Generative models for calorimeters response simulation - from GANs through VAE to e2e SAE

Kamil Deja, Jan Dubiński, Piotr Nowak, Sandro Wenzel, Tomasz Trzciński

https://arxiv.org/abs/2006.06704

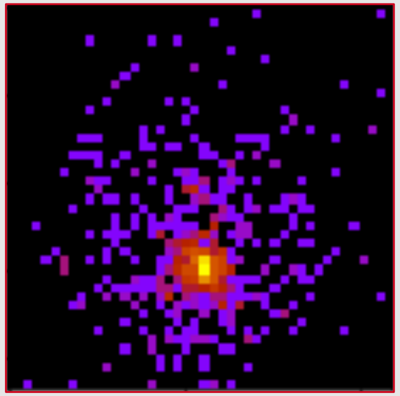
Agenda

- ZDC Calorimeters response simulation
- Generative Adversarial Network and Variational Autoencoder approach and results
- Problems of generative autoencoders
- End-to-end Sinkhorn Autoencoder
- Final Results

Calorimeters response simulation

- Two sets of images simulated with GEANT
- 11 attributes describing particle (pdg, Energy, momentum, primary vertex…)
- Two dataset with over **100 000/ 500 000** pairs of images
- Pixel values [0-1000] (mean ~0.23)

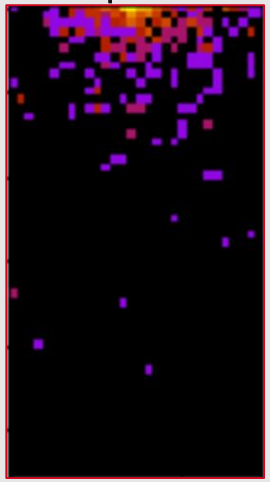
44 pixels



44 pixels

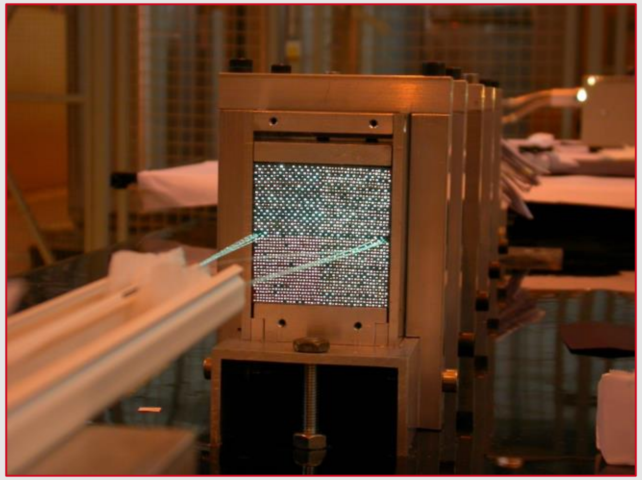
Neutron detector response

30 pixels



63 pixels

Proton detector response



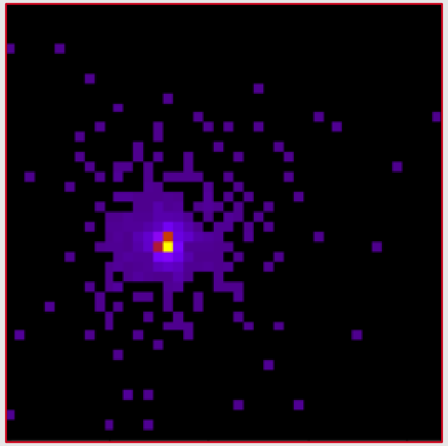
ZDC Calorimeter

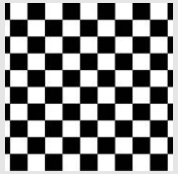
Basic solution and evaluation

- Deep conditional generative model based on GAN or VAE

Evaluation

- Comparison of sums over specific pixels for generated and real data distributions using „channel summing" procedure



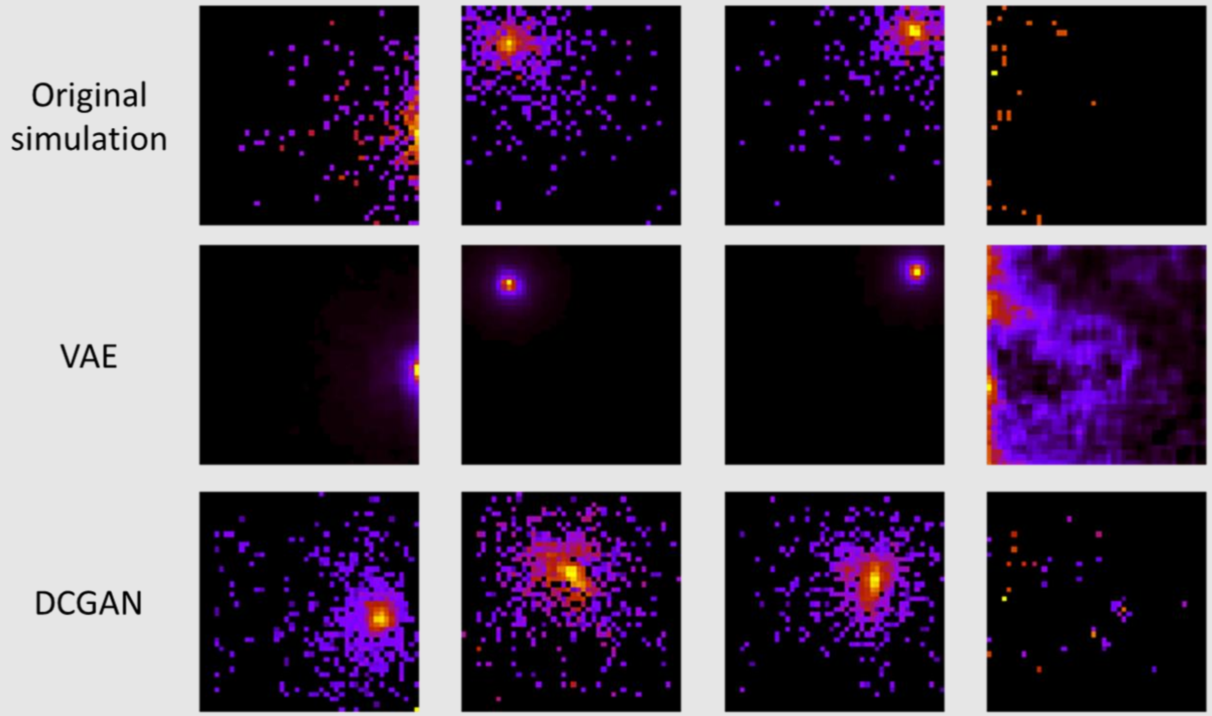


Pixel count

Sum 1	Sum 2	Sum 3	Sum 4	Sum 5
8	22	6	197	791

https://cds.cern.ch/record/381433

Exemplar results



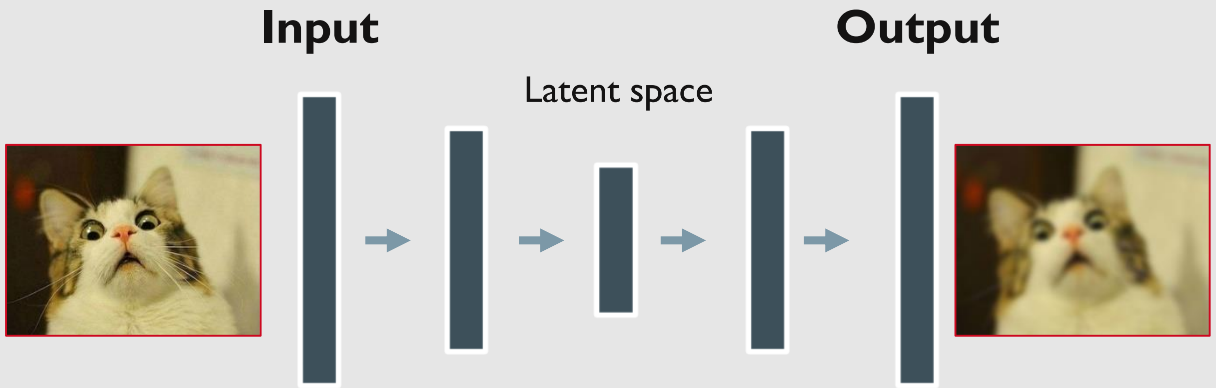
Center of the response

Distribution of all responses

Model	MAE (distance from original)	Wasserstein (distribution comparison)
Cond VAE	**23,13**	14,92
Cond WAE (MMD)	43,54	34,46
Cond DCGAN	68,27	**6,95**
Original simulation	6,59	2,89

From VAE/WAE to e2e Sinkhorn autoencoder

Autoencoder

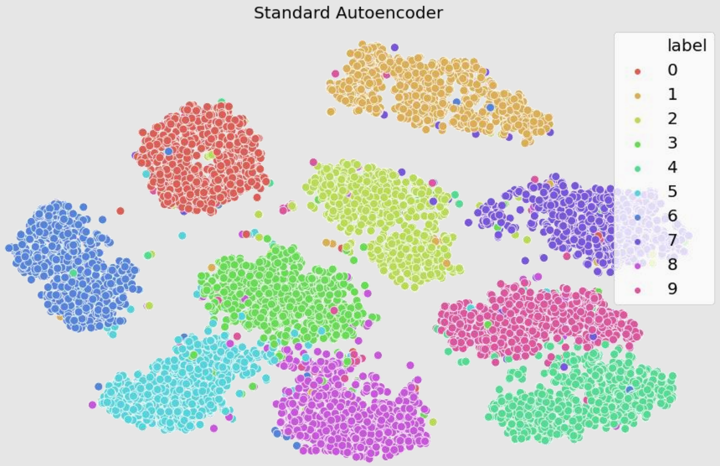


1	2	2	...
1	3	2	...
1	4	3	...
...	...	...	...

MSE loss

1	3	2	...
1	3	3	...
1	4	3	...
...	...	...	...

$$Loss = \|x - dec(enc(x))\|^2$$



TSNE visualisation of 10 –dimensional standard autoencoder's latent space

Variational Wasserstein and Sinkhorn Autoencoders

Regularisation on latent space to the Normal distribution

Variational Autoencoder (VAE):

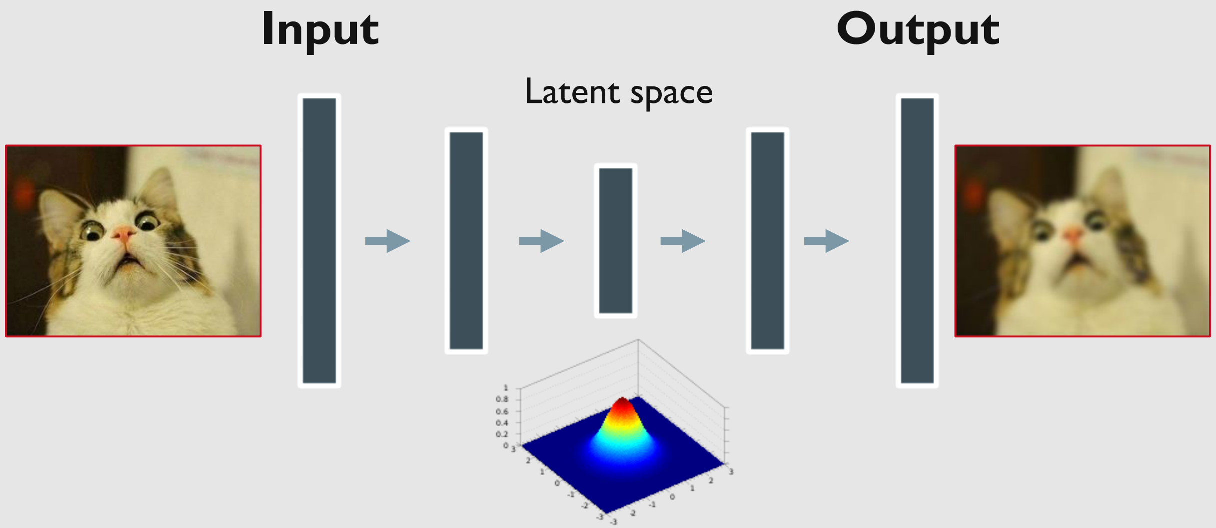
- KL Divergence

Wasserstein Autoencoder (WAE) :

- Maximum Mean Discrepancy

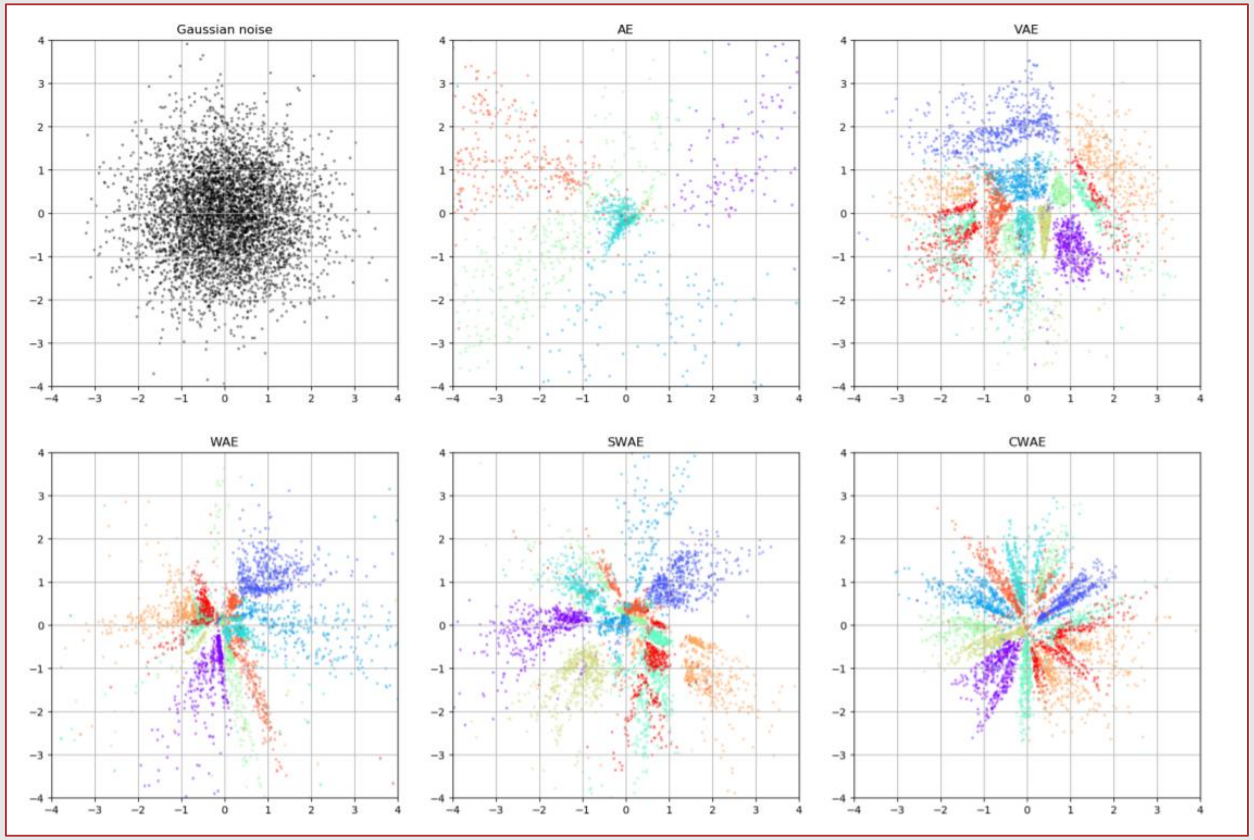
Sinkhorn Autoencoder (SAE)

- Soft Wasserstein distance

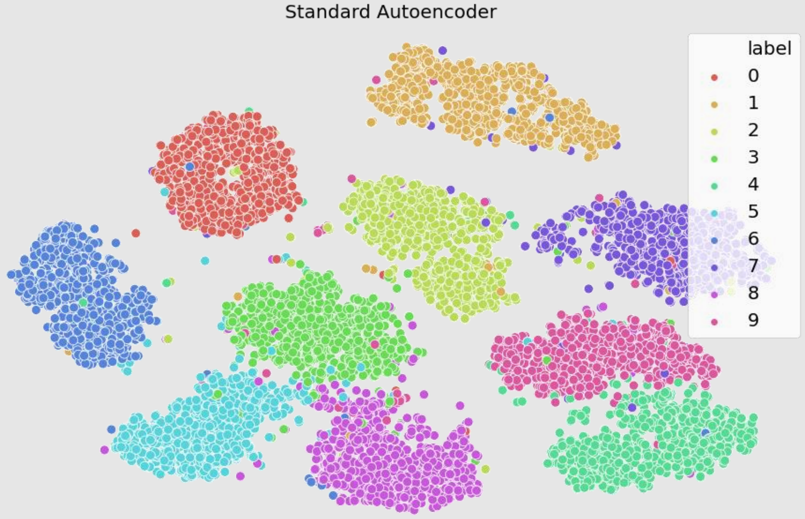


$$Loss = \|x - dec(enc(x))\|^2 + D[enc(x), \mathcal{N}(0,1)]$$

Latent space in different generative autoencoders (MNIST)



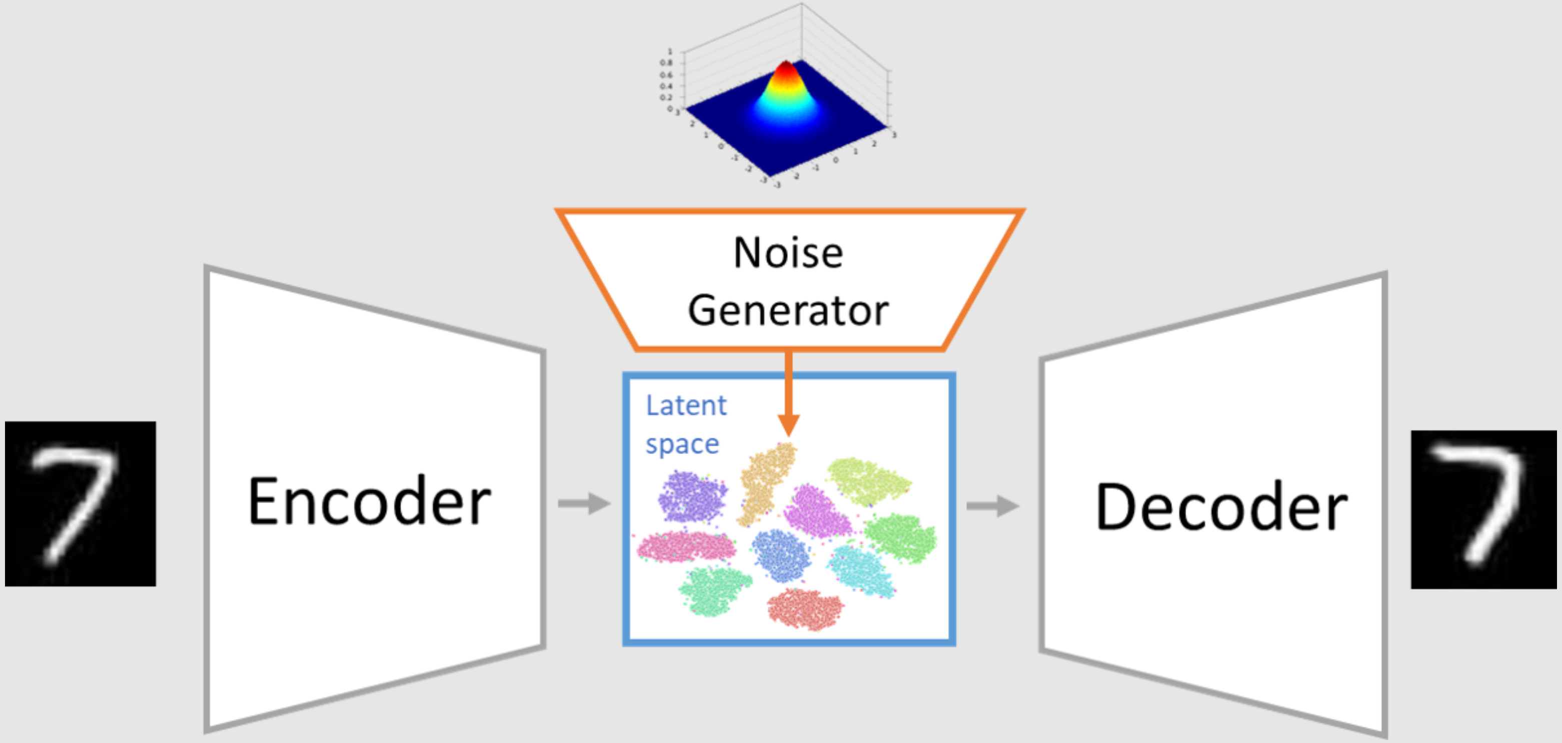
Tabor et.al. Cramer-Wold AutoEncoder



TSNE visualisation of 10 –dimensional standard autoencoder's latent space

Regularisation and separation required at the same time!

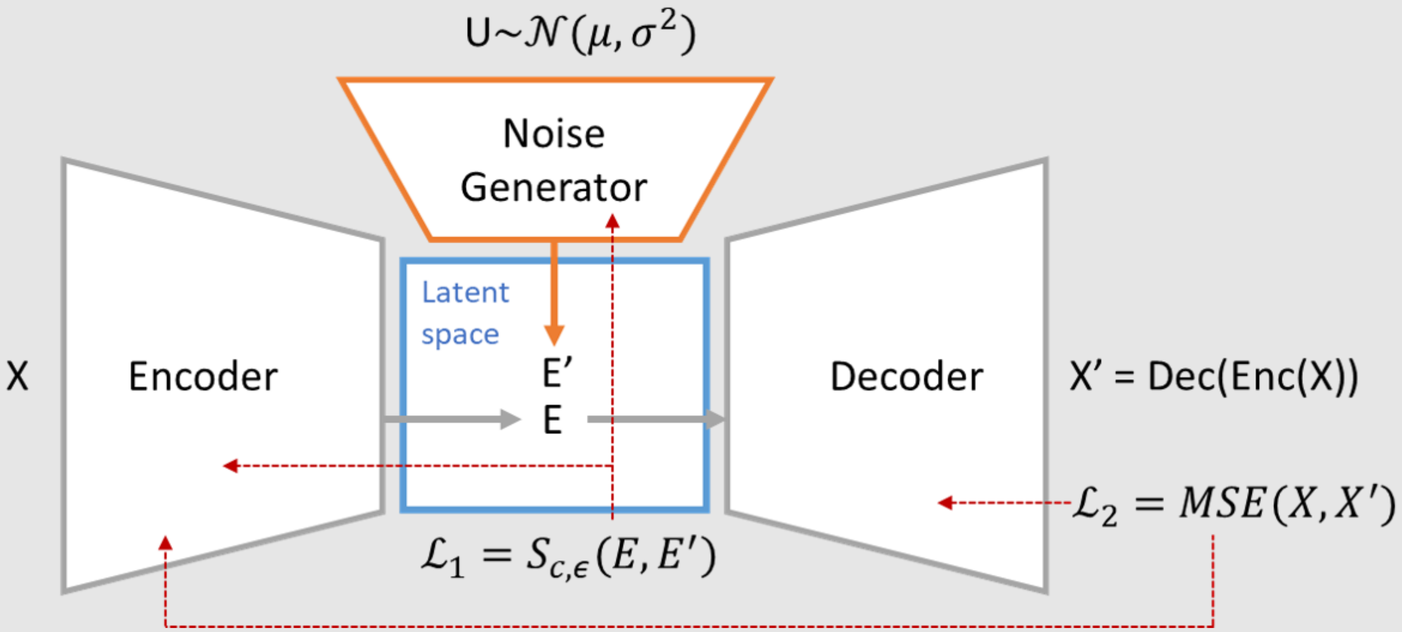
End-to-end Sinkhorn Autoencoder with Noise Generator



No regularisation on latent space – additional noise mapper

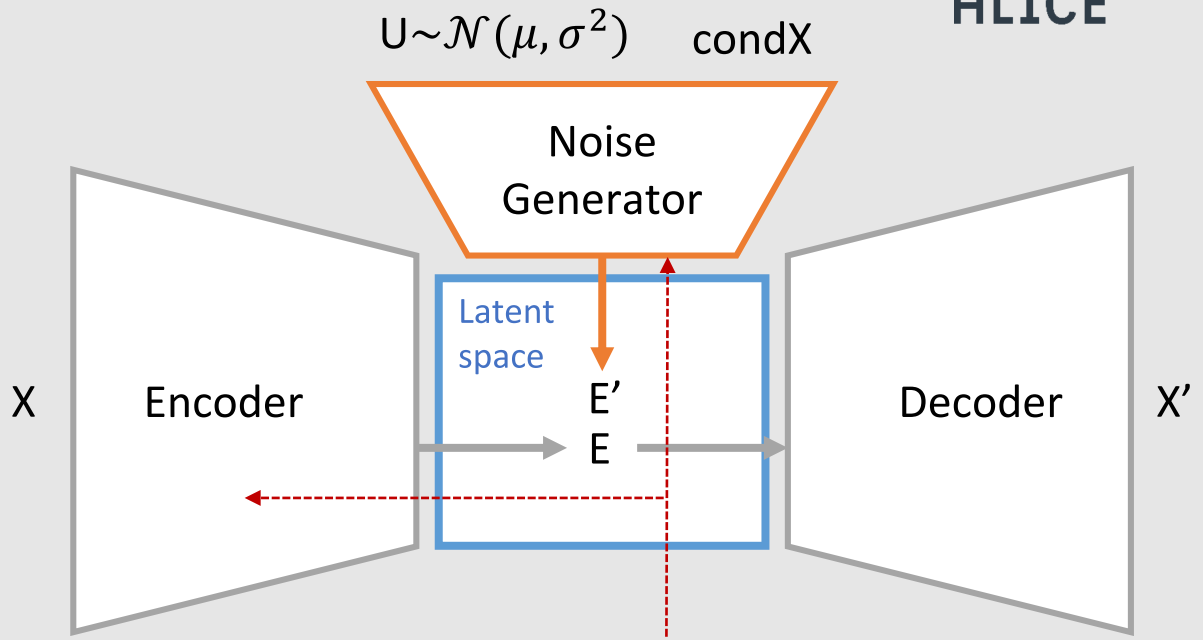
End-to-end Sinkhorn Autoencoder

- No implicit regularisation of autoencoder's latent space
- Approximation of original data embeddings with deterministic neural network
- Joint optimisation of both neural networks



Conditional e2e-SAE

- Conditional information added to noise generator
- Wasserstein distance between embeddings concatenated with conditional values
- Original data distribution on latent space



$$\mathcal{L}_1 = S_{c,\epsilon}([E, condX], [E'(condX), condX])$$

Encoded data

E1	E2	CondX
0.1	0.2	0
0.9	0.8	1
0.15	0.1	0

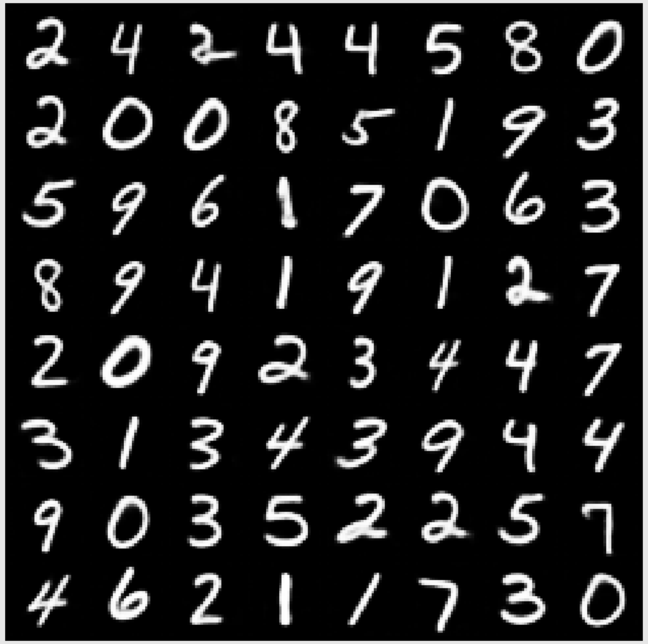
Wasserstein dist.

Encoded noise

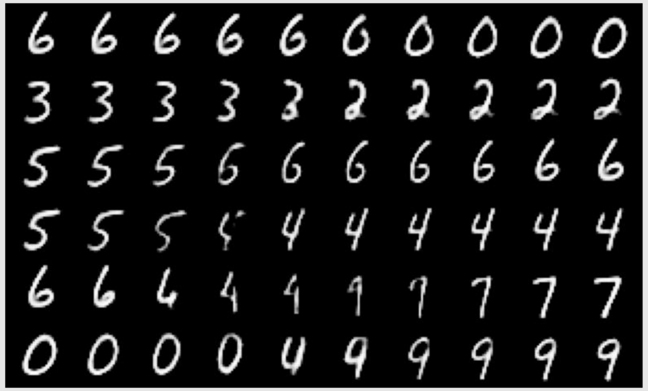
E'1	E'2	CondX
0.16	0.12	0
0.88	0.82	1
0.11	0.18	0

Experimental results

Benchmark of conditional models



Generated samples



Interpolations (on input data)

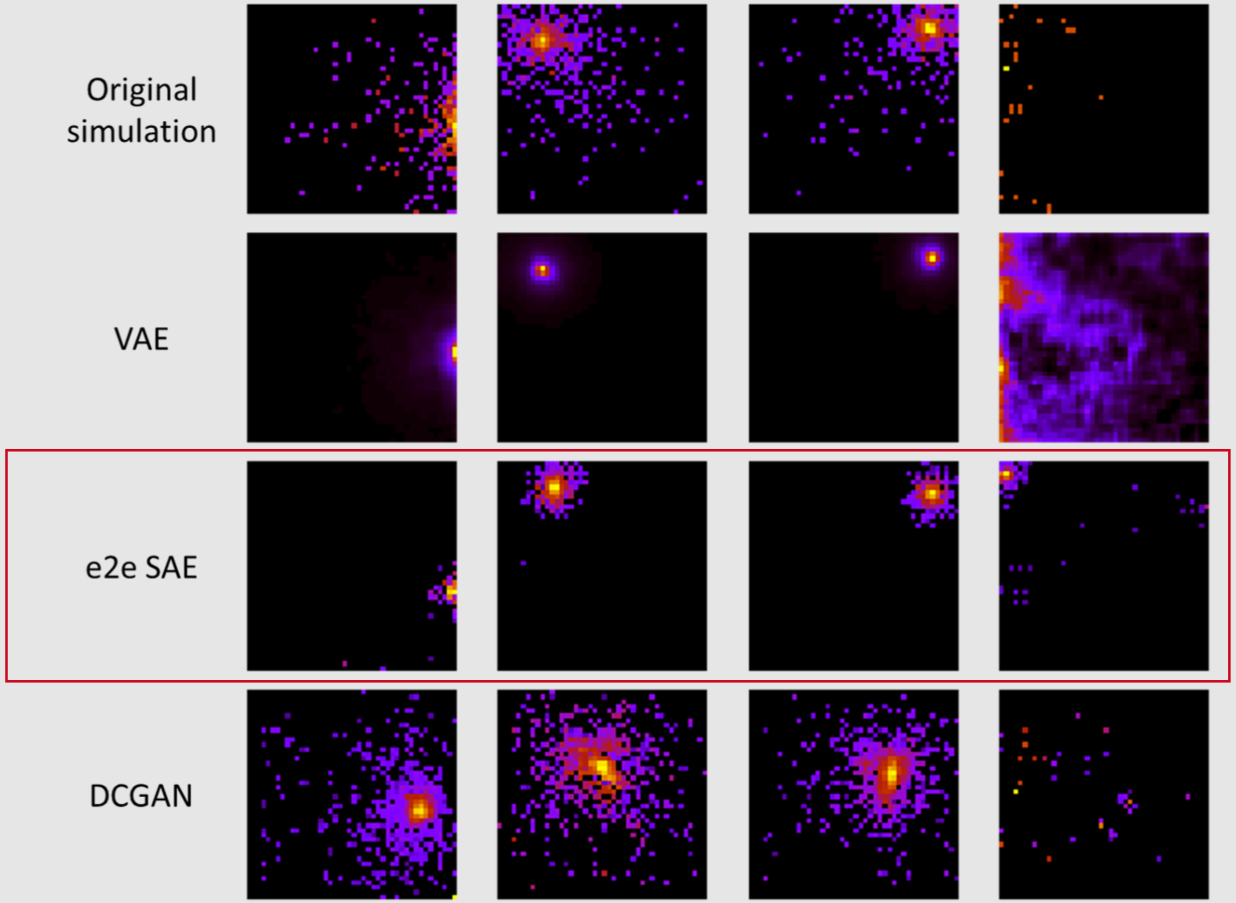
Metoda	FID
Cond VAE	6,61
Cond WAE (MMD)	34,73
Cond e2e SAE	**4,11**
Cond DCGAN	0,93
Original data	0,33

FID – Fréchet Inception Distance – [Hausel et al.](#)

Benchmark - CelebA

Model	FID
VAE	55
WAE (MMD)	55
SWAE	64
SAE (H)	56
e2e SAE	**54.5**
bigWAE (MMD)	37
Style GAN	5.06

Calorimeter response simulation



Model	MAE	Wasserstein
Cond VAE	23,13	14,92
Cond WAE (MMD)	43,54	34,46
Cond e2e SAE	**13.50**	7,91
Cond DCGAN	68,27	**6,95**
Original simulation	6,59	2,89

Summary

- First experiments with ZDC fast simulation with generative models
- New method based on generative autoencoder without regularisation on latent space
  - More stable then Generative Adversarial Networks
  - Better results then other Generative Autoencoders

Future steps:

- Further evaluation of simulations (e.g. comparison on the real physical analysis)

Backup slides

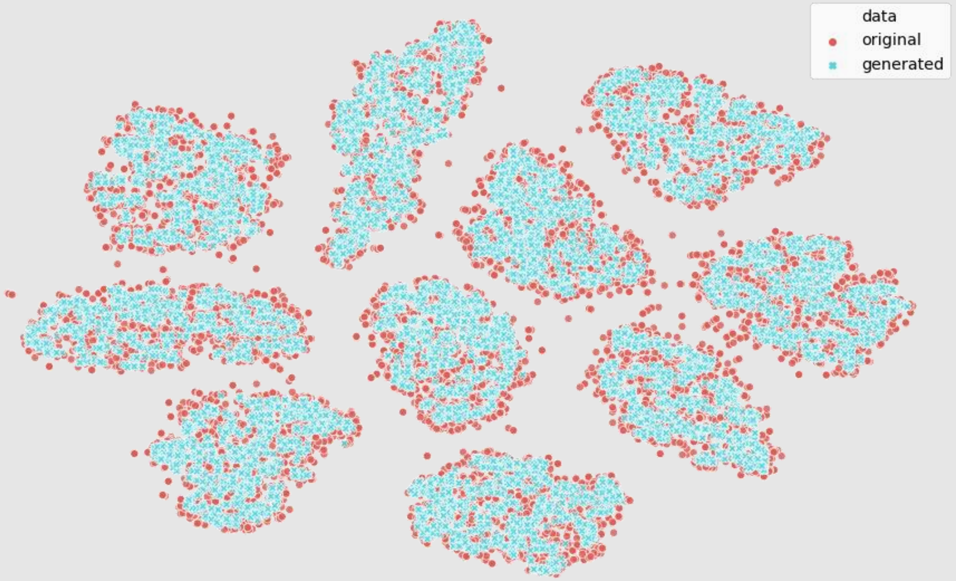
Generative models for fast simulations

- Multiple approaches with GANs:
  - CaloGAN (Paganini et al.)
  - 3D convolutional GAN (Vallecorsa et al.)
  - Fast and accurate simulations of particle detectors (Musella et al.)
- VAE
  - Variational Autoencoders for Jet Simulation (Dohi et al.)

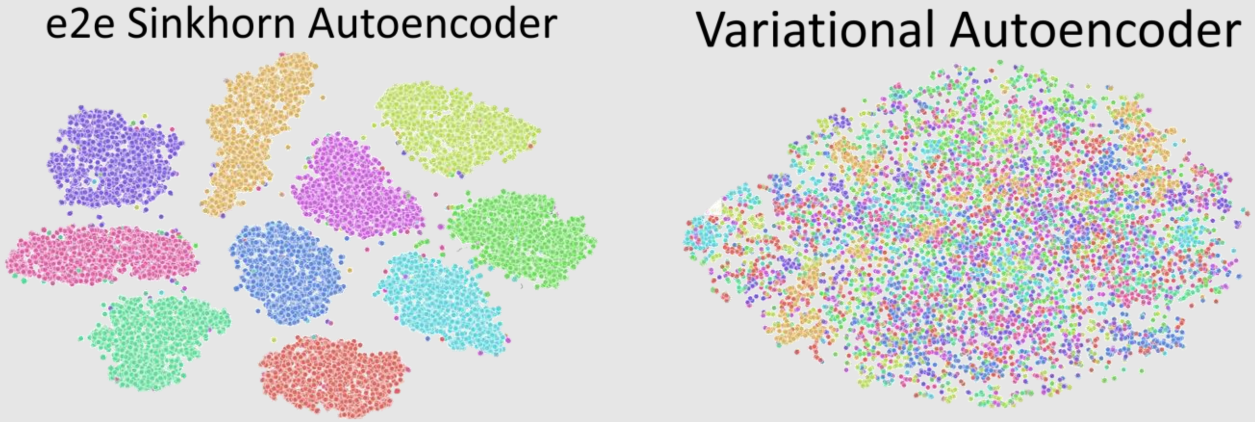
Both:

- Generative Models for Fast Cluster Simulations in the TPC for the ALICE Experiment (Deja et al.)
- Using Generative Models for Fast Cluster Simulation in the ALICE Experiment LHC (Deja et al.)

Latent space visualisation



TSNE visualisation of autoencoders latent space with original data examples (red) and generated noise (blue)



TSNE visualisation of conditional models - classes

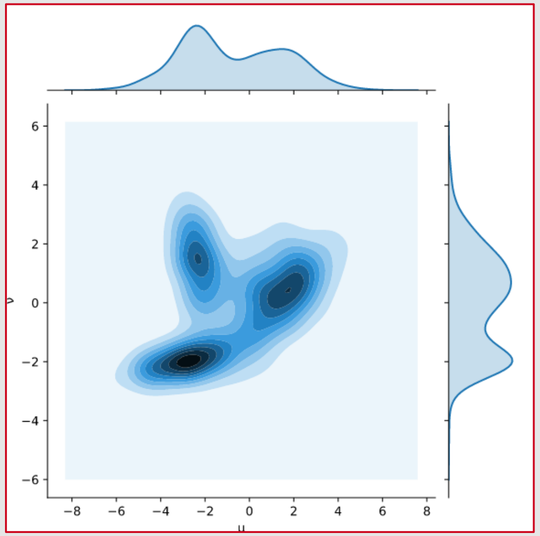
Wasserstein distance

- Total cost of moving one probability distribution into another
  (one pile of earth into another)
- Taking into account distance to cover (as cost)

$$OT_c[p,q] = \inf_\gamma \int \gamma(x,y)c(x,y)dxdy \ ,$$

For Euclidean distance as a cost:

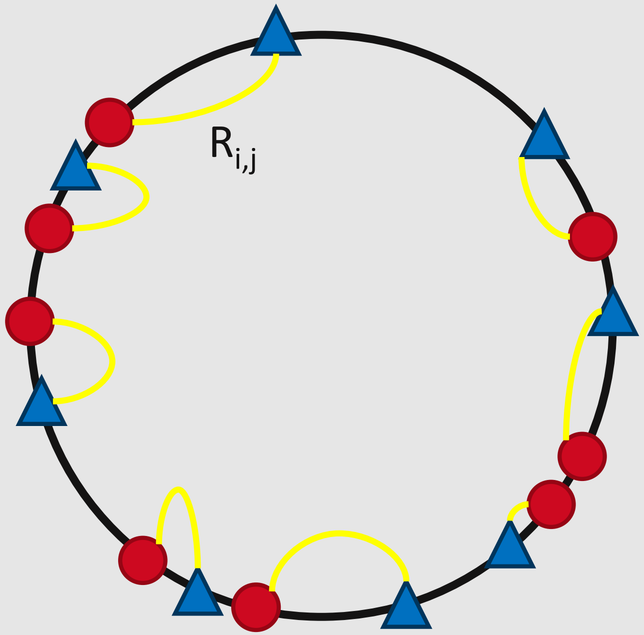
$$\mathcal{W}_2[p,q]^2 = \inf_\gamma \int \gamma(x,y)\,\|x-y\|_2^2\,dxdy \ .$$



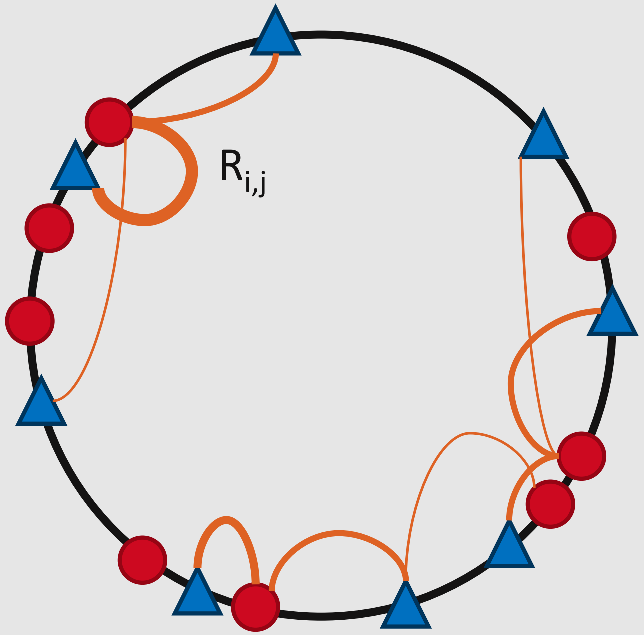
Soft Wasserstein Distance

$$D_S[enc(x), \mathcal{N}(0,1)] = \min_R \sum_{i,j} R_{i,j} \| enc(x), \mathcal{N}(0,1) \| - \epsilon H(R)$$

R: Soft transport route
H: Entropic regularisation



Original min R



Exemplar soft R

- Additional Entropic regularisation allows to use Sinkhorn algorithm [Cuturi, NeurIPS 2013]
- Differentiability [Genevay et al. 2018]