Anomaly Detection for BSM searches

Jennifer Curran, University of Edinburgh

Why use Anomaly Detection?

- Many BSM searches perform **dedicated** searches that target specific signal models.

- Traditional searches optimise the event selection such that the separation between the signal and background is maximised.

- **BUT**… what if the signal model is not known? Then dedicated searches will miss them due to the **limited phase spaces** they consider.

- To extend the search to wider phase spaces, **model-independent** searches are conducted (do not depend on a specific signal model).

- **Model-independent** searches can be performed by e.g., searching for a localised excess ("bump") in the data compared to simulated SM backgrounds, but such methods have **limitations.**

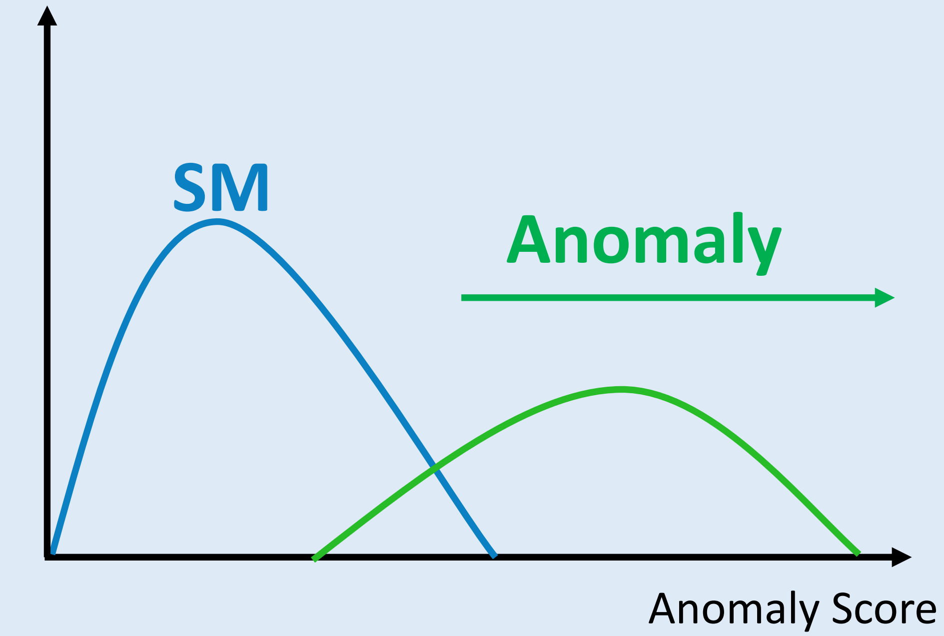
    - Reduction in ability to supress backgrounds when using less stringent event selection

    - The significance of any signals is reduced by the number of analysis channels (look-elsewhere effect)

    - Limited by the accuracy of the simulations

Why use Anomaly Detection?

- To overcome these limitations, we use **Anomaly Detection** techniques which are:
  - **model-independent**
  - Search for anything that looks different to the SM data by defining an **Anomaly score**
  - Use sophisticated analysis methods called **Unsupervised Machine Learning (ML)**
  - Extend sensitivity of searches and therefore the discovery reach



Unsupervised Machine Learning

Supervised ML:

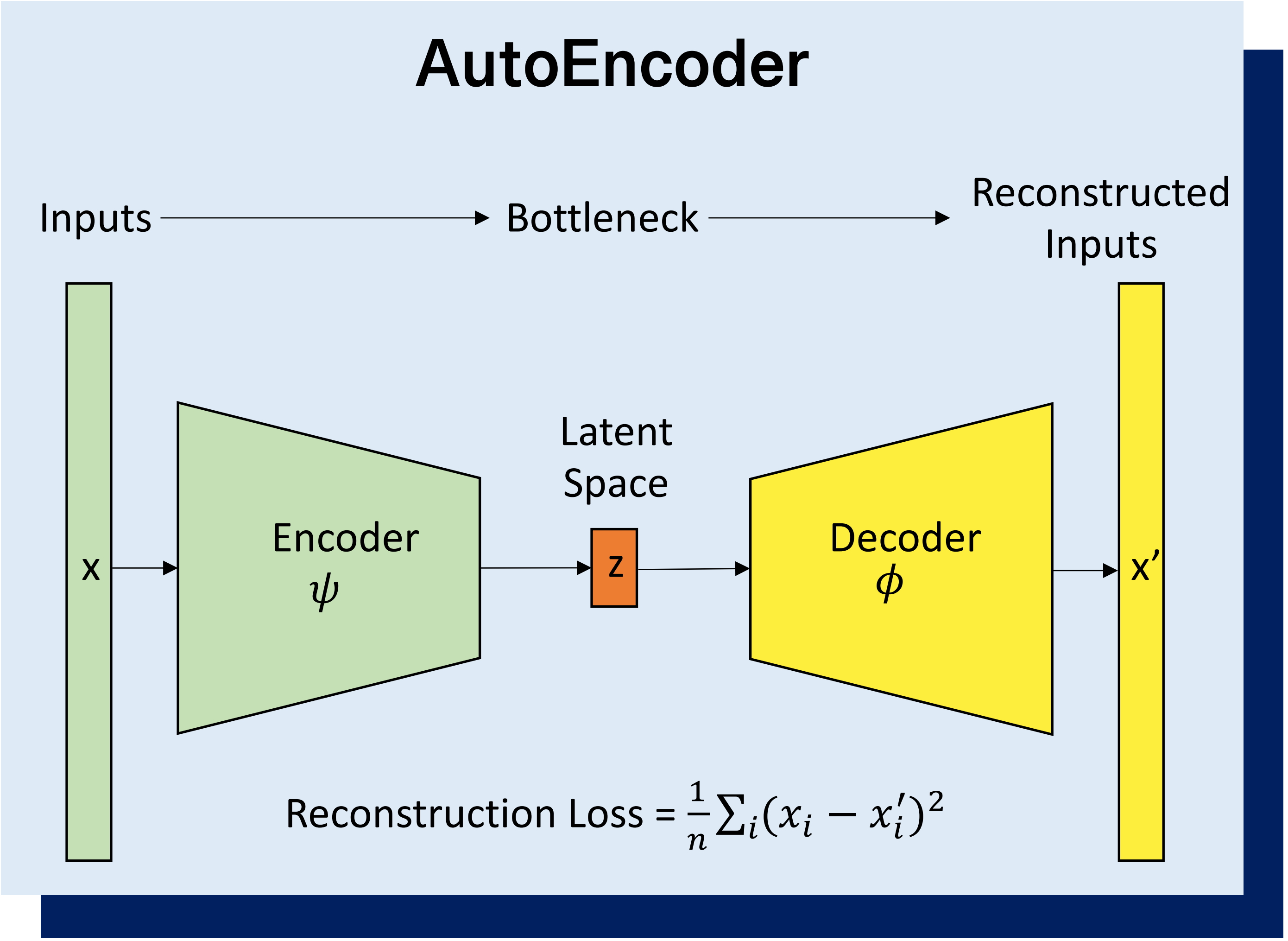
- This is the most common type of machine learning in which a model is trained to predict labels, e.g., Signal or Background.
- Used to optimise the separation between specific signal models and SM backgrounds.
- This requires knowledge of the signal model being searched for.

Unsupervised ML:

- This method of ML is used for Anomaly Detection.
- There are no labels in the training data and so knowledge of the signal being studied is not explicitly required.
- They are trained on Monte Carlo simulated backgrounds only or SM dominated data.
- The models explicitly or implicitly learn to estimate the probability density.

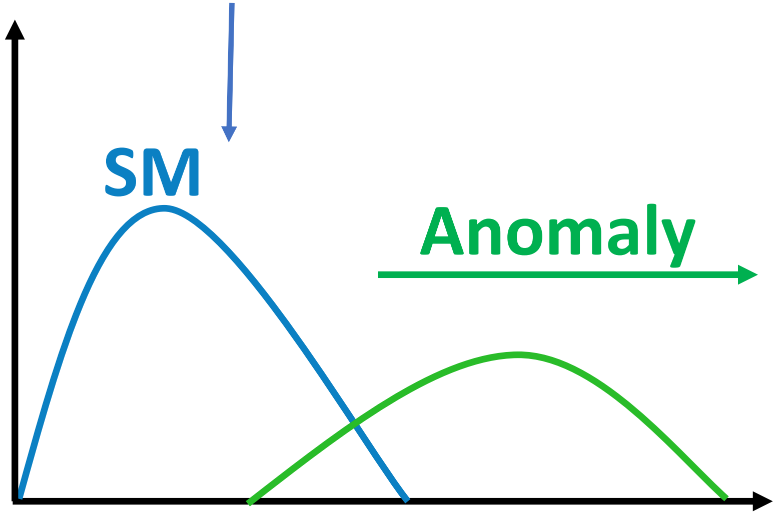
Unsupervised Machine Learning Models – AutoEncoder (AE)

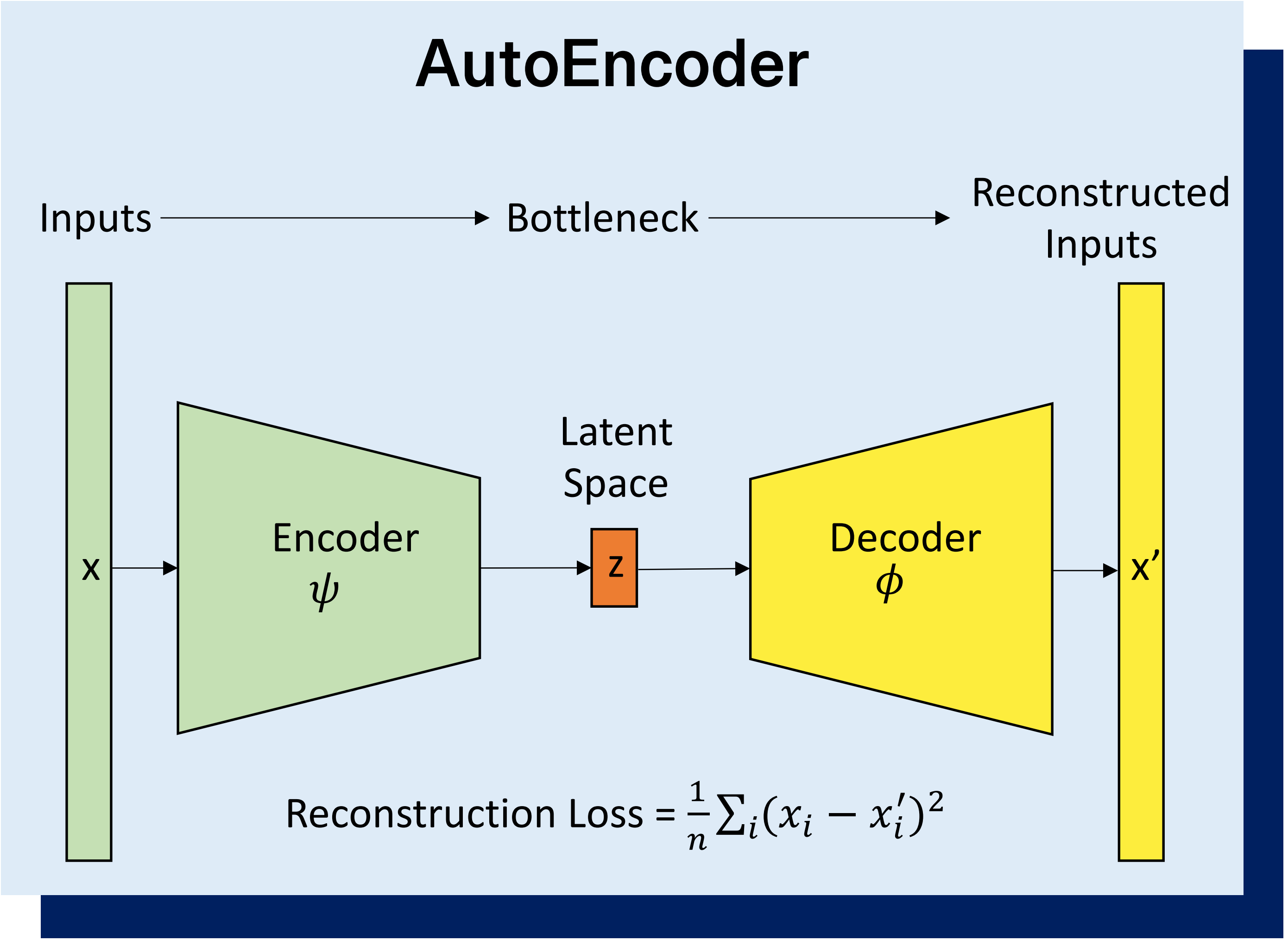
- An example of a common Unsupervised ML model is the: **AutoEncoder (AE)**.
- Designed to automate the process of optimising the representation of the input feature space.
- Aims to learn a pair of functions: the encoder (ψ) and decoder ($\emptyset$) such that the **error** on the reconstructed data ($\sum_i \|x_i - \emptyset(\psi(x_i))\|^2$) is minimised.



Unsupervised Machine Learning Models – AutoEncoder (AE)

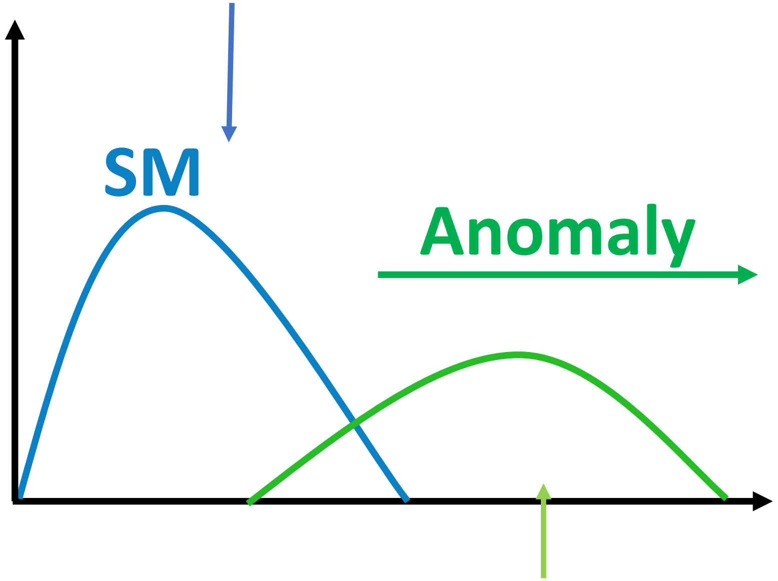
- If **x'** is close to **x** (**small loss**) : events likely to be common – i.e., SM background events on which the AE is trained.



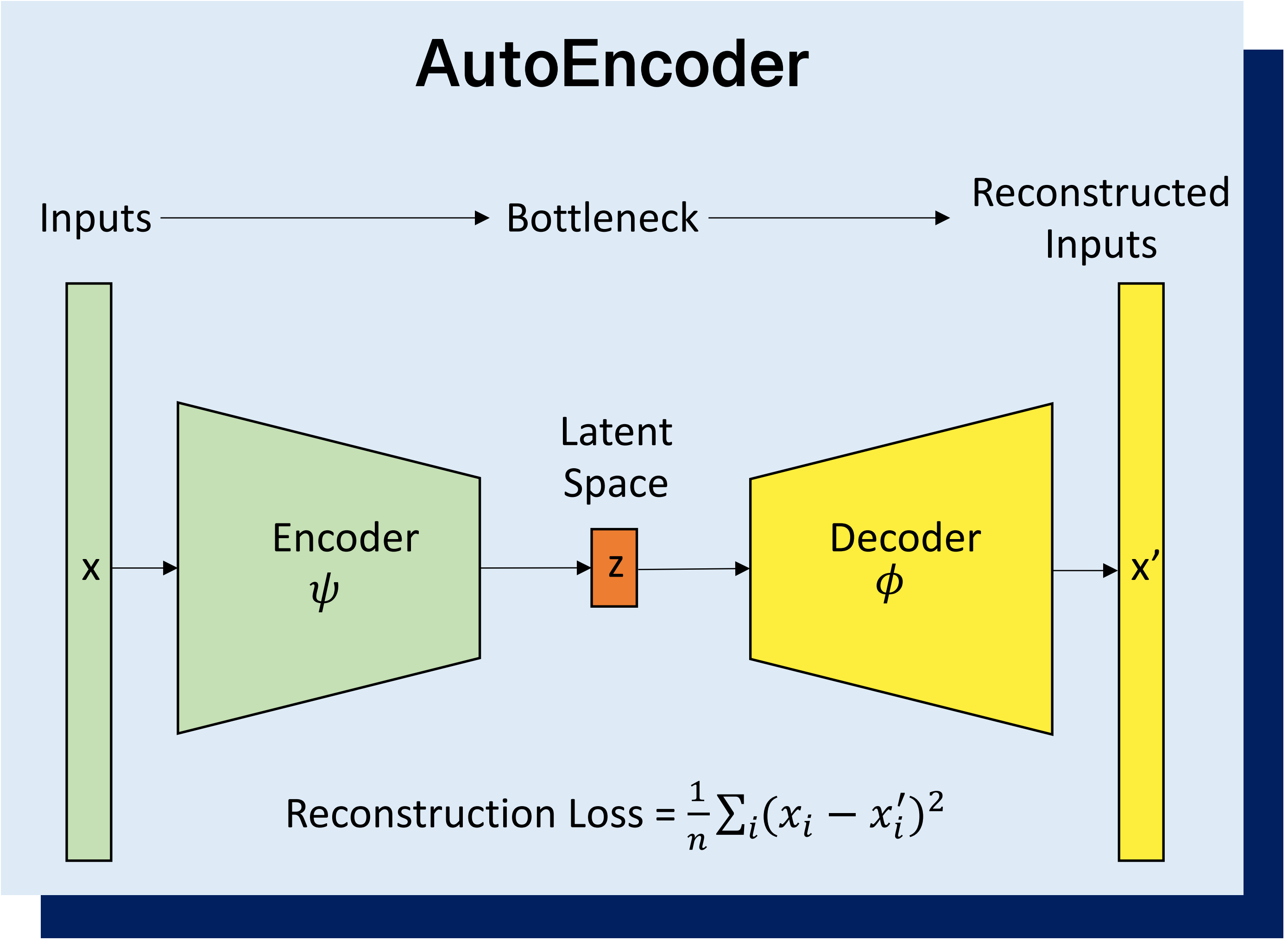


Unsupervised Machine Learning Models – AutoEncoder (AE)

- If **x'** is close to **x** (**small loss**) : events likely to be common – i.e., SM background events on which the AE is trained.



- If **x'** is different to **x** (**large loss**) : events should be less common – i.e., BSM signal events.



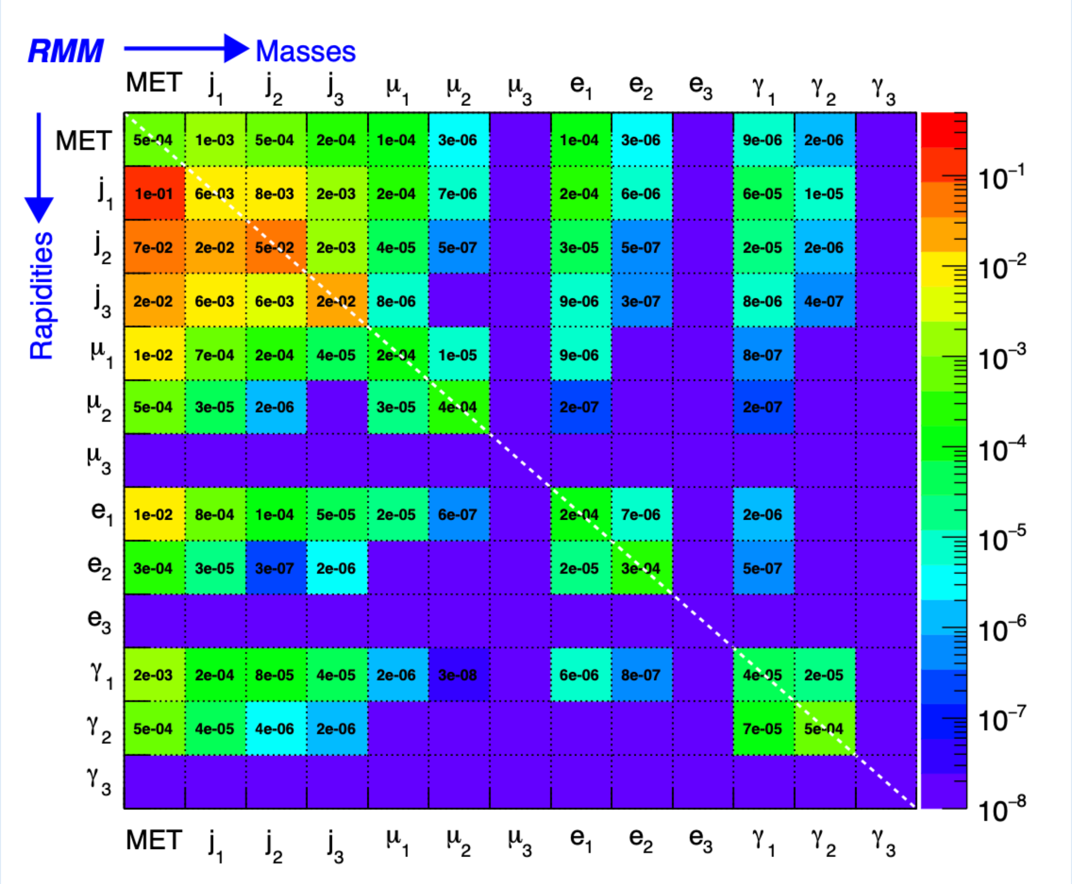
Anomaly Detection at the ATLAS experiment

Anomaly Detection at the ATLAS experiment

- A recent analysis at the ATLAS experiment applied Anomaly Detection algorithms on data collected during Run 2 (2015-2018) of the LHC (ATLAS, arXiv:2307.01612v1).
- The search was performed for any 2-body final states of the form jet+Y where jet was a light jet or b-jet and Y was allowed to be a lepton (electron or muon), photon or another light jet or b-jet.
- The Auto-Encoder was trained on 1% of Run 2 collision data (following preselection cuts).
- It is assumed that such a subset of the data would not contain a statistically significant number of BSM events and so can be assumed to represent SM backgrounds.
- The analysis was performed using high level kinematic information (in the form of the Rapidity Mass Matrix (**RMM**)) to train an Auto-Encoder.

Rapidity Mass Matrix (RMM)

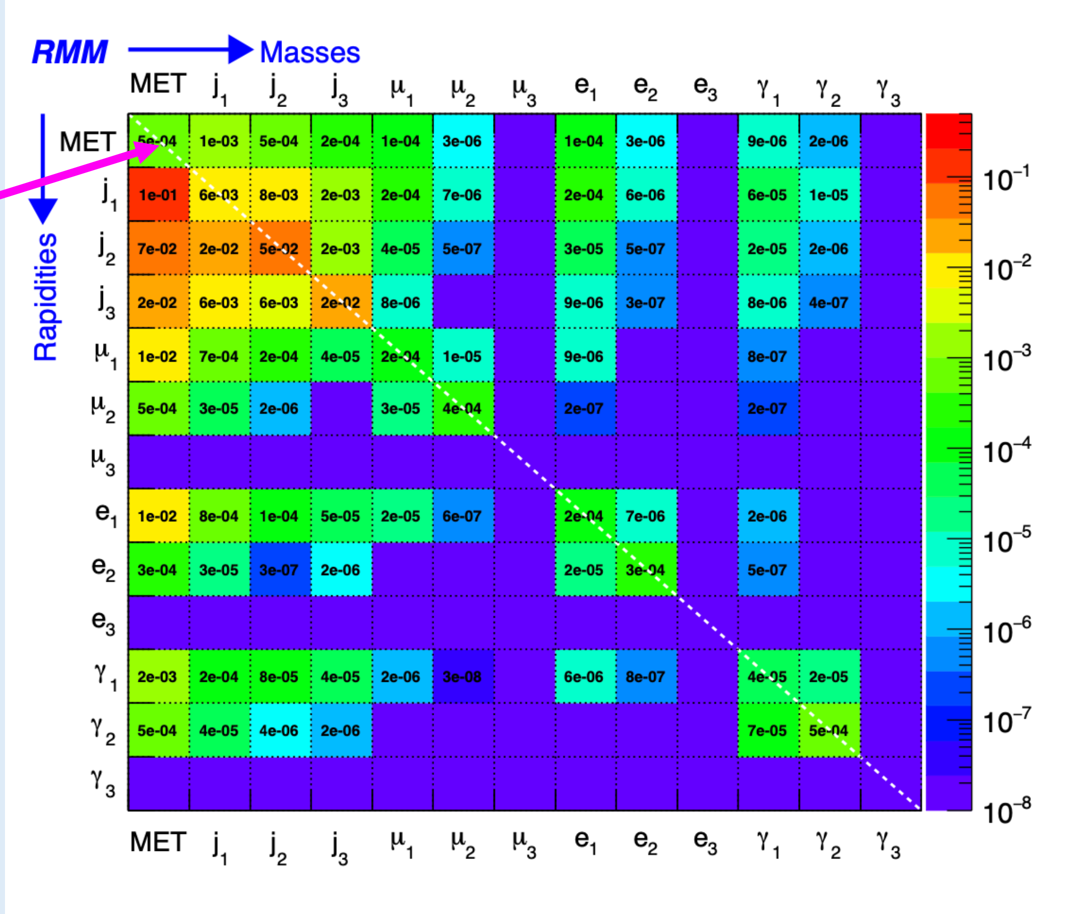
- An image-like representation of the kinematic information for the data in an event.
- Components are normalized by the center of mass energy – allows for comparisons between Run2 (13 TeV) and Run3 (13.6 TeV) results.
- Encodes information that could be useful for NP searches.



S. V. Chekanov, arXiv:1805.11650

Rapidity Mass Matrix (RMM)

- Encapsulates many variables that are favored for BSM searches.
- For example:
  - Missing Transverse Energy (MET)

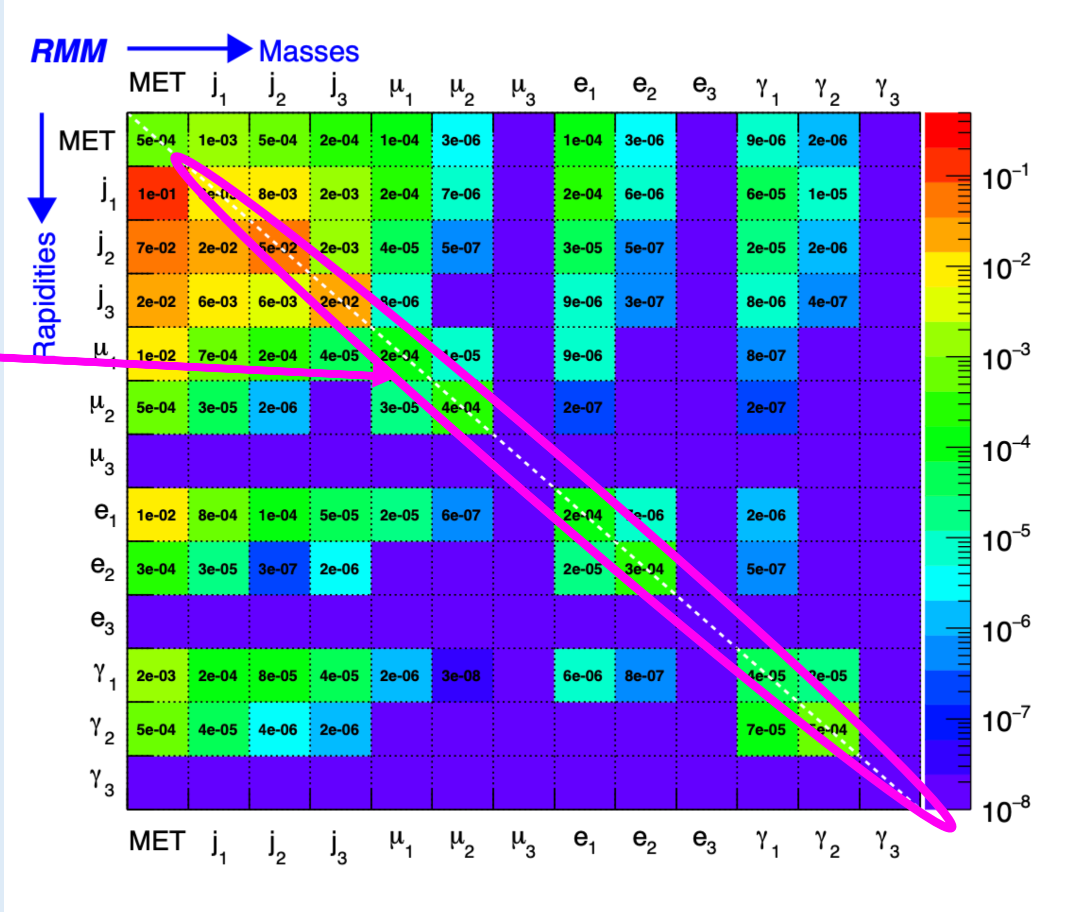


S. V. Chekanov, arXiv:1805.11650

Rapidity Mass Matrix (RMM)

- Encapsulates many variables that are favored for BSM searches

- For example:
  - Missing Transverse Energy (MET)
  - Transverse energies



S. V. Chekanov, arXiv:1805.11650

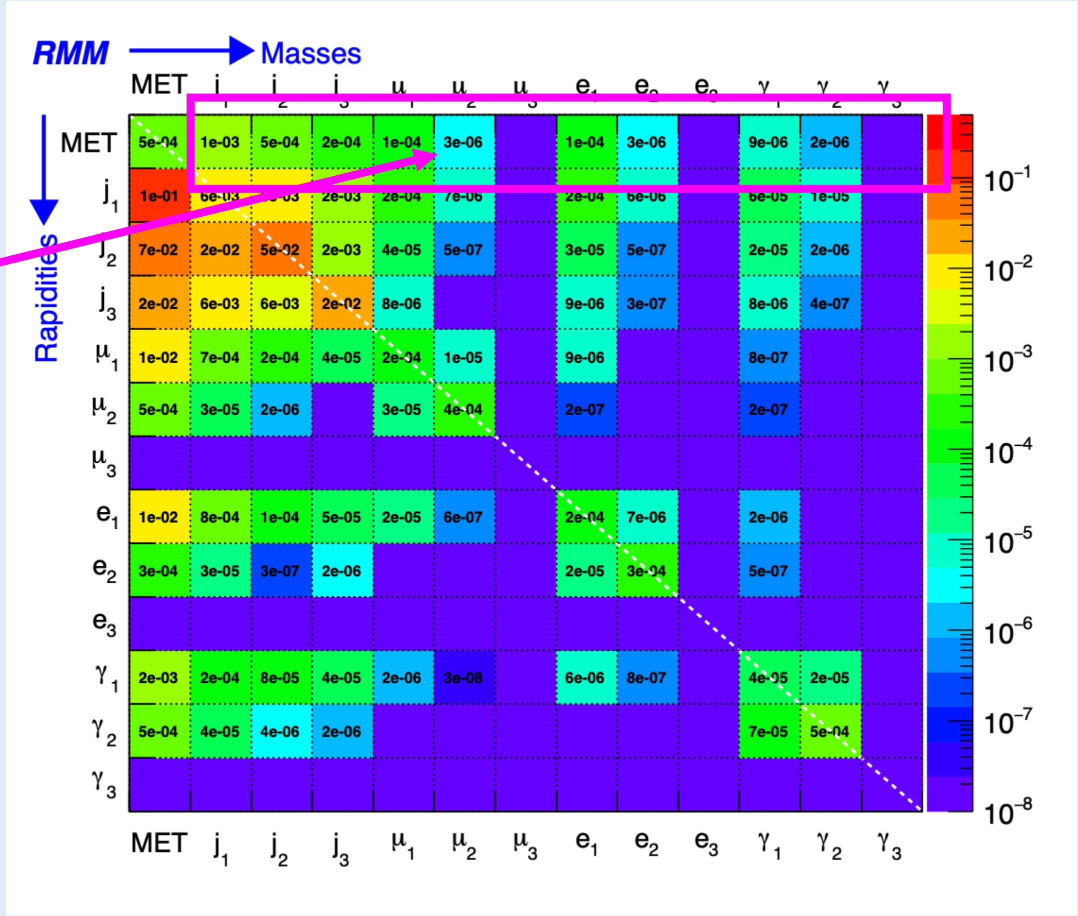
Rapidity Mass Matrix (RMM)

- Encapsulates many variables that are favored for BSM searches

- For example:
  - Missing Transverse Energy (MET)
  - Transverse energies
  - Transverse masses

$$M_T(i_n) = \sqrt{(E_T + E_T^{miss})^2 - \left(p_T + E_T^{miss}\right)^2}$$

Encodes information on masses of particles that decay to invisible particles, e.g., $W \rightarrow l\nu$



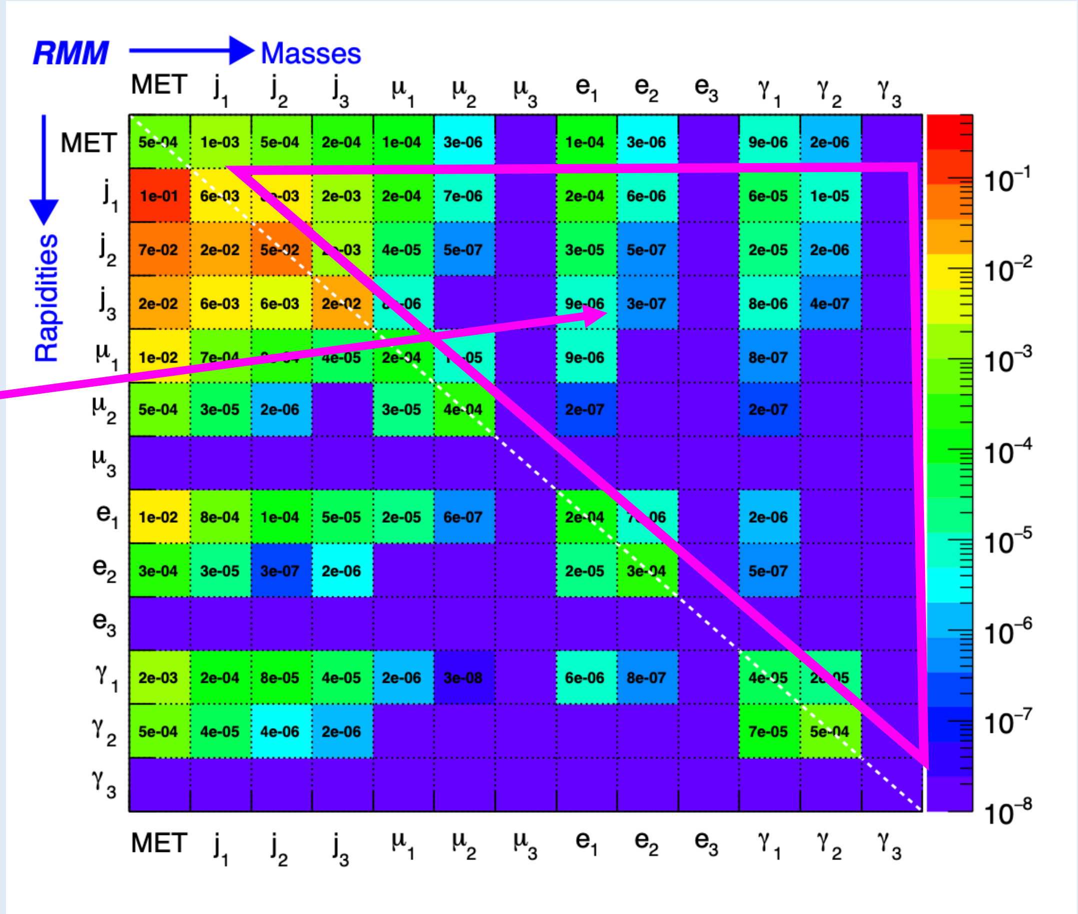
S. V. Chekanov, arXiv:1805.11650

Rapidity Mass Matrix (RMM)

- Encapsulates many variables that are favored for BSM searches

- For example:
  - Missing Transverse Energy (MET)
  - Transverse energies
  - Transverse masses
  - 2-body Invariant masses

Information on masses of resonances that decay to 2 particles, e.g., $Z \rightarrow ll$



S. V. Chekanov, arXiv:1805.11650

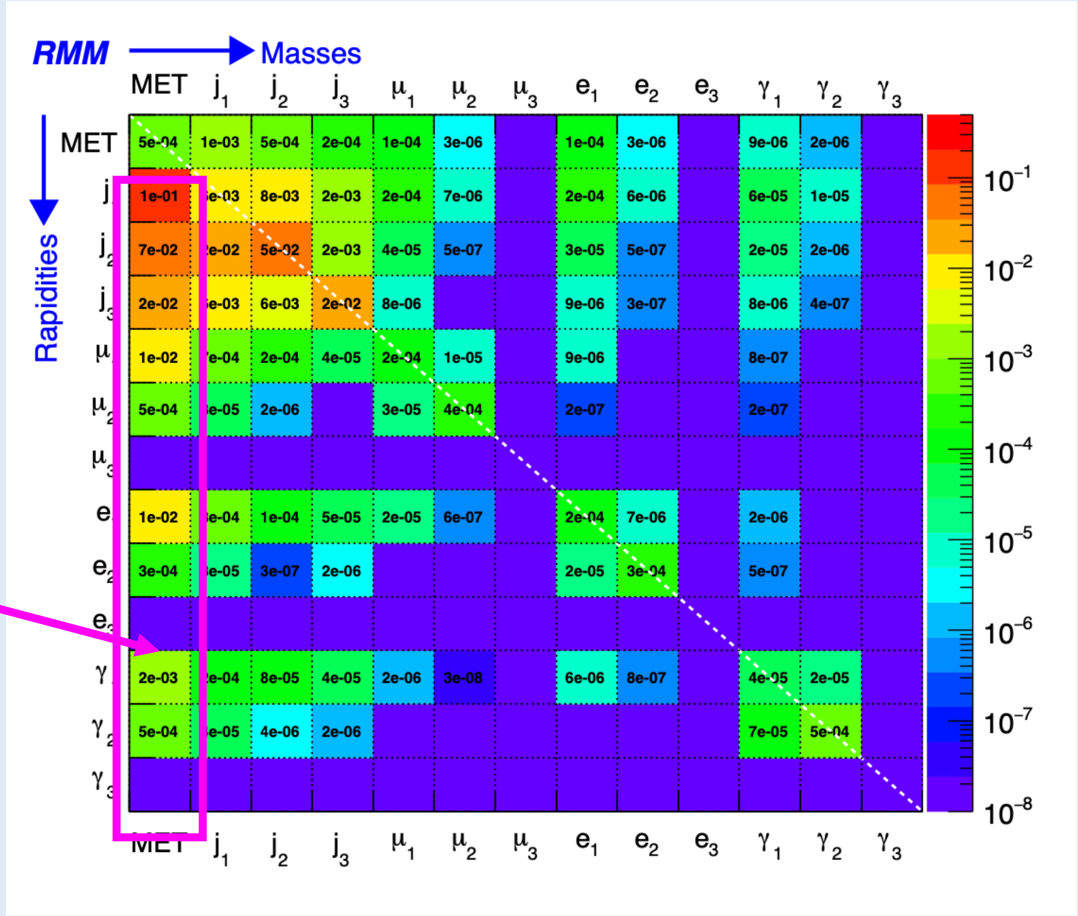
Rapidity Mass Matrix (RMM)

- Encapsulates many variables that are favored for BSM searches

- For example:
  - Missing Transverse Energy (MET)
  - Transverse energies
  - Transverse masses
  - 2-body Invariant masses
  - Rapidity (shifted to [0, 1])

$$h_L(i_n) = C(cosh(y) - 1); \text{ y = rapidity}$$

Indicates the region of the detector where particle was located



S. V. Chekanov, arXiv:1805.11650

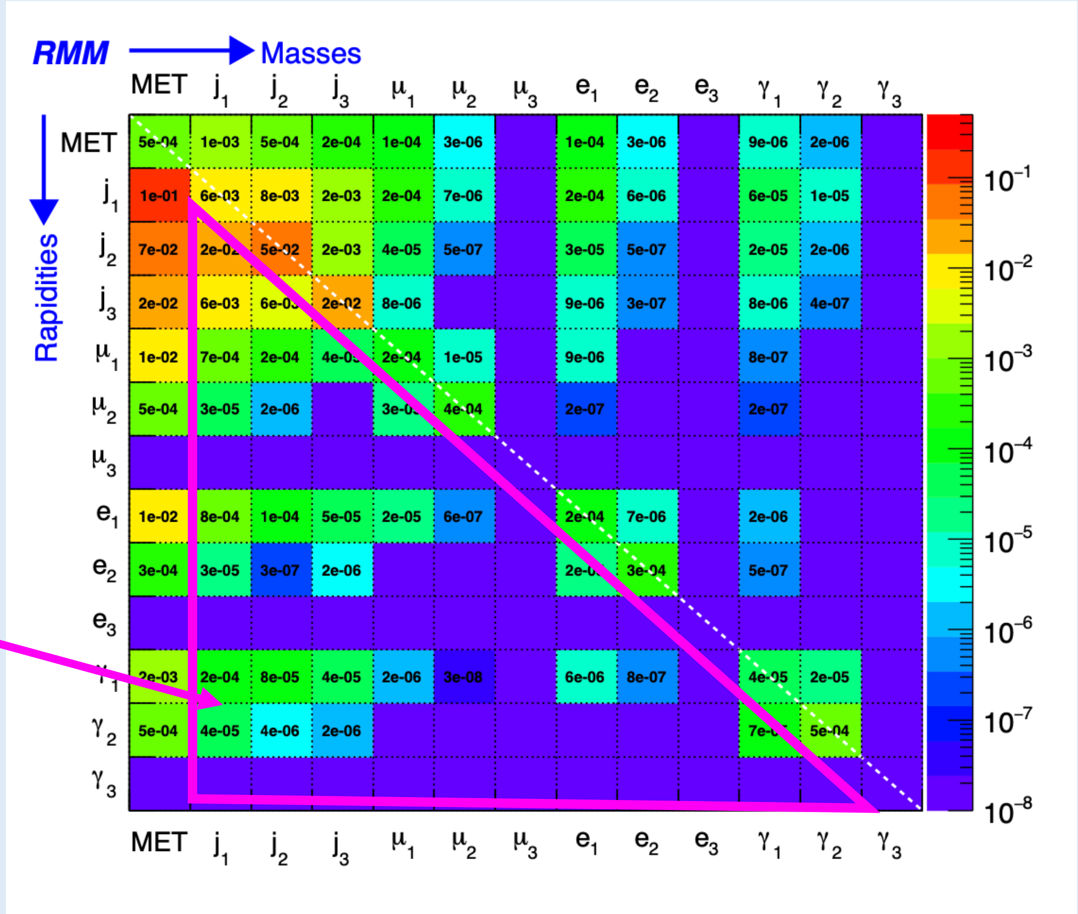
Rapidity Mass Matrix (RMM)

- Encapsulates many variables that are favored for BSM searches

- For example:
  - Missing Transverse Energy (MET)
  - Transverse energies
  - Transverse masses
  - 2-body Invariant masses
  - Rapidity (shifted to [0, 1])
  - 2-particle rapidity differences

$$h_L(i_n, j_k) = C(cosh(y_{i_n} - y_{j_k}) - 1)$$; y = rapidity

Indicates how collimated particles are
(small ⇒collimated)



S. V. Chekanov, arXiv:1805.11650

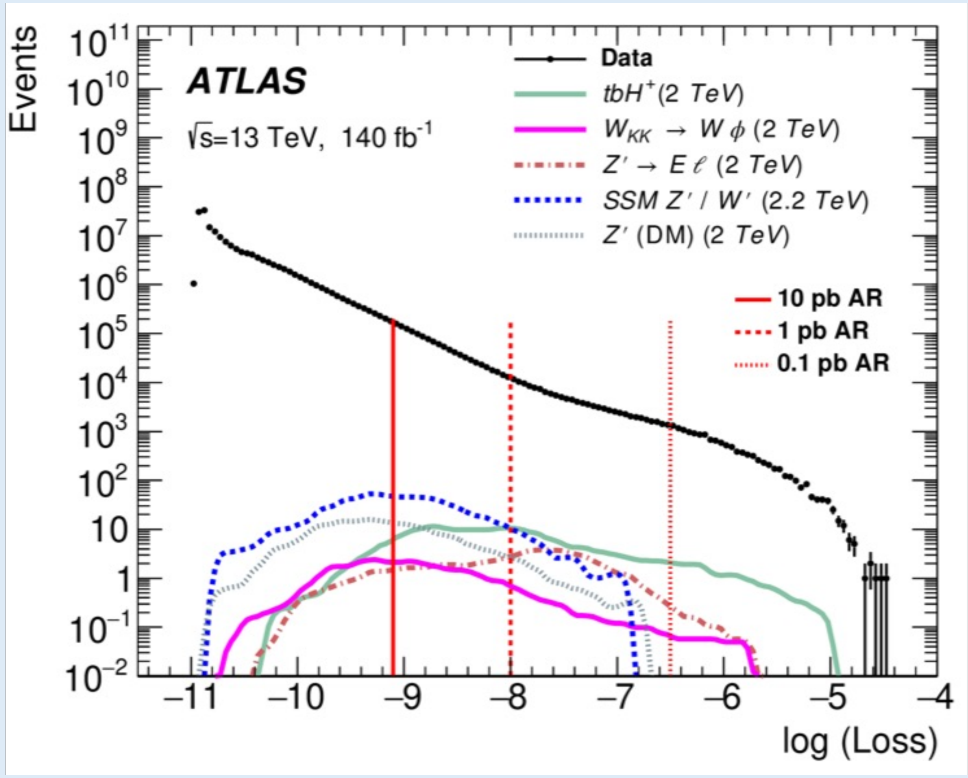
Anomaly Detection at the ATLAS experiment

- Loss and anomaly region cuts defined by the trained AE for a variety of benchmark signals and the data.

- The predictions for the BSM models represent the expected number of events for masses of about 2 TeV from the Run 2 data.

- The BSM models considered were:

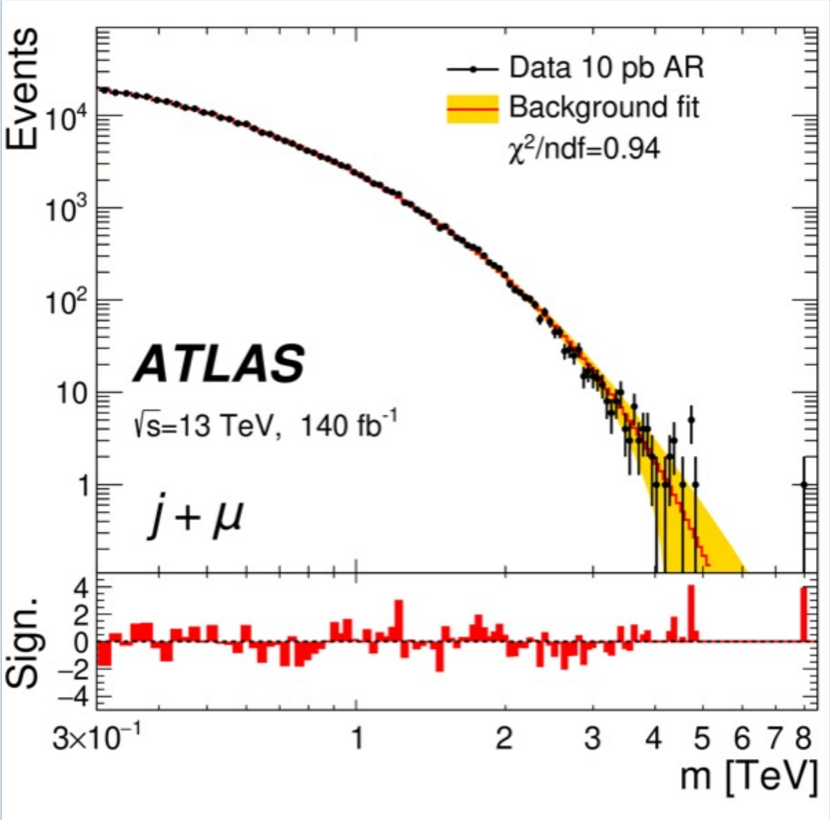
  - Charged Higgs boson produced with a top quark, tbH^+
  - Kaluza-Klein Gauge-boson, W_{KK}
  - A Z'-boson, Z'
  - Sequential-SM Z'-boson, SSM Z'
  - A simplified DM model with an axial-vector Z'-boson mediator



ATLAS, arXiv:2307.01612v1

Anomaly Detection at the ATLAS experiment

- Following training of the AE on the RMM data, anomaly regions were defined based on the reconstruction loss of the AE.
- 9 invariant masses were investigated in each anomaly region: $m_{jj}, m_{bb}, m_{jb}, m_{je}(m_{be}), m_{j\mu}(m_{b\mu}), m_{j\gamma}(m_{b\gamma})$.
- Deviations from the expected SM backgrounds were searched for each of the nine regions.
- No significant deviations from the SM were found.
- Largest deviation of 2.9 sigma seen for a mass of approximately 4.8 TeV shown for the jet + muon mass in the figure, where it was observed.



ATLAS, arXiv:2307.01612v1

Anomaly Detection at the ATLAS experiment

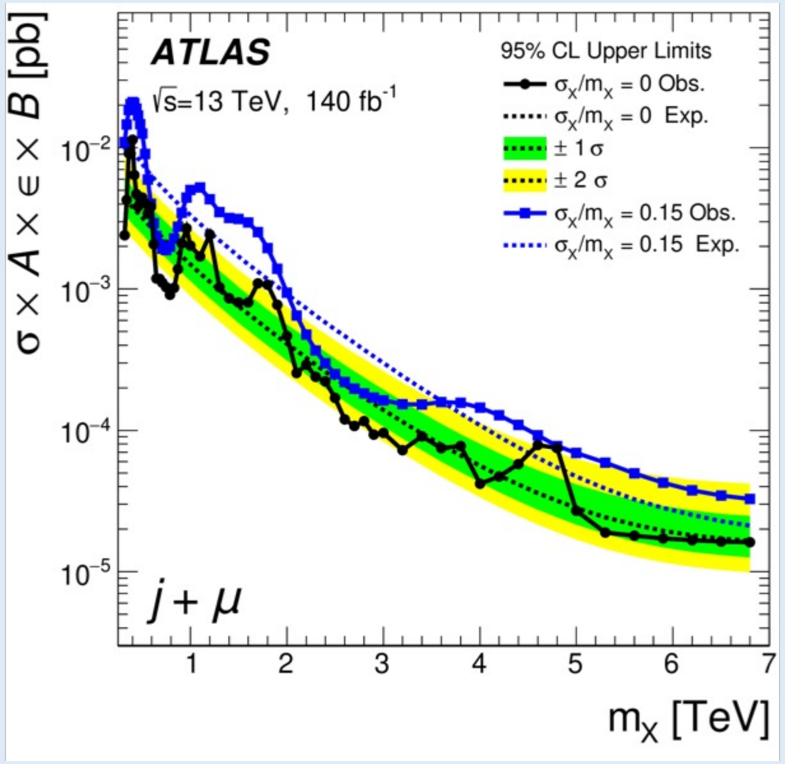
- 95% Confidence limits were determined on : $\boldsymbol{\sigma}\times\boldsymbol{A}\times\boldsymbol{\epsilon}\times\boldsymbol{B}$ for the 9 invariant masses in each Anomaly Region.

  - $\boldsymbol{\sigma}$ = **cross-section** (probability of collision and interaction)
  - $\boldsymbol{A}$ = **acceptance** (indicates the number of particles that would be present the detector)
  - $\boldsymbol{\epsilon}$ = **efficiency** (indication of number of particles detected)
  - $\boldsymbol{B}$ = **branching ratio** (frequency of decays to the given final state).

- These limits are more stringent than the previous results on data with the same preselection.

- The AE acceptance is particularly good for high mass BSM models.

- Now focusing on performing a similar analysis combining the Run-2 data with that collected so-far in Run 3 of the LHC.



ATLAS, arXiv:2307.01612v1

Ongoing Investigations

- Investigations into a number of different AD algorithms are ongoing.

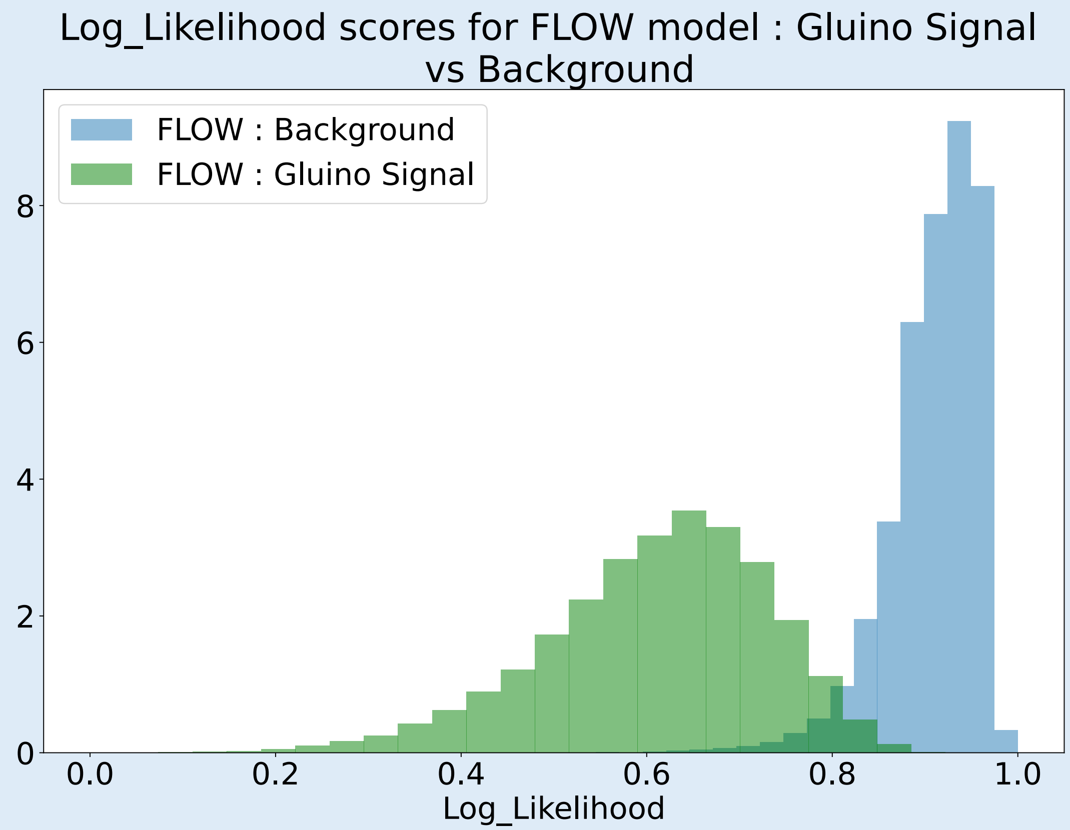
- For example:

  - Deep SVDD
  - Autoregressive Flow

  (S. Caron et. Al., arXiv:2106.10164)

- An example of the Likelihood score for SM backgrounds vs. a BSM Gluino particle for an Autoregressive Flow model is shown in igure.

- The aim is to find an algorithm that is best able to identify a variety of BSM signal models over a large phase space .



T. Aarrestad et. al., arXiv:2105.14027

Summary - Why Anomaly Detection?

- Anomaly Detection with the use of unsupervised machine learning allows for model-independent searches that can extend the sensitivity, and so the discovery-reach, of BSM searches.

- Anomaly Detection analyses have started to be performed at the ATLAS experiment at the LHC, CERN.

- These analyses produce more stringent exclusion limits on the BSM models studied when compared to traditional selections to optimise the signal and background separation.

- Many other AD algorithms and BSM models are being/still to be investigated.

Back-Up Slides

Rapidity Mass Matrix (RMM)

Transverse Energy of leading particle

Transverse Mass of particle :

$$M_T(i_n) = \sqrt{(E_T + E_T^{miss})^2 - (p_T + E_T^{miss})^2}$$

Encodes information on masses of particles that decay to invisible particles, e.g., W $\rightarrow l\nu$

Missing Transverse Energy

$$\begin{pmatrix}
e_T^{\text{miss}} & m_T(j_1) & m_T(j_2) & ...m_T(j_N) & m_T(e_1) & m_T(e_2) & ...m_T(e_N) & ... \\
h_L(j_1) & e_T(j_1) & m(j_1,j_2) & ...m(j_1,j_N) & m(j_1,e_1) & m(j_1,e_2) & ...m(j_1,e_N) & ... \\
h_L(j_2) & h(j_1,j_2) & \delta e_T(j_2) & ...m(j_2,j_N) & m(j_2,e_1) & m(j_2,e_2) & ...m(j_2,e_N) & ... \\
... & ... & ... & ... & ... & ... & ... & ... \\
h_L(j_N) & h(j_1,j_N) & ... & \delta e_T(j_N) & m(j_N,e_1) & m(j_N,e_2) & ...m(j_N,e_N) & ... \\
h_L(e_1) & h(e_1,j_1) & h(e_1,j_2) & ...h(e_1,j_N) & e_T(e_1) & m(e_1,e_2) & ...m(e_1,e_N) & ... \\
h_L(e_2) & h(e_2,j_1) & h(e_2,j_2) & ...h(e_2,j_N) & h(e_2,e_1) & \delta e_T(e_2) & ...m(e_2,e_N) & ... \\
... & ... & ... & ... & ... & ... & ... & ... \\
h_L(e_N) & h(e_N,j_1) & h(e_N,j_2) & ...h(e_N,j_N) & h(e_N,e_1) & h(e_N,e_2) & \delta e_T(e_N) & ... \\
... & ... & ... & ... & ... & ... & ... & ...
\end{pmatrix}$$

2-particle invariant mass

Info on masses of resonances that decay to 2 particles, e.g., Z $\rightarrow ll$

Relative transverse energy between same-type particles

$$\delta e_T(i_n) = \frac{E_T(i_{n-1}) - E_T(i_n)}{E_T(i_{n-1}) + E_T(i_n)}$$

$$h_L(i_n) = C(\cosh(y) - 1);$$

y = rapidity

$$h_L(i_n, j_k) = C(\cosh(y_{i_n} - y_{j_k}) - 1)$$

Indicates how collimated particles are (small -> collimated)

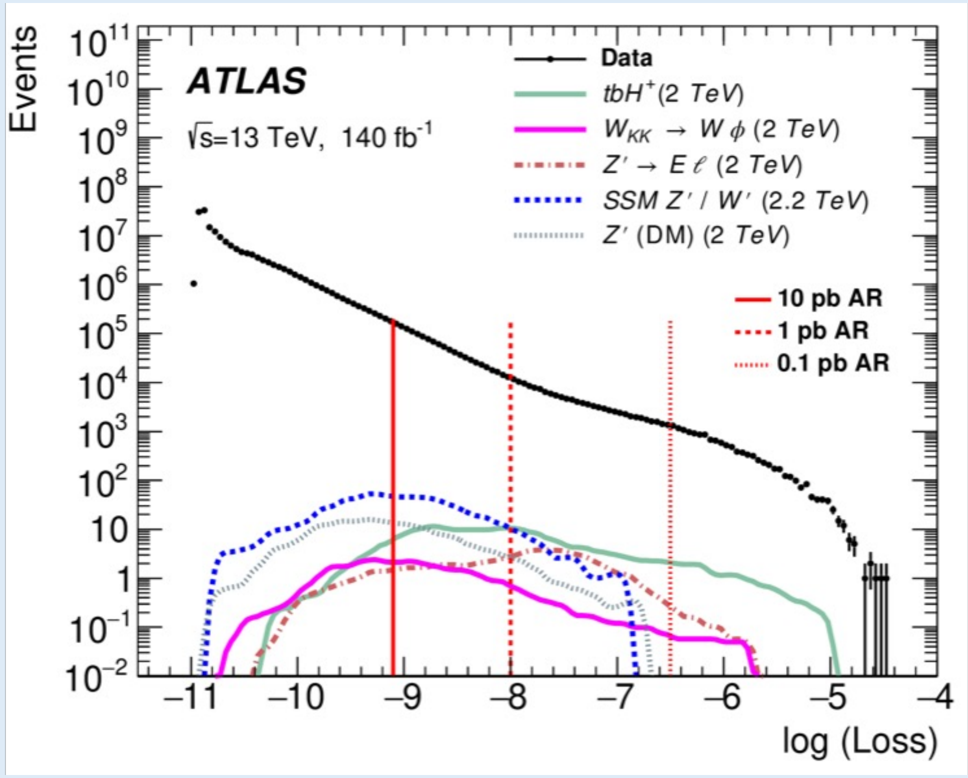
Anomaly Detection at the ATLAS experiment

- Loss and anomaly region cuts defined by the trained AE for a variety of benchmark signals and the data.

- The predictions for the BSM models represent the expected number of events for masses of about 2 TeV from the Run 2 data.

- The BSM models considered were:

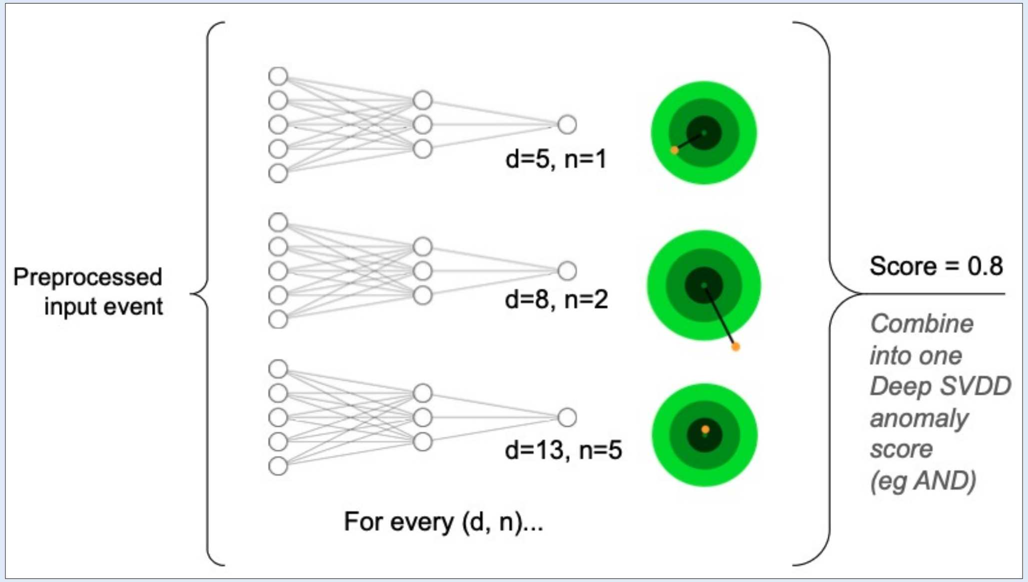
  - Charged Higgs boson produced with a top quark, tbH^+
  - Kaluza-Klein Gauge-boson, W_{KK}
  - A Z'-boson, Z'
  - Sequential-SM Z'-boson, SSM Z'
  - A simplified DM model with an axial-vector Z'-boson mediator



ATLAS, arXiv:2307.01612v1

Other Unsupervised Machine Learning Models

Unsupervised Machine Learning Models – Deep SVDD



S. Caron et. Al., arXiv:2106.10164

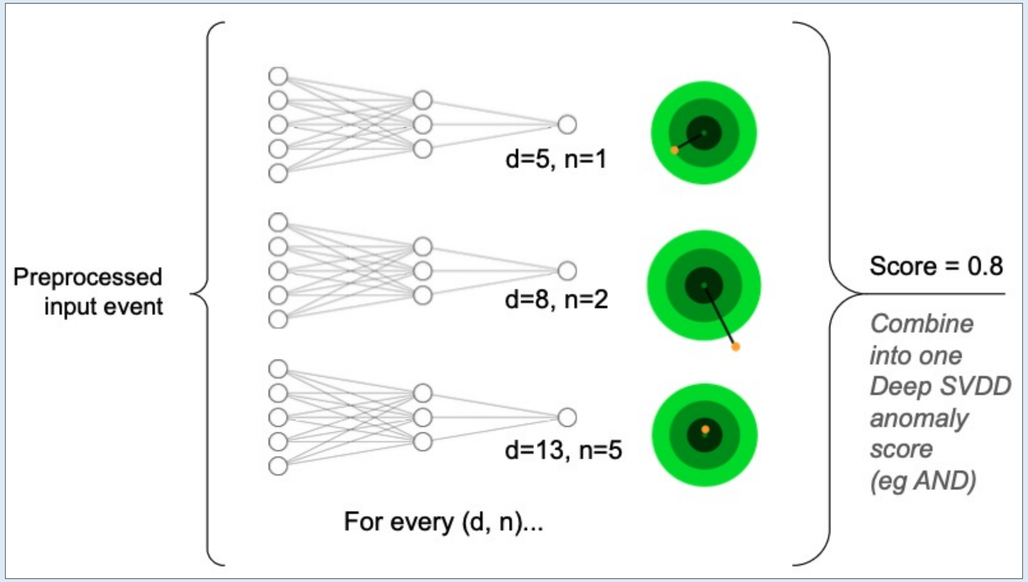
- **Deep SVDD** model [refs] maps input data to a latent space defined by a multidimensional point of a defined target value, e.g., d=5, n= 1 -> (1, 1, 1, 1, 1).

- The latent space is 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.

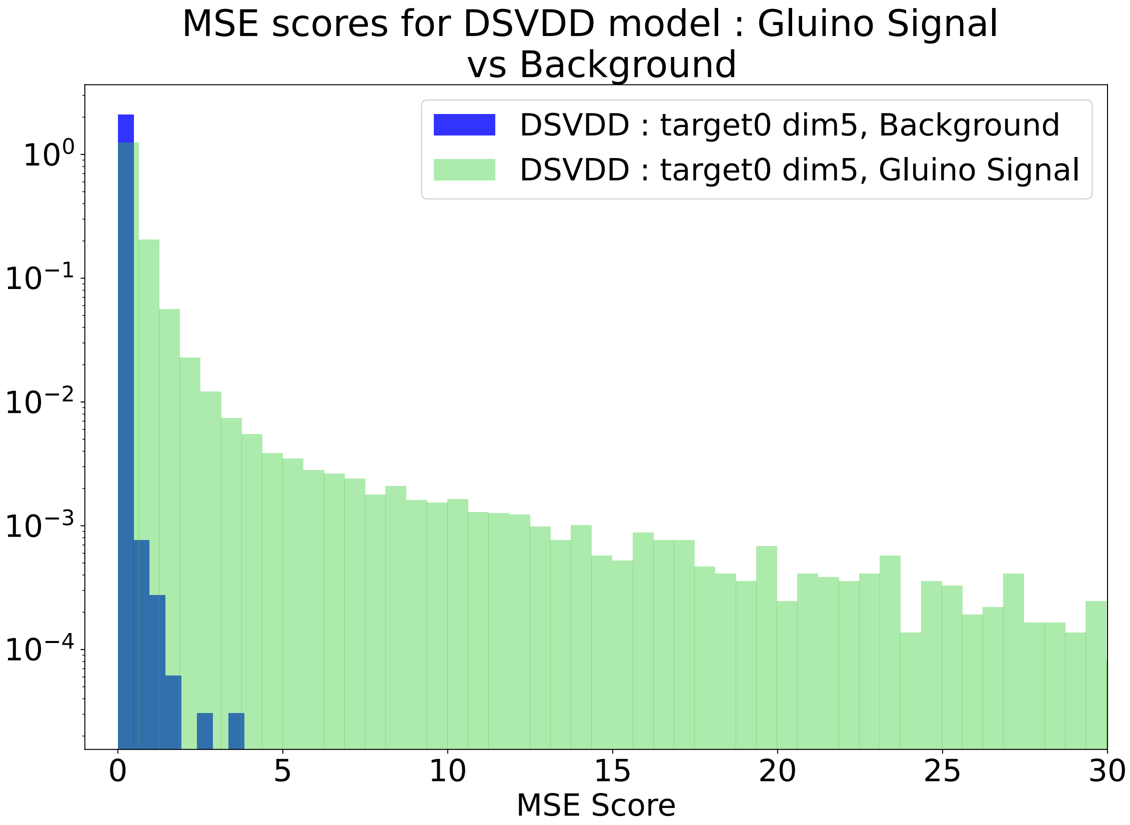
- The anomaly score is then defined as the distance to the defined multidimensional point.

- Data similar to that on which it was trained (here the SM) should lie within the defined multidimensional region, whilst anomalous (BSM) data should fall outside this region.

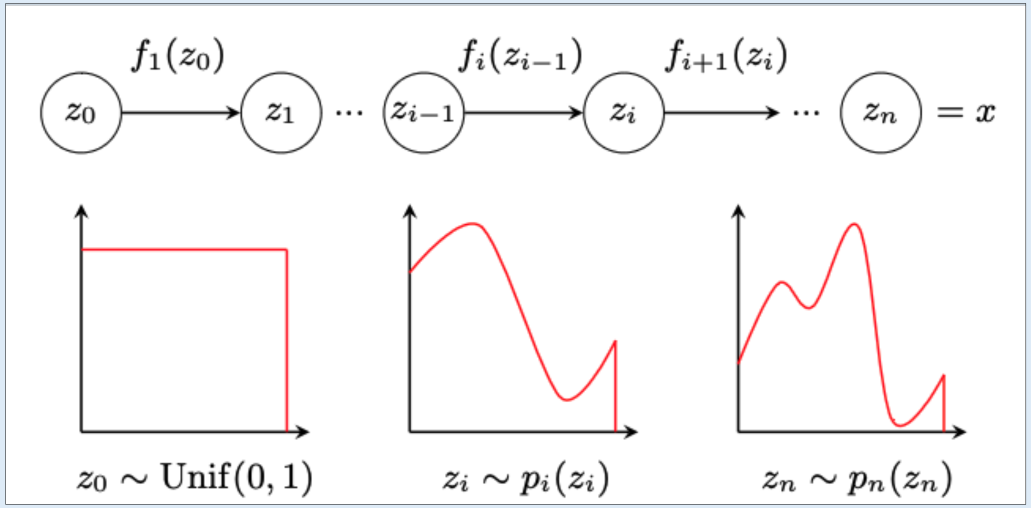
Unsupervised Machine Learning Models – Deep SVDD



S. Caron et. Al., [arXiv:2106.10164]



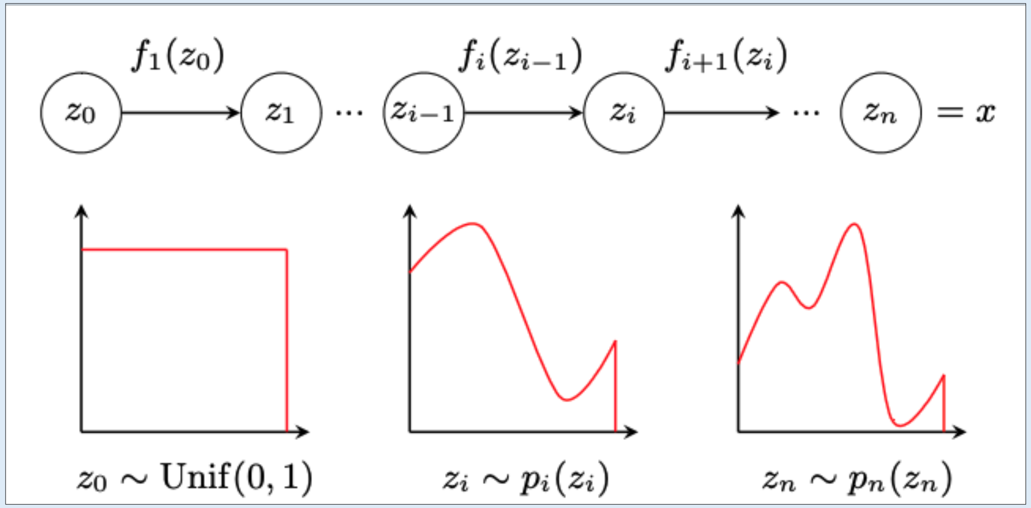
Unsupervised Machine Learning Models – Autoregressive Flow



S. Caron et. Al., arXiv:2106.10164

- The Autoregressive Flow model [ref] 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.
- Useful for detecting **rare** anomalies.

Unsupervised Machine Learning Models – Autoregressive Flow



S. Caron et. Al., [arXiv:2106.10164]