

The effects of assembly bias on galaxy clustering predictions

Arnau Pujol*

*Institut de Ciències de l'Espai (ICE, IEEC/CSIC), E-08193 Bellaterra (Barcelona), Spain
E-mail: pujol@ice.cat*

Enrique Gaztañaga

Institut de Ciències de l'Espai (ICE, IEEC/CSIC), E-08193 Bellaterra (Barcelona), Spain

The Halo Occupation Distribution (HOD) model is frequently used in surveys to predict the mass of the dark matter haloes from the clustering of galaxies. On the other hand, semi-analytical models (SAMs) of galaxy formation are often used to populate simulations according to some physical prescriptions and merger trees. We compare galaxy bias measured in SAMs with the bias reconstructed from the halo bias and HOD measured in the same simulations. We find that the reconstruction underestimates the bias by 5 – 10%, which translates in 50% overestimation of the halo mass. We attribute this failure to assembly bias. The clustering of haloes with $M \lesssim 3 - 5 \times 10^{11} h^{-1} M_{\odot}$ depends strongly on environment (number of subhaloes) and this also reflects in different clustering for galaxies with different luminosity but equal halo mass. Thus we need to include properties other than halo mass to do a proper HOD reconstruction.

*Frontiers of Fundamental Physics 14 - FFP14,
15-18 July 2014
Aix Marseille University (AMU) Saint-Charles Campus, Marseille*

*Speaker.

1. Introduction

The Halo Occupation Distribution (HOD) (e.g. [1]) formalism uses the Halo Model (e.g. [2]) to describe the population of galaxies in haloes according to the properties of the host haloes. In many cases we assume that the properties and population of galaxies depend only on the halo mass. The population of galaxies is then described by the probability $P(N|M)$ that a halo of virial mass M contains N galaxies of a given type. One can then calculate galaxy clustering from the combination of the HOD with the clustering of halos if we assume that the clustering of haloes depends only on the halo mass. If these assumptions are valid we can use galaxy surveys to obtain the relations between properties of galaxies and halo mass, to measure the clustering of dark matter haloes, as well as halo masses (e.g. [3, 4]). However, some studies indicate that several properties of galaxy and halo clustering depend on properties of dark matter haloes other than mass, such as halo formation time, density concentration or subhalo occupation number [5, 6].

On the other hand, semi-analytical models (SAM) populate galaxies in the dark matter haloes by modelling baryonic processes according to the potentials of dark matter [7]. These processes contain free parameters that can be constrained by observations. Because of these processes, semi-analytical models of galaxy formation follow the evolution of the dark matter haloes, mergers, and they are more physical than HOD models in terms of environmental dependences and evolution.

Here we study the consequences of the assumption that the galaxy population and clustering only depends on the halo mass in SAMs. We study the public SAMs of galaxies of the Millennium Simulation to see if we can reproduce the galaxy bias from this HOD assumption, thereby assuming that the clustering of galaxies only depends on the mass of the host halo.

2. Simulation data

We use the Millennium Simulation [8], that corresponds to a Λ CDM cosmology with the parameters: $\Omega_m = \Omega_{dm} + \Omega_b = 0.25$, $\Omega_b = 0.045$, $h = 0.73$, $\Omega_\Lambda = 0.75$, $n = 1$ and $\sigma_8 = 0.9$. Haloes are identified as Friends-of-Friends (FOF) groups. Then, subhaloes are identified in the FOF groups using the SUBFIND [9] algorithm. We studied several SAMs, although here we will focus on [10] (hereafter G11). For the study of the rest of the models we refer to [11].

3. Bias

We estimate the 2-Point Correlation Function (2PCF) using density pixels and the expression

$$\xi(r_{12}) = \langle \delta(r_1) \delta(r_2) \rangle, \quad (3.1)$$

where $\delta(r)$ refers to the density fluctuation defined by $\delta(r) = \rho(r)/\bar{\rho} - 1$ in pixels. From that, we measure the bias using the local bias model:

$$b(r) = \sqrt{\frac{\xi_g(r)}{\xi_m(r)}} \quad (3.2)$$

where $\xi_g(r)$ corresponds to the 2PCF of the studied object (haloes or galaxies), $b(r)$ is the bias factor, and $\xi_m(r)$ is the 2PCF of the dark matter field. We define the mean value by fitting $b(r)$ to a

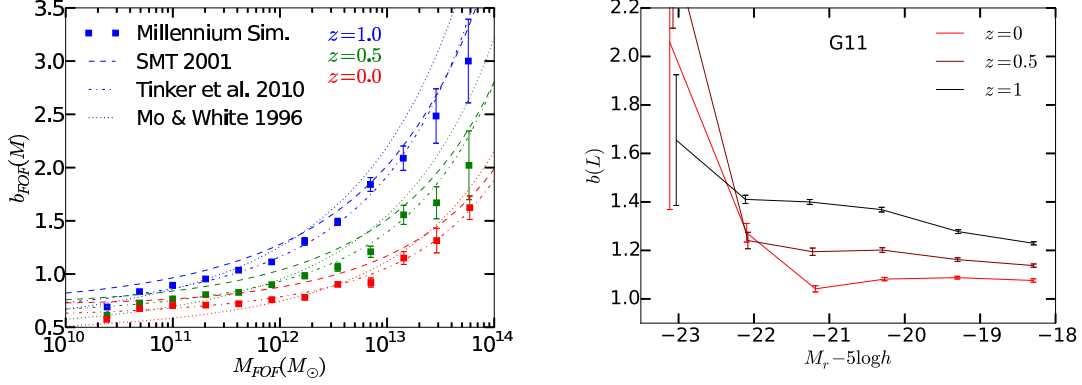


Figure 1: Left: halo bias as a function of mass at 3 different redshifts compared to theoretical expressions. The squares show the measurements of bias from the Millennium Simulation, while the different lines represent different theoretical expressions. Each colour represents a different redshift, as specified. Right: Luminosity dependence (in absolute r band magnitude) of galaxy bias at different redshifts as specified.

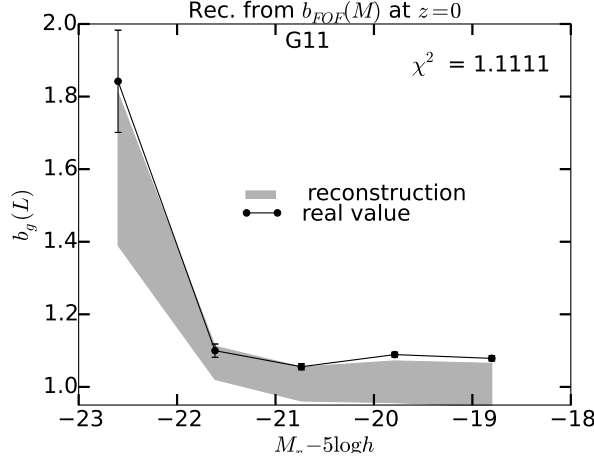


Figure 2: Reconstruction of $b_g(L)$ from $b_{FOF}(M)$ at $z = 0$. The grey shaded region corresponds to the predicted $b_{rec}(L) \pm 1\sigma$, while solid line represents the real measured value of $b_g(L)$.

constant in the scale range $20h^{-1}Mpc < r < 30h^{-1}Mpc$. The errors are measured with a Jack-Knife method of this measurement of b , using 64 cubic subsamples. In the left panel of Fig. 1 we present the measured halo bias as a function of mass and we compare the results with some analytical models [12, 13, 14]. Note that the agreement with the model of [14] expression is remarkable. Right panel of Fig. 1 shows galaxy bias as a function of luminosity, $b_g(L)$ at 3 different redshifts, as specified.

4. Galaxy bias reconstructions

In this section we want to measure if we can recover galaxy bias by assuming that the HOD and galaxy clustering depend only on halo mass. To do this we make a reconstruction of $b_g(L)$, that we will call $b_{rec}(L)$, from the measurements of $b_{FOF}(M)$ and the occupation of galaxies in

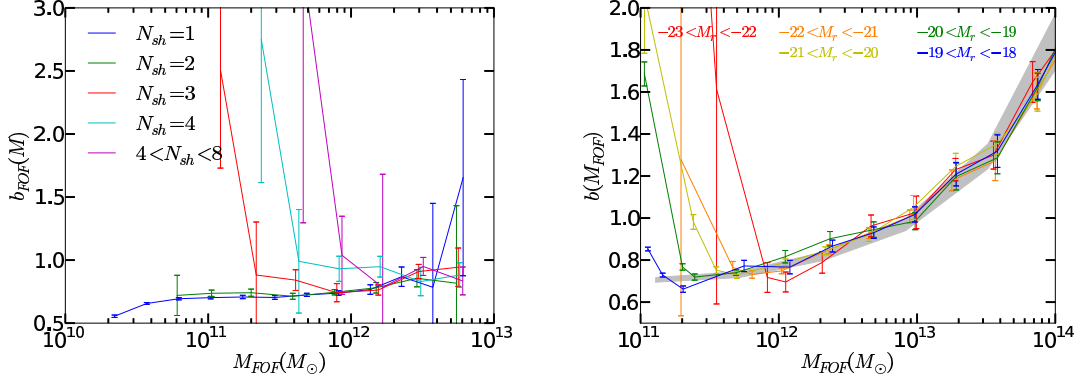


Figure 3: Left: $b_{FOF}(M)$ for different samples according to their number of subhaloes N_{sh} inside at $z = 0$. Right: $b_g(M_{FOF})$ and $b_{rec}(M_{FOF})$ for different luminosity bins. Solid lines represent galaxy bias as a function of FOF mass of their host haloes with different luminosities represented by different colours. The grey shaded zone refer to the range of $b_{rec}(M) \pm 1\sigma$ from b_{FOF} . As the reconstructions are binned by halo mass, $b_{rec}(M) = b_{FOF}M$.

these haloes. If we are in the linear regime (as we are) and the occupation of galaxies is only halo mass dependent (e.g. there is no assembly bias), then the value of $b_g(L)$ must coincide with the reconstruction $b_{rec}(L)$ obtained from the following expression:

$$b_{rec}(L) = \int dM b_h(M) n_h(M) \frac{N_g(M, L)}{n_g(L)} \quad (4.1)$$

where n corresponds to the number density of the galaxies or haloes and $N_g(M, L)$ is the mean number of galaxies per halo of mass M , which we measure directly in the simulation. The error is obtained by calculating the Jack-Knife error of the reconstruction using 64 cubic subsamples.

Fig. 2 shows the reconstructions of $b_g(L)$ from haloes at $z = 0$. We see that the reconstructions tend to be different from the SAM measurements by a factor of 6–7% at the level of 1σ . Although not shown here, this is common for all the SAMs and halo definitions [11]. These differences in clustering corresponds to a 50% difference in halo mass (see Fig. 1).

For equation (4.1) to be accurate we need to satisfy one of the following to conditions. First, all the haloes of the same mass have the same clustering. If this is the case, all the galaxies in these haloes have the same clustering and then we are assigning the correct clustering for the galaxies. Second, galaxies populate haloes only according to their mass. If this is the case, even if the first condition is not satisfied the galaxies in the same masses must statistically have the same mean clustering. We have seen that the reconstructions differ from the measurement of $b_g(L)$, so this means that both conditions fail. So, for a fixed halo mass, different haloes must have different clustering (assembly bias). Moreover, the population of galaxies in haloes of the same mass must be correlated with the halo bias.

In the left panel of Fig. 3 we can see $b_{FOF}(M)$ separating the samples of FOFs according to the number of subhaloes in the halo. We see that for a fixed mass (at low masses at least), the dependence of b on subhalo occupation, N_{sh} , is very strong, while in the reconstruction we are wrongly assuming that bias only depends on mass. This is an indication of assembly bias. As

the haloes with higher N_{sh} have more galaxies, the reconstructions produce an underestimation of $b_g(L)$, since we are assuming the same mean bias for these galaxies, while the haloes with more galaxies have a higher bias than the mean value of these masses.

To study the mass dependence of the success in the reconstructions, we measured $b_g(M_{FOF})$ in several luminosity bins. Right panel of Fig. 3 shows $b_g(M_{FOF})$ for galaxies at different luminosities. This figure, then, allows us to see explicitly how the reconstructions of galaxy bias work at different masses. We can see two different behaviours. At high masses, $b_g(M_{FOF})$ is close to $b_{rec}(M_{FOF})$. This means that the reconstruction of $b_g(L)$ at these masses works, and the halo mass gives sufficient information to predict galaxy clustering. In the low mass regions there is a strong disagreement between $b_g(M_{FOF})$ and $b_{rec}(M_{FOF})$, especially for the brightest galaxies. The bias of the brightest galaxies is much higher than the mean one of the haloes of the corresponding mass. So, in these low masses, the galaxies are populated precisely in a way that the brightest galaxies are in the most clustered haloes of the corresponding mass. This means, again, that the clustering of these galaxies does not only depend on mass, and this is also another indication of assembly bias, since haloes of the same mass must have a different clustering.

5. Summary and discussion

The standard implementation of the HOD assumes that the clustering of galaxies only depends on the mass of the haloes. If for a fixed halo mass there are more galaxies in the halos with more subhalos, then the mean clustering of these galaxies will be higher than the mean clustering of the haloes and we will wrongly conclude that they are in more massive haloes. This results in an overestimation of the halo mass using clustering. We have shown that this is the case for SAMs, but we expect this to also be true in any other models of galaxy formation where the number of galaxies correlate with the number of subhalos.

The reconstructions of b_g from haloes work better at high masses, but some disagreements with b_g come from the low mass haloes, where the assembly bias effect is stronger. This effect occurs for $M_h \lesssim 3 - 5 \times 10^{11} M_\odot$. This indicates that the SAMs agree with our assumptions at large masses. However, other dependences than mass are needed to predict the clustering of the SAMs on small masses, then care must be taken when assuming the HOD model at masses below $3 - 5 \times 10^{11} M_\odot$, especially when assuming that galaxy clustering only depends on mass.

Frequently, the HOD is assumed to relate galaxy properties and halo masses [3, 4]. In the case of [3], they measure the HOD parameters from the clustering of galaxies assuming that the clustering only depends on the halo mass. But these results can be affected by the halo assembly bias since this can affect galaxy clustering.

Acknowledgements

Funding for this project was partially provided by the Spanish Ministerio de Ciencia e Innovación (MICINN), project AYA2012-39559, Consolider-Ingenio CSD2007- 00060, European Commission Marie Curie Initial Training Network CosmoComp (PITN-GA-2009-238356), research project 2009- SGR-1398 from Generalitat de Catalunya. A.P. was supported by beca FI from Generalitat de Catalunya.

References

- [1] A. A. Berlind, D. H. Weinberg, *The Halo Occupation Distribution: toward an empirical determination of the relation between galaxies and mass*, 2002, ApJ, 575, 587, [astro-ph/0109001]
- [2] A. Cooray, R. K. Sheth, *Halo models of large scale structure*, 2002, PhysRep, 372, 1, [astro-ph/0206508]
- [3] I. Zehavi, Z. Zheng, D. H. Weinberg, M. R. Blanton, N. A. Bahcall, A. A. Berlind, J. Brinkmann, J. A. Frieman et al., *Galaxy Clustering in the Completed Sdss Redshift Survey: the Dependence on Color and Luminosity*, 2011, ApJ, 736, 59, [arXiv:1005.2413]
- [4] J. Coupon, M. Kilbinger, H. J. McCracken, O. Ilbert, S. Arnouts, Y. Mellier, U. Abbas, S. de la Torre et al., *Galaxy clustering in the CFHTLS-Wide: the changing relationship between galaxies and haloes since $z \sim 1.2^*$* , 2012, A & A, 542, A5, [arXiv:1107.0616]
- [5] L. Gao, V. Springel, S. D. M. White, *The age dependence of halo clustering*, 2005, MNRAS, 363, L66, [astro-ph/0506510]
- [6] R. H. Wechsler, A. R. Zentner, J. S. Bullock, A. V. Kravtsov, B. Allgood, *The Dependence of Halo Clustering on Halo Formation History, Concentration, and Occupation*, 2006, 652, 71, [astro-ph/0512416]
- [7] S. Cole, C. G. Lacey, C. M. Baugh, C. S. Frenk, *Hierarchical galaxy formation*, 2000, MNRAS, 319, 168, [astro-ph/0007281]
- [8] V. Springel, S. D. M. White, A. Jenkins, C. S. Frenk, N. Yoshida, L. Gao, J. Navarro, R. Thacker et al., *Simulations of the formation, evolution and clustering of galaxies and quasars*, 2005, Nature, 435, 629, [astro-ph/0504097]
- [9] V. Springel, S. D. M. White, G. Tormen, G. Kauffmann, *Populating a cluster of galaxies - I. Results at $z = 0$* , 2001, MNRAS, 328, 726, [astro-ph/0012055]
- [10] Q. Guo, S. D. M. White, M. Boylan-Kolchin, G. De Lucia, G. Kauffmann, G. Lemson, C. Li, V. Springel, S. Weinmann, *From dwarf spheroidals to cD galaxies: simulating the galaxy population in a Λ CDM cosmology*, 2011, MNRAS, 413, 101, [arXiv:1006.0106]
- [11] A. Pujol, E. Gaztañaga, *Are the halo occupation predictions consistent with large-scale galaxy clustering?*, 2014, MNRAS, 442, 1930, [arXiv:1306.5761]
- [12] H. J. Mo, S. D. M. White, *An analytic model for the spatial clustering of dark matter haloes*, 1996, MNRAS, 282, 347, [astro-ph/9512127]
- [13] R. K. Sheth, H. J. Mo, G. Tormen, *Ellipsoidal collapse and an improved model for the number and spatial distribution of dark matter haloes*, 2001, MNRAS, 323, 1, [astro-ph/9907024]
- [14] J. L. Tinker, B. E. Robertson, A. V. Kravtsov, A. Klypin, M. S. Warren, G. Yepes, S. Gottlöber, *The Large-scale Bias of Dark Matter Halos: Numerical Calibration and Model Tests*, 2010, ApJ, 724, 878, [arXiv:1001.3162]