dm2gal: Mapping Dark Matter to Galaxies with Neural Networks

Noah Kasmanoff[1], Jeremy Tinker[1], Francisco Villaescusa-Navarro[2], Shirley Ho[3]

[1]New York University, [2]Princeton University, [3]Flatiron Institute

Motivation

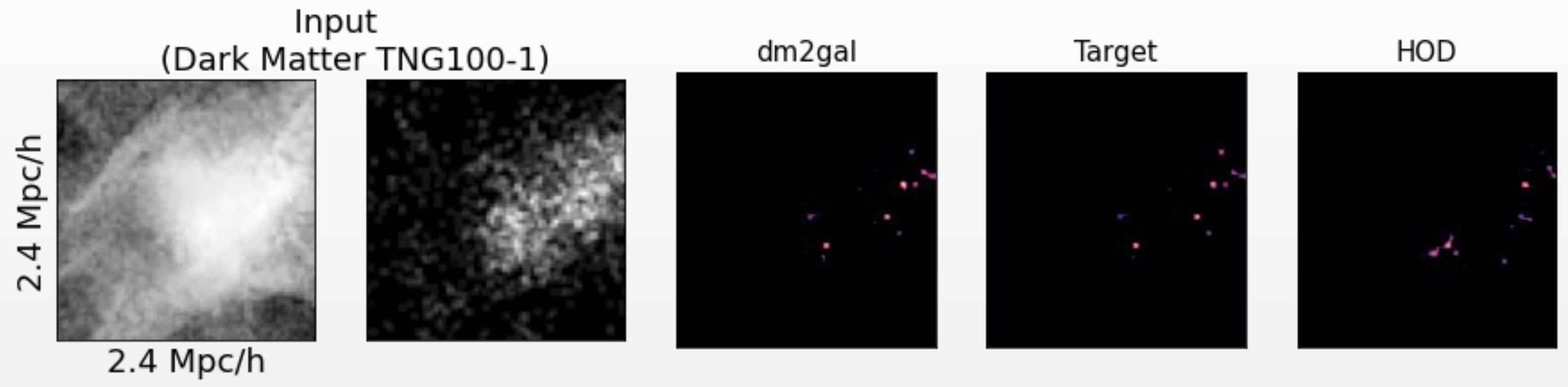
- Galaxies are not randomly distributed in the sky, but follow a particular pattern known as the *cosmic web*.
- Cosmological hydrodynamic simulations are probably the best way to simulate this distribution; however, due to their large computational cost (millions of CPU hours), they only allow predictions of very small volumes.
- Our purpose is to show that neural networks can learn to *paint* galaxy properties on top of the computationally cheaper gravity-only simulations.
  - We compare our findings with benchmark model known as halo occupation distribution (HOD) [1].

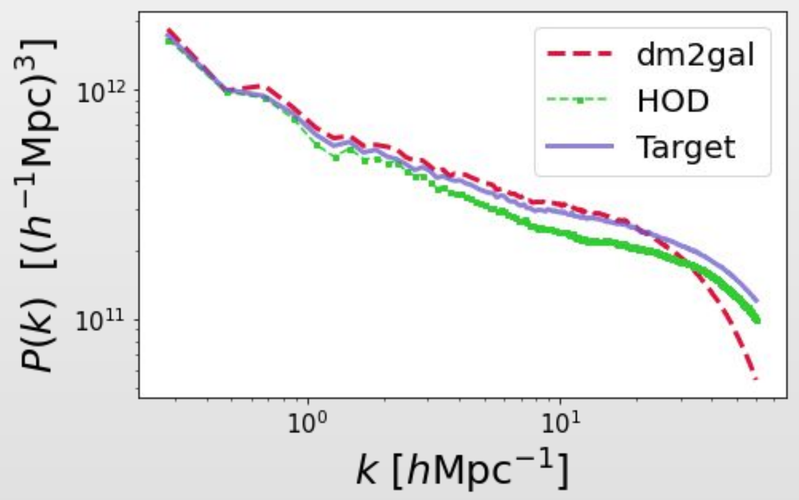
Methods

Data. We use data from the state-of-the-art magneto-hydrodynamic simulation TNG100-1 and its gravity-only counterpart, TNG100-1-Dark, at present time [2]. We select as input-target pairs, regions of size (2.42 h^{-1} Mpc)^3 centered on subhalos with stellar mass ≥ 10^8.

Model. a two-channel 3D volume with 65^3 voxels each: the dark matter and subhalos fields. Our architecture consists of a series of blocks composed of convolutional, batch normalization, and ReLU activation layers while efficiently down-sampling to a single value. The latent representation is flattened into a vector that is passed through two fully connected layers which produces a predicted stellar mass.

Results

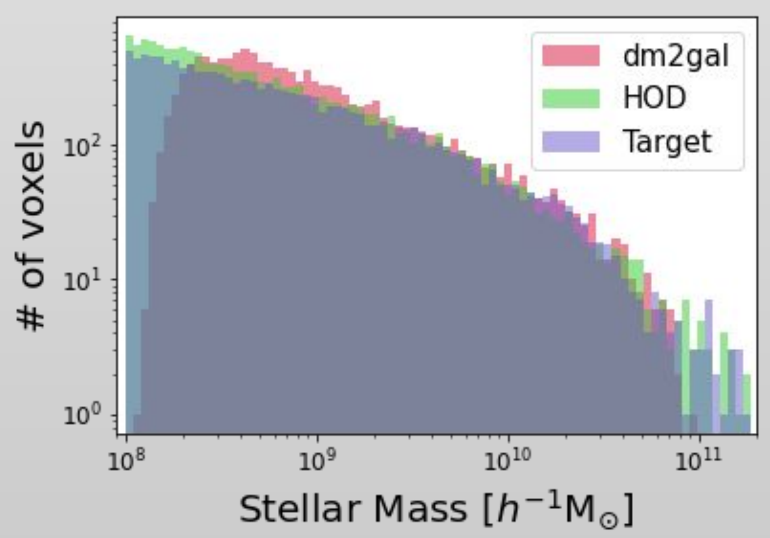




The power spectrum is one of the most important quantities in cosmology, as it describes the cosmic web on large, linear, scales.

We find strong agreement on large scales (low values of k) between all fields for the power spectrum. This is expected for the HOD, but is a prediction for dm2gal.

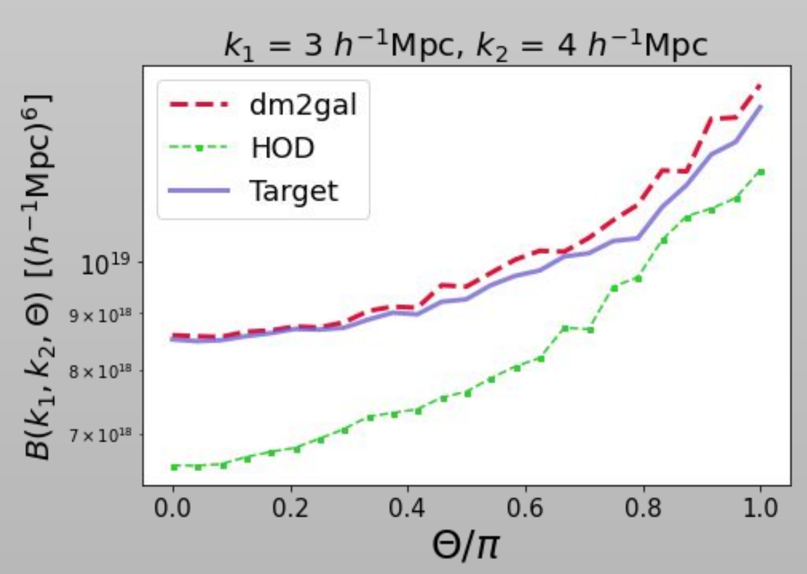
On smaller scales, dm2gal outperforms the HOD, with the exception of scales k ≥ 30



We consider the raw probability distribution function [3], that we compute as the number of voxels with a certain stellar mass, as a function of the stellar mass value.

We find that for stellar masses > $10^{8.5}$, both dm2gal and the HOD outputs a distribution similar to the one from the hydrodynamic simulation.

At the low mass end of the stellar mass PDF, the HOD outperforms dm2gal. We note that it is expected for the HOD model to work very well for the PDF, as it is built to reproduce this statistic.



The bispectrum is a higher-order statistic that contains non-Gaussian information from density fields [4].

We have taken a configuration with k_1=3 and k_2=4 and show the results of the bispectrum, as a function of Θ.

In this case, we find that dm2gal outperforms the HOD for all angles.

We have repeated the exercise for other triangle configurations, finding similar results.

Conclusion

- Convolutional neural networks can be used to *paint* stellar mass into the dark matter field of computationally cheap gravity-only simulations.
  - Generating these fields using hydrodynamic simulations will have a computational cost between 10 to 100x higher than our method.
- Our model outperforms the traditional HOD method for cosmologically relevant scales.
- With additional training and tuning of the hyperparameters, further agreement between ground truth and dm2gal can be reached.

Future Work

- Generalize our network to models with different cosmologies and astrophysics.
- Quantify mapping dependence on time using merger trees.
- Predict other galactic properties such as metallicity, luminosity, and radius.

Acknowledgements

We thank Jacky Yip, Carlos Fernandez-Granda, Gabriella Contardo, Yin Li, and Sigurd Naess for insightful discussions. This work was conducted using the computational resources at New York University and Princeton University

References

[1]: Andreas A. Berlind and David H. Weinberg. The Halo Occupation Distribution: Toward an Empirical Determination of the Relation between Galaxies and Mass.
[2]: https://www.tng-project.org/
[3]:Cora Uhlemann, Oliver Friedrich, Francisco Villaescusa-Navarro, Arka Banerjee, and Sand rine Codis. Fisher for complements: extracting cosmology and neutrino mass from the counts-in-cells PDF.
[4]: ChangHoon Hahn, Francisco Villaescusa-Navarro, Emanuele Castorina, and Roman Scoccimarro. Constraining M_ν with the bispectrum. Part I. Breaking parameter degeneracies.