Generative Models for Physics

MMCP 2024

Yerevan, October 24, 2024

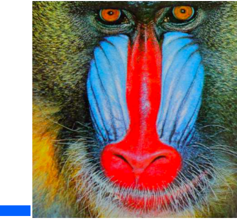
Fedor Ratnikov

NRU Higher School of Economics,

ML for Natural Science

- ML (a.k.a. AI) is widely used nowadays in many different areas
- Models available on the market are tuned for different tasks and different data specifics
  - Wolpert, D. H.; Macready, W. G. (1997). "No Free Lunch Theorems for Optimization". doi:10.1109/4235.585893
- Most problems are human driven:
  - get me from A to B by the fastest way
  - paint a pleasant picture for me
  - translate a text to/from my language
  - evaluate a credit score of my client
  - …
- Problems of the natural science are driven by the **Mother Nature**
  - problems, conditions, limitations, data specifics etc. are driven externally

Example: Science vs Aesthetics



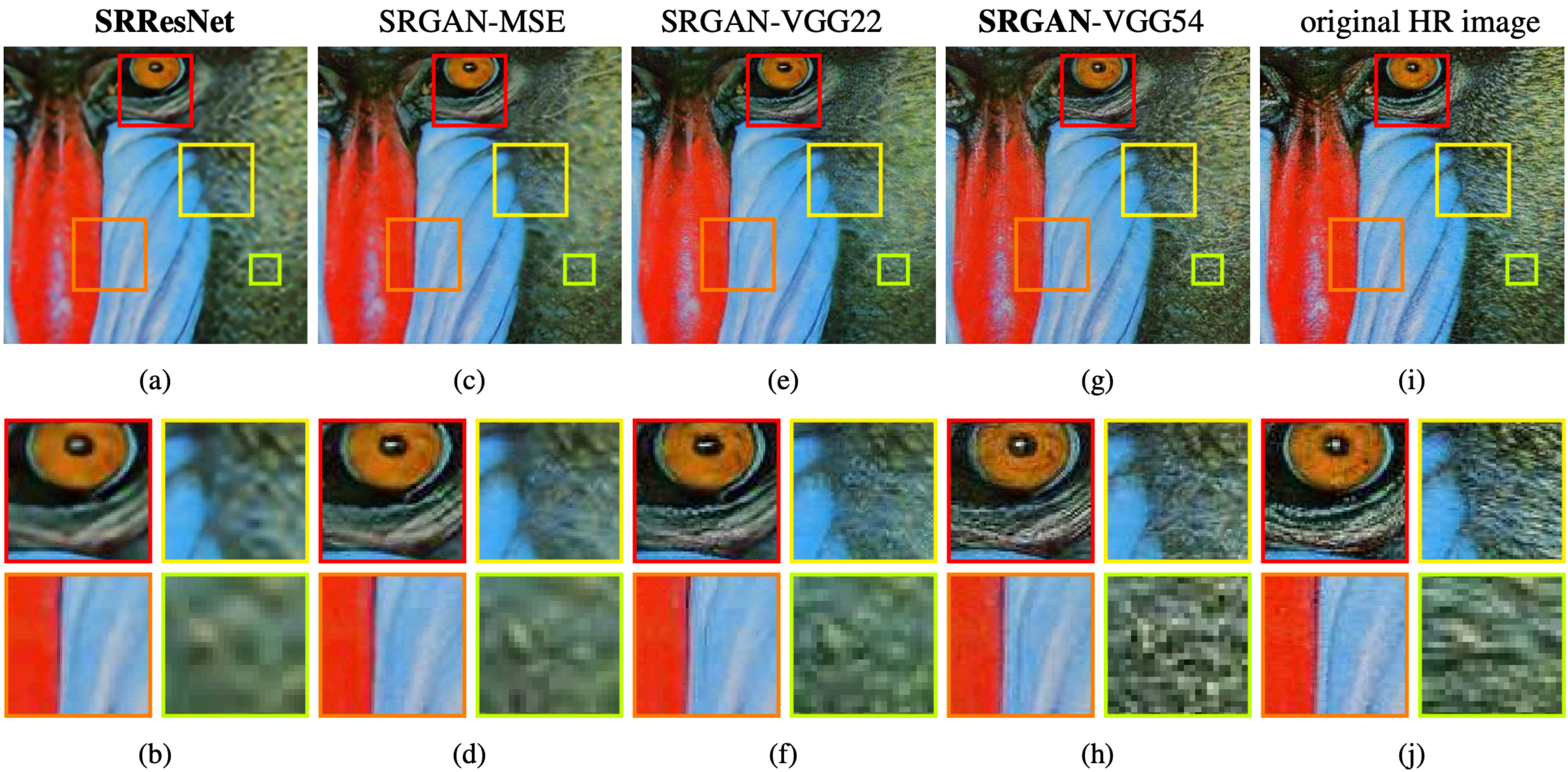


Figure 6: **SRResNet** (left: a,b), SRGAN-MSE (middle left: c,d), SRGAN-VGG2.2 (middle: e,f) and **SRGAN**-VGG54 (middle right: g,h) reconstruction results and corresponding reference HR image (right: i,j). [4× upscaling]

C. Ledig et al., ", doi: 10.1109/CVPR.2017.19.

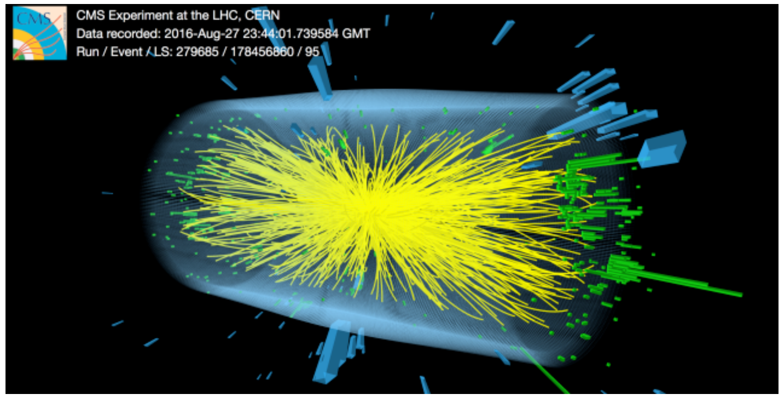
- SuperResolution - not how it actually looks, but how it **could look like**

Why Generative Models?

- The scientific research is essentially an inference
  - mathematically it is an **inverse problem**
    - we evaluate intrinsic parameters from external observables
- The Maximum Likelihood is a practical approach for inference
  - requires a likelihood function to maximize
- No explicit likelihood function for complicated problems
  - too many intermediate intrinsic stated to marginalise
- Solve **inverse problem iteratively via direct problem**
  - simulate and sample stochastic external observables for a given intrinsic parameters
    - evaluate the best intrinsic parameters by best matching between simulated and experimentally obtained observables

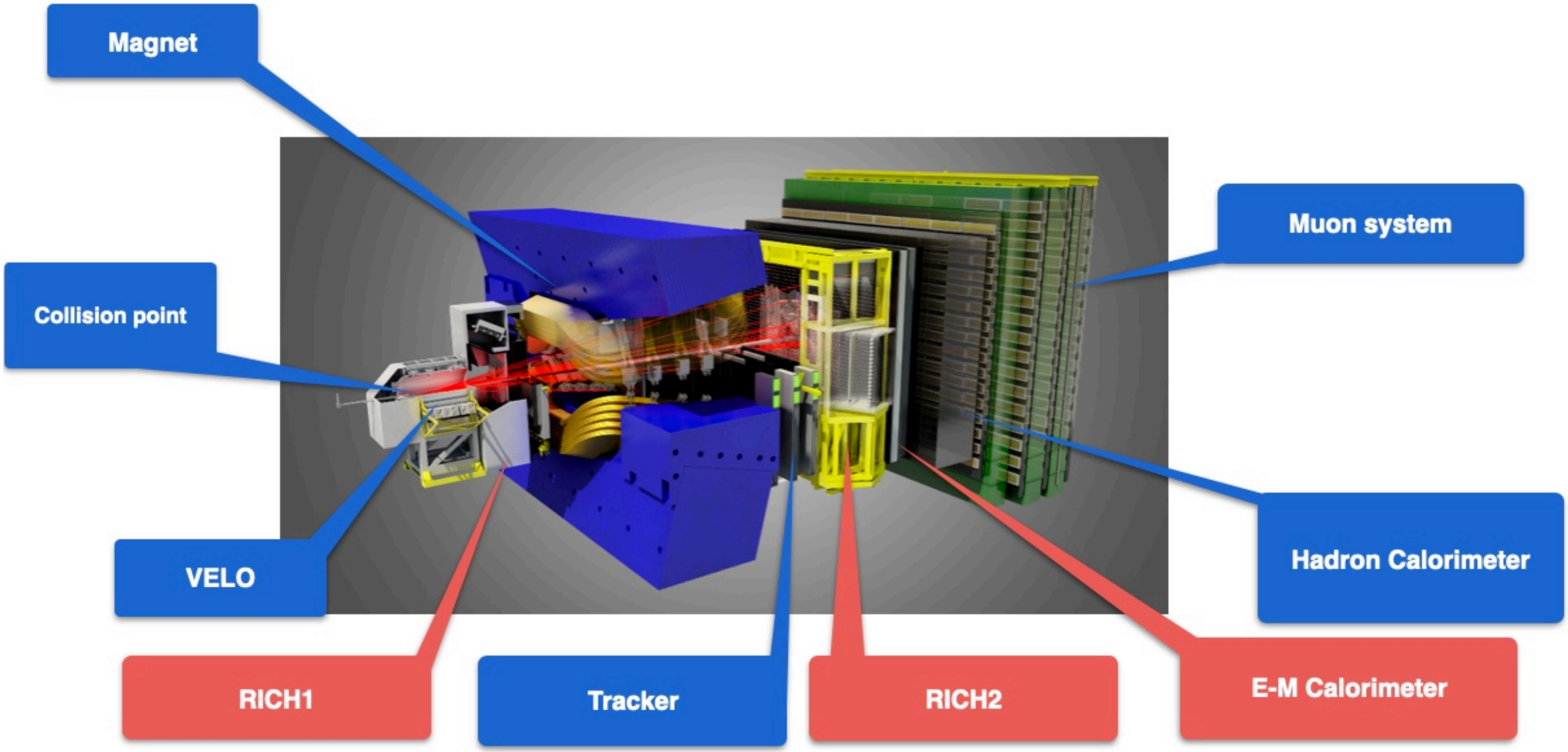
Generative Models in HEP

- More than 80% of HEP computing resources is dedicated to simulation of experimental data
- Moore's law is saturated since 2010th
  - new approaches are required



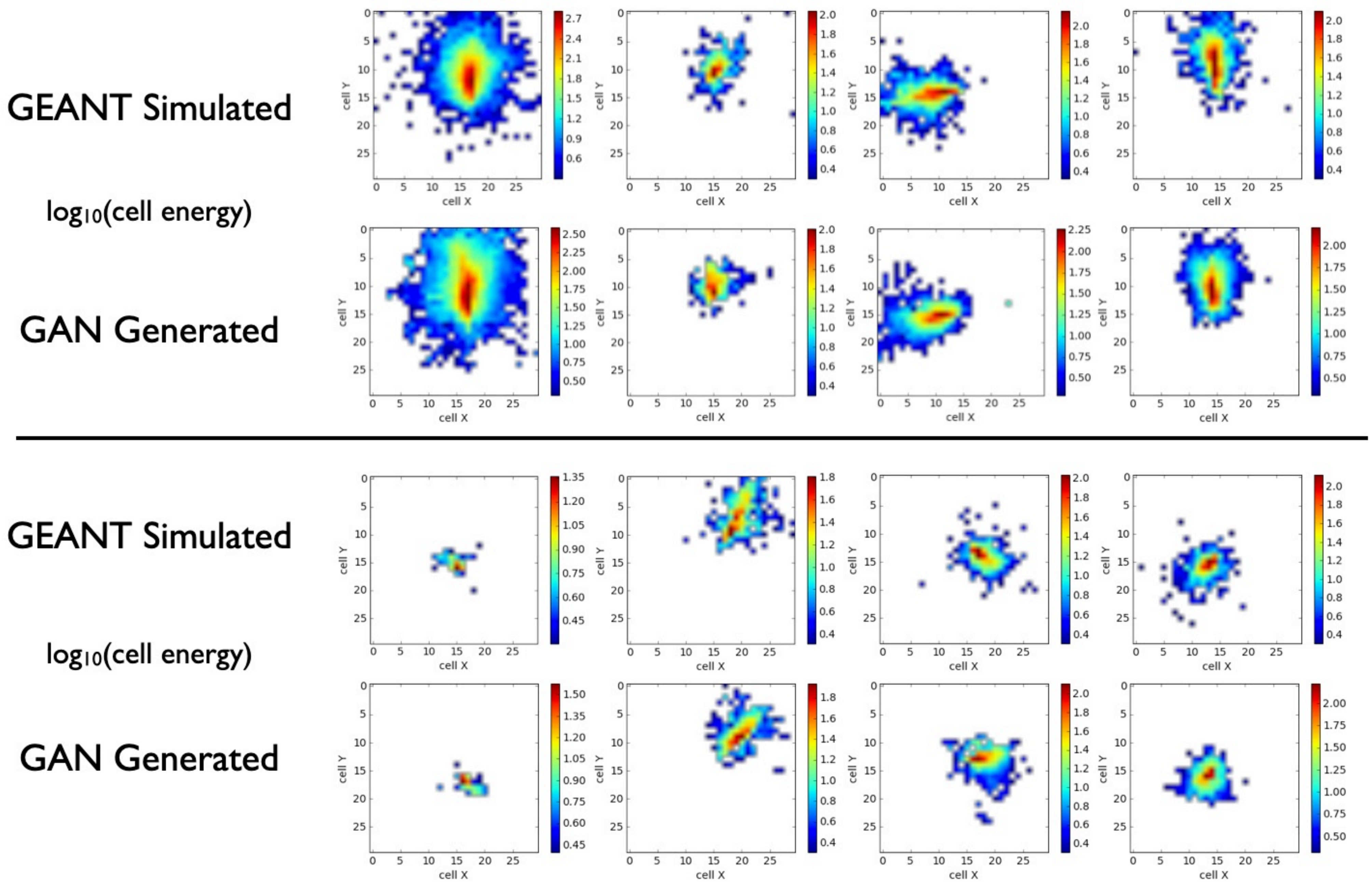
- Physics simulation i.e. GEANT for HEP is very detailed and thus slow
- Physics simulation it is just a stochastic function to describe macroscopic response of the detector for a particle with very few parameters
  - which is calculated on the microscopic simulation level
- The idea then is to train a **simple and fast surrogate generative model** to reproduce that **macroscopic stochastic function**

Physics



Magnet
Collision point
VELO
RICH1
Tracker
RICH2
Muon system
Hadron Calorimeter
E-M Calorimeter

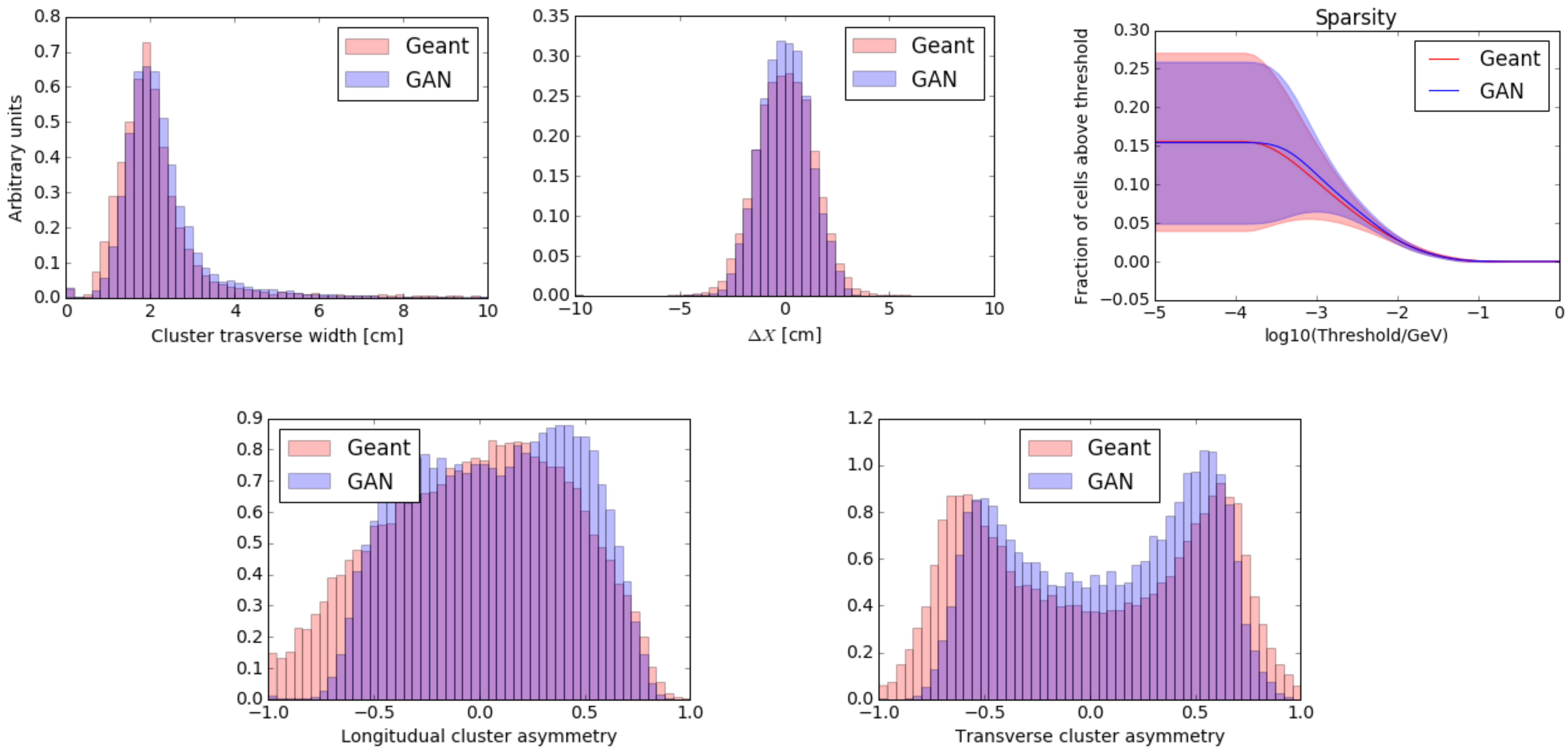
Physics vs Aesthetics



V.Chekalina et al,
doi: 10.1051/epjconf/201921402034

- Visually pleasant images
- What about **physics goodness?**

Primary and Marginal Distributions



- **Is hard to fit marginal distributions**

Scientific Requirements

- For image generation we are usually happy if the result **looks** like it is desired
- In science we need the result to match the given set of requirements reasonably well. Requirements are driven by **physics considerations** closely connected to the ultimate physics goal

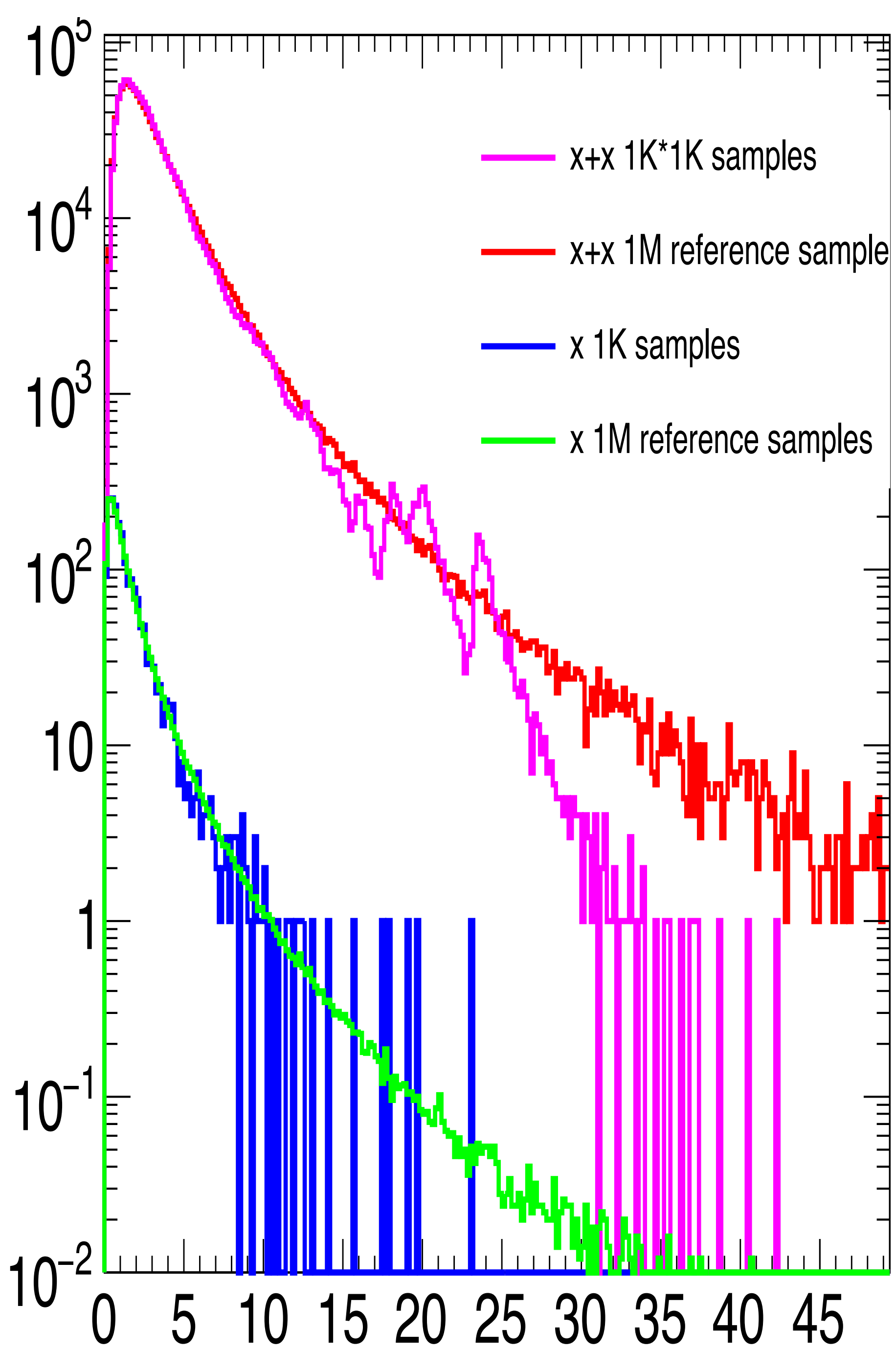
Physics-driven Model Training Specifics

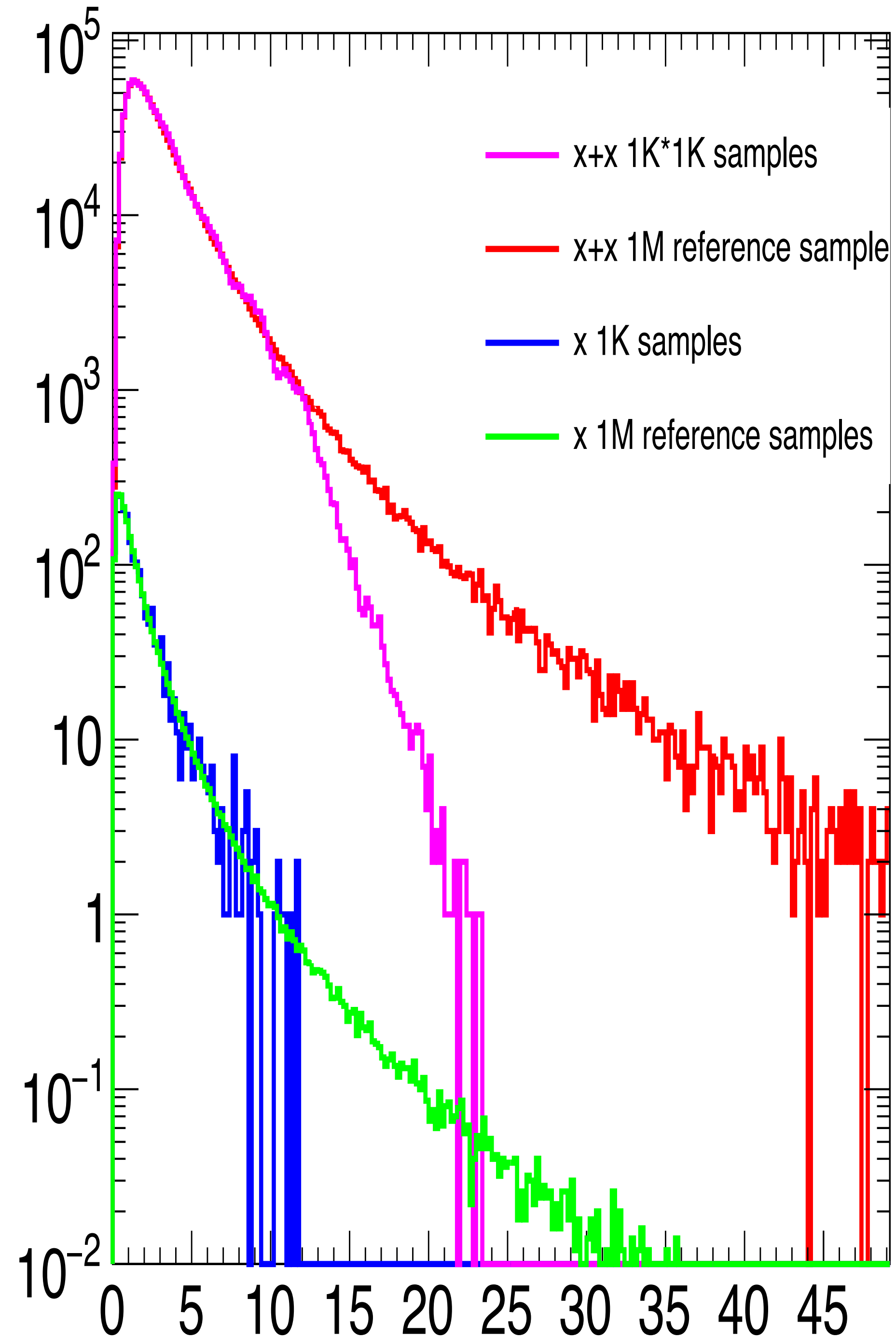
- For scientific use, generative models are required to ensure high quality for specific physics-driven metrics
- Neither generative model is ideal
  - the training procedure is agnostic, thus it doesn't care of physics metrics
  - some of them may be reproduced by the model well, some may not
- How can a good quality for the **specific metrics** be ensured?

Surrogate models for Physics Simulations

- Why do we need fast simulation?
  - to sample many data for reasonable computing resources
- Why do we need many data?
  - to study very fine and/or rare effects in data
- Fine and rare effects mean much information (entropy) in data
- But the surrogate generative model trained on the given dataset contains only information from this dataset ab initio
  - thus train sample needs to be big enough
- **To train a surrogate model to describe data with necessary precision, one would need more train data than it would be necessary to directly provide that precision**

Surrogate models for Physics Simulations





- Quality of the generative models is **limited by the size of the train data sample**

Statistical Limitations

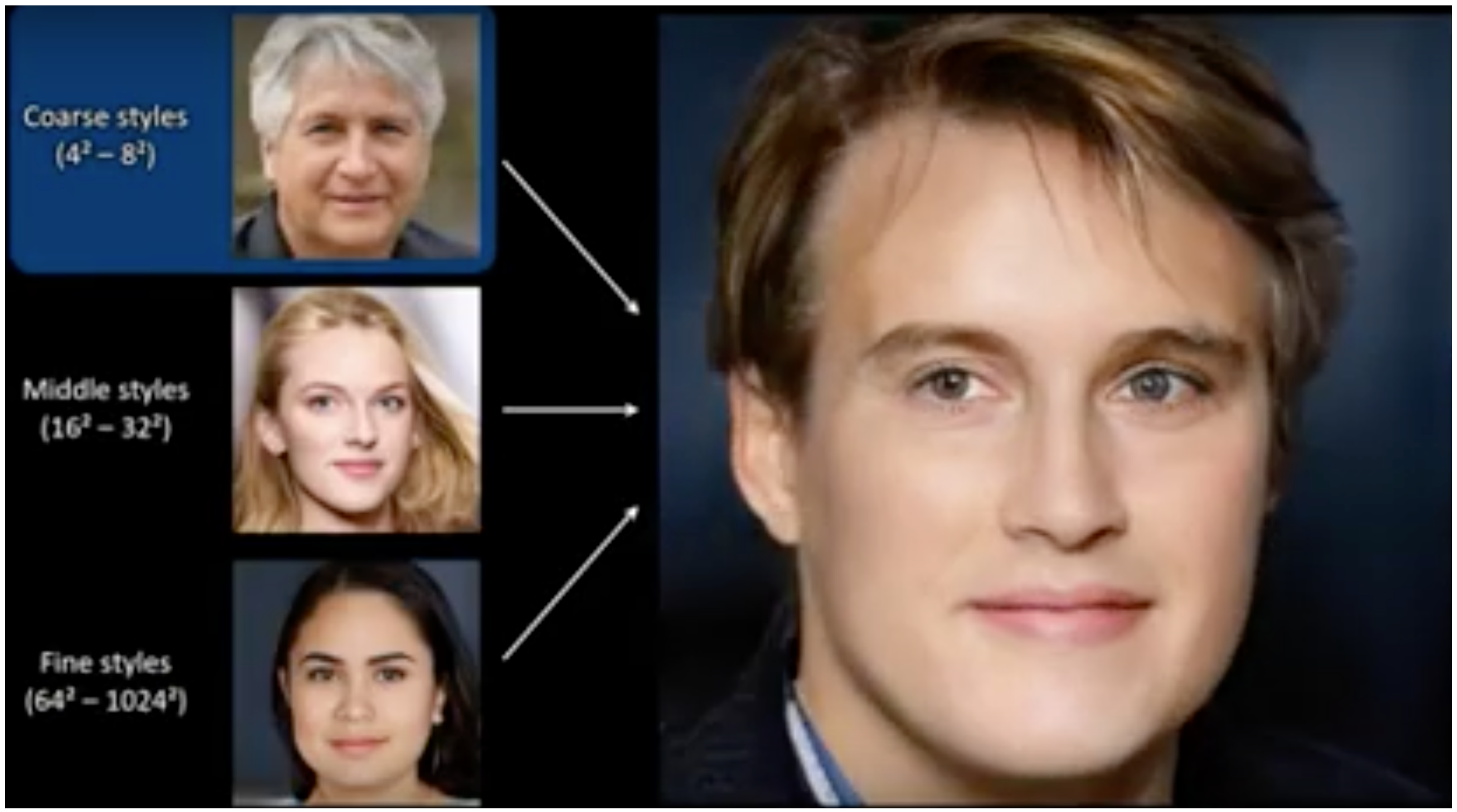
- Quality of the generative models is limited by the size of the train data sample
  - generative models can not give a profit for producing statistically correct big data sets:
    - **no extra information** beyond the train sample is available
    - model systematics corresponds to the **train sample statistics**

Discussion

- Does all these mean than generative models being used for high level scientific researches are **dead in the water**?

Decomposition

- No information beyond the train sample is available



- Not quite if we can decompose generative model into separate components (yes, it is an extra *a priori* information)
  - random combinations of different components can **drastically increase variability**

Decomposition

- E.g. fast simulation of the calorimeter response
  - generator is trained on 10^6 incident particles
  - $\sim$50 particles in the calorimeter per event
  - total variability $\sim(10^6)^{50} = 10^{300}$! (NB intrinsic correlation)

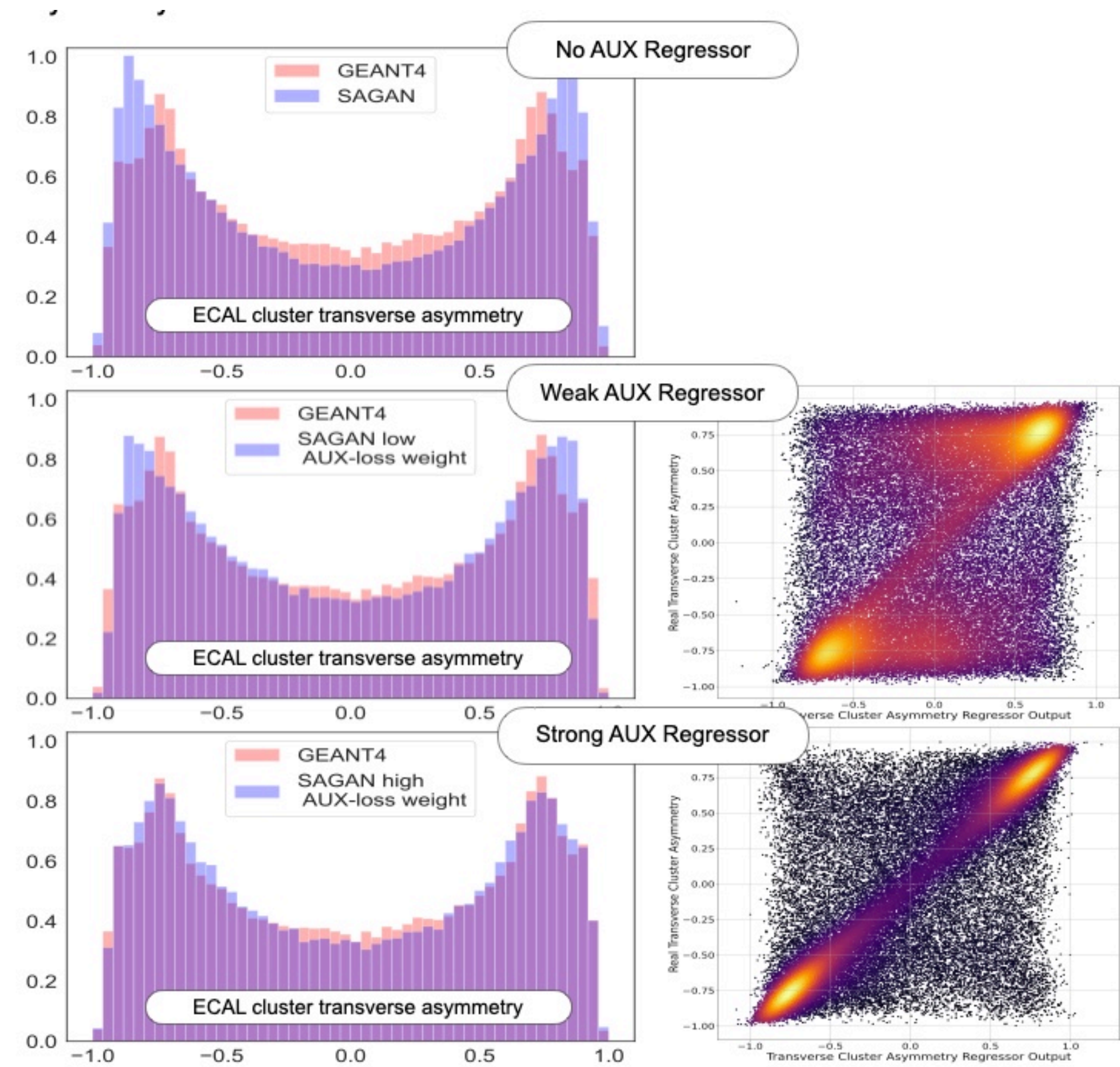
Training Prioritization

- How can a good quality for specific metrics be ensured?
    - if metric is simple, just add it to the training loss
- What if metrics can not be converted into computational graph?

Surrogate Regressor

A.Rogachev, FR,
doi: 10.1088/1742-6596/2438/1/012086

- If metrics is complicated, substitute it by the surrogate regressor trained on this metrics
    - this regressor may be incorporated into e.g. GAN discriminator
- To improve model metrics quality the surrogate regressor doesn't need to be very good
    - regressor is used for both train and generated data, thus its errors cancels



Conclusion

- Generative models provide powerful tools of the modern civilization
- Use of ML driven surrogate models may help to significantly reduce computing resources needed for different studies in natural sciences
- There are specific requirements to such models which are not addressed by the CS community beyond the scope of the scientific use
- These specifics are not show-stoppers but require extra knowledge from the field, i.e. require interdisciplinary efforts