ANOMALY DETECTION & EVENT GENERATION USING (V)AE MODELS

LUC HENDRIKS

RADBOUD UNIVERSITY, NIJMEGEN (NL)

OUTLINE

- Short introduction me & ML
- Unsupervised collider searches using (beta-)(V)AEs
  - Latent spaces
  - Combined models
- Generative models as event generators
  - B-VAE

ABOUT ME

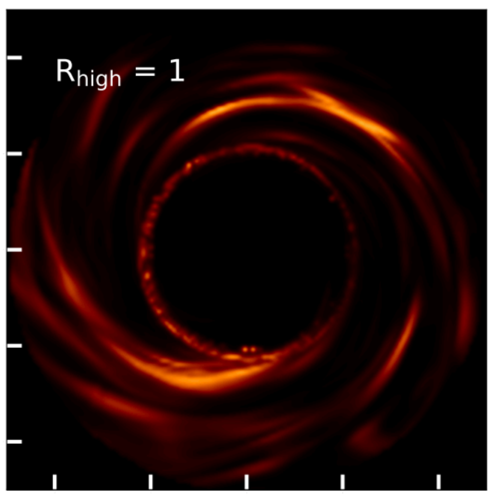
- Finishing my PhD in machine learning applied to dark matter related problems
  - At HEP in Radboud University Nijmegen (supervisor: Sascha Caron)
- Cofounded Resnap that automatically creates photo books using AI (acquired by Albelli 3 years ago)
- Main topics around 50% HEP and 50% astroparticle physics

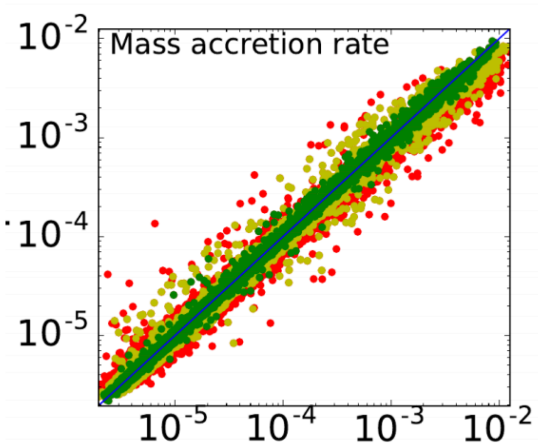
https://www.linkedin.com/in/luchendriks/ - list of projects

ABOUT ME

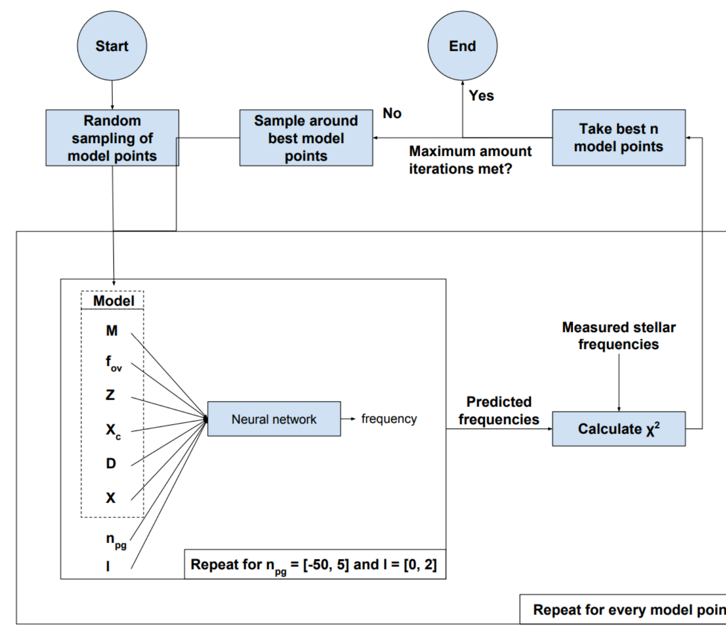
- Subset of astroparticle physics projects

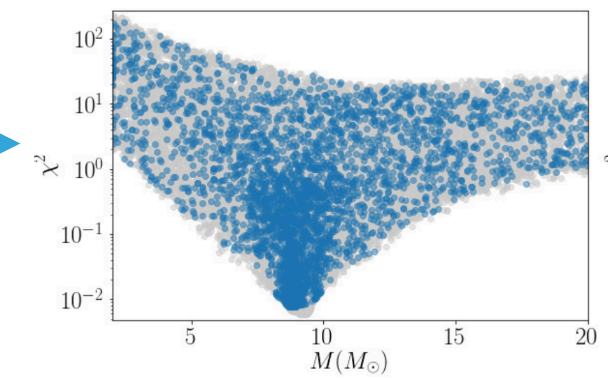
Bayesian deep learning on EHT simulations
https://arxiv.org/abs/1910.13236 (A&A)



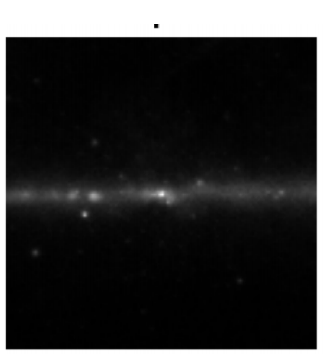


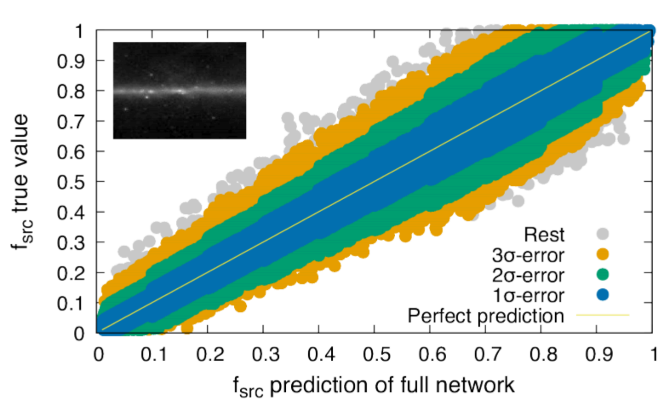
Deep learning applied on asteroseismological data
https://arxiv.org/abs/1811.03639 (PASP)



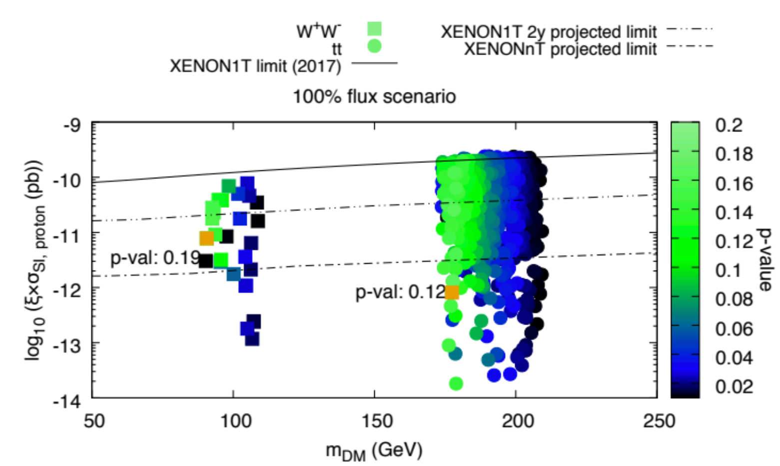


Parameter inference on γ-rays from the GC
https://arxiv.org/abs/1708.06706 (JCAP)
(follow ups almost done: PS detection & BDL)

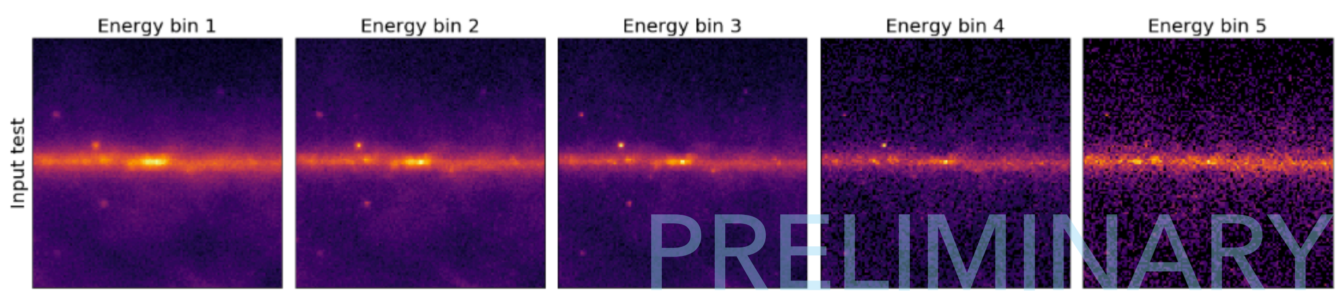


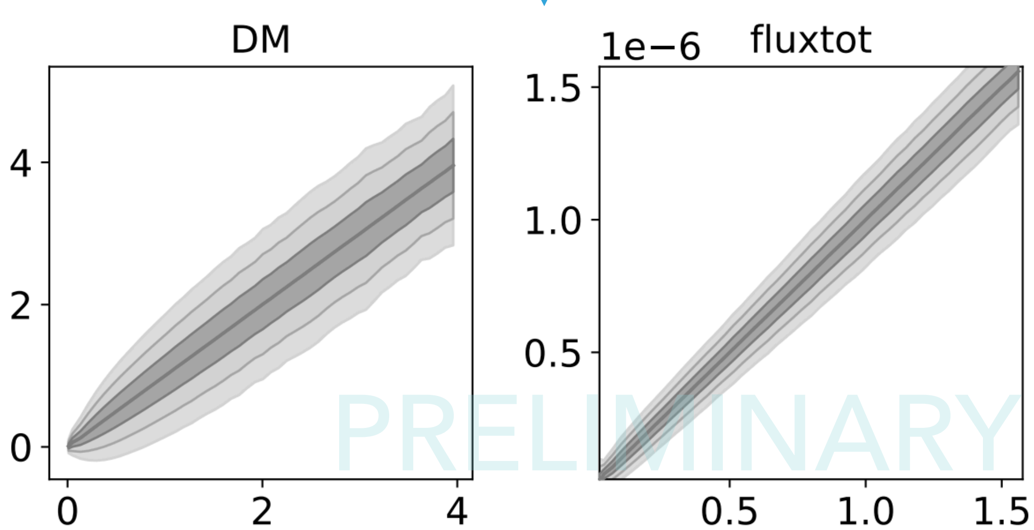


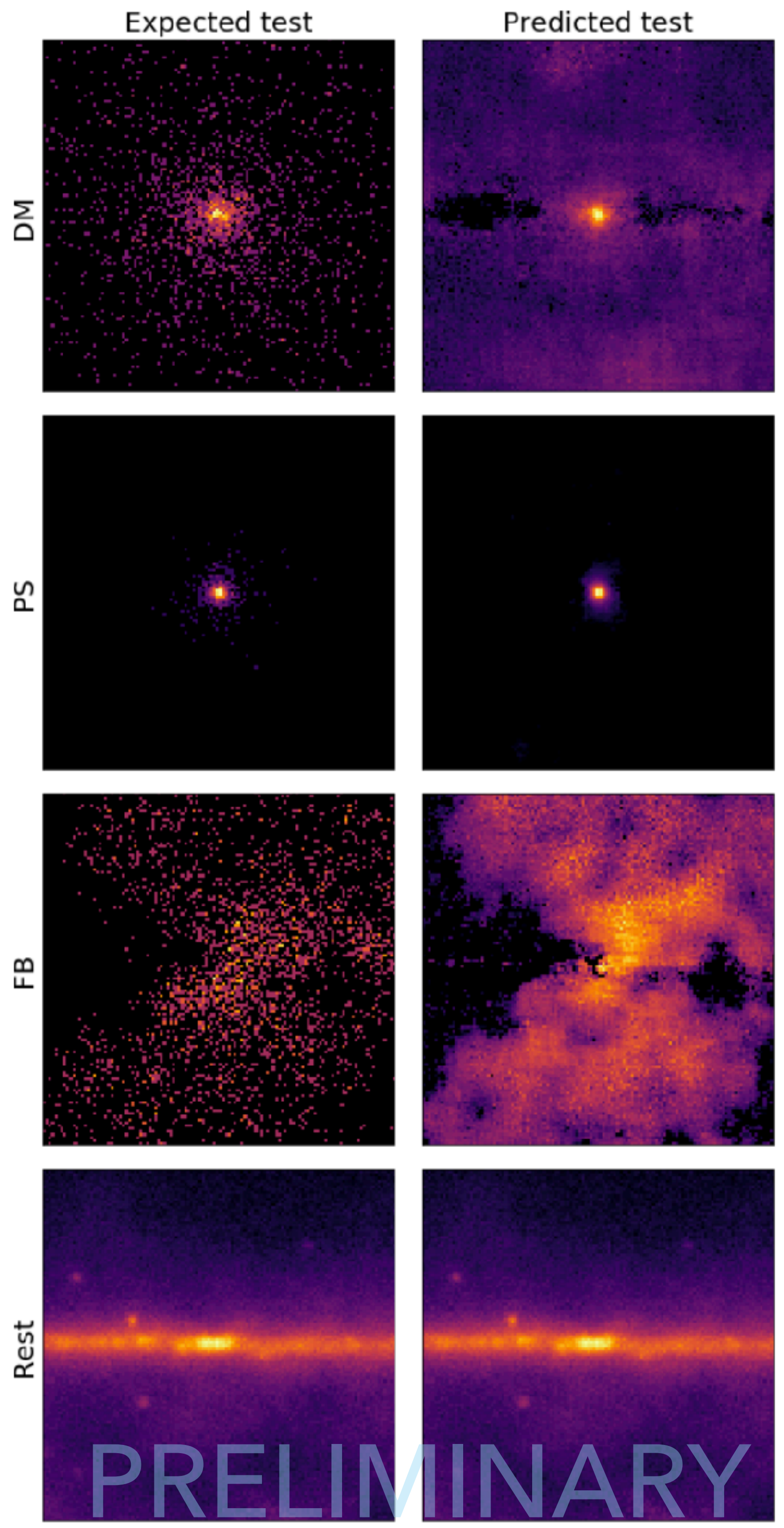
Likelihood fit of pMSSM models on GC excess
https://arxiv.org/abs/1709.10429 (JCAP)
follow up of master thesis - featured in nature news



- Subset of astroparticle physics projects







INTRODUCTION

- ML methods are “discipline-independent”

MULTIDISCIPLINARY

- ML methods are “discipline-independent”
- When you strip away the (astro)physics, almost any problem becomes a data science problem
  - Interpret satellite data -> computer vision
  - Finding new physics in particle collisions -> anomaly detection
  - Gravitational wave detection -> time-series analysis
  - …

MULTIDISCIPLINARY

- ML methods are "discipline-independent"
- When you strip away the (astro)physics, almost any problem becomes a data science problem
  - Interpret satellite data -> computer vision
  - Finding new physics in particle collisions -> anomaly detection
  - Gravitational wave detection -> time-series analysis
  - …
- Methods from one field can be re-used in another field
- eg DarkMachines was founded with this in mind (darkmachines.org): experts in one field can contribute their methods in another

UNSUPERVISED COLLIDER SEARCHES

- Topic: unsupervised collider searches

UNSUPERVISED COLLIDER SEARCHES

- Topic: unsupervised collider searches
- One approach: theory finds a particle candidate, then experiment tries to find or exclude it

UNSUPERVISED COLLIDER SEARCHES

- Topic: unsupervised collider searches
- One approach: theory finds a particle candidate, then experiment tries to find or exclude it
- Curse of dimensionality!
- Many models, every models has many parameters – you can keep search until the end of the universe

UNSUPERVISED COLLIDER SEARCHES

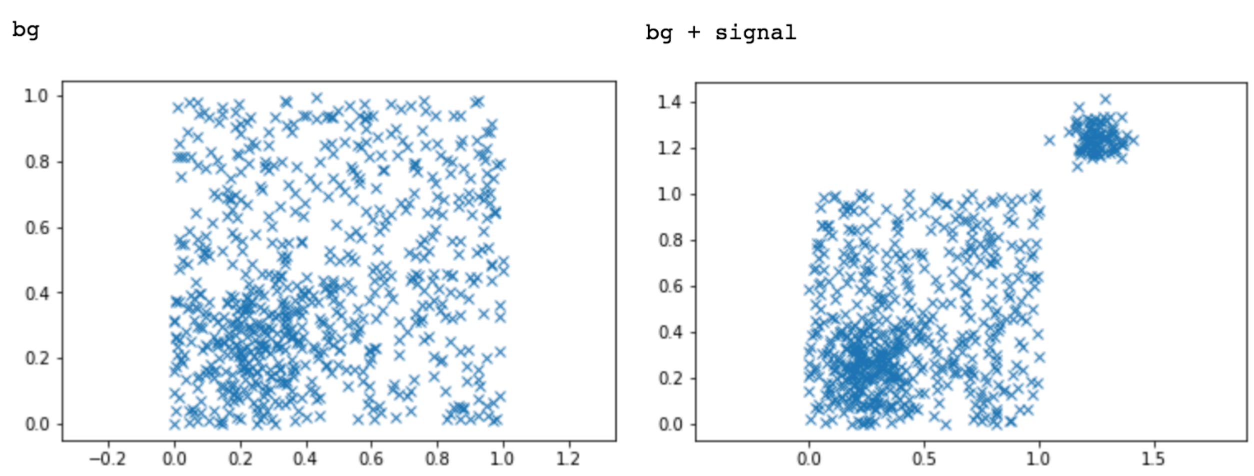
- Topic: unsupervised collider searches
- One approach: theory finds a particle candidate, then experiment tries to find or exclude it
- Curse of dimensionality!
- Many models, every models has many parameters – you can keep search until the end of the universe
- Alternative approach:
  - The experiment records data
  - Compare with expectation from only SM hypothesis
  - If rejected -> look at the events that reject that hypothesis and try to explain

UNSUPERVISED COLLIDER SEARCHES

- Topic: unsupervised collider searches
- One approach: theory finds a particle candidate, then experiment tries to find or exclude it
- Curse of dimensionality!
- Many models, every models has many parameters – you can keep search until the end of the universe
- Alternative approach:
  - The experiment records data
  - Compare with expectation from only SM hypothesis
  - If rejected -> look at the events that reject that hypothesis and try to explain
  - (=unsupervised search of new physics)

UNSUPERVISED COLLIDER SEARCHES

- Typical setup of the experiment:
  - Compare experiment data (real data) to expected data from only SM (simulated data)
  - Real data contains SM plus possible, but unknown, signal
  - Two datasets:
    - SM only (from simulation) – background
    - SM + possible signal (from real data) – background + signal



UNSUPERVISED COLLIDER SEARCHES

- For evaluating performance, simulate also signals and pretend you don't know. Gives two datasets:
  - Train on SM only simulated data
  - Test on SM+signal simulated data

- For evaluating performance, simulate also signals and pretend you don't know. Gives two datasets:
  - Train on SM only simulated data
  - Test on SM+signal simulated data
- Counting experiment:
  - From SM only hypothesis you expect λ events
  - You measure k events

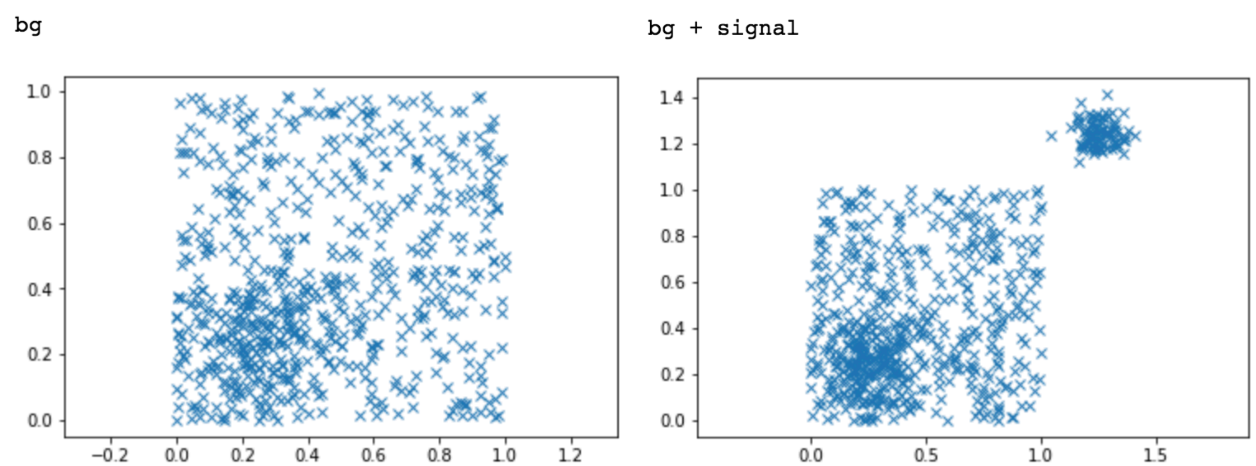
$$f(k; \lambda) = \Pr(X = k) = \frac{\lambda^k e^{-\lambda}}{k!}$$

  - (simplified)
- Filter the data such that you "cut" out the SM so only signal is left, using only SM as your training data

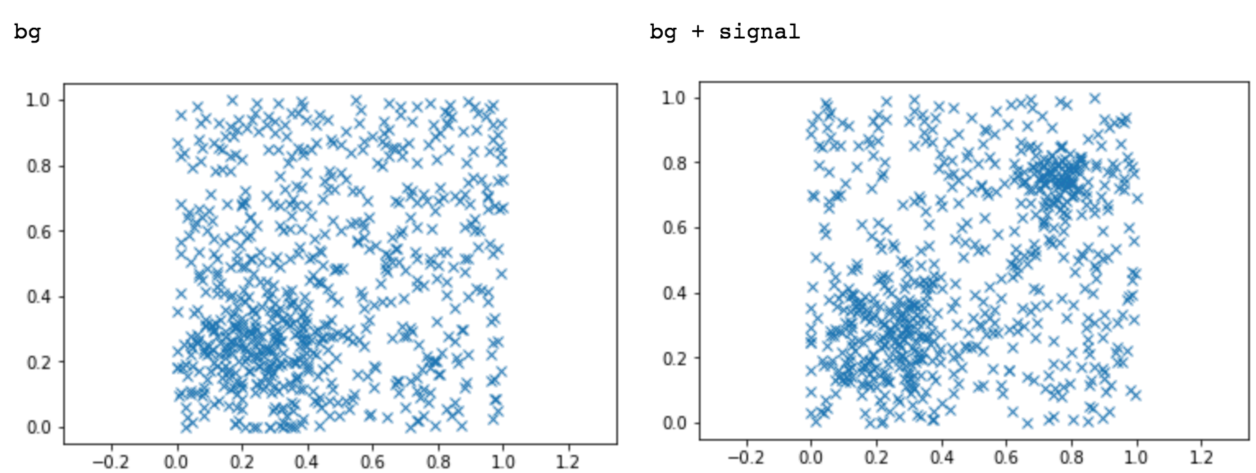
UNSUPERVISED COLLIDER SEARCHES

Two types of signals:

Outlier detection



Density estimation

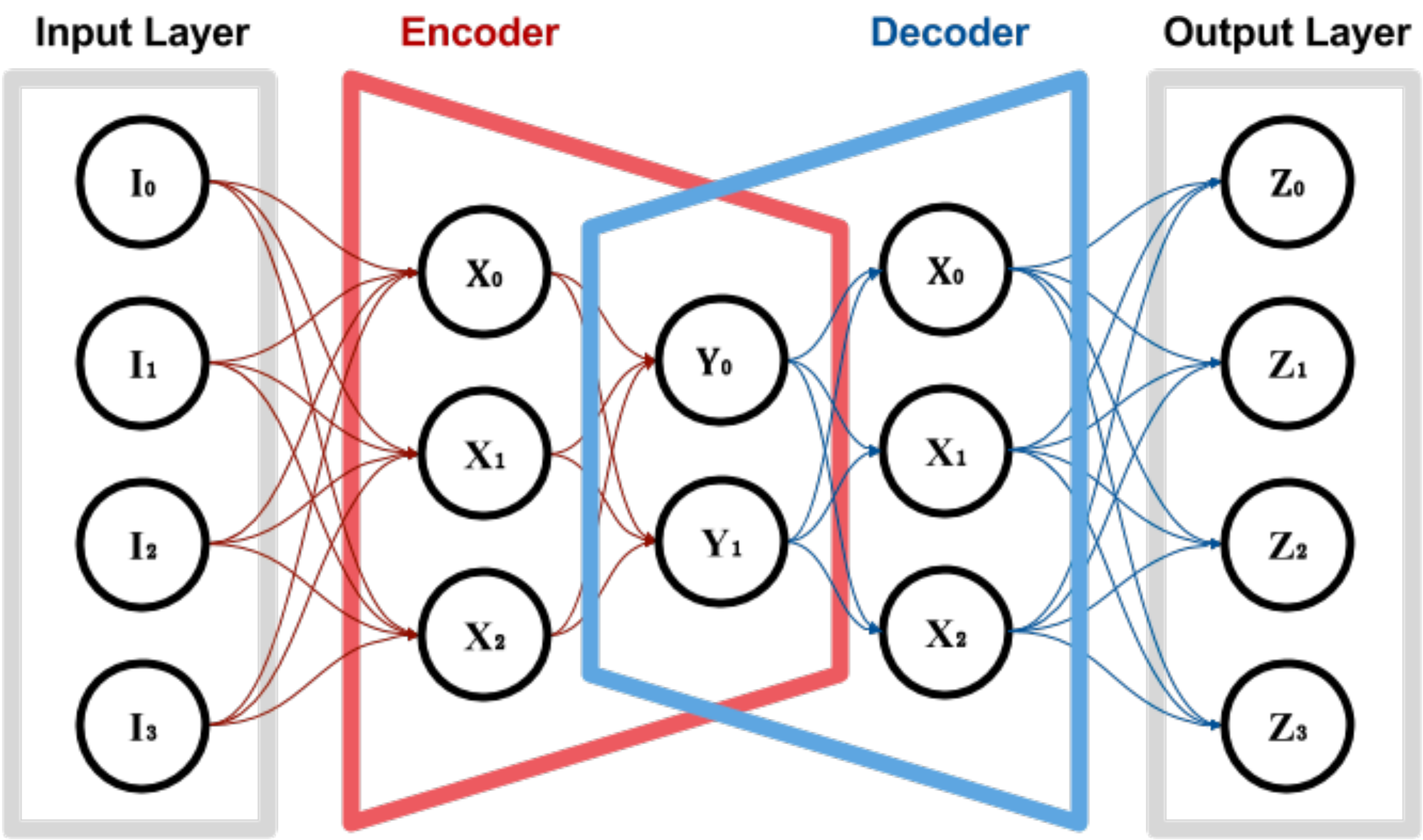


UNSUPERVISED COLLIDER SEARCHES

- Focus on outlier detection
- Huge field with many methods, focus on autoencoders & latent spaces

UNSUPERVISED COLLIDER SEARCHES

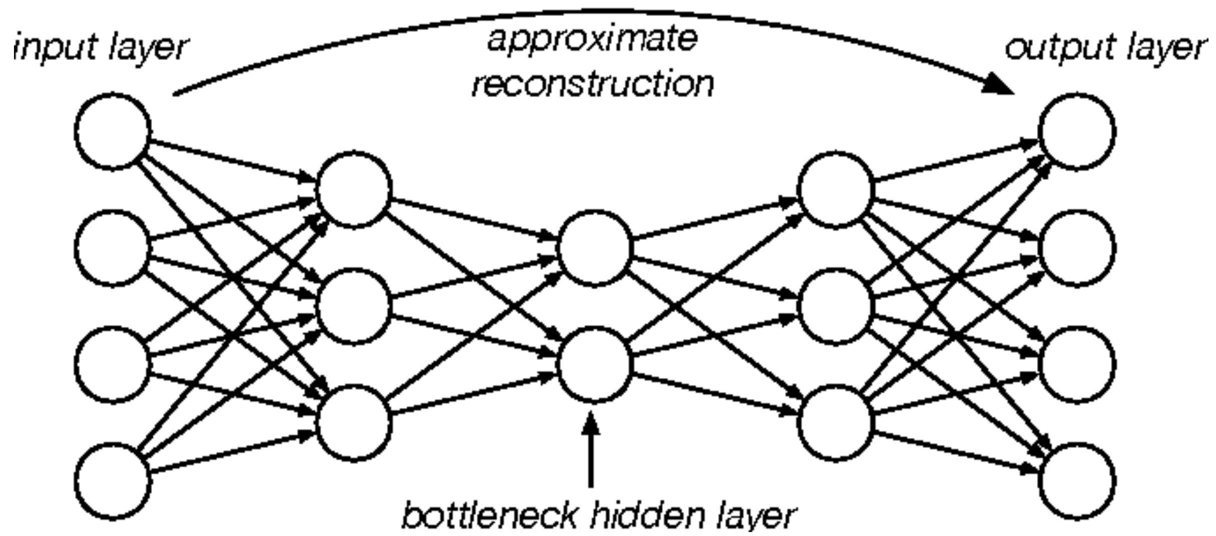
- Focus on outlier detection
- Huge field with many methods, focus on autoencoders & latent spaces

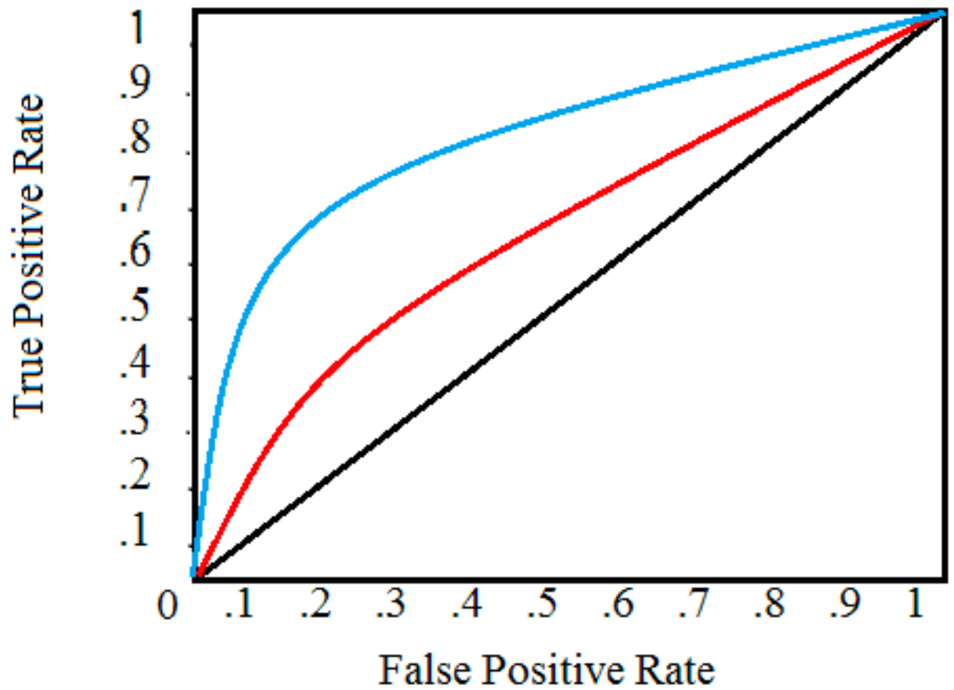


Anomaly score typically is normalised reconstruction loss (eg MSE)

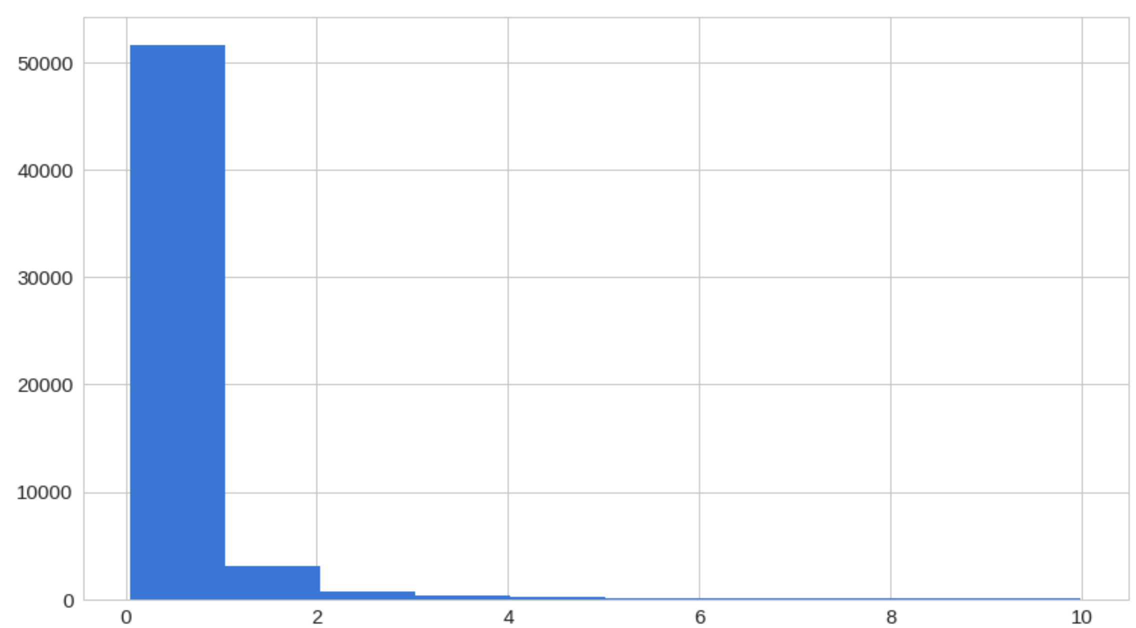
UNSUPERVISED COLLIDER SEARCHES

- Example: credit card fraud detection with autoencoder

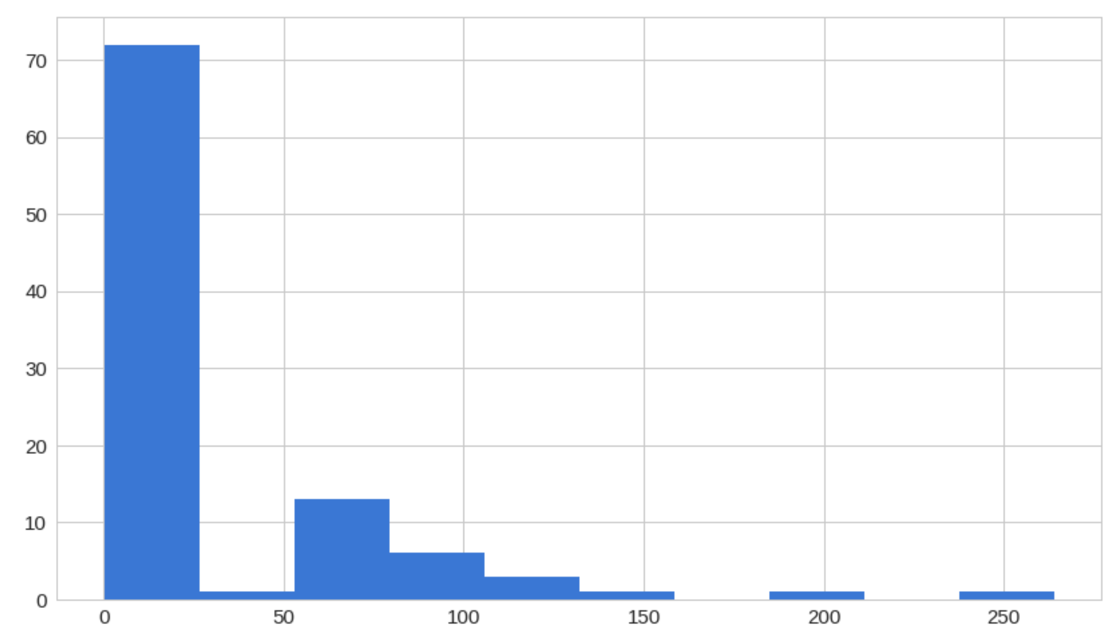




No fraud



Fraud



https://medium.com/@curiousily/credit-card-fraud-detection-using-autoencoders-in-keras-tensorflow-for-hackers-part-vii-20e0c85301bd

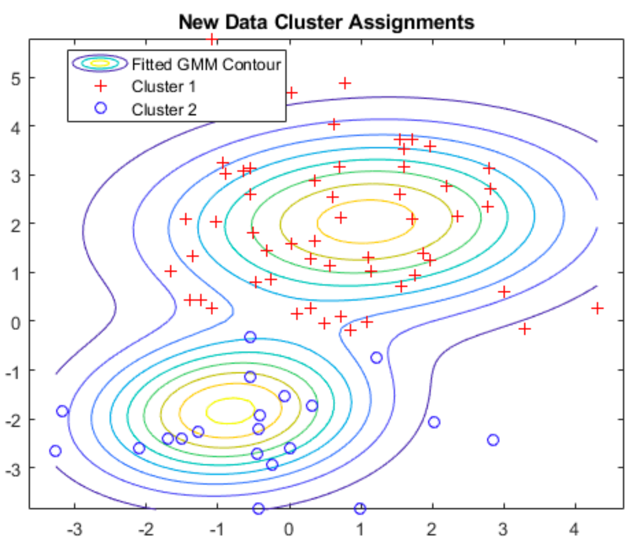
UNSUPERVISED COLLIDER SEARCHES

- Dataset: www.phenomldata.org
  - Accompanying paper: https://arxiv.org/abs/2002.12220
- Contains >30GB of simulated LHC events
- Separated in background and various signals
- Event structure:

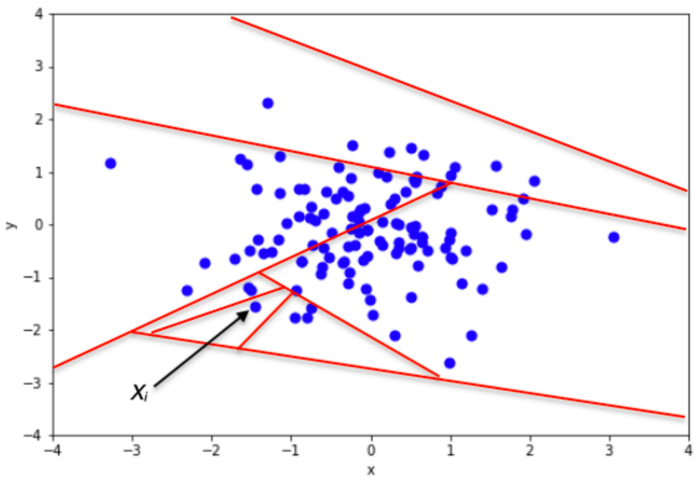
$$\mathbf{x} = \left(N, \begin{bmatrix} c_0 \\ c_1 \\ \vdots \\ c_{19} \end{bmatrix}, \begin{bmatrix} (p_T, \eta, \phi)_0 \\ (p_T, \eta, \phi)_1 \\ \vdots \\ (p_T, \eta, \phi)_{19} \end{bmatrix} \right)$$

UNSUPERVISED COLLIDER SEARCHES

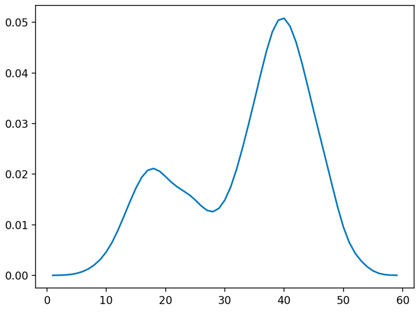
- Use different outlier detection methods and compare



GMM

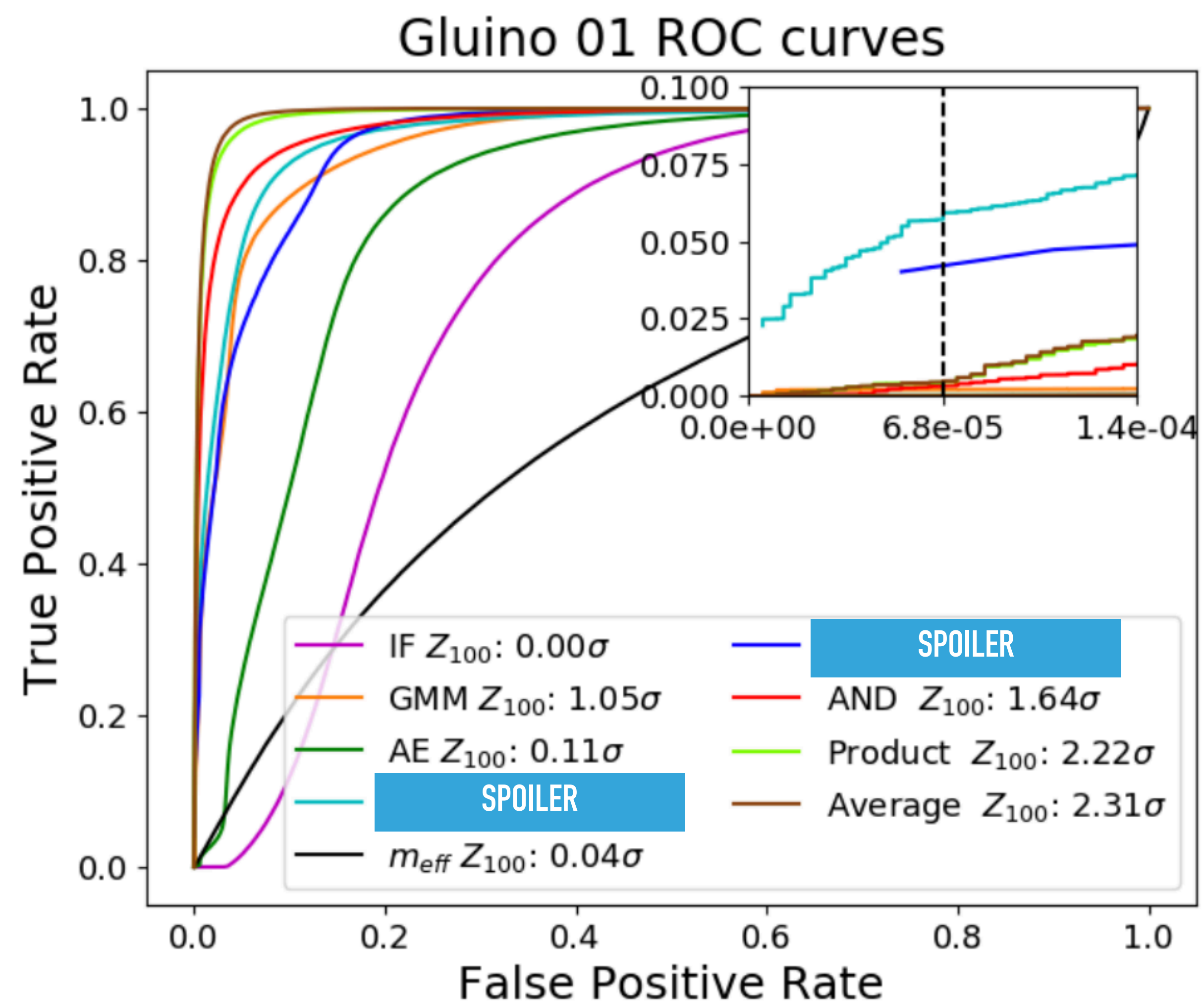


IF



UNSUPERVISED COLLIDER SEARCHES

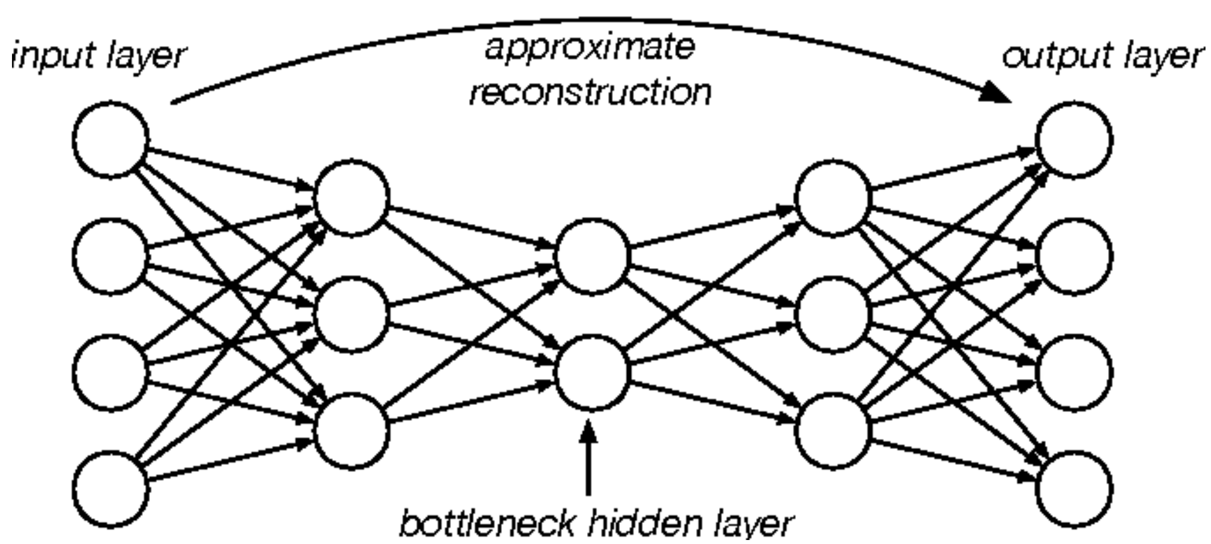
- You can use a ROC curve and the AUC to determine how well an algorithm does
- Additionally, determine signal efficiency at a predetermined background efficiency



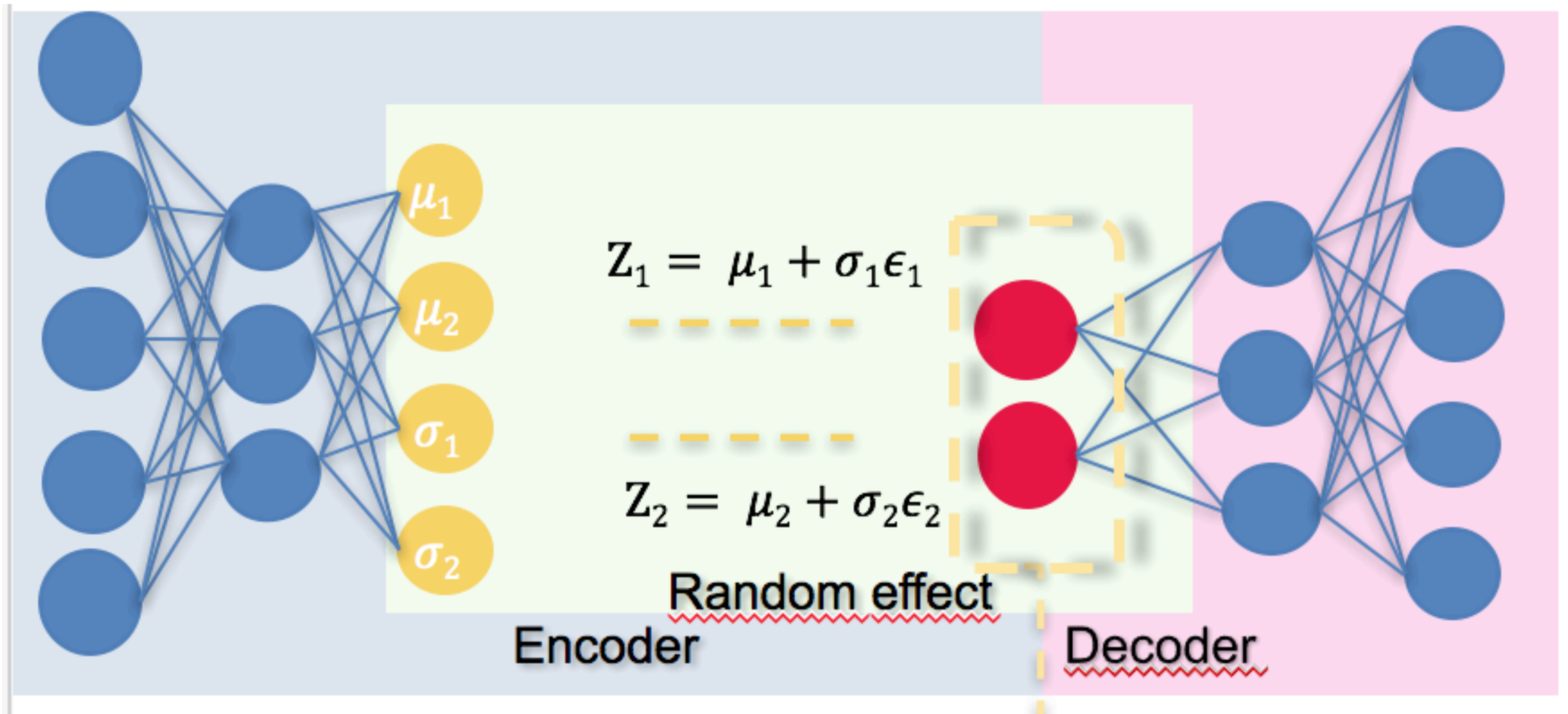
https://arxiv.org/abs/2010.07940

AUTOENCODERS

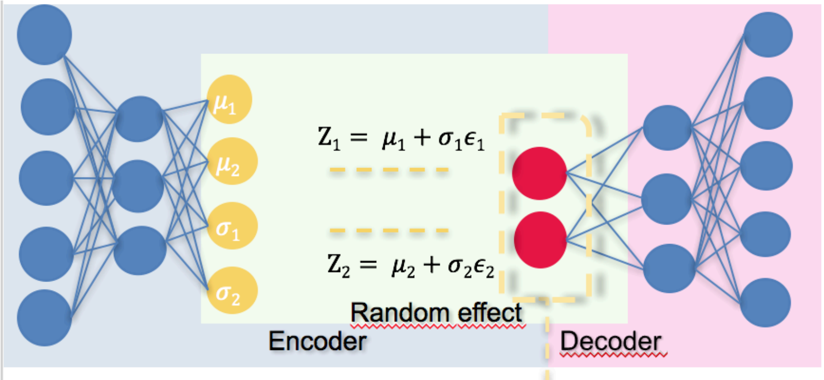
- If autoencoders are bad, why are they so popular?
- The bottleneck layer is interesting
- Transforms 4D to 2D
- Latent space



VARIATIONAL AUTOENCODERS

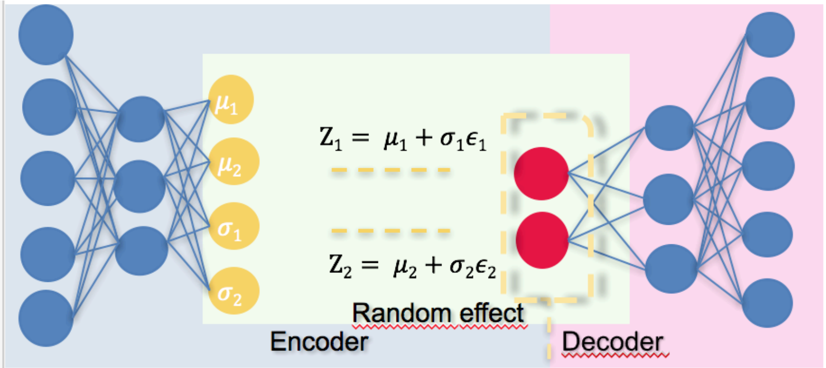


VAE

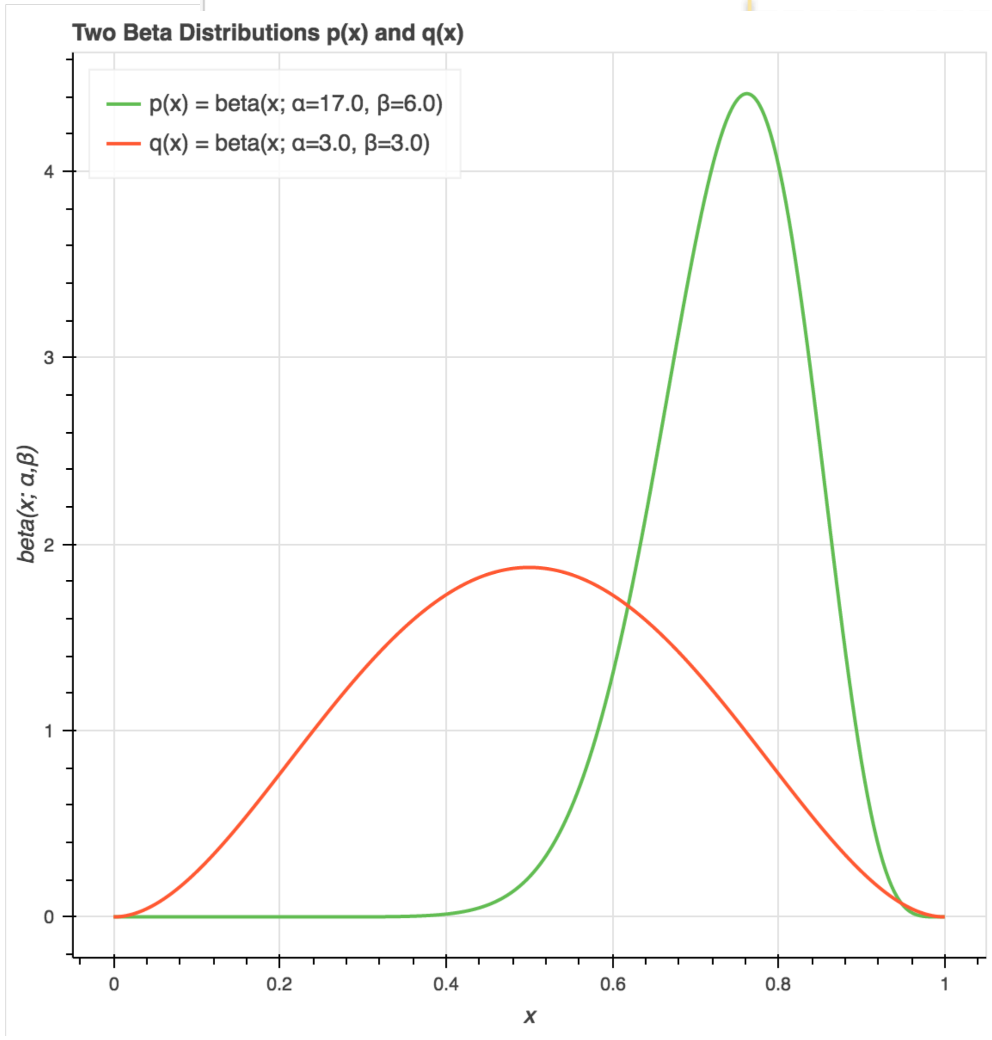


- Force ordering in latent space
- During training, you are minimising some loss function
- For regression (normal AE): MSE(*output - input*)

VAE



- Force ordering in latent space
- During training, you are minimising some loss function
- For regression (normal AE): MSE(*output - input*)
- Add KL-divergence term: $\Sigma_i \mathrm{KL}(\mathcal{N}(\mu_i, \sigma_i), \mathcal{N}(0,1)) := \mathrm{KL}(\mu,\sigma)$

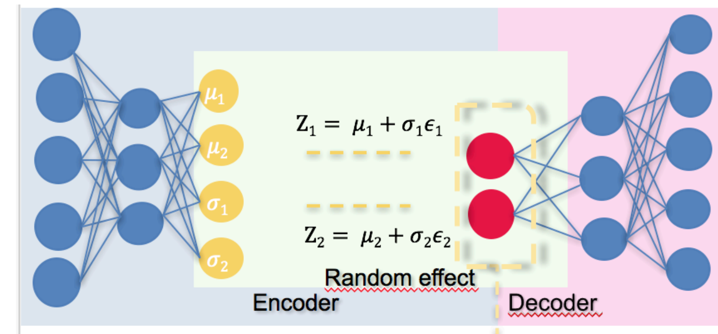


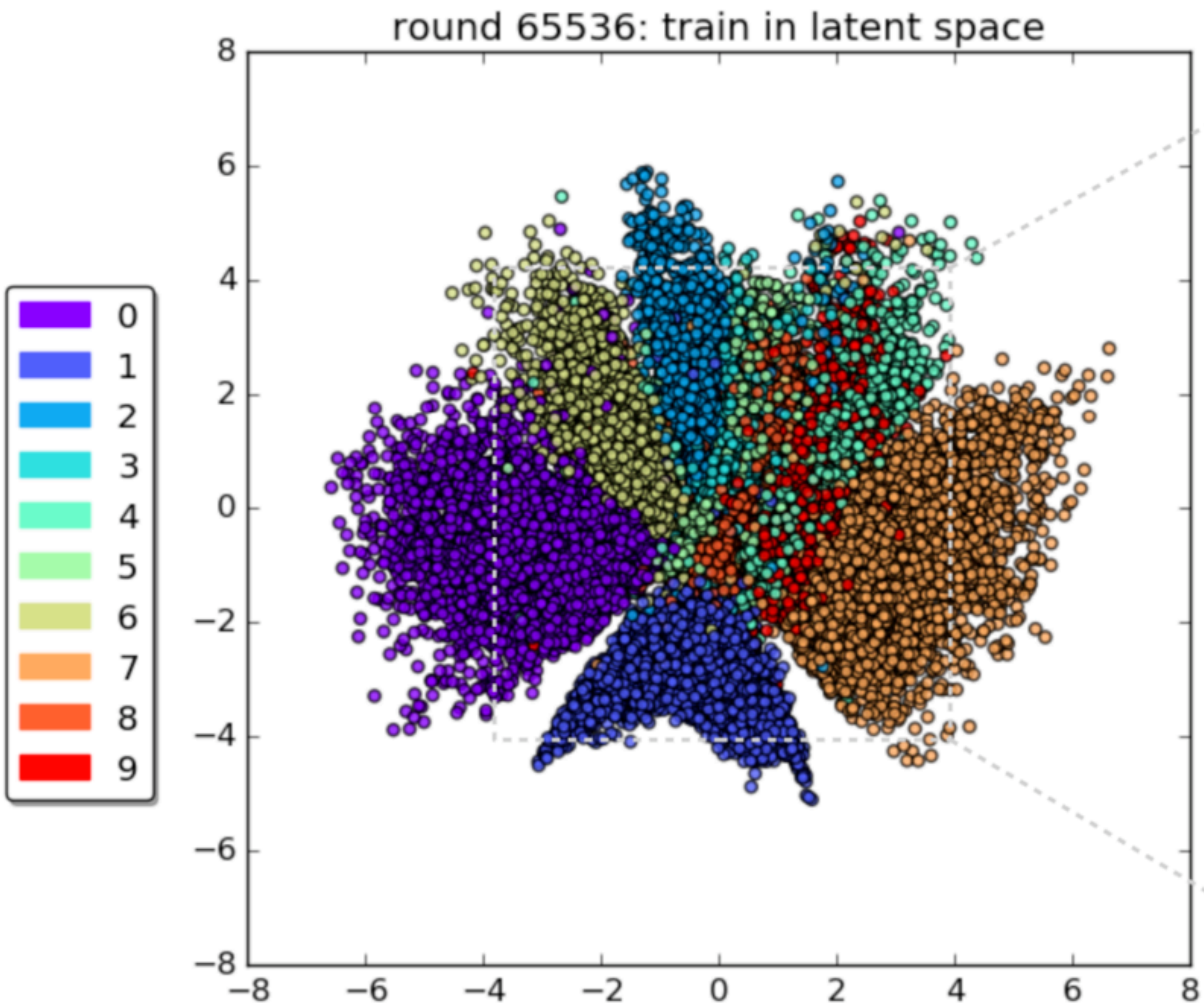
- So $\mathscr{L}$ = MSE(*output - input*) + KL(μ,σ)

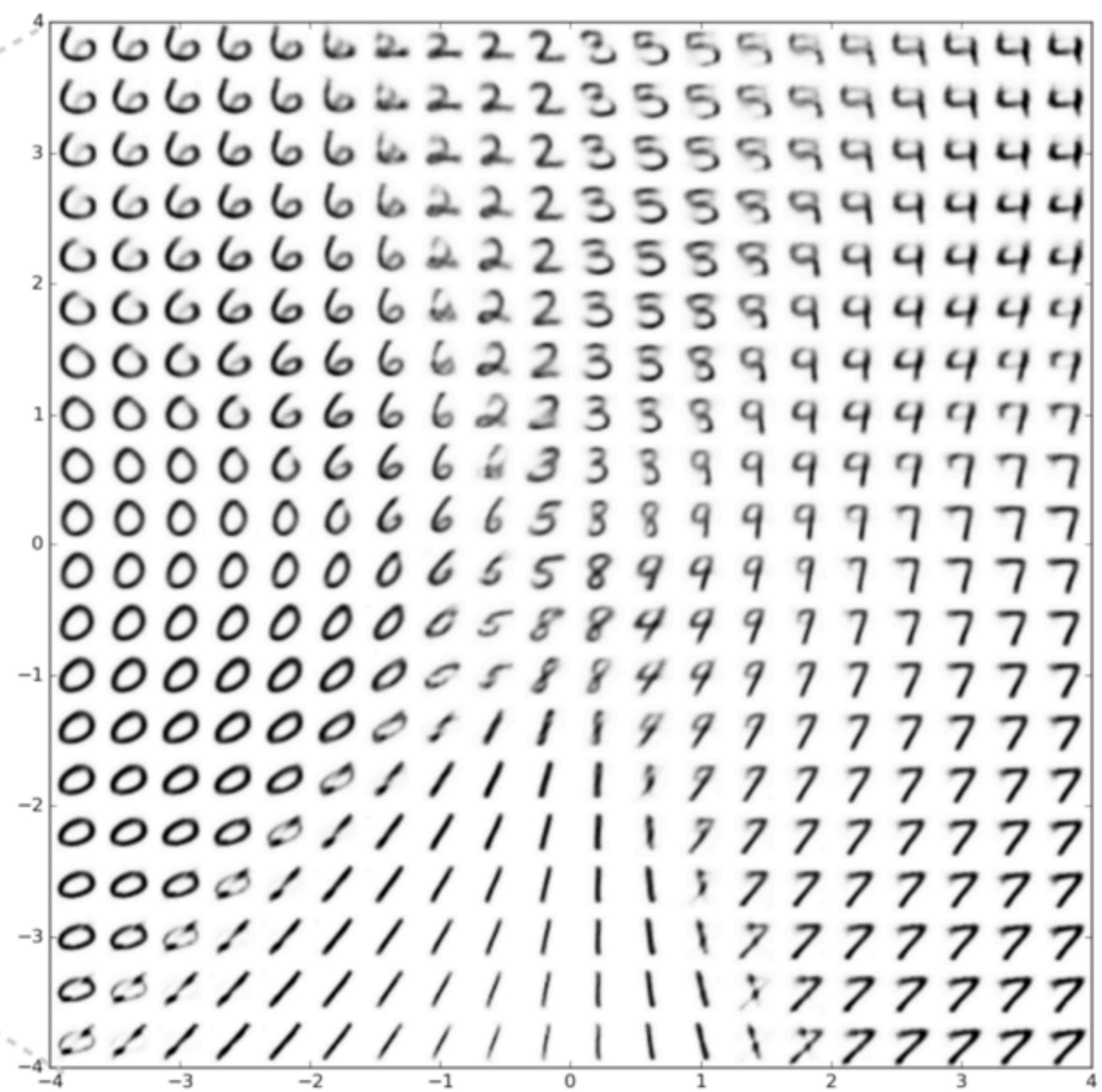
VAE

- The KL divergence punishes latent space values far away from the center
- Balance MSE and KL –> group similar structures around the center while keeping RL in check

LATENT SPACE



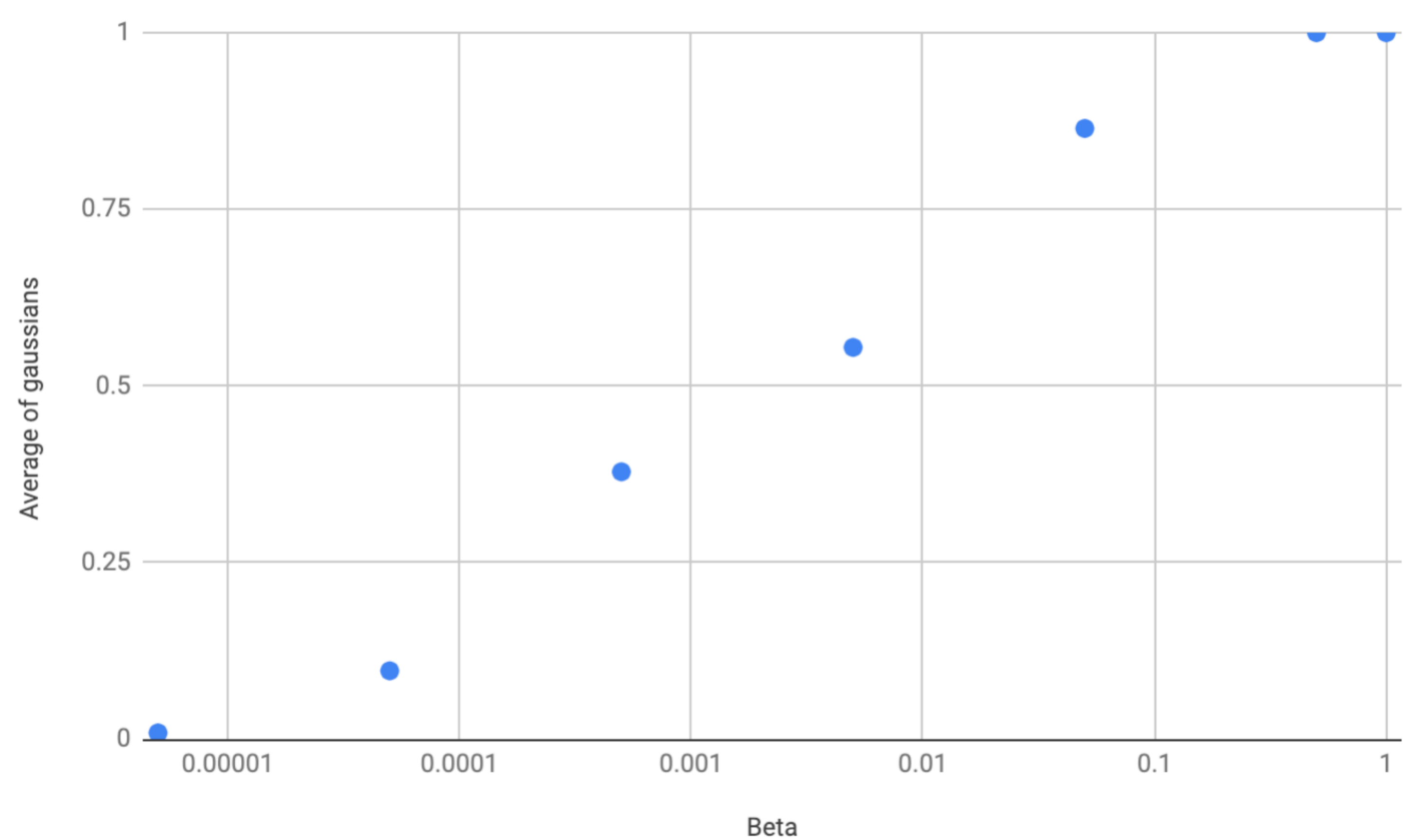




VAE

- Balancing MSE and KL is tricky
- Balance using another hyperparameter **β**
    - $\mathscr{L} = (1-\beta) * \mathrm{MSE}(\textit{output} - \textit{input}) + \beta * \mathrm{KL}(\mu, \sigma)$
- β-VAE

β	Avg var	Avg mean
1	1	1.89E-09
5E-01	0.99999905	2.35E-07
5E-02	0.86448085	…
5E-03	0.554529	
5E-04	0.3784553	
5E-05	0.09676677	
5E-06	0.008932933	
0	0.0000442	

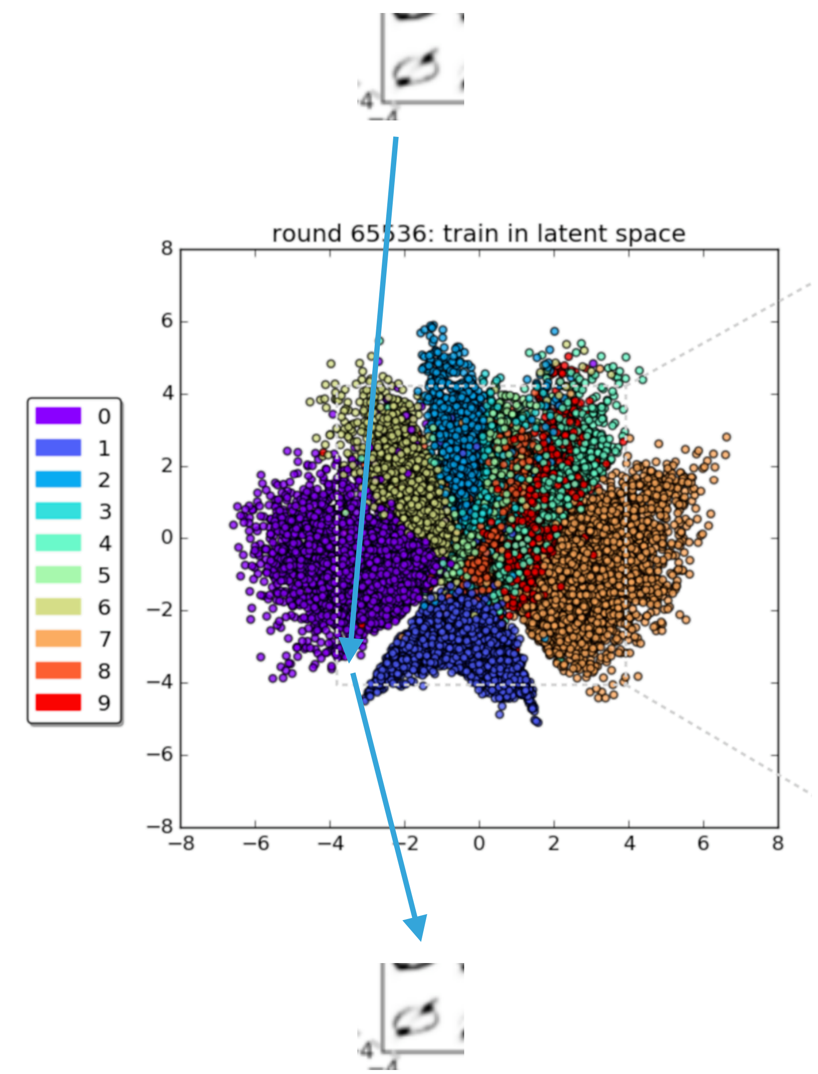


VAE

- Typical anomaly score of a VAE is the reconstruction loss
- This can be a bad anomaly variable

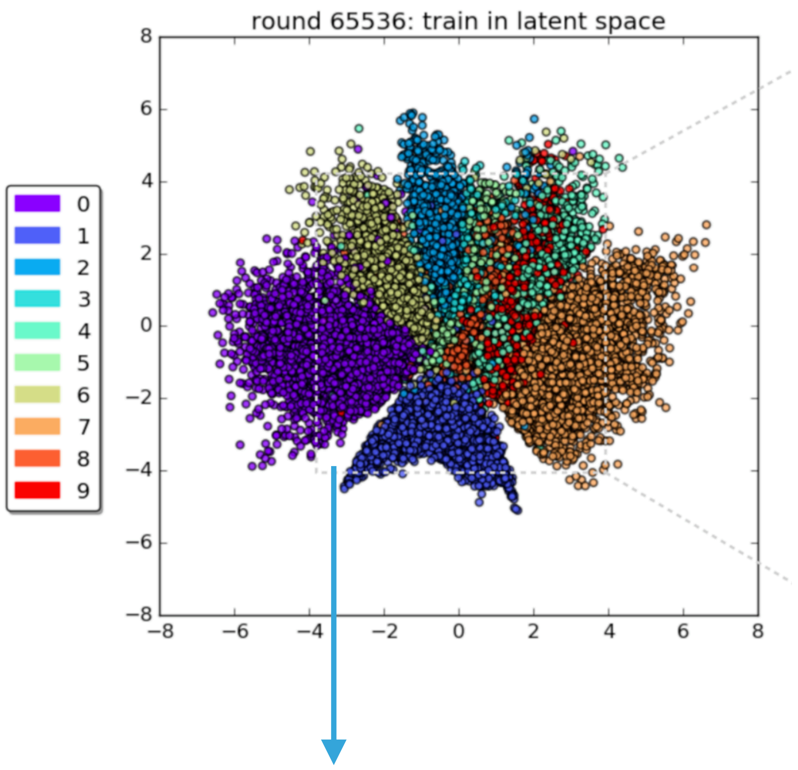
VAE

- Typical anomaly score of a VAE is the reconstruction loss
- This can be a bad anomaly variable
  - The anomalies could still be reconstructed well



VAE

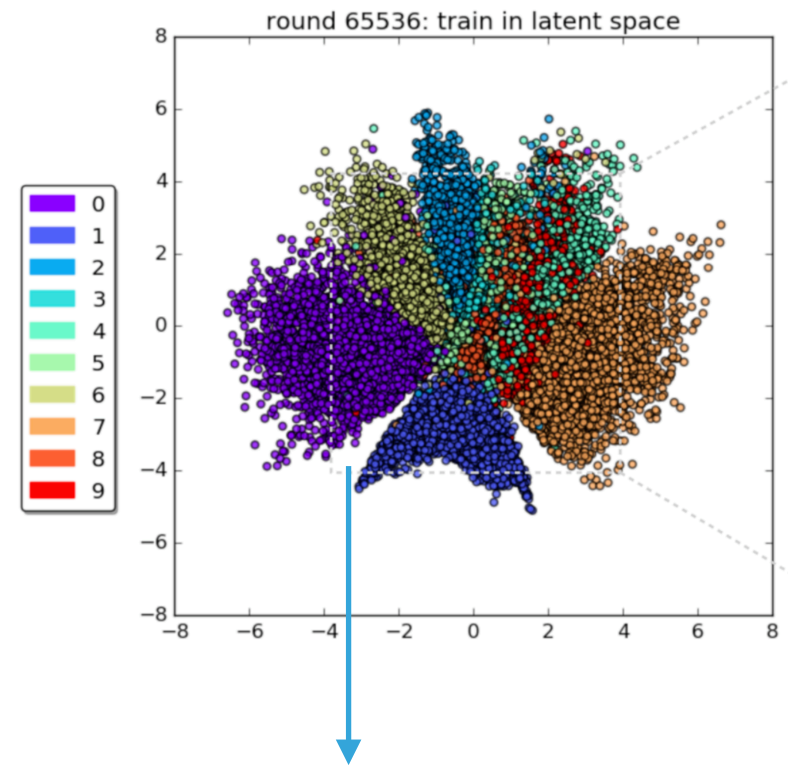
- Another approach is the likelihood of an event in a position in latent space



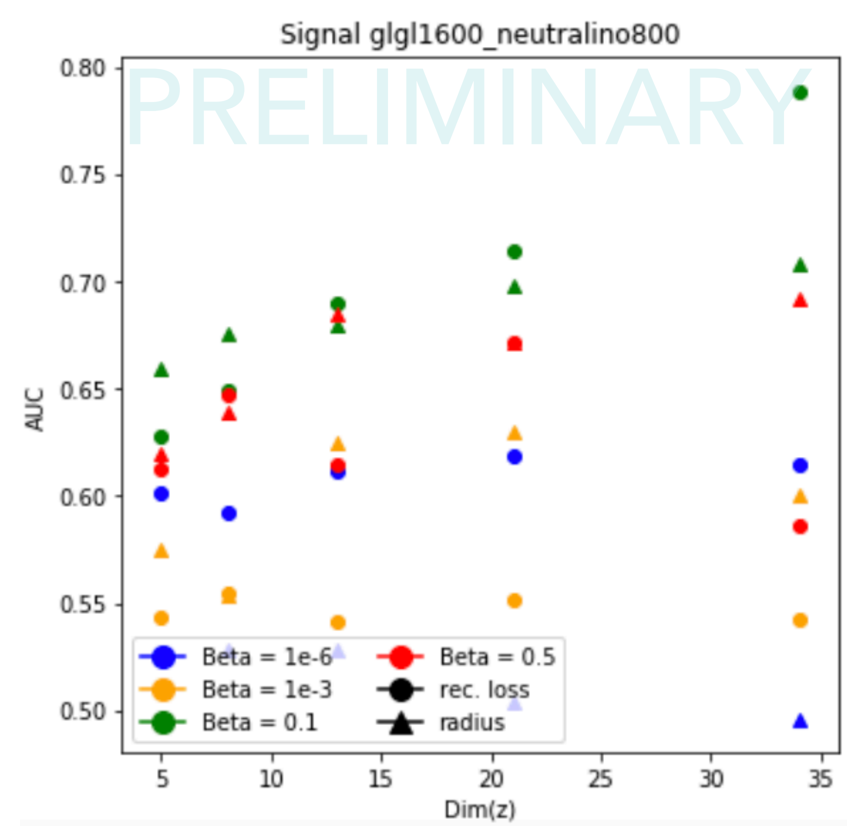
Likelihood that encoder puts a point here is low

VAE

- Another approach is the likelihood of an event in a position in latent space
    - Expensive to calculate
    - Simple approximation: radius

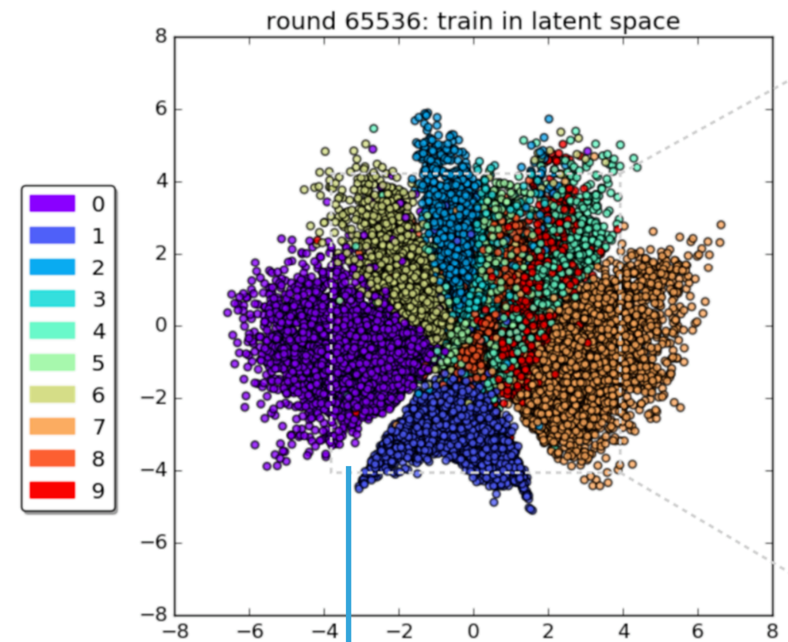


Likelihood that encoder puts a point here is low

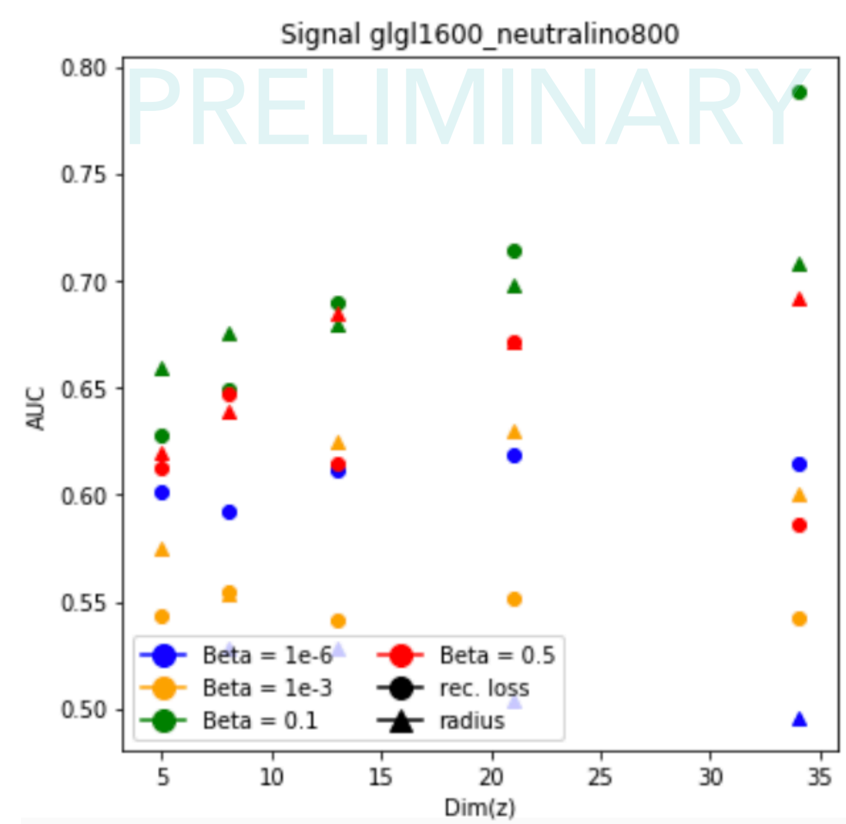


VAE

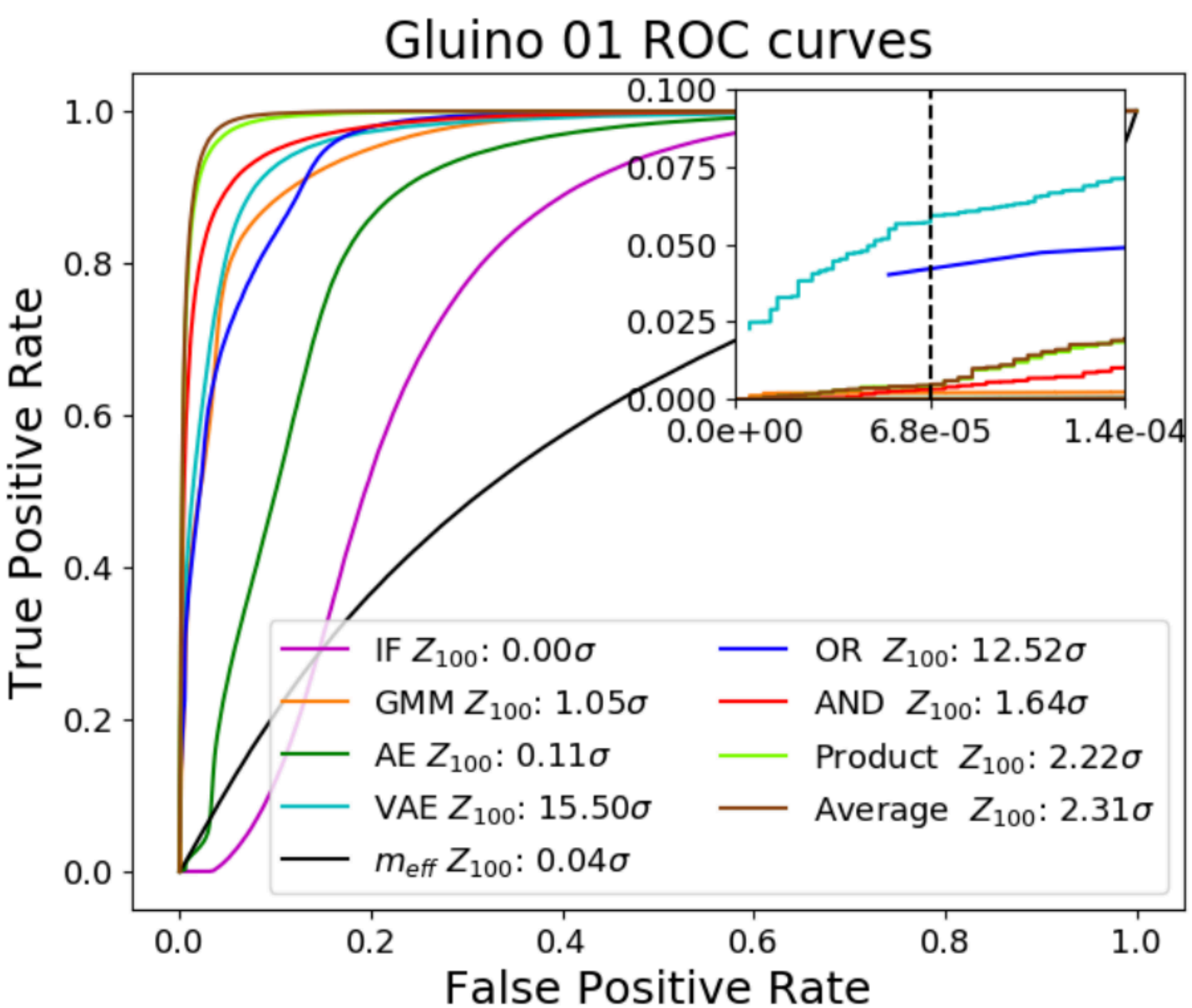
- Another approach is the likelihood of an event in a position in latent space
  - Expensive to calculate
  - Simple approximation: radius
  - Lower beta terms do worse, also for reconstruction loss
    - Even though they can reconstruct events better
    - Radius works ~equally well



Likelihood that encoder puts a point here is low



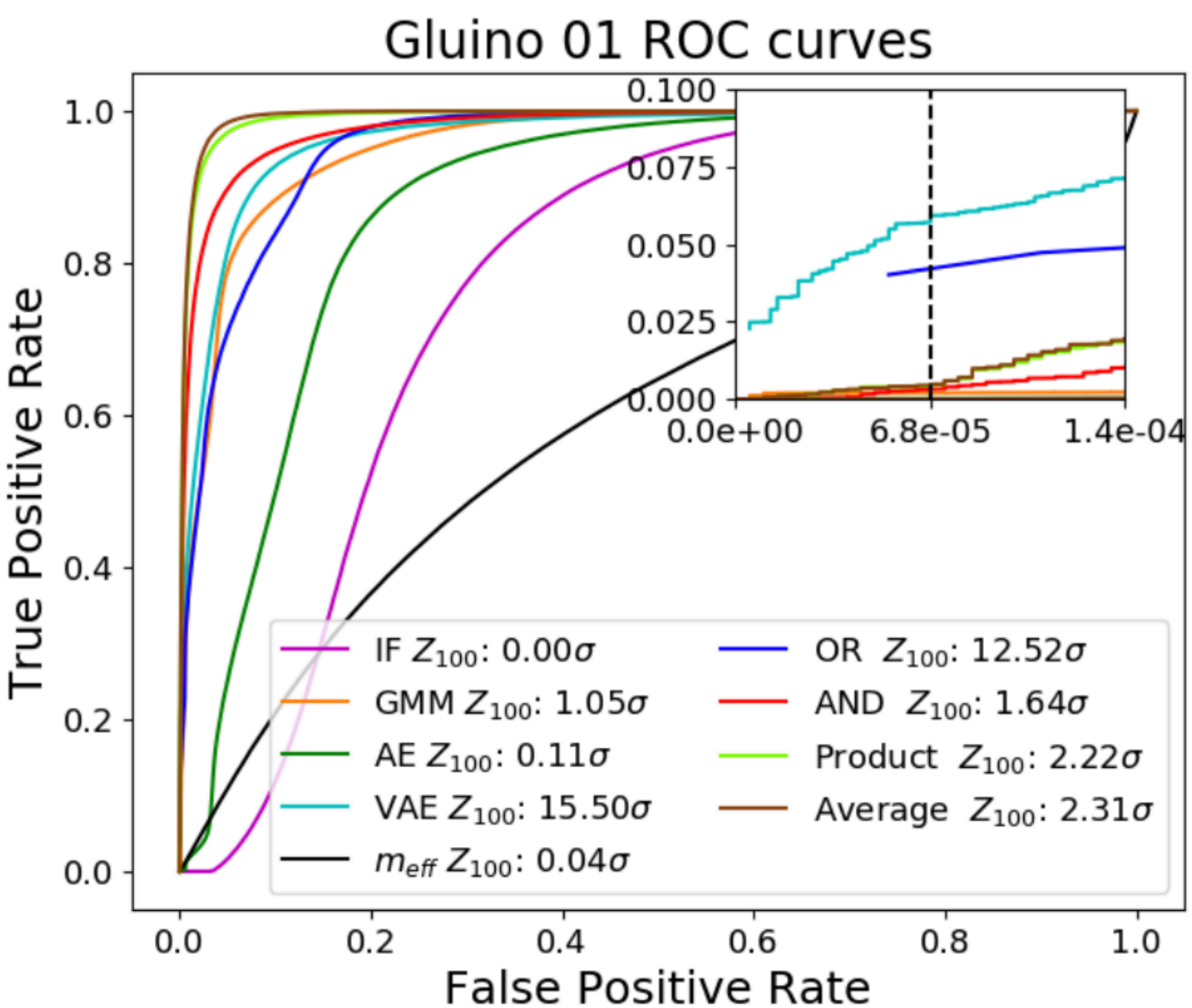
- Apply the AE/IF/GMM from before on the latent space of a VAE trained on the background events



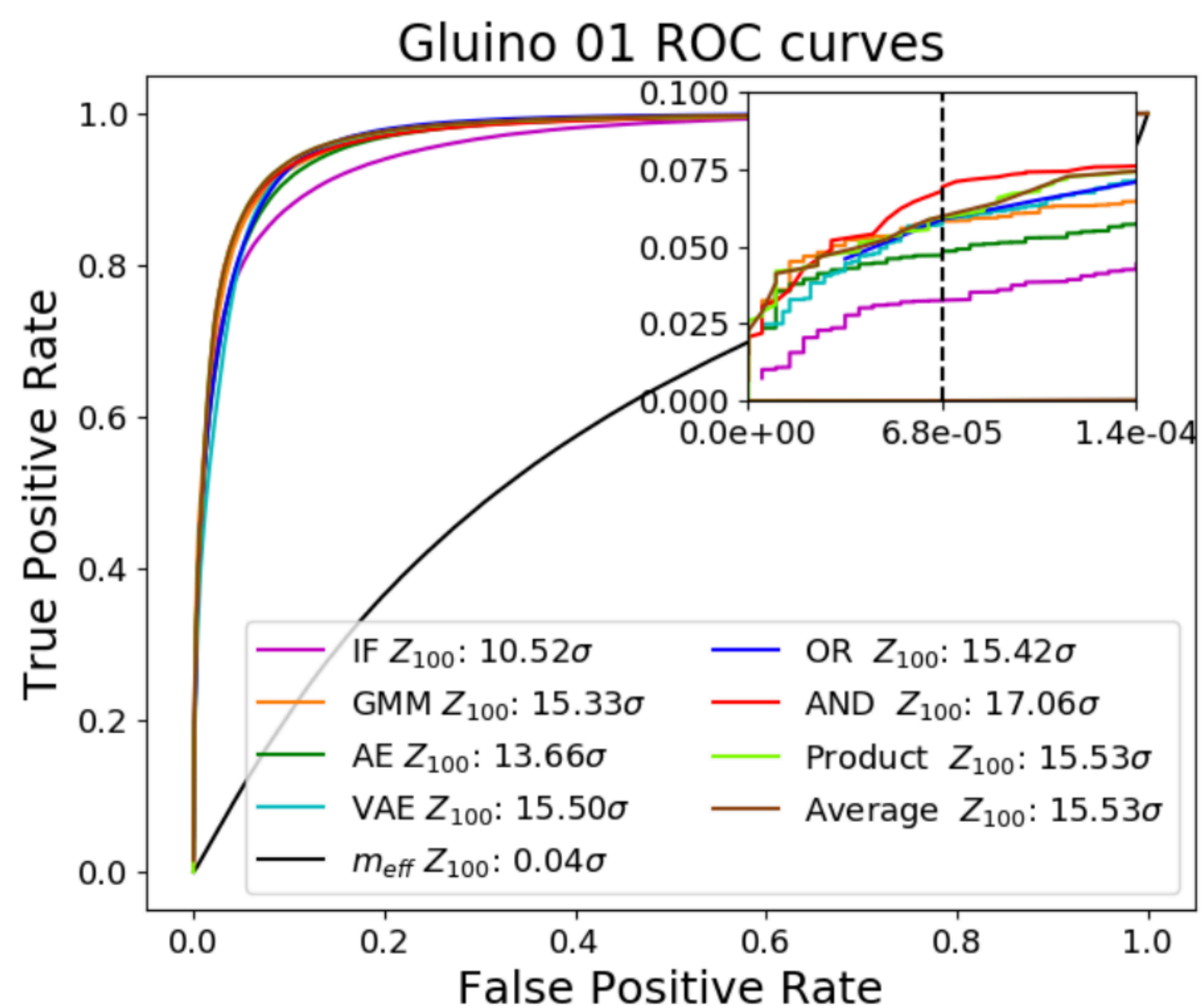
4-vector space

https://arxiv.org/abs/2010.07940

- Apply the AE/IF/GMM from before on the latent space of a VAE trained on the background events



4-vector space



Latent space of VAE

https://arxiv.org/abs/2010.07940

UNSUPERVISED CHALLENGE

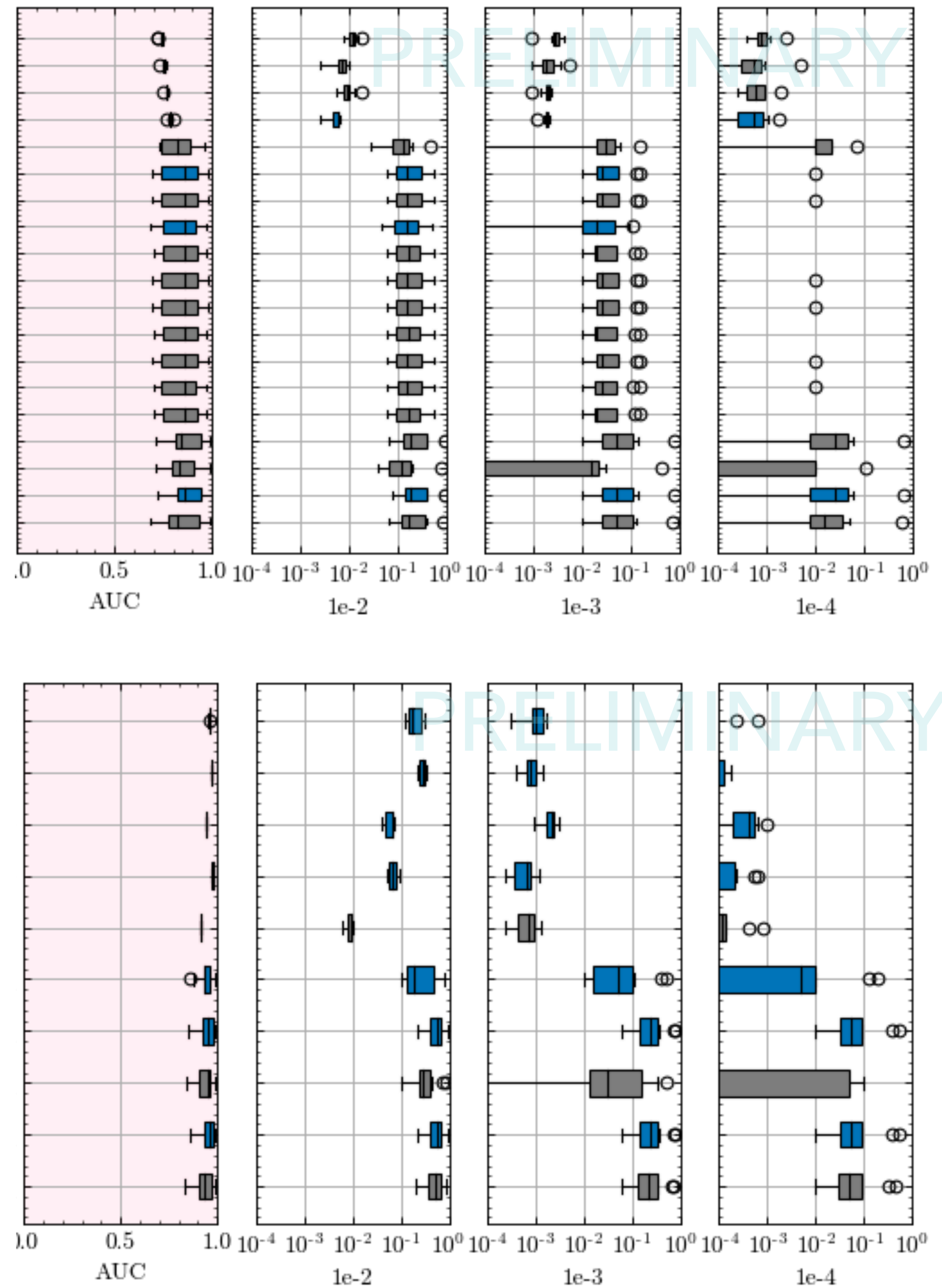
- Latent spaces change your input parameter space to an abstract space

UNSUPERVISED CHALLENGE

- Latent spaces change your input parameter space to an abstract space
- Work with methods inside the latent space of a VAE can vastly improve performance
- Within darkmachines collaboration we are close to publishing a comparison paper on the phenomldata dataset
    - Challenge
    - Given a training and test set with various known signals
    - Test models to a secret test with various blinded signals

UNSUPERVISED CHALLENGE

- Latent spaces change your input parameter space to an abstract space
- Work with methods inside the latent space of a VAE can vastly improve performance
- Within darkmachines collaboration we are close to publishing a comparison paper on the phenomldata dataset
  - Dozens of models (traditional, beta-VAE, CNN-VAE, flows, flows in VAE latent spaces, combined, …)
  - 1000s of hyperparameter combinations
  - Multiple anomaly score definitions

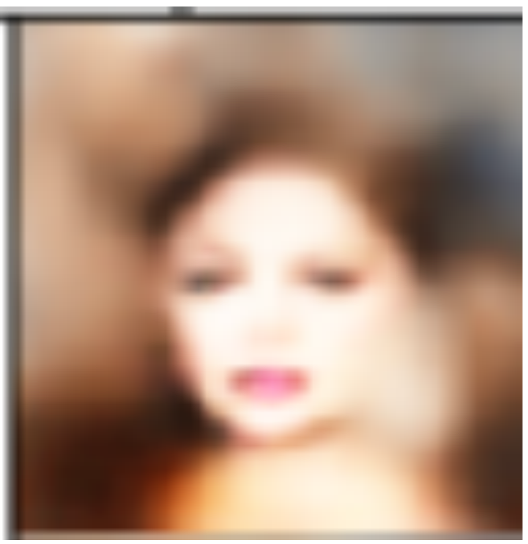


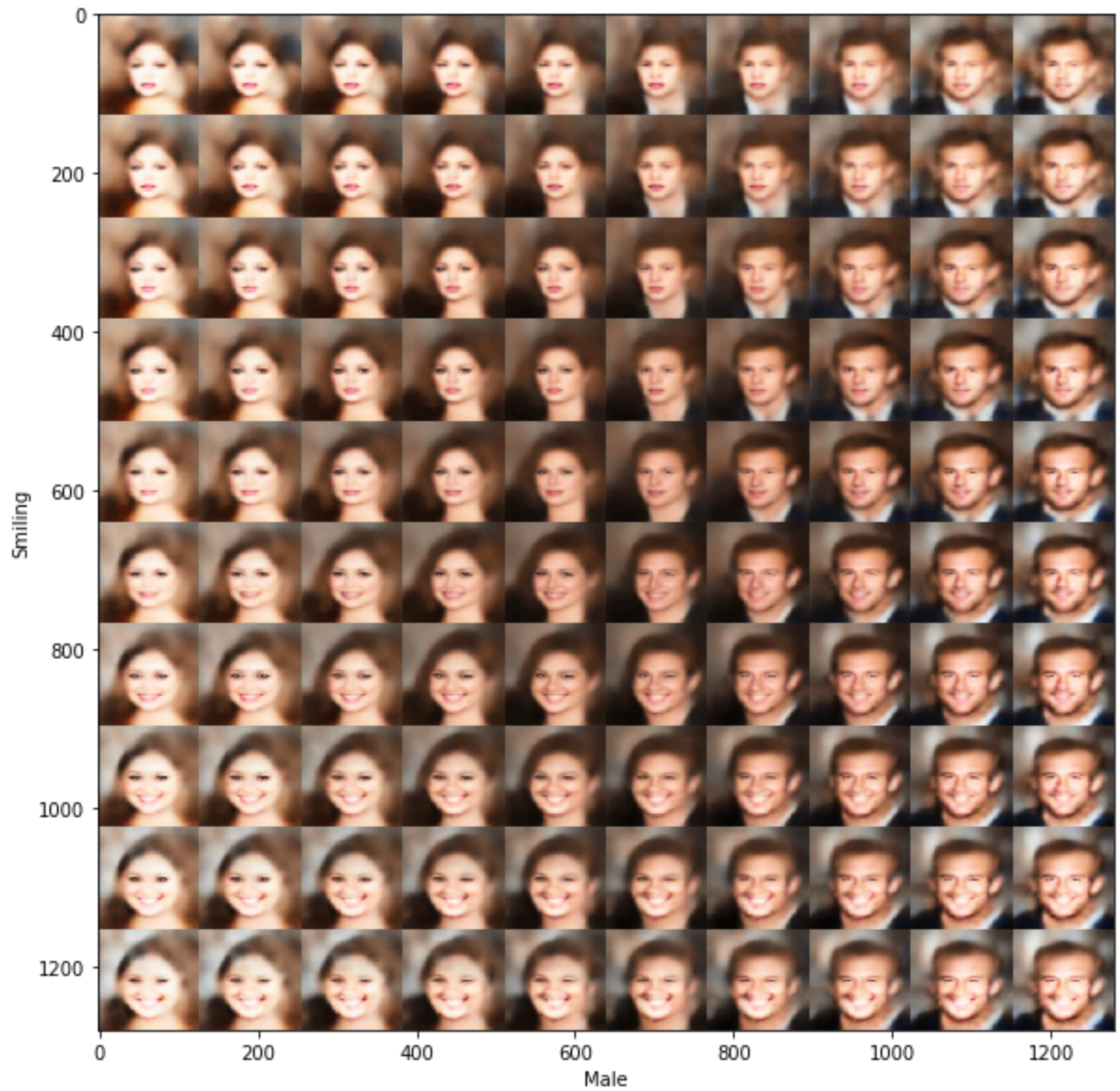
GENERATIVE MODELS AS EVENT GENERATORS

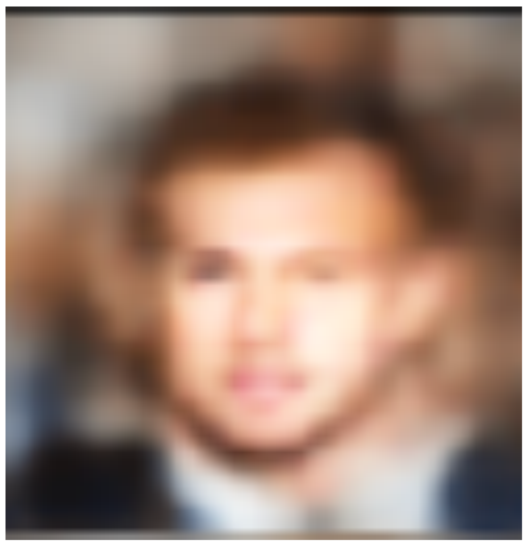
- To be able to do the previous, need lots of events
- Event generation is slow, especially if you need billions of events and need to run the whole LHC simulation pipeline
- You can also use VAEs as event generators

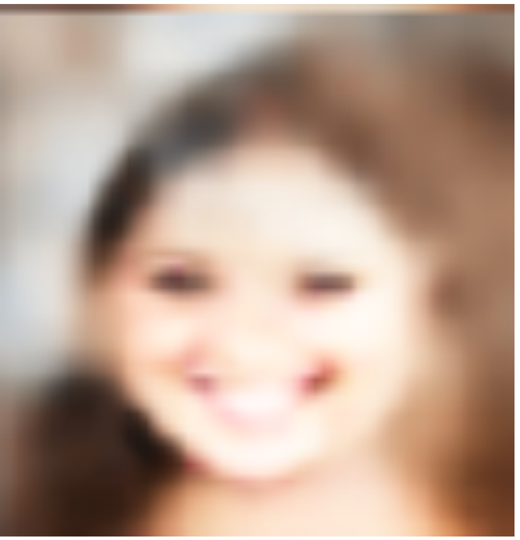
PLAYING WITH LATENT SPACES

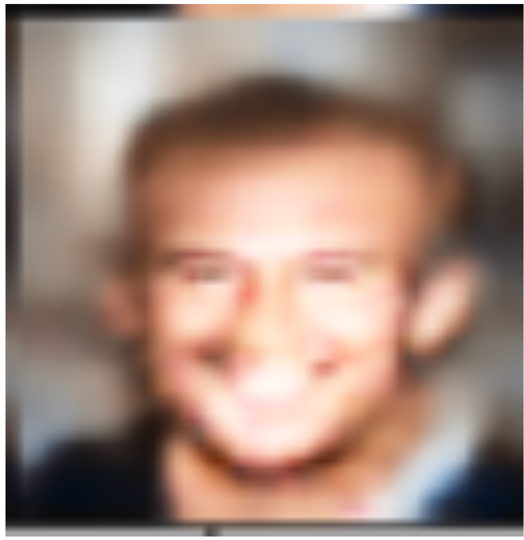
- Train VAE on face images
- PCA, then change the latent space variables









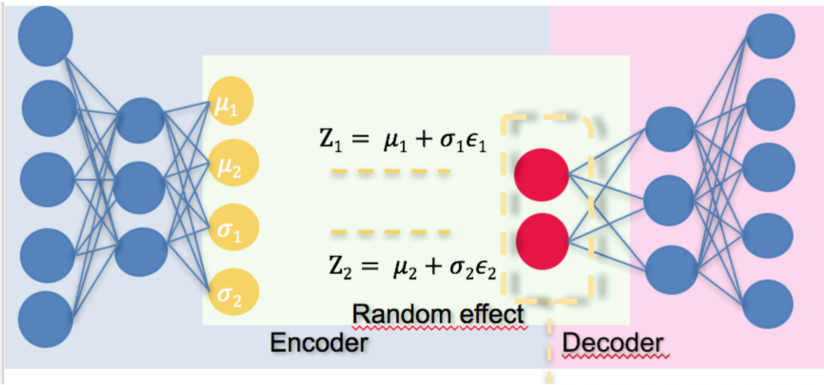


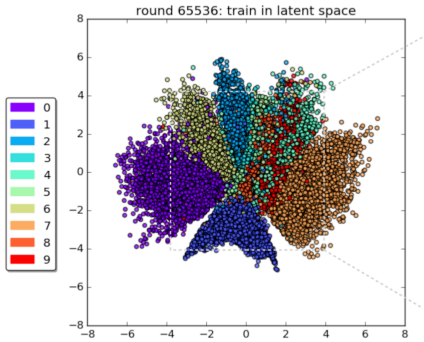
EVENT GENERATORS USING VAE

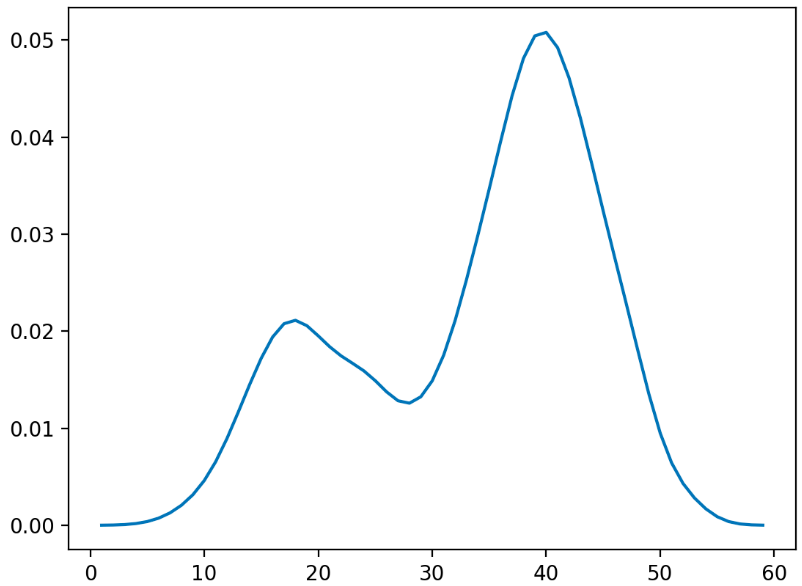
- Set up a VAE, train on the events you want to generate

EVENT GENERATORS USING VAE

- Set up a VAE, train on the events you want to generate
- Run representative set (=buffer) through trained encoder to get PDF of the dataset in latent space
  - (=sum of gaussians)

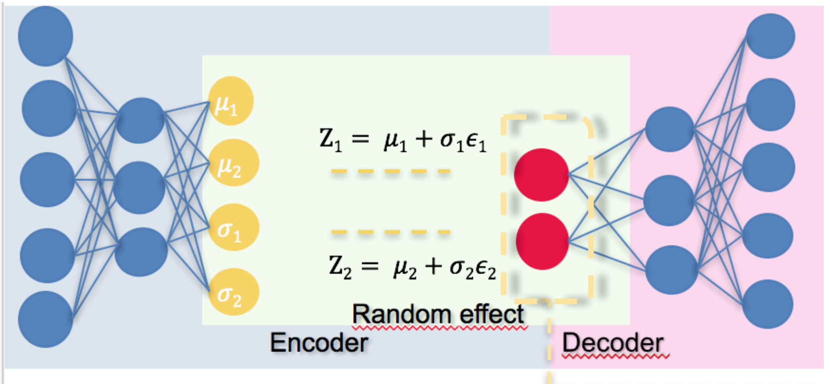




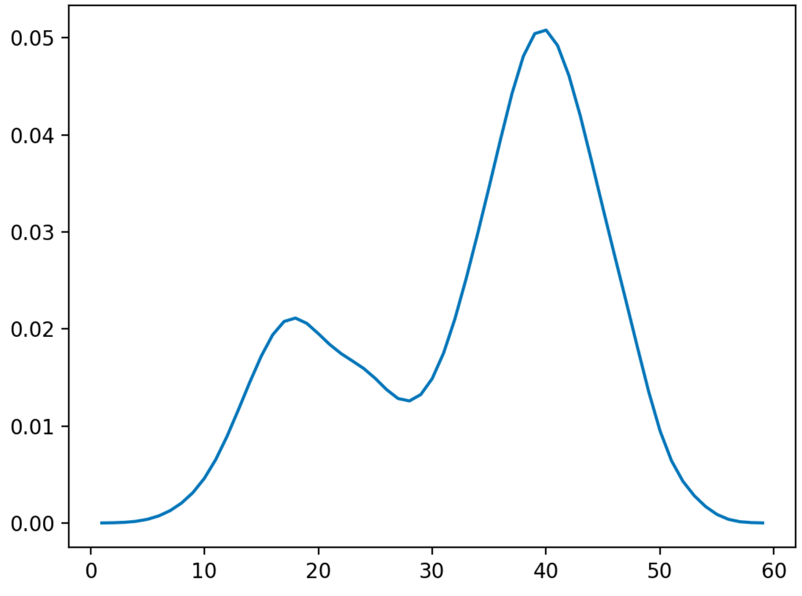


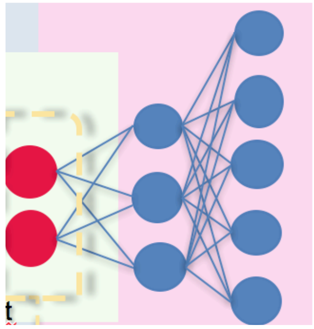
EVENT GENERATORS USING VAE

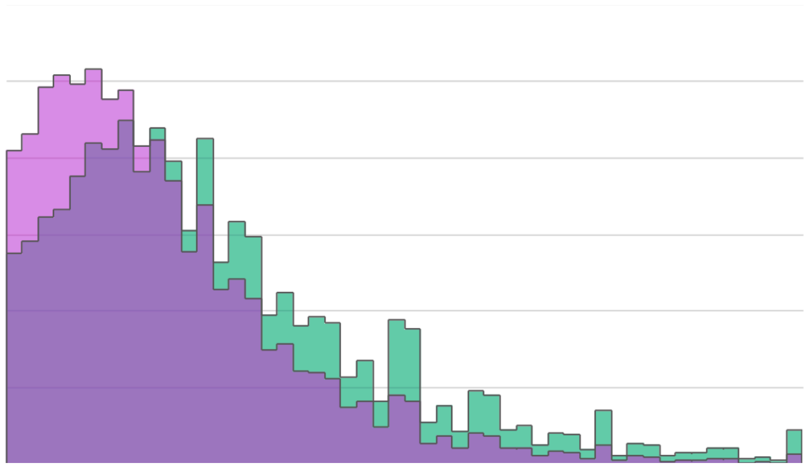
- Set up a VAE, train on the events you want to generate
- Run representative set (=buffer) through trained encoder to get PDF of the dataset in latent space
  - (=sum of gaussians)



- Sample from the PDF, run through decoder

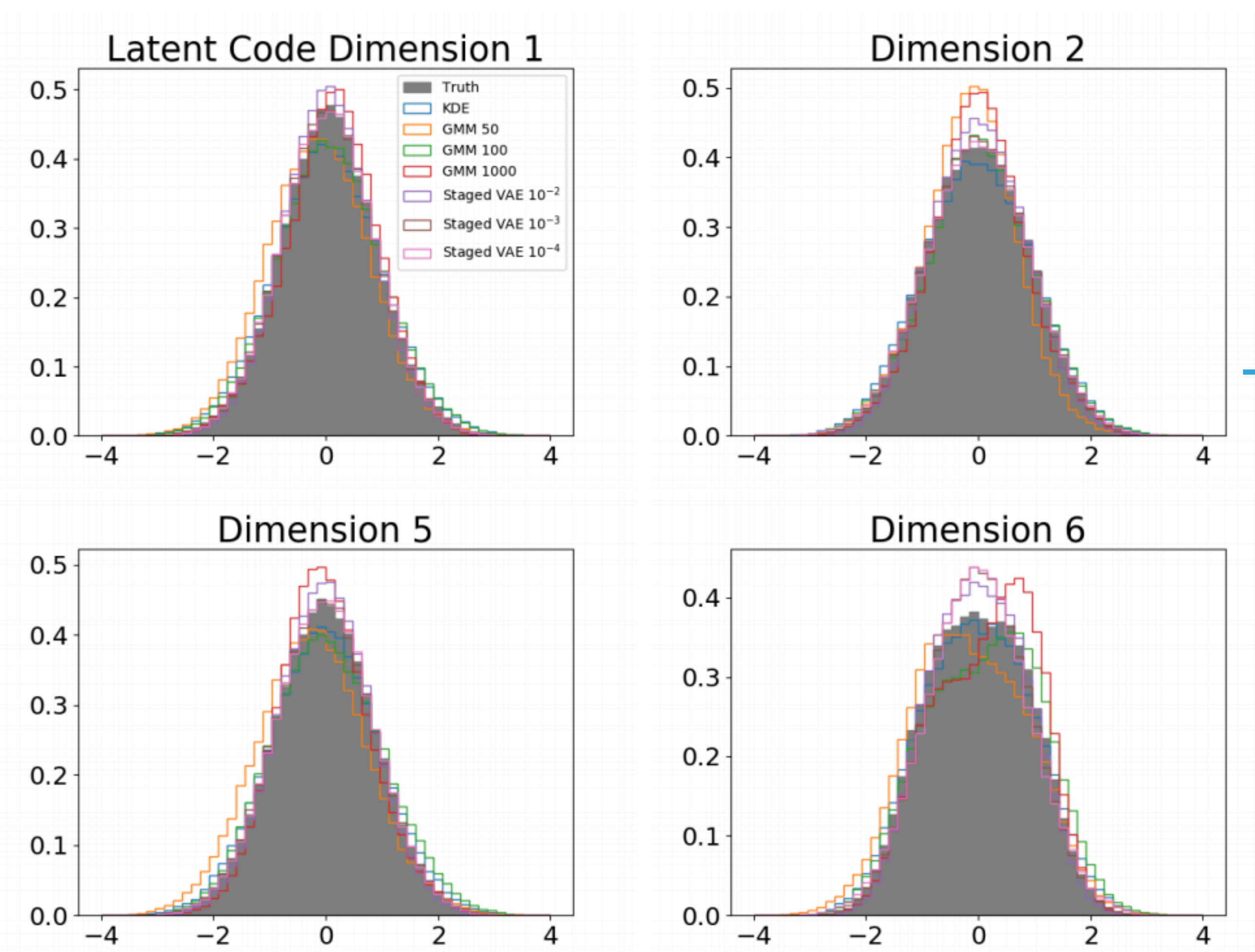






EVENT GENERATORS USING VAE

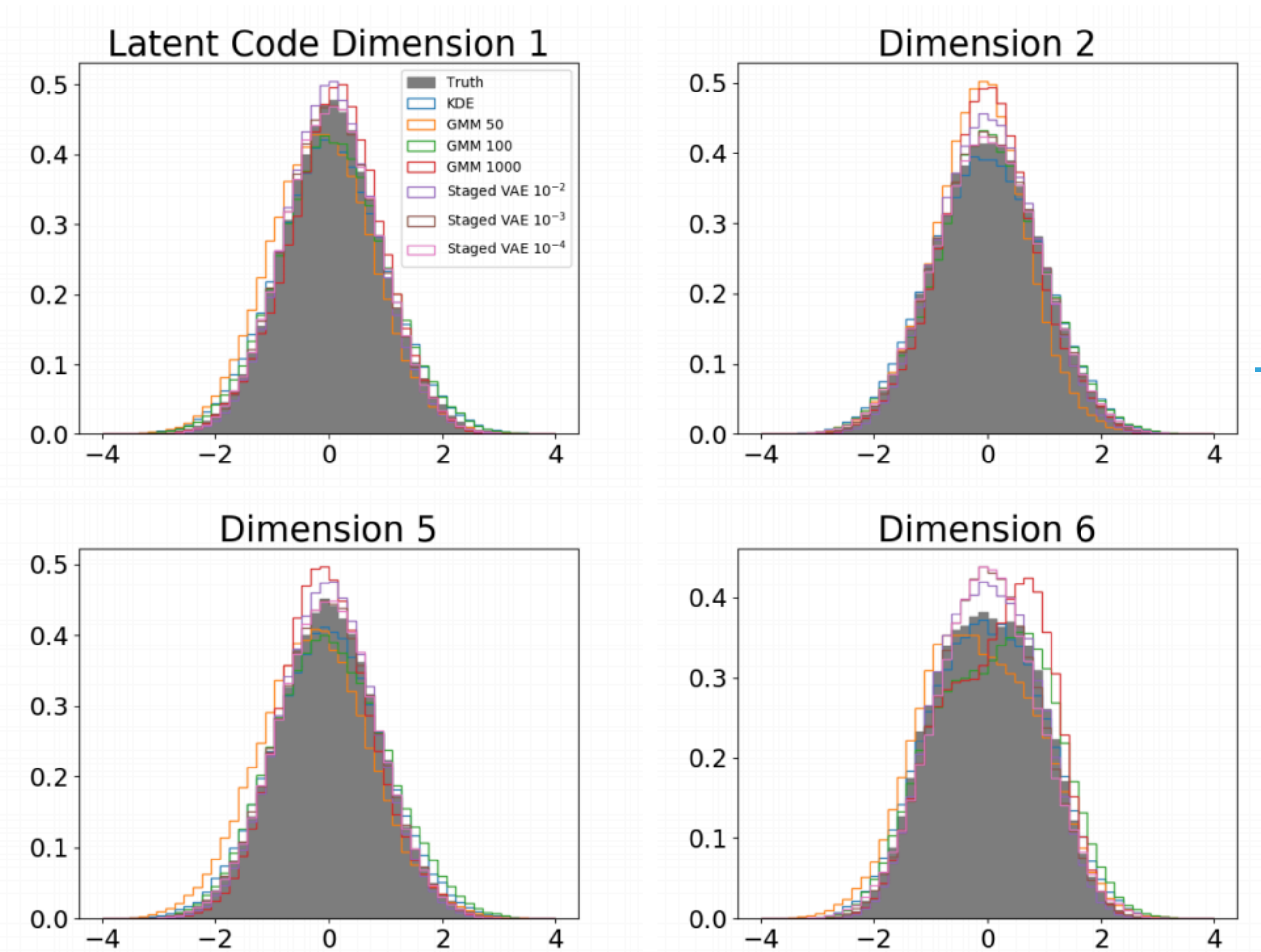
- It generates events in 28D, show 8
- Z=20, show 4

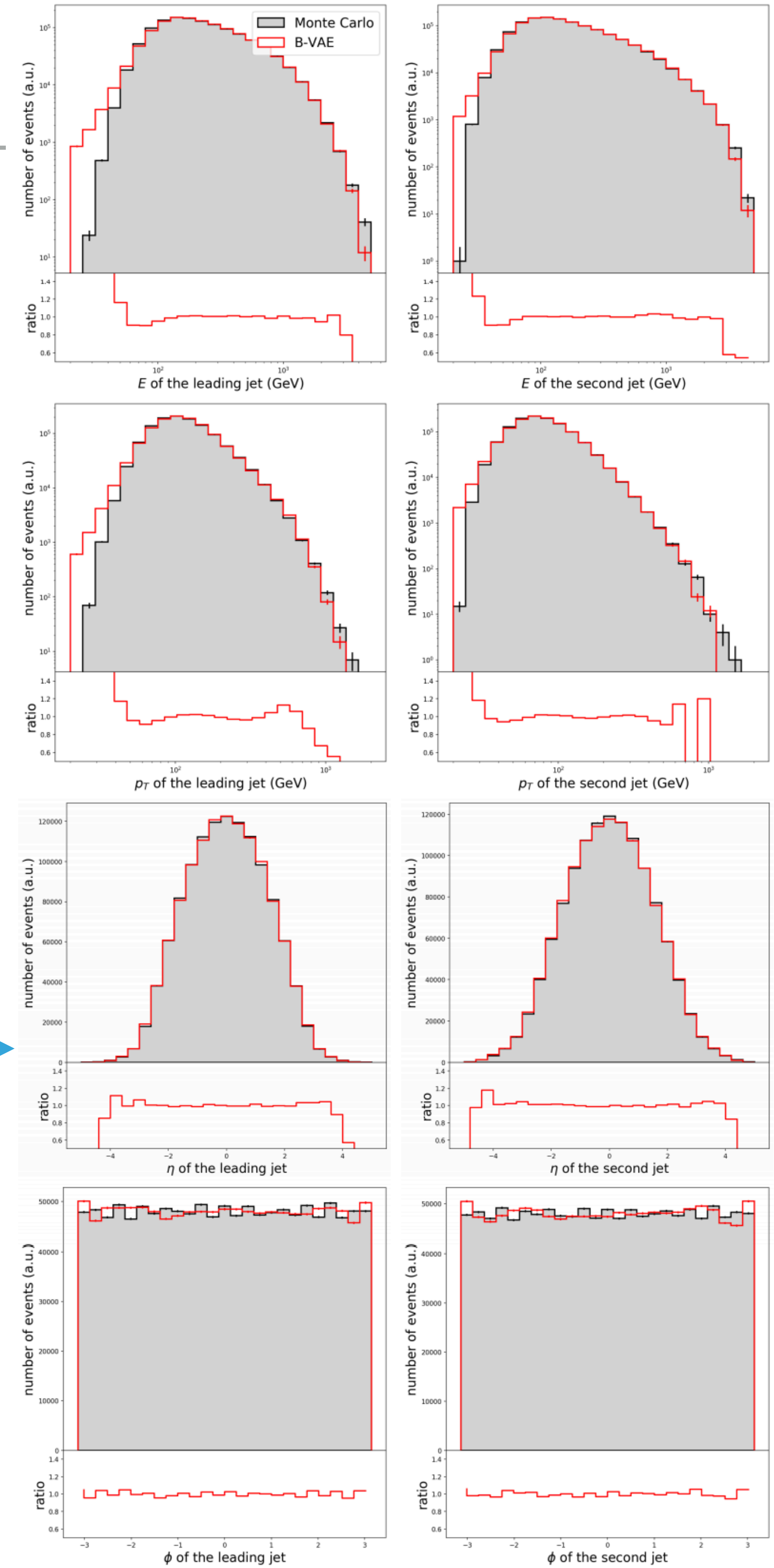


https://arxiv.org/abs/1901.00875

EVENT GENERATORS USING VAE

- It generates events in 28D, show 8
- Z=20, show 4
- Using B-VAE is orders of magnitude faster (10 million events in 3 minutes)

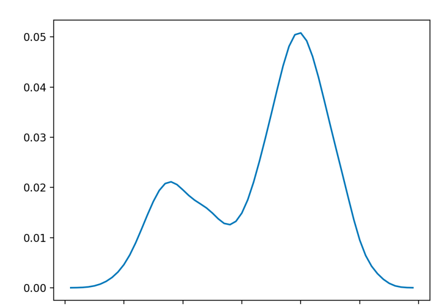


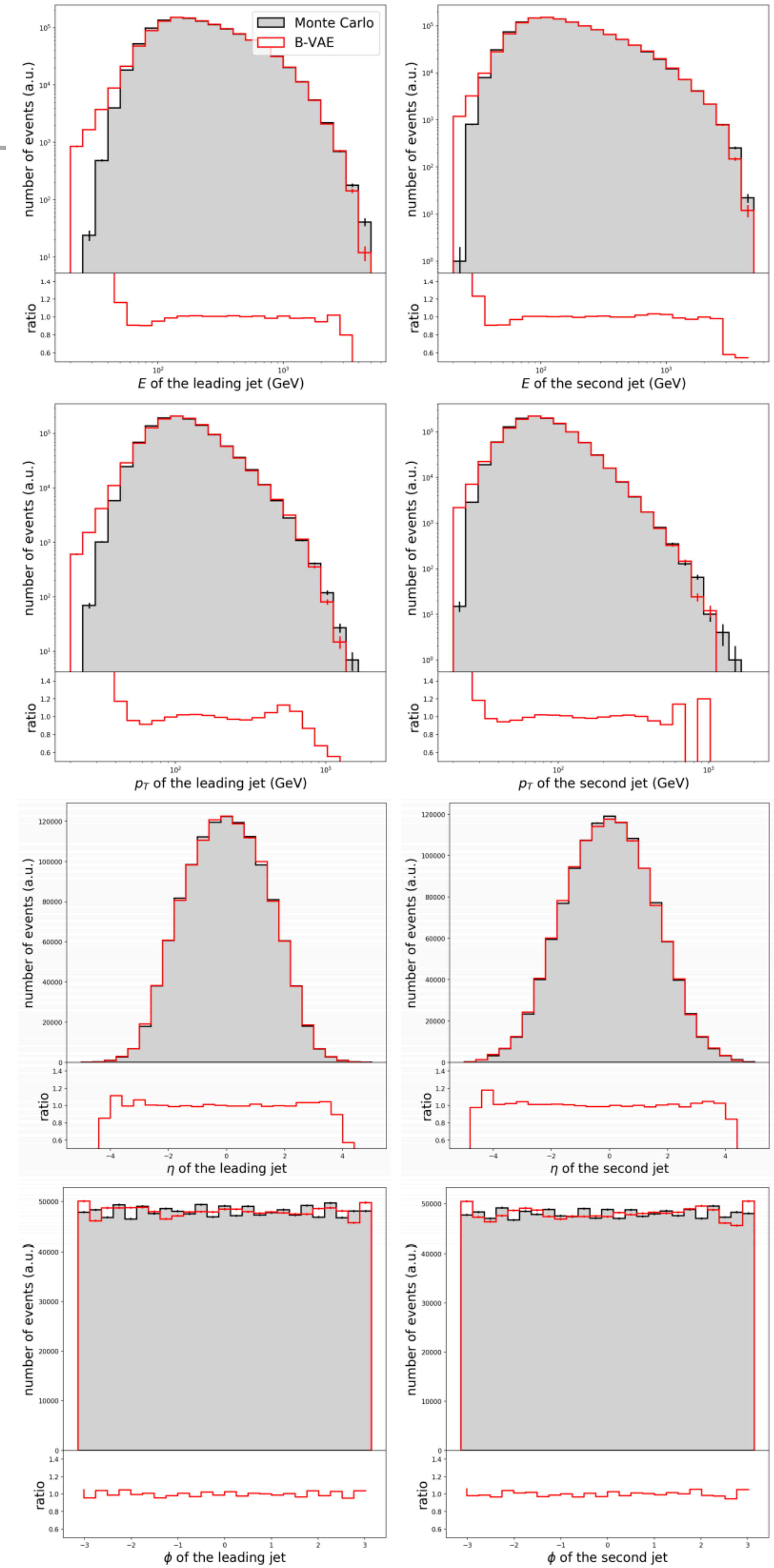


https://arxiv.org/abs/1901.00875

EVENT GENERATORS USING VAE

- Use cases:
  - Artificial data generation
  - Create events for hybrid/new signals
  - Condition to generate rare events in only a specific angle or MET
    - Modify PDF





CONCLUDING REMARKS

- Anomaly detection using latent spaces of VAEs can be a very good method to find new physics
- Can also be used as event generators
- Try your own methods using data from www.phenomldata.org (comparison paper in the works with darkmachines)
- Quick visualisation tool we have built: http://spot.phenomldata.org/#session=https://s3-eu-west-1.amazonaws.com/www.phenomldata.org/session.json
    - See also https://arxiv.org/abs/2011.03801

BACKUP

HOW GOOD IS A GENERATIVE MODEL?

- Use two figures of merit:
  - 1D kinematic distributions to quantify how well events are reconstructed $\delta_{kin} = -\sum_{i=1}^{27} \log(\chi_i^2) + \log(W_i) + \log(JSD_i)$
  - 2D density measure: calculate a fraction of "holes" in generated samples $\delta_{de} = |f_{MC} - f_{NN}| \cdot (\chi^2 + JSD)$

(does the model just generate the same thing multiple times?)

  - Useful for artificial data generation
  - VAEs are also good generative models