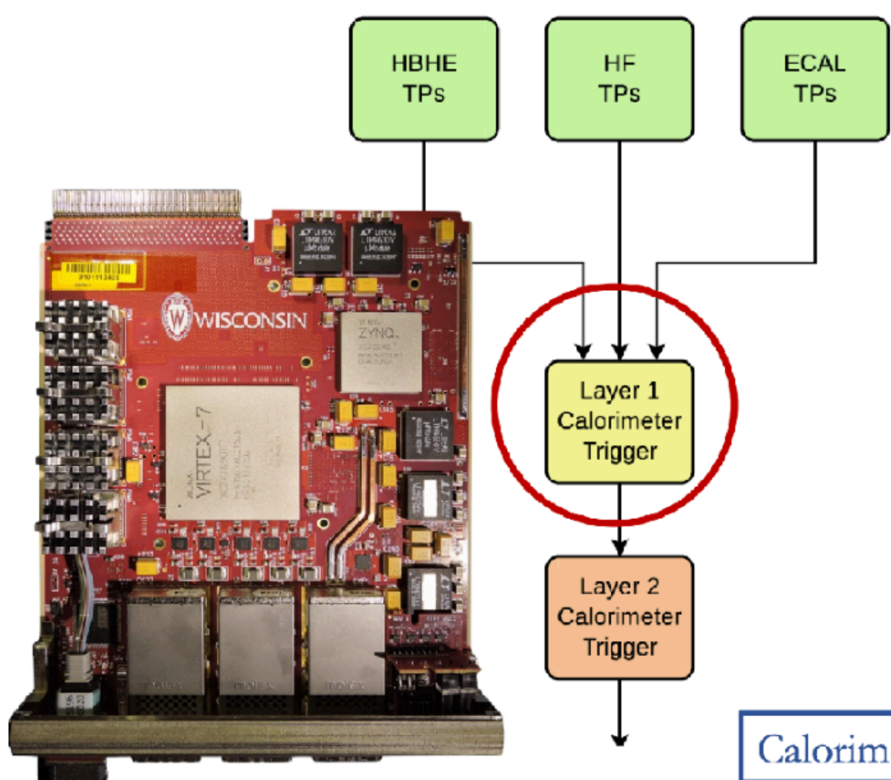




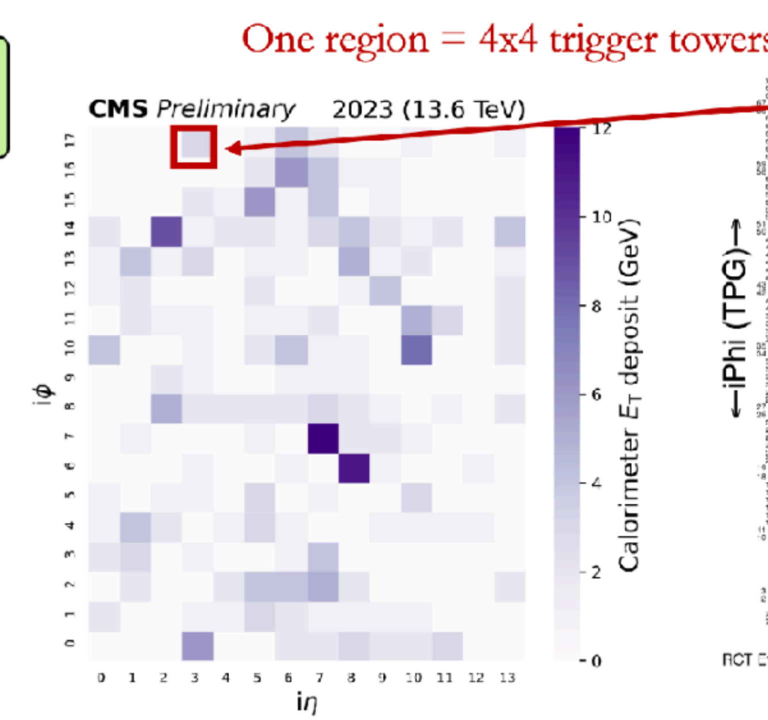
CICADA Design

Public Site

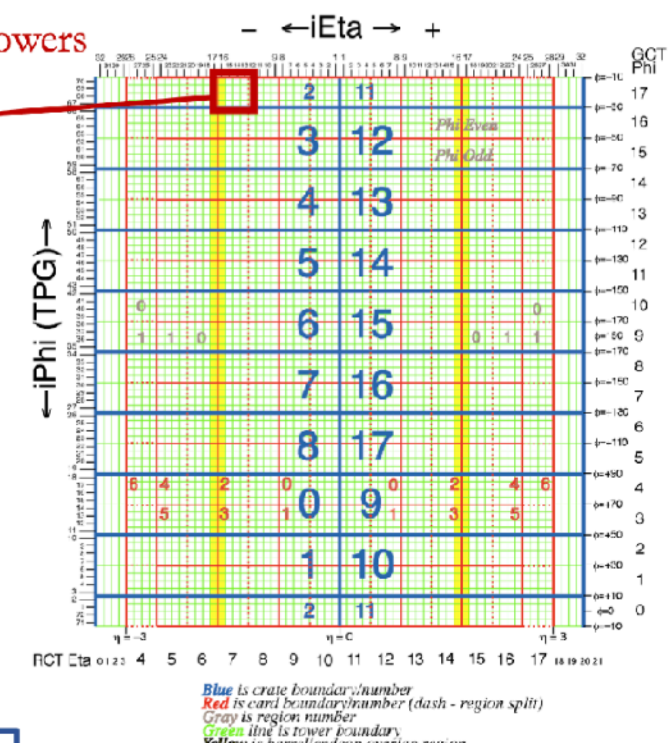
- "**C**alorimeter **I**mage **C**onvolutional **A**nomaly **D**etection **A**lgorithm"
- Anomaly detection using custom board deployed at calorimeter layer-1 subsystem for L1 trigger
- Autoencoder that takes in low-level calorimeter energy deposit info (both ECAL and HCAL)
- Deployed this month!!



Inputs from CaloLayer1



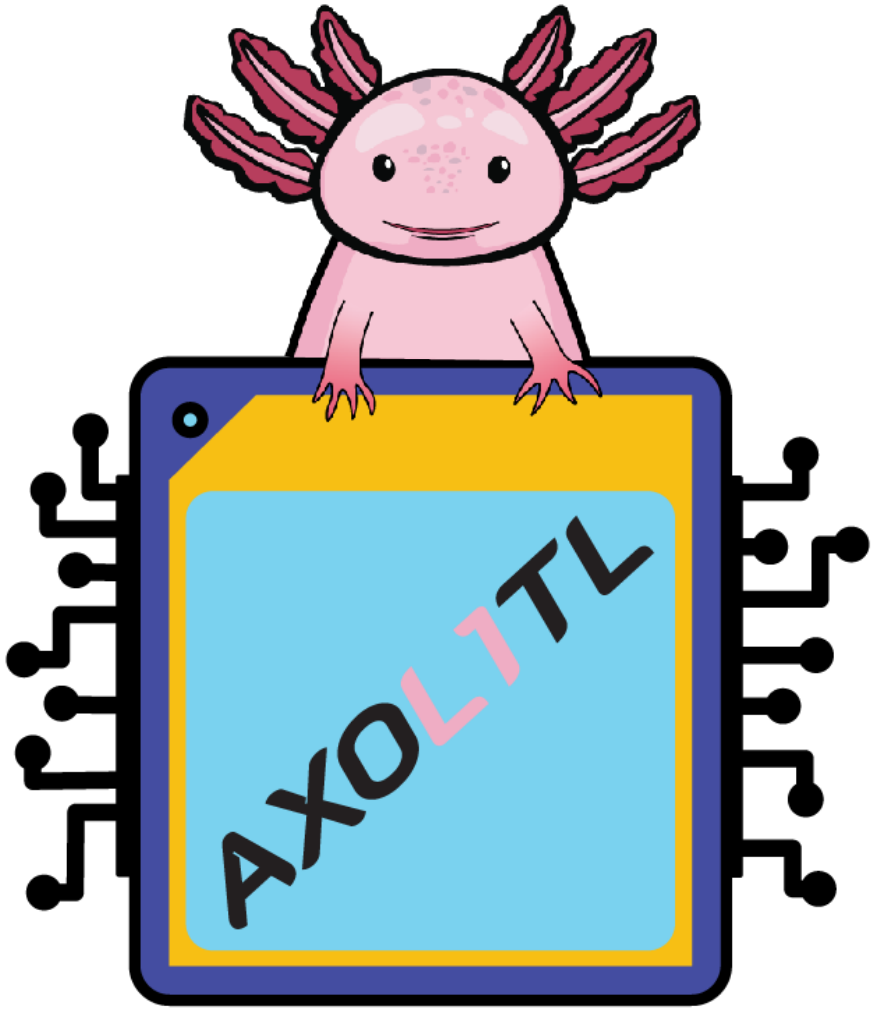
Calorimeter E_T deposit from one Zerobias event



Calorimeter trigger tower-region map

Ho Fung Tsoi

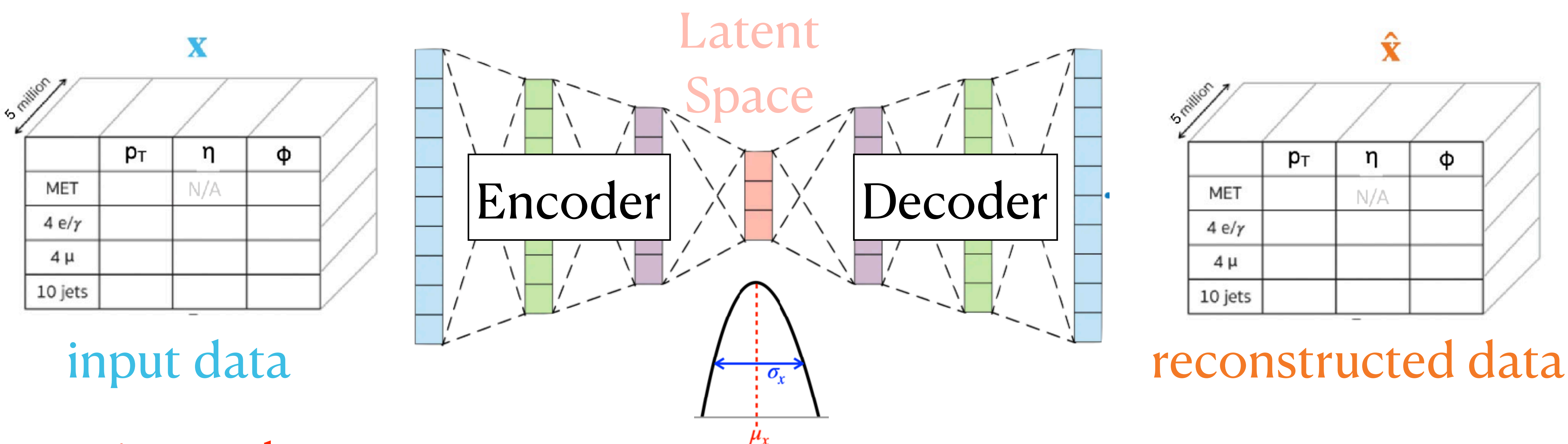
AXOL1TL



- "**A**nomaly e**X**traction **O**nline **L1 T**rigger **L**ightweight" → Anomaly detection at L1 Global Trigger in CMS
    - Unsupervised variational autoencoder that inputs L1 Global trigger objects
    - First ML-based and anomaly detection trigger in CMS
    - Currently deployed and taking data as of May 2024!

AXOL1TL Design

- **Inputs L1 trigger objects**: (pT, η, ϕ) of MET, 4 electron/photons, 4 muons, 10 jets
- **Train on ZeroBias data** collected by CMS in 2023/2024
- **Variational autoencoder**: Additional KL-divergence loss term regularizes latent space to be normal gaussian (N(μ_x, σ_x) → N(0,1)) to prevent overfitting



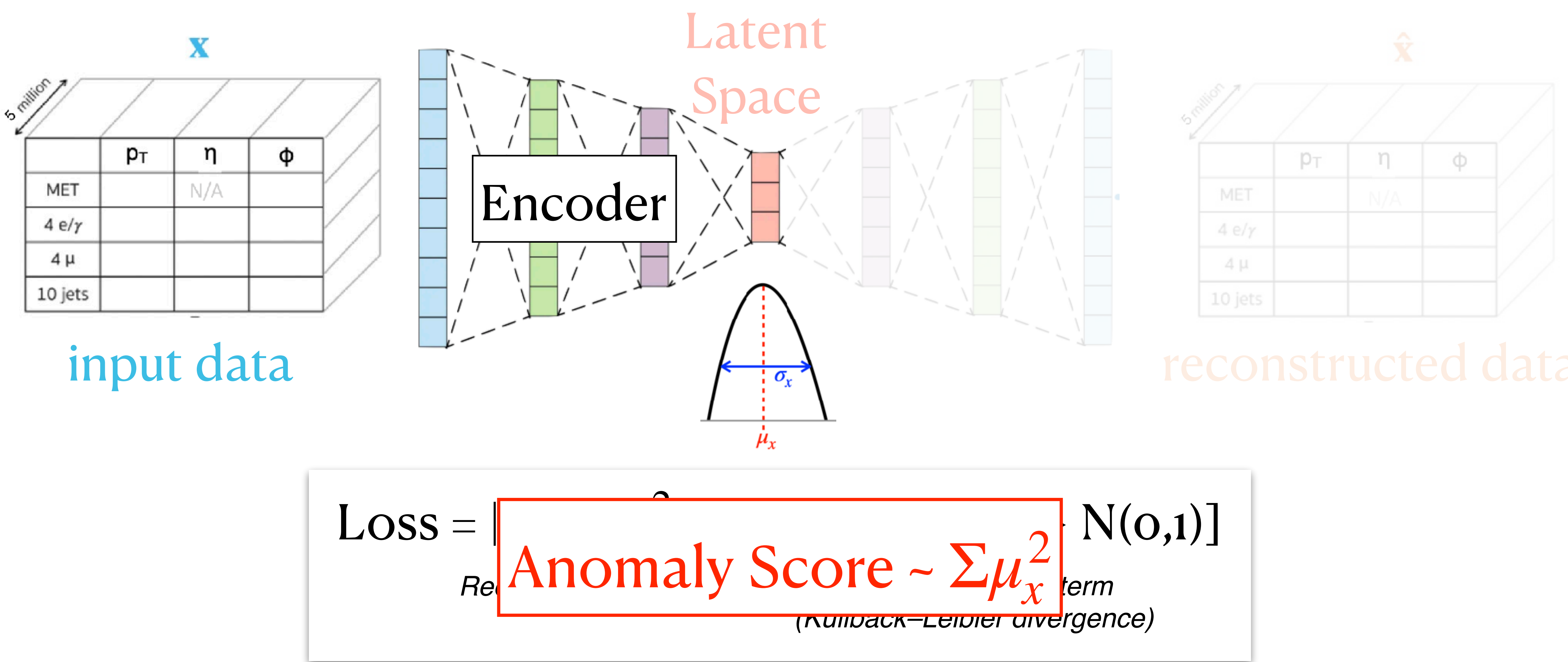
Anomaly Score

$$\text{Loss} = ||x - \hat{x}||^2 + \text{KL}[\text{N}(\mu_x, \sigma_x) \rightarrow \text{N}(0,1)]$$

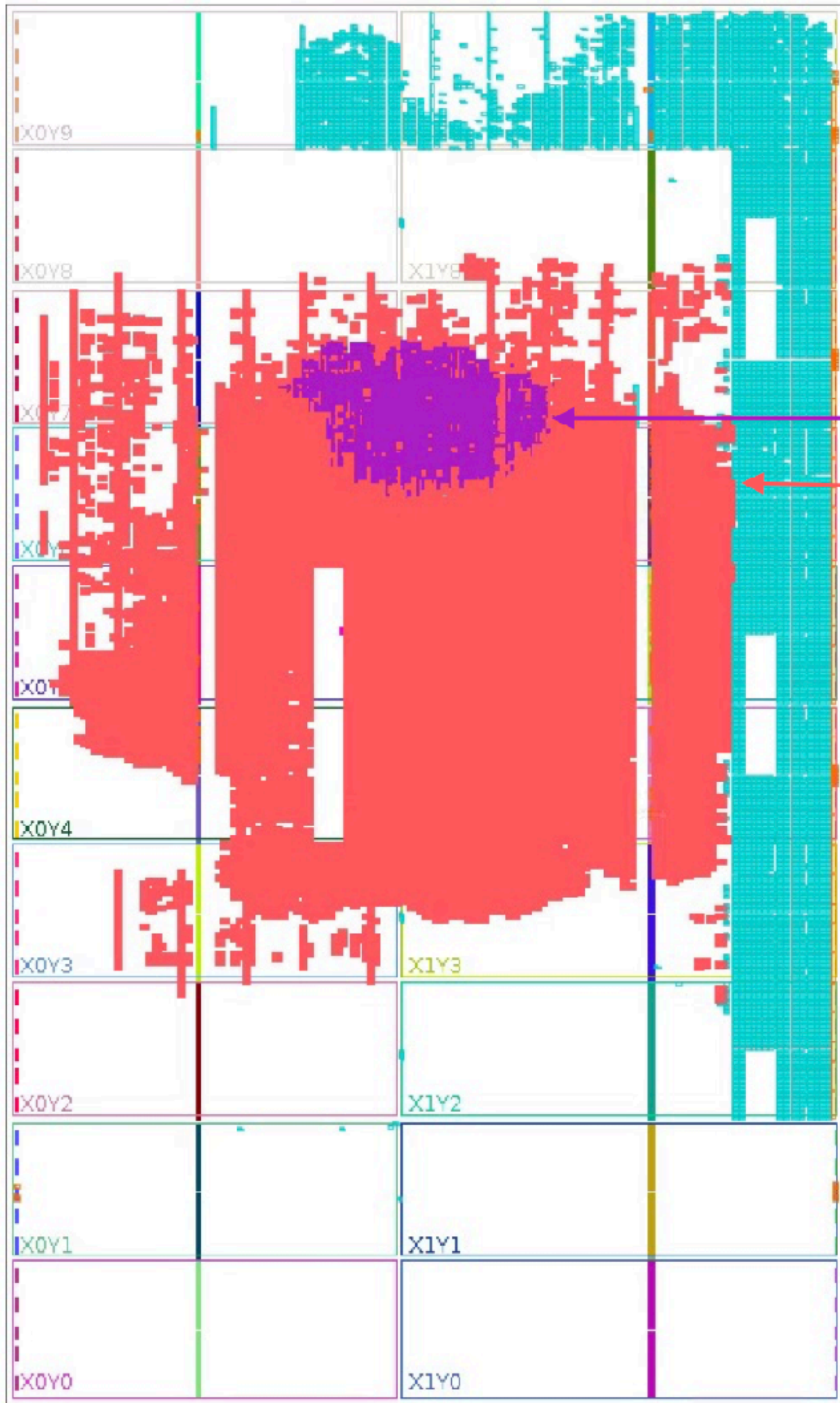
Reconstruction loss *Full regularization term (Kullback–Leibler divergence)*

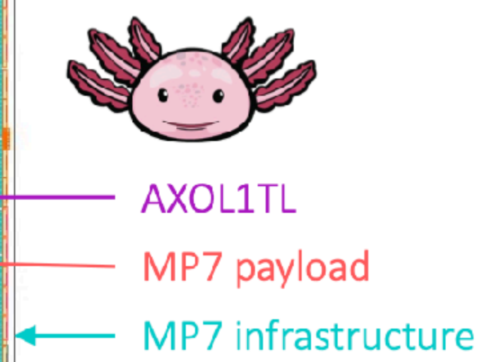
AXOL1TL Design

- Only deploy encoder half of the network, compute degree of abnormality from latent space directly → **Halves the network size and latency!**
- Small, fully connected network architecture
  - **Satisfies strict μGT requirements**: latency of 50ns, low resource utilization

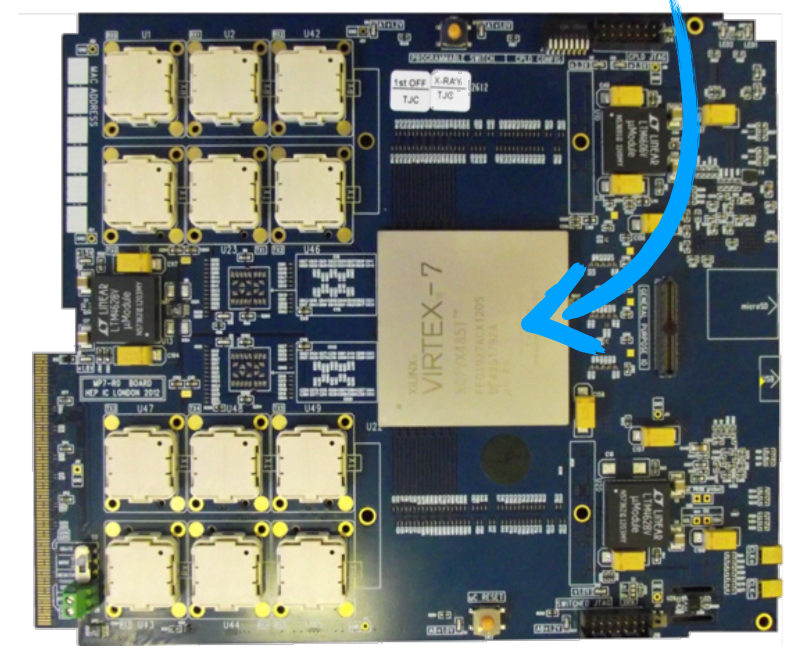


AXOL1TL Implementation





Xilinx Virtex-7 FPGA on MP7 board used at Level-1 Global Trigger



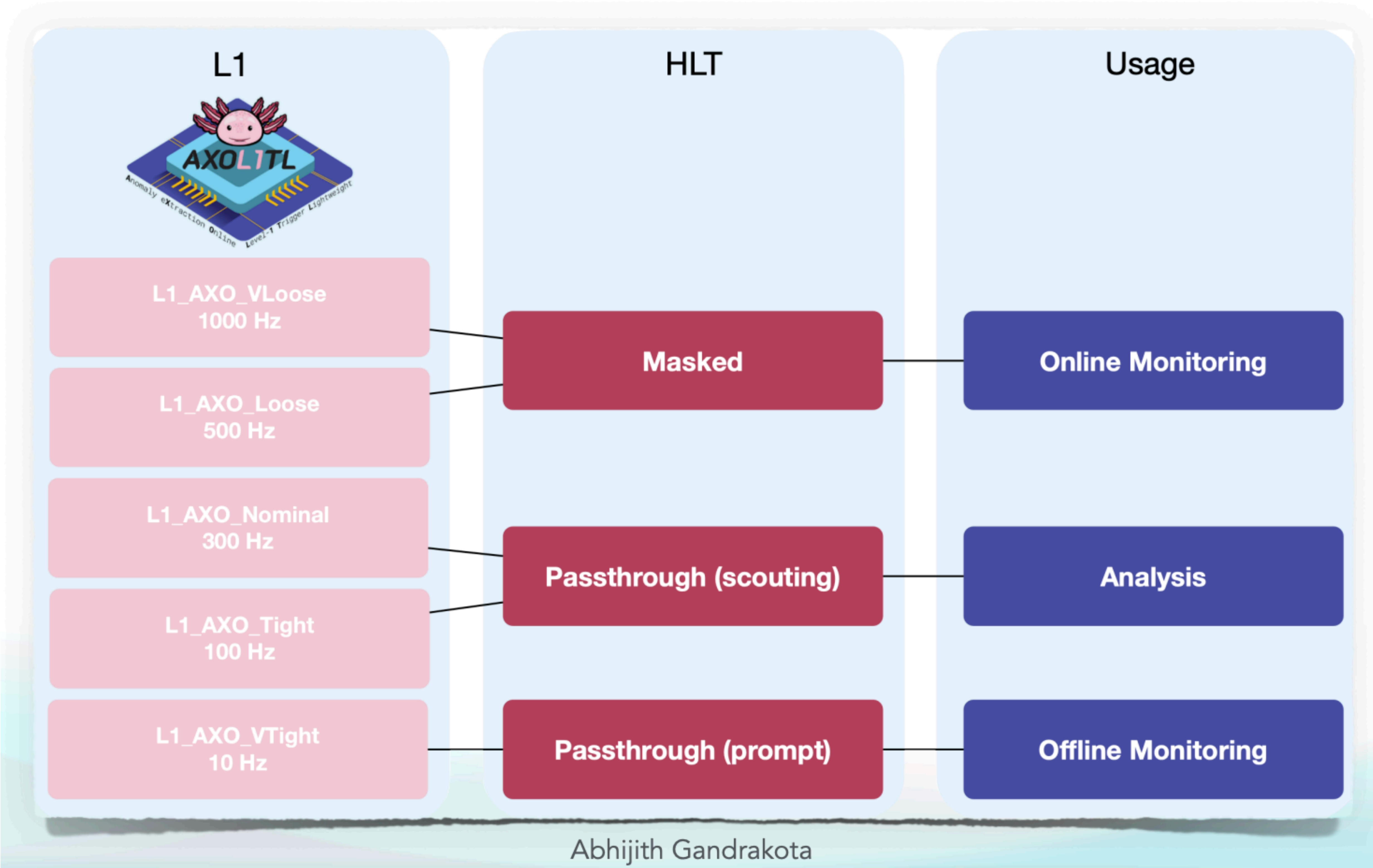
- Implemented on L1 FPGAs
- **Satisfies strict μGT requirements**: latency of 50ns, low resource utilization on FPGAs

hls4ml

Resource utilization of Virtex-7 FPGA chip on Imperial College MP7 µGT board

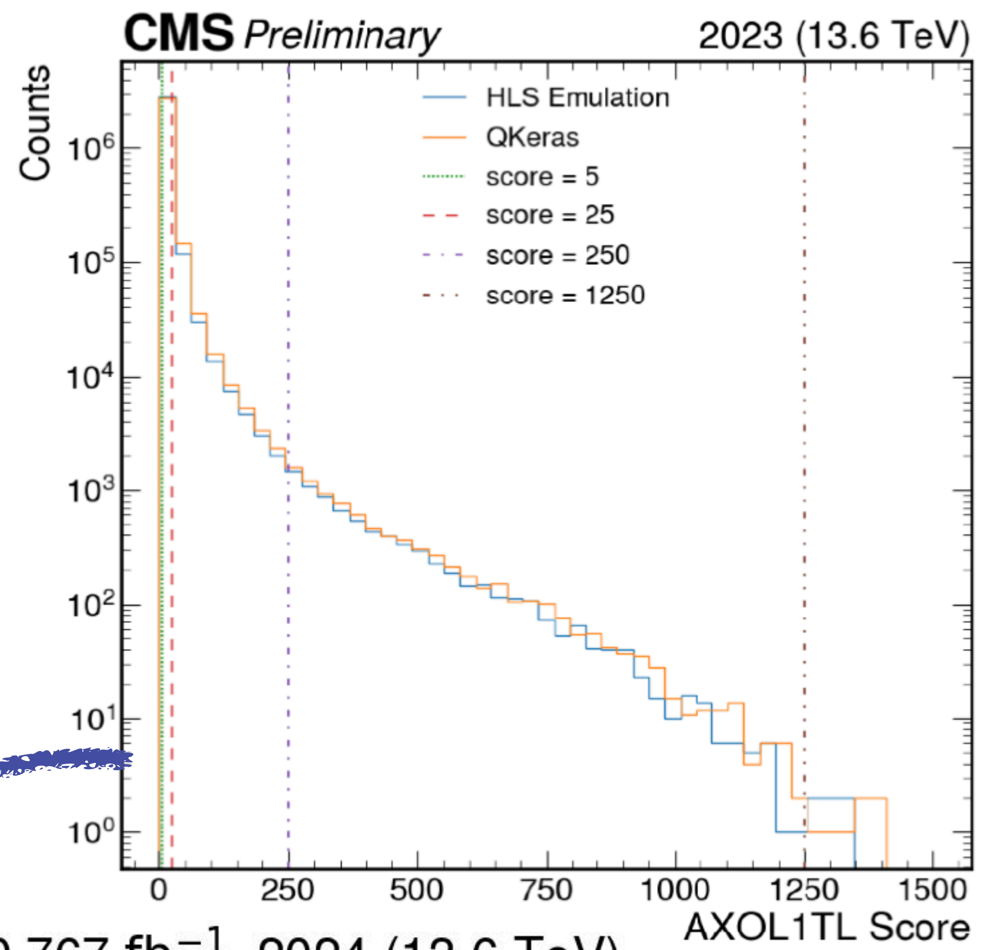
	Latency	LUTs	FFs	DSPs	BRAMs
AXOL1TL	2 ticks 50 ns	2.1%	~0	0	0

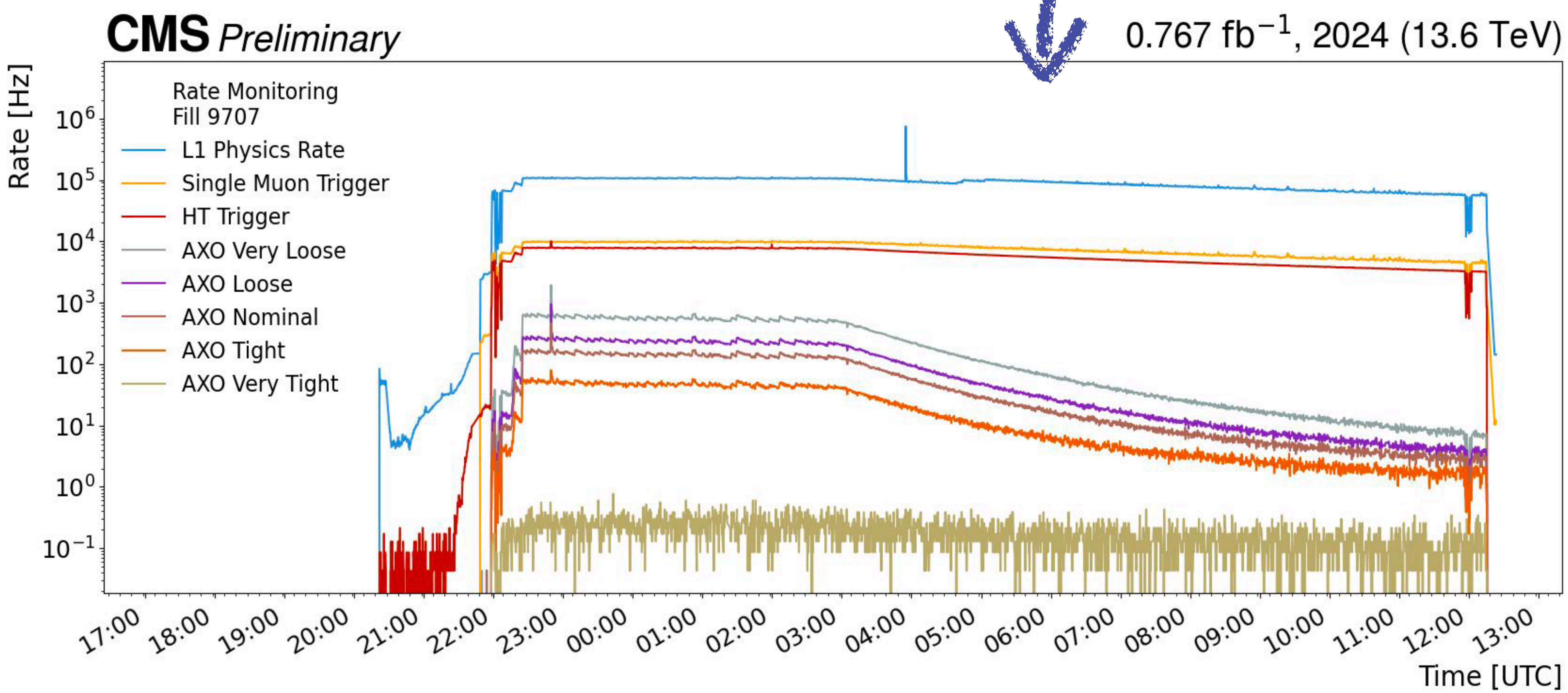
Current L1 Trigger Structure



AXOL1TL Implementation

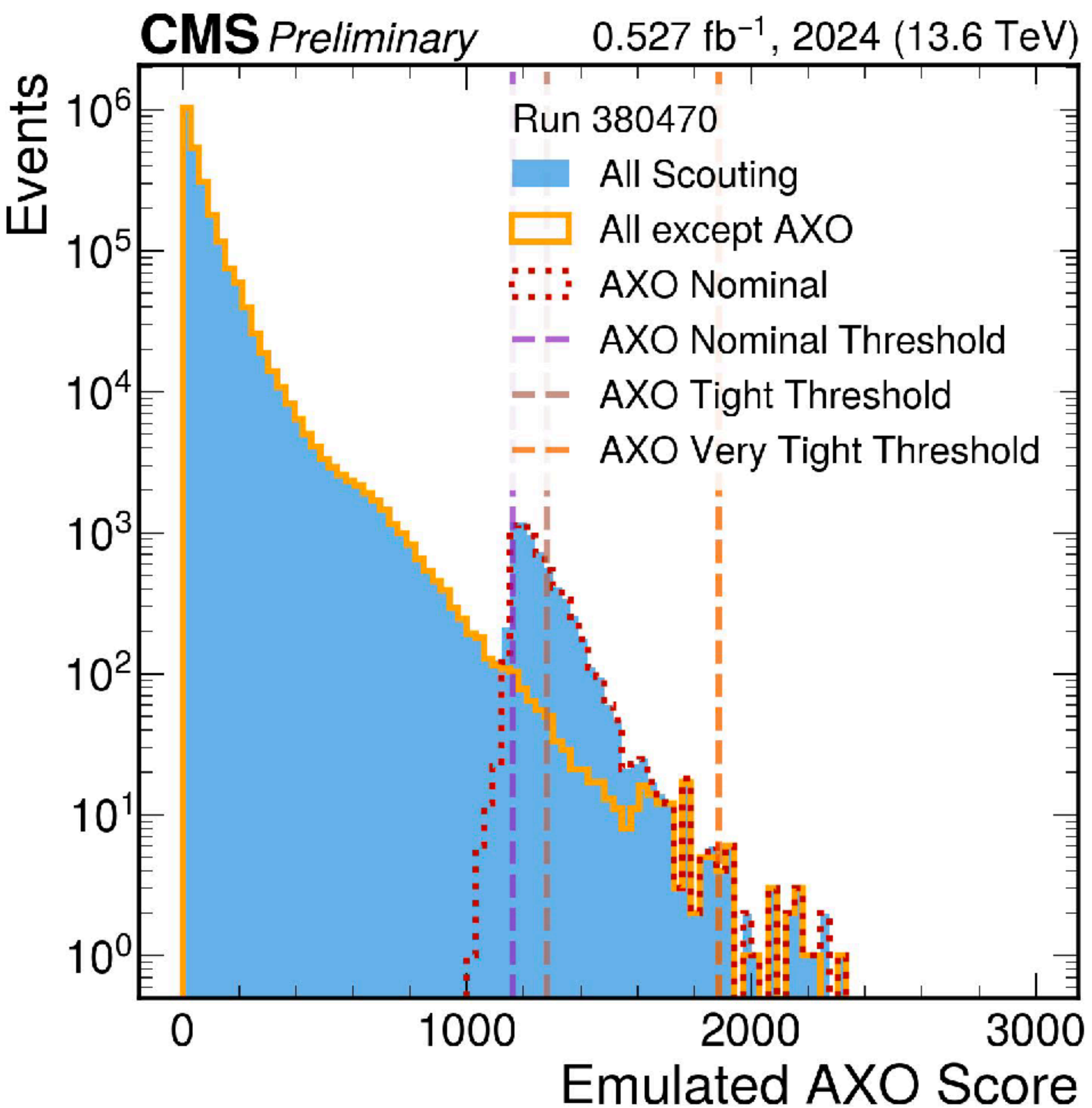
- Score thresholds target different trigger rates
- Rates stable relative to other L1 triggers
- Model updated for new data and model improvement





Preliminary Results

- Improved physics performance relative to existing L1 menu
  - Large signal efficiency increase for low rate increase



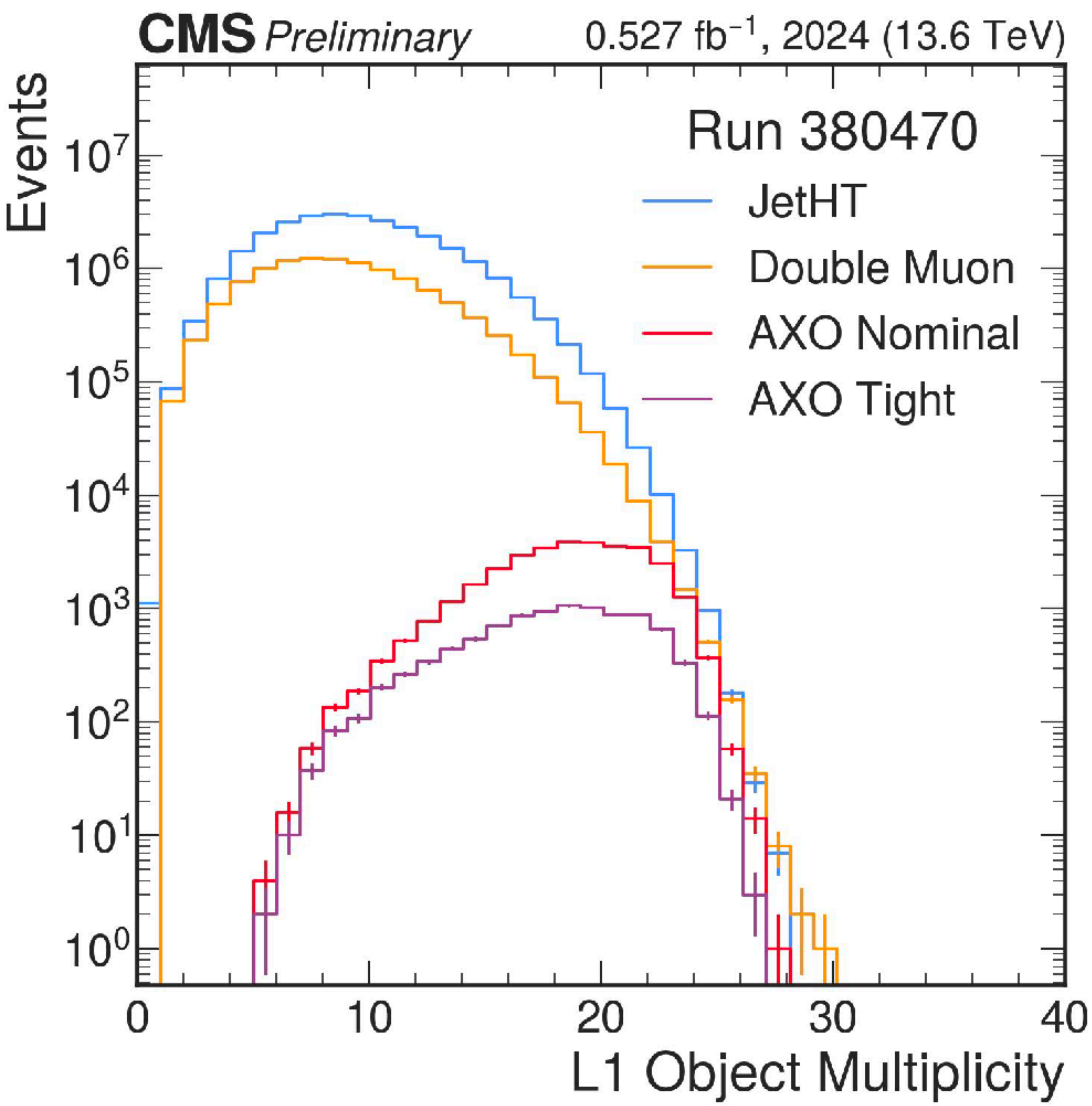
AXOL1TL rate	1 kHz	5 kHz
Signal Efficiency gain	***46%***	***100%***

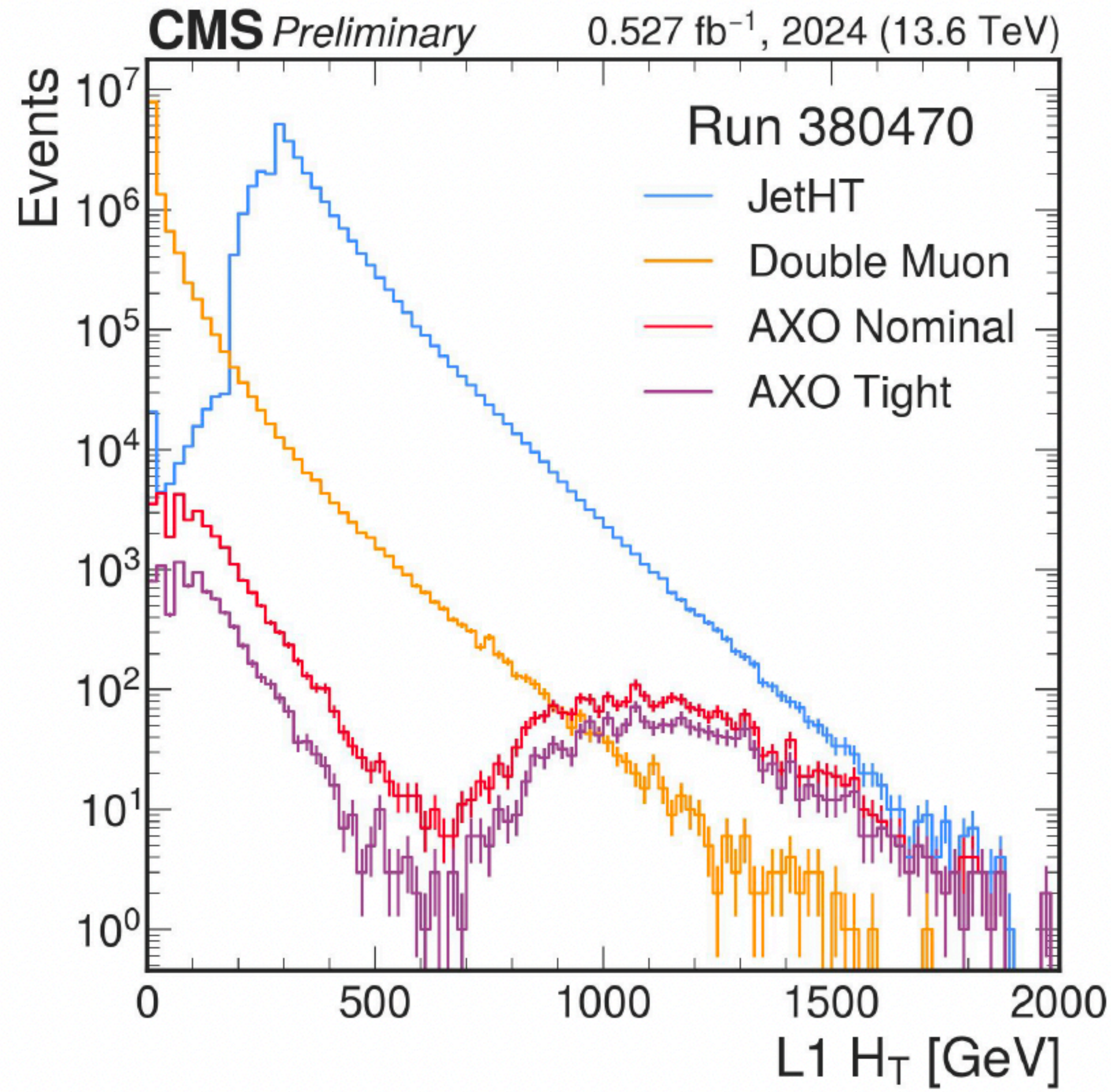
Abhijith Gandrakota

Evaluated on HH → aa → bbbb

→ Many AXOL1TL events are unique vs. existing triggers

Preliminary Results





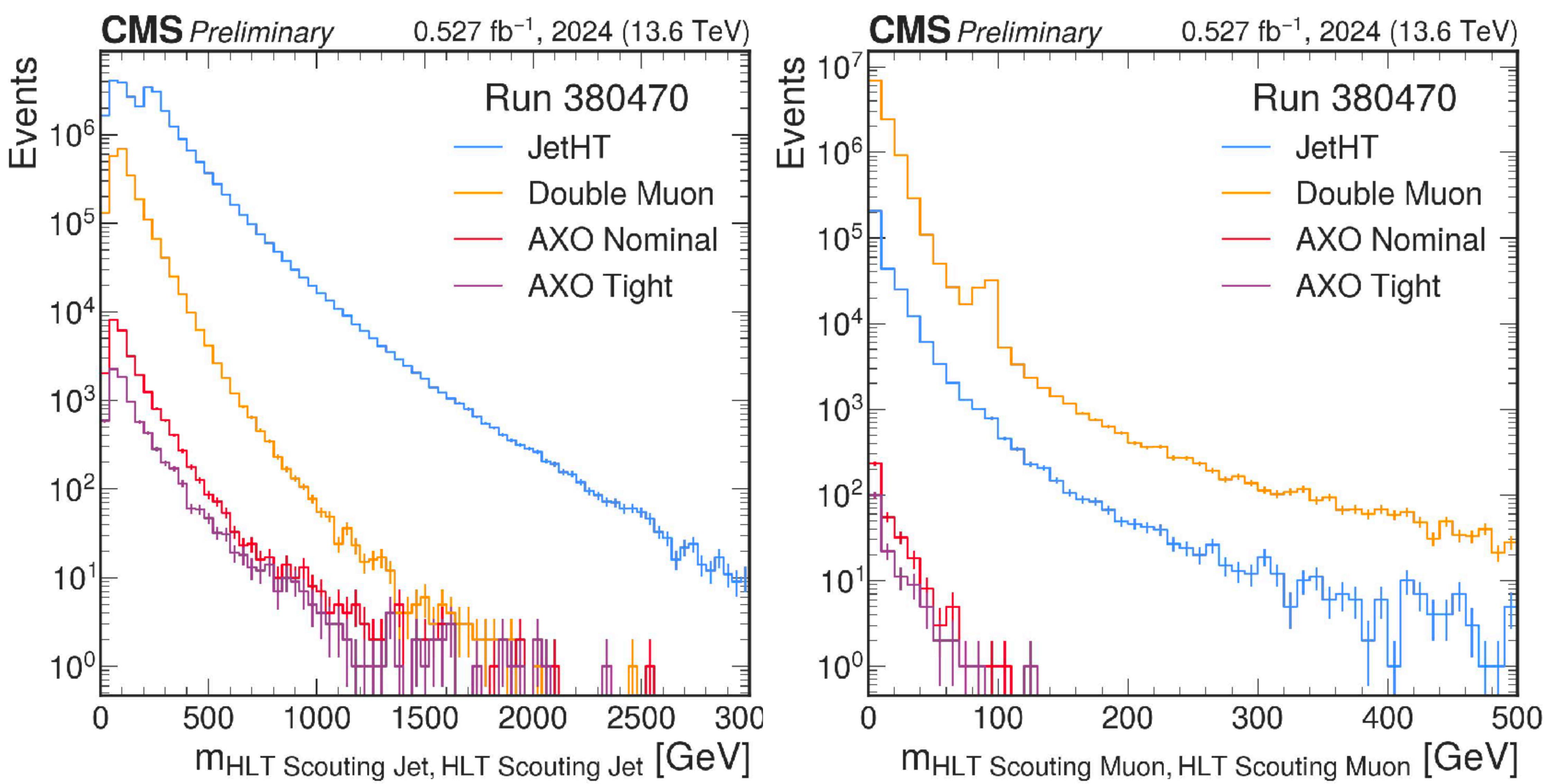
→ Preference for higher object multiplicity vs. existing triggers

→ Preference for higher jet multiplicity & p_T

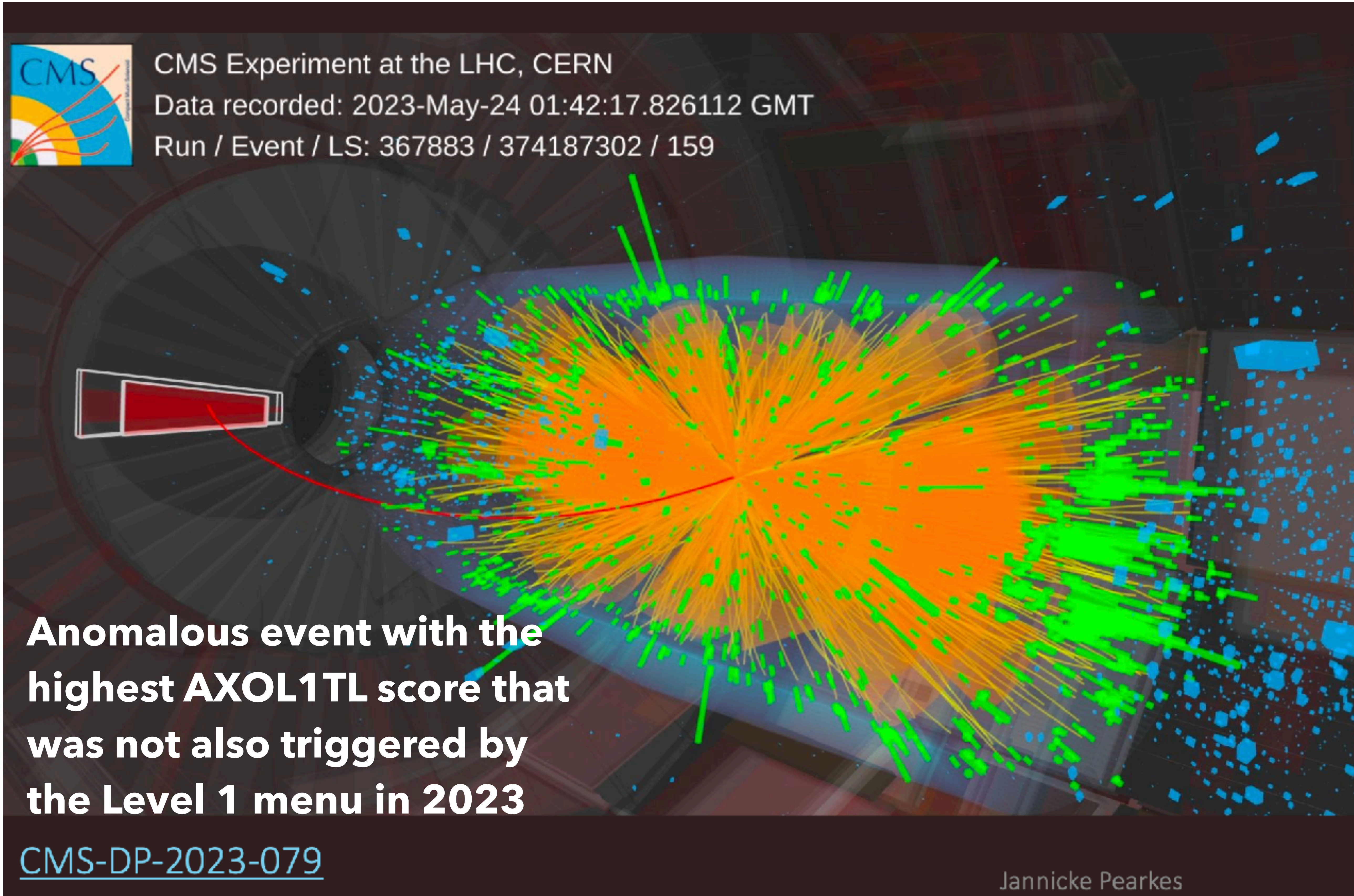
Preliminary Results

→ Preliminary invariant mass distributions recorded in partial scouting dataset (May 2024)

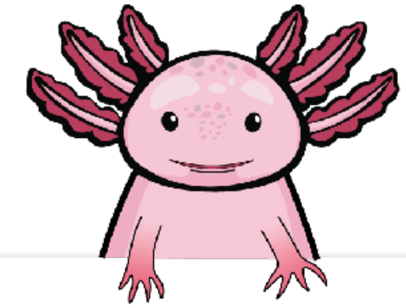
- AXO triggers: Generally smoothly falling distributions, no sculpting
- Ideal for BSM/SM searches



Anomalous Event Example



Conclusion



Publications/Presentations:

[DP Note 1: 2023 Test Crate Implementation]

[DP Note 2: 2024 Data Taking]

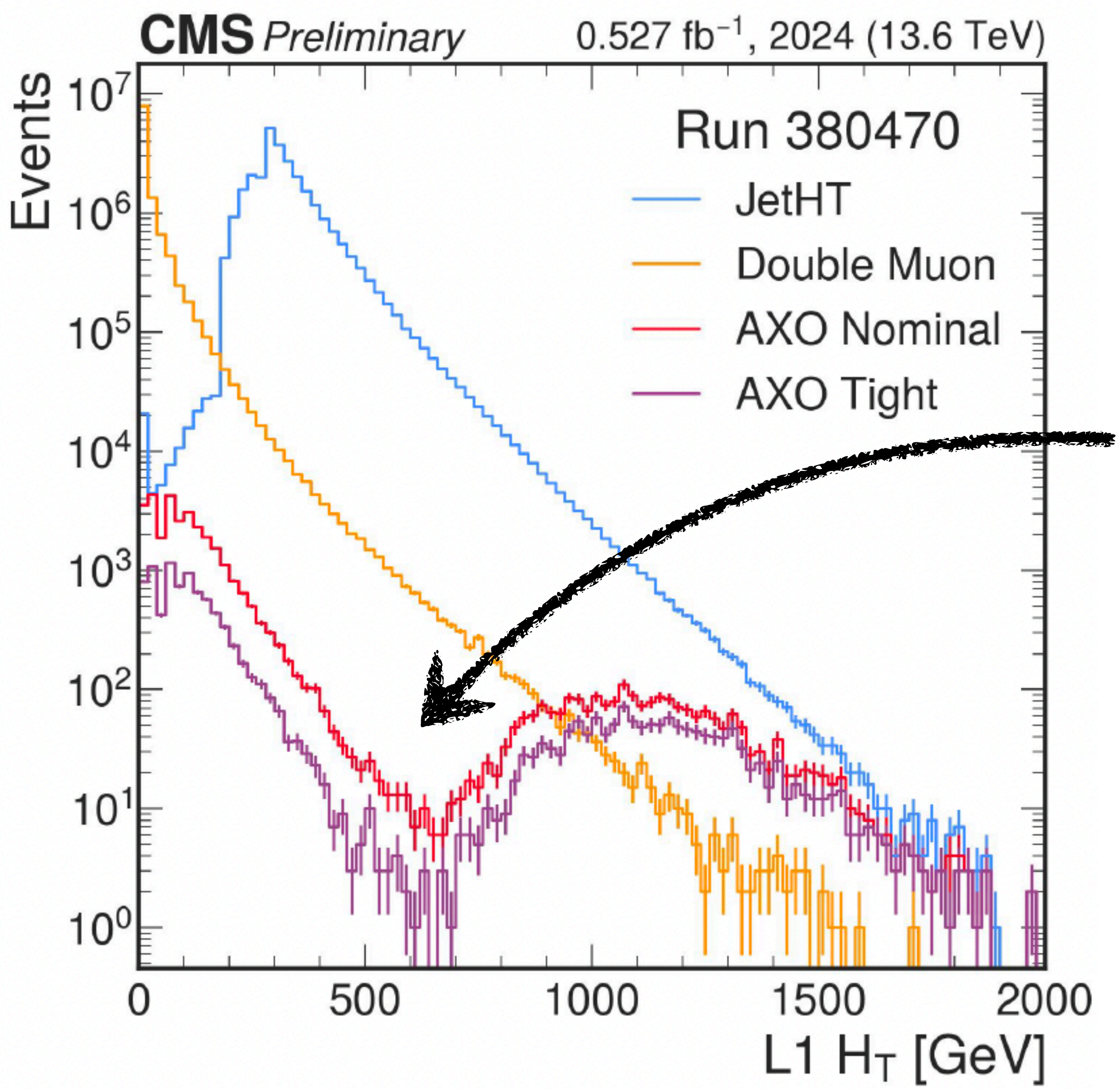
[ICHEP 2024 Presentation (A. Gandrakota)]

- First ML-based anomaly detection trigger deployed in CMS
  - Sensitive to rare SM + potential new physics despite low rates!
- Currently taking data!
  - Improving model for next year + HL-LHC
    - See Axel's talk later today!
  - Developing analysis strategy
- Possible due to support from CMS L1 Trigger community!

AXOL1TL

Anomaly eXtraction Online Level-1 Trigger Lightweight

Preliminary Results

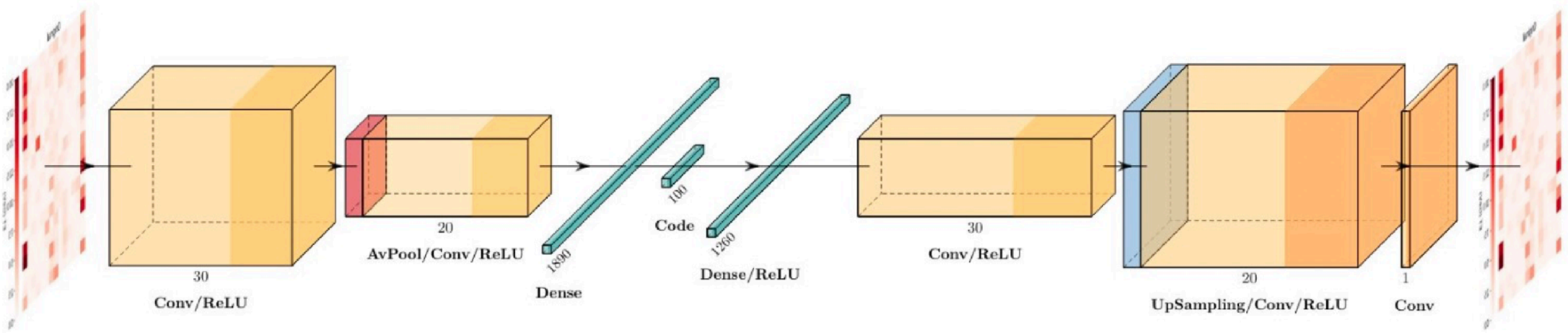


Dip in AXO-selected events due to the abundance of energetic dijet events seen by the L1T, making them less anomalous with respect to other features of interest

→ Preference for higher jet multiplicity & p_T

CICADA Design

Model architecture: calo input → encoder → latent space → decoder → reconstructed input



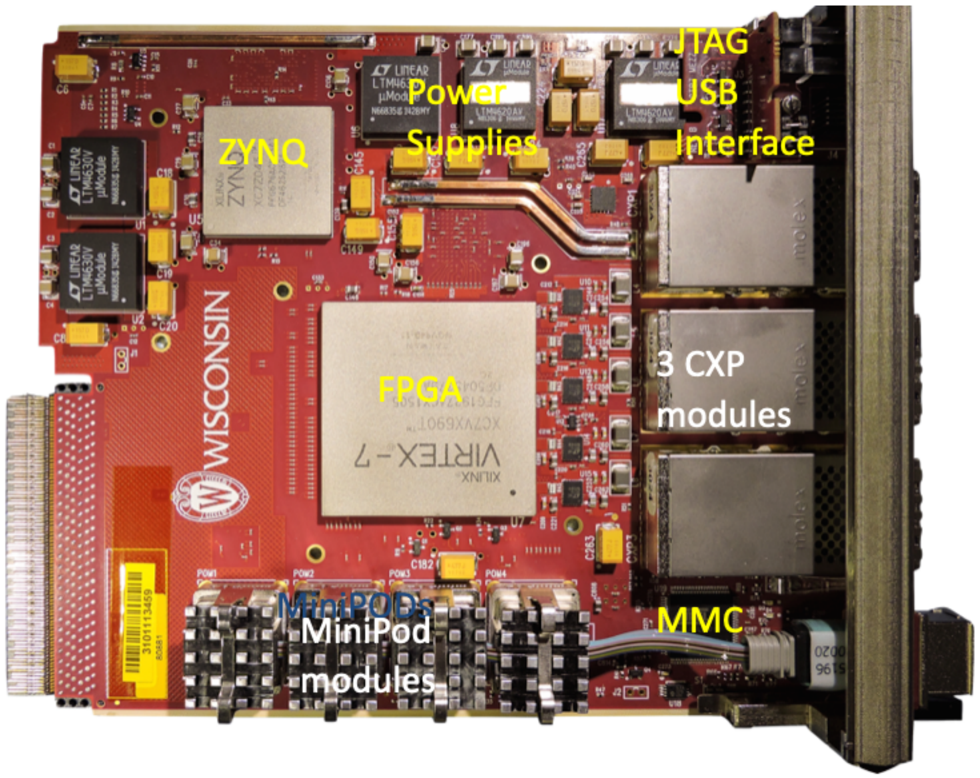
- Unsupervised autoencoder trained on ZeroBias data

Ho Fung Tsoi

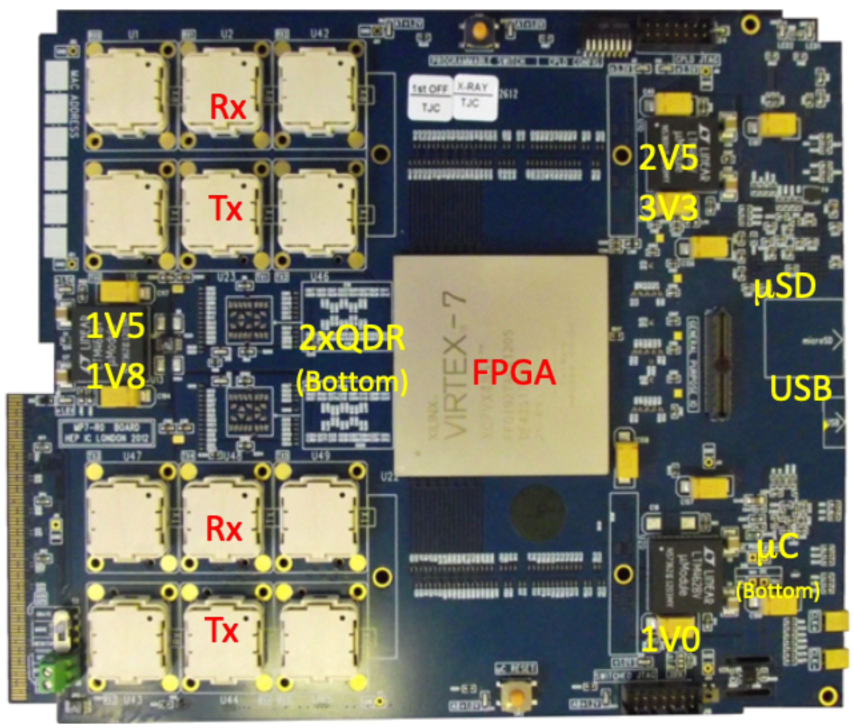
  - Input 2D tensor calorimeter energy maps
  - Encoder and decoder are convolutional neural networks
  - Learn reconstruction of input energy maps
  - Uses mean-squared error MSE(input, output) as anomaly score to trigger on anomalous events
  - Uses knowledge distillation to train a smaller "student" model to directly regress MSE to replicate "teacher" performance to fit strict L1T constraints

CMS FPGAs

CTP7



MP7



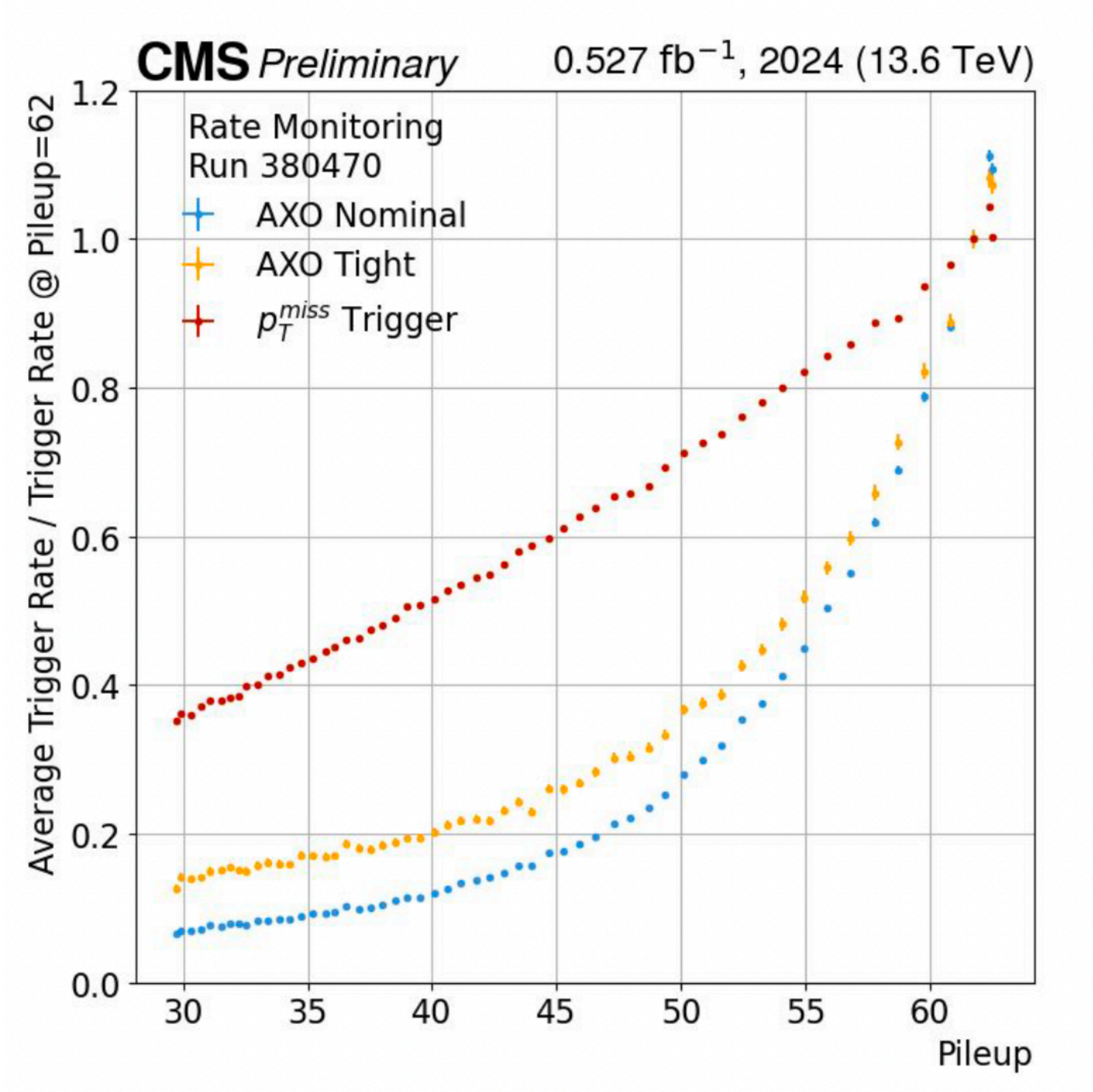
Calorimeter Trigger Processor(CTP7 – left), and Master Processor (MP7 - right)

- CTP7 (Layer-1) – μTCA Single Virtex 7 FPGA, 67 optical inputs, 48 outputs, 12 RX/TX backplane
  - Virtex 7 allows 10 Gb/s link speed on 3 CXP(36 TX & 36 RX) and 4 MiniPODs (31 RX & 12 TX)
  - ZYNQ processor running Xilinx PetaLinux for service tasks, including virtual JTAG cable
- MP7 (Layer-2) – μTCA Single Virtex 7 FPGA, up to 72 input & output links
  - Virtex 7 has 72 input and output links at 10 Gb/s
  - Dual 72 or 144MB QDR RAM clocked at 500 MHz

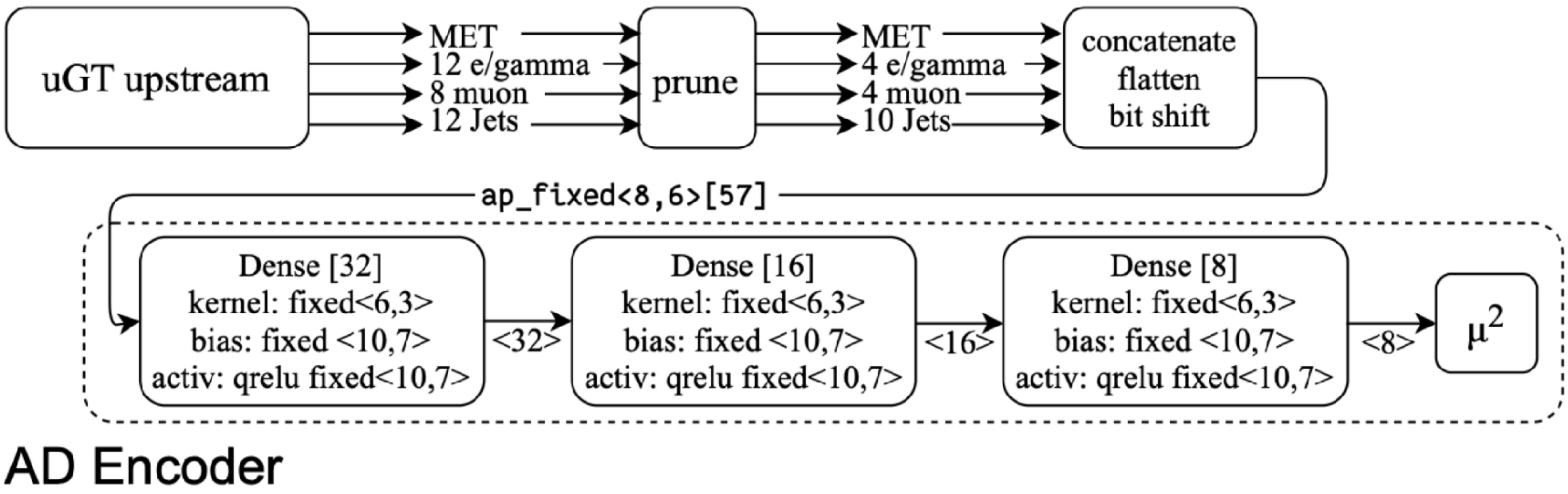
Sridhara Dasu

Pileup Dependence

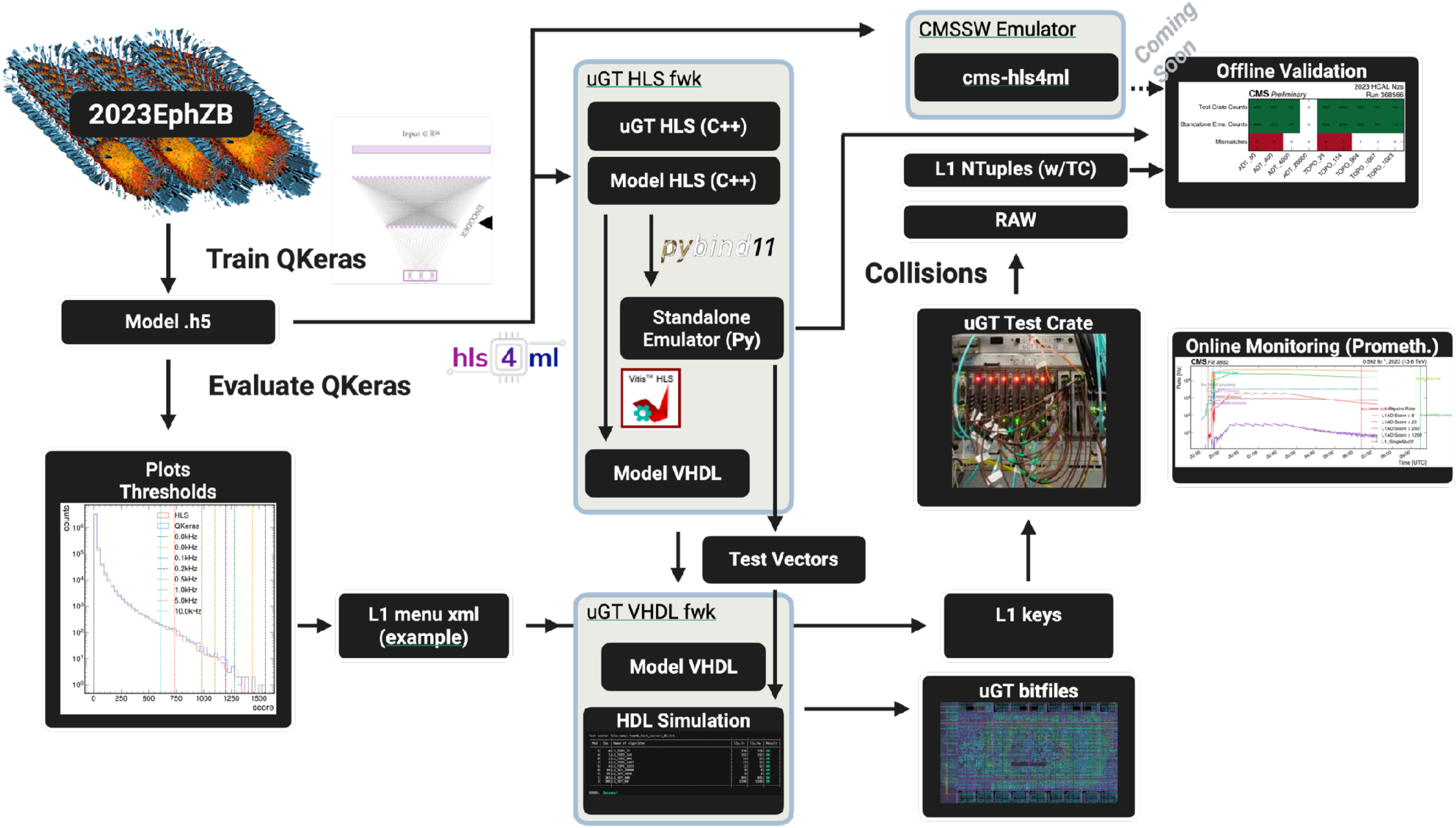
- Strong pileup correlation due to correlation with high object multiplicity and high total energy, which increase at higher pileup.
  - Pileup mitigation studies underway



AD Encoder Implementation



Implementation



Anomaly Detection with CICADA

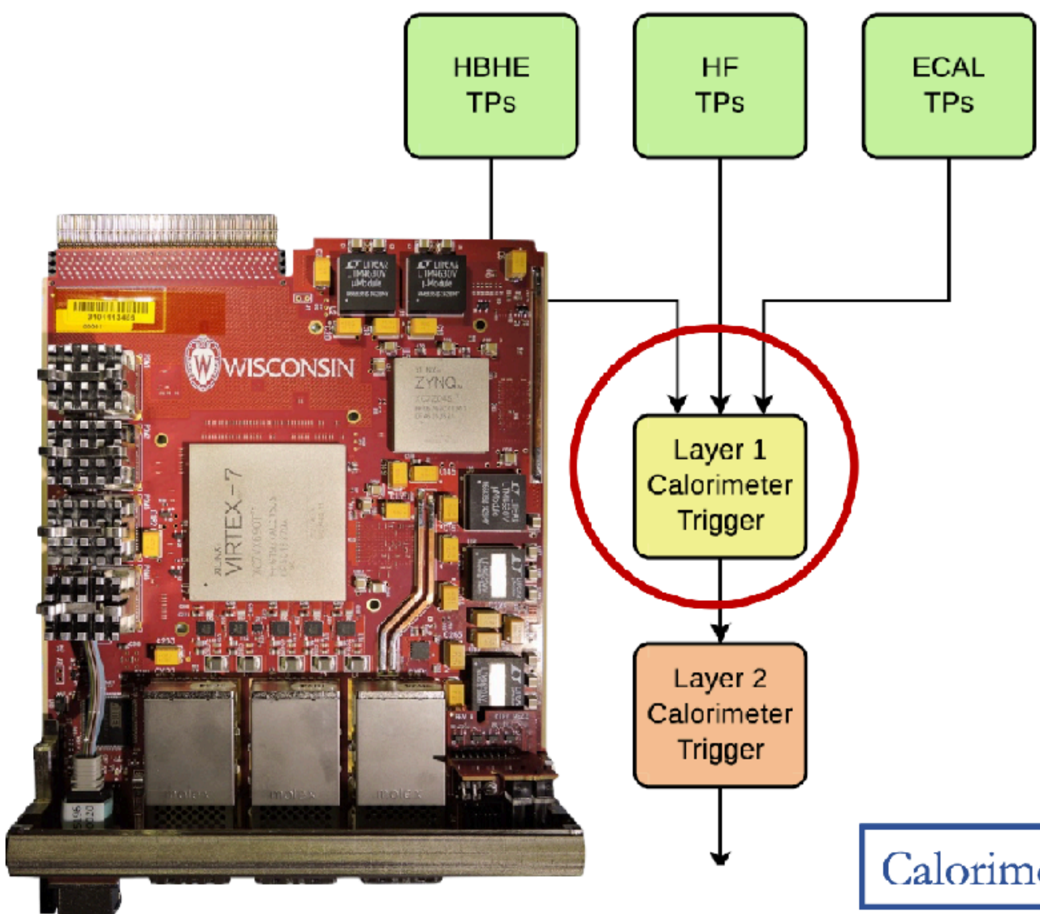
- "**C**alorimeter **I**mage **C**onvolutional **A**nomaly **D**etection **A**lgorithm" → Anomaly detection at calorimeter layer-1 subsystem for L1 trigger
  - Autoencoder that takes in low-level calo trigger inputs
  - Uses custom board deployed at L1 calo layer 1
  - Deployed ~last week!!

CICADA

Public Site

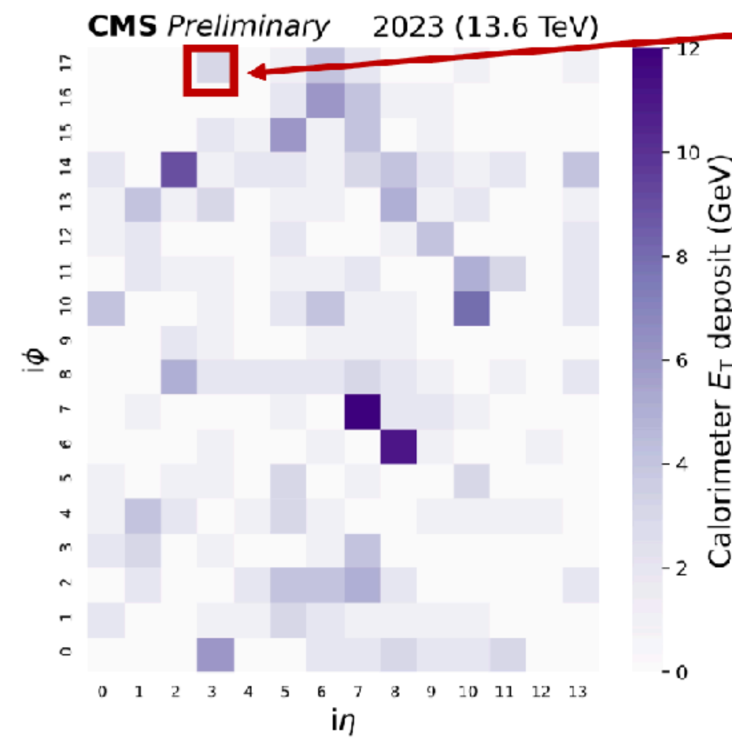
CICADA Design

- Inputs low level calorimeter energy deposit information from calo layer 1 (both ECAL and HCAL)
- $18\phi \times 14\eta = 252$ calo tower regions = ~2D summary of energy distribution profile for each region

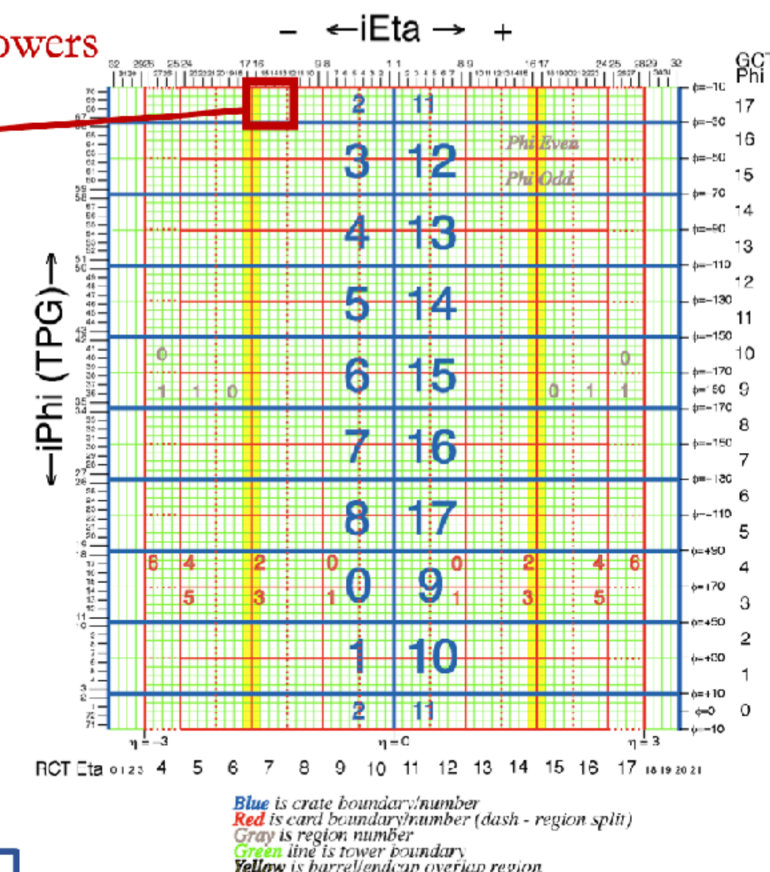


Inputs from CaloLayer1

One region = 4x4 trigger towers



Calorimeter E_T deposit from one Zerobias event

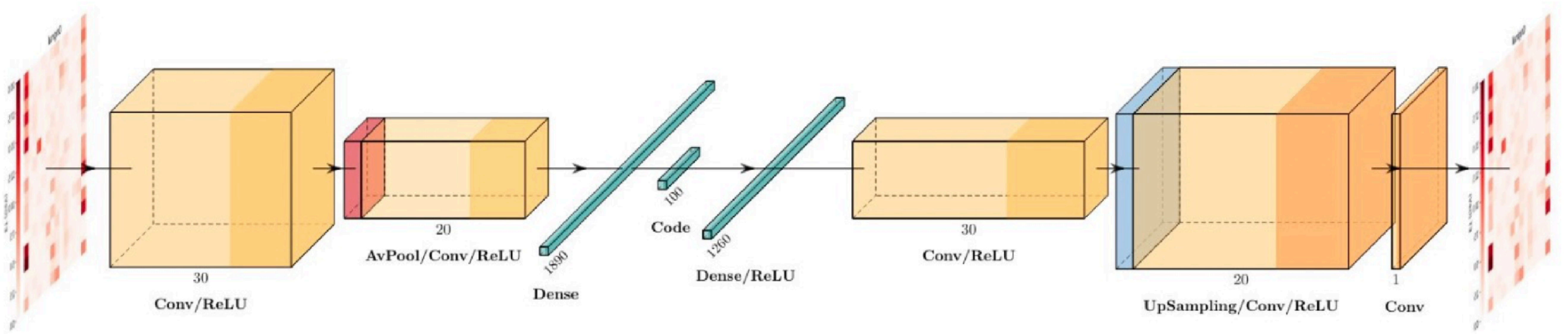


Calorimeter trigger tower-region map

Ho Fung Tsoi

CICADA Design

Model architecture: calo input → encoder → latent space → decoder → reconstructed input



Ho Fung Tsoi

- Unsupervised autoencoder trained on ZeroBias data
  - Input 2D tensor calorimeter energy maps
  - Encoder and decoder are convolutional neural networks
  - Learn reconstruction of input energy maps
  - Uses mean-squared error MSE(input, output) as anomaly score to trigger on anomalous events
  - Uses knowledge distillation to train a smaller “student” model to directly regress MSE to replicate “teacher” performance to fit strict L1T constraints