A Variational AutoEncoder for model independent searches of new physics at LHC

Giulia Lavizzari[1,2], Giacomo Boldrini[1,2], Simone Gennai[1], Pietro Govoni[1,2]

[1] INFN sezione di Milano Bicocca
[2] Università degli Studi di Milano Bicocca

Aim of the work

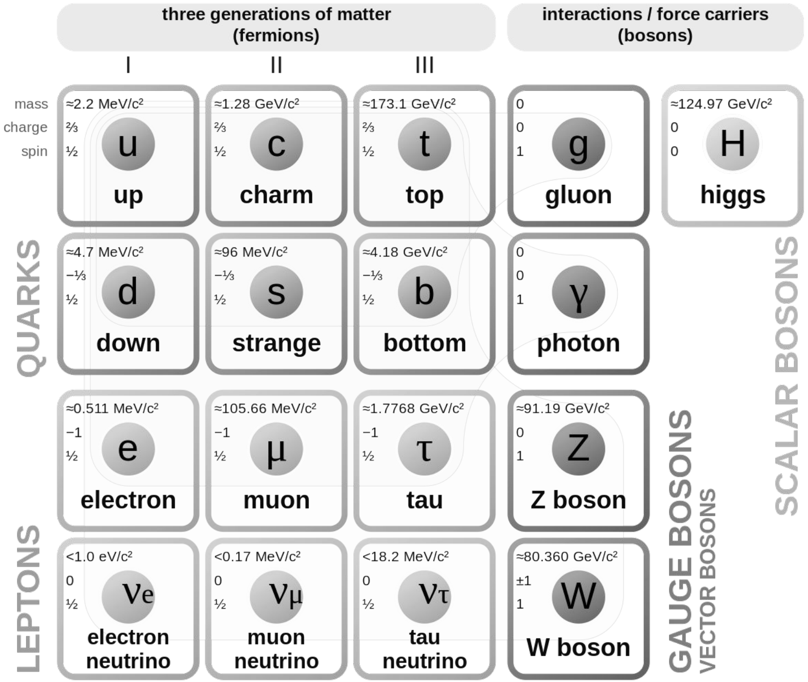
Unsupervised learning methods (Variational AutoEncoders) for anomaly detection to search for new physics at the LHC

Summary:

- The search of new physics at LHC
- The physics use-case: an effective field theory interpretation of Vector Boson Scattering
- Autoencoders
- Variational AutoEncoders
- Our model and its performance

The Standard Model and its limitations

The SM encodes our understanding of the fundamental structure of matter:

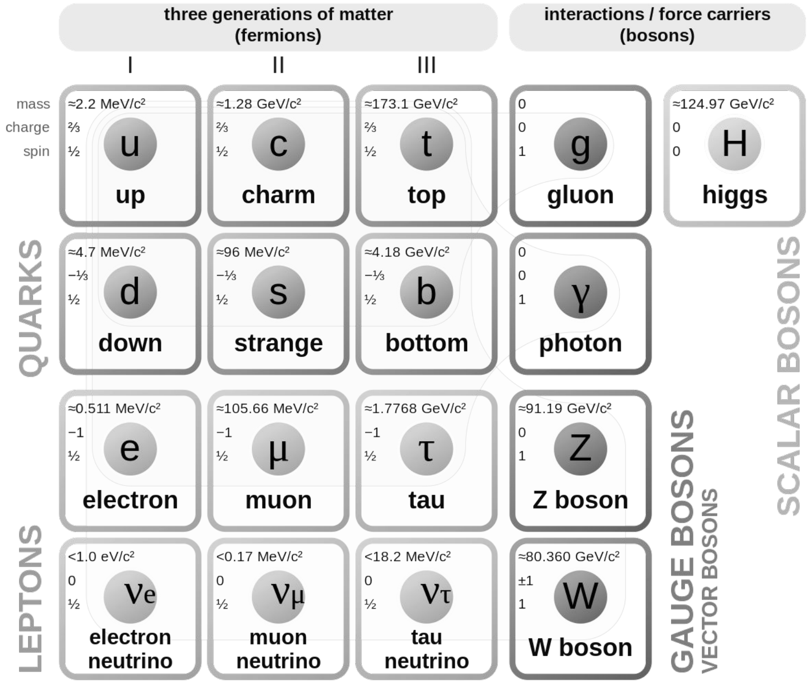


It describes:

- All the know particles that constitute matter
- Three of the four fundamental forces that govern their interactions
- The Higgs Boson

The Standard Model and its limitations

The SM encodes our understanding of the fundamental structure of matter:



So far it was **extremely successful in providing experimental predictions and theoretical explanations**

→ e.g. discovery of the Higgs Boson

However, **many questions remain unanswered** e.g.

- dark matter
- matter/antimatter asymmetry
- hierarchy problem

→ need for **new physics models**

The search for new physics at the LHC:

Direct searches:

target specific signatures (e.g. SUSY…)

- very effective if the model is correct
- they are only sensitive to the model they target

despites all the data collected at the LHC, no new physics was found

We are looking in the wrong direction!

The search for new physics at the LHC:

Direct searches:

target specific signatures (e.g. SUSY…)

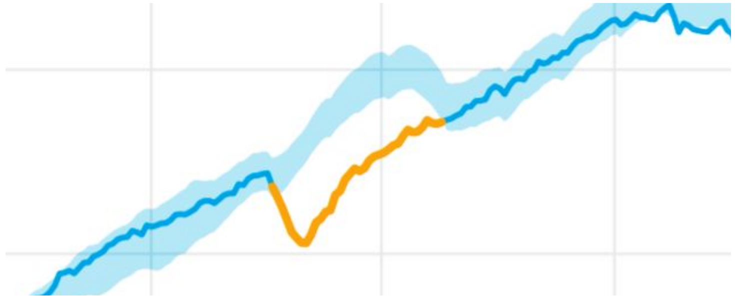
- very effective if the model is correct
- they are only sensitive to the model they target

despites all the data collected at the LHC, no new physics was found

We are looking in the wrong direction!

Model independent searches:

Aimed at finding unusual patterns in data, regardless of the new physics responsible for such anomalies



- less effective on specific signatures
- broader search

Modeling the anomalies: Effective Field Theories

We need simulations of physics Beyond the Standard Model (BMS) to test our strategy:

- a **general** but **still predictive** theory
- a theory that can **regroup a large number of BSM processes**

SM Effective Field Theory (SMEFT) approach → taylor expansion of SM Lagrangian

$$\mathcal{L}_{EFT} = \mathcal{L}_{SM} + \sum_{i,d>4} \frac{c_i}{\Lambda^{d-4}} \mathcal{O}^{(d_i)}$$

Λ – new physics scale

$\mathcal{O}^{(d_i)}$ – EFT operator of dimension d_i

c_i – Wilson coefficient

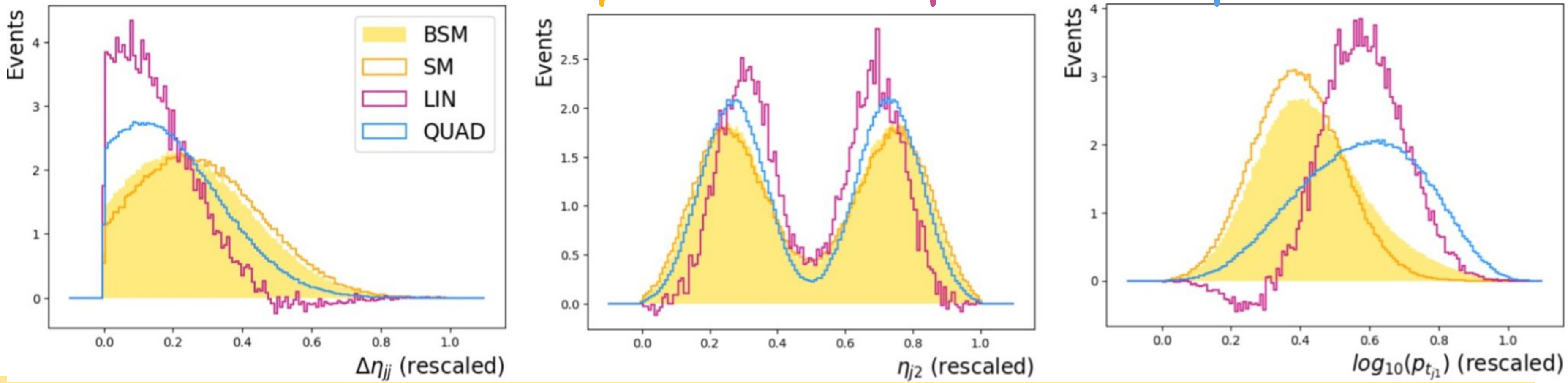
- The **SM** is seen as a **low energy approximation** of a more complete theory
- The BSM effects are parametrized as **higher order operators**

Modeling the anomalies: Effective Field Theories

The EFT operators modify the distributions of the variables, that now comprise:

- A pure **SM contribution**
- Additional terms with **linear** and **quadratic** dependence on the EFT operator

$$|A_{EFT}|^2 = \underbrace{|A_{SM}|^2} + \underbrace{2Re(A_{SM}A^*_{op})} + \underbrace{|A_{op}|^2}$$



The strategy: anomaly detection with VAEs

EFT is a complex, multidimensional problem:

- o (2500) parameters to constrain
- each operator affects differently each variable
  - hard to define a single observable to detect all operators

The strategy: anomaly detection with VAEs

EFT is a complex, multidimensional problem:

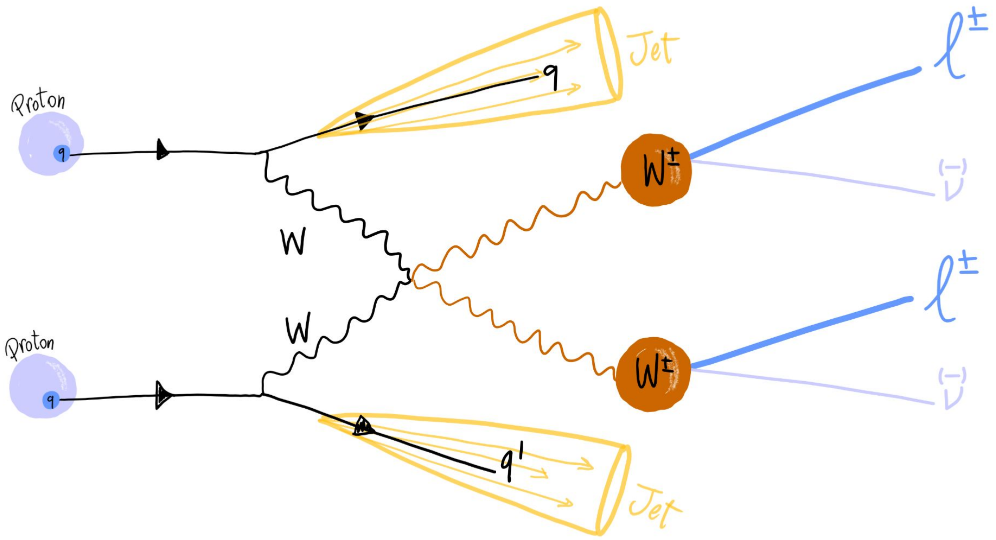
- o (2500) parameters to constrain
- each operator affects differently each variable
  - hard to define a single observable to detect all operators

→ We want to **build a strategy that maximizes the observation of anything that is not Standard Model** (in principle we should see all the operators):

- Variational AutoEncoders
  - Unsupervised learning is an increasingly popular choice [2101.08320]
- idea: **train a model on know physics, and later use it to detect outliers** (anomaly detection task)

The physics use-case: Vector Boson Scattering

Takes place at the LHC when two quarks radiate vector bosons, which in turn interact

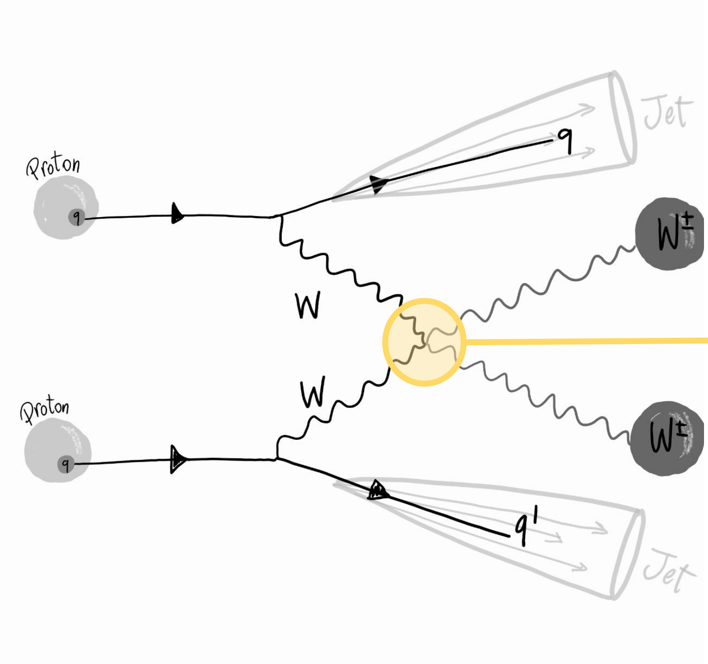


Same sign WW scattering: a very clean signature in the detector

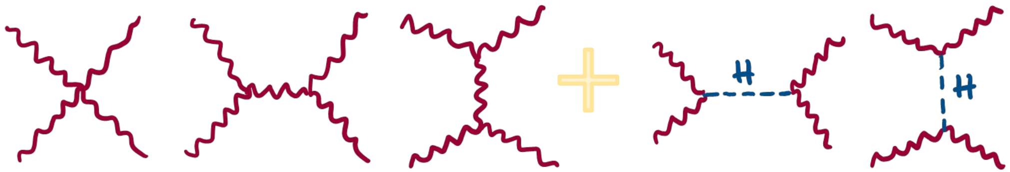
- two jets
- two same sign charged leptons
- Missing Transverse Energy (neutrinos)

The physics use-case: Vector Boson Scattering

It is the **perfect place to search for modifications in the higgs and electroweak sector**



Delicate set of cancellations between diagrams with and without Higgs boson:



It's a delicate equilibrium:

→ any deviations would signal new physics, independently of the theory considered!

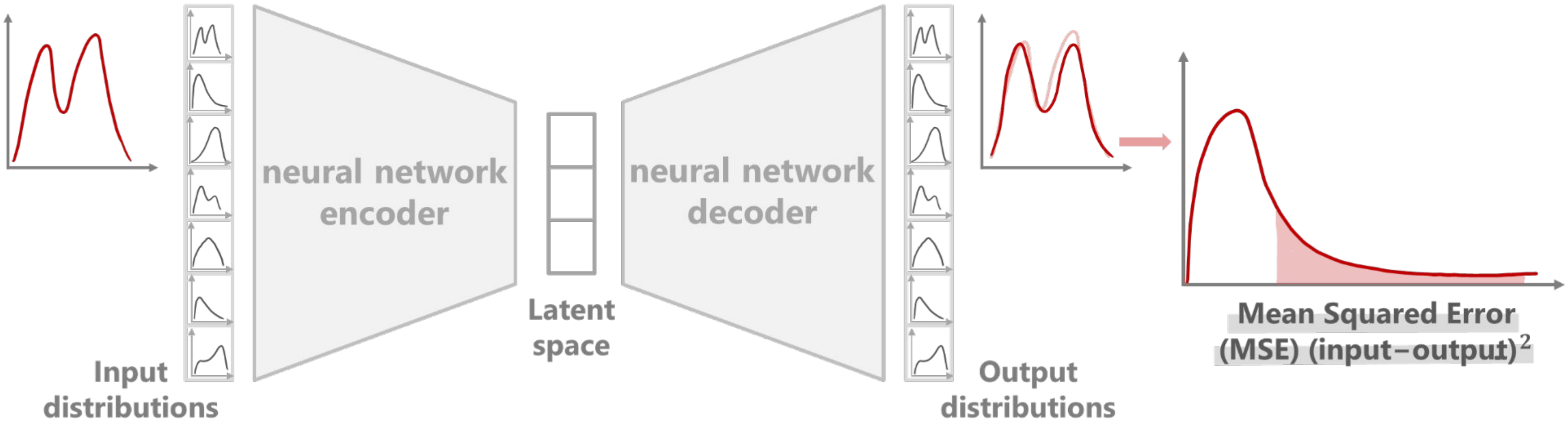
AutoEncoders

Encoder (DNN): operates a dimensionality reduction (maps inputs to latent space)

+

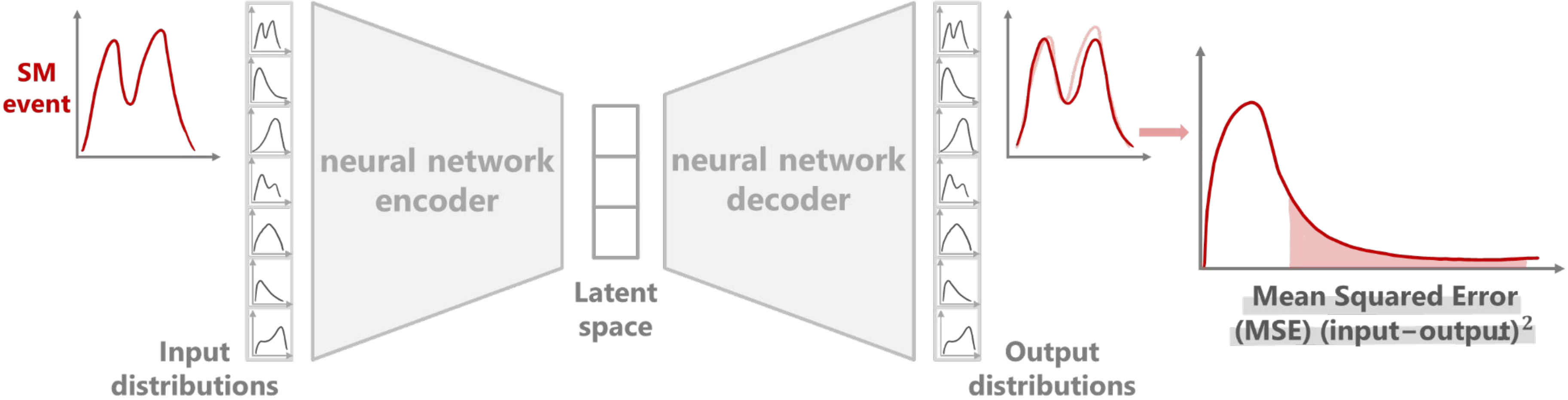
Decoder (DNN): maps the latent points back to the input space

- Trained via minimization of a "reconstruction loss" e.g. MSE = (in-out)^2



AutoEncoders for Anomaly Detection

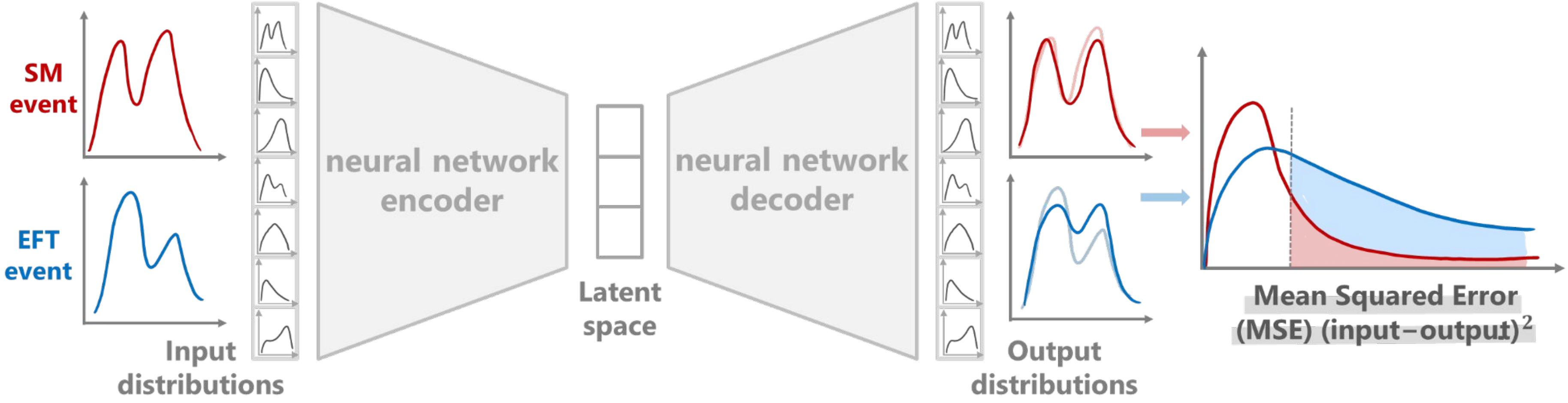
- **Trained only on SM physics** (no knowledge of new physics!)



AutoEncoders for Anomaly Detection

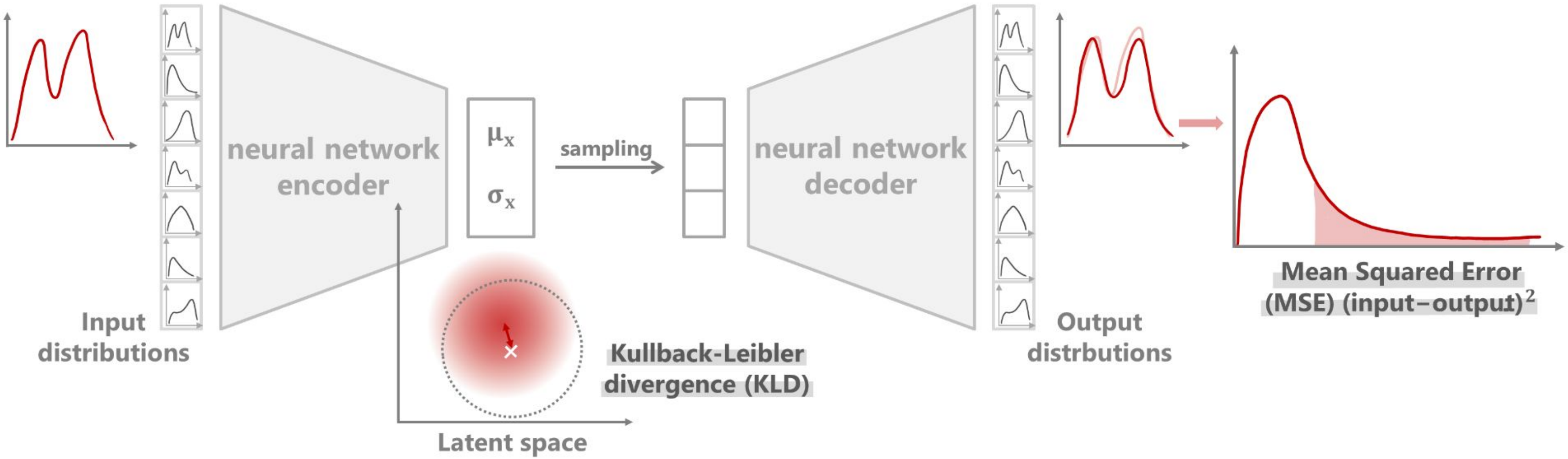
- **Trained only on SM physics** (no knowledge of new physics!)
- **Later run on BSM contributions**: BSM events are reconstructed worse

→ **anomalies lie in the tails of the loss function**



Variational AutoEncoders

- The latent space is forced to be regular, namely described by a multidimensional **gaussian distribution**
  - via minimization of a **regularization loss (KLD)** + **reconstruction loss (MSE)**
- A point is sampled from the latent space and decoded

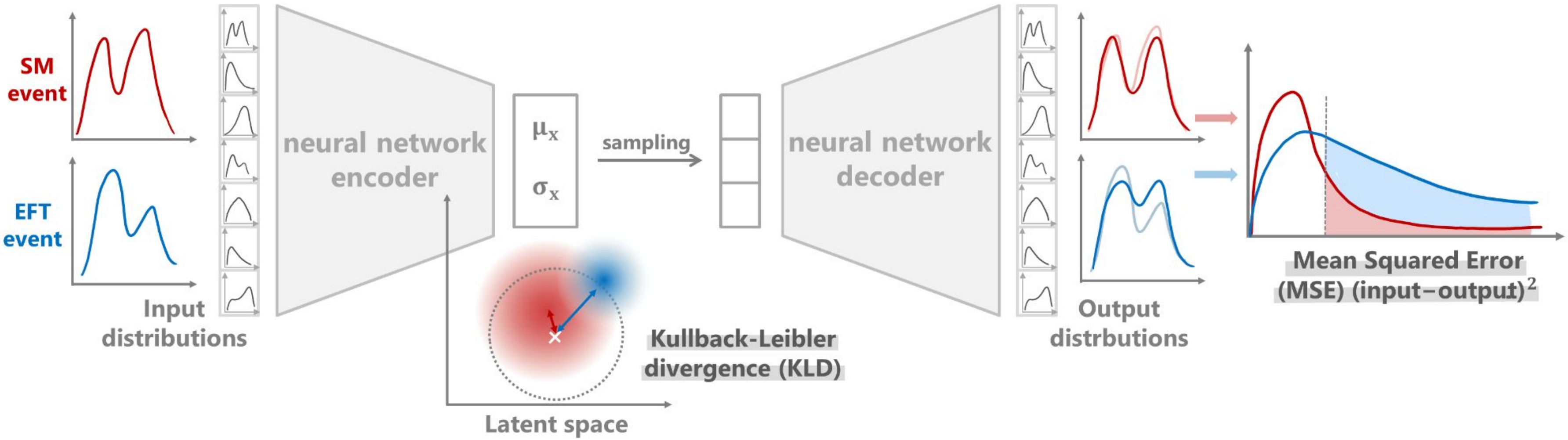


Variational AutoEncoders for Anomaly Detection

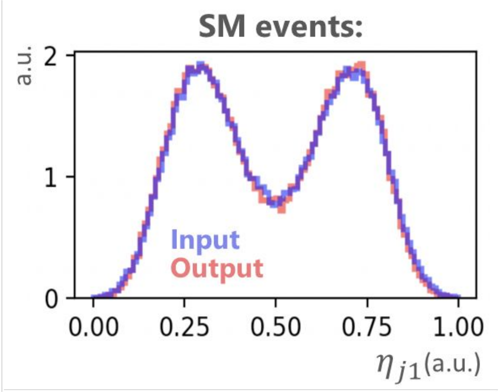
Generative model: it **learns to decode samples drawn from the same probability distribution of the original dataset**

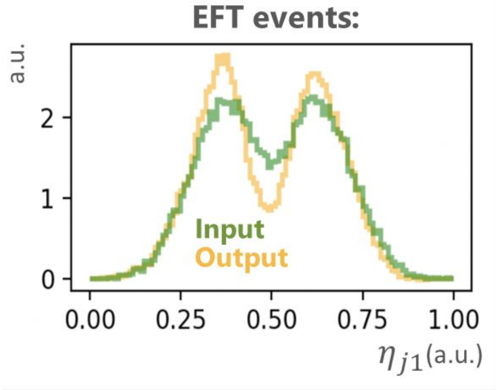
→ noise reduction, content generation
→ anomaly detection

More **robust** and **variation-tolerant** AD strategy compared to AEs

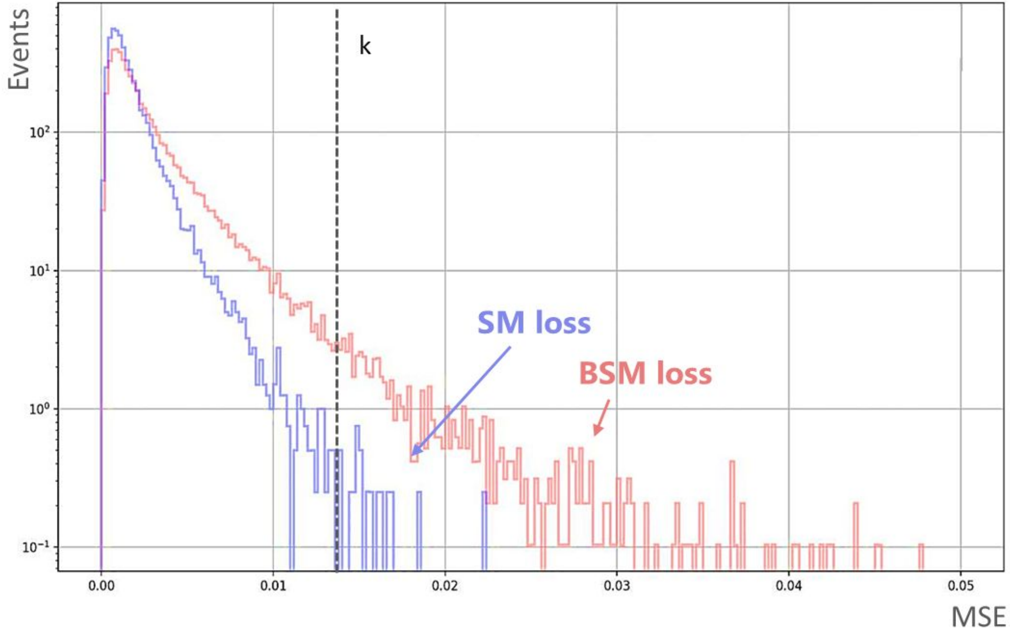


VAE performances





EFT events are reconstructed worse than SM ones, and end up in the tails of the loss function (as expected!)



Adding a supervised NN classifier to the VAE

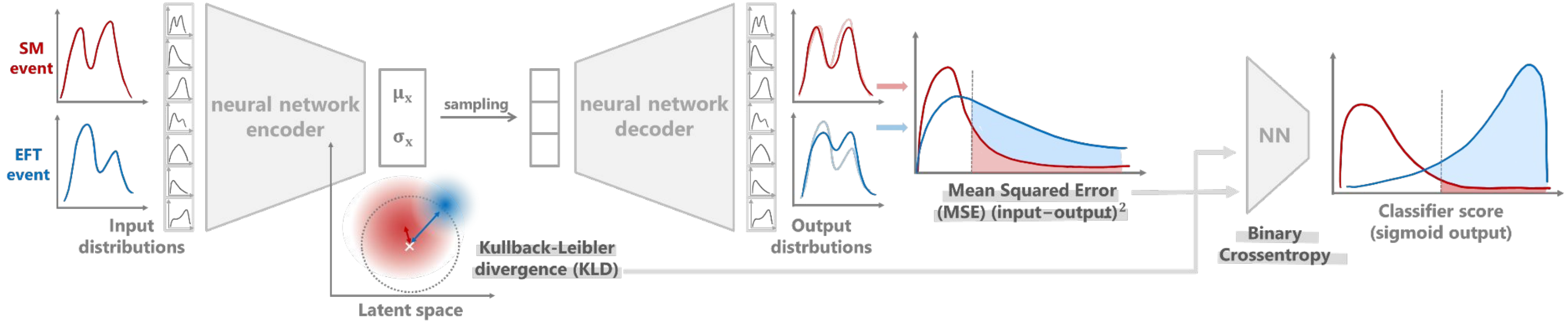
The VAE is only trained to reconstruct a SM sample, while our goal is to isolate EFT events.

We want to **embed discrimination in the training**

Adding a supervised NN classifier to the VAE

The VAE is only trained to reconstruct a SM sample, while our goal is to isolate EFT events.

We want to **embed discrimination in the training** → **VAE + NN classifier**

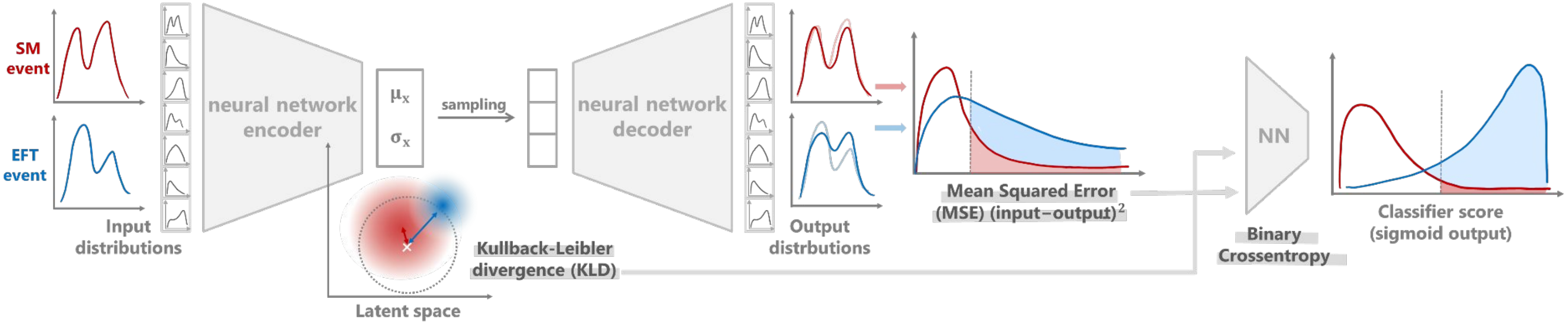


- Trained by minimization of MSE + KLD + **Binary Crossentropy**
- Input data are divided between purely SM and SM + EFT:
  - MSE and KLD coming from **SM** events are added to the model loss
  - MSE and KLD coming from a set of **SM+EFT** events are given to the classifier
  - the binary crossentropy is added to the model loss

Adding a supervised NN classifier to the VAE

The VAE is only trained to reconstruct a SM sample, while our goal is to isolate EFT events.

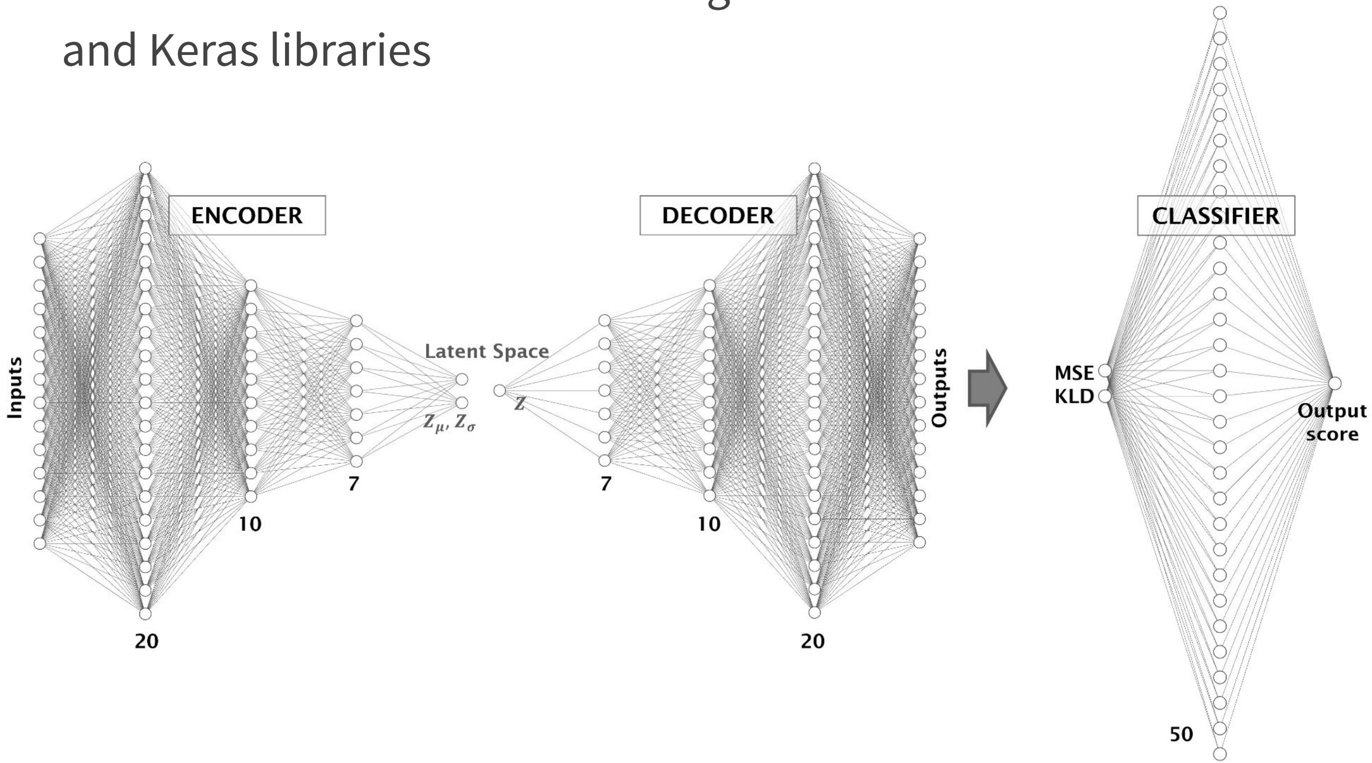
We want to **embed discrimination in the training** → **VAE + NN classifier**



- More effective SM-EFT discrimination
- Reduce the model independence
    - Keep the model as general as possible by using a single EFT operator during the training

VAE + classifier: the model

The model is built via subclassing on Tensorflow and Keras libraries

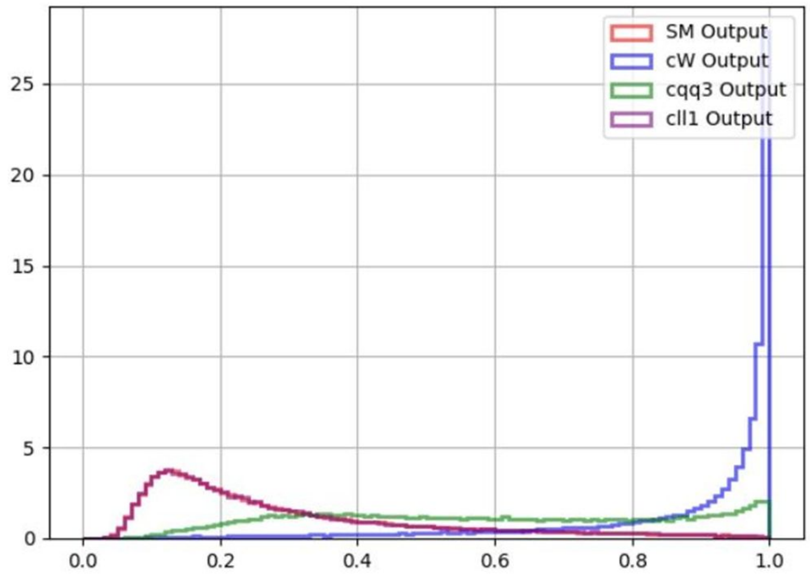


- ➔ Deeply connected layers (keras Dense layers)
- ➔ Leaky ReLU (VAE)
- ➔ Hard Sigmoid (classifier)
- ➔ optimizer: ADAM
- ➔ Epochs: 100 for convergence
- ➔ Batch size: 32/64

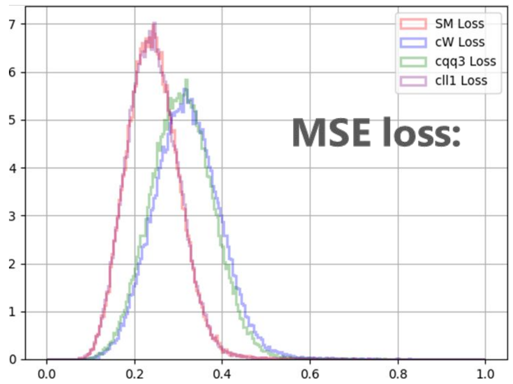
Results

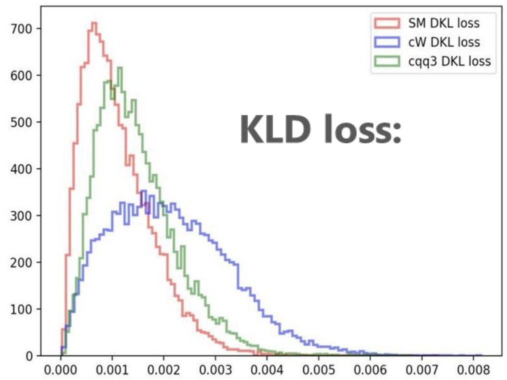
The model is able to discriminate between SM and EFT events:

Classifier score:



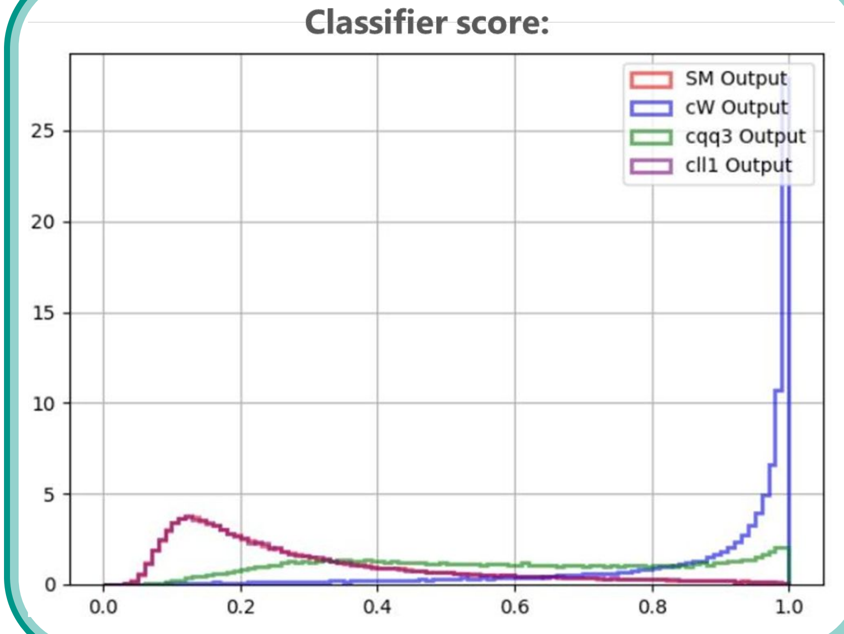
- SM events are reconstructed better than EFT
- Discrimination between SM and EFT is best for the operator the model was trained on (cW)
- Other operators are also recognized
- Some operators are not singled out (shapes similar to SM)



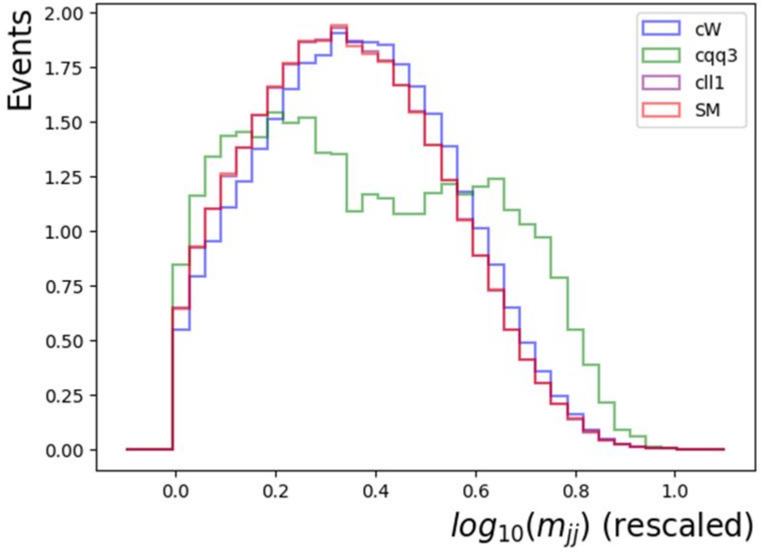


Results

The model collects information from various inputs → it provides a variable (**output score**) whose **shape maximizes the separation between EFT and SM**



(e.g. wrt a **simple kinematic variable**)



Defining a metric

The model is sensitive to various different operators: we define a proxy metric for the significance σ, which depends on the Wilson coefficients of the operator considered during testing:

$$\sigma(c_{op}) = \frac{|BSM(c_{op}) - SM|}{\sqrt{SM}} = \frac{|LIN(c_{op}) + QUAD(c_{op}^2)|}{\sqrt{SM}}$$

We consider the model sensitive to an operator if σ reaches the value of 3:

operator	c_W	c_{qq}^{1}	$c_{qq}^{1,1}$	c_{qq}^{3}	$c_{qq}^{3,1}$	c_{Hq}^{1}	c_{HW}
$c_{op} : \sigma(c_{op}) = 3$	0.13	0.17	0.18	0.11	0.11	0.61	0.65

→ The smallest this value, the more sensitive the VAE to the operator!

Conclusions and future perspectives

- The strategy **allows to isolate EFT contributions in a mostly model independent way**
- The goal is to isolate a signal-enriched region, on which further and more specific analyses can be performed

Further steps:

- Test of the strategy on **fully reconstructed events** and application to **Run 2 data**
- Inclusion of the background processes (mainly due to QCD production and fake leptons)