Calculating Gravitational Wave Bias With TNG simulation

Taisiia Karasova[1], Nikshay Chugh[2], Supervisors: Dorsa Sadat Hosseini[3,4], Ghazal Geshnizjani[3]

[1]Massachusetts Institute of Technology, [2]Indian Institute of Science, [3] Perimeter Institute for Theoretical Physics, [4] University of Waterloo

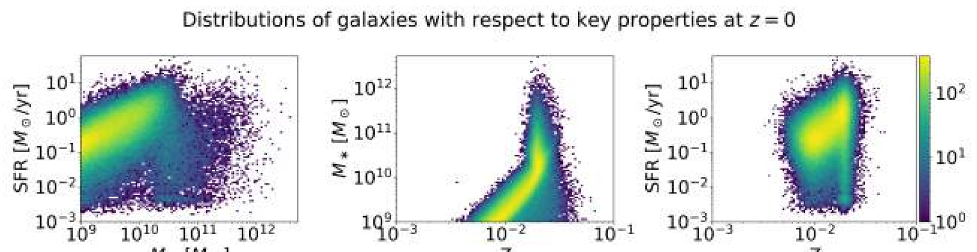
karasova@mit.edu, nikshaychugh@iisc.ac.in

Motivation

The growing number of gravitational wave (GW) detections from binary black hole (BBH) mergers offers a unique opportunity for cosmological inference, including measuring the Hubble parameter through standard sirens. Statistical distributions of GW events differ from those of galaxies due to their relative sparsity. This distinction necessitates careful modeling to optimize cosmological estimators. We use cosmological dark matter simulations to generate mock catalogs of GW events and their associated host galaxies.

Data Preprocess and Validation

Galaxies. We extracted galaxy (subhalo) properties from the TNG300 simulation for each snapshot from $z = 0$ to $z = 1$, selecting only cosmological subhalos. For each, we retain key quantities including stellar mass, star formation rate (SFR), stellar metallicity, and position. Subhalos with stellar mass $> 10^9 M_\odot$ are selected for analysis.



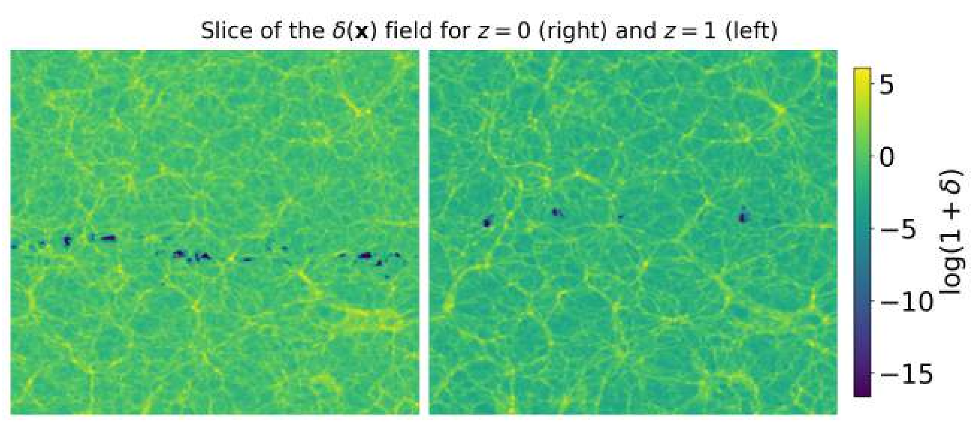
We validated our dataset against literature [1]-[4] (e.g., specific SFR and stellar metallicity vs stellar mass) and found good agreement.

Dark Matter Particles. We extracted the spatial positions of dark matter particles within the periodic simulation domain for a set of 10 snapshots spanning range of redshifts from $z = 0$ to $z = 1$.

Theory and Computation

Density Field. We first compute the matter density field $\rho(\mathbf{x})$ for dark matter particles, galaxies, and GW sirens. We use the Cloud-In-Cell (CIC) mass assignment scheme implemented in `Pylians`, which interpolates each particle's mass to its eight nearest 3D grid points based on relative position within the cell. We define the overdensity field as

$$\delta(\mathbf{x}) = \rho(\mathbf{x})/\bar{\rho} - 1 \tag{1}$$



Power Spectrum. The power spectrum quantifies the statistical distribution of density fluctuations as a function of spatial scale. Defined in Fourier space, it is the variance of the overdensity field:

$$P(k) = \langle|\tilde{\delta}(\mathbf{k})|^2\rangle \tag{2}$$

where $\tilde{\delta}(\mathbf{k})$ is the Fourier transform of $\delta(\mathbf{x})$, and k is the wavenumber corresponding to scale $\lambda = 2\pi/k$. We compute the power spectra $P(k)$ for dark matter, galaxies, and GW sirens using `Pylians`.

Gravitational Wave / Galaxy Bias. To characterize how well galaxies and GW sirens trace the underlying matter distribution, we compute their scale-dependent bias:

$$b(k) = \sqrt{P_{\rm obj}(k)/P_{\rm dm}(k)} \tag{3}$$

where $P_{\rm obj}(k)$ is the power spectrum of galaxies or GW sirens, and $P_{\rm dm}(k)$ is that of dark matter. To extract a constant linear bias, we fit Eq. 3 over the quasi-linear range $0.04 < k < 1.0$ by minimizing the χ^2 with `scipy.optimize.minimize`.

Modelling

Number of Mergers. We simulate the GW siren population by assigning binary mergers to galaxies in simulation snapshots. We estimate the expected number of sirens at redshift z via [6]:

$$N_{GW}(z) = V_{box} \cdot \frac{R_{GW}(z)t_{obs}}{1+z} \tag{4}$$

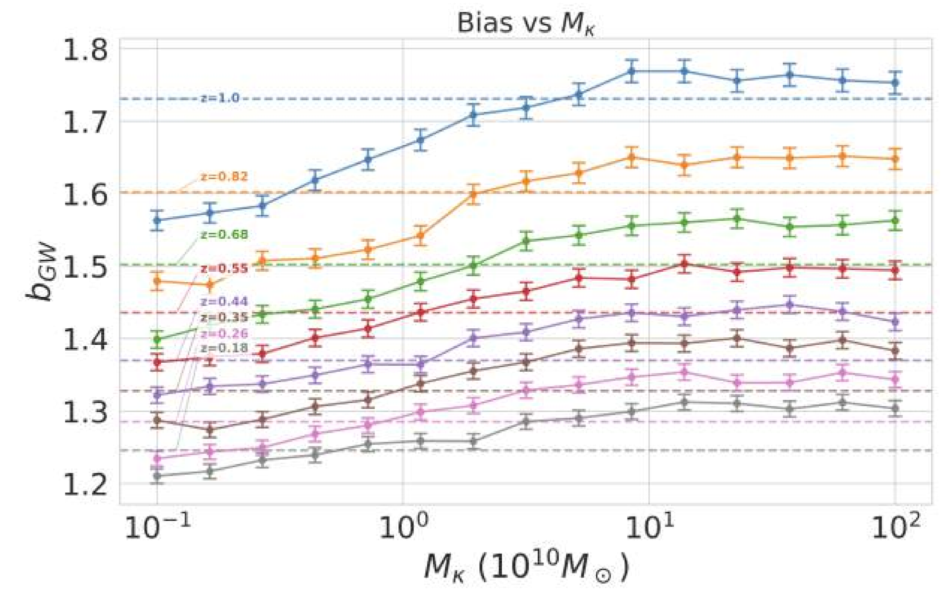
where $V_{box} = 306.2\ \mathrm{Mpc}^3$, $R(z)$ is the merger rate density based on the star formation history and a delay time distribution $P(t_d) \propto t_d^{-1}$, and t_{obs} is the observation time.

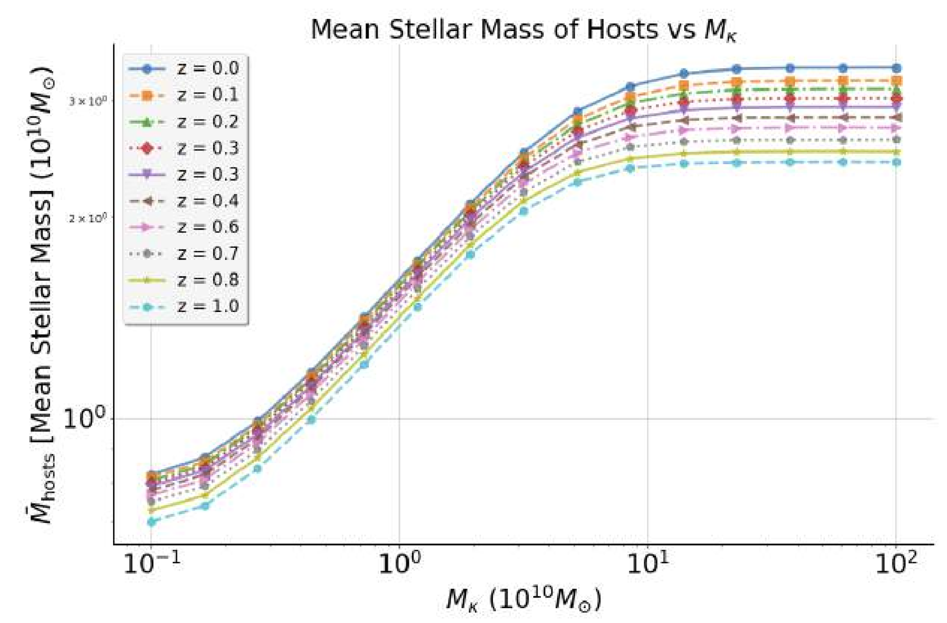
GW Host Probability Function. We select host galaxies using a broken power-law probability distribution based on stellar mass [6]:

$$\begin{cases} M_*^{1/\delta_l} & \text{if } M_* < M_K \\ M_*^{-1/\delta_h} \cdot M_K^{1/\delta_l + 1/\delta_h} & \text{if } M_* \geq M_K \end{cases} \tag{5}$$

Results

Simulation results follow trends similar to those of the observational data. Below, dashed lines indicate the galaxy bias.





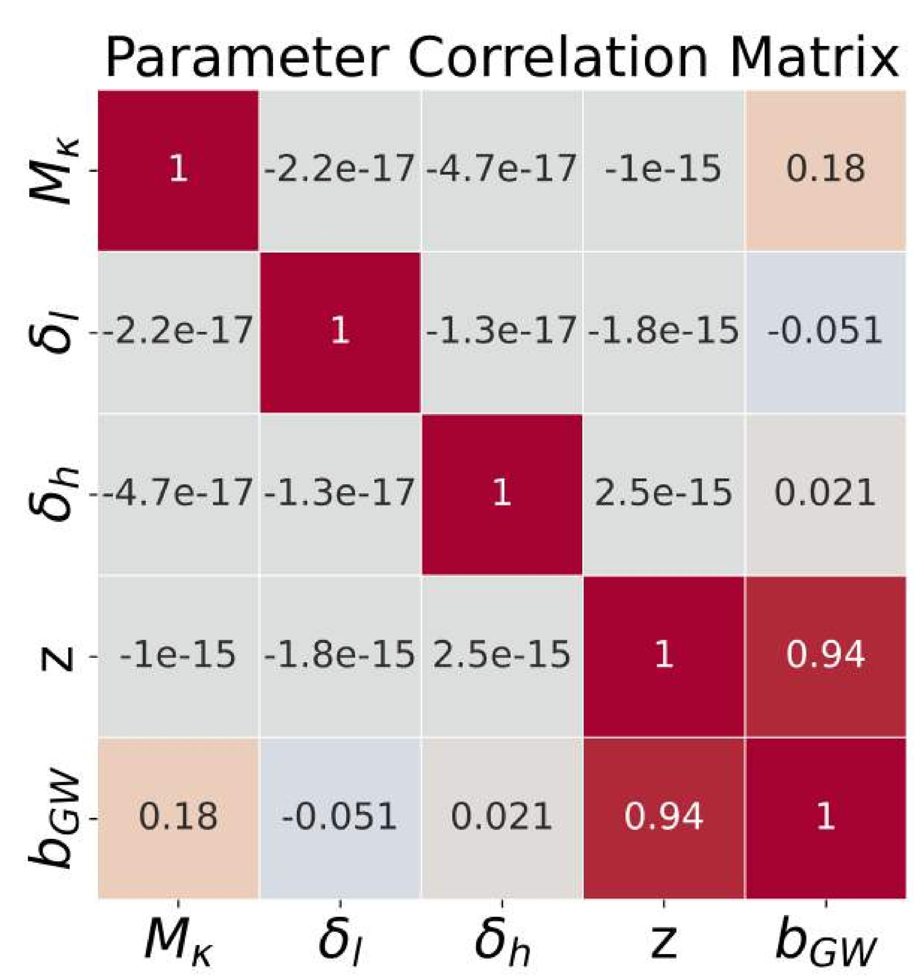
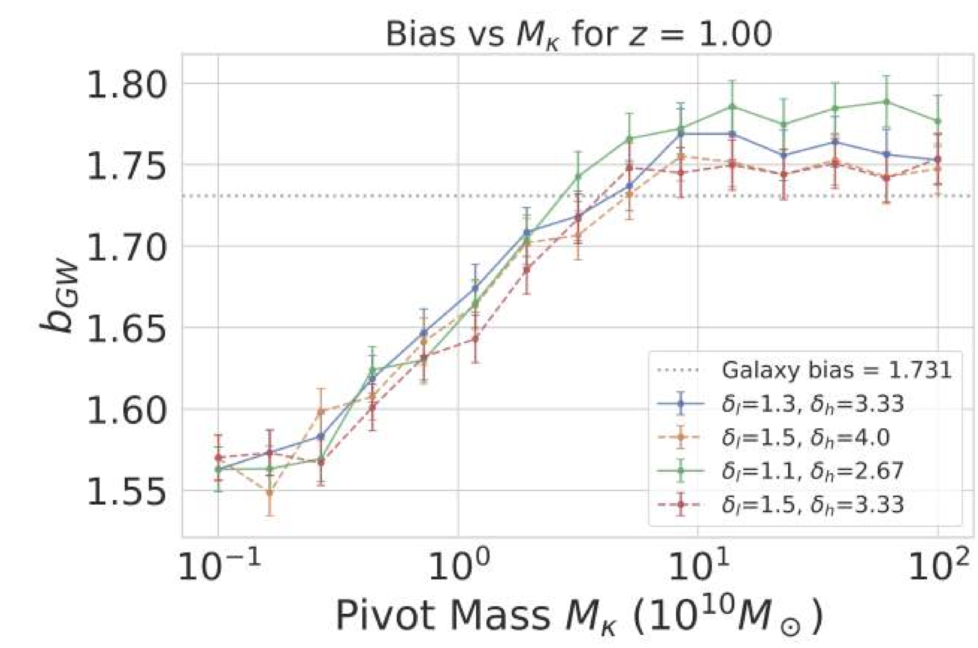
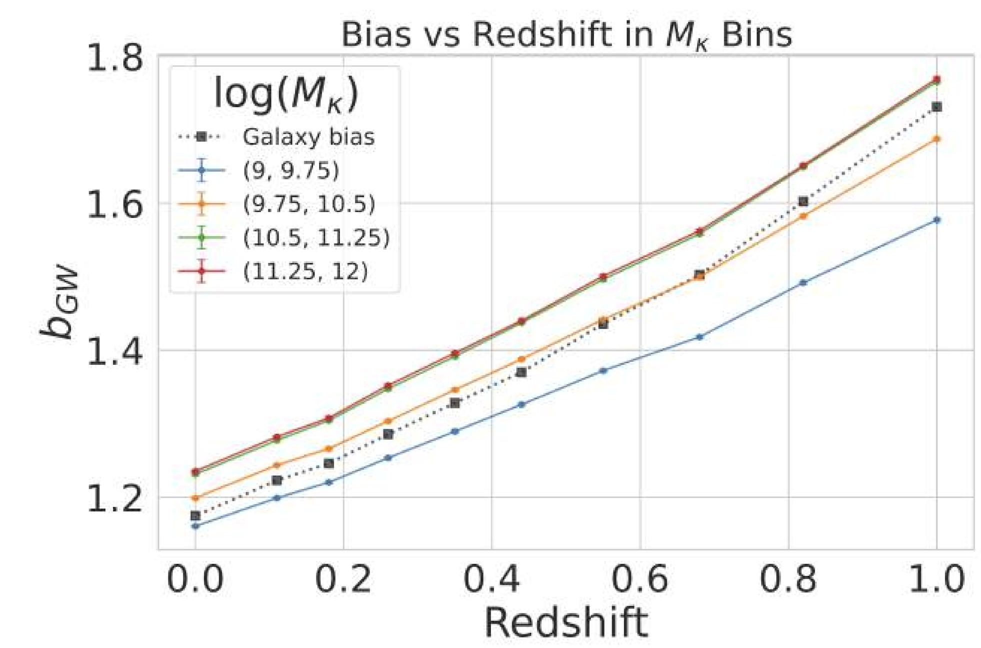


Table: Statistical Properties of b_{GW} w.r.t z

z	$\langle b_{GW}\rangle$	$\sigma^{b_{GW}}$	Med(b_{GW})	$\langle\chi^2_r\rangle$	Med(χ^2_r)
0.00	1.2069	0.0333	1.2153	0.0983	0.0792
0.11	1.2457	0.0369	1.2584	0.0846	0.0727
0.18	1.2679	0.0386	1.2808	0.0814	0.0745
0.26	1.3062	0.0429	1.3199	0.1157	0.1001
0.35	1.3456	0.0460	1.3651	0.1518	0.1421
0.44	1.3864	0.0493	1.4057	0.2014	0.1857
0.55	1.4408	0.0551	1.4609	0.2298	0.2153
0.68	1.4973	0.0605	1.5203	0.2201	0.2047
0.82	1.5816	0.0672	1.6079	0.2803	0.2701
1.00	1.6849	0.0799	1.7187	0.2900	0.2745





Discussion and Conclusion

Interpretation

- b_{GW} increases with redshift in good agreement with the galaxy bias. The power spectrum of the dark matter also matches the CAMB power spectrum to high accuracy indicating correctness of results.
- Inflexion point of the mean stellar mass of host galaxies follows a similar trend as the pivot mass of the broken power law for galaxy count suggesting some probabilistic "resonance" at these values of M_K.
- b_{GW} is (weakly) negatively correlated with δ_l because a higher value of δ_l implies that low mass galaxies are more likely to be hosts. δ_h and b_{GW} follow the opposite trend.
- b_{GW} shows a non-trivially positive correlation with M_K suggesting that increasing the pivot mass leads to more clustering of sirens with weak dependence on the slopes in the power law.
- Binning in mass bins of the pivot mass indicates that there is little difference in the qualitative features of host galaxies in mass range $10^{10.5} - 10^{12} M_\odot$. The reasons are unclear yet.

Limitations

- TNG 300-1 simulation has a small size compared to observational data.
- Results are specific to ΛCDM cosmology assumed in TNG.

References

1. Dylan Nelson et. al. The illustristng simulations: Public data release, 2021.
2. Jubee Sohn et. al. Velocity dispersions of quiescent galaxies in illustristng, 2024.
3. Harry George Chittenden and Rita Tojeiro. Modelling the galaxy–halo connection with semi-recurrent neural networks. Monthly Notices of the Royal Astronomical Society, 518(4):5670–5692, November 2022.
4. Jayashree Behera et. al. Optimised neural network predictions of galaxy formation histories using semi-stochasti corrections, 2024
5. Francisco Villaescusa-Navarro. Pylians: Python libraries for the analysis of numerical simulations. Astrophysics Source Code Library, record ascl:1811.008, November 2018
6. Dorsa Sadat Hosseini et. al. Modeling gravitational wave bias from 3d power spectra of spectroscopic surveys, 2025.

Acknowledgments

We thank our supervisors Dorsa Sadat Hossein and Ghazal Geshnizjani for their guidance throughout the project, PSI Start program for this unique opportunity, Dustin Lang for assistance with Python packages installation, and our fellow interns for making this journey so enjoyable and memorable.