Particle-Based Representation Learning for Anomaly Detection in the CMS High-Level Trigger

Mehrnoosh Moallemi

Prof. Claire Shepherd-Themistocleous (RAL), Prof. Stefano Moretti (U. Southampton)
Dr. Jennifer Ngadiuba (Fermilab), Dr. Maciej Glowacki (CERN)

IOP Joint APP and HEPP Annual Conference

Science & Technology Facilities Council
Rutherford Appleton Laboratory

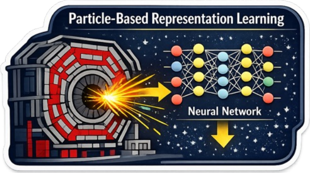
University of Southampton

NexTGen
Next Generation Triggers

Particle-Based Representation Learning for Anomaly Detection in the CMS High-Level Trigger

CMS
Compact Muon Solenoid
Science & Technology Facilities Council
Rutherford Appleton Laboratory
University of Southampton
NexTGen
Next Generation Triggers

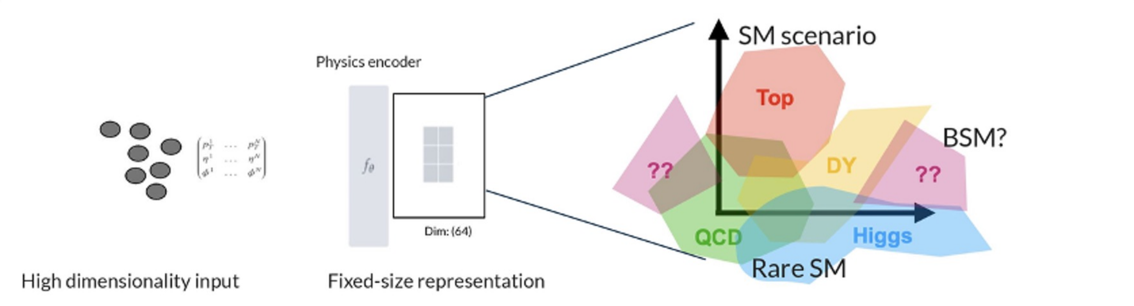
What is Particle-Based Representation Learning?



Representation Learning allows a model to transform raw, high-dimensional particle data into a compact latent space.

- By training on a diverse set of simulated Standard Model processes, the model constructs a **'physics-informed'** representation space (latent space) from raw particle kinematics.

- This latent space serves as a compact **physics summary**, helping the model distinguish signal from background even under extreme pileup.



Particle-Based Representation Learning for Anomaly Detection in the CMS High-Level Trigger

Science & Technology Facilities Council
Rutherford Appleton Laboratory

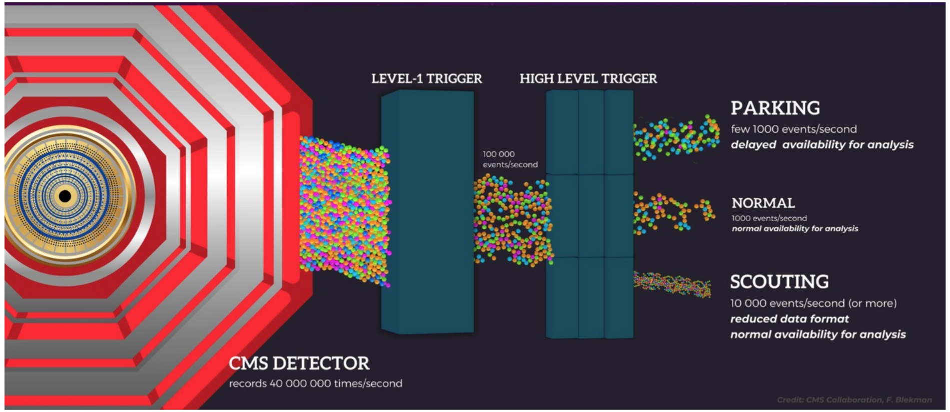
University of Southampton

NexTGen
Next Generation Triggers

What is CMS High-Level Trigger?

CMS has a two-level trigger system that selects potentially interesting events and sits between the detector readout and permanent storage.

- **Level-1 Trigger (L1T)** is implemented in fast **hardware** (including FPGAs) to make decisions within microseconds.
- **High-Level Trigger (HLT)** is a **software-based** system that forms the final stage of the CMS trigger and makes the final “accept or discard” decision.



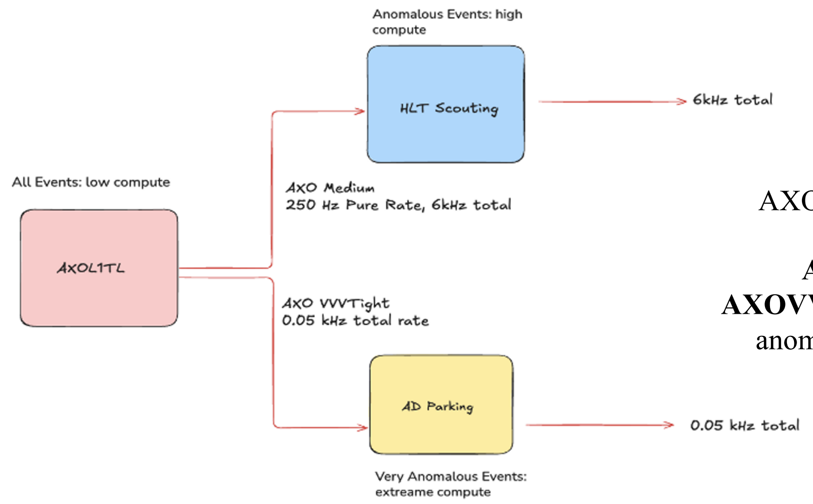
Particle-Based Representation Learning for Anomaly Detection in the CMS High-Level Trigger

Science & Technology Facilities Council
Rutherford Appleton Laboratory
University of Southampton
NexTGen Next Generation Triggers

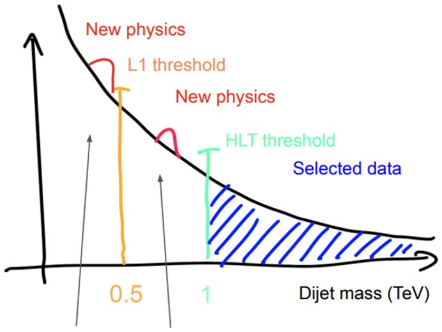
What is Anomaly Detection?

Rare events cannot be discovered if they are not recorded!

- The anomaly detection goal is to detect events that deviate significantly from known physics.
- The CMS Level-1 anomaly trigger, the **A**nomaly e**X**traction **O**nline **L**evel-**1** **T**rigger a**L**gorithm (AXOL1TL), has been deployed in data taking since May 2024.
    - This allows the most anomalous events to receive the full offline reconstruction.



AXOL1TL provides several predefined working points, such as **AXOMedium**, **AXOTight**, and **AXOVVTight**, which correspond to different anomaly score thresholds used to define specific L1 seeds.



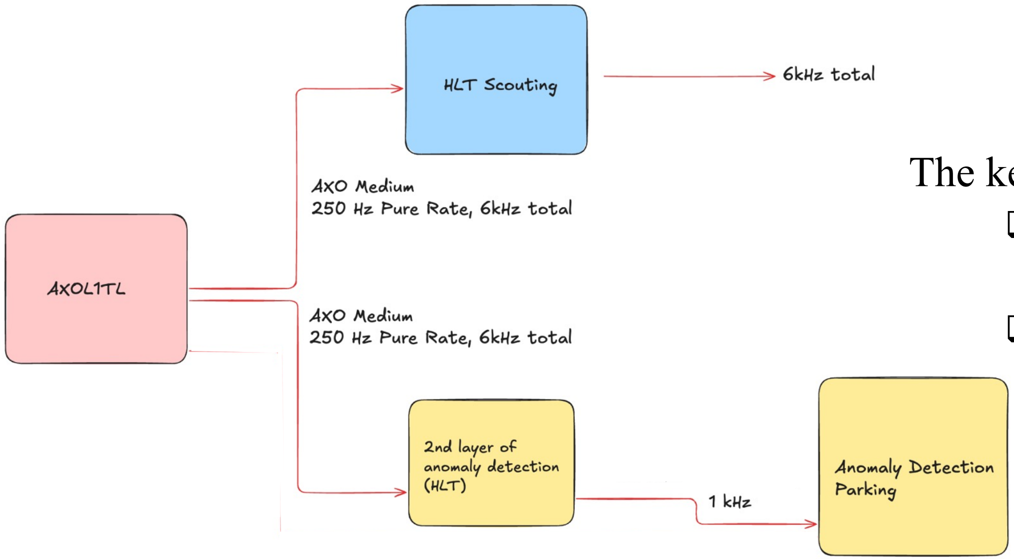
Particle-Based Representation Learning for Anomaly Detection in the CMS High-Level Trigger

CMS Compact Muon Solenoid
Science & Technology Facilities Council
Rutherford Appleton Laboratory
University of Southampton
NexTGen Next Generation Triggers

Anomaly Detection @ HLT

We proposed a second layer of anomaly detection at the CMS High-Level Trigger (HLT):

- The output of AXOMedium serves as the input to this model, which processes at a rate of 6 kHz.
- It then feeds the Anomaly Detection Parking stream at 1 kHz.



The key motivation is:

- ❑ Enhance the purity of anomalous events in the AD parking stream.
- ❑ Provide an opportunity to test the L1T strategy for CMS HL-LHC (Phase-II) using the current HLT.

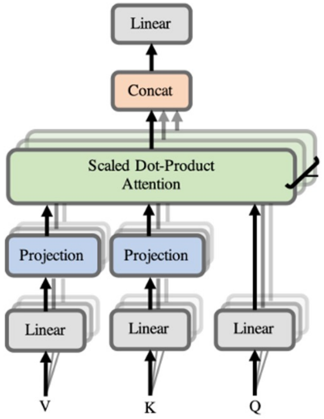
Anomaly Detection @ HLT: Training Step

- The model is trained on a diverse set of simulated SM processes, including Drell-Yan, QCD, $t\bar{t}$, and W+jets.
- Particle Flow (PF) candidates and their kinematic features (p_T , φ, d_{xy}, d_{xy}sig, particle Id) are used as input objects.
- We trained the model using the Linformer encoder architecture with use of InfoNCE and cross entropy losses:

draw decision boundaries

$$\mathcal{L} = \beta \cdot \mathcal{L}_{\text{NCE}} + \mathcal{L}_{\text{CE}}$$

shape the latent space



Model Performance: Evaluation Step

New samples are needed to evaluate the model:

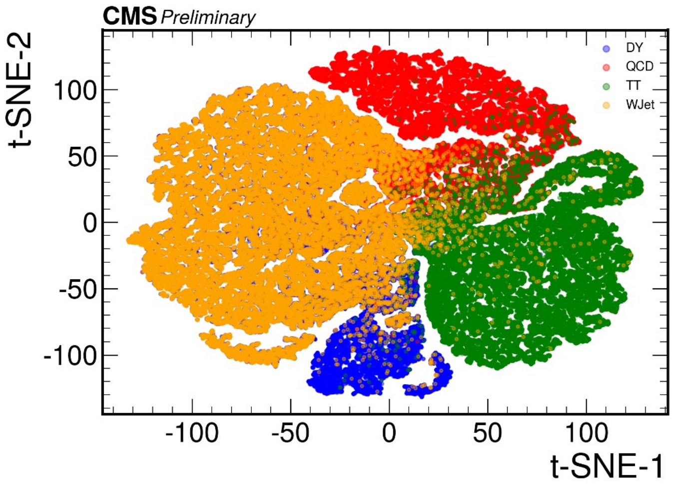
- A set of unseen events from SM
- VBF events

passed the AXOMedium seeds

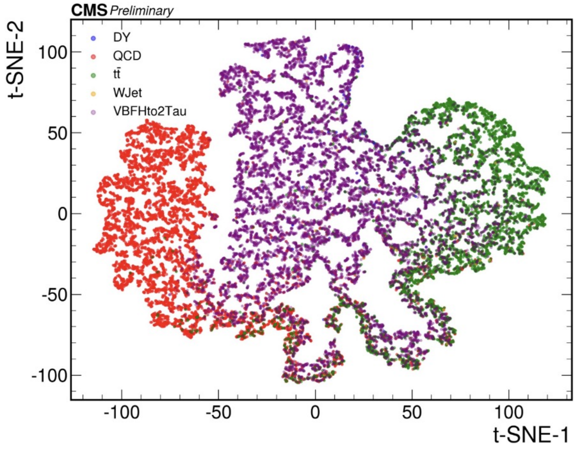
The t-SNE plot visualises the model's latent space.

It discriminates between SM backgrounds and anomalous signals by mapping them into distinct regions of the latent space.

Training samples



Evaluation samples

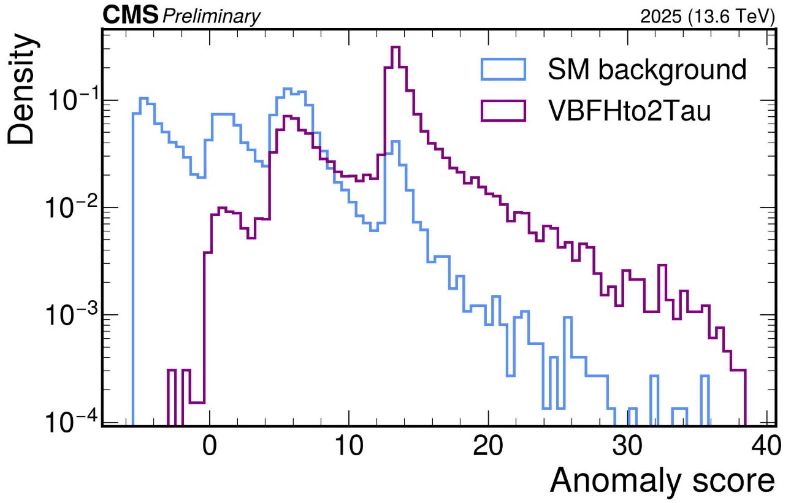


Model Performance: Anomaly Score

An **anomaly score** is a single numerical value that represents how "unusual" an event is compared to the baseline data the model was trained on.

To define the anomaly score:

- Fit a **Gaussian Mixture Model** (**GMM**) on the SM background
- Compute the likelihood of each event under the SM background
- Events farther from the SM distribution have a lower likelihood and are more likely to be anomalous.



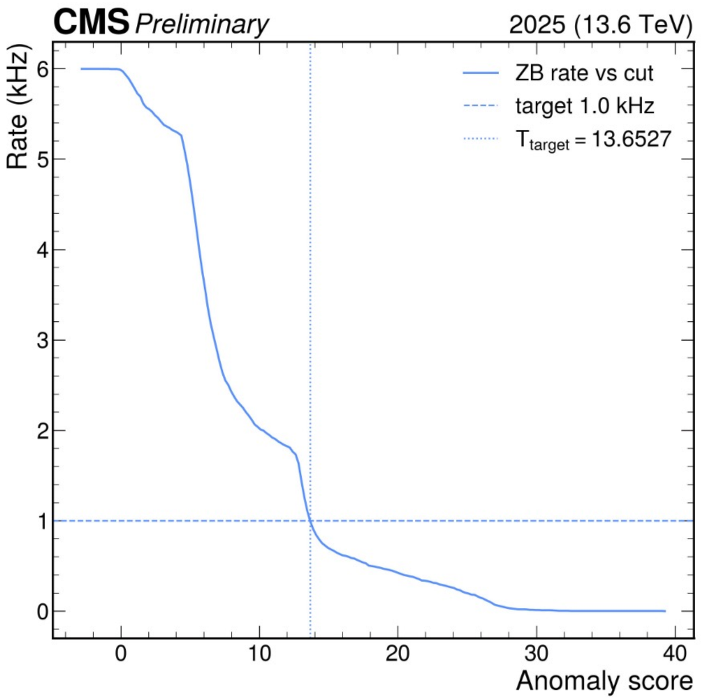
The anomaly score is defined as the negative log-likelihood from the GMM:

$$\mathcal{A}(z) = -\log p(z)$$

Model Performance: Anomaly Score Threshold

To feed the most anomalous events into the anomaly detection parking stream, the rate should be reduced from 6 kHz to 1 kHz.

- Plot the HLT output rate vs anomaly score
- The target threshold is chosen such that rate = 1 kHz
- The anomaly score values $T_{target} = 13.6527$ is chosen.



Model Performance: Signal Purity

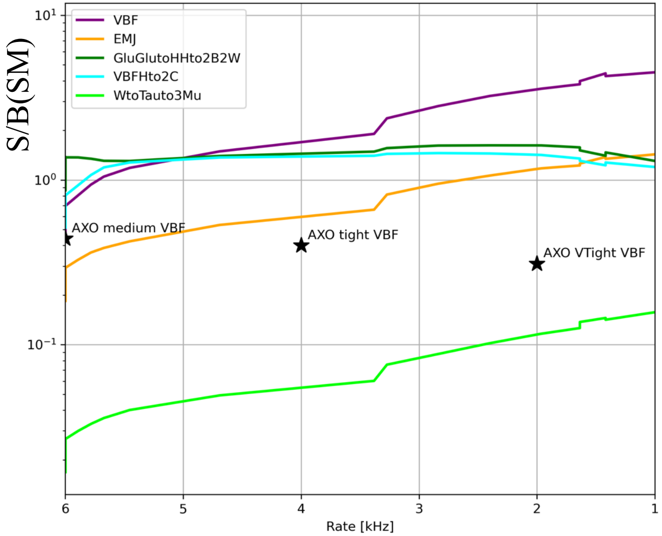
The purity study is done by computing the S/B ratio using the HLT rate.

- Background (B): Standard Model background
- Signal (S): the anomalous signal.

The purity for most of the samples increases as the rate decreases!
The opposite trend is observed with AXOL1TL working points!

Timing Studies completed:
Total Latency: ~4ms/event

Impact on HLT throughput:
With AD@HLT: 477.7 ± 0.4
without: 481.6 ± 1.8



To evaluate the intrinsic performance of the model architecture, all events are treated with equal weight, and no physics-motivated reweighting (such as cross-section normalization) was applied.

Conclusion & Future plans:

✓ Implement a second layer of anomaly detection in the CMS High-Level Trigger.

✓ Train the model on the SM samples and evaluate it with anomalous samples.

✓ Compute the anomaly score with respect to the SM cluster.

✓ The purity of the signals increases as the rate decreases.

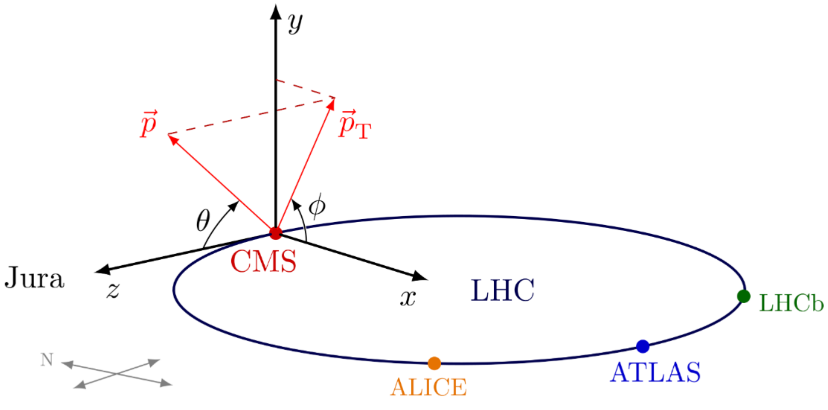
Some future ideas:

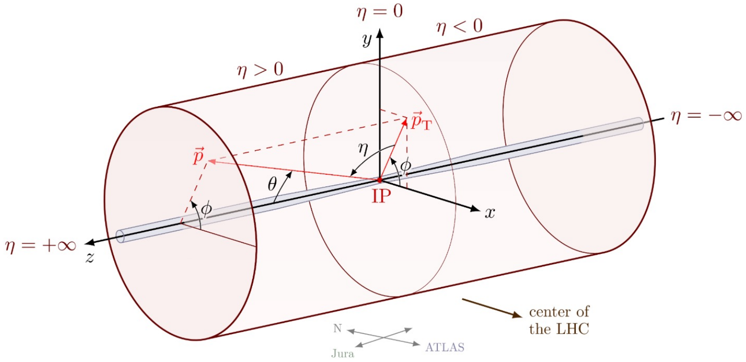
- Test the framework with the Level-1 trigger objects for Phase-II.
    - Train and test with PUPPI and PF candidate objects
- Simplify the architecture to reduce complexity → suitable for hardware implementation
    - Train the model with a Multilayer Perceptron (MLP) architecture

Backups

CMS Coordinate system

- p_T: transverse momentum is the component of a particle's momentum measured perpendicular to the beam axis in a collider experiment.
- φ is the azimuthal angle measured around the beam axis in the transverse plane, ranging from −π to +π.
- d_{xy} : Track displacement is the distance of closest approach of a particle's track to the primary vertex, measured in the transverse (x–y) plane — i.e., perpendicular to the beam axis. It tells you how far "off-centre" a track is from where the original proton-proton collision happened.
- d_{xy}sig: d_{xy} significance is simply the ratio dxy/σ(dxy), where σ(dxy) is the measurement uncertainty on dxy

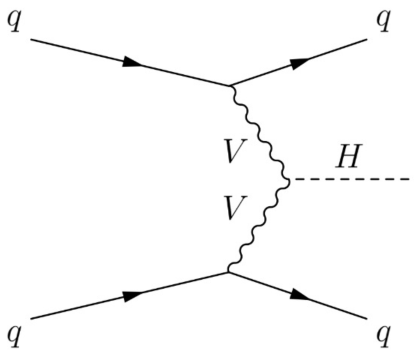




Signal Definitions

- **Vector-Boson-Fusion(VBF):** A quark from each of the incoming LHC protons radiates off a heavy vector boson (V), which are either W or Z bosons. These bosons interact or “fuse” to produce a particle, such as a Higgs boson. The initial quarks that first radiated the vector bosons are deflected only slightly and travel roughly along their initial directions. They are then detected as particle "jets" in the different hemispheres of the detector.

- **Emerging Jets (EMJ):** A signature from "Dark QCD" models where dark quarks decay into Standard Model particles after traveling a finite distance, resulting in jets with multiple displaced vertices and high-impact parameter tracks.



Model Performance: Anomaly Score Threshold

To feed events into the anomaly detection parking stream, the rate should be reduced from 6 kHz to 1 kHz.
The Zero Bias (ZB) sample is used to translate an anomaly-score cut into an expected HLT output rate.

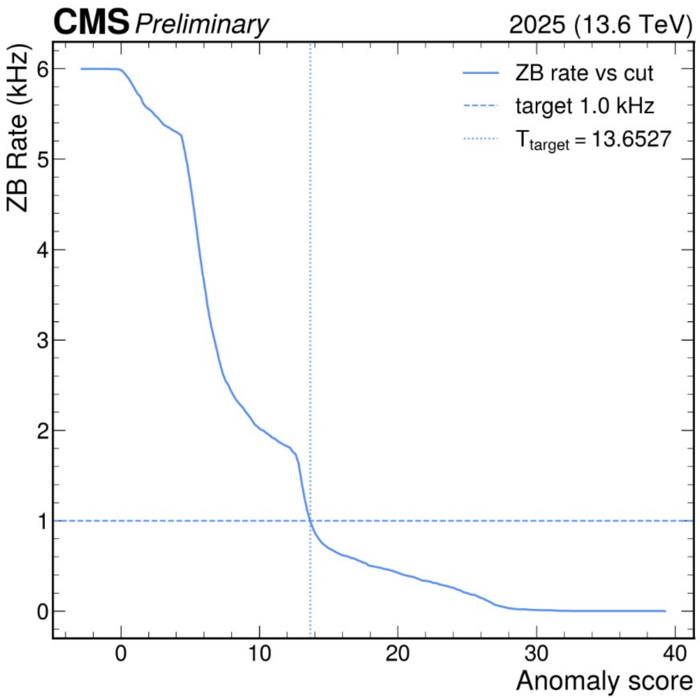
- Start to scan the possible anomaly score (t_{score}) threshold from the AXOMedium baseline rate of 6 kHz
- For each t_{score}, the events with anomaly score > t_{score} are being kept and measuring the surviving fraction of ZB events by:

$$\epsilon(t_{score}) = \frac{N_{ZB}(score > t_{score})}{N_{ZB,total}}$$

- converted into a rate by:

$$R(t_{score}) = 6\ \text{kHz} \times \epsilon(t_{score})$$

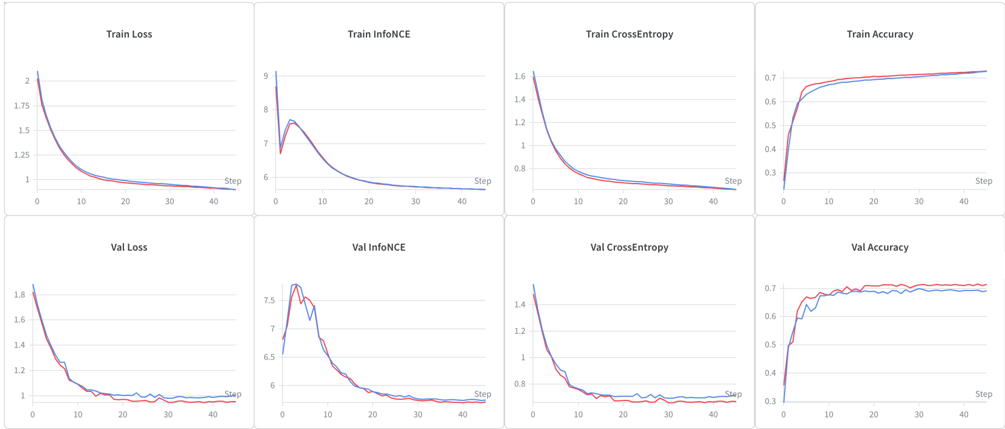
- The target threshold is chosen such that rate = 1 kHz (keeping 1/6 of the ZB events after anomaly score cut)



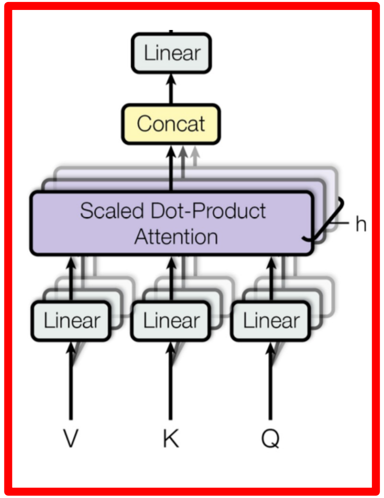
The Model Architecture

The model was initially based on a Transformer. To reduced computation resources, migrated to the Linformer Encoder architecture [4] while maintaining the physics performance comparable to the full Transformer.

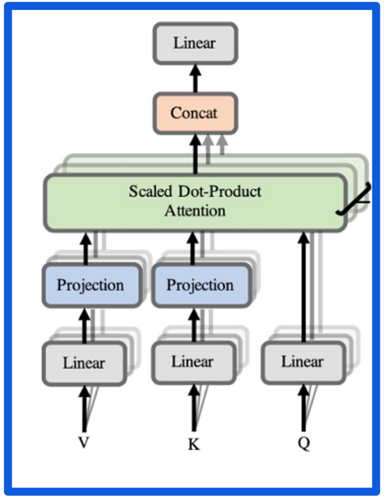
Train with Transformer
Train with Linformer



Transformer architecture



Linformer Architecture



The Model Architecture

A mix of the objective of InfoNCE loss and cross entropy was used:

$$\mathcal{L} = \beta \cdot \mathcal{L}_{\mathrm{NCE}} + \mathcal{L}_{\mathrm{CE}}$$

Where InfoNCE [5] is shaping the latent space, and cross entropy draws the decision boundaries.

The timing study was conducted by integrating the model into CMSSW. The model introduces a total HLT latency of approximately 4 ms per event. Consequently, the HLT throughput decreases from 481.6 ± 1.8 to 477.7 ± 0.4 events per second after implementing this second layer of anomaly detection.

Hyperparameters

Parameter	Value
Embed Size	128
Latent Dim	6
Proj Dim	6
Batch Size	256
Num Layer	4
Num Heads	8
Linear Dim	16

Datasets and Parameters

It is trained on a diverse set of simulated SM process, including: Drell-Yan, QCD, $t\bar{t}$, WJets

PF candidates and their kinematic features (p_T ,η, φ, d_{xy}, d_{xy}sig, particle Id) are used for object selections. A selection of 50k events per sample and 500 objects per class was imposed.

Training sample	description
DY	Decay mode: DY→ 2 leptons
QCD	p_T bin: 15 to 7000
$t\bar{t}$	Decay mode: hadronic + lepton
WJets	Decay mode: WJets → lepton + ν + jets

All Samples used LHC 2025 Run conditions

Model Performance: Signal Efficiency

To quantify the signal acceptance as a function of the ZB rate, the anomaly score values t_{score} is scanned. The signal efficiency is defined relative to the pre-AXOMedium signal sample.

S_0 is taken as the number of signal events **before** applying AXOMedium, while S(t_{score}) is taken as the number of signal events that **pass AXOMedium** and also satisfy the anomaly score requirement (e.g. score > t_{score}). Therefore:

$$\frac{S(t_{score})}{S_0} = \frac{N_{sig}(\text{pass AXOMedium} \cap \text{score} > t_{score})}{N_{sig}(\text{before AXOMedium})}$$

For each t_{score} , the corresponding output rate is estimated using the Zero Bias sample:

$$R(t_{score}) = R_0 \times \frac{N_{ZB}(\text{score} > t_{score})}{N_{ZB,total}}$$

Where R_0 denotes the baseline rate after AXOMedium (6 kHz).

Model Performance: Signal Efficiency

Plotting signal efficiency shows that, for the same output rate, higher signal efficiency is retained with the anomaly score cut than with fixed AXOL1TL working points alone (AXOMedium/Tight/VTight). The “raw AXO” points are obtained by counting how many signal events remain after each AXOL1TL working point without applying the anomaly score selection.