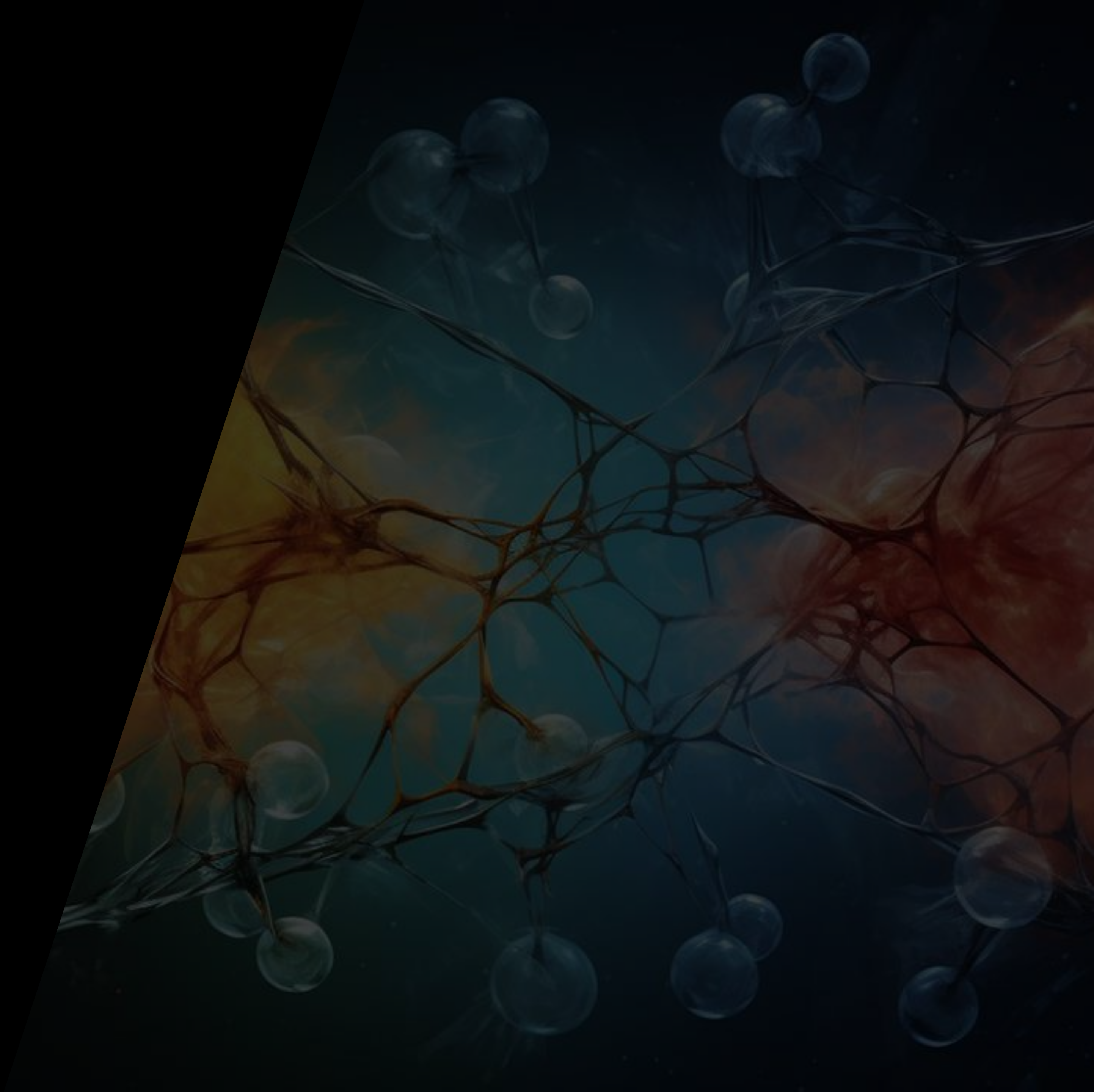
Projects

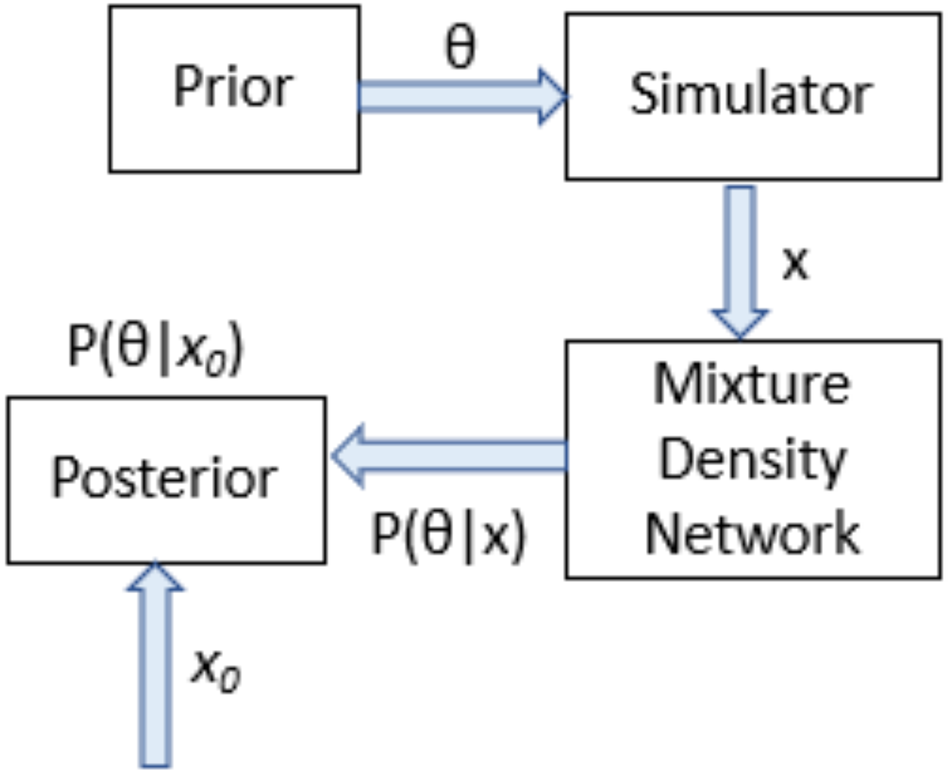


I. Simulation-based inference

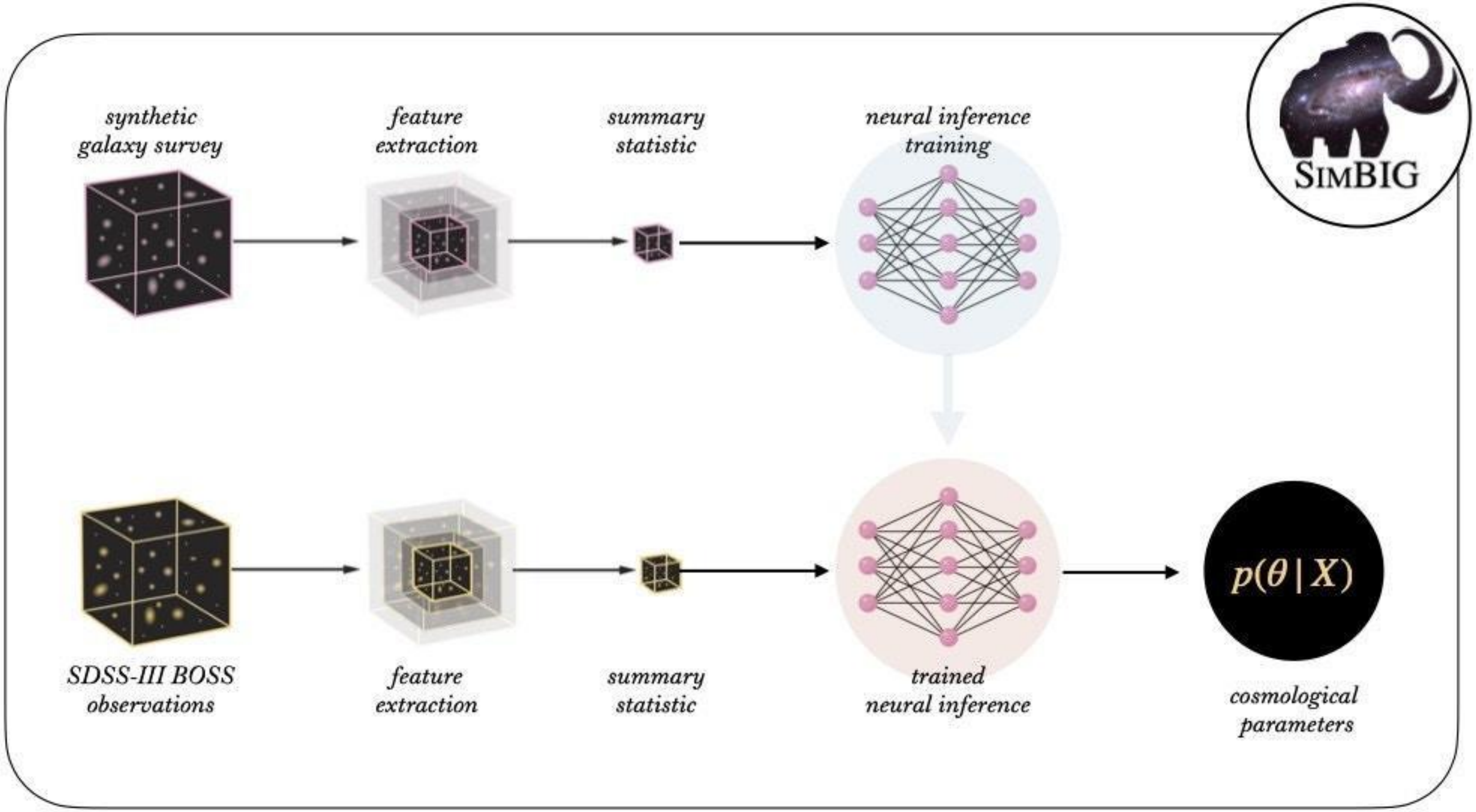
Bayesian inference : Observe an observation of a physical system x_o. This system depends on some parameters θ. Try to find the posterior distribution of $p(\theta|x_o)$

How : Access to a simulator of the system $\theta \rightarrow x$. Use samples from this simulator to train a density estimation network

Goal : Understand how to use the SBI pipeline on toy model and apply it on a physical problem of your choice



I. Simulation-based inference

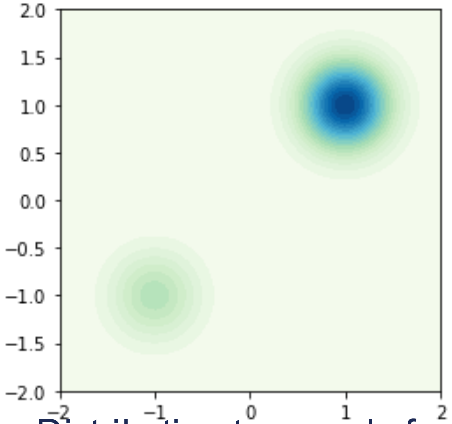


II. Enhancing sampling with machine learning

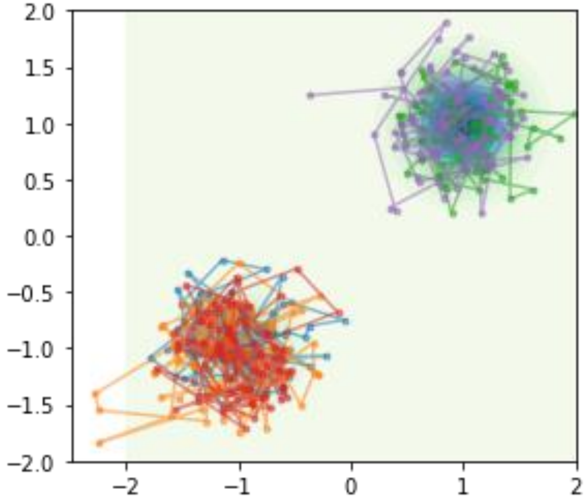
• ***Sampling*** : Wants to draw samples $x_1, \dots, x_n \sim \rho(x)$ with $\rho(x) = \frac{\exp(-U(x))}{Z}$. Difficulty when $\rho(x)$ has two or many modes

How : Use an adaptative algorithm that will use the samples created at each step to train a normalizing flow to match $\rho(x)$. This flow will then be used to create new samples

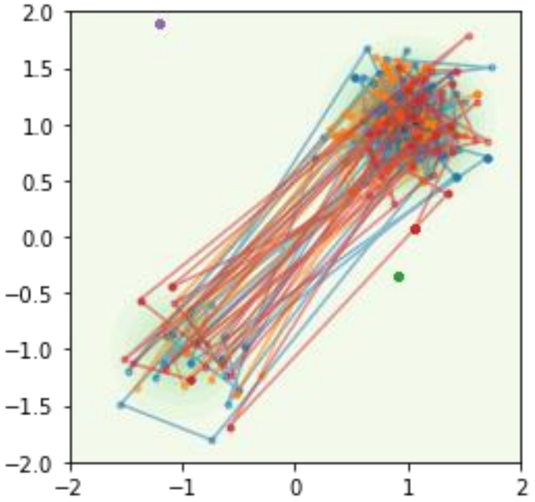
Goal : Code this method on toy models and identify its drawbacks and advantages. Try to use it on a multimodal physical system



Distribution to sample from



Sampling with local methods : no jumps between modes



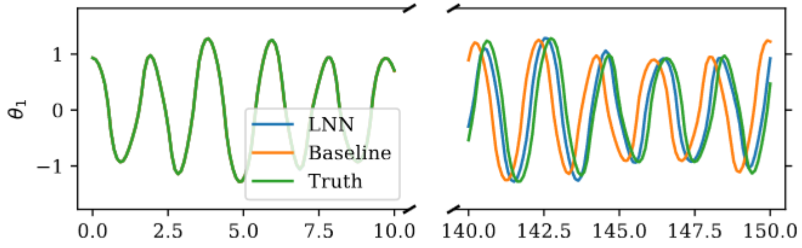
Sampling with adaptative algorithm: jumps between modes

III. Physics-inspired neural networks

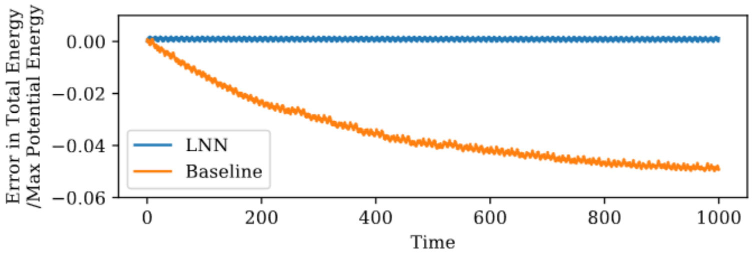
- ***Why ?*** : NNs are black box function. Sometimes yields non-physical results. Requires large amount of data that may not be available.

  How : Incorporate domain knowledge from physics (symmetries, equations, conservations, ...) directly in the architecture of NNs to restrict the space of admissible solutions.

  Goal : Understand the differences between physics inspired NNs and classical NNs. Try to reproduce the result from a paper and apply the technique to another physical system.



Lagrangian neural net on double pendulum problem Cranmer+2020



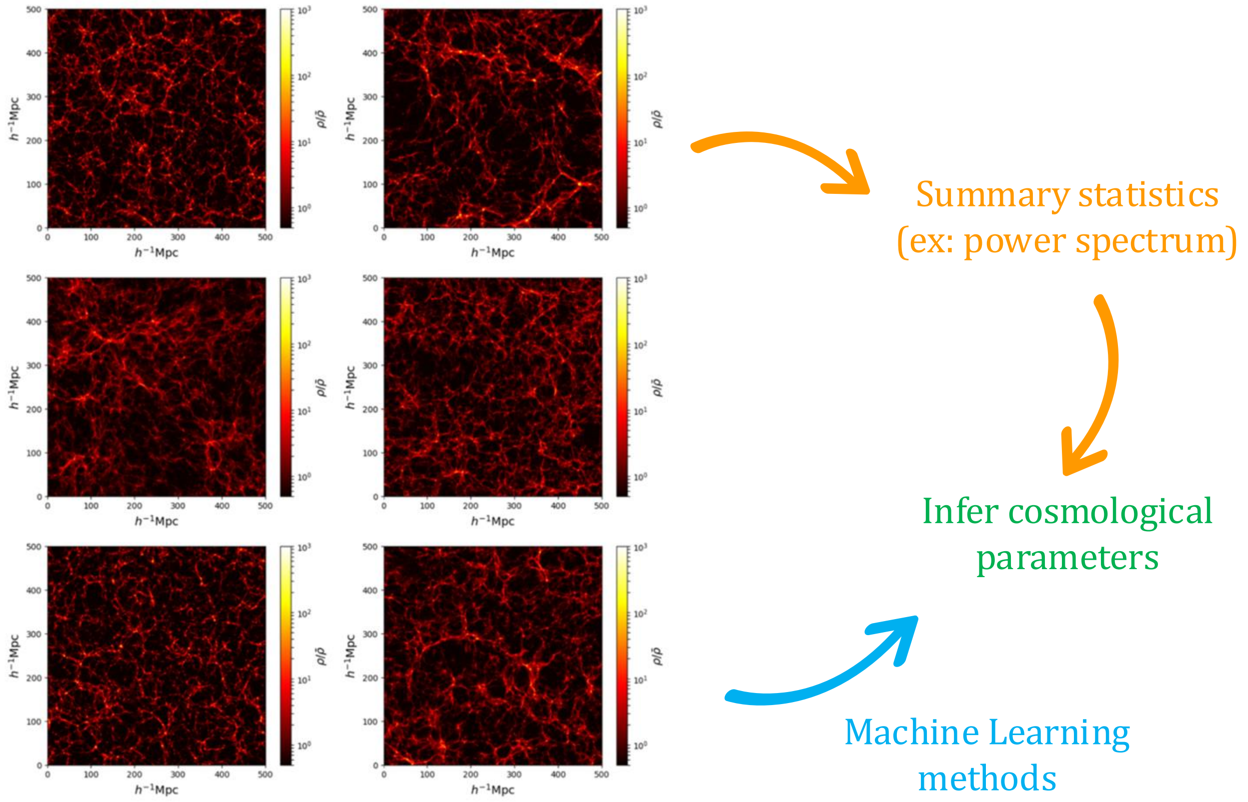
Error in energy for LNN and regular NN Cranmer+2020

IV. Estimating cosmological parameters from matter distribution in cosmological simulations

Keywords: Cosmology, supervised learning, regression, 2D images (or 3D fields).

What you'll do:

- **Find suitable data** to tackle the question,
- Identify and **implement** a machine-learning based solution for predicting cosmological parameters based on 2D (or 3D) fields,
- **Discuss** possible **improvements** and implement them,
- Then different directions possible: how to handle uncertainties, improve the model, add more parameters to infer, etc.



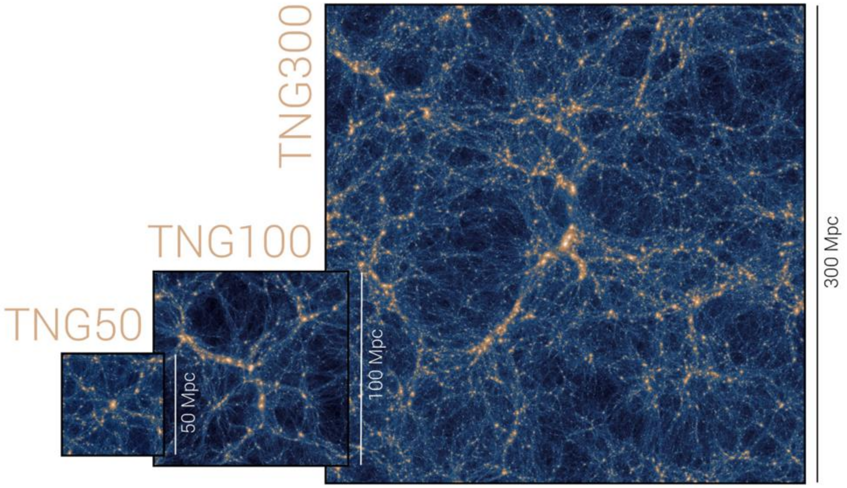
References: Ravanbakhsh+2016, Pan+20, Quijote dataset, CAMELS dataset.

V. Unsupervised Learning by Diffusion

Keywords: Unsupervised learning, sampling, 2D images (or 3D fields).

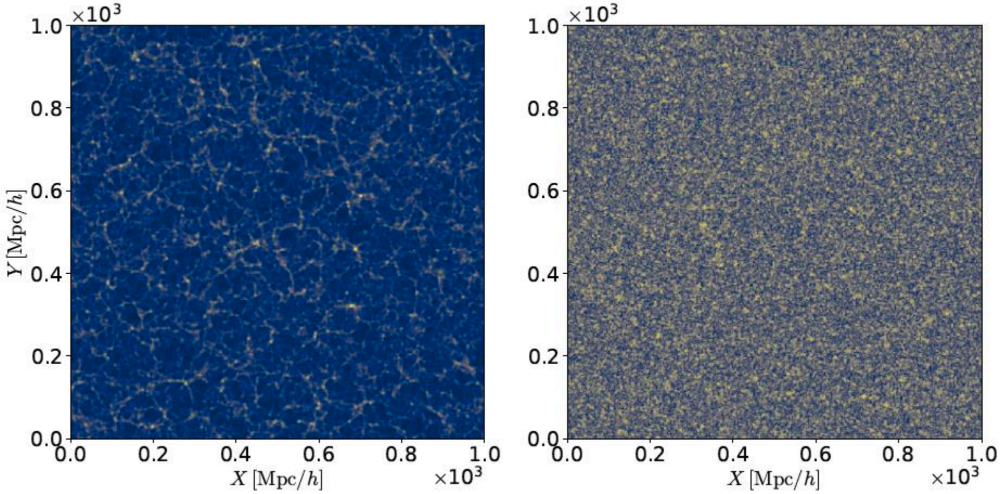
What you'll do:

- Understand the problematic of **unsupervised learning** for sampling,
- **Discuss its application** in physics,
- Understand the maths of **Diffusion Models** and implement a neural-network-based solution for generating samples of 2D fields (cosmology or something else?),
- Discuss shortcomings, evaluation metrics, possible improvements, and alternative architectures.



Several months of runtime!

Naïve power spectrum reconstruction cannot work!



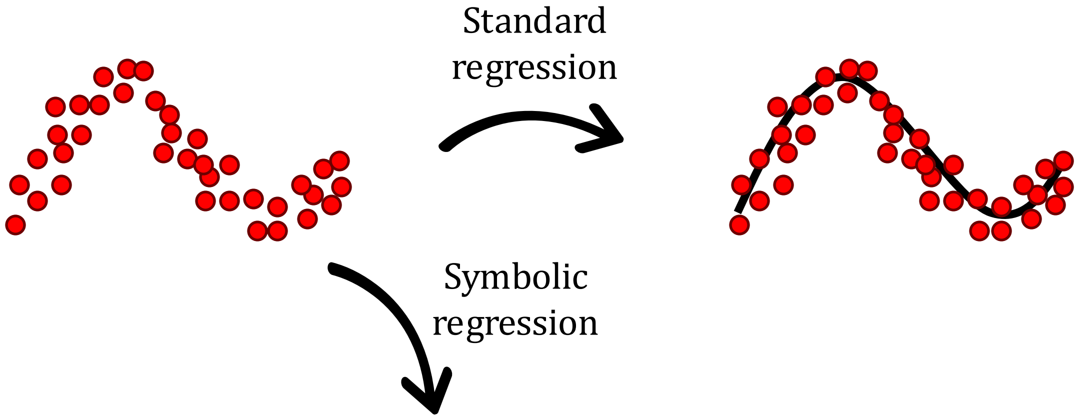
References: Ho+21, Villaescusa Navarro+21, Mudur+23.

VI. (Deep) Symbolic regression for new physics

Keywords: Mathematical physics, Supervised (or unsupervised) learning, symbolic regression.

What you'll do:

- Understand the problematic of **symbolic regression**,
- Discuss its applications in physics,
- Learn how to handle **text data**
- **Implement** a solution for symbolic regression using neural networks,
- **Discuss** possible **extensions** and state-of-the-art methods.



$f(x) = 3\sin(2x+1)$

Applications in all fields of physics, from quantum physics to cosmology

Ideas of data: [Physics expressions]() or Feynman diagrams in quantum physics

References: [Kimmieny+21](), [Alnyqaydan+23]()