Anomaly Detection techniques for New Physics searches at the ATLAS experiment

Jennifer Curran

The University of Edinburgh

Overview

- **In this talk:** overview of an unconventional analysis technique that is beginning to be used more widely by the ATLAS experiment called **Anomaly Detection**

- **Covering the topics:**
  - What is Anomaly Detection and why is it an appealing choice of analysis method?
  - What machine learning models are used for Anomaly Detection?
  - How is Anomaly Detection being used in this analysis that targets di-lepton final states?

Motivation

- Despite extensive efforts by ATLAS and other experiments, we have not discovered any New Physics since the discovery of the Higgs boson

  → **This begs the question - are we looking in the correct places?**

- Traditional analysis methods - target a specific BSM model and perform searches in regions where the target signal dominates over the backgrounds (**S>>B**)

- Many phase spaces exist that have not yet been covered by ATLAS analyses
  - Often because they contain difficult to model backgrounds

How can we extend these BSM searches?

→ One approach is to use **Anomaly Detection** to perform signal **model-independent** searches to target phase spaces that are not attainable or have not yet been considered by traditional analysis methods

Motivation

Benefits

- **Anomaly Detection** searches therefore have several **benefits** over traditional analysis methods:

  - Model agnostic searches → search for New Physics without signal model assumptions
  - Not sensitive to mis-modelling of MC simulated backgrounds → data-driven background estimations
  - Allow for targeting of complicated phase spaces, e.g., containing high multiplicities
  - **Simultaneously target many BSM signatures**

Challenges

- However, there are also some **challenges** to face when performing such an analysis, for example:

  - It can be difficult to design an analysis that is completely unbiased to potential BSM signals
  - Designing robust validation methods can be more challenging than traditional searches
  - Performing a model-independent analysis → subject to the look-elsewhere effect

What is Anomaly Detection?

What is Anomaly Detection?

- Anomaly Detection refers to the use of **unsupervised machine learning (ML)** algorithms to identify rare and different events that could be attributed to new physics

What is Unsupervised ML?

- Refers to training the ML algorithm without labels

- 2 main types of ML used for New Physics searches: **Supervised ML** and **Unsupervised ML**

Supervised ML:

- **common** type of machine learning
- requires assumptions on the signal model being searched for
- model is trained to predict labels on data
- for example, to classify an event as a Signal or SM Background

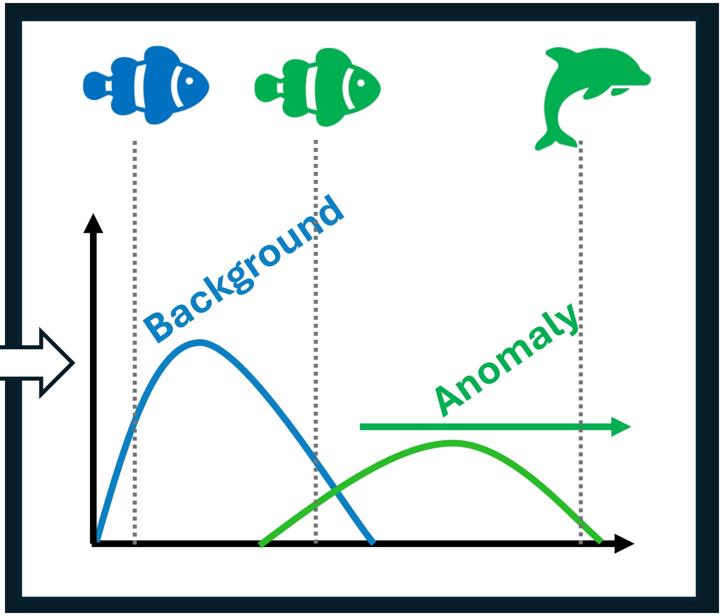
Unsupervised ML:

- used for **Anomaly Detection**
- train on Monte Carlo backgrounds only or SM dominated data
- no labels in the training data
- **NO knowledge of the signal required**
- explicitly or implicitly learn to estimate the probability density

What is Anomaly Detection?

Why use Unsupervised ML?

- These ML algorithms are trained on data (or MC simulated backgrounds)
- No knowledge of the potential BSM signal is required during training
  → **signal model independent**
- Search for anything that looks different to the SM data by defining an **Anomaly Score**
- Defining possible signal regions using the anomaly score allows for the simultaneous targeting of **many BSM signatures**

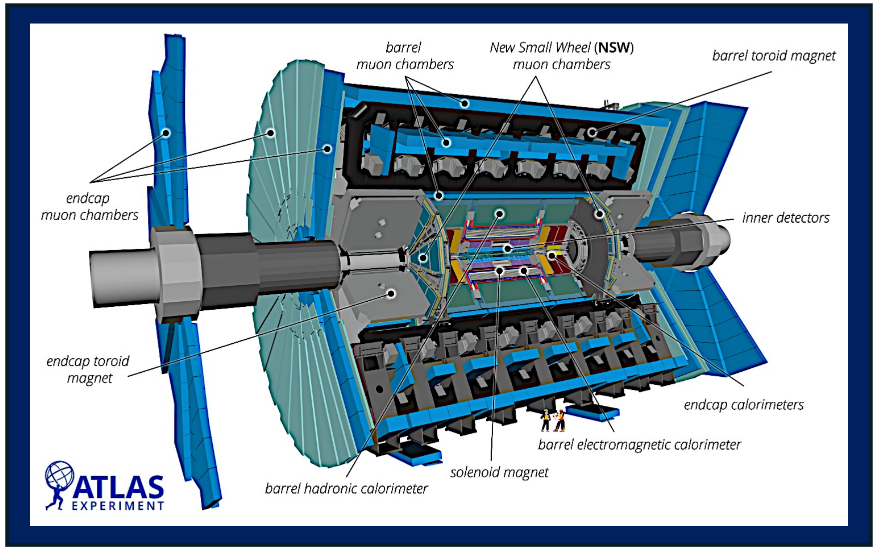


Anomaly Detection in di-lepton searches

- This analysis performs **Event-based** Anomaly Detection for outlier detection → Identify rare events that differ to the SM dominated data

- Targeting events with **≥ 2 leptons (electrons, muons) + X** in the ATLAS detector
  → record an event with 2 charged leptons and any other object e.g. jets, b-jets or photons

- **Why search for di-lepton BSM events in the ATLAS detector in the LHC?**
  - Good reconstruction of high energy leptons with the electromagnetic calorimeters and muon chambers
  - Forward-backward symmetry means missing energy in direction transverse to beam (${E_T}^{miss}$) can be determined
  - Allows us to target a wide variety of possible **BSM** signatures, for example:
    - ➢ High mass dilepton resonances: Z' E6 interpretation, RS gravitons etc.
    - ➢ Heavy Scalar (Higgs) to l^+l^- + ${E_T}^{miss}$: $S \rightarrow ZZ \rightarrow \bar{\nu}\nu l^+l^-$
    - ➢ SUSY stop pairs decaying to $\bar{b}bl^+l^-$+ ${E_T}^{miss}$



Anomaly Detection Methods

How do we use Anomaly Detection in the analysis?

Step 1:
Basic object selections

Select events that contain at least 2 leptons:

- Leptons = electrons or muons
- With leading 2 leptons, $p_T > 27$ GeV

Loose selections → target a wide range of BSM signatures

Step 2:
Train AD algorithms on 1% data

Train AD algorithms using event-based information from small percentage of the data including:

- 4-momenta of objects such as leptons, jets, b-jets, photons (E, p_T, η, ϕ)
- ${E_T}^{miss}, {E_T}^{miss}(\phi)$
- Angular separation of leptons (l): ΔR_{ll}, $\Delta\phi_{ll}$

Why training is performed on a small percentage of data:

- Remain unbiased/blinded to the bulk of the dataset
- New Physics events are expected to be rare
- If present in small % data
  → should be in tails of anomaly score
  → not statistically impactful in the Control Regions

Anomaly Detection Methods

How do we use Anomaly Detection in the analysis?

Step 1: Basic object selections

Select events that contain at least 2 leptons:

- Leptons = electrons or muons
- With leading 2 leptons, $p_T > 27$ GeV

Loose selections → target a wide range of BSM signatures

Step 2: Train AD algorithms on 1% data

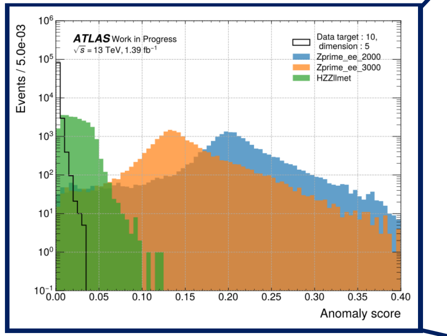
Train AD algorithms using event-based information from small percentage of the data including:

- 4-momenta of objects such as leptons, jets, b-jets, photons (E, p_T, η, ϕ)
- $E_T^{miss}, E_T^{miss}(\phi)$
- Angular separation of leptons (l): $\Delta R_{ll}, \Delta\phi_{ll}$

Step 3: Test performance on MC BSM signals

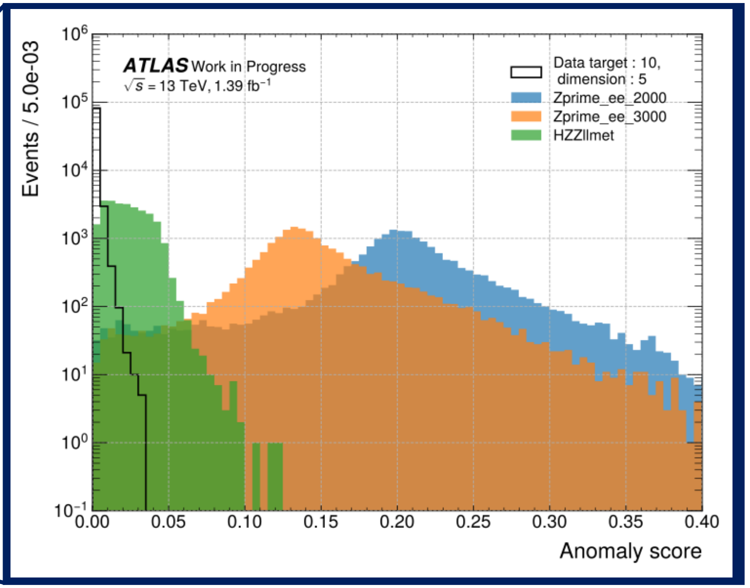
Evaluate the performance of the AD algorithm by testing on MC BSM signals

- E.g., Observe separation in the Anomaly score of the model



black = test data vs Z' 2TeV, Z' 3TeV, heavy Higgs 400GeV

- Trained on **1% Run2 (2015-2018) data**



Anomaly Detection Methods

How do we use Anomaly Detection in the analysis?

Step 1:
Basic object selections

Select events that contain at least 2 leptons:

- Leptons = electrons or muons
- With leading 2 leptons, p_T > 27 GeV

Loose selections → target a wide range of BSM signatures

Step 2:
Train AD algorithms on 1% data

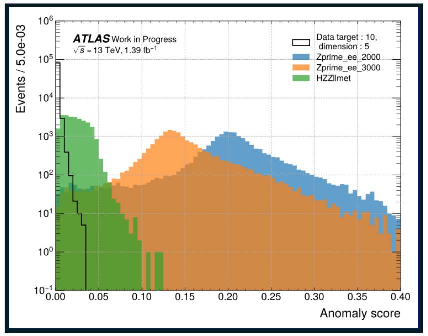
Train AD algorithms using event-based information from small percentage of the data including:

- 4-momenta of objects such as leptons, jets, b-jets, photons (E, p_T, η, ϕ)
- $E_T{}^{miss}, E_T{}^{miss}(\phi)$
- Angular separation of leptons (l): $\Delta R_{ll}, \Delta\phi_{ll}$

Step 3:
Test performance on MC BSM signals

Evaluate the performance of the AD algorithm by testing on MC BSM signals

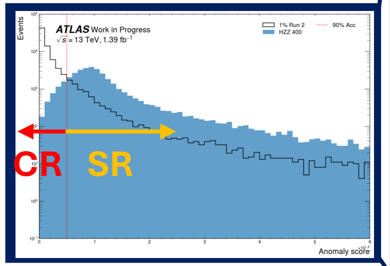
- E.g., Observe separation in the Anomaly score of the model

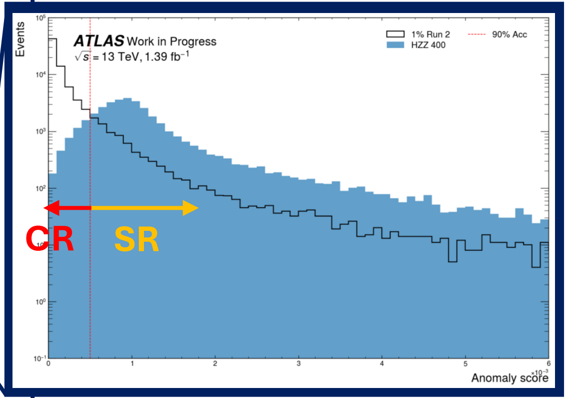


black = test data vs Z' 2TeV, Z' 3TeV, heavy Higgs 400GeV

Step 4:
Define SR/CR based on anomaly score

Use the anomaly score to help define model agnostic Signal and Control regions





e.g., based on 90% background acceptance vs rejection

Anomaly Detection Methods

How do we use Anomaly Detection in the analysis?

Step 1: Basic object selections

Select events that contain at least 2 leptons:

- Leptons = electrons or muons
- With leading 2 leptons, p_T > 27 GeV

Loose selections → target a wide range of BSM signatures

Step 2: Train AD algorithms on 1% data

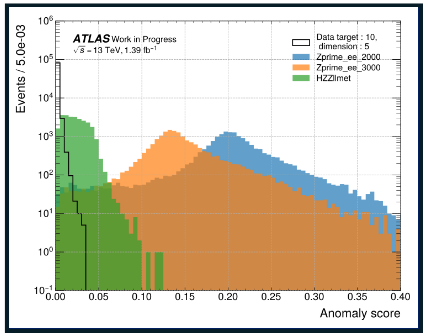
Train AD algorithms using event-based information from small percentage of the data including:

- 4-momenta of objects such as leptons, jets, b-jets, photons (E, p_T, η, ϕ)
- $E_T{}^{miss}, E_T{}^{miss}(\phi)$
- Angular separation of leptons (l): $\Delta R_{ll}, \Delta\phi_{ll}$

Step 3: Test performance on MC BSM signals

Evaluate the performance of the AD algorithm by testing on MC BSM signals

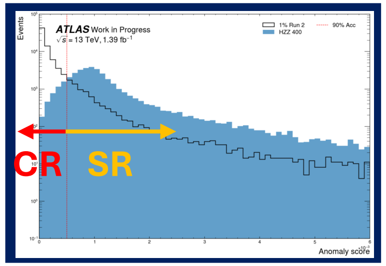
- E.g., Observe separation in the Anomaly score of the model



black = test data vs Z’ 2TeV, Z’ 3TeV, heavy Higgs 400GeV

Step 4: Define SR/CR based on anomaly score

Use the anomaly score to help define model agnostic Signal and Control regions



e.g., based on 90% background acceptance vs rejection

Step 5: Perform search for BSM physics

Perform a search for New Physics using the full Run 2 / Run 3 data in the SR of the chosen final state

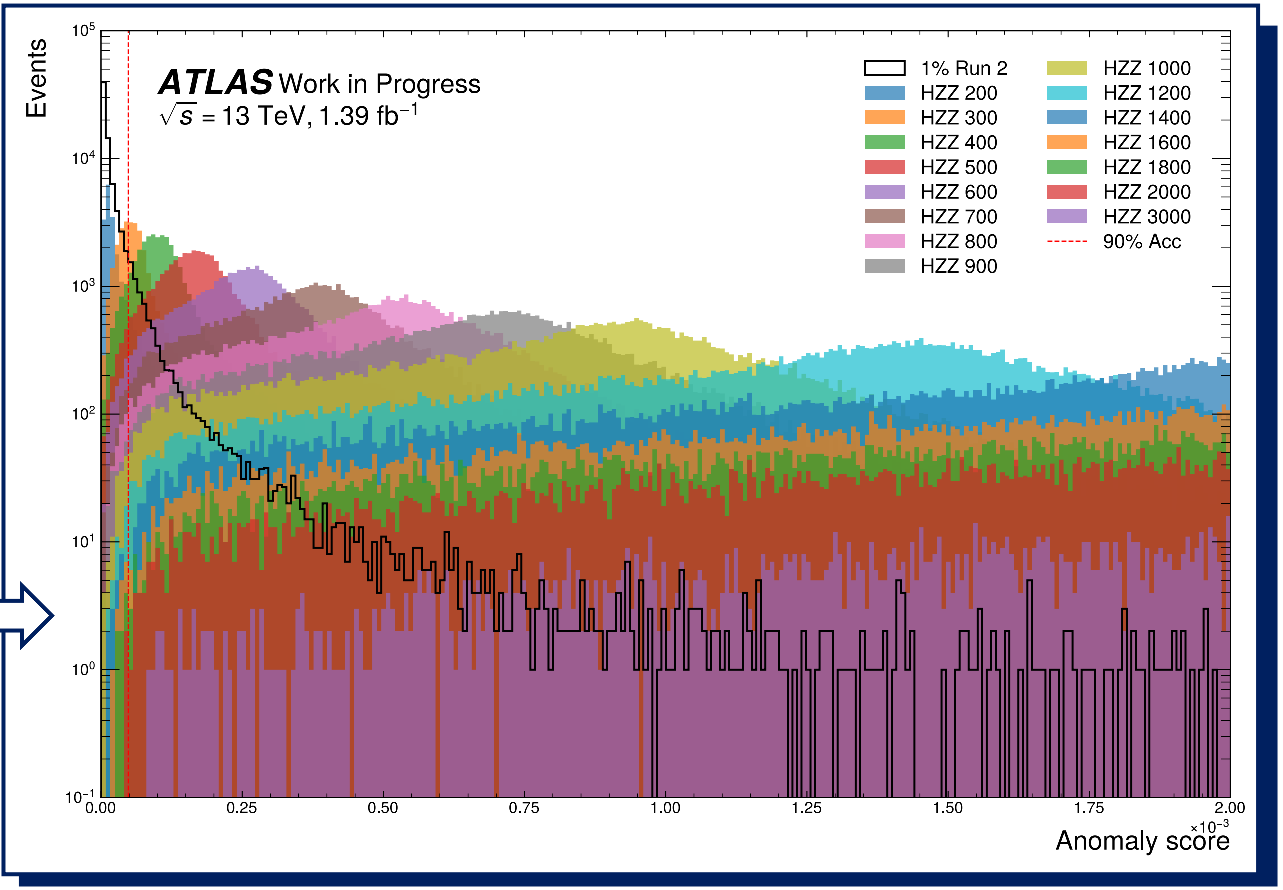
If no deviations are found: set model-dependent and model-independent limits

Anomaly Detection Algorithms

- Investigating different unsupervised ML algorithms to evaluate which is most effective at identifying a wide range of BSM signatures

One Example: the AutoEncoder (AE)

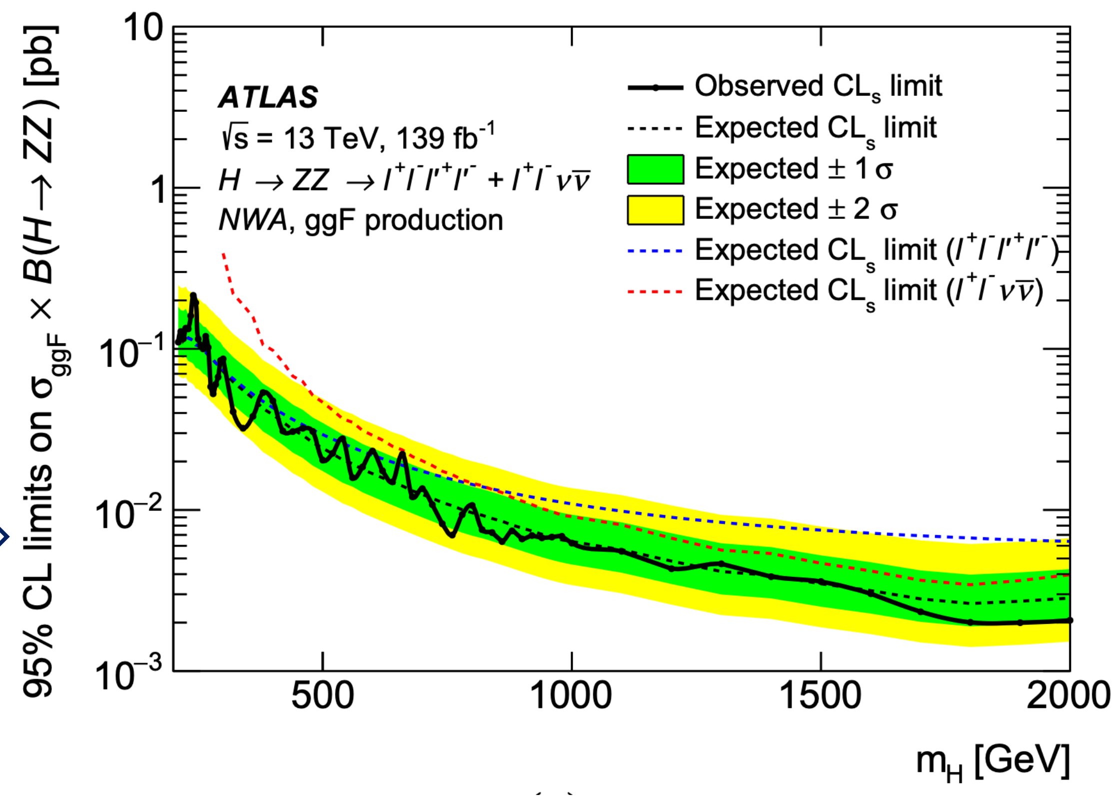
- Common Unsupervised ML
- Designed to automate the process of optimising the representation of the input feature space
- **AE** Trained on **1% Run2 (2015-2018) data**
- Example **Anomaly Score (MSE)** output comparing a small percentage of Run 2 test data to the BSM signal:
  - Heavy Scalar (Higgs) to 2 leptons + E_T^{miss}:
    - $S \to ZZ \to \bar{\nu}\nu l^+ l^-$
  - Mass range 200 GeV – 3 TeV



Anomaly Detection Algorithms

Example: AutoEncoder (AE)

- These results are currently being used to produce expected model-independent and model-dependent limits for this BSM signal for final states containing **2 leptons and $E_T{}^{miss}$**
- Example of the limits set for the **dedicated Run 2** analysis for the BSM signal – heavy Higgs boson to 4 leptons and 2 leptons + $E_T{}^{miss}$
- Aim is to produce comparable limits using the Anomaly Detection approach



Eur. Phys. J. C 81, 332 (2021)

Summary

- **Anomaly Detection (AD)** methods provide promising new model-independent analysis techniques

- Such methods can be sensitive to a broad scope of BSM signatures **simultaneously**

- In this analysis, we perform **event-based** anomaly detection, allowing for the exploitation of correlations in the kinematics of an event that have not necessarily been considered before

- As we are targeting **di-lepton** final states, the AD approach could be sensitive to a variety of BSM signatures, including heavy di-lepton resonances, such as Z', and the heavy Higgs boson

- By producing model-independent limits in Signal Regions defined by these AD scores, we extend the phase space of traditional searches to potentially uncovered regions

Back-Up

Anomaly Detection in ATLAS

- Several Anomaly Detection (AD) searches have already been published by the ATLAS collaboration

- For example, jet-level AD was used to set **upper limits on the production cross-section**: $\sigma(pp \to Y \to HX \to \bar{b}b\bar{q}q)$ of a high mass particle (Y) decaying to SM Higgs boson (H) and a new particle X of 1.5 < m_Y < 6 TeV & 65 < m_X 3000 GeV [DOI: 10.1103/PhysRevD.108.052009]

- Additionally, event-based AD was performed in the jet+X final state and upper limits on the production cross section of generic resonant signals of varying mass were produced [DOI: 10.1103/PhysRevLett.132.081801]

- Further weakly supervised AD searches have been performed, e.g., [DOI: 10.1103/PhysRevLett.125.131801]

- Many more ongoing

- In addition to targeting many BSM signatures, the applications and interpretations of AD are also vast

Event Selection

- Using **full Run 2** and **partial Run 3 (2022, 2023) data : ~196 fb^{-1}**
- **3 channels defined**: di-electron, di-muon and electron + muon

Channel	N lep	pT threshold	pT requirement	1% Run 2	1% Run 3 (2022/23)
Di-electron	$\geq$ 2 electrons	$>$ 27GeV	2 electrons highest pT leptons	391,781	149,296
Di-muon	$\geq$ 2 muons	$>$ 27GeV	2 muons highest pT leptons	495,291	152,398
El-mu	$\geq$ 1 muon and $\geq$ 1 electron	$>$ 27GeV	electron and muon are highest pT leptons	8772	3118

- Possibility to **combine** di-electron + di-muon into a single channel (electron + muon is kept separate due to background composition)

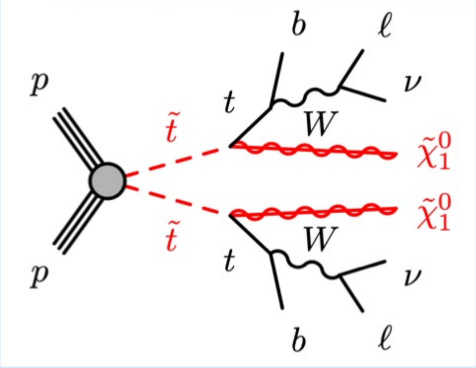
Event Selection

- Targeting ll+X final states → Initial focus on ll + MET (+X) final states

- Using mc20/mc23, main backgrounds under consideration:
  - **Zjets, $\bar{t}t$, Single Top, Diboson, Triboson, Wjets**

- 1% Run 2 / Run 3 used for training of ML algorithms: ~1.39fb^{-1}/0.56fb^{-1}

- Signals currently being used as benchmarks:
  - High mass dilepton resonances (ee, mumu) – **Z'** E6 interpretation, RS gravitons, MSSM Higgs etc.
  - Heavy Scalar (Higgs) to ll + $E_T{}^{miss}$ (ee, mumu) - **S→ZZ →llvv** (range of masses 200-2400GeV)
  - SUSY **stop pairs to bbll + MET** (emu)



Anomaly Detection Algorithms

What Anomaly Detection methods are used in the analysis?

- Investigating a number of different unsupervised ML algorithms to evaluate which is most effective at identifying a wide range of BSM signatures

- 3 main algorithms are currently under investigation:
  - **Autoencoder:** a common unsupervised ML algorithm that is designed to automate the process of optimising the representation of the input feature space. The differences between the input and reconstructed data are used to form an anomaly score.

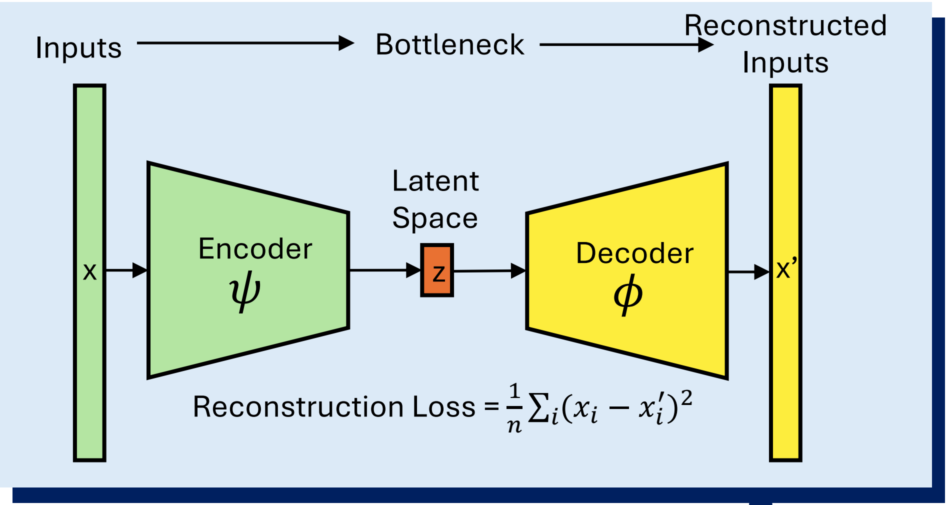
  - **DSVDD (Deep Set Vector Data Descriptor):** like the Autoencoder, the DSVDD learns a different representation of the input phase space that is more optimal for identifying anomalies. These models should be able to identify outliers from the SM data [arXiv:2106.10164].

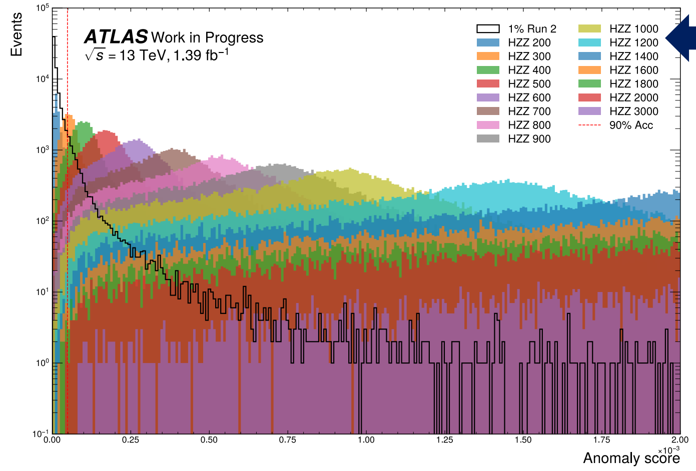
  - **Autoregressive Flow:** tries to explicitly learn the likelihood of an event which can then be used to form the anomaly score. Therefore, these models should be able to identify less probable or rarer events [arXiv:2106.10164].

AD Algorithms

AutoEncoder (AE)

- Common Unsupervised ML
- Designed to automate the process of optimising the representation of the input feature space
- Aims to learn a pair of functions: the encoder (ψ) and decoder ($\emptyset$) such that the **error** on the reconstructed data ($\sum_i \|x_i - \emptyset(\psi(x_i))\|^2$) is minimised (the **mean square error (MSE)**)
- Deciding if an event is anomalous:
    - If **MSE** is small → good reconstruction → likely **SM**
    - If **MSE** is large → poor reconstruction → could be **BSM**
      → **Investigate further**

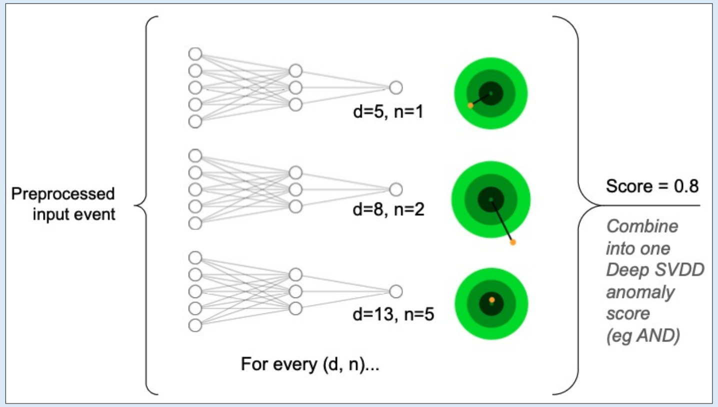




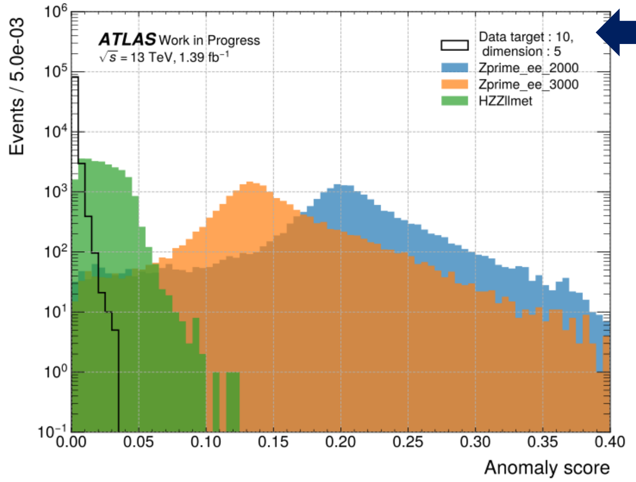
AD Algorithms

Deep Set Vector Data Descriptor (DSVDD)

- Similar to an AE
- Maps input data to a multidimensional point of a defined target value e.g., d=5, n= 1 -> (1, 1, 1, 1, 1).
- The latent space = a multidimensional space that can be thought of as a compressed representation of the input feature space that encodes meaningful information on this input space
- The anomaly score = the distance to the multidimensional point
- The SM data should lie within the defined multidimensional region
- Anomalous (**BSM**) data should fall **outside** this region



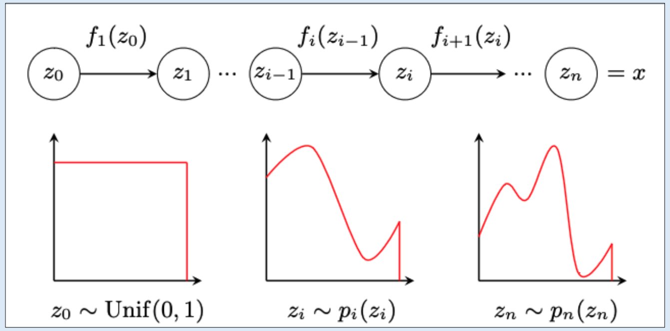
arXiv:2106.10164



AD Algorithms

Autoregressive Flow

- Attempts to evaluate the likelihood of each event and convert this to an anomaly score.
- These models start from a uniform prior distribution and try to determine a probability distribution for the known data (the SM) through transformations of parameterised variables.
- SM events that the model is trained on should have a high likelihood, whilst **BSM** events should have a **low** likelihood.



arXiv:2106.10164