HIGlow: High Fidelity Invertible Generative Model for HI Maps

R. Friedman[1] and S. Hassan[2,3]

1. The Rachel and Selim Benin School of Computer Science and Engineering, The Hebrew University of Jerusalem, Israel
2. Center for Computational Astrophysics, Flatiron Institute, New York
3. Department of Physics & Astronomy, University of the Western Cape, South Africa

roy.friedman@mail.huji.ac.il
github.com/friedmanroy

FLATIRON INSTITUTE

THE HEBREW UNIVERSITY OF JERUSALEM

Introduction

Extracting the maximum amount of cosmological and astrophysical information from upcoming large-scale surveys remains a challenge. This includes evaluating the exact likelihood and generating new diverse synthetic examples of the incoming high-dimensional data sets.

In this work, we propose the use of Normalizing Flow (NF) as a generative model of the neutral hydrogen (HI) maps from the CAMELS project [1]. We choose the HI maps since many experiments aim to detect the HI emission as a tracer of large-scale density fluctuations. NF has been very successful at parameter inference and generating new, realistic examples. Unlike previous work (HIFlow [2] using Masked Autoregressive Flow), we use 1X1 invertible convolutions as an inductive bias to utilize the spatial structure of the HI maps following the Glow model [3].

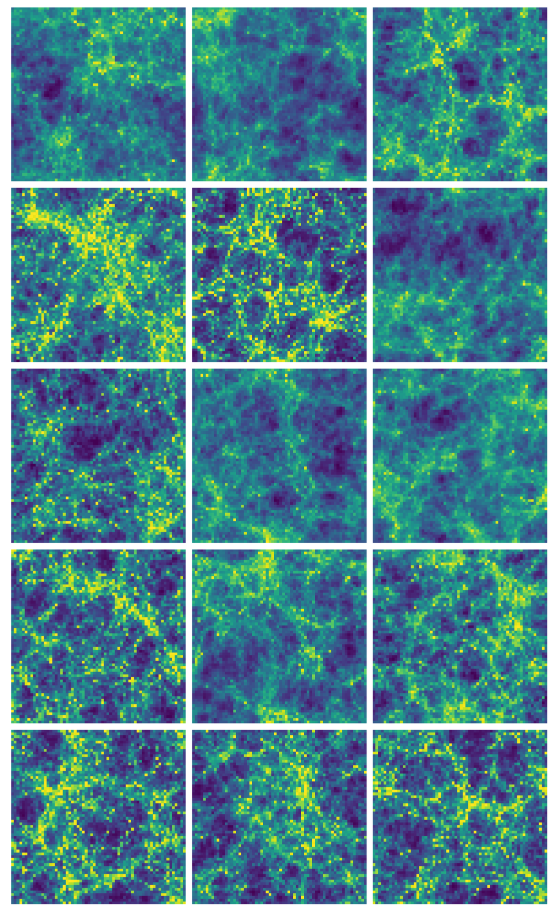

Fig. 1: simulated HI maps using CAMELS

MAIN GOALS Using conditional NFs that take spatial information into account for parameter inference and generation of new samples

HIGlow: a (Conditional) Glow

Normalizing Flows: $p_y(y) = \mathcal{N}(0, I) \quad x = f_\theta^{-1}(y)$

Training:

$$p_x(x) = p_y\left(f_\theta(x)\right) \cdot |\det \frac{\partial f_\theta(x)}{\partial x}|$$

$$\hat{\theta} = \arg\min_\theta \{-\log p_x(x)\}$$

Glow [3] architecture:

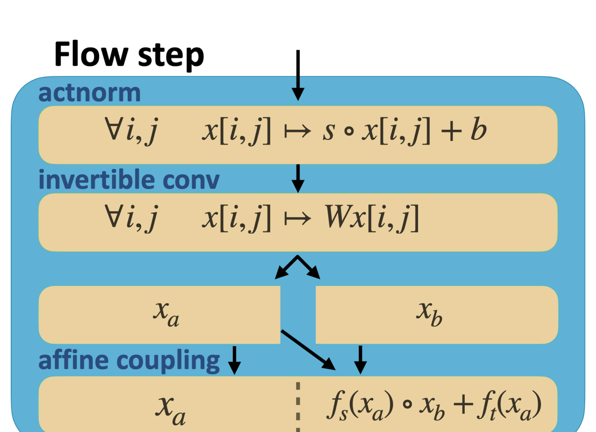

Fig. 2a: a flow step in the Glow architecture, including layer types

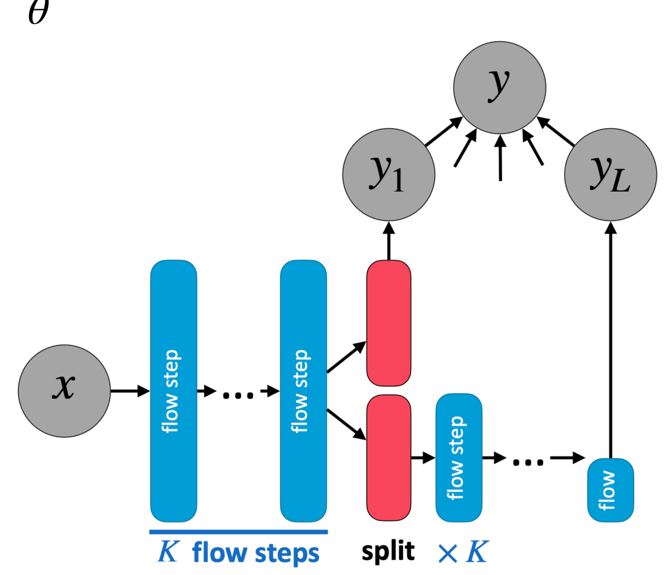

Fig. 2b: full Glow architecture

Conditional flows: $p_x(x \mid z) = p_y\left(f_\theta(x, z)\right) \cdot |\det \frac{\partial f_\theta(x, z)}{\partial x}|$

To make Glow conditional, we can change the layers in the architecture [4]:

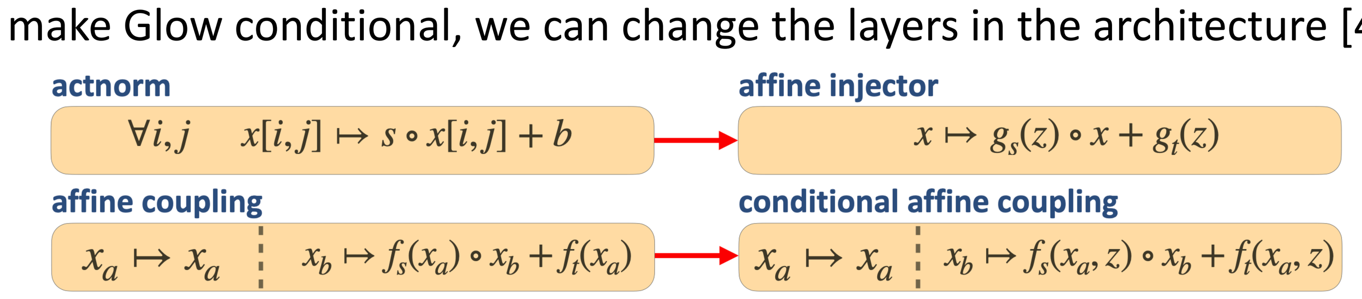

Fig. 3: layer types used in HIGlow and connection to layers in Glow

Generating Conditional HI Maps Using HIGlow

Using the conditional Glow (HIGlow), conditional samples x_z^* can be generated:

$$y^* \sim \mathcal{N}(0, I) \longrightarrow x_z^* = f_\theta^{-1}\left(y^*, z\right)$$

Generating samples is simple, and results in visually similar HI maps:

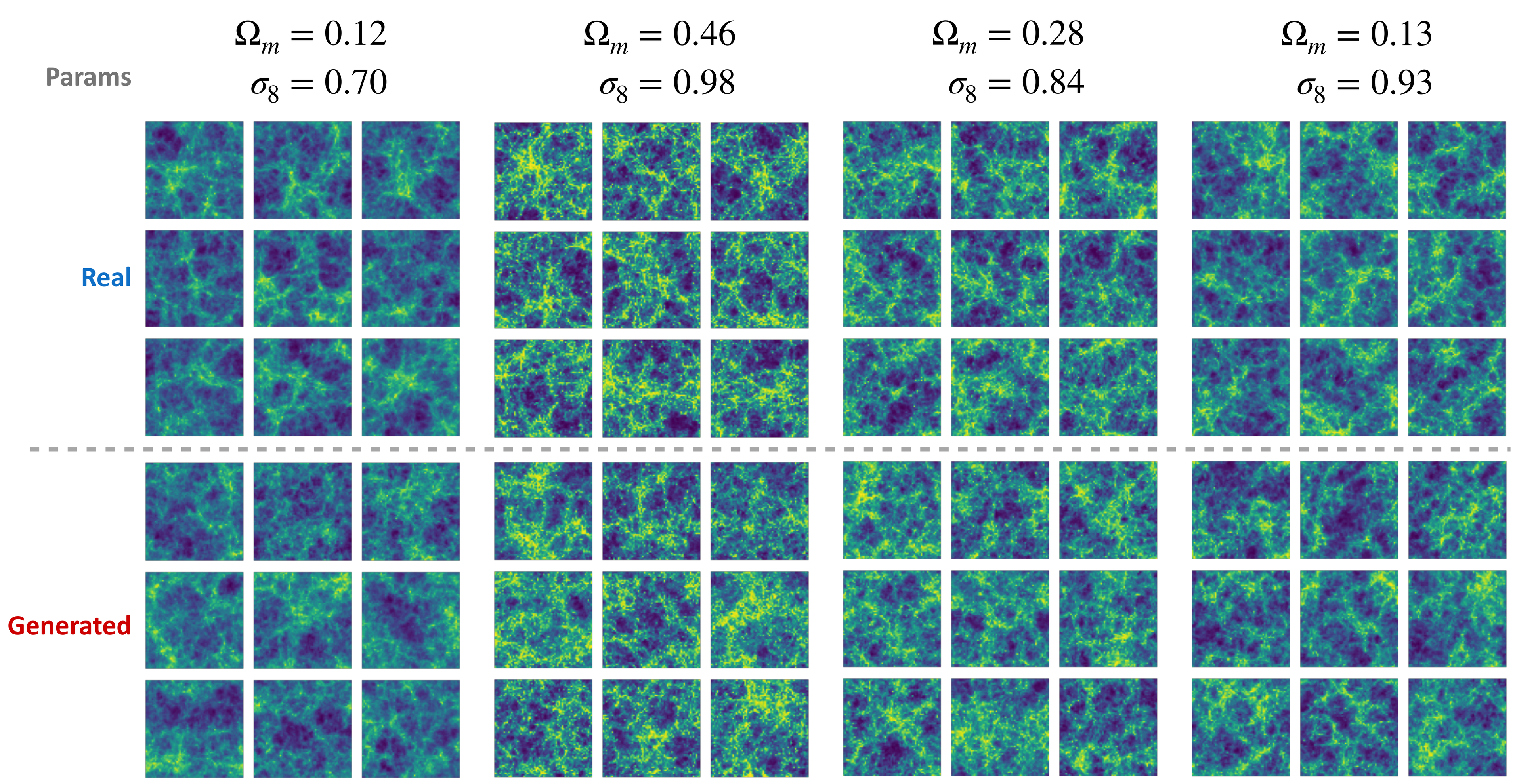

Fig. 4: generated HI maps using HIGlow (bottom) next to CAMELS (top) under various cosmology parameters

Model Evaluation

Ensuring that the generated samples are from the same distribution as the data itself is crucial. To do so, we calculate the mean power spectrum and the standard deviation from this mean for 500 of the CAMELS data as well as images generated by HIGlow and HIFlow:

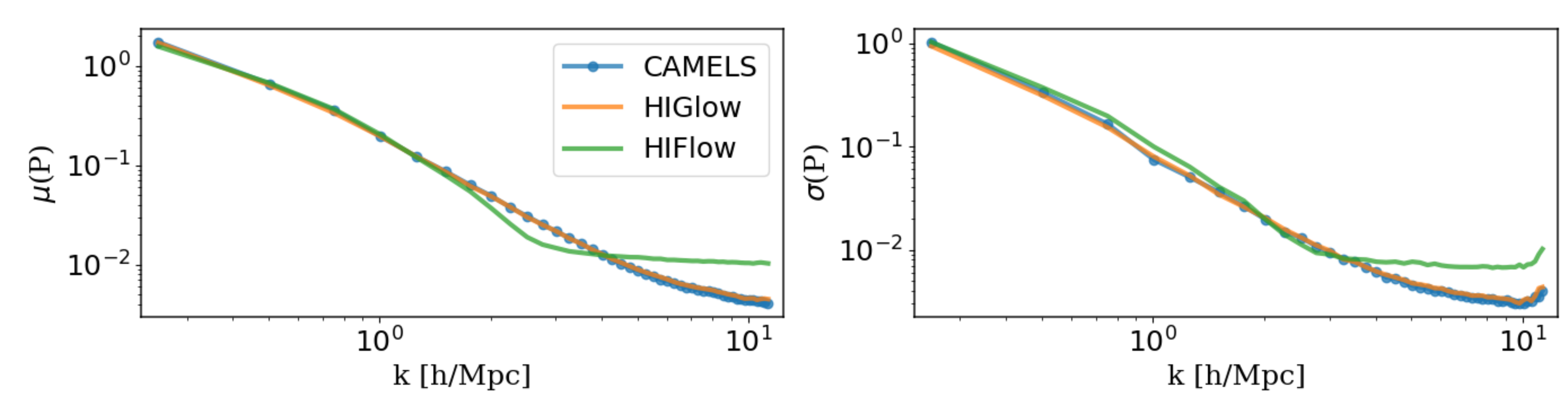

Fig. 5: comparison of the HIGlow marginal $p_x(x)$ to CAMELS data and HIFlow; HIGlow models high frequencies better than HIFlow

The power spectra can vary for different parameter values, and as such the model should generate new samples that also follow the same changes in the power spectrum.

To check the fidelity of conditional samples to the CAMELS data, we compare the mean and standard deviation of the power spectra of CAMELS and HIGlow under various parameter settings.

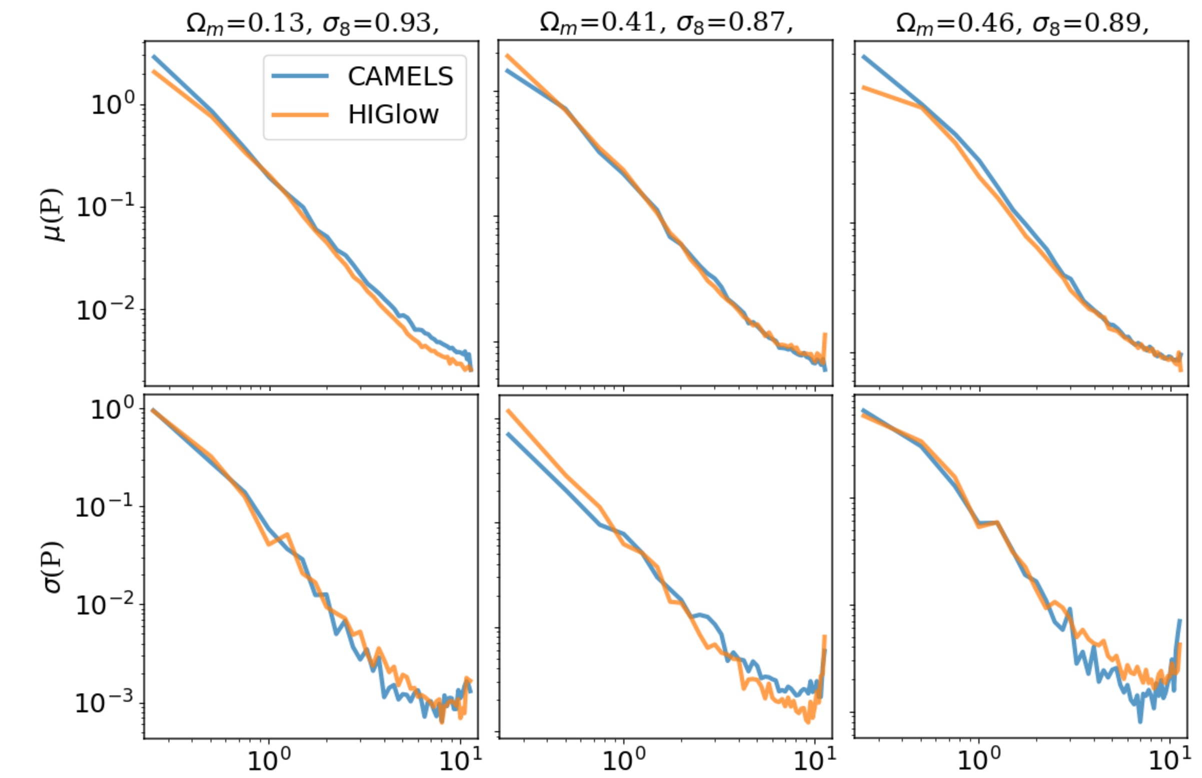

Fig. 6: comparison of the conditional $p_x(x \mid z)$ to CAMELS

Parameter Inference

Calculating the exact likelihood is possible with NFs, which allows estimation of the posterior of the cosmological parameters:

$$p_x(z \mid x) = \frac{p_x(x \mid z)p(z)}{\int p_x(x \mid z)p(z)dz} \approx \frac{p_x(x \mid z)p(z)}{\sum_i p_x(x \mid z_i)p(z_i)}$$

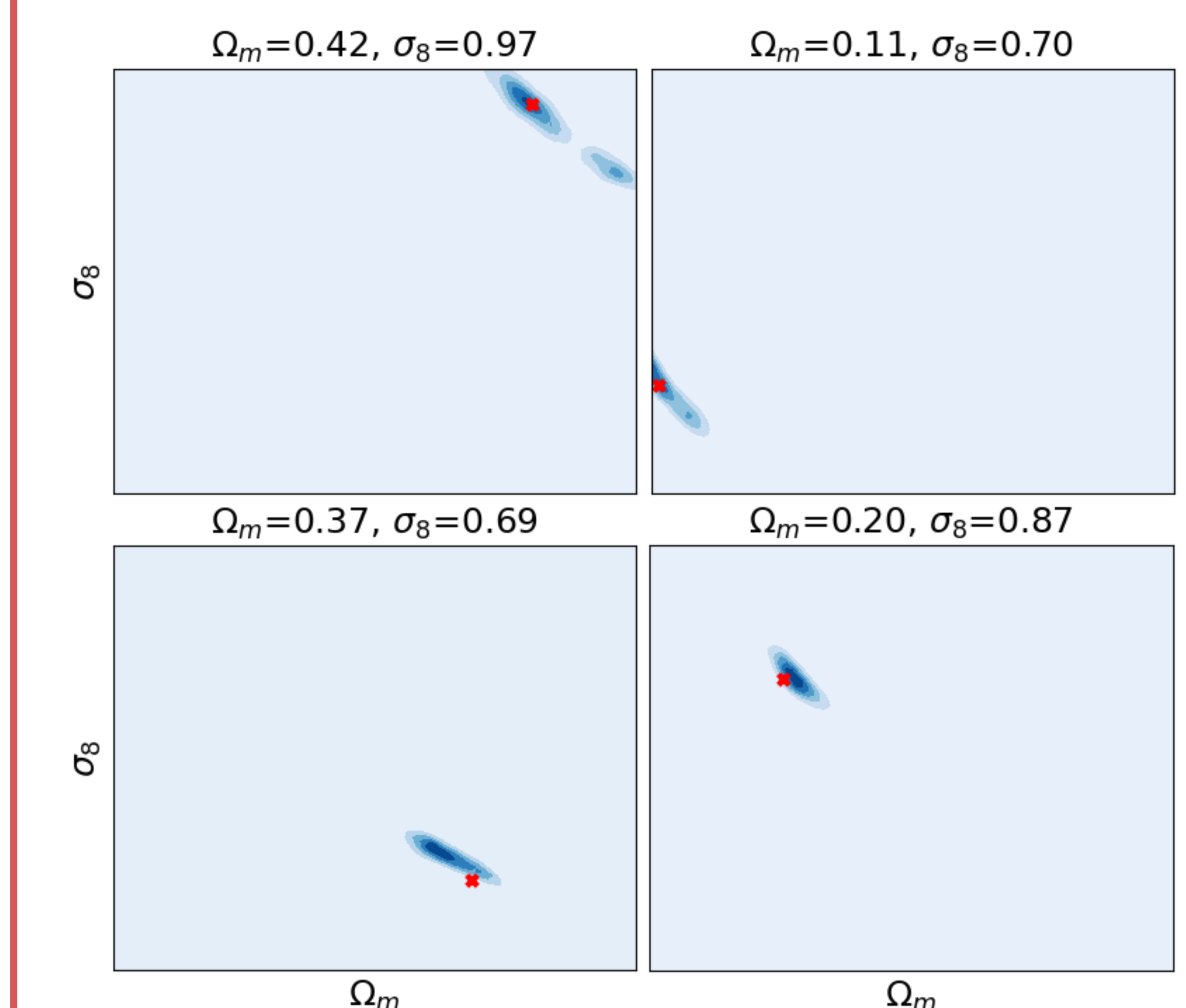

Fig. 7: parameter estimation using HIGlow

Conclusion

In this work, we have trained a conditional Glow on HI maps; a generative model we call HIGlow. HIGlow is capable of generating high-fidelity HI maps and allows for exact conditional likelihood estimation of high-dimensional data points.

HIGlow is able to generate new maps conditional on cosmological parameters. In contrast to the earlier work of HIFlow, our model accurately captures the high-frequency details of the CAMELS simulation data. Additionally, generating samples **conditional on cosmological parameters** gives rise to samples with the same power spectrum distribution as CAMELS.

By utilizing the ability of this model to calculate the exact likelihood of maps conditional on a set of parameters, parameter inference is made possible, which may prove helpful in upcoming large-scale surveys.

References

[1] ***The CAMELS Project***; Villaescusa-Navarro et al., 2021
[2] ***HIFlow***; Hassan et al., 2022
[3] ***Glow***; Kingma and Dhariwal, 2018
[4] ***SRFlow***; Lugmayr et al., 2020