Using conditional variational autoencoders to generate images from atmospheric Cherenkov telescopes

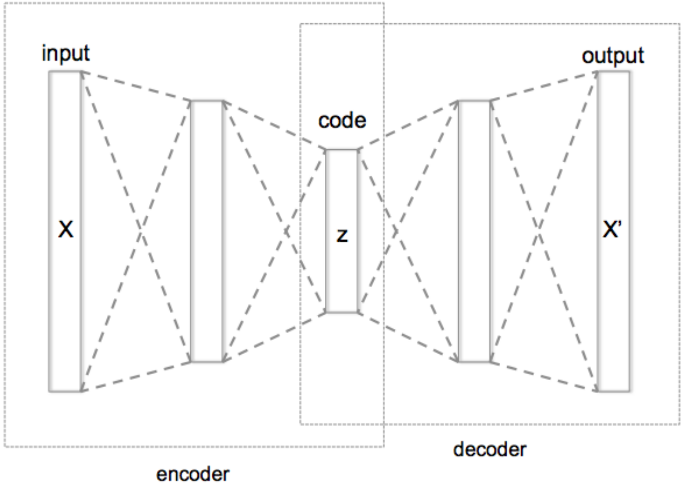
Stanislav Polyakov[1], Alexander Kryukov[1], Andrey Demichev[1], Yulia Dubenskaya[1], Elizaveta Gres[2], Anna Vlaskina[3]

[1] *Skobeltsyn Institute of Nuclear Physics, Lomonosov Moscow State University*
[2] *Irkutsk State University*
[3] *Lomonosov Moscow State University*

The work was funded by the Russian Science Foundation (grant No. 22-21-00442).

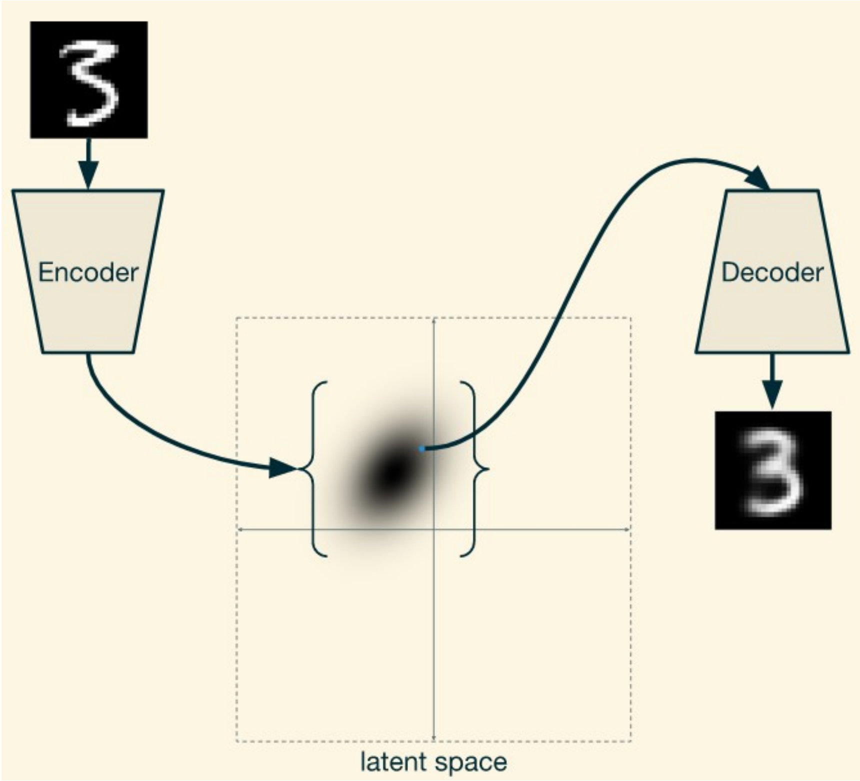
Autoencoders

Autoencoders are artificial neural networks used to learn efficient codings of the input data. An autoencoder consists of an encoder that maps the input into a low-dimensional code, and a decoder that attempts to reconstruct the input based on its code. Autoencoders learn to ignore insignificant data and are useful in e.g. image denoising or image compression.



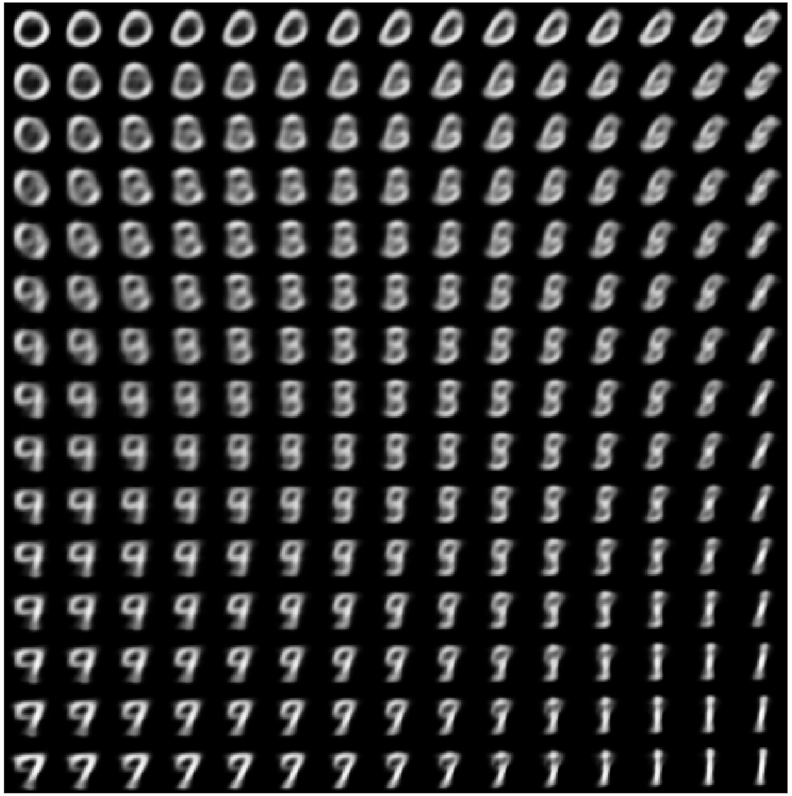
Variational autoencoders

Variational autoencoders use a similar idea but instead of a specific code the encoder finds a set of parameters of a distribution which is then sampled to get a code for the decoder.



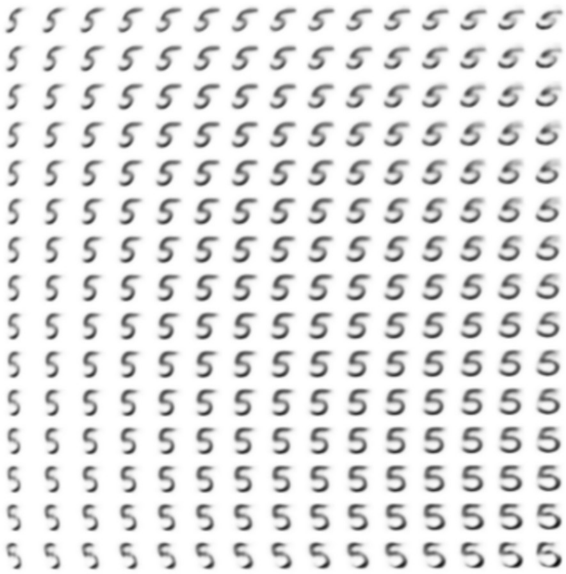
Variational autoencoders

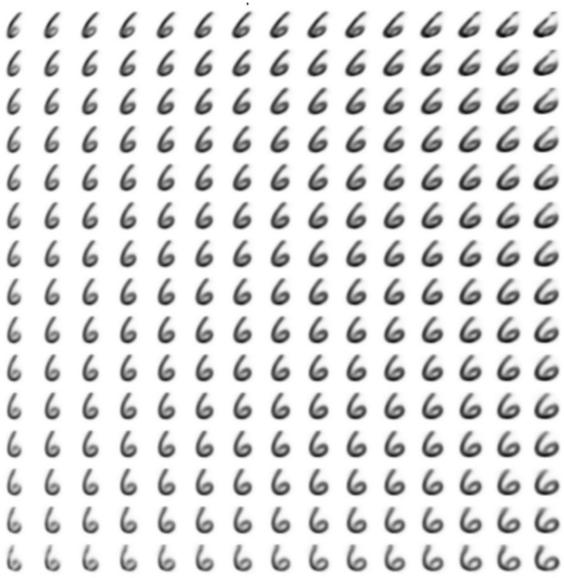
Variational autoencoders can be used to generate new data by interpolating the input data.



Conditional variational autoencoders (CVAE)

Variational autoencoders can be trained with additional parameters passed to both encoder and decoder. The parameters can be discrete (e.g. the digit or the particle type) or continuous (e.g. energy). Parameter values can be then passed to the decoder to generate new data.



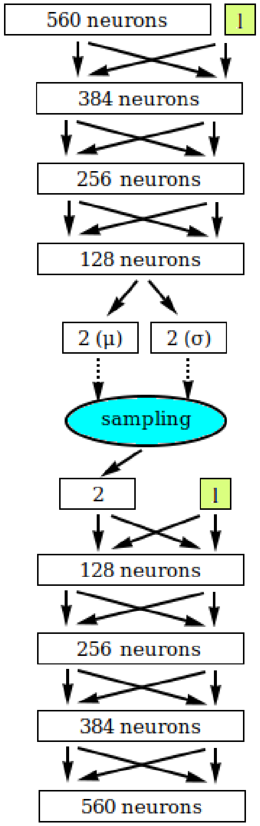


Input data

As our input data, we used Monte Carlo-simulated images of atmospheric events caused by high-energy gamma quanta and recorded by Cherenkov telescopes of the TAIGA experiment. The sample included 39443 images from a single telescope. The energy of the gamma quanta was between 1.5 TeV and 60 TeV.

The TAIGA IACT located in Tunka valley, Russia

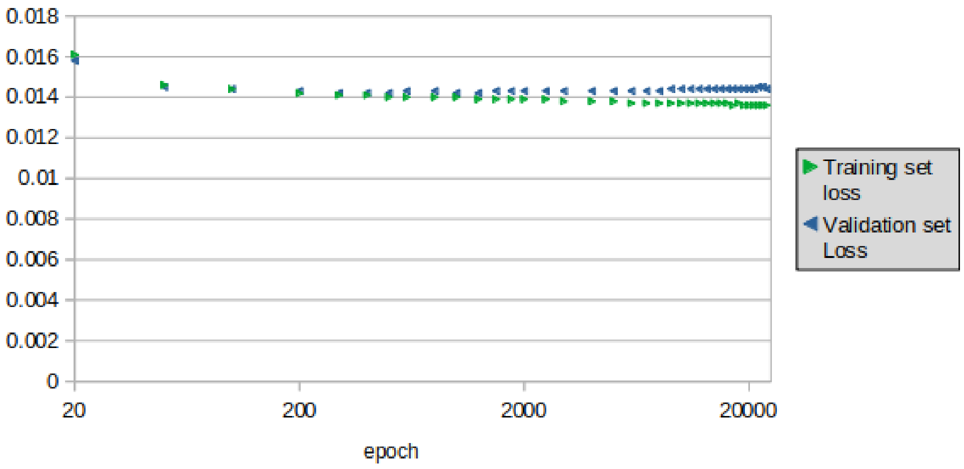
CVAE architecture example



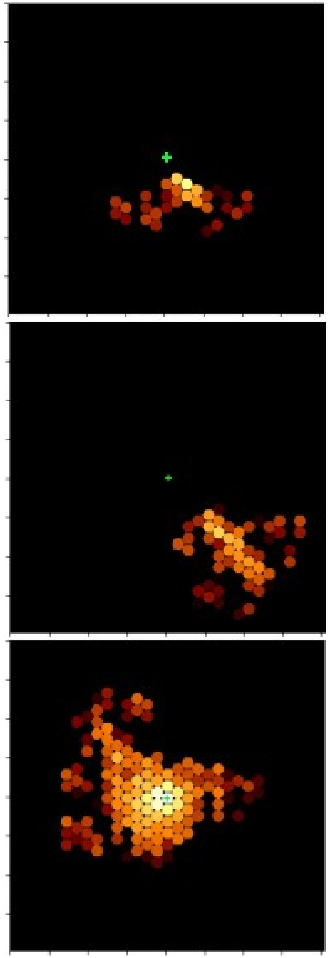
Conditional variational autoencoder training

The parameter we passed to the conditional variational autoencoders during training was the logarithm of the sum of the pixel values (called size).

The autoencoders were implemented in Keras and trained on 29568 gamma event images, with 9856 images used for validation.

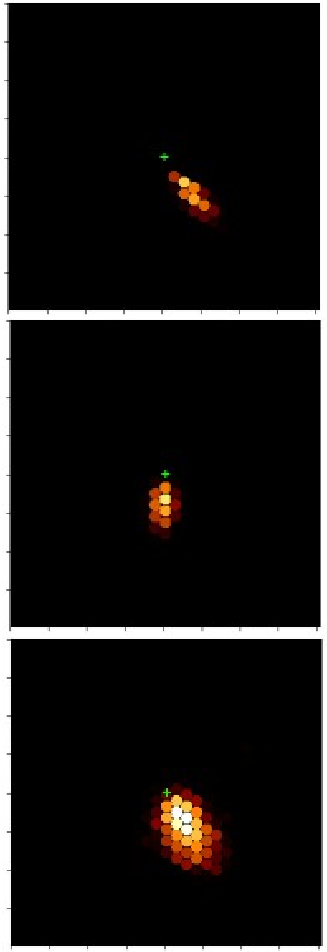


Images



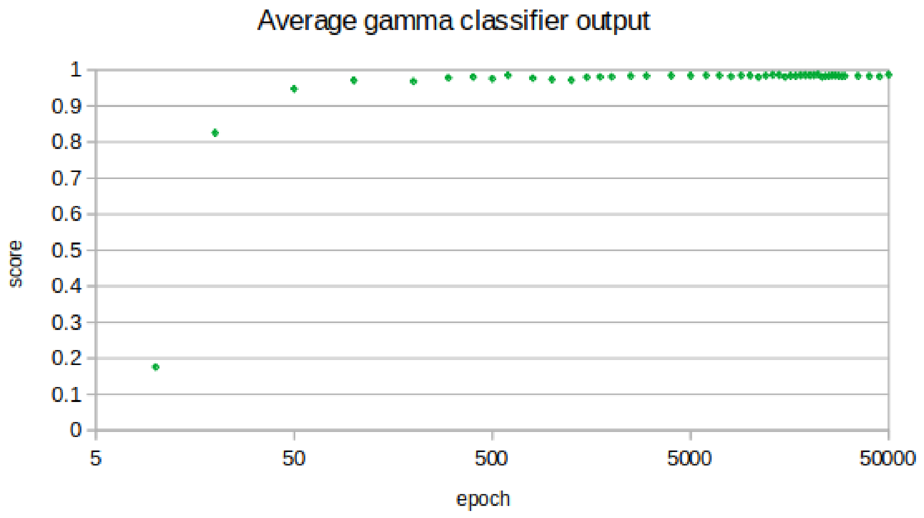
Monte Carlo-simulated events (left)

CVAE-generated events (right)



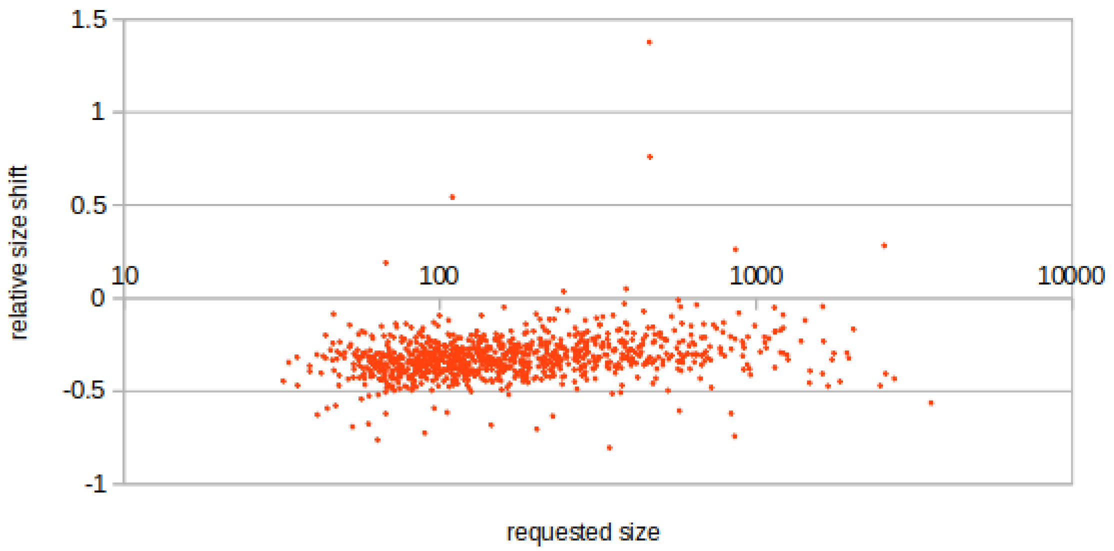
Classifier results

We used a convolutional neural network classifier trained to distinguish between gamma events and proton events to evaluate the generated images. For one of the CVAEs, the classifier score for generated images was between 0.98 and 0.987. Monte Carlo-simulated images of gamma events have the average score of 0.99.

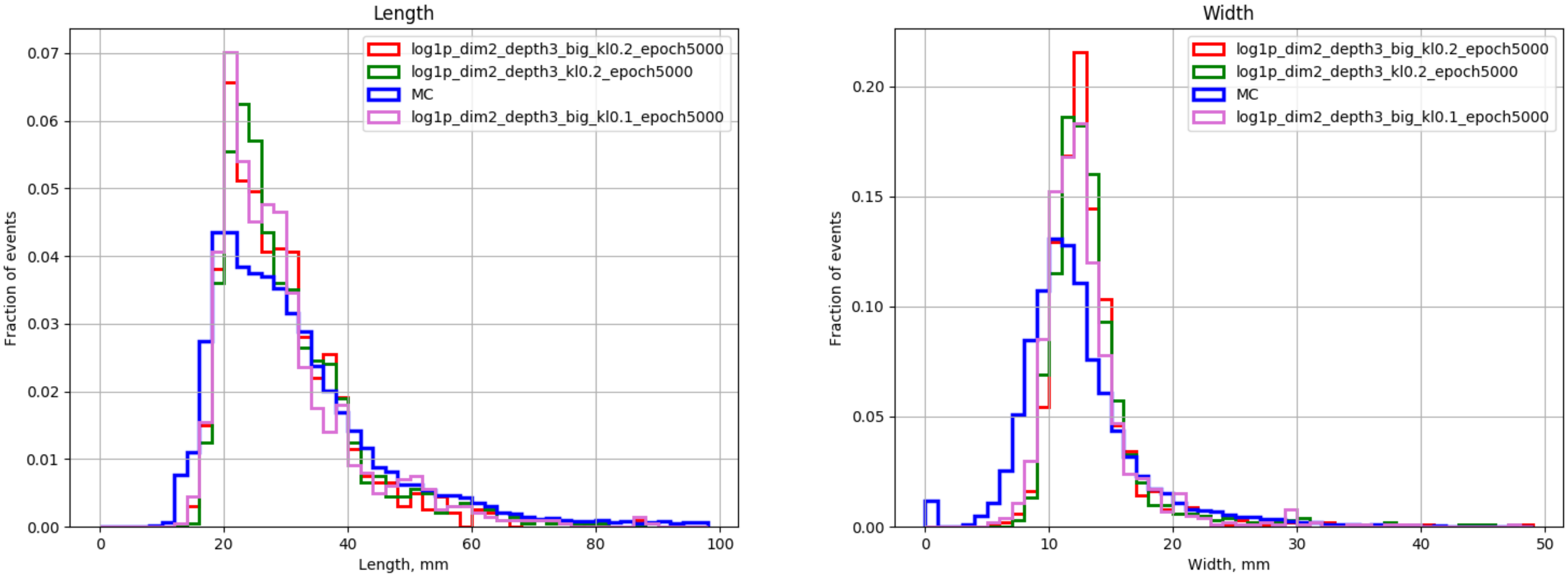


Sizes of generated images

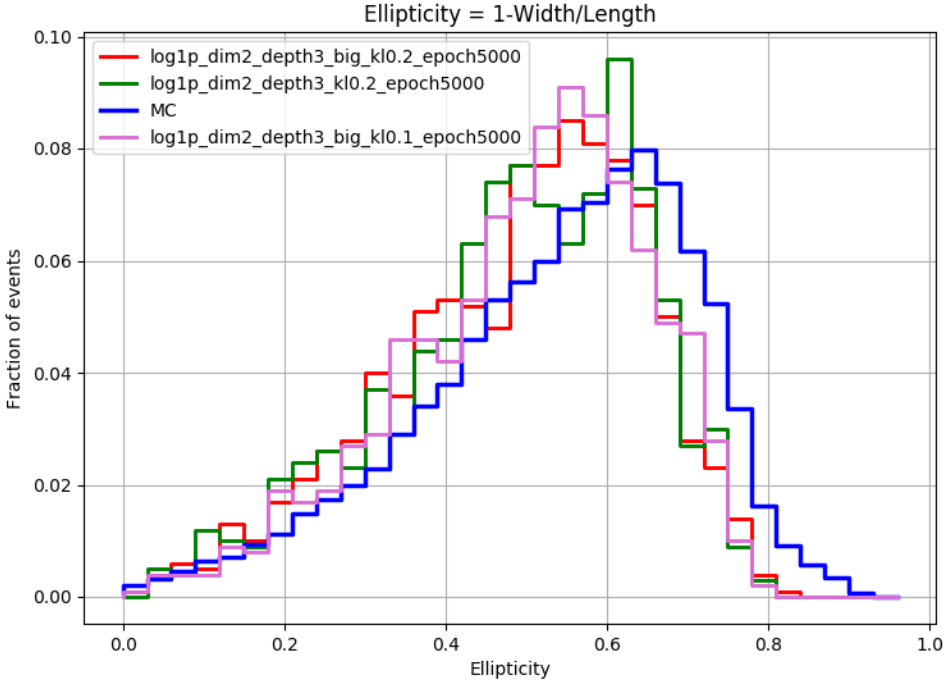
So far there is a mismatch between the sizes passed to CVAE decoders as a parameter and the size of the generated images.

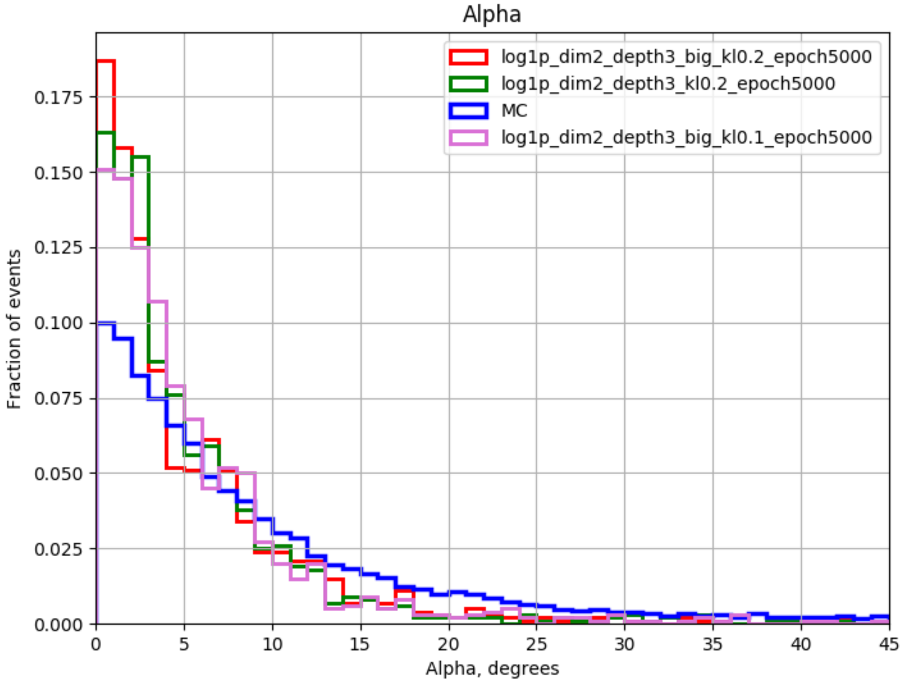


Hillas parameters



Hillas parameters





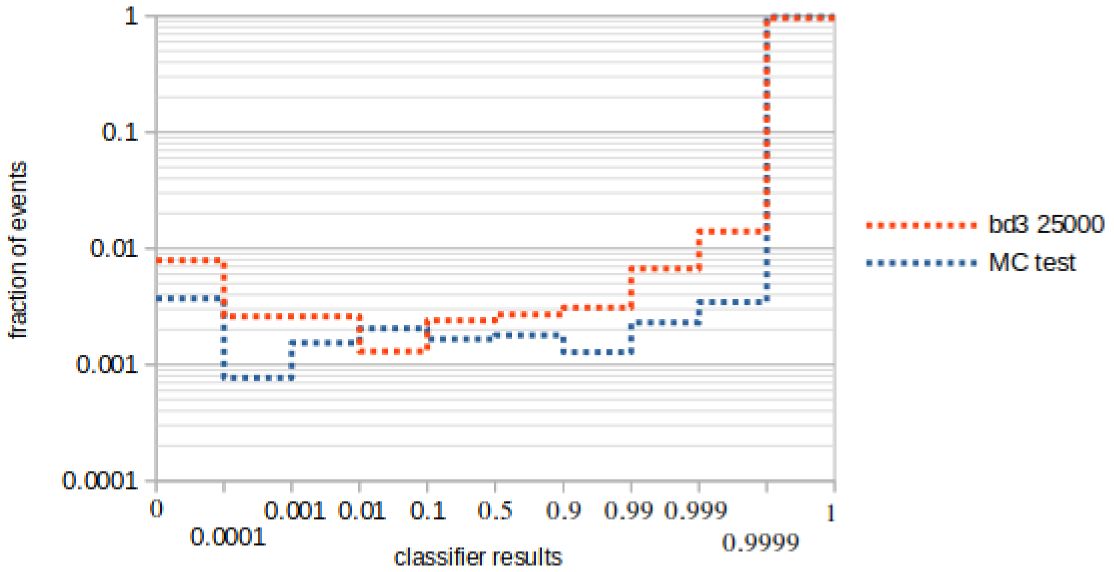
Results

We trained conditional variational autoencoders (CVAEs) using a set of Monte Carlo-simulated images of gamma events.

- Images generated by the trained CVAEs are similar to the Monte Carlo images, including very high score by a classifier trained to distinguish images of gamma events.
- However, some of the parameters of the generated images are distinct from the parameters of the Monte Carlo-simulated images.

Autoencoder codes

Classifier scores



Hillas parameters