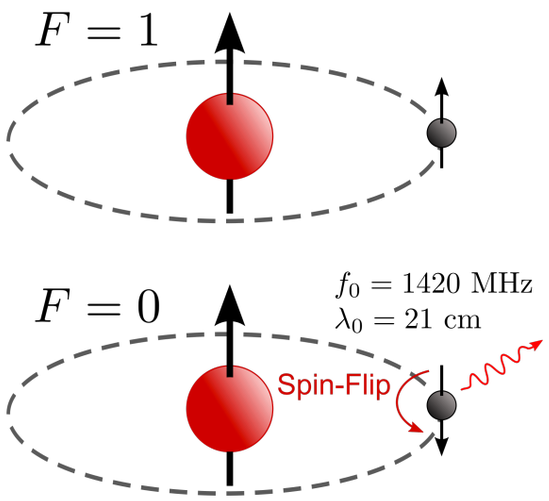
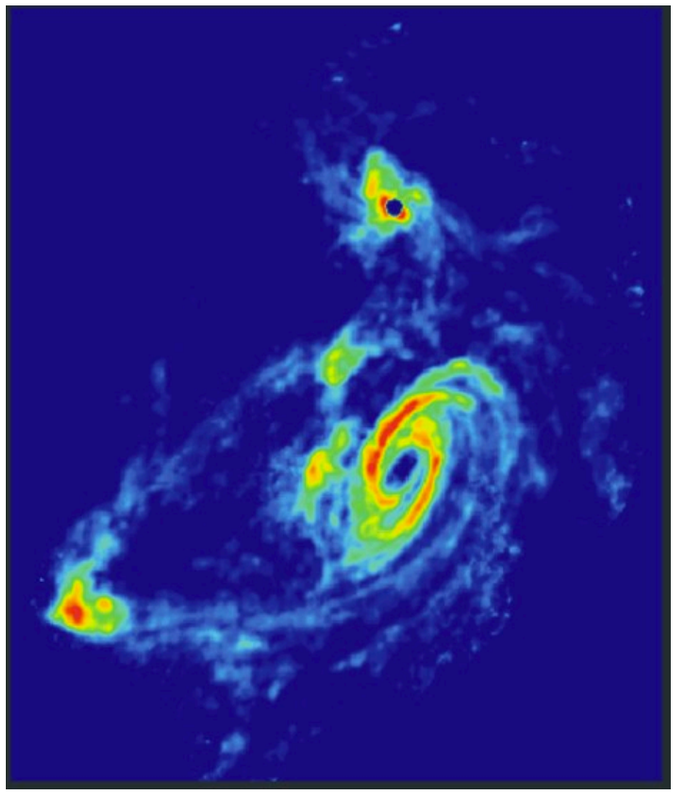
HIGAN: COSMIC NEUTRAL HYDROGEN WITH GENERATIVE ADVERSARIAL NETWORKS

https://arxiv.org/abs/1904.12846

Cosmic HI





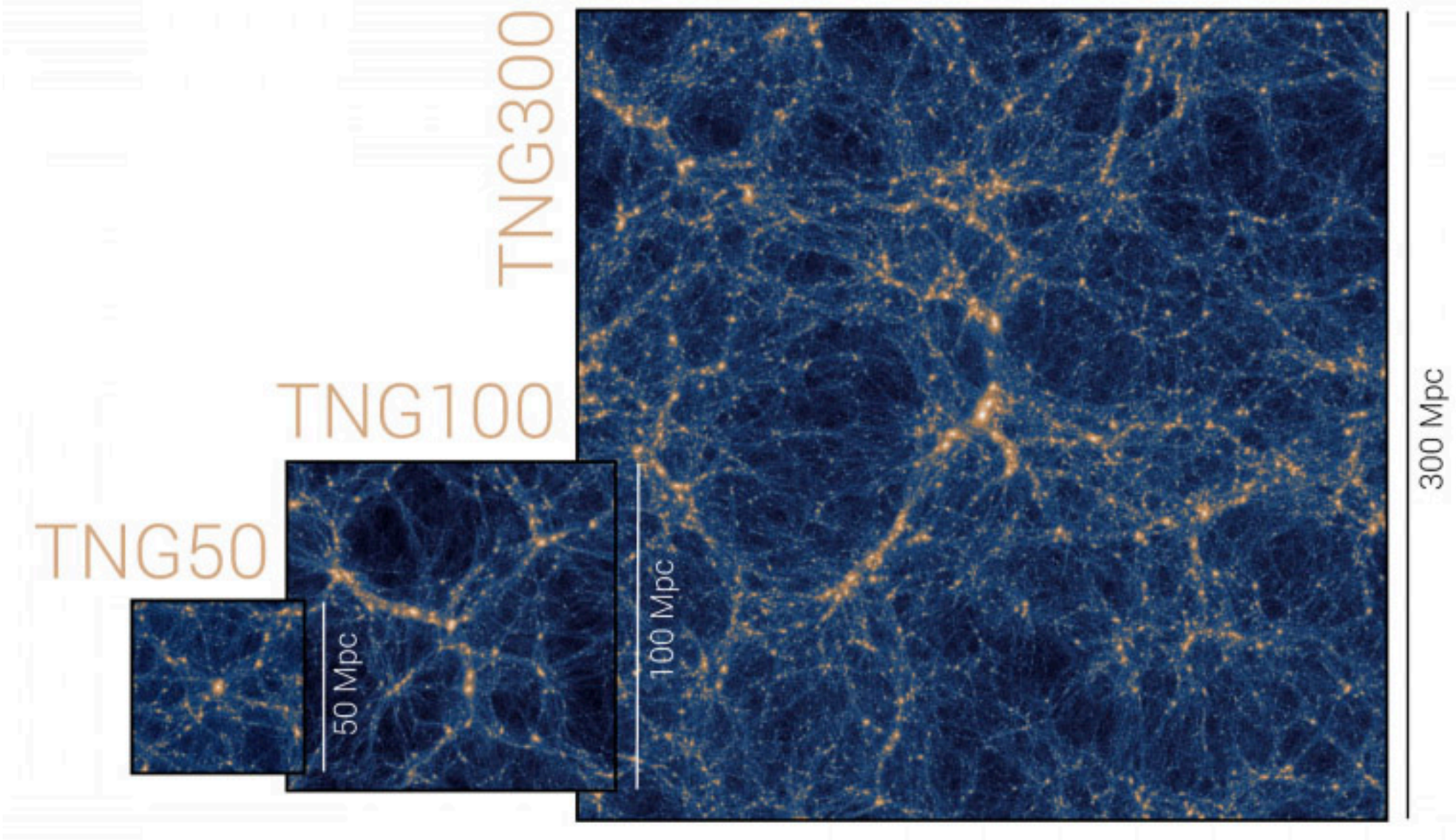
21 cm hydrogen line

Intergalactic medium,
galaxy evolution,

spatial fluctuations

On large scale, linear predictions can be obtained analytically
On small scale, non-linear regime, dominated by complicated physics

Simulation



IllustrisTNG the current state-of-the-art
hydrodynamical simulations
to study galaxy formation and evolution

Simulation

Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY

MNRAS **475,** 648–675 (2018) doi:10.1093/mnras/stx3112
Advance Access publication 2017 December 1

First results from the IllustrisTNG simulations: the stellar mass content of groups and clusters of galaxies

Annalisa Pillepich,[1,2★] Dylan Nelson,[3★] Lars Hernquist,[2] Volker Springel,[4,5★] Rüdiger Pakmor,[4] Paul Torrey,[6†] Rainer Weinberger,[4] Shy Genel,[7,8] Jill P. Naiman,[2] Federico Marinacci[6] and Mark Vogelsberger[6‡]

Run name	L_{box} (cMpc)	N_{GAS} –	N_{DM} –	m_{baryon} (10^6 $M_\odot$)	m_{DM} (10^6 $M_\odot$)	$\epsilon^{z=0}_{DM,stars}$ (kpc)	$\epsilon_{gas,\,min}$ (ckpc)	CPU time (Mh)	N_{cores} –
TNG100(-1)	110.7	1820^3	1820^3	1.4	7.5	0.74	0.19	18.0	10 752
TNG100-2	110.7	910^3	910^3	11.2	59.7	1.48	0.37	0.6	2688
TNG100-3	110.7	455^3	455^3	89.2	477.7	2.95	0.74	$\ll 1$	336
TNG300(-1)	302.6	2500^3	2500^3	11	59	1.48	0.37	34.9	24 000
TNG300-2	302.6	1250^3	1250^3	88	470	2.95	0.74	1.3	6000
TNG300-3	302.6	625^3	625^3	703	3764	5.90	1.47	$\ll 1$	768

Computation expensive

First results of IllustrisTNG is 18 M hours,

Is there a way to create new samples much faster, but keep the same statistical properties

GANs

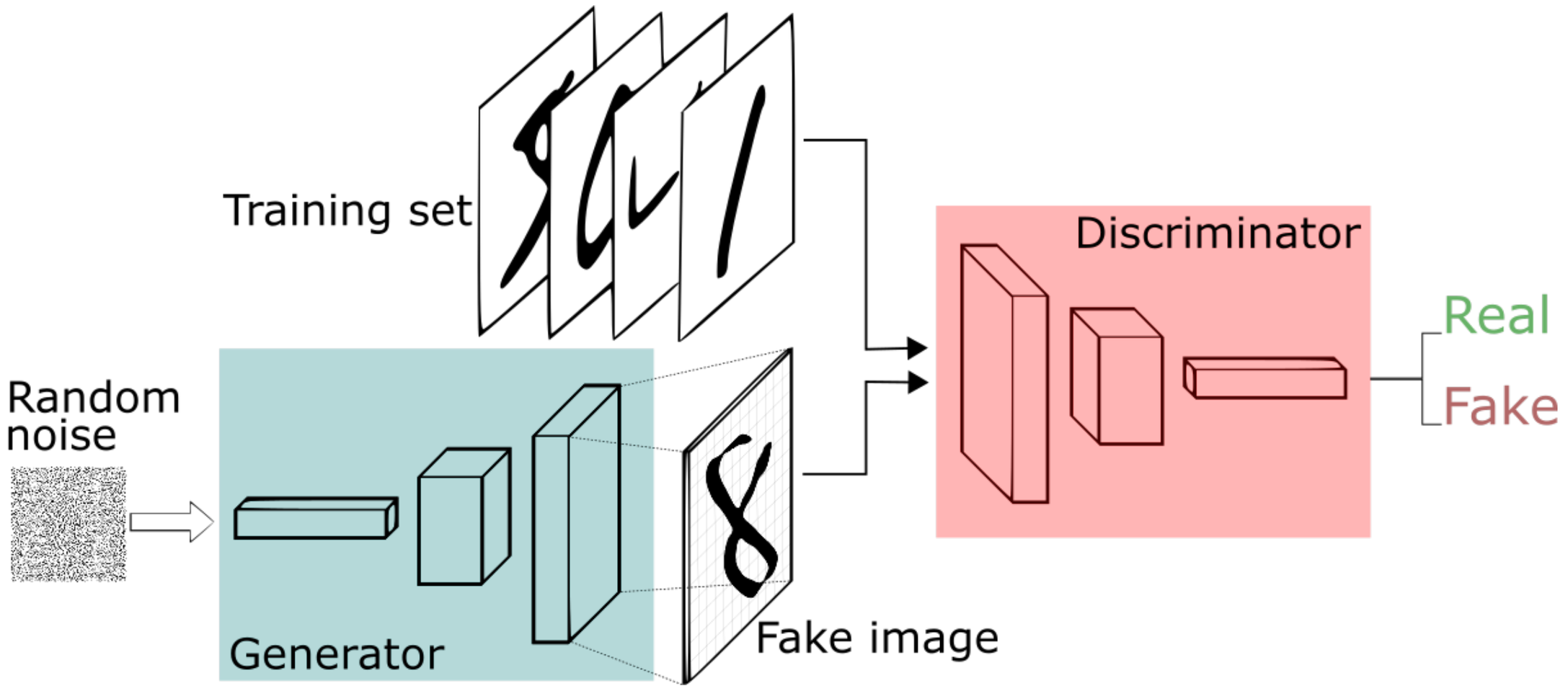
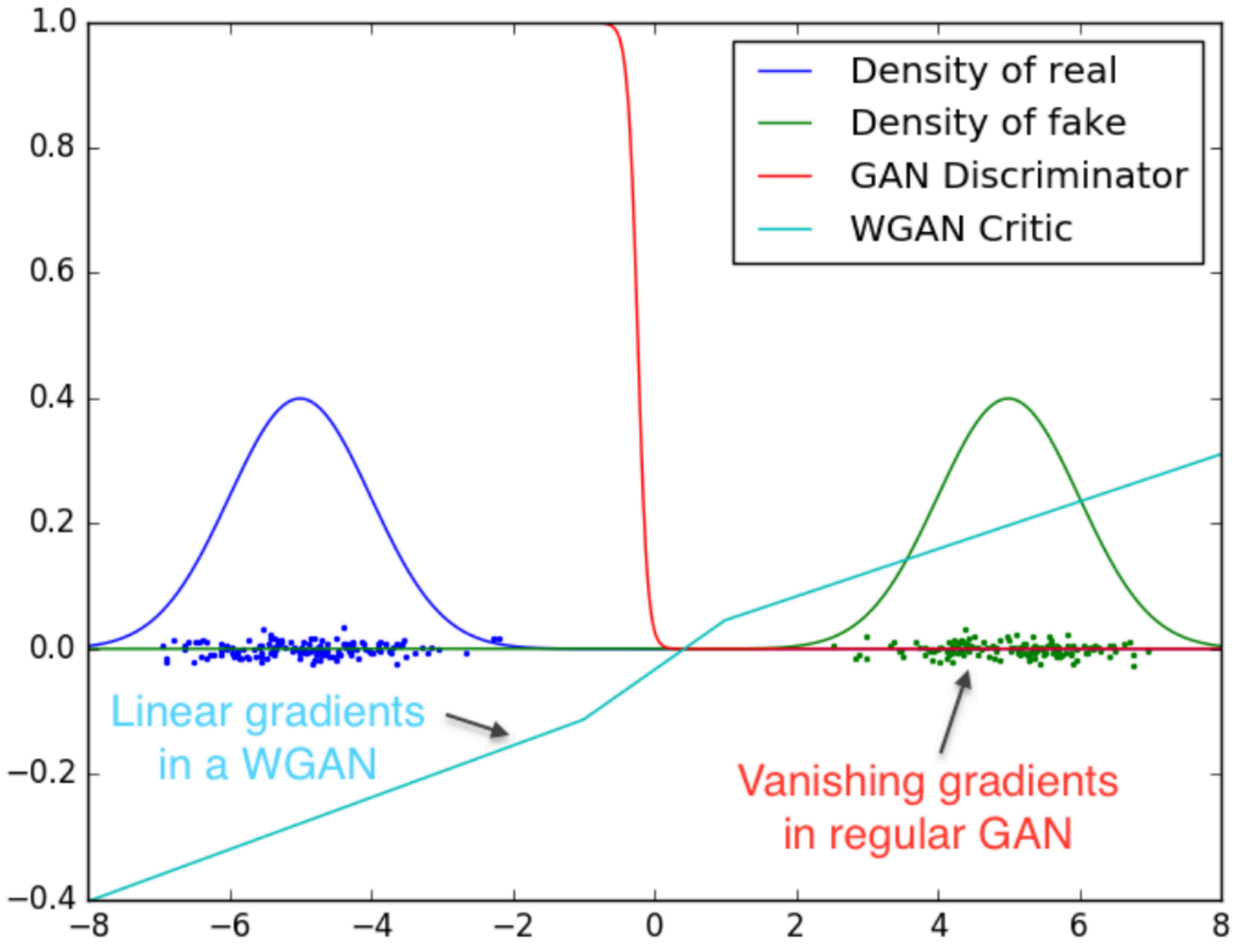


Image credit: Thalles Siva

Hard to train

WGAN



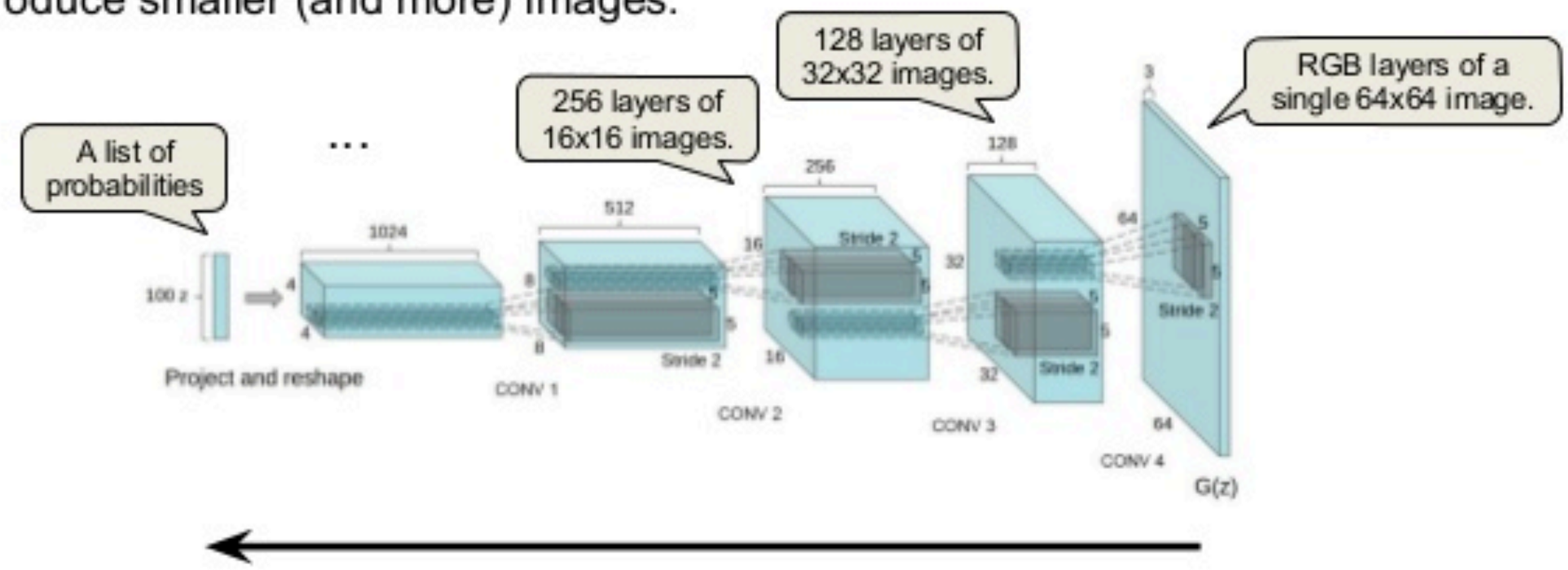
Arjovsky et al. 2017

Instead of differentiate between real and generated samples, provides an approximation of how far the generated samples are from real ones.

Architecture

Typical CNN Filtering Layers

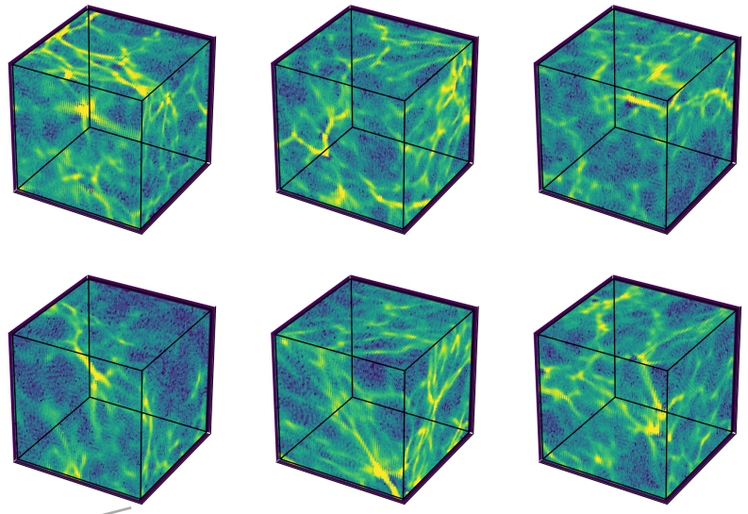
- Starting from a single RGB image on the right, multiple filtering layers are applied to produce smaller (and more) images.



http://arxiv.org/abs/1511.06434

Data

Cubes from IllustrisTNG

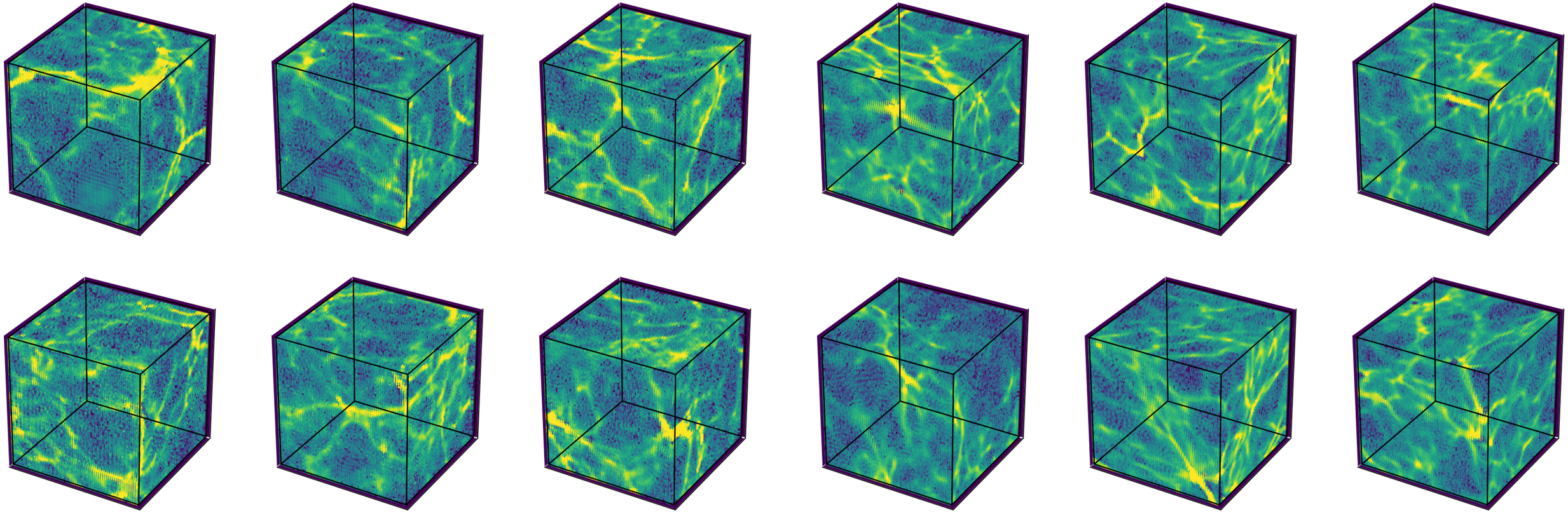


64x64x64

2048x2048x2048

7/8 as training set
1/8 as test set

3D HI distribution



WGAN

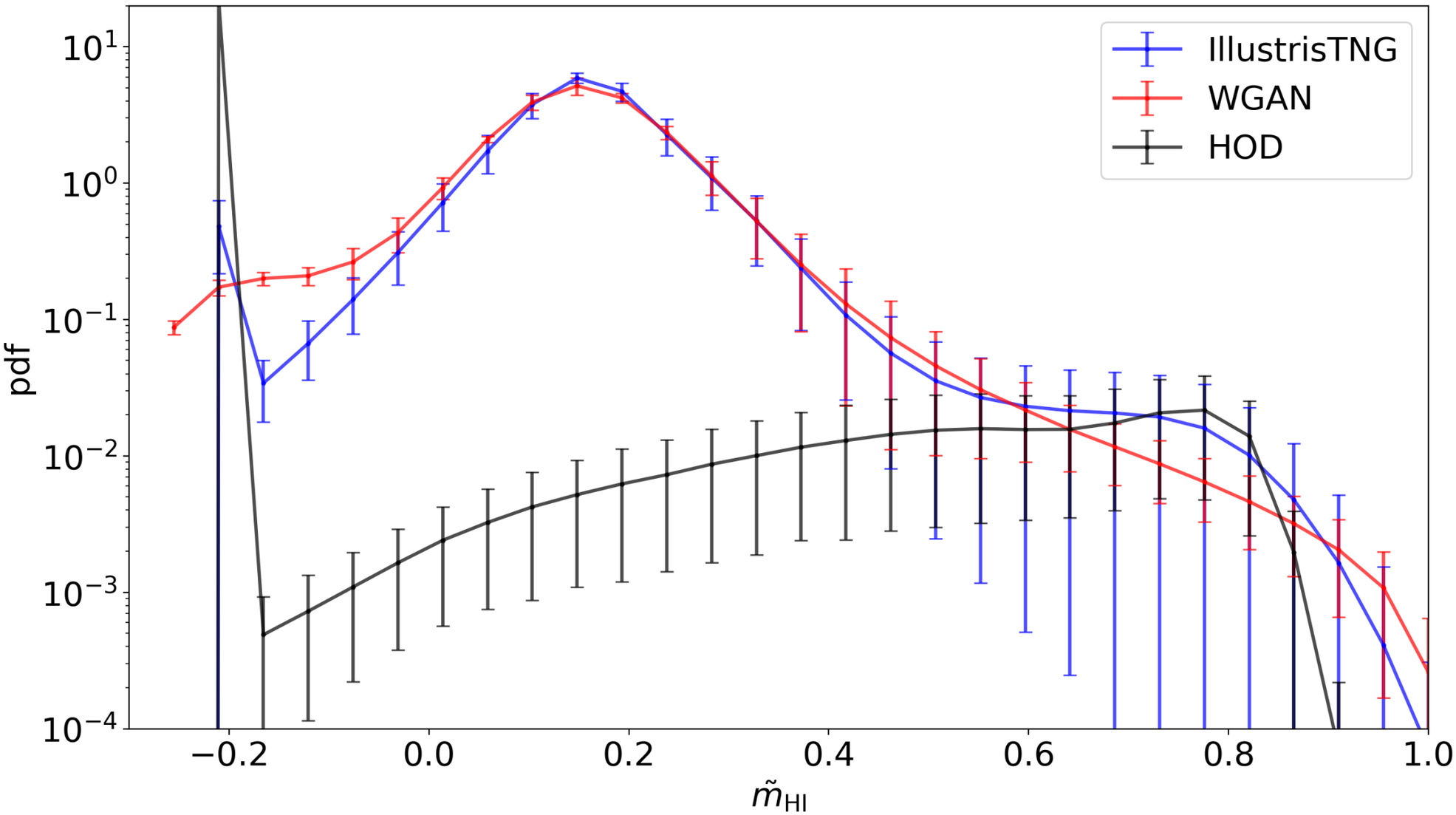
IllustrisTNG

WGAN successfully produce with all the elements of the HI web: filaments, voids and dense regions.

Validation Metrics

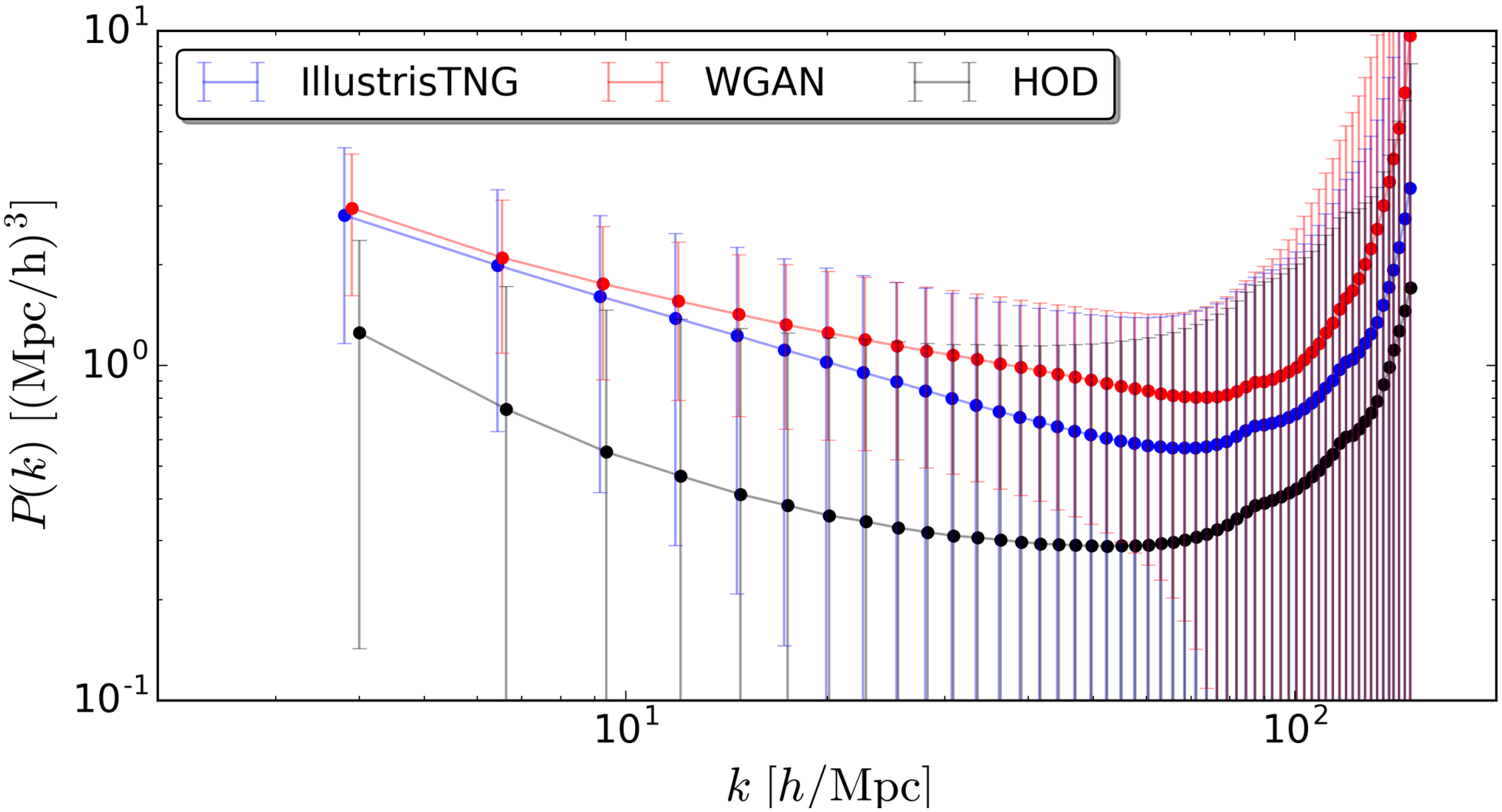
- Gravity-only model: Halo Occupation Distribution
- 1D Probability mass
- Power Spectrum
- Bispectrum
- Voids Abundance

1D Probability mass function



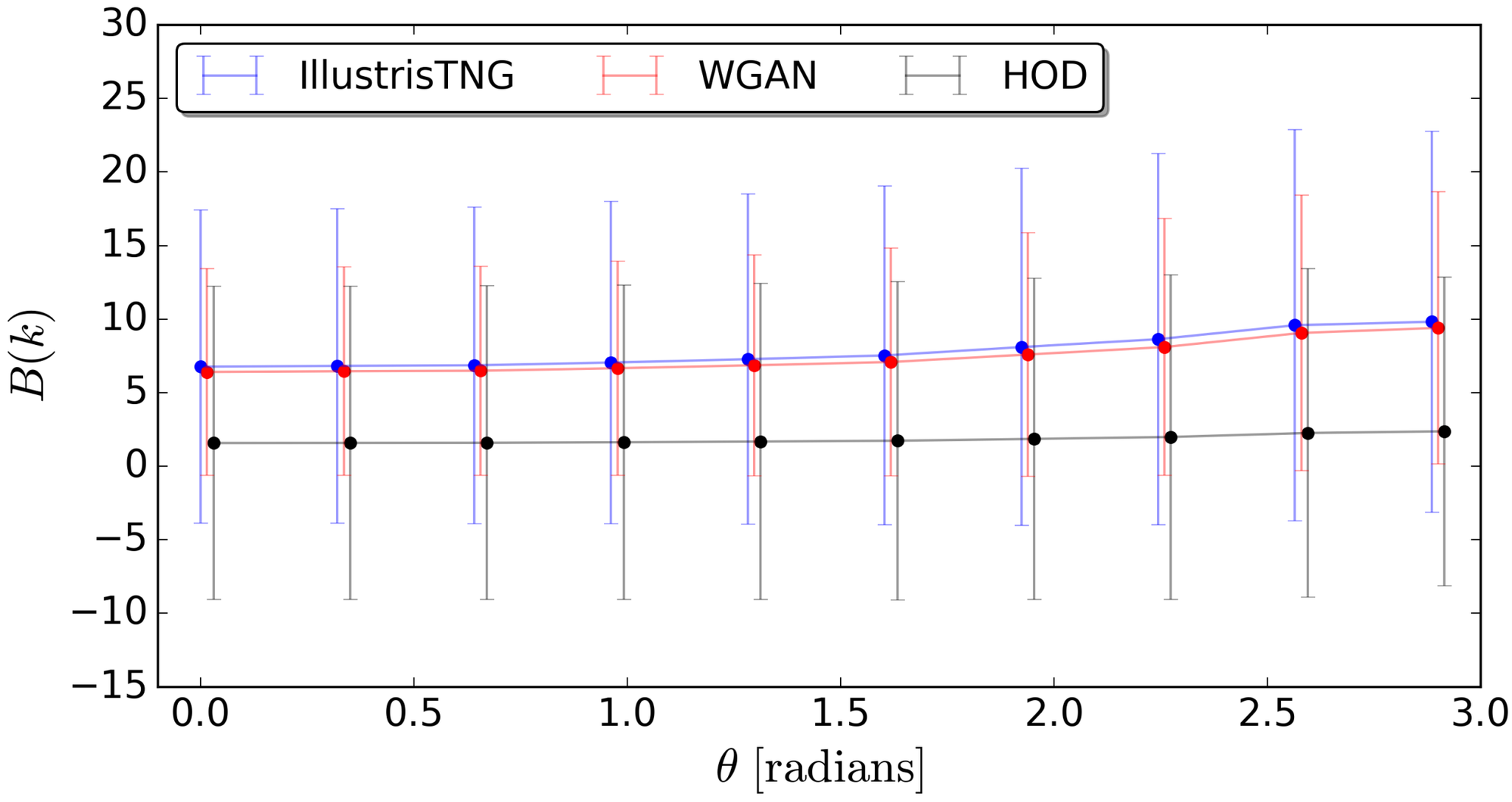
The WGAN model reproduce well the HI PDF from IllustrisTNG,
While the HOD fails to produce low HI density cells.

Power Spectrum



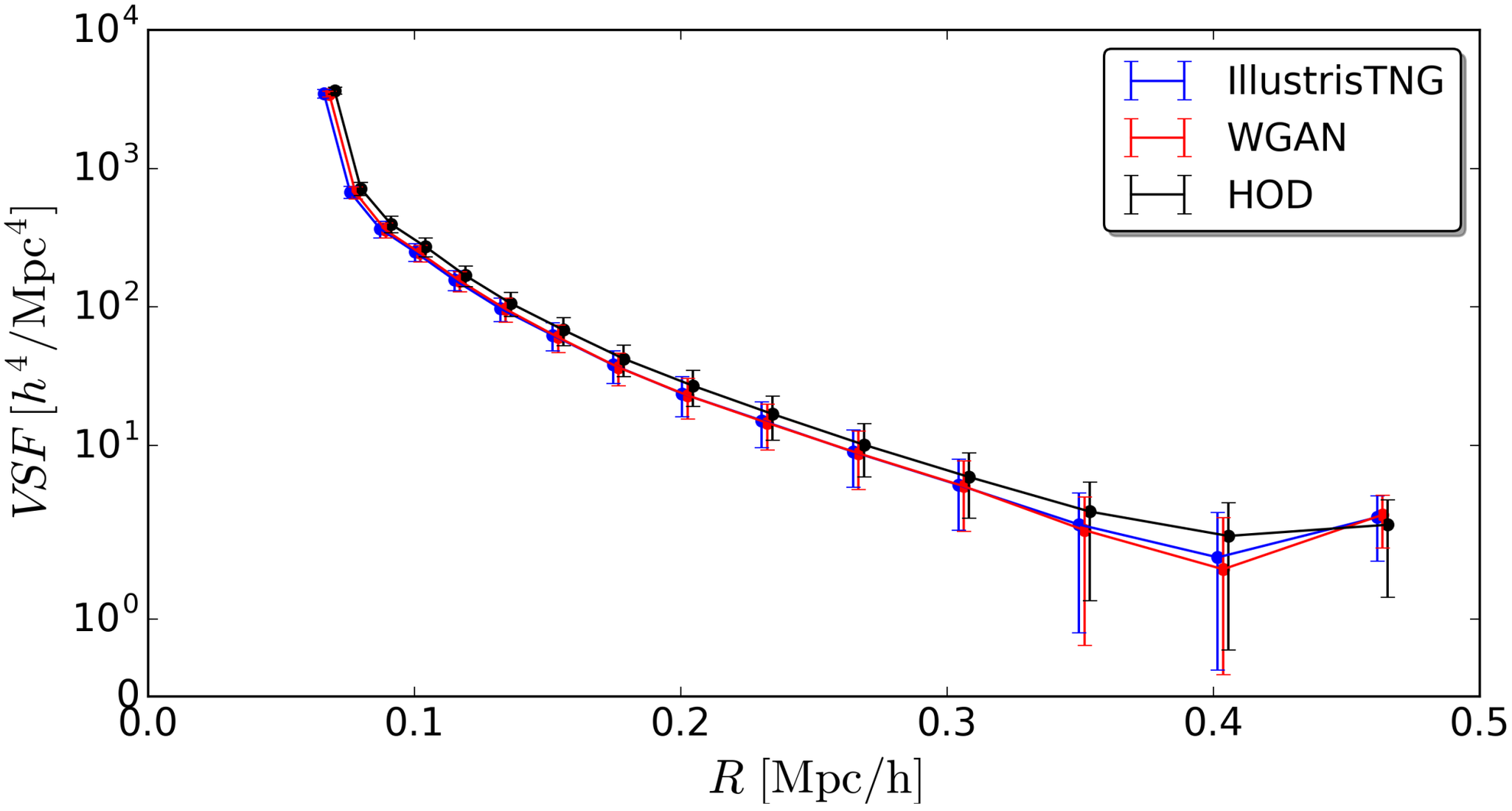
The WGAN samples have a similar clustering amplitude and shape as those from IllustrisTNG. The HOD model is not able to reproduce the power spectrum because the scales here are highly non-linear.

Bispectrum



The bispectra of the WGAN samples are closer to those of IllustrisTNG than to those of the HOD.

Voids Abundance



The number density of voids per unit of radius as a function of void radius. Both the WGAN and HOD model produce fields with the same abundance of voids.

Summary

- Investigate GANs ability to create HI web;
- WGAN model successfully produces spatial distribution of HI with all elements;
- Statistical properties match well with IllustrisTNG

Discuss

- Which features are learned by CNN layers
- Conditional input
- Apply to different redshifts?