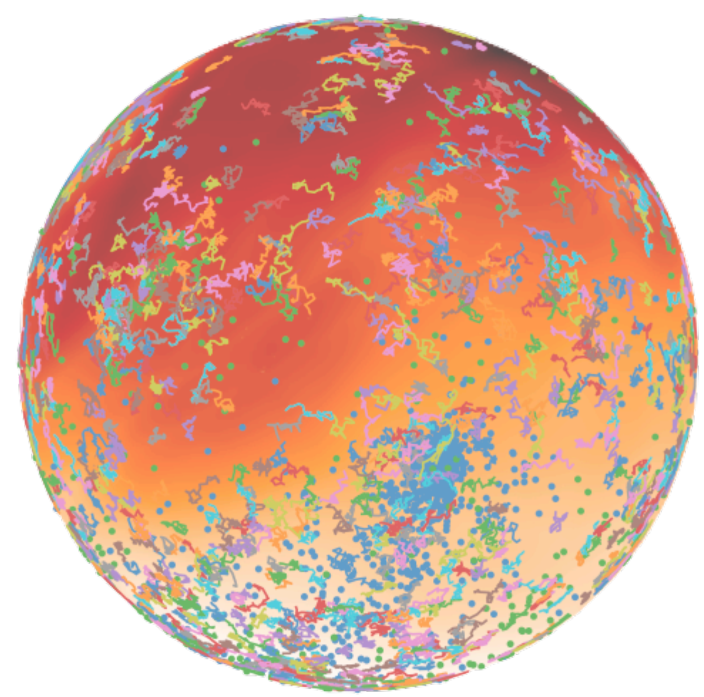
Anomaly searches for new physics at the LHC

Barry Dillon

Universität Heidelberg
Zukunft. Seit 1386.



ACAT - Bari - 25/10/2022

Finding new physics with machine-learning

[see talks by G. Kasieczka, A. Wulzer, A. Gandrakota, …]
[and several interesting posters]

Traditional searches

- specific theory hypotheses & targeted search strategies
- many many possible hypotheses..

Anomaly detection → CMS (MUSiC) & ATLAS (General search)

- compare simulation to data
- typically low-dimensional (high-level observables)

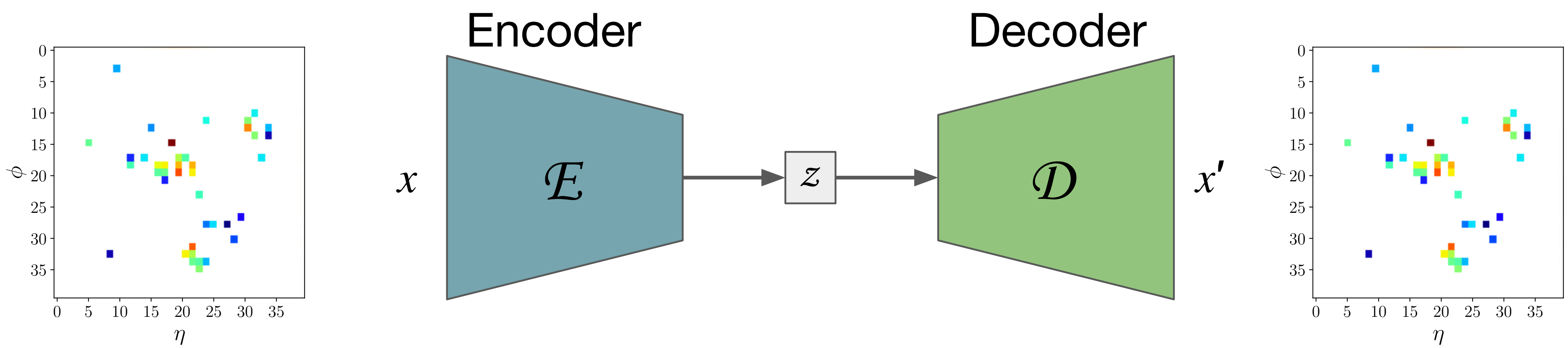
Anomaly detection with machine-learning

- more powerful, can use higher-dimensional data
- more difficult, more complex tools

Dijet resonance search with weak supervision using 13 TeV pp collisions in the ATLAS detector - 2020
arxiv:2005.02983

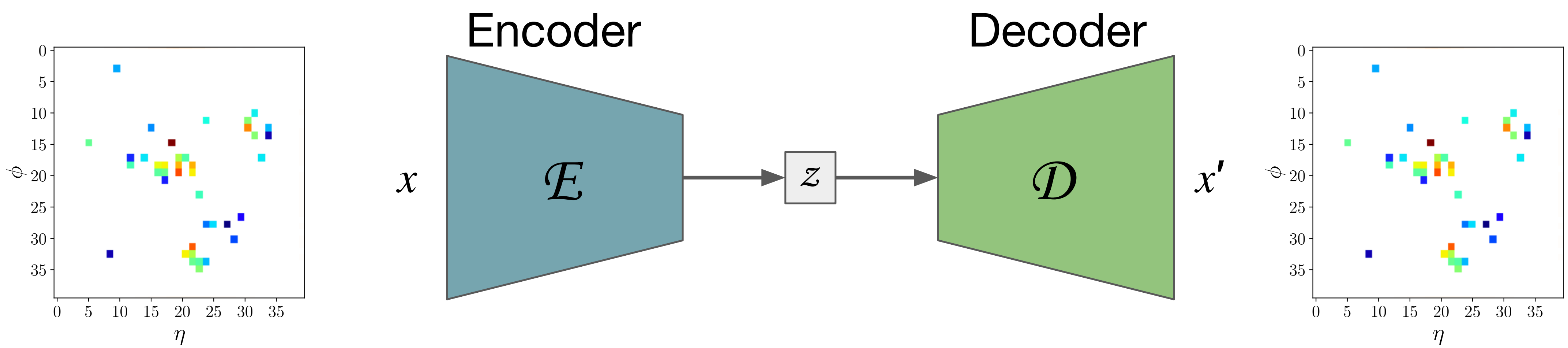
CMS L1 trigger anomaly detection challenge
arxiv:2107.02157

AutoEncoder networks



- Trained to reconstruct the data they are trained on
- Encode the most general features of the data in a latent space z
- Optimised on background-dominant data
- Unsupervised ⟶ model-agnostic, no labels
- Reconstruction loss: $\mathcal{L} = ||x - x'||^2$
- Anomalous data ⇒ data the network has seen least ⇒ larger reconstruction loss

AutoEncoder networks



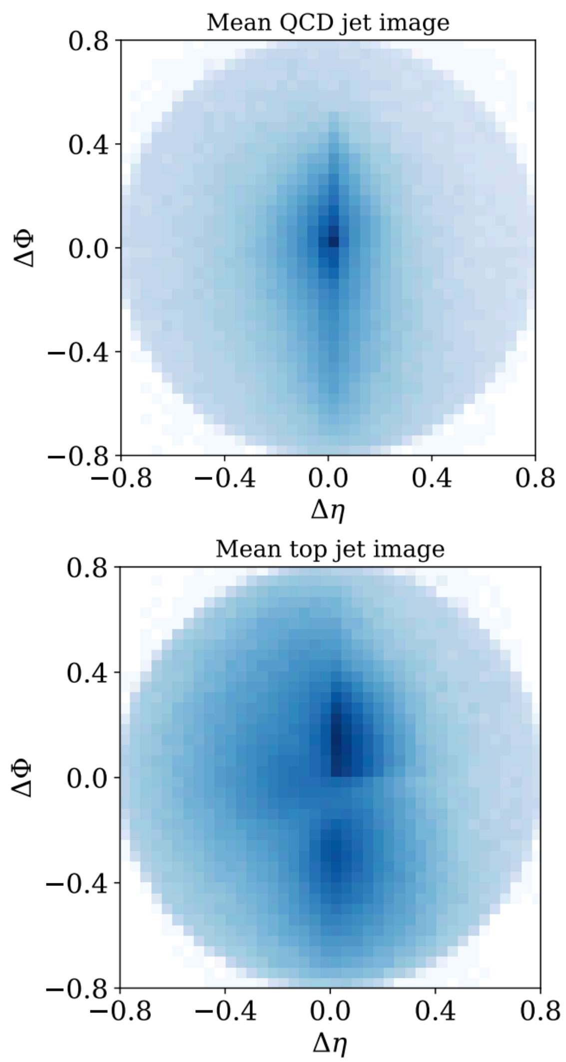
- Trained to reconstruct the data they are trained on
- Encode the most general features of the data in a latent space z
- Optimised on background-dominant data
- Unsupervised → model-agnostic, no labels
- Reconstruction loss: $\mathcal{L} = ||x - x'||^2$
- Anomalous data ⇒ data the network has seen least ⇒ larger reconstruction loss

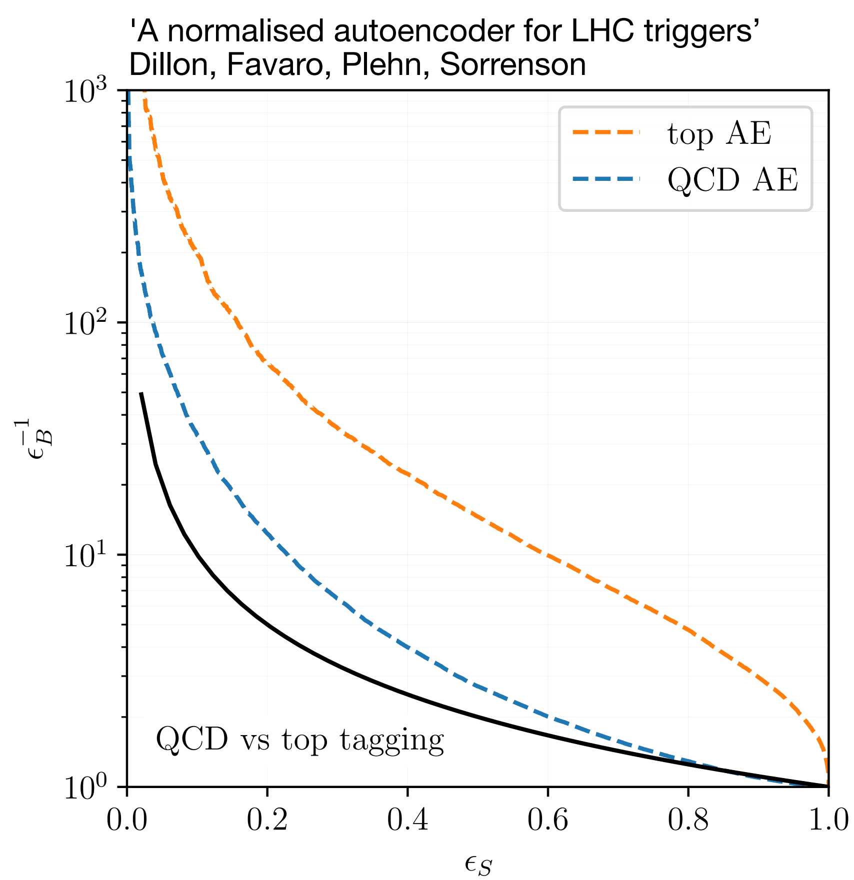
Has proved quite successful, but…

AutoEncoder networks - the problems

They don’t robustly identify anomalous jets.

They do robustly identify complex jets,   e.g anomalous top/QCD jets





An AE trained on only top jets learns to reconstruct QCD jets…

AutoEncoder networks - the problems

They don’t robustly identify anomalous jets.

They do robustly identify complex jets,   e.g anomalous top/QCD jets

Very sensitive to the choice of representation / observables

e.g. under re-mapping of p_T’s, $p_T \to p_T^n$

the results vary a lot [‘What’s anomalous in LHC jets?’ Buss et al]

[‘Anomaly detection under coordinate transformations’ Kasieczka et al]

AutoEncoder networks - the problems

They don't robustly identify anomalous jets.

They do robustly identify complex jets,   e.g anomalous top/QCD jets

Very sensitive to the choice of representation / observables

e.g. under re-mapping of p_T's, $p_T \to p_T^n$

the results vary a lot ['What's anomalous in LHC jets?' Buss et al]
['Anomaly detection under coordinate transformations' Kasieczka et al]

Not invariant to physical symmetries in the problem.

AE can't reconstruct something the latent space is invariant to…

Preprocessing is necessary, but approximate.

AutoEncoder networks - the problems

They don’t robustly identify anomalous jets.

They do robustly identify complex jets,   e.g anomalous top/QCD jets

Very sensitive to the choice of representation / observables

e.g. under re-mapping of p_T’s, $p_T \rightarrow p_T^n$

the results vary a lot [‘What’s anomalous in LHC jets?’ Buss et al]

[‘Anomaly detection under coordinate transformations’ Kasieczka et al]

Normalised AutoEncoder

Not invariant to physical symmetries in the problem.

AE can’t reconstruct something the latent space is invariant to…

Preprocessing is necessary, but approximate.

What is anomalous?

['What's anomalous in LHC jets?' Buss et al]

Reconstruction is a very vague way to define anomalies

More accurately: anomalies are events/jets in low density regions of the feature space

$\Rightarrow$ not invariant to transformations in feature space

Machine-learned density estimation:

1 - some parameterisation of the density $p_{\text{data}}(\overrightarrow{x})$

2 - a scheme to minimise $-\log p_{\text{data}}(\overrightarrow{x})$ wrt to the parameters

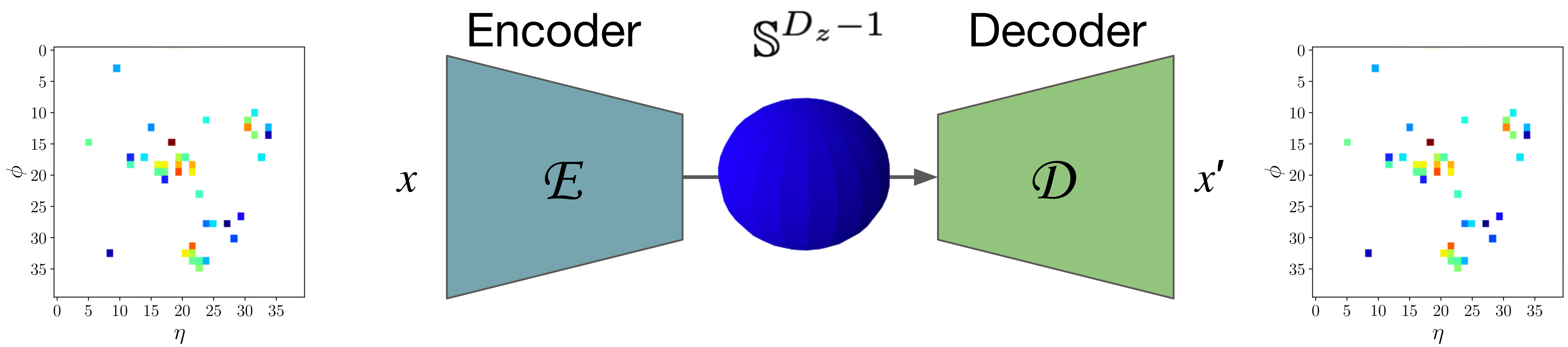
Can be difficult in high-dimensions

Use the dimensional reduction of an AE latent space to initiate the density estimation

$\rightarrow$ **the Normalised AutoEncoder**

The Normalised AutoEncoder

[‘Autoencoding under normalization constraints’ - Yoon, Noh, Park]



Energy-based model:

$$p_\theta(x) = \frac{1}{Z_\theta} e^{-E_\theta(x)}, \quad E_\theta(x) = (x - x')^2$$

with

$$Z_\theta = \int_x e^{-E_\theta(x)} dx$$

Loss function: $\mathscr{L} = -\log p_\theta(x)$

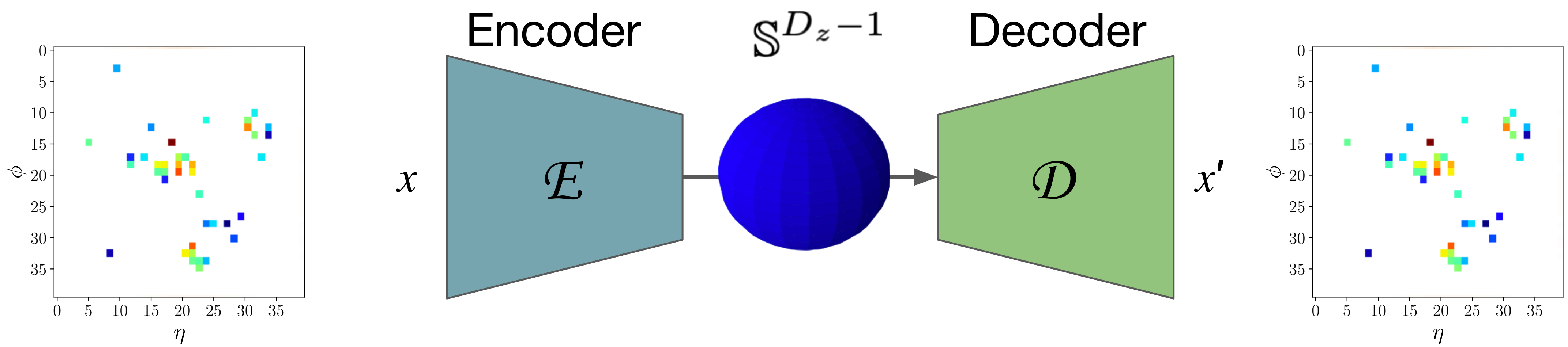
$\Rightarrow$ learning a density model for the data

The anomaly score is just the density, or equivalently, the energy:

$$E_\theta(x) = (x - x')^2$$

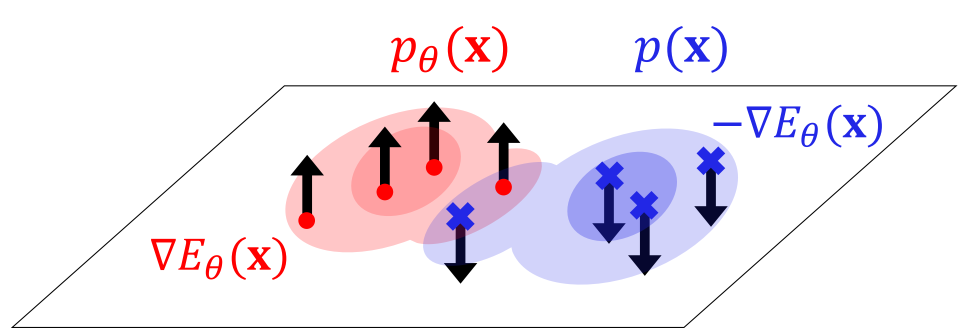
The Normalised AutoEncoder

[‘Autoencoding under normalization constraints’ - Yoon, Noh, Park]



Just the AE loss, plus an additional term:

$$\mathscr{L} = -\log p_\theta(x) = E_\theta(x) + \log Z_\theta$$



Taking gradients:

$$\nabla \mathscr{L}(x) = \nabla_\theta E_\theta(x) - \langle \nabla_\theta E_\theta(x) \rangle_{x \sim p_\theta(x)}$$

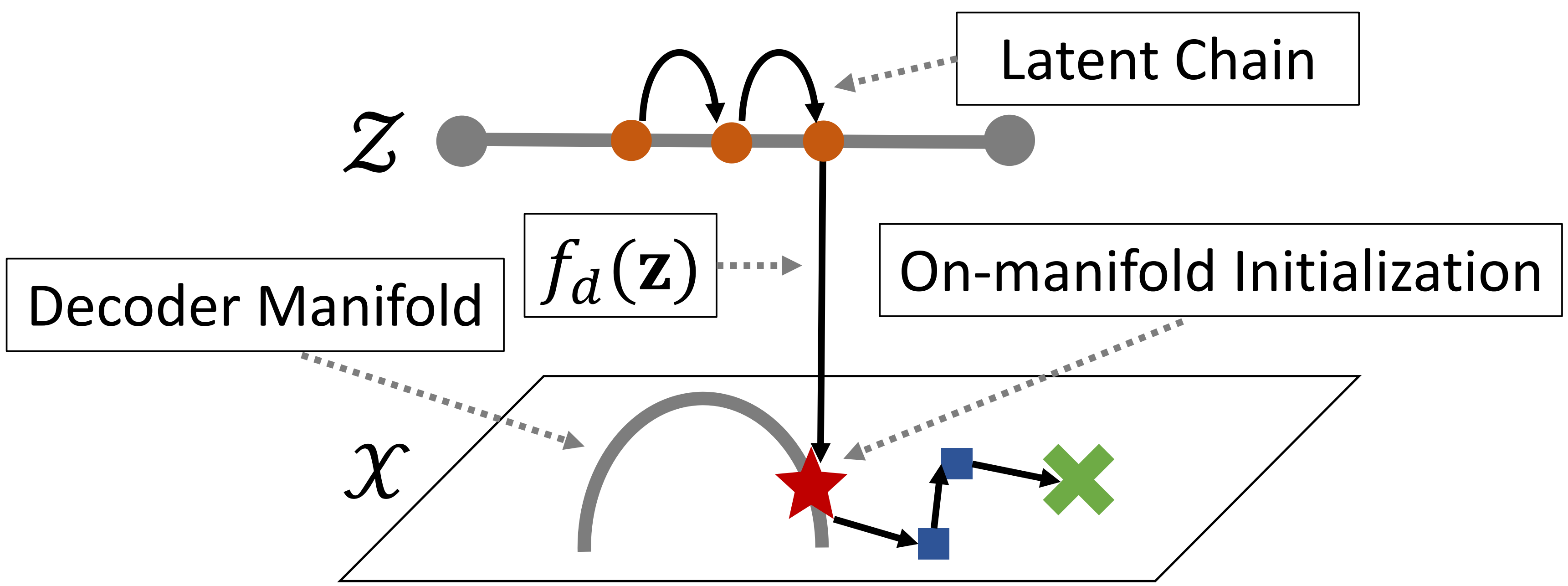
This second term is intractable → approximated via MCMC

Plays an important (more intuitive) role: penalises out of distribution samples!

The Normalised AutoEncoder

[‘Autoencoding under normalization constraints’ - Yoon, Noh, Park]

Sampling from $p_\theta(\overrightarrow{x})$ with On-Manifold-Initialisation



$$x_{t+1} = x_t + \lambda \nabla_x \log p_\theta(x) + \sigma \epsilon_t$$

Reduces the workload of density estimation in high-dimensional spaces

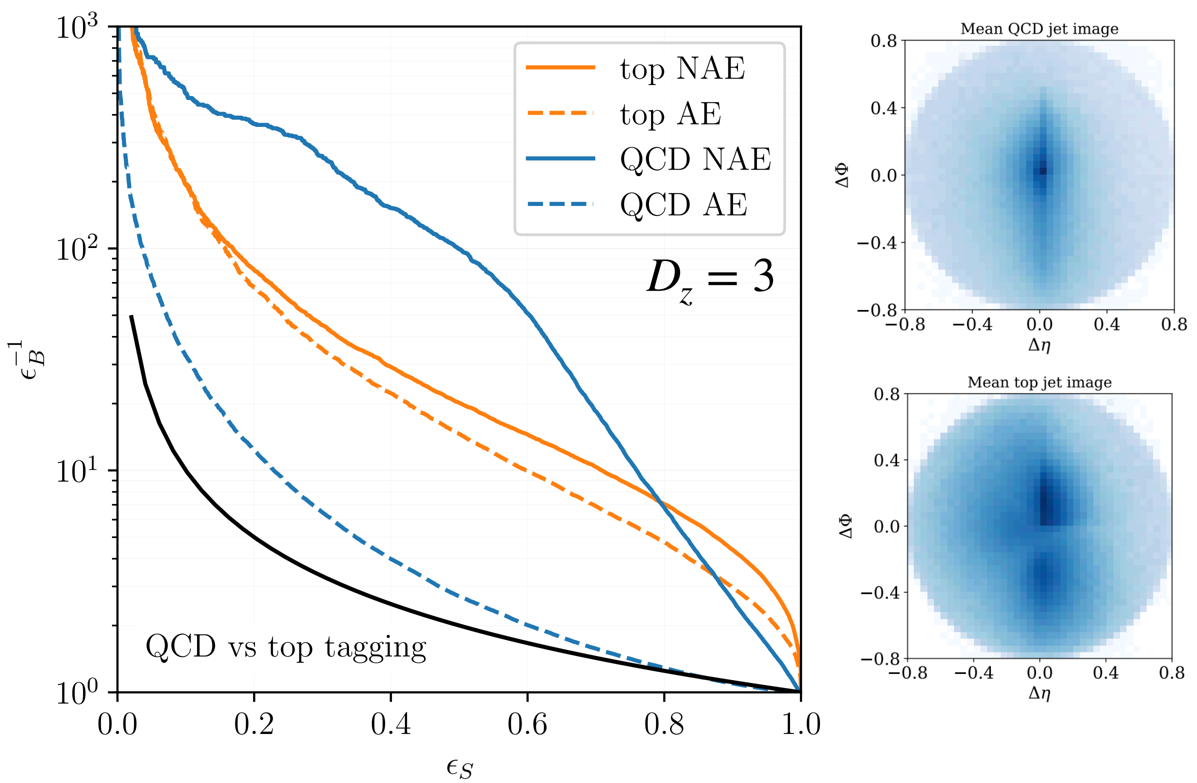
Training is time-consuming $\rightarrow$ inference time is the same as a regular AE

The Normalised AutoEncoder

['A normalised autoencoder for LHC triggers' - Dillon, Favaro, Plehn, Sorrenson, Krämer]

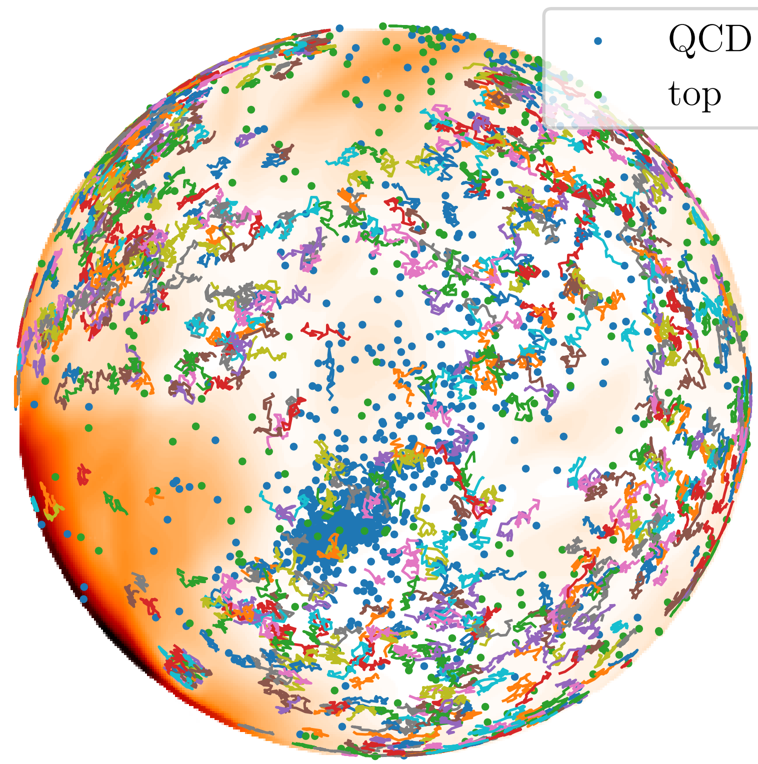
No complexity bias!

More rubust and reliable anomaly detection



Visualisation of the latent space

Looks like a mess, but very useful for interpreting the results and diagnosing problems with the training!



The Normalised AutoEncoder

['A normalised autoencoder for LHC triggers' - Dillon, Favaro, Plehn, Sorrenson, Krämer]

CMS challenge: anomaly detection on L1 triggers

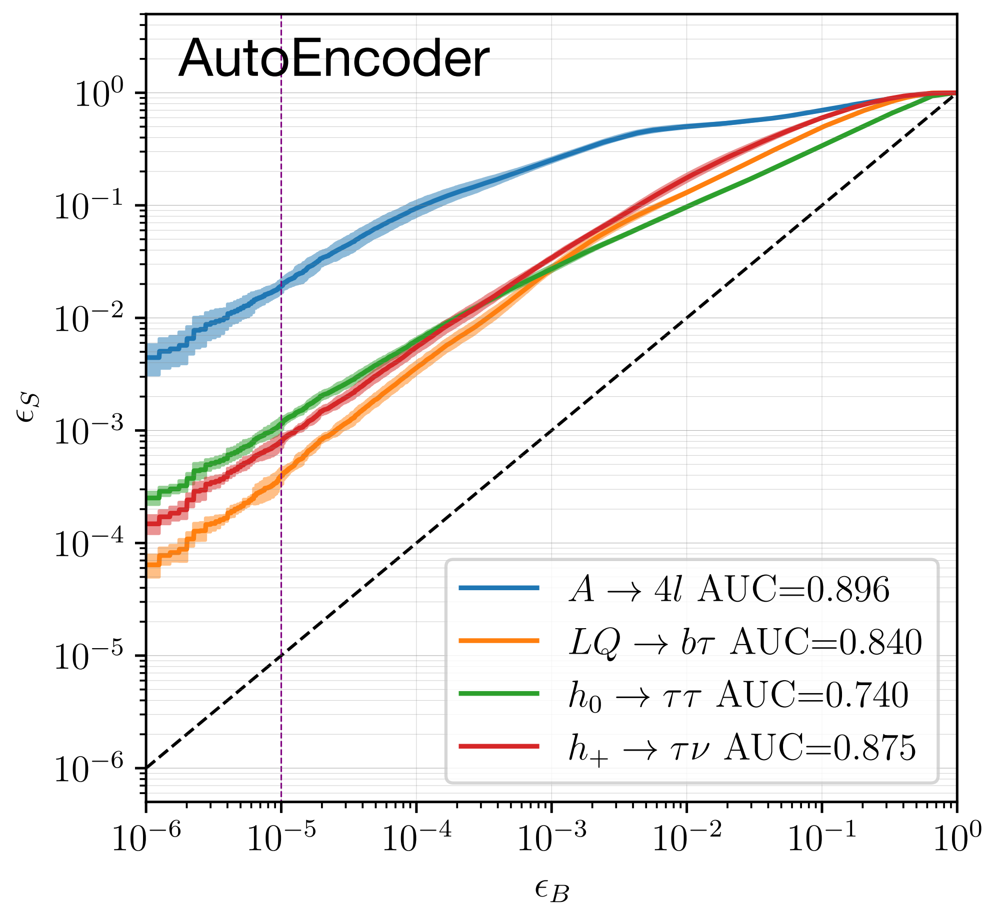
'LHC physics dataset for unsupervised New Physics detection at 40 MHz' - Govorkova, Puljak, Aarrestad, Pierini, Woźniak, Ngadiuba

Benchmark data

selection cut:
e or μ with $p_T > 23$ GeV

objects:
10 jets, 4 e, 4 μ, MET
(p_T, η, ϕ) for each

Preliminary



Backgrounds & signals

- SM bkgs:
  - inclusive W and Z
  - $t\bar{t}$
  - QCD
- BSM sigs:
  - $A \to 4l$
  - $LQ \to b\tau$
  - $h_0 \to \tau\tau$
  - $h_+ \to \tau\nu$

Results at low ϵ_B very sensitive to data preprocessing!

The Normalised AutoEncoder

[‘A normalised autoencoder for LHC triggers’ - Dillon, Favaro, Plehn, Sorrenson, Krämer]

CMS challenge: anomaly detection on L1 triggers

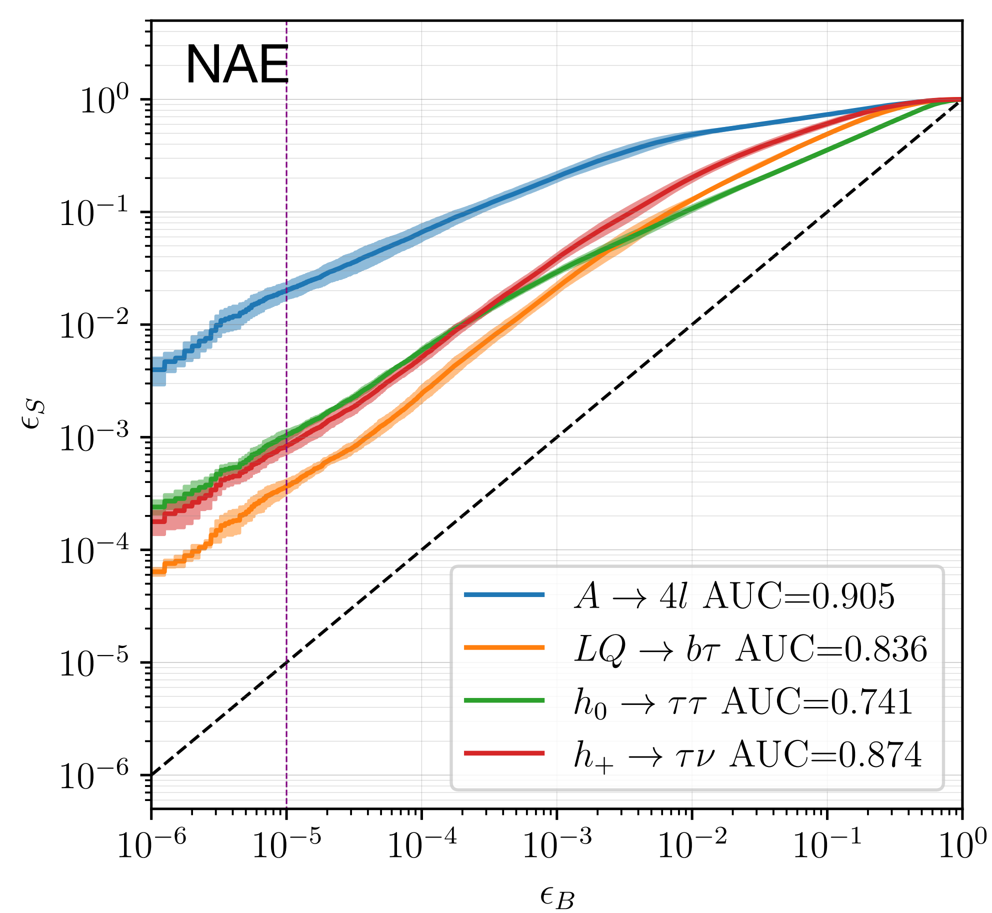
'LHC physics dataset for unsupervised New Physics detection at 40 MHz’ - Govorkova, Puljak, Aarrestad, Pierini, Woźniak, Ngadiuba

Benchmark data

selection cut:
e or μ with $p_T > 23$ GeV

objects:
10 jets, 4 e, 4 μ, MET
(p_T, η, ϕ) for each

Preliminary



What about the NAE?

- much the same performance on benchmark data
- much better performance when we flip signals and backgrounds

  → better density estimation
  → no complexity bias
  → more signal coverage

- more precise results coming soon

  see Luigi Favaro’s talk at ML4Jets2020!

Results at low ϵ_B very sensitive to data preprocessing!

Conclusions & outlook

The Normalised AutoEncoder gives much more robust and interpretable anomaly detection than regular AutoEncoders.

Further work still on-going in the study of the NAE.

1 - Anomaly scores at very low background efficiencies

→ results at low ϵ_B very sensitive to re-mappings of data, not model-independent

2 - Representation-learning for anomaly detection

→ we need representations of data that are invariant to symmetries and expressive

['Symmetries, Safety, & Self-supervision' - Dillon et al]
['Self-supervised anomaly detection' - Dillon et al]

→ symmetries ⇒ more efficient use of data, especially important in the tails…