Calorimeter Shower Simulations using VAEs

Sascha Diefenbacher

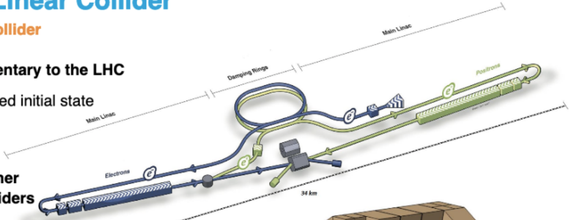
Universität Hamburg

October 28th, 2019

CLUSTER OF EXCELLENCE
QUANTUM UNIVERSE

The International Linear Collider

A planned electron-positron collider

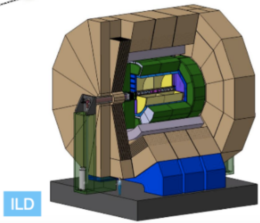


Lepton colliders are complementary to the LHC

- cleaner environment, controlled initial state
- coupling to leptons is tested

Advantages of the ILC over other planned electron-positron colliders

- mature technology
- centre-of-mass energy can be tuned and increased: 250 GeV in initial stage, upgrades to 500 GeV and 1 TeV
- polarisation of both beams: *P(e-)*=±80%, *P(e+)*=∓30%
- triggerless operation
- hermeticity of detector down to lowest angles

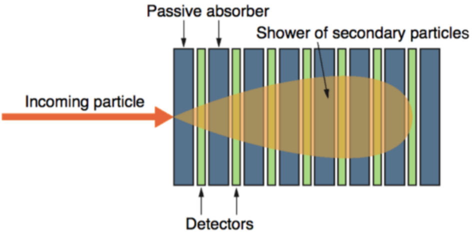


DESY. I Particle Discovery Opportunities at the ILC I M. Habermehl, 12 July 2019 Page 2

Image taken from EPS-HEP 2019 Conference, presentation by M.Habermehl

Sampling Calorimeter

- Energetic particles interact with matter
- Causes shower of secondary particles

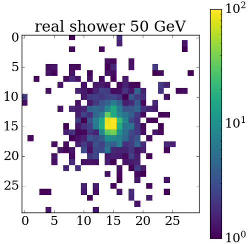


- Detector layers intermixed with passive absorber
- Several shower layers are sampled

Image taken from:
https://www.hephy.at/fileadmin/user_upload/VO-6-Calorimeters.pdf

Goals

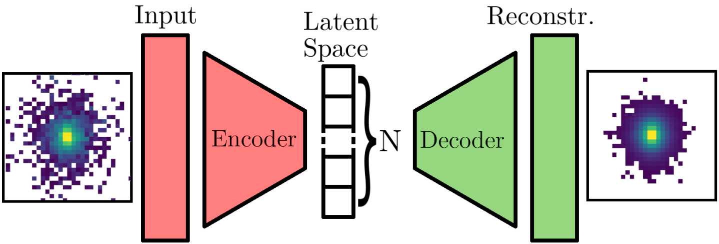
- Simulating showers form first-principle time intensive
- Generate particle showers using generative neural network
  - **G**enerative **A**dversarial **N**etwork (Engin)
  - **V**ariational **A**uto**E**ncorder



- Showers as $30 \times 30 \times 30$ matrices
- For now: 30×30 images, summed along beam line

Auto Encoders

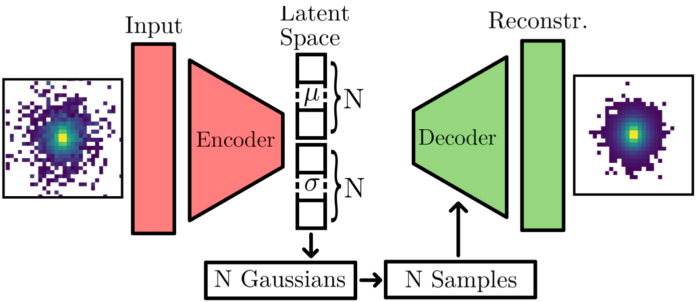
- Unsupervised learning method
- AE's learn images features by compression and reconstruction



- Used (in physics) for classification/anomaly detection
- Complex, irregular latent space, ill suited for generation

Variational Auto Encoders

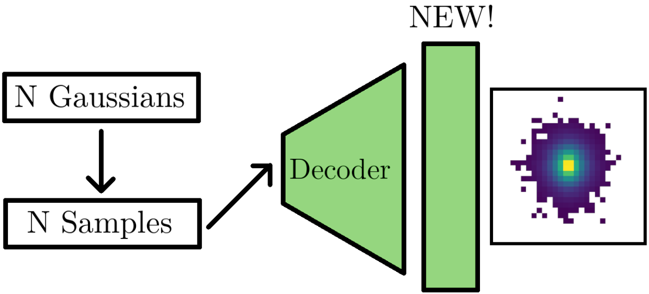
- Latent space set of Gaussian distributions



- Regularized latent space allows for image generation

Generation with VAE's

- Input random samples from normal distributions



- Get new image with similar properties to training data

Loss Function

- Measure the network tries to minimize during training
- Default VAE loss:

$$L = L_{MSE} + L_{KLD}$$

- Mean Squared Error: Per pixel comparison
  $\sum(image - recon)^2$, facilitates reconstruction
- KL-Divergence: Forces latent space to be Gaussian, with width 1, mean 0

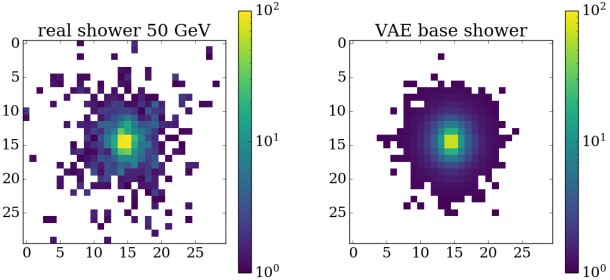
Criteria

- Generated images should be indistinguishable from real ones
- Simple by eye comparison not sufficient
  - Total energy in shower (sum of all (non-log) pixels)
  - Energy per calorimeter layer (not applicable here)
  - Radial energy distribution
  - Occupancy (number of non zero pixels)

Default VAE

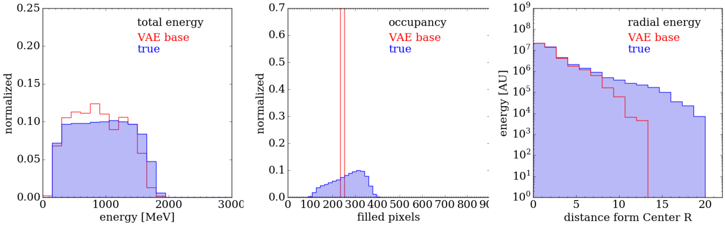
- Standard loss function

$$L = L_{MSE} + L_{KLD}$$



- Center region rather well reproduced
- Sparse structure problematic

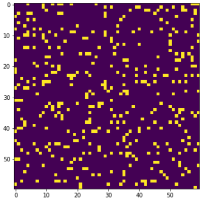
Default VAE

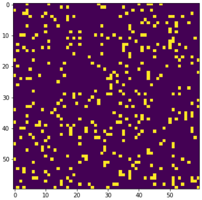


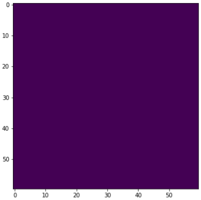
- Energy distribution good
- Fails for occupancy and radial distribution

MSE and sparsity

- Take two random, sparse images
- Take empty image
- Calculate MSE(*sparse*, *sparse*)
- Calculate MSE(*sparse*, *empty*)

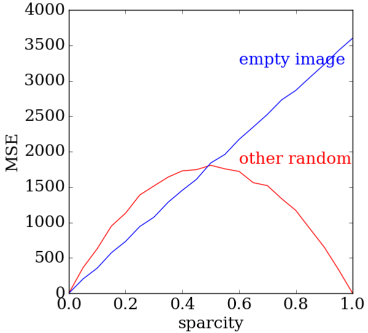






MSE and sparsity

- For very sparse images: Empty gives smaller loss



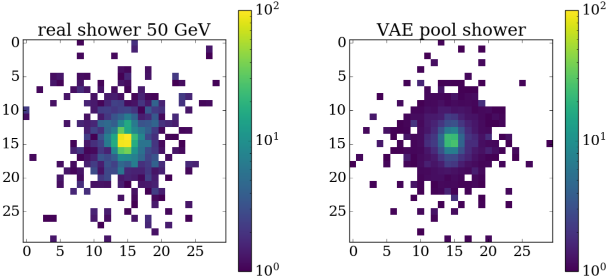
- MSE alone can't reproduce sparse, radial structure

Pooling MSE VAE

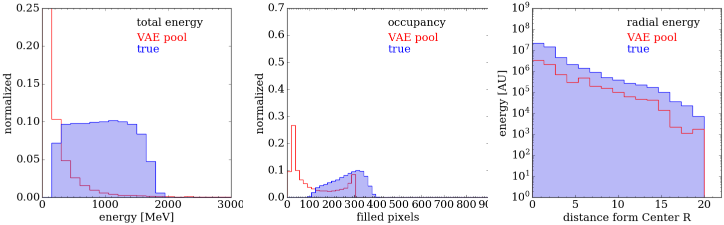
- Apply pooling before MSE

$$L = L_{MSE} + L_{KLD} + L_{\Delta E} + L_{MSE_{3\times3}} + L_{MSE_{5\times5}} + L_{MSE_{10\times10}}$$

- $L_{MSE_{10\times10}}$: 10×10 avg-pooling before MSE calculation
- Additional term for depos. energy

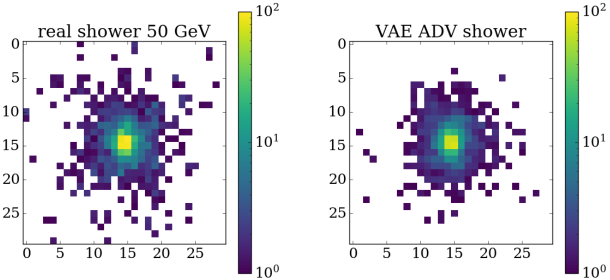


Pooling MSE VAE

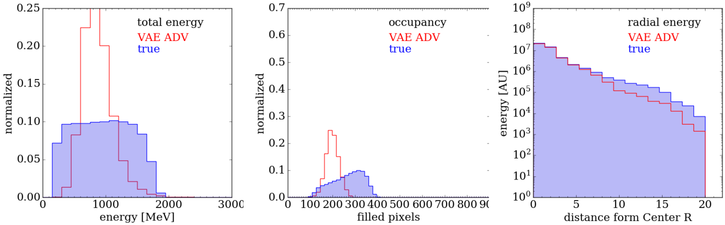


- Improved occupancy and radial distribution
- Energy distribution way off

- Use adversarial network trained to differentiate original and reconstructed image as loss function



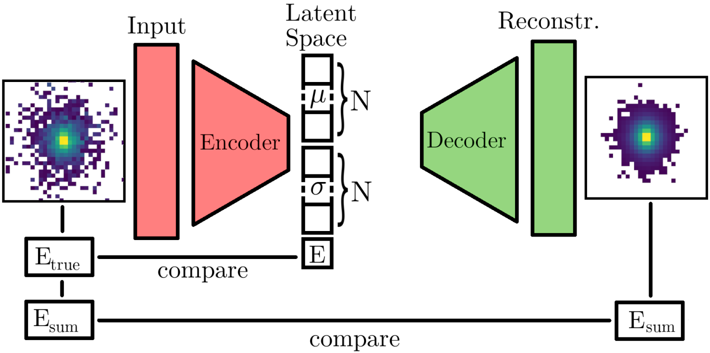
Adversarial VAE



- Occupancy and radial distribution decent
- Energy distribution now approx Gaussian

Energy Conditioned VAE

- Additional latent variable for Energy

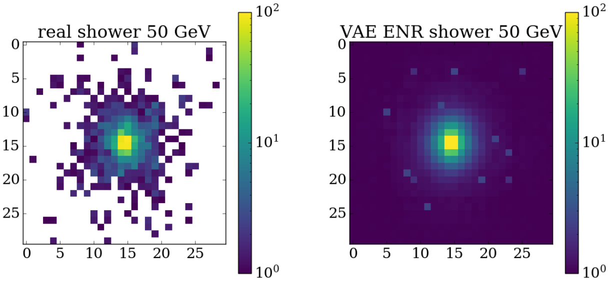


- Energy can be any distribution
- Can request shower for given energy

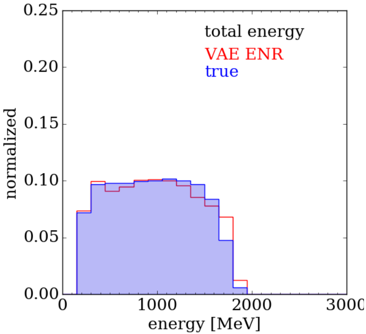
Energy Conditioned VAE

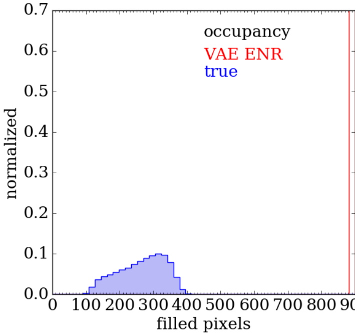
- Convolutional/transpose-convolutional architecture

$$L = L_{MSE} + L_{Pool_{10\times10,s=5}} + K_{KLD} + L_{MMD_{radial}}$$
$$+ L_{\Delta(E_{incoming},E_{predict})} + L_{\Delta(E_{sum},E_{sum-recon})}$$

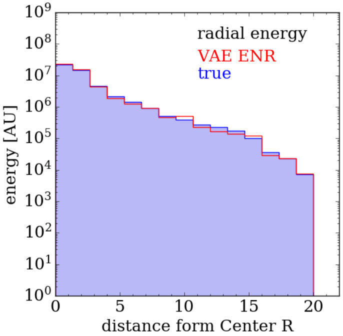


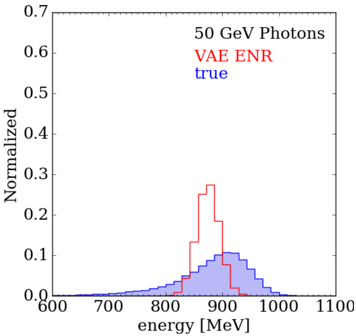
Energy Conditioned VAE





Energy Conditioned VAE





Outlook

- Low energy pixel cutoff
- Variance per pixel in loss
- Loss term for occupancy
- Adversarial plus conditioning