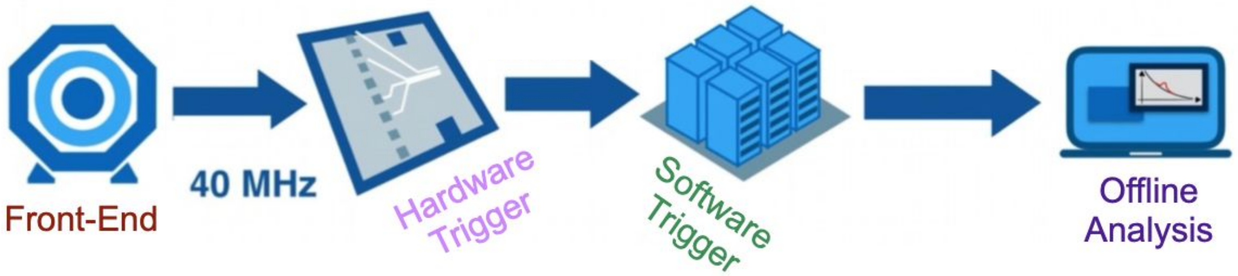
GELATO: a Generic Event-Level Anomalous Trigger Option for ATLAS

Max Cohen on behalf of ATLAS

Introduction

- Anomaly Detection (AD) has continued to gain popularity in the ATLAS experiment in the search for new physics
- It remains a possibility that such new physics evades the standard set of triggers



- In parallel, extensive developments in *fast* ML tools have made additional real-world applications possible
- We introduce GELATO: AD triggers in both hardware (L1) and software (HLT) triggers, providing a unique opportunity to enhance detection efficiency for unknown signatures

GELATO Trigger Pipeline

40 MHz →

GELATO L1

- VAE GAN architecture
- 44 input features: (p_T, η, φ) from 15 objects (6 jets, 4 taus, 4 muons, MET)
- Quantized, pruned, and chopped to fit in 25 ns inference time
- Commissioned to run on L1 backend FPGA
- Uses 26k out of ~700k active LUTs, 3 out of ~1.3k DSPs
- Two WPs: **T** (physics) and **L** (control) based on anomaly score Threshold

1 kHz unique (target) →

GELATO HLT

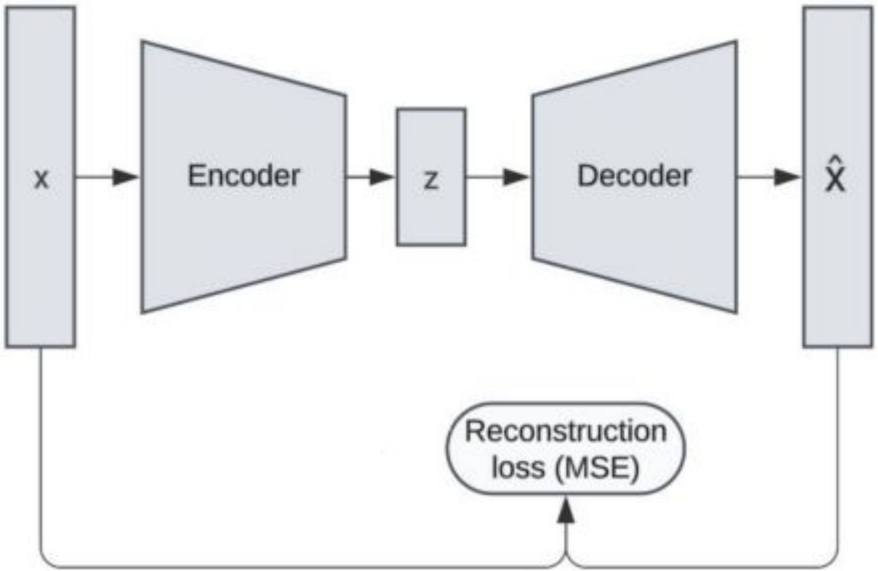
- AE architecture
- 47 input features: (p_T, η, φ) from 16 objects (6 jets, 3 electrons, 3 muons, 3 photons, MET)
- Only uses jets above 50 GeV, electrons, muons, photons above 30 GeV
- Only runs on events passing GELATO L1
- Three WPs: **T** (fallback), **M** (physics), **L** (control) based on anomaly score threshold

10 Hz unique (target) →

Implementation

The HLT network uses the Autoencoder (AE) architecture with mean squared error (MSE) loss

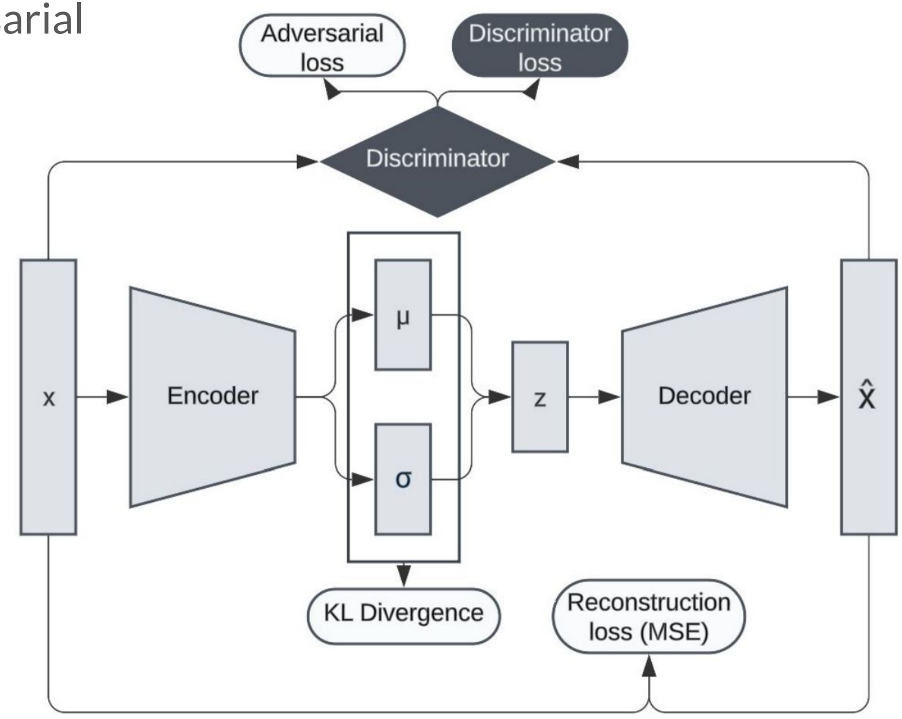
- Loss = masked MSE
  - Like standard MSE, but does not take zero values into account
- AD Score = masked MSE
- Layer sizes = 47 → 100 → 100 → 64 → 32 → 4



Implementation

The L1 network uses a VAE Generative Adversarial Network (GAN) architecture

- Discriminator attempts to distinguish inputs and outputs
- VAE attempts to trick the discriminator (adversarial loss)
- Loss_VAE = MSE + β*KL + γ*ADV
- AD Score = clipped KL
    - Clipped KL = $\mu^2\cancel{-1+\sigma^2-\log(\sigma^2)}$
- Layer sizes: 44 → 32 → 16 → 3

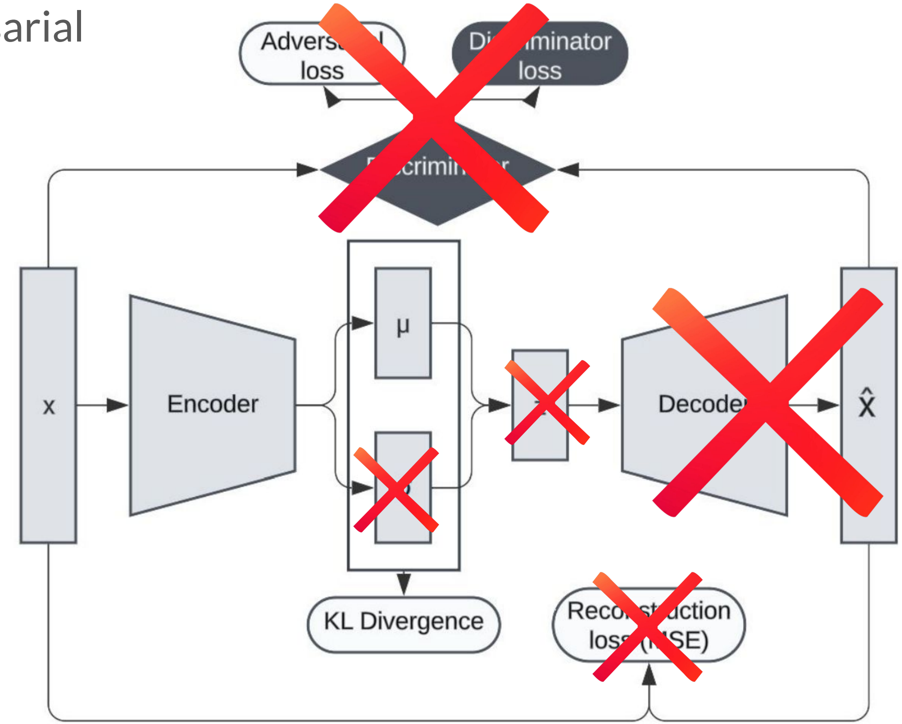


During Training

Implementation

The L1 network uses a VAE Generative Adversarial Network (GAN) architecture

- Discriminator attempts to distinguish inputs and outputs
- VAE attempts to trick the discriminator (adversarial loss)
- Loss_VAE = MSE + β*KL + γ*ADV
- AD Score = clipped KL
  - Clipped KL = $\cancel{\mu^2-1+\sigma^2-\log(\sigma^2)}$
- Layer sizes: 44 → 32 → 16 → 3



When Implemented [1]

Training and Testing Data

Both the L1 and HLT networks are trained using Enhanced Bias (EB) data:

- EB data is collected via a minimal set of triggers spanning all energies with known prescales
- Event weights are applied to correct for these prescales and restore an unbiased spectrum
- EB therefore retains an unbiased spectrum for estimating rates while also containing more rare events to help reduce uncertainties

Evaluation datasets:

- $HNL \to e^{\pm}\,\mu^{\mp}\,\nu,\ m = 7.5\,\mathrm{GeV},\ c\tau = 1\,\mathrm{mm}$
- $HAHM$ (ggF): $h \to Z_d\,Z_d \to 2\ell\,2\nu,\ m_{Z_d} = 28\,\mathrm{GeV}$
- ggF: $h \to \mathrm{SUEP} \to$ full-had
- VBF $h \to a\,a \to 4b,\ m_a = 55\,\mathrm{GeV},\ \tau_a = 1\,\mathrm{ns}$
- $Z \to \nu\bar{\nu}$ (b filter)
- ggF: $h \to a\,a \to 4b,\ m_a = 16\,\mathrm{GeV},\ \tau_a = 10\,\mathrm{ns}$
- VBF $h \to 4b$

- We use a broad selection of signal models to exhibit the model independent nature of AD.

- Prioritized cases where standard ATLAS triggers struggle

Simulation, Firmware, and Validation

[1] https://fastmachinelearning.org/hls4ml/

- GELATO L1 converted to HLS using hls4ml[1], synthesized with Vitis
- GELATO HLT network converted to ONNX model for use in HLT CPUs
- Trigger decisions must be reproducible in Athena (ATLAS software framework) for MC simulations + monitoring
    - Athena simulations were built for both algorithms
    - Exact agreement between hardware and simulation (up to sorting ambiguities) for GELATO L1!
    - Simulated HLT AD scores have been successfully verified against offline calculations

Commissioning Strategy in 2025

- A gradual ramp up commissioning sequence was decided with experts to ensure safe deployment
- Early LHC ramp-up: ☑
  - GELATO L1 decisions monitored but not used for actual trigger decision
- First stable beams: ☑
  - Use GELATO L1 for trigger decisions, but use a pass-through HLT chain (1 Hz prescaled rate)
  - No GELATO HLT activation yet
- Soon:
  - Enable GELATO L1 at a reduced rate (10 Hz). Activate GELATO HLT at 1 Hz to study trigger costs + rejection
- Soon:
  - Once everything is fully understood and operations look stable, enable GELATO chain at full capacity

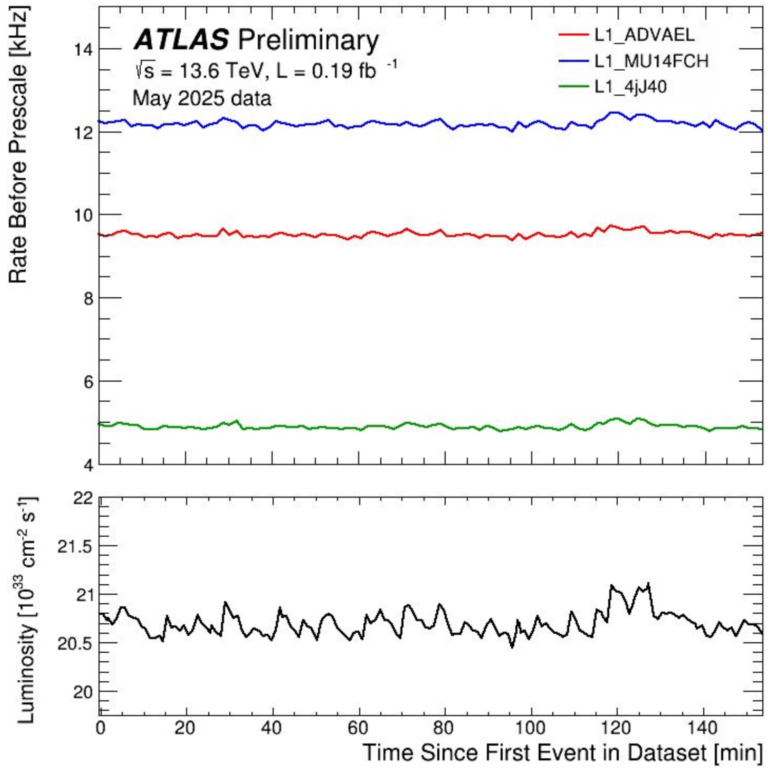
GELATO

Rate Monitoring

- Rates of GELATO L1 (L1_ADVAEL) compared to that of L1_MU14FCH (14 GeV single muon trigger) and L1_4jJ40 (40 GeV 4-jet trigger)

- All triggers exhibit a small uptick at 120 minutes, which can be attributed to a slight increase in the luminosity rather than an issue with GELATO itself

- The stable 9.5 kHz total rate is consistent with our calculations for 1 kHz unique rate



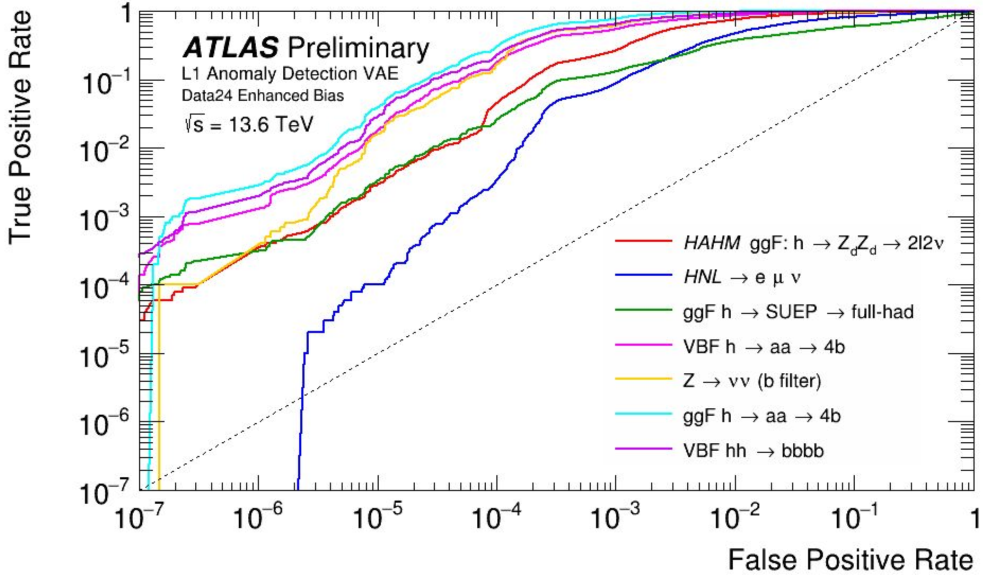
https://twiki.cern.ch/twiki/bin/view/AtlasPublic/CombinedTriggerPublicResults

Physics Performance – L1

- ROC curves for GELATO L1 for each MC test signal

- All MC signals lie above the random guess line, indicating discrimination power across a wide variety of signatures

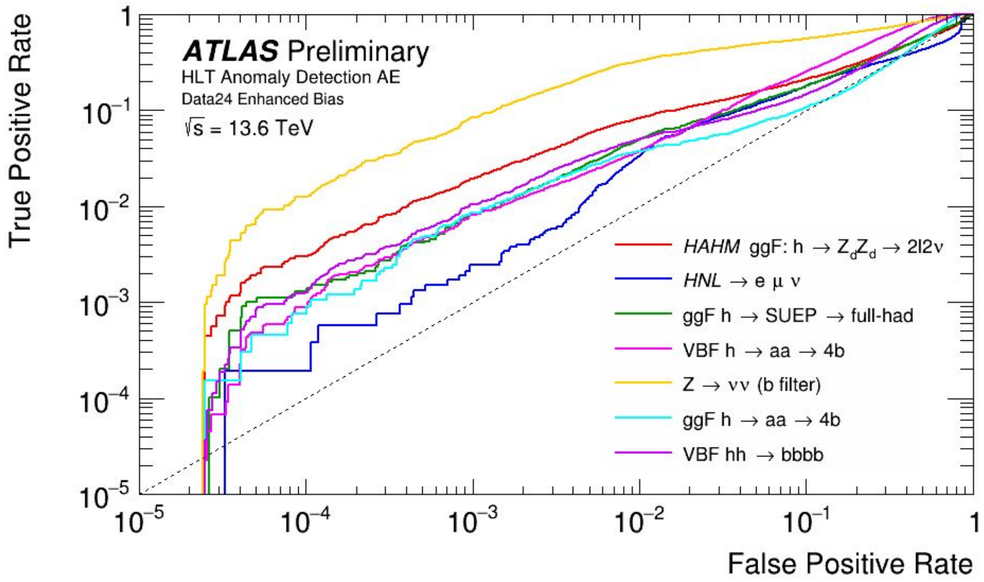
- In our calculations, 1 kHz corresponds to an FPR of 3.7×10^{-4}



https://twiki.cern.ch/twiki/bin/view/AtlasPublic/CombinedTriggerPublicResults

Physics Performance – HLT

- ROC curves for GELATO HLT for each MC test signal
- Only events passing GELATO L1 were used in this plot
- All MC signals lie above the random guess line, indicating discrimination power across a wide variety of signatures
- In our calculations, 10 Hz corresponds to an FPR of 9.5×10^{-3}



https://twiki.cern.ch/twiki/bin/view/AtlasPublic/CombinedTriggerPublicResults

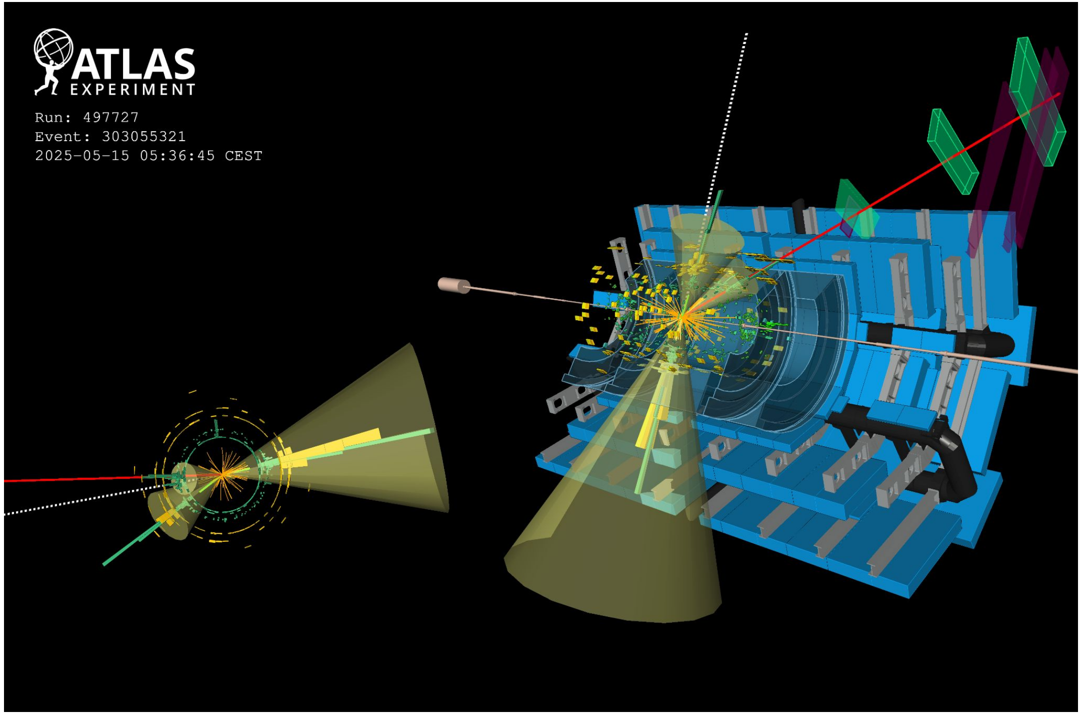
Overview | Algorithm Details | **Performance**

First Look at 2025 Data

- Event display of a GELATO L1 fired event with the highest GELATO HLT AD score during May 2025 at 13.6 TeV.

- Four jets above 50 GeV, three electrons and one muon above 30 GeV, and 215 GeV of MET

- We're excited to study unique events once the full chain becomes enabled!



https://twiki.cern.ch/twiki/bin/view/AtlasPublic/CombinedTriggerPublicResults

Summary and Next Steps

- Two Autoencoder-based AD trigger algorithms have been added to ATLAS for 2025 data taking at L1 and HLT
  - Very exciting to see the first use of AD triggers in ATLAS!!
- The GELATO chain is sensitive to a variety of physics signatures with different final states
- We are in the process of ramping up operations of these triggers to full capacity
  - Closely monitoring rates and performance
  - Things look good so far, and we're excited to move forward!
- We've started exploring ideas for the physics analysis strategy, more to come soon!

Backup

L1Topo

- 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