Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **496,** 1182–1196 (2020)
Advance Access publication 2020 June 16

doi:10.1093/mnras/staa1624

The manifestation of secondary bias on the galaxy population from IllustrisTNG300

Antonio D. Montero-Dorta,[1]★ M. Celeste Artale,[2] L. Raul Abramo,[1] Beatriz Tucci,[1] Nelson Padilla,[3,4] Gabriela Sato-Polito,[5] Ivan Lacerna,[6,7] Facundo Rodriguez[8,9] and Raul E. Angulo[10,11]

[1]*Departamento de Física Matemática, Instituto de Física, Universidade de São Paulo, Rua do Matão 1371, CEP 05508-090, São Paulo, Brazil*
[2]*Institut für Astro- und Teilchenphysik, Universität Innsbruck, Technikerstrasse 25/8, A-6020 Innsbruck, Austria*
[3]*Instituto de Astrofísica, Pontificia Universidad Católica de Chile, 8970117, Santiago, Chile*
[4]*Centro de Astro-Ingeniería, Pontificia Universidad Católica de Chile, 8970117, Santiago, Chile*
[5]*Department of Physics & Astronomy, Johns Hopkins University, 3400 N. Charles St, Baltimore, MD 21218, USA*
[6]*Instituto de Astronomía y Ciencias Planetarias, Universidad de Atacama, Copayapu 485, Copiapó, Chile*
[7]*Instituto Milenio de Astrofísica, Av. Vicuña Mackenna 4860, Macul, Santiago, Chile*
[8]*Observatorio Astronómico de Córdoba, Universidad Nacional de Córdoba, Laprida 854, X5000BGR, Córdoba, Argentina*
[9]*CONICET, Instituto de Astronomía Teórica y Experimental, Laprida 854, X5000BGR, Córdoba, Argentina*
[10]*Donostia International Physics Center (DIPC), Paseo Manuel de Lardizábal, 4, E-20018 Donostia-San Sebastián, Spain*
[11]*IKERBASQUE, Basque Foundation for Science, E-48013 Bilbao, Spain*

Accepted 2020 May 24. Received 2020 May 12; in original form 2020 January 4

ABSTRACT

We use the improved IllustrisTNG300 magnetohydrodynamical cosmological simulation to revisit the effect that secondary halo bias has on the clustering of the central galaxy population. With a side length of 205 h^{-1} Mpc and significant improvements on the subgrid model with respect to previous Illustris simulations, IllustrisTNG300 allows us to explore the dependencies of galaxy clustering over a large cosmological volume and halo mass range. We show at high statistical significance that the halo assembly bias signal (i.e. the secondary dependence of halo bias on halo formation redshift) manifests itself on the clustering of the galaxy population when this is split by stellar mass, colour, specific star formation rate, and surface density. A significant signal is also found for galaxy size: at fixed halo mass, larger galaxies are more tightly clustered than smaller galaxies. This effect, in contrast to the rest of the dependencies, seems to be uncorrelated with halo formation time, with some small correlation only detected for halo spin. We also explore the transmission of the spin bias signal, i.e. the secondary dependence of halo bias on halo spin. Although galaxy spin retains little information about the total halo spin, the correlation is enough to produce a significant galaxy spin bias signal. We discuss possible ways to probe this effect with observations.

Key words: methods: numerical – galaxies: formation – galaxies: haloes – cosmology: theory – dark matter – large-scale structure of Universe.

1 INTRODUCTION

In the standard model of cosmology, dark matter (DM) clusters along density peaks that were generated during inflation, collapsing later on to form DM *haloes*. It is inside these collapsing structures that galaxies form, when gas falls into their potential wells (e.g. White & Frenk 1991). The relationship between galaxies, haloes, and the underlying matter distribution is therefore crucial to our understanding of the galaxy formation process, and to our ability to test cosmological models against observations.

The clustering of DM haloes is commonly characterized by *halo bias*, which can be broadly defined as the relation between the spatial distribution of haloes and the underlying matter density field. In its simplest description, the linear halo bias can be assumed to depend only on halo mass, with more massive haloes being more strongly clustered than less massive haloes (e.g. Kaiser 1984; Sheth & Tormen 1999). However, halo clustering is a very complex process that is known to depend on a variety of secondary halo properties. Among these secondary dependencies, the one that has drawn more

★ E-mail: amonterodorta@gmail.com

attention is the dependence on the assembly history of haloes, an effect called *halo assembly bias*[1]. Lower mass haloes that assemble a significant portion of their mass early on are more tightly clustered than haloes that assemble at later times, *at fixed halo mass* (see e.g. Sheth & Tormen 2004; Gao, Springel & White 2005; Wechsler et al. 2006; Gao & White 2007; Angulo, Baugh & Lacey 2008; Li, Mo & Gao 2008; Lazeyras, Musso & Schmidt 2017; Han et al. 2019; Mao, Zentner & Wechsler 2018; Salcedo et al. 2018; Johnson et al. 2019; Sato-Polito et al. 2019). At higher halo masses ($M_{\rm vir} \gtrsim 10^{14}$ $h^{-1}\,{\rm M}_{\odot}$), the signal seems to depend strongly on the definition of halo age (Li et al. 2008; Chue, Dalal & White 2018). Besides halo assembly bias, a number of other secondary dependencies have been identified for halo clustering (on, e.g. concentration, spin, shape, and environment), but a comprehensive theory for the physical origins of these effects is yet to be established (see e.g. Dalal et al. 2008; Paranjape, Hahn & Sheth 2018; Ramakrishnan et al. 2019).

A question that has stirred up debate in recent years is whether the above secondary halo bias effects manifest themselves on the galaxy population, and, if so, whether they can actually be detected with current data. The term *galaxy assembly bias* is often used to refer to the dependence of galaxy clustering on secondary halo properties beyond halo mass (see e.g. Croton et al. 2006; Zhu et al. 2006; Lacerna, Padilla & Stasyszyn 2014; Zentner, Hearin & van den Bosch 2014; Hearin, Watson & van den Bosch 2015; Miyatake et al. 2016; Lin et al. 2016; Zentner et al. 2019; Zu et al. 2017; Montero-Dorta et al. 2017b; Romano-Díaz et al. 2017; Niemiec et al. 2018; Walsh & Tinker 2019). This signal could be simply a direct manifestation of secondary halo bias on the galaxy population, but could also include other contributions that are not directly related to halo formation. In the context of halo occupation distribution (HOD) modelling, galaxy assembly bias is described as the combination of halo assembly bias and the so-called *occupancy variations*, i.e. the dependencies of the galaxy content of haloes on secondary halo properties at fixed halo mass (see e.g. Artale et al. 2018; Zehavi et al. 2018; Bose et al. 2019).

With the Sloan Digital Sky Server (SDSS; York et al. 2000) at low redshift, Miyatake et al. (2016) and More et al. (2016) claimed the detection of the so-called concentration bias on galaxy clusters, i.e. the secondary dependence of clustering on concentration (which is related to halo formation time[2]), for high-mass haloes. This claim was subsequently refuted by the evidence of projection effects that affected significantly the identification of cluster members (Zu et al. 2017; Sunayama & More 2019). Also using the SDSS, Lin et al. (2016) found little evidence of any dependence of galaxy clustering on star formation history (SFH) beyond what is expected from the measured halo–mass difference. These results appear in some tension with findings for luminous red galaxies (LRGs) at $z = 0.55$ (see Montero-Dorta et al. 2017b and Niemiec et al. 2018, who used the Baryon Oscillation Spectroscopic Survey, BOSS, Dawson et al. 2013). In these works, a dependence of the amplitude of clustering on SFH is shown for LRGs, which also seems to occur at fixed halo mass (within the weak-lensing errors). Several other works have addressed the galaxy assembly bias question, producing mixed results. Lacerna et al. (2014), for instance, used both SDSS and mock galaxies to claim a weak, although significant detection of assembly bias for central galaxies.

Despite these contradictory results, a convincing proof of the existence of galaxy assembly bias seems attainable with upcoming cosmological surveys, which will reduce significantly the errors on the weak-lensing and clustering measurements, and will allow a better identification and characterization of galaxy clusters (see discussion in, e.g. Wechsler & Tinker 2018; Zentner et al. 2019). Note that, although most of the previous galaxy assembly bias analyses have focused on properties related to halo accretion history, it is conceivable that other secondary halo bias dependencies leave an imprint on the galaxy population. The dependence on halo spin seems particularly appealing in this context, given its large effect at the high-mass end (see e.g. Johnson et al. 2019; Sato-Polito et al. 2019).

In this paper, we use the IllustrisTNG simulation at redshift $z = 0$ to study how the secondary halo bias signal transmits to the galaxy population. IllustrisTNG is an ongoing suite of magnetohydrodynamical cosmological simulations that model the formation and evolution of galaxies within the lambda cold dark matter (ΛCDM) paradigm. Multiple refinements in the subgrid model have improved significantly the performance of the simulation with respect to Illustris, in terms of reproducing important observational constraints such as the stellar mass function, along with halo–galaxy relations such as the stellar-to-halo mass relation (SHMR). These updates range from a new implementation of galactic winds, to black hole-driven kinetic feedback at low accretion rates and the inclusion of magnetohydrodynamics (Weinberger et al. 2017; Pillepich et al. 2018). In the context of assembly bias, hydrodynamical simulations provide predictions that are less model dependent than those coming from semi-analytic models (SAMs), even though these have proven to be excellent tools to investigate multiple aspects of the effect (see e.g. Croton et al. 2006; Lacerna et al. 2018; Zehavi et al. 2018, 2019; Contreras et al. 2019; Padilla et al. 2019).

Among the different IllustrisTNG boxes, we choose the IllustrisTNG300 box, which presents an additional advantage as compared to previous Illustris versions. With a side length of 205 h^{-1} Mpc, it starts to be statistically comparable to some of the *N*-body numerical simulations that have been used in the context of the measurement of secondary halo bias, such as the MultiDark boxes (Klypin & Prada 2017). This allows us to increase the statistical significance of the measurement presented in Xu & Zheng (2020), who used the 75 h^{-1} Mpc Illustris-2 box.

The aforementioned analysis from Xu & Zheng (2020) shows that the halo assembly bias signal is indeed reflected on the clustering of the galaxy population (a prediction that, again, has not been convincingly confirmed with observations). Their results from Illustris-2 also expose a strong correlation between the stellar mass of central galaxies and the peak maximum circular velocity of their hosting haloes ($V_{\rm peak}$). The maximum circular velocity has, in fact, been proposed as a more efficient property in terms of encapsulating the assembly bias effect, as compared to halo mass (see e.g. Xu & Zheng 2020; Zehavi et al. 2005, 2019).

Using the 67 h^{-1} Mpc EAGLE simulation, Chaves-Montero et al. (2016) quantified the assembly bias clustering effect in 20 per cent with respect to the standard mass-based subhalo (hereafter *subhalo*) abundance matching (SHAM) framework. More recently, Hadzhiyska et al. (2020) investigated the effect of adding secondary dependencies of halo clustering to the standard HOD framework, showing that the local environment, the velocity dispersion anisotropy, and the product of the half-mass radius and the velocity

[1]Throughout this paper, halo assembly bias is considered a particular case of secondary halo bias where the secondary halo property considered is directly related to halo formation time or halo accretion history.

[2]Note that, nevertheless, the secondary halo bias signal found for halo age is different to that found for concentration (see e.g. Chue et al. 2018; Salcedo et al. 2018; Sato-Polito et al. 2019).

dispersion of the halo produces the best agreement with respect to the clustering of IllustrisTNG300 galaxies.

Several other aspects that are tightly related to assembly bias have been addressed with hydrodynamical simulations. These include the effect of *galactic conformity*, a term that generally refers to the correlations in the colours and star formation rates of neighbouring galaxies (see Bray et al. 2016 for an analysis on Illustris), and the aforementioned occupancy variations (Artale et al. 2018; Bose et al. 2019). The analysis on occupancy variations presented in Bose et al. (2019) can, in fact, be seen as complementary to our work. Their results, also obtained with IllustrisTNG300, indicate a dependence of the number of satellites on halo properties, at fixed halo mass. It is shown that haloes tend to harbour more satellites when they are less concentrated or younger, live in dense environments, and have higher angular momenta. The probability of hosting a central galaxy, on the other hand, is enhanced for low-mass high-concentration haloes and for low-mass haloes that live in overdense regions.

The paper is organized as follows. Sections 2 and 3 provide, respectively, brief descriptions of both the simulation boxes and the halo and galaxy properties analysed in this work. The methodology for the computation of the relative bias between subsets of haloes/galaxies is explained in Section 4. The main results of our analysis in terms of the manifestation of secondary halo bias on the galaxy population are presented in Section 5. Finally, Section 6 is devoted to discussing the implications of these results and providing a brief summary of the paper. The IllustrisTNG300 simulation adopts the standard ΛCDM cosmology (Planck Collaboration XIII 2016), with parameters $\Omega_{\rm m} = 0.3089$, $\Omega_{\rm b} = 0.0486$, $\Omega_{\Lambda} = 0.6911$, $H_0 = 100\,h\,{\rm km\,s^{-1}Mpc^{-1}}$ with $h = 0.6774$, $\sigma_8 = 0.8159$, and $n_{\rm s} = 0.9667$.

2 SIMULATIONS

In this paper, we use the galaxy and DM halo catalogues from *The Next Generation* Illustris (IllustrisTNG)[3] magnetohydrodynamical cosmological simulations, which represent an updated version of the Illustris simulations (Genel et al. 2014; Vogelsberger et al. 2014a, b). The IllustrisTNG simulations are performed with the AREPO moving-mesh code (Springel 2010) and include subgrid models that account for radiative metal-line gas cooling, star formation, chemical enrichment from SNII, SNIa, and AGB stars, stellar feedback, supermassive-black hole formation with multimode quasar, and kinetic black hole feedback. The main updates with respect to the Illustris simulation are: a new implementation of black hole kinetic feedback at low accretion rates, a revised scheme for galactic winds, and the inclusion of magnetohydrodynamics (see Pillepich et al. 2018; Weinberger et al. 2017, for further details).

In this work, we analyse the IllustrisTNG300-1 run and its *DM-only* counterpart IllustrisTNG300-1-DMO (hereafter IllustrisTNG300 and IllustrisTNG300-DMO, respectively), which are the largest simulated boxes from the IllustrisTNG suite featuring the highest resolution level. These runs adopt a cubic box of side $205\,h^{-1}$ Mpc with periodic boundary conditions. The IllustrisTNG300 run follows the evolution of 2500^3 DM particles of mass $4.0 \times 10^7\,h^{-1}\,{\rm M}_{\odot}$, and 2500^3 gas cells of mass $7.6 \times 10^6 h^{-1}\,{\rm M}_{\odot}$. The IllustrisTNG300-DMO boxes contain 2500^3 DM particles with mass $7.0 \times 10^7 h^{-1}\,{\rm M}_{\odot}$.

The IllustrisTNG300 simulation has proven to be a powerful and self-consistent tool to investigate the distribution of galaxies and DM haloes on large scales, one of the main focuses of our work. Springel et al. (2018) analysed the two-point correlation function of galaxies (and DM haloes) in IllustrisTNG300, finding good agreement with observations in terms of the stellar-mass and colour dependence of galaxy clustering. Also complementary to our analysis is the study presented in Bose et al. (2019), who explored occupancy variations of DM haloes at a fixed halo mass as a function of environment and secondary properties such as concentration, formation time, and angular momentum. Several other constraints related to the properties of galaxies have been shown to agree with observations, including the size distributions of star-forming and quiescent galaxies (Genel et al. 2018), the evolution of the galaxy stellar mass function (Pillepich et al. 2018), and the location and shape of the red sequence and blue cloud of the $z = 0$ SDSS galaxy population (Nelson et al. 2018). These results lay a solid foundation for our analysis.

3 HALO AND GALAXY PROPERTIES

In IllustrisTNG300, DM haloes are identified using a friends-of-friends (FOF) algorithm with a linking length of 0.2 times the mean interparticle separation (Davis et al. 1985). The gravitationally bound substructures called subhaloes are subsequently identified using the SUBFIND algorithm (Springel et al. 2001; Dolag et al. 2009). Among all subhaloes, those containing a non-zero stellar component are considered galaxies. Each DM halo typically contains multiple galaxies, including a central galaxy and several satellites, where the positions of centrals coincide with the FOF centres.

In order to compute the formation time of haloes, we use the subhalo merger trees that were obtained with the SUBLINK algorithm (Rodriguez-Gomez et al. 2015) and are publicly available at the IllustrisTNG data base.

In our analysis, we focus on the following halo properties:

(i) Virial mass, $M_{\rm vir}$ $[h^{-1}\,{\rm M}_{\odot}]$, computed by adding up the mass of all gas cells and particles enclosed within a sphere of radius $R_{\rm vir}$ that is defined so that the enclosed density equals 200 times the critical density. Note that throughout this work all virial quantities, denoted by the subscript 'vir', are computed within this radius (i.e. assuming a density threshold $\Delta = 200$).

(ii) Formation redshift, $z_{1/2}$, defined as the redshift at which half of the present-day halo mass has been accreted into a single subhalo for the first time. For this, we use the progenitors of the main branch of the subhalo merger tree computed with SUBLINK, which is initialized at $z = 6$.

(iii) Spin, $\lambda_{\rm halo}$, defined as in Bullock et al. (2001)

$$\lambda_{\rm halo} = \frac{|{\rm J}|}{\sqrt{2} M_{\rm vir} V_{\rm vir} R_{\rm vir}}, \tag{1}$$

where J is the angular momentum of the halo and $V_{\rm vir}$ is its circular velocity at the virial radius $R_{\rm vir}$. Note that some works use the Peebles (1969) halo spin definition in the context of secondary bias (see e.g. Lacerna & Padilla 2012), but this choice is known to have little impact on the secondary bias signal (Johnson et al. 2019).

(iv) Concentration, $c_{\rm vir}$, defined as

$$c_{\rm vir} = \frac{R_{\rm vir}}{R_{\rm s}}, \tag{2}$$

where $R_{\rm s}$ is the scale radius, obtained from fitting the DM density profiles of individual haloes with an NFW profile (Navarro, Frenk & White 1997).

[3] http://www.tng-project.org

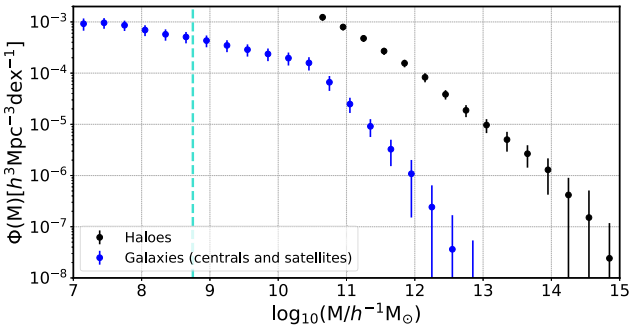

Figure 1. The halo and stellar mass functions in the IllustrisTNG300 simulation at $z = 0$. In this paper, only haloes of $\log_{10}(M_{\rm vir}/h^{-1}\,{\rm M}_\odot) > 10.5$ are considered. In this plot, the mass labelled 'M' corresponds to the virial halo mass ($M_{\rm vir}$) for haloes and the stellar mass (M_*) for galaxies. Error bars show the box-to-box variation for a set of 64 subboxes. Note that the mass resolution is $4.0 \times 10^7\ h^{-1}\,{\rm M}_\odot$ for DM particles and $7.6 \times 10^6\ h^{-1}\,{\rm M}_\odot$ for baryonic particles. The vertical dashed line marks our resolution limit of 50 gas particles for galaxies.

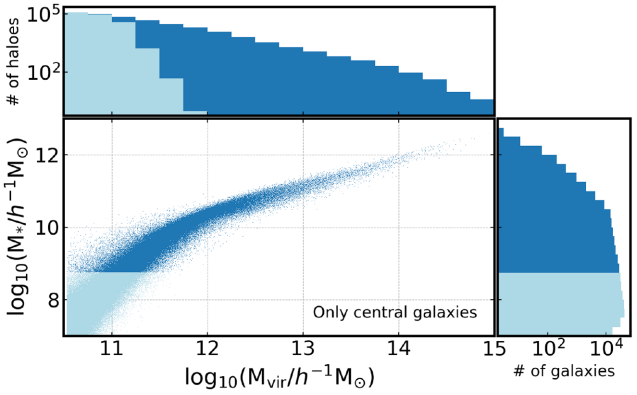

Figure 2. The relation between stellar mass and halo mass for central galaxies in the IllustrisTNG300 simulation. The histograms above and to the right of the scatter plot show the distributions of halo virial mass and stellar mass, respectively, in logarithmic units. Lighter colours mark the distribution of galaxies (and their hosting haloes) below our resolution limit of 50 gas particles.

For the simulated galaxies (i.e. subhaloes with non-zero stellar components), we investigate the secondary bias effect coming from:

(i) Star formation rate, SFR [$\rm M_\odot\ yr^{-1}$], computed as the sum of the star formation rate of all gas cells in each subhalo.

(ii) Stellar mass, M_* [$h^{-1}\,{\rm M}_\odot$], defined as the total mass of all stellar particles bound to each subhalo.

(iii) Specific star formation rate, sSFR [$h\ \rm yr^{-1}$], defined as sSFR = SFR/M$_*$.

(iv) Stellar half-mass radius, $R^{(*)}_{1/2}$ [$h^{-1}\,$kpc], defined as the comoving radius containing half of the stellar mass of each subhalo.

(v) Subhalo half-mass radius, $R^{\rm (sh)}_{1/2}$ [$h^{-1}\,$kpc], defined as the comoving radius containing half of the total (stellar + gas + DM) mass of each subhalo.

(vi) Galaxy colour (g–i), computed using the magnitudes provided at the IllustrisTNG data base. The magnitudes are computed by summing up the luminosities of the stellar particles of each subhalo (based on the procedure of Buser 1978). These magnitudes are intrinsic, i.e. the attenuation produced by dust is not included.

(vii) Galaxy total spin, $\lambda_{\rm galaxy}$, defined as in equation (1) but using all the particles (DM + stellar + gas components) inside the stellar half-mass radius of the galaxy, $R^{(*)}_{1/2}$.

(viii) Galaxy stellar spin, λ_*, defined as in equation (1), but using only the stellar component inside $R^{(*)}_{1/2}$.

(ix) Surface density [$h\ \rm M_\odot\ kpc^{-2}$], defined as the stellar mass divided by the square of the stellar half-mass radius.

(x) Velocity dispersion, σ [$\rm km\,s^{-1}$], defined as the one-dimensional velocity dispersion of all the particles and cells of each subhalo.

Fig. 1 shows the halo and stellar mass functions in the IllustrisTNG300 $z = 0$ box (including both centrals and satellites). In order to ensure good resolution, only haloes above $\log_{10}(M_{\rm vir}/h^{-1}\,{\rm M}_\odot) > 10.5$ and central galaxies with $\log_{10}({\rm M}_*/h^{-1}\,{\rm M}_\odot) > 8.75$ are considered in this analysis. This imposes a threshold of at least ∼1000 particles per halo and ∼50 gas cells for central galaxies.

In Fig. 2, we show the relation between halo mass and the stellar mass of central galaxies in the simulation (the SHMR), along with the distributions of these two quantities (in logarithmic scale).

4 METHODOLOGY: RELATIVE BIAS MEASUREMENT

In order to quantify the dependence of clustering on a secondary property S, we measure the relative bias, $b_{\rm rel}$, between a subset of objects selected according to S and all objects in the same primary bias property (B) bin. Note that in this analysis, the primary property B is always the halo virial mass, $M_{\rm vir}$, whereas S can be either a halo or a central galaxy property. We use subsets of 50 per cent of the entire sample in each halo mass bin, following Xu & Zheng (2020). This choice slightly attenuates the magnitude of the secondary bias effect with respect to the standard way of presenting the measurement (based on 25 per cent quartiles), but it allows us to improve the statistics at the high-mass end significantly.

In order to further increase the signal-to-noise ratio in the computation of the relative bias, we use the cross-correlation with the entire sample (i.e. all mass bins), for both the S-subset and the B-bin. Namely, for a given halo-mass bin B and scale r

$$b_{\rm rel}(r, {\rm S|B}) = \frac{\xi_{\rm [S,all]}(r, {\rm S})}{\xi_{\rm [B,all]}(r, {\rm B})}, \tag{3}$$

where $\xi_{\rm [S,\,all]}$ is the cross-correlation between all objects in the S-subset and all objects in the sample, and $\xi_{\rm [B,\,all]}$ is the cross-correlation between all objects in the halo-mass bin and the entire sample. Note, again, that $B \rightarrow M_{\rm vir}$.

The computation of errors is based on a standard jackknife technique. The IllustrisTNG300 box at $z = 0$ is divided in 8 subboxes with $L_{\rm sub\,-\,box} = L_{\rm box}/2 = 102.5\,h^{-1}\,$Mpc. The relative bias of equation (3) is measured in 8 different configurations of equal volume, obtained by subtracting one subbox at a time. The errors on the measured relative bias correspond to the standard deviation of all individual configurations.

The 2-point correlation function is measured using CORRFUNC (Sinha & Garrison 2017) in a range of scales between 5 and 12 $h^{-1}\,$Mpc. We choose this range of scales due to the higher signal-to-noise ratio in the assembly bias detection. The maximum scale is set at 12 $h^{-1}\,$Mpc in order to avoid problems derived from the relatively small size of the simulation box. For simplicity, to analyse the secondary bias signals, we use the relative bias $b_{\rm rel}$(S|B) (or simply $b_{\rm rel}$) averaged over the aforementioned scale range $5 < r[h^{-1}\,{\rm Mpc}] < 12$.

5 RESULTS

The manifestation of secondary halo bias on the clustering and properties of the central galaxy population is addressed in this section. We use several galaxy properties to split the galaxy population at fixed halo mass and measure the relative bias between subsets, as described in Section 4. Before showing the results derived from this analysis, we will present the secondary halo bias effect measured from IllustrisTNG300 (which serves as the *reference signal*) and discuss the correlations between halo and galaxy properties.

5.1 Secondary halo bias

The dependence of halo clustering on multiple secondary halo properties at fixed halo mass (or velocity) has been measured extensively from different suites of *N*-body numerical simulations (see e.g. Gao et al. 2005; Wechsler et al. 2006; Gao & White 2007; Han et al. 2019; Salcedo et al. 2018; Sato-Polito et al. 2019). As a sanity check, we begin by measuring the effect from haloes in both the IllustrisTNG300 and IllustrisTNG300-DMO boxes. From this comparison, the effect of baryons on the measured secondary bias effect can be assessed.

Fig. 3 presents the secondary bias signal for age (i.e. formation redshift, $z_{1/2}$), concentration, and spin. We recall that in order to maximize signal-to-noise ratio, subsets of 50 per cent of the entire distribution in each halo-mass bin are employed (instead of the standard 25 per cent quartiles). We have checked that this choice has little effect, qualitatively, on the main conclusions of this paper.

In order to further reduce the noise of the measurement at the high-mass end, where haloes are scarce, a varying bin size is adopted (see also Xu & Zheng 2020). At $\log_{10}(M_{\rm vir}/h^{-1}\,{\rm M}_\odot) < 12.75$ we use $\Delta\log_{10}(M_{\rm vir}) = 0.15$ (the typical value in these analyses), but we choose $\Delta\log_{10}(M_{\rm vir}) = 0.3$ above (with a larger single bin for very high mass haloes).

For concentration and age, Fig. 3 displays very similar trends. Older/more-concentrated haloes are more tightly clustered than younger/less-concentrated haloes below $M_{\rm vir} \sim 10^{13}\ h^{-1}\ {\rm M}_\odot$, as expected (with the effect being slightly larger for halo age). Above this halo mass, the trend appears to invert. Interestingly, this behaviour, the inversion of the signal, is very well documented for concentration, but not for $z_{1/2}$, for which most measurements show the signal vanishing (this is only strictly true for $z_{1/2}$, since Chue et al. 2018 show that the amount of assembly bias at the high-mass end depends on the fraction of already formed mass used to define halo age). These are, however, the mass ranges for which we must be particularly careful, due to the low-number statistics of the sample.

Also interesting in Fig. 3 is the result for so-called *spin bias*, i.e. the secondary dependence on spin. Throughout the entire mass range, faster-rotating haloes are more tightly clustered than lower spin haloes. This result is in some degree of tension with recent findings from *N*-body numerical simulations such as MultiDark or Vishnu (Sato-Polito et al. 2019 and Johnson et al. 2019, respectively), which show a statistically significant inversion of the trend at the low-mass end (i.e. 'the spin-bias inversion'). Interestingly, as we discuss in Section 5.4 below, the crossover is indeed found when the selection is performed on the basis of the central galaxy spin. These apparent inconsistencies deserve further investigation, since they can potentially reveal clues on the origins of this secondary bias effect itself (Tucci et al. in preparation).[4]

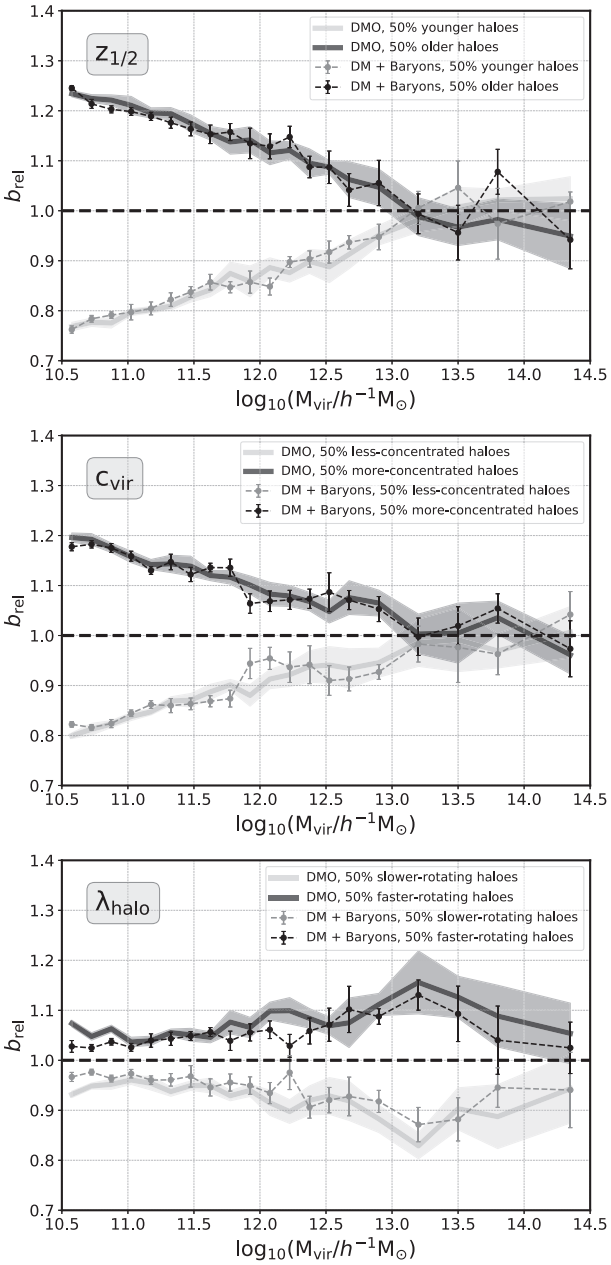

Figure 3. Secondary halo bias for halo age (described by $z_{1/2}$), concentration (c_{200}), and spin ($\lambda_{\rm halo}$), for IllustrisTNG300 (dashed lines and dots) and IllustrisTNG300-DMO (solid lines) at $z = 0$. In each panel, darker symbols/lines represent the 50 per cent subset of haloes with higher values of the corresponding secondary property, whereas lighter symbols/lines show the remaining 50 per cent lower value subset. Error bars display the jackknife error on the relative bias obtained from a set of 8 subboxes, as described in Section 4.

The scope of this analysis is to show how secondary halo bias transmits to the galaxy population, and thus we will not concentrate on the particular details of the halo bias signal. We note also that the apparent decrease in the spin bias signal at the very

[4]Note that in all previous works ROCKSTAR (Behroozi, Wechsler & Wu 2013) was used to identify haloes, whereas IllustrisTNG300 employs an FOF algorithm. The *spin-bias inversion* was found for haloes of $\log_{10}(M_{\rm vir}/h^{-1}\,{\rm M}_\odot) \simeq 11.5$, a range of masses where small differences in the halo definition could potentially have significant effects on the clustering measurements.

high mass end is probably an artefact due to the low number of haloes.

Fig. 3 also shows that the effect of baryons on the secondary bias signal is not significant (the solid lines in each panel display the measurement in the DMO box). As expected, the trends are qualitatively very similar, with only small fluctuations that are consistent with the measured uncertainties.

Finally, it is worth reminding the reader at this point that the different secondary halo bias trends are not mutually independent. Since the different halo properties are correlated with each other, it is still unclear whether these effects have a common origin or we are detecting the entangled manifestation of multiple physical mechanisms (see e.g. Dalal et al. 2008; Han et al. 2019; Johnson et al. 2019; Ramakrishnan et al. 2019; Sato-Polito et al. 2019; Mansfield & Kravtsov 2020 for more discussion).

5.2 The correlations between halo and galaxy properties

We analyse now the correlation between the properties of central galaxies and those of their hosting haloes. In Fig. 4, central galaxy properties are displayed as a function of halo mass. Galaxies are also colour coded according to halo formation time, $z_{1/2}$, which means that the halo assembly bias effect (i.e. the secondary bias on halo age) is completely characterized. In the subplots, the relative differences in the median $z_{1/2}$ in the top and bottom 50 per cent subsets of the corresponding galaxy property are shown for reference, i.e. $\mathrm{diff}(< z_{1/2} >) = \left(< z^{(2)}_{1/2} > - < z^{(1)}_{1/2} > \right) / < z^{(1)}_{1/2} >$ (where 1 and 2 represent the bottom and top subsets, respectively). We focus here on stellar mass, velocity dispersion (σ), (g–i) colour, SFR, sSFR, subhalo half-mass radius $R^{(\mathrm{sh})}_{1/2}$, galaxy half-mass radius $R^{(*)}_{1/2}$, and surface density (see Section 3 for details).

The following conclusions are drawn from this exercise:

(i) As expected, more massive central galaxies typically inhabit more massive haloes. In addition, older haloes at fixed halo mass typically harbour more massive galaxies. Above $\sim 10^{12}$–$10^{12.5}$ $h^{-1}\,\mathrm{M}_{\odot}$, the SHMR is flatter and exhibits less scatter. Also, the dependence on halo formation redshift at fixed halo mass appears to weaken (also see Matthee et al. 2017). Note that this effect is noticeable in the range of ages (colours) displayed as a function of halo mass, but it gets diluted in the fractional differences.

(ii) The total velocity dispersion σ (which takes into account all matter particles in the subhalo) is clearly the property that more tightly correlates with halo mass across the entire mass range. This is a reflection of DM being the dominant component. At fixed halo mass below $M_{\mathrm{vir}} \lesssim 10^{13}\ h^{-1}\,\mathrm{M}_{\odot}$, galaxies with higher σ tend to reside in older haloes.

(iii) In lower mass haloes (below the characteristic halo mass mentioned above), central galaxies exhibit larger scatter in colours, whereas higher mass haloes are only inhabited by very red galaxies. At fixed halo mass, older haloes tend to host redder galaxies, but, again, the correlation seems stronger at the low-mass end.

(iv) Below the characteristic halo mass, there is a clear correlation (anticorrelation) between SFR (sSFR) and halo mass. Above this halo mass, the correlation appears weaker. The large scatter in the SFR for haloes above $\sim 10^{12.5} h^{-1}\,\mathrm{M}_{\odot}$ might be explained by the effect of active galactic nuclei (AGN) feedback, which prevents star formation in massive systems (this is also noticeable in the galaxy colours). The dependence on halo age at fixed halo mass is slightly more noticeable for sSFR than for SFR.

(v) There is a tight correlation between the size of the baryonic + DM component of central galaxies (i.e. subhalo size) and the mass of the halo, with larger systems living in more massive haloes. At fixed halo mass, larger systems tend to prefer younger haloes. The strong correlation is again expected since subhalo size is mostly determined by the DM component and not by the baryonic component.

(vi) When only the stellar component is considered (what is mostly measured in observations, i.e. $R^{(*)}_{1/2}$), the behaviour changes significantly. Galaxy size shows little dependence on halo mass for galaxies in haloes of $M_{\mathrm{vir}} \sim 10^{10.5}$–$10^{12.5}\ h^{-1}\,\mathrm{M}_{\odot}$, while it increases with halo mass above $10^{12.5}\ h^{-1}\,\mathrm{M}_{\odot}$. These results are in agreement with those presented by Genel et al. (2018). We also note the slight upturn for haloes with $M_{\mathrm{vir}} \lesssim 10^{11}\ h^{-1}\,\mathrm{M}_{\odot}$. Genel et al. (2018), who also report this effect, speculate that the size difference might be environment related, since some of the central galaxies in this halo-mass range could be 'splash-back galaxies.[5] Although the analysis by Genel et al. (2018) is performed using the higher resolution IllustrisTNG100 box, resolution effects must not be discarded just yet. At fixed halo mass, there is almost no dependence on halo formation time.

(vii) The surface (stellar) mass density of galaxies increases with halo mass below the characteristic halo mass range, but it remains fairly constant above. The dependence on halo formation time shows up again due to the inclusion of stellar mass.

The different panels of Fig. 4 illustrate the importance of the range of halo masses around $\sim 10^{12}$–$10^{12.5}\ h^{-1}\,\mathrm{M}_{\odot}$. Above this mass range, the efficiency in the production of stellar mass decreases (see the SHMR in the first panel of Fig. 4, and, e.g. Rodríguez-Puebla et al. 2015 for comparison), which transmits to many other halo–galaxy relations. This loss of efficiency is directly related to AGN feedback, which in these haloes becomes dominant with respect to other less-effective feedback mechanisms such as stellar feedback (see e.g. McNamara & Nulsen 2007; Matthee et al. 2017; Kravtsov, Vikhlinin & Meshcheryakov 2018; Pillepich et al. 2018; Donnari et al. 2019; Terrazas et al. 2020). Note that 'halo quenching', a process in which gas inflows are shock heated to virial temperature preventing the accreted gas from fueling star formation, might also play a role in massive haloes (see e.g. Zu & Mandelbaum 2016).

The size of the stellar component is also notably affected by the different feedback mechanisms (see Pillepich et al. 2018). While galaxy size in haloes below $\sim 10^{12}$–$10^{12.5}\ h^{-1}\,\mathrm{M}_{\odot}$ is driven by stellar winds, galaxies in massive haloes are mostly affected by the much more efficient AGN feedback. This process acts by sweeping large amounts of baryonic matter towards the outer regions of the galaxy, thus expanding their size and increasing their half-mass radius. This mechanism explains the steep slope in galaxy size–halo mass relation for massive haloes. As expected, the relations for subhalo size and total velocity dispersion are not affected by these processes, since these quantities are mostly determined by the DM component.

Another important question is which galaxy properties better trace halo mass. Among the stellar population properties considered, stellar mass emerges as the most efficient tracer, although the scatter is large at the low halo mass end. When the DM component is included, both size and, especially, the total velocity dispersion are the properties that exhibit less scatter. Note that several studies point to the *stellar* velocity dispersion as the best tracer of halo mass (e.g. Zahid et al. 2016, 2019). This motivates the use of the

[5]Galaxies identified as 'centrals' at $z = 0$ that were satellites in the recent past, since they passed through the virial radius of a distinct halo. Splashback galaxies/haloes are known to suffer strong tidal disruption during these encounters.

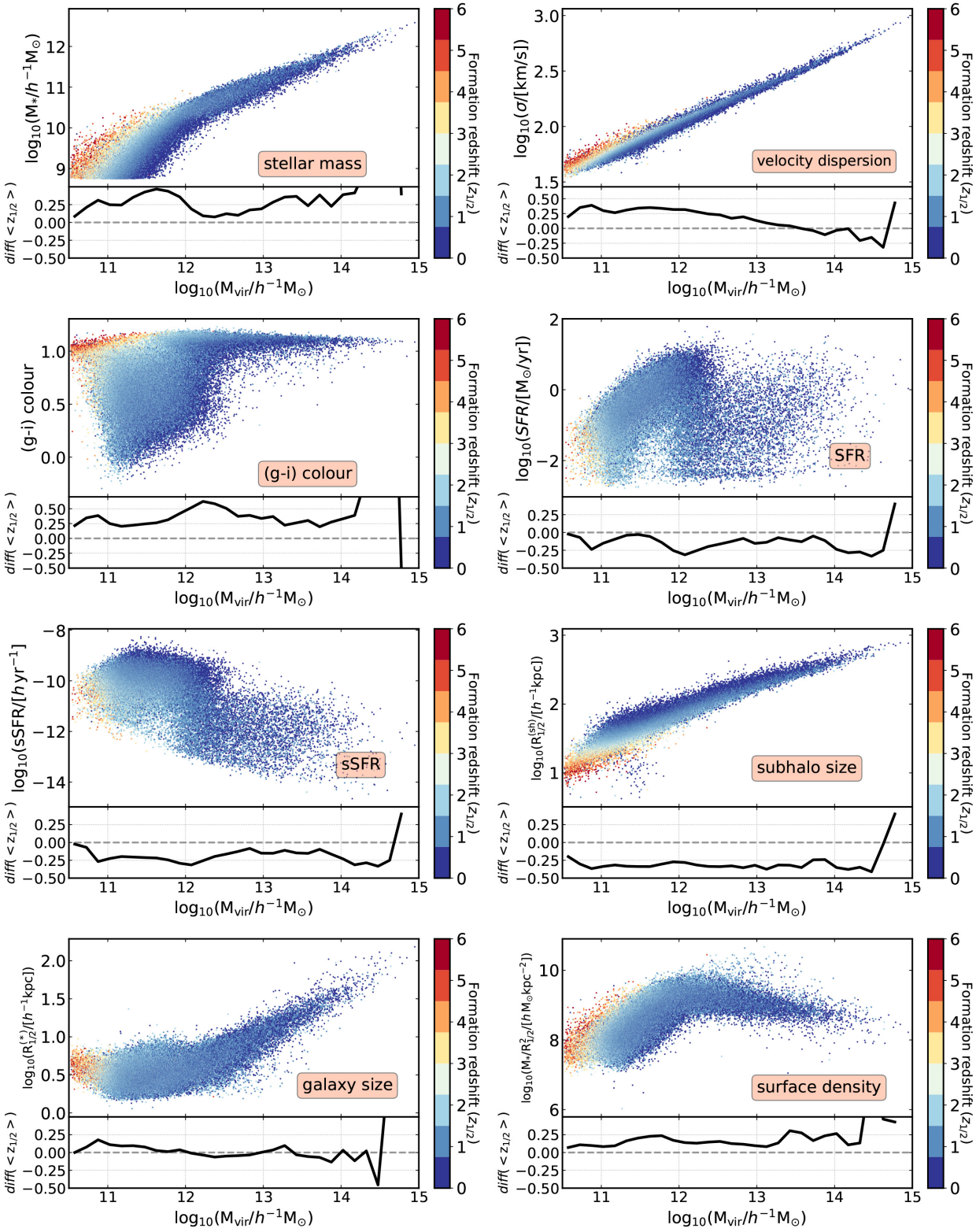

Figure 4. The distributions of central galaxy properties as a function of halo mass. The colour code indicates the age of the halo where each galaxy lives, which is described by the halo formation redshift, $z_{1/2}$. In the subplots, the relative difference between the median $z_{1/2}$ in 50 per cent subsets of the corresponding galaxy property is shown for reference, i.e. $\mathrm{diff}(< z_{1/2} >) = \left(< z^{(2)}_{1/2} > - < z^{(1)}_{1/2} > \right) / < z^{(1)}_{1/2} >$ (where 1 and 2 represent the bottom and top subsets, respectively).

velocity dispersion function (VDF, e.g. Montero-Dorta, Bolton & Shu 2017a) in the context of halo–galaxy connection models.

It is also noteworthy that halo age typically varies diagonally across the plane formed by a given galaxy property and halo mass. Haloes of the same age but higher mass tend to harbour typically more massive and redder galaxies. In other words, there is no particular galaxy property that completely determines halo formation redshift. On the other hand, the galaxy size, $R_{1/2}^{(*)}$, separates from the rest of the properties, as it is the only one that exhibits no correlation with halo age at fixed virial mass.

5.3 Galaxy clustering at fixed halo mass

The correlations and scatters between halo and galaxy properties along with the distribution of biases dictate how the secondary bias signal transmits to the central galaxy population when different galaxy selections are applied. In Fig. 5, we select galaxies according to the same properties discussed in the previous sections, and measure relative galaxy bias at fixed halo mass as described in Section 4. For reference, Fig. 5 displays the secondary halo bias on concentration and age in the background (i.e. halo assembly bias) which are the halo secondary properties more likely to be connected to the galaxy properties used here (the possible manifestation of the spin bias signal will be addressed in Section 5.4).

From the very first glimpse at Fig. 5, it becomes evident that the clustering of galaxies at fixed halo mass depends on multiple galaxy properties. Statistically significant signals are detected for all properties (although the measurement is very noisy for SFR). The comparison with the halo assembly bias signal on the background and the halo–galaxy relations of Fig. 4 clearly indicates that the galaxy clustering difference is, in most cases, due to the mapping between galaxy and halo properties, at fixed halo mass. These measurements thus represent a prediction of galaxy assembly bias, *according to the definition adopted in this analysis*. This galaxy assembly bias signal was already presented, albeit with less statistical significance, in Xu & Zheng (2020), who used the 75 h^{-1} Mpc Illustris-2 box. As mentioned before, no conclusive, irrefutable evidence has been found with real data. Among the main issues that beset this measurement are: the large uncertainties in the determination of halo masses, the contamination from satellites, and even the uncertainties in stellar population synthesis (see Miyatake et al. 2015; Lin et al. 2016; Sunayama et al. 2016; Montero-Dorta et al. 2017b; Niemiec et al. 2018 for more discussion).

Among the different galaxy properties featured in Fig. 5, the stellar mass, (g–i) colour, sSFR, and subhalo size are the ones that more clearly follow the halo assembly bias signal (note that subhalo size is considered a galaxy property here for simplicity, but it is actually determined by the DM component). Weaker/noisier results are found for velocity dispersion, surface density, and especially SFR, although the dependence at fixed halo mass is still compatible with the background signal.

The galaxy size emerges in Fig. 5 as a special case. While smaller subhaloes are more tightly clustered than larger subhaloes, this trend inverts when only the stellar component (galaxy size) is considered. Larger central galaxies are significantly more clustered than their smaller counterparts up to, at least, $\log_{10}(M_{\rm vir}/h^{-1}\,{\rm M}_{\odot}) \sim 12.5$. Interestingly, Fig. 4 shows little to no correlation between galaxy size and halo age at fixed halo mass. This could indicate that the difference in clustering displayed in Fig. 5 is not due to halo assembly bias.

In order to illuminate the secondary dependencies shown for galaxy clustering in Fig. 5, we have analysed the overlap between galaxy and halo subsets. In essence, the fraction of haloes that are selected using central galaxy properties that are also selected using secondary halo properties, at fixed mass, characterizes the transmission of the signal. Note that we measure the overlap between corresponding subsets (i.e. high 50 per cent galaxy subset with high 50 per cent halo subset, low 50 per cent galaxy subset with low 50 per cent halo subset). These results are summarized in Table 1, where the average subset overlap for different galaxy properties is listed with respect to subsets selected by halo formation redshift and spin. Note that the average value reported in Table 1 is the mean value between corresponding subsets averaged over the halo mass range where the assembly bias signal is statistically more significant ($11 < \log_{10}(M_{\rm vir}/h^{-1}\,{\rm M}_{\odot}) < 12.5$ for all properties except for stellar spin – see caption). Note that a value around 50 per cent indicates that the galaxy selection is not biased towards any particular type of haloes, which implies little transmission of the signal. A value $>$ 50 per cent indicates some correlation (i.e. for stellar mass), whereas a value $<$ 50 per cent means anticorrelation (e.g. for SFR/sSFR); in both cases some secondary halo dependency is traced.

Several consequences can be extracted from the above quantification analysis. First, a relatively small overlap can seemingly produce a significant secondary galaxy bias signal. Secondly, the dependence on formation time (i.e. halo assembly bias) is, as expected, generally stronger than that on halo spin. Finally, a non-negligible overlap with respect to halo spin is detected for galaxy size, which might be, as anticipated previously, the cause of the signal in Fig. 5. Recall, also, that both halo age and halo spin are intrinsically and strongly correlated (older haloes have lower spin, see e.g. fig. 3 from Johnson et al. 2019) so it is not trivial to disentangle these effects in the galaxy population.

It seems clear, nevertheless, that the scatter in the halo–galaxy connection does not wash out the assembly bias signal (as long as halo mass is known with precision). Another interesting question that Fig. 5 poses is whether the signal measured for several properties at fixed halo mass also depends on stellar mass. As shown in the upper left panel of Fig. 5, more massive galaxies are more tightly clustered than less massive galaxies in haloes of similar mass. In order to investigate the stellar mass effect on galaxy assembly bias, we take advantage of the statistics provided by IllustrisTNG300 to additionally split the population in each halo mass bin by stellar mass. Fig. 6 displays the galaxy assembly bias signal for velocity dispersion, colour, sSFR, half-mass radius (using both definitions), and surface density for the 50 per cent higher stellar mass subpopulation (orange/cyan colours) and the 50 per cent lower stellar mass subpopulation (red/blue colours).

Fig. 6 reveals that the implicit clustering dependence on stellar mass has different effects depending on the galaxy property used to perform the subset selection. Although results are unavoidably noisier than those presented in Fig. 5, the following conclusions can be drawn, qualitatively, from this analysis:

(i) The galaxy assembly bias signal for properties such as colour, sSFR, and subhalo size displays a clear dependence on stellar mass: the signal is always stronger (i.e. larger clustering difference between S-subsets) for higher mass galaxies than for lower mass galaxies. Relative to the magnitude of the signal itself, the dependence is, as expected, weaker for subhalo size, which is very connected to the DM content of the halo.

(ii) The effect of stellar mass seems insignificant for galaxy size: higher and lower mass galaxies exhibit statistically similar assembly bias signals.

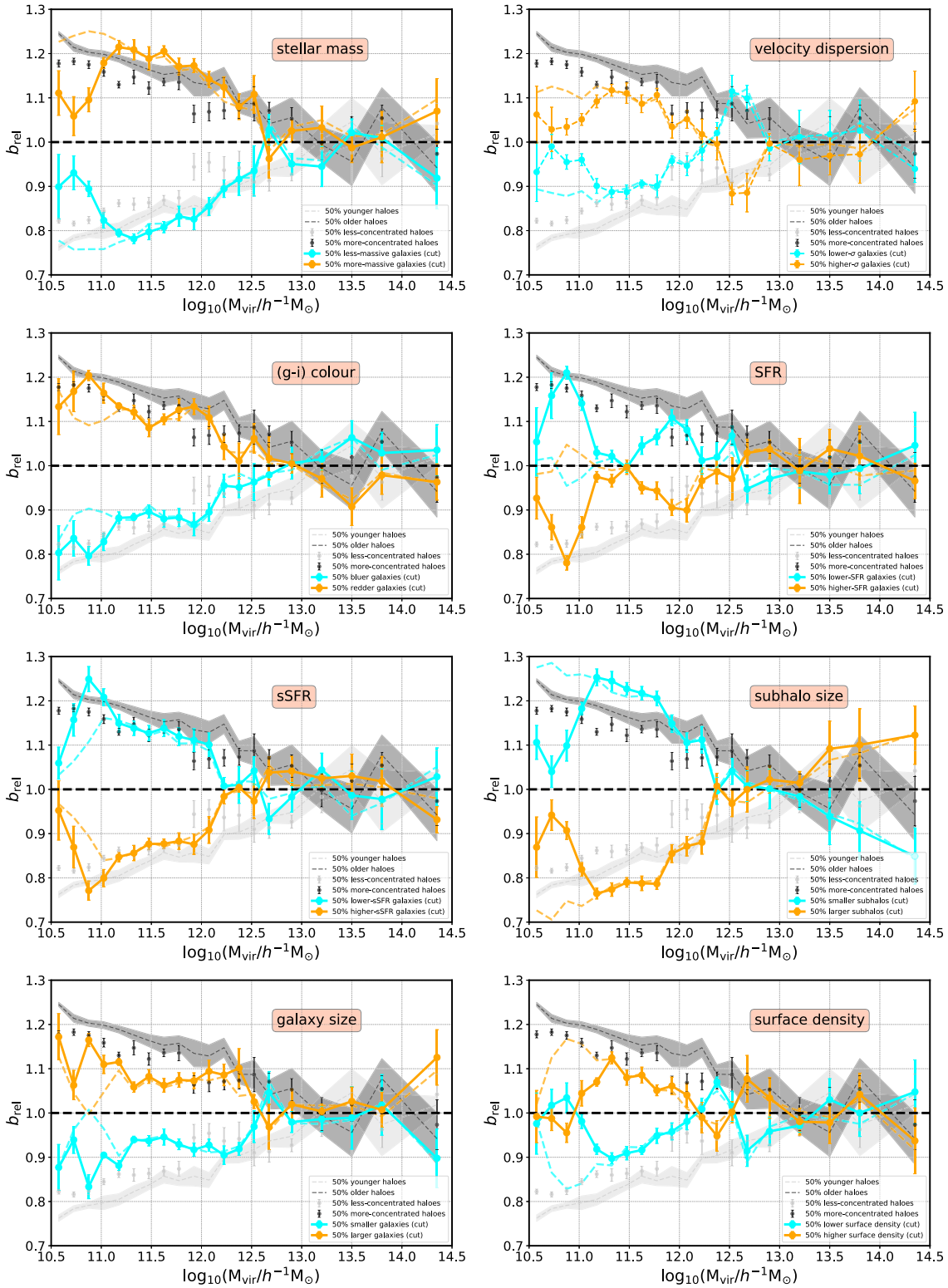

Figure 5. The dependence of galaxy bias on several galaxy properties as a function of halo mass. In each panel, orange dots represent the 50 per cent subset of galaxies with higher values of the corresponding galaxy property, whereas cyan dots show the remaining 50 per cent lower value subset. These results are obtained assuming a stellar-mass cut of $\log_{10}(M_*/h^{-1}\,\mathrm{M}_\odot) > 8.75$, which implies a minimum subhalo-mass resolution of 50 particles. Orange/cyan dashed lines show the same measurements assuming no stellar-mass cut. For reference, the secondary halo bias results for concentration (black/grey dots with error bars) and formation redshift (black/grey dashed lines with error bands) are shown in the background. Errors bars/bands show the jackknife uncertainties obtained from a set of 8 equal-volume subboxes.

Table 1. Average subset overlap (in percentage) within the entire mass range considered between central galaxy properties and either halo formation redshift (second column) or halo spin (third column). For the first block of galaxy properties, the measurement is performed within the halo mass range $11 < \log_{10}(M_{\rm vir}/h^{-1}\,{\rm M}_\odot) < 12.5$, whereas the stellar and galaxy spin measurements are performed for haloes with masses $11.5 < \log_{10}(M_{\rm vir}/h^{-1}\,{\rm M}_\odot) < 13$. See text for more details.

Central galaxy property	Halo $z_{1/2}$	Halo spin
Stellar mass	63.4 ± 6.4	42.6 ± 4.7
Colour	66.1 ± 4.9	42.5 ± 7.2
SFR	43.4 ± 6.8	56.0 ± 7.7
sSFR	38.6 ± 3.9	59.8 ± 5.0
Velocity dispersion	64.6 ± 1.6	45.2 ± 3.2
Subhalo size	32.7 ± 2.2	54.6 ± 0.6
Galaxy size	52.9 ± 3.8	58.7 ± 2.4
Surface density	58.1 ± 2.1	39.3 ± 4.3
Stellar spin	49.2 ± 3.9	55.1 ± 2.6
Galaxy spin	52.16 ± 2.3	57.2 ± 2.4

(iii) Surface density shows an interesting inversion of the signal for lower mass galaxies (this is also found at low statistical significance for velocity dispersion). For lower mass galaxies, the relative bias is higher for objects with low surface density, with the opposite trend happening for higher mass galaxies.

We have checked that the above dependencies reflect, to a large extent, the correlation between stellar mass and formation redshift at fixed halo mass (see Fig. 4). When the same analysis is performed on the basis of halo mass and $z_{1/2}$, similar trends are obtained. We regard Fig. 6 as a first glimpse into these trends that go beyond the secondary dependencies, but better statistics are needed to characterize them with precision.

Fig. 5 and 6 show some particular cases where the galaxy assembly bias signal exceeds the halo assembly bias signal in some mass ranges (i.e. for stellar mass and subhalo size). Note that this does not necessarily imply any additional signal that only depends on galaxies (i.e. on the galaxy formation process). In a given halo mass bin, by choosing half the population by stellar mass, we could be essentially selecting slightly older/younger haloes than what a halo-age-based selection would produce (i.e. it depends on the function that maps stellar mass and age, at fixed halo mass). It is also possible that by selecting on the basis of these galaxy properties we are effectively mapping a combination of multiple secondary halo bias effects (e.g. age, concentration, spin, etc.).

5.4 The effect of spin bias

The dependence of galaxy clustering on secondary halo properties has mostly been probed with real data using galaxy properties that are expected to be related to the assembly history of haloes (i.e. colour or SFH, the so-called galaxy assembly bias effect). However, the large magnitude of the secondary dependence measured from *N*-body simulations on other halo properties such as spin (see the recent measurements of Sato-Polito et al. 2019 and Johnson et al. 2019) suggests alternative routes for the manifestation of the secondary halo bias effect on the galaxy population. In this section, we discuss the potential existence of a *galaxy spin bias* effect and its detectability with real data.

The first aspect to investigate is the transmission of the angular momentum or spin of the halo to the central galaxy. Galaxies form when baryons collapse towards the centres of rotating haloes (and subsequently cool and condensate), so it is conceivable that their angular momenta carry some information about the angular momenta of the hosting haloes.[6] This initial correlation between the angular momentum of baryons and DM is, however, threatened by several physical processes that can take place during the lifetime of galaxies, including 'wet compaction' and mergers (see e.g. Stevens et al. 2017; Jiang et al. 2019).

Fig. 7 displays, in the upper panel, the relation between the total spin of central galaxies ($\lambda_{\rm galaxy}$, which takes into account DM, gas, and stars) and the spin of their hosting haloes, for the youngest and oldest haloes (subsets of 20 per cent of the entire population). Fig. 7 shows that, as expected, the halo–galaxy correlation is conserved, to some extent, when DM is taken into account: faster-rotating haloes clearly tend to harbour faster-rotating central galaxies, but the scatter in this correlation is significant (see Table 1, the average overlap with respect to the halo-spin subsets is 57.2 ± 2.4). The separation of the hosting halo population by age indicates that: (1) younger haloes typically rotate faster, (2) the scatter in the $\lambda_{\rm galaxy}$–$\lambda_{\rm halo}$ relation is similar for older and younger haloes, and (3) the $\lambda_{\rm galaxy}$–$\lambda_{\rm halo}$ relation approximates the 1:1 relation, albeit with large scatter, for younger haloes. Note that the third point might be a consequence of older haloes having more time to undergo merging processes that can weaken the correlation. The same mechanism could also explain the occupancy variations of haloes measured in Illustris and IllustrisTNG. At fixed halo mass, younger haloes tend to host a larger number of satellites than older haloes, which, again, might be due to the effect of mergers (Artale et al. 2018; Bose et al. 2019).

The fact that we find some correlation between halo spin and total galaxy spin is not surprising (since this is dominated by DM). In the lower panel of Fig. 7, we show how this correlation appears to vanish when only the stellar component of the galaxy is taken into account (i.e. the stellar spin, λ_*). This result is in agreement with recent findings from Jiang et al. (2019), who reported no correlation between these quantities using the NIHAO and VELA zoom-in hydrocosmological simulations.[7] These simulations, however, only contain 13 and 34 moderate-to-high-mass central galaxies across a wide redshift range, respectively. Our results thus add a strong statistical sense to their claim.

Although the correlation shown in the lower panel of Fig. 7 is very weak, in Fig. 8 we show that it is enough to produce a small separation in $\lambda_{\rm halo}$ when subsets are selected in λ_*. Fig. 8 employs the same format as Fig. 4 to show λ_* versus $M_{\rm vir}$, but it plots $\lambda_{\rm halo}$ as a secondary halo dependency. The subplot displays a non-negligible positive $\Delta\log_{10}(\lambda_{\rm halo})$ between subsets that opens the door for a transmission of the halo spin bias effect to the galaxy population. This small correlation shows up, although mildly, in the subset-overlap analysis of Table 1. The overlap between stellar-spin and halo-spin subsets is 55.1 ± 2.6, within the halo mass range $11.5 < \log_{10}(M_{\rm vir}/h^{-1}\,{\rm M}_\odot) < 13$.

In Fig. 9, we show the relative bias for high- and low-λ_* galaxies as a function of halo mass, in the same format as Fig. 5. In the background, we show the spin bias effect on haloes, which in IllustrisTNG300 is characterized by a progressive increase of signal as a function of halo mass. Fig. 9 presents a clear measurement of

[6] In SAMs, in fact, it is common to assume that $\lambda_{\rm galaxy} \propto \lambda_{\rm halo}$ (see e.g. Somerville et al. 2008; Guo et al. 2011; Benson 2012).

[7] Jiang et al. (2019) actually assume a radius of 0.1 $R_{\rm vir}$ in the computation of the galaxy stellar spin, under the assumption that this radius typically corresponds to the stellar half-mass radius. We have checked that our conclusions remain unaltered when this radius is adopted.

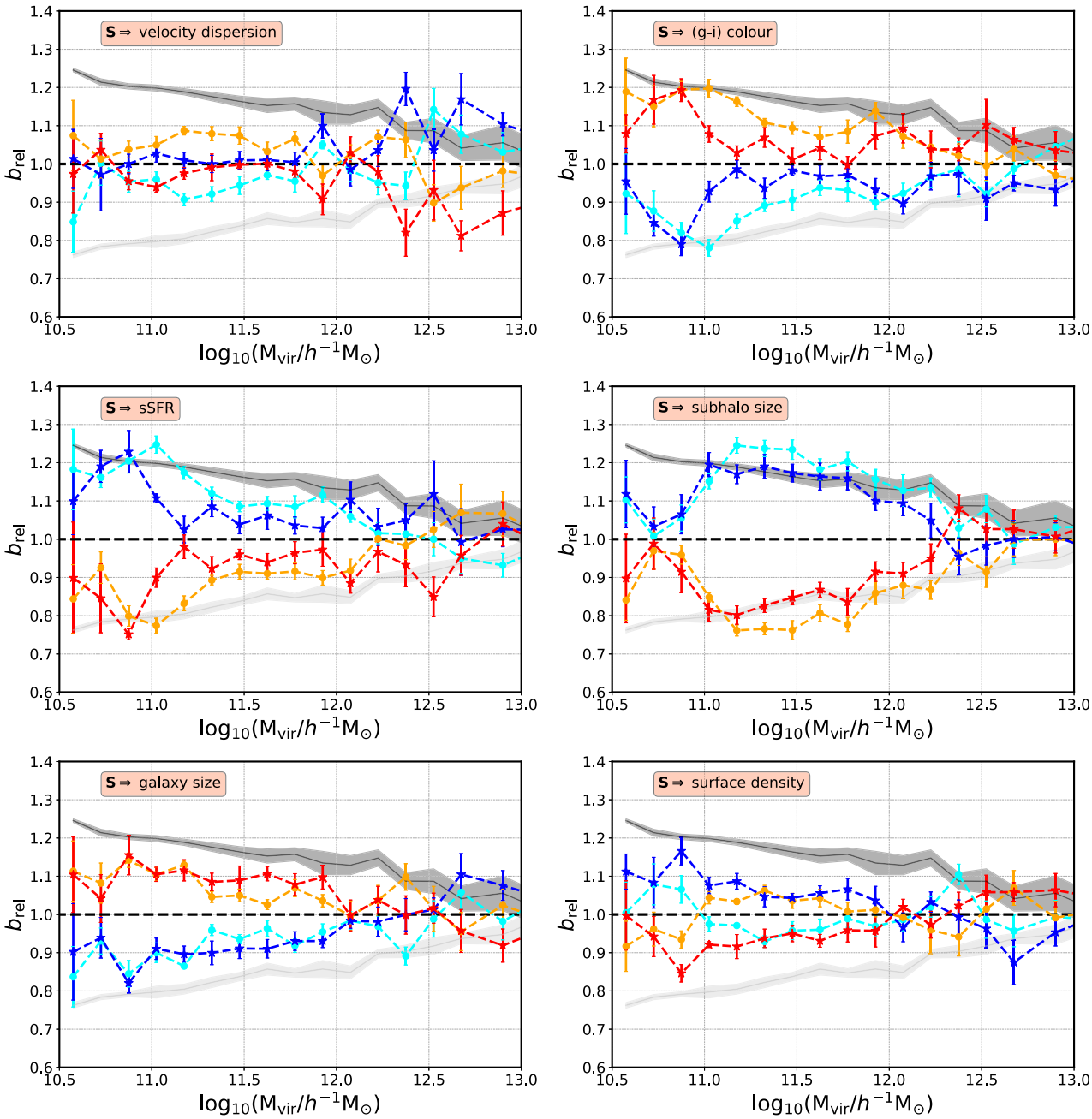

Figure 6. The dependence of galaxy bias on several galaxy properties as a function of halo mass and stellar mass. Orange/cyan symbols show the secondary galaxy bias signal for high-mass galaxies (50 per cent of the distribution), whereas red/blue symbols show results for lower mass galaxies. The secondary halo bias effect on halo age is kept in the background for reference (black/grey solid lines). The error bars/bands represent the jackknife uncertainties obtained from a set of 8 divisions.

galaxy spin bias that overall follows the halo spin bias signal. Note that, despite the similarities with the halo spin bias trend, a better understanding of the origins of secondary halo bias is necessary to determine the contribution of halo spin, concentration, and age to the galaxy spin bias signal. All these different properties are strongly correlated with each other, and it is still unclear whether the different secondary bias trends originates from the same effect or not (see e.g. Johnson et al. 2019; Ramakrishnan et al. 2019; Sato-Polito et al. 2019; Mansfield & Kravtsov 2020 for discussion).

Only at the very low mass end does the galaxy spin bias signal deviate from the halo spin bias signal in IllustrisTNG300. An inversion in galaxy spin bias is observed at $\log_{10}(M_{vir}/h^{-1}\,M_\odot) \sim 11.5$, in agreement, interestingly, with previous halo spin bias results (Johnson et al. 2019; Sato-Polito et al. 2019). Note that, apart from the relation between λ_* and λ_{halo}, by selecting only haloes that contain a central galaxy, we are effectively removing unoccupied haloes and haloes that harbour satellite galaxies (along with haloes whose central galaxies fall below our resolution limit). Further investigation is required to understand this feature in IllustrisTNG300. The physical origins of the halo spin bias inversion at the low-mass end will be addressed in an upcoming paper (Tucci et al., in preparation).

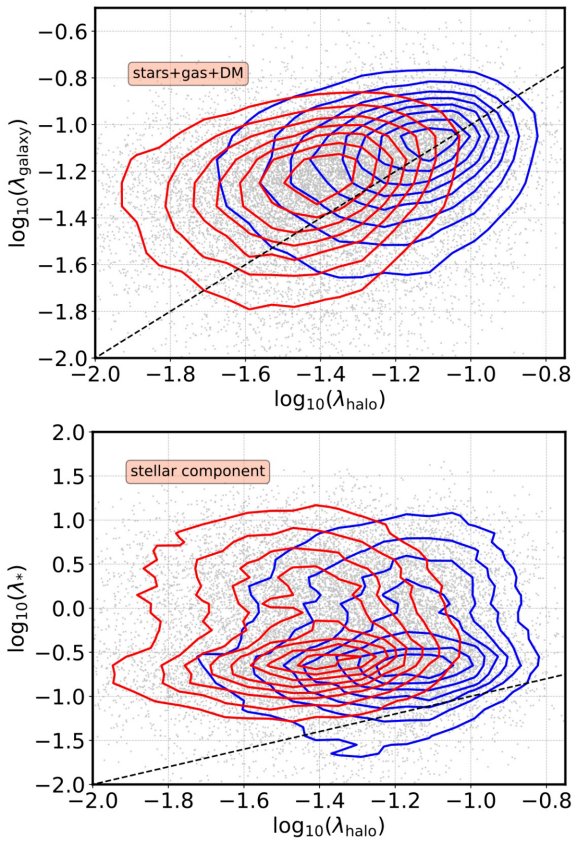

Figure 7. The correlation between halo spin and galaxy spin for central galaxies. In the upper panel, the total galaxy spin, λ_{galaxy}, is computed with all DM, gas, and stellar-mass particles. In the lower panel, only the stellar component is considered (λ_*). In both panels, red contours show the distribution for the oldest haloes, whereas blue contours represent the youngest population (20 per cent subsets).

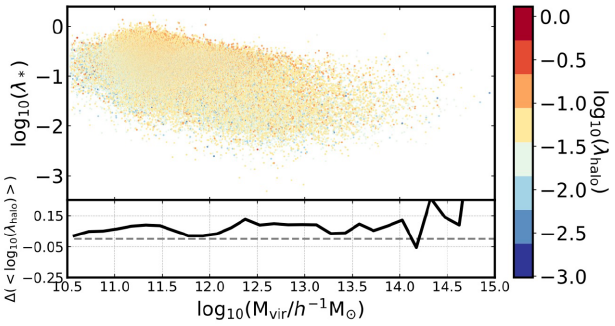

Figure 8. The distribution of the stellar spin, λ_*, as a function of halo mass. The colour code shows the spin of the hosting halo, λ_{halo}. In the subplot, the difference between the median $\log_{10}(\lambda_{halo})$ in 50 per cent subsets of λ_* is shown for reference, i.e. $<\log_{10}(\lambda_{halo})^{(2)}> - <\log_{10}(\lambda_{halo})^{(1)}>$ (where 1 and 2 represent the bottom and top subsets, respectively).

6 DISCUSSION AND CONCLUSIONS

In this paper, we use the IllustrisTNG300 simulation box at redshift $z = 0$ to evaluate the transmission of the secondary halo bias signal to the central galaxy population. The IllustrisTNG simulations incorporate important improvements as compared to the previous Illustris boxes, such as a new implementation of galactic winds, black hole-driven kinetic feedback at low accretion rates, and

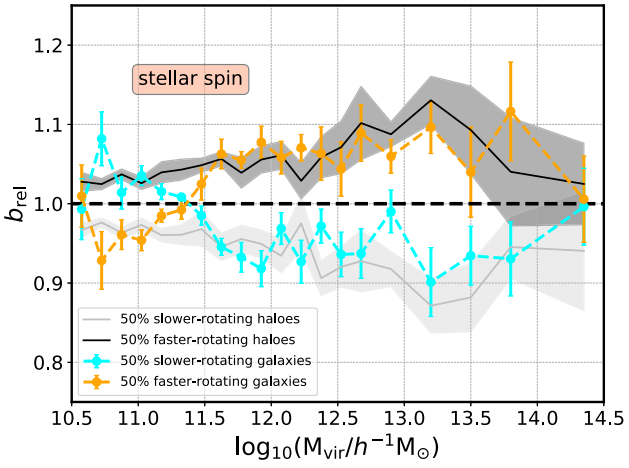

Figure 9. The dependence of galaxy bias on the galaxy stellar spin (λ_*) as a function of halo mass in the same format of Fig. 5. Orange symbols represent the 50 per cent subset of galaxies with higher values of spin, whereas cyan symbols show the remaining 50 per cent lower spin subset. For reference, the secondary halo spin bias effect is shown in the background. The error bars represent the jackknife uncertainties obtained from a set of 8 divisions.

the inclusion of magnetohydrodynamics (Weinberger et al. 2017; Pillepich et al. 2018). Importantly, the IllustrisTNG300 box that we use here spans 205 h^{-1} Mpc on a side with a mass resolution below $10^8 h^{-1}\,M_\odot$ for both baryons and DM particles, which increases considerably the statistical significance of our analysis as compared to previous measurements (e.g. Xu & Zheng 2020).

On the halo side, our analysis of IllustrisTNG300 yields an interesting result regarding the secondary dependence of halo bias on halo spin. Contrary to recent findings using several *N*-body simulations (Johnson et al. 2019; Sato-Polito et al. 2019), we detect no inversion of the signal at the low-mass end (i.e. faster-rotating haloes are in IllustrisTNG300 more tightly clustered than slower rotators across the entire mass range). This discrepancy is not likely to be caused by resolution effects, since the resolution of IllustrisTNG300 for DM particles is comparable to that of the Small MultiDark and Vishnu boxes where the 'crossover' was previously found. Given the range of halo masses where the discrepancy occurs, it is possible, however, that this could be due to differences in the way substructure and/or splashback haloes (i.e. distinct haloes at $z = 0$ that were recently subhaloes because they passed by a bigger halo) are identified in the simulations. The physical origins of spin bias at the low-mass end and the connection with other secondary bias trends will be addressed in an upcoming paper (Tucci et al., in preparation).

In this work, we adopt a simple definition of *galaxy assembly bias* that directly reflects the halo assembly bias effect: the dependence of the clustering of central galaxies on the formation history (or concentration) of hosting haloes at fixed halo mass. We have shown that this effect exists in various degrees when the central galaxy population is split by stellar mass, total (including DM) velocity dispersion, colour, sSFR, subhalo size, and surface density, over a wide halo mass range. We have also analysed the dependence of the signal on stellar mass at fixed M_{vir} (to the extent that the IllustrisTNG300 statistics permit). This implicit dependency produces different effects depending on the galaxy property used to split the galaxy population, which overall reflect the correlation between stellar mass and halo formation redshift, at fixed halo mass. The galaxy assembly bias signal is stronger in higher mass galaxies

for colour, sSFR, and subhalo size, and fairly independent of stellar mass for galaxy size. The surface density, on the other hand, exhibits an inversion of the galaxy assembly bias signal.

Among the different galaxy properties considered, galaxy size emerges as a special case. At fixed halo mass, larger galaxies are more tightly clustered than smaller galaxies, but this effect, in contrast to the rest of the dependencies, seems to be uncorrelated with halo formation time. Interestingly, we do find some correlation, although small, with halo spin, which is a property that is more closely related to the dynamics of the haloes. Our results (even at 'fixed' stellar mass) seem also in tension with $z = 0$ measurements from Hearin et al. (2019) indicating that small SDSS galaxies cluster much more strongly than large galaxies of the same stellar mass. It is noteworthy, however, that even though the difference observed between the two SDSS populations is strong on small scales, it seems to vanish at larger distances (i.e. $\gtrsim$10 h^{-1} Mpc). Note also that several other authors have found weak or no dependence of the mass–size relation on environment (see e.g. Huertas-Company et al. 2013). Further investigation is required to understand these comparisons and the origin of the size dependence. Interestingly, we do find the inverse trend (smaller objects being more biased) when the subhalo size is used to split the population.

Our results are complementary to those presented in Bose et al. (2019), who also use lllustrisTNG300. The authors study the dependence of the galaxy content of haloes (i.e. *halo occupancy variations*) on several secondary halo properties at fixed halo mass. Their results indicate a strong dependence of the average number of satellites on formation time, concentration, and environment, at fixed halo mass. These occupancy variations of satellites imply the existence of galaxy assembly bias on the 1-halo term, whereas our study on central galaxies reflects the 2-halo term component of the effect.

Our findings also align well with the galactic and halo conformity signal found in the Illustris simulation. Bray et al. (2016) show the existence of a strong correlation in the colours of galaxies residing in neighbouring haloes (between 3 and 10 h^{-1} Mpc) at either fixed stellar or halo mass (i.e. the so-called 2-halo conformity). This environmental dependence is connected to the galaxy assembly bias signal shown in our Fig. 5, where redder galaxies are more tightly clustered than bluer galaxies at fixed halo mass, due to the different formation histories of their hosting haloes.

We have also addressed for the first time the transmission of the halo spin bias signal to the central galaxy population, an effect that we call *galaxy spin bias*. We have shown that, even though central galaxy spin (the baryonic component) seemingly retain little (but not null) information about the total halo spin, the correlation is enough to produce a significant galaxy spin bias signal. Note that the existence of galaxy spin bias does not imply that the effect comes exclusively from halo spin; a better understanding of the physical origins of secondary halo bias is necessary to disentangle the different contributions (age, spin, and concentration) to the signal (see e.g. Sato-Polito et al. 2019; Johnson et al. 2019; Ramakrishnan et al. 2019; Mansfield & Kravtsov 2020 for more discussion). Ramakrishnan et al. (2019), for instance, claim that the halo assembly bias trend is a consequence of the multiscale connection between internal halo properties, the anisotropy of the tidal field and the large-scale environment. In this sense, by selecting different subsets of haloes (through halo properties or central-galaxy properties) we could be tracing in non-trivial ways more fundamental environmental effects.

Although IllustrisTNG300 predicts the existence of central galaxy spin bias (something that could be directly probed by mapping the velocities of stars and gas), the idea of an observational detection of the halo spin bias signal is even more appealing. The effect of halo spin bias is particularly large at the high-mass end, i.e. for group- and cluster-sized haloes. An observational detection of the halo spin bias signal would require a sizeable sample of well-identified groups/clusters with good measurements of their total masses (halo masses), along with a reliable proxy for halo spin. Ideally, one would also like the sample to expand a wide range of cluster masses for statistical reasons, even though the effect must be studied 'at fixed halo mass'. In this sense, smaller groups are particularly challenging, both in terms of their identification and their mass determination through dynamical or weak-lensing methods.

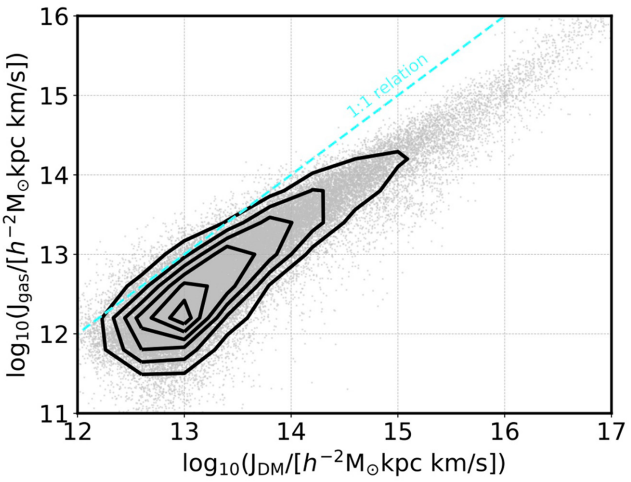

Figure 10. The correlation between the gas-only and the total angular momenta of haloes within the virial radius R_{vir}. Contours show the distribution of all haloes in IllustrisTNG300 with $M_{vir} > 10^{10.5}\ h^{-1}\ M_\odot$, whereas dots represent a randomly selected subset containing 10 per cent of this population. The dashed line indicates the one-to-one relation.

Fig. 10 displays the tight correlation that we find between the total angular momentum of the halo and the angular momentum of the gas that orbits inside it. This correlation suggests that the rotation of the intracluster gas can be used to determine halo spin. In fact, several methods have been proposed to measure this rotating intracluster gas. On the one hand, the neutral component of the gas can be observed through the 21-cm emission (i.e. the 21-cm hyperfine transition of neutral hydrogen), which will be accessible for millions of objects in the near future thanks to the Square Kilometer Array project (SKA[8]) and other similar radio surveys. Moreover, the rotation of the ionized component (i.e. the hot gas in the intracluster medium, ICM) can be measured through the kinetic Sunyaev–Zeldovich (kSZ) effect, i.e. the 'Comptonization' of the photons of the cosmic microwave background (CMB) as they propagate through galaxy clusters due to the motion of the cluster as a whole. A purely rotational motion of the ICM would produce a detectable dipole pattern, which again can be used as a proxy for halo spin (see Chluba & Mannheim 2002; Cooray & Chen 2002; Baldi et al. 2018).

Upcoming surveys such as *Euclid*,[9] the Dark Energy Spectroscopic Instrument (DESI[10]), the Javalambre Physics of the Accel-

[8]https://www.skatelescope.org
[9]https://www.euclid-ec.org
[10]https://www.desi.lbl.gov

erated Universe Astrophysical Survey (J-PAS[11]), the Prime Focus Spectroscopy survey (PFS[12]), the Large Synoptic Spectroscopic Telescope (LSST[13]), or SPHEREx,[14] to name but a few, will map the large-scale structure of the Universe (LSS) over huge cosmological volumes using a variety of galaxy populations as LSS tracers. The enhanced photometry and tracer density that these experiments will provide can improve the quality of galaxy measurements and the detection and characterization of clusters. More importantly, their detailed weak-lensing maps can crucially reduce the uncertainties in the determination of halo masses, one of the most delicate aspects involved in probing the assembly bias effect (see e.g. Niemiec et al. 2018).

ACKNOWLEDGEMENTS

The authors thank the anonymous referee for his/her useful comments.

ADMD and BT thanks the Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) for financial support. MCA acknowledges financial support from the Austrian National Science Foundation through Fonds zur Förderung der wissenschaftlichen Forschung (FWF) stand-alone grant P31154-N27. LRA thanks both FAPESP and the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) for financial support. NP acknowledges support from the Fondo Nacional de Desarrollo Científico y Tecnológico (Fondecyt) Regular 1191813 and from the Comisión Nacional de Investigación Científica y Tecnológica (CONICYT) project Basal AFB-170002. FR has been supported by Agencia Nacional de Promoción Científica y Tecnológica (PICT 2015-3098), the Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET, Argentina), and the Secretaría de Ciencia y Tecnología de la Universidad Nacional de Córdoba (SeCyT, UNC, Argentina).

We also thank S. Vitenti, M. Penna-Lima, and C. Doux for the use of their Numerical Cosmology package, *NumCosmo*.[15]

REFERENCES

Angulo R. E., Baugh C. M., Lacey C. G., 2008, MNRAS, 387, 921
Artale M. C., Zehavi I., Contreras S., Norberg P., 2018, MNRAS, 480, 3978
Baldi A. S., De Petris M., Sembolini F., Yepes G., Cui W., Lamagna L., 2018, MNRAS, 479, 4028
Behroozi P. S., Wechsler R. H., Wu H.-Y., 2013, ApJ, 762, 109
Benson A. J., 2012, New A, 17, 175
Bose S. et al., 2019, MNRAS, 2192
Bray A. D. et al., 2016, MNRAS, 455, 185
Bullock J. S., Dekel A., Kolatt T. S., Kravtsov A. V., Klypin A. A., Porciani J. R., Primack J. R., 2001, ApJ, 555, 240
Buser R., 1978, A&A, 62, 411
Chaves-Montero J., Angulo R. E., Schaye J., Schaller M., Crain R. A., Furlong M., Theuns T. M., 2016, MNRAS, 460, 3100
Chluba J., Mannheim K., 2002, A&A, 396, 419
Chue C. Y. R., Dalal N., White M., 2018, J. Cosmol. Astropart. Phys., 10, 012
Contreras S., Zehavi I., Padilla N., Baugh C. M., Jiménez E., Lacerna I., 2019, MNRAS, 484, 1133
Cooray A., Chen X., 2002, ApJ, 573, 43
Croton D. J. et al., 2006, MNRAS, 365, 11
Dalal N., White M., Bond J. R., Shirokov A., 2008, ApJ, 687, 12
Davis M., Efstathiou G., Frenk C. S., White S. D. M., 1985, ApJ, 292, 371
Dawson K. S. et al., 2013, AJ, 145, 10
Dolag K., Borgani S., Murante G., Springel V., 2009, MNRAS, 399, 497
Donnari M. et al., 2019, MNRAS, 485, 4817
Gao L., White S. D. M., 2007, MNRAS, 377, L5
Gao L., Springel V., White S. D. M., 2005, MNRAS, 363, L66
Genel S. et al., 2014, MNRAS, 445, 175
Genel S. et al., 2018, MNRAS, 474, 3976
Guo Q. et al., 2011, MNRAS, 413, 101
Hadzhiyska B., Bose S., Eisenstein D., Hernquist L., Spergel D. N., 2020, MNRAS, 493, 5506
Han J., Li Y., Jing Y., Nishimichi T., Wang W., Jiang C., 2019, MNRAS, 482, 1900
Hearin A. P., Watson D. F., van den Bosch F. C., 2015, MNRAS, 452, 1958
Hearin A., Behroozi P., Kravtsov A., Moster B., 2019, MNRAS, 489, 1805
Huertas-Company M., Shankar F., Mei S., Bernardi M., Aguerri J. A. L., Meert A., Vikram V., 2013, ApJ, 779, 29
Jiang F. et al., 2019, MNRAS, 488, 4801
Johnson J. W., Maller A. H., Berlind A. A., Sinha M., Holley-Bockelmann J. K., 2019, MNRAS, 486, 1156
Kaiser N., 1984, ApJ, 284, L9
Klypin A., Prada F., 2017, preprint (arXiv:e-print)
Kravtsov A. V., Vikhlinin A. A., Meshcheryakov A. V., 2018, Astron. Lett., 44, 8
Lacerna I., Padilla N., 2012, MNRAS, 426, L26
Lacerna I., Padilla N., Stasyszyn F., 2014, MNRAS, 443, 3107
Lacerna I., Contreras S., González R. E., Padilla N., Gonzalez-Perez V., 2018, MNRAS, 475, 1177
Lazeyras T., Musso M., Schmidt F., 2017, J. Cosmol. Astropart. Phys., 2017, 059
Li Y., Mo H. J., Gao L., 2008, MNRAS, 389, 1419
Lin Y.-T., Mandelbaum R., Huang Y.-H., Huang H.-J., Dalal N., Diemer B., Jian H.-Y., Kravtsov A., 2016, ApJ, 819, 119
Mansfield P., Kravtsov A. V., 2020, MNRAS, 493, 4763
Mao Y.-Y., Zentner A. R., Wechsler R. H., 2018, MNRAS, 474, 5143
Matthee J., Schaye J., Crain R. A., Schaller M., Bower R., Theuns T., 2017, MNRAS, 465, 2381
McNamara B. R., Nulsen P. E. J., 2007, ARA&A, 45, 117
Miyatake H. et al., 2015, ApJ, 806, 1
Miyatake H., More S., Takada M., Spergel D. N., Mandelbaum R., Rykoff E. S., Rozo E., 2016, Phys. Rev. Lett., 116, 041301
Montero-Dorta A. D., Bolton A. S., Shu Y., 2017a, MNRAS, 468, 47
Montero-Dorta A. D. et al., 2017b, ApJ, 848, L2
More S. et al., 2016, ApJ, 825, 39
Navarro J. F., Frenk C. S., White S. D. M., 1997, ApJ, 490, 493
Nelson D. et al., 2018, MNRAS, 475, 624
Niemiec A. et al., 2018, MNRAS, 477, L1
Padilla N., Contreras S., Zehavi I., Baugh C. M., Norberg P., 2019, MNRAS, 486, 582
Paranjape A., Hahn O., Sheth R. K., 2018, MNRAS, 476, 3631
Peebles P. J. E., 1969, ApJ, 155, 393
Pillepich A. et al., 2018, MNRAS, 473, 4077
Planck Collaboration XIII, 2016, A&A, 594, A13
Ramakrishnan S., Paranjape A., Hahn O., Sheth R. K., 2019, MNRAS, 489, 2977
Rodriguez-Gomez V. et al., 2015, MNRAS, 449, 49
Rodríguez-Puebla A., Avila-Reese V., Yang X., Foucaud S., Drory N., Jing Y. P., 2015, ApJ, 799, 130
Romano-Díaz E., Garaldi E., Borzyszkowski M., Porciani C., 2017, MNRAS, 469, 1809
Salcedo A. N., Maller A. H., Berlind A. A., Sinha M., McBride C. K., Behroozi P. S., Wechsler R. H., Weinberg D. H., 2018, MNRAS, 475, 4411

[11] http://www.j-pas.org
[12] https://pfs.ipmu.jp
[13] https://www.lsst.org
[14] http://spherex.caltech.edu
[15] https://numcosmo.github.io/

Sato-Polito G., Montero-Dorta A. D., Abramo L. R., Prada F., Klypin A., 2019, MNRAS, 487, 1570
Sheth R. K., Tormen G., 1999, MNRAS, 308, 119
Sheth R. K., Tormen G., 2004, MNRAS, 350, 1385
Sinha M., Garrison L., 2017, Corrfunc: Blazing Fast Correlation Functions on the CPU
Somerville R. S., Hopkins P. F., Cox T. J., Robertson B. E., Hernquist L., 2008, MNRAS, 391, 481
Springel V., 2010, MNRAS, 401, 791
Springel V., White S. D. M., Tormen G., Kauffmann G., 2001, MNRAS, 328, 726
Springel V. et al., 2018, MNRAS, 475, 676
Stevens A. R. H., Lagos C. d. P., Contreras S., Croton D. J., Padilla N. D., Schaller M., Schaye J., Theuns T., 2017, MNRAS, 467, 2066
Sunayama T., More S., 2019, MNRAS, 490, 4945
Sunayama T., Hearin A. P., Padmanabhan N., Leauthaud A., 2016, MNRAS, 458, 1510
Terrazas B. A. et al., 2020, MNRAS, 493, 1888
Vogelsberger M. et al., 2014a, MNRAS, 444, 1518
Vogelsberger M. et al., 2014b, Nature, 509, 177
Walsh K., Tinker J., 2019, MNRAS, 488, 470
Wechsler R. H., Tinker J. L., 2018, ARA&A, 56, 435
Wechsler R. H., Zentner A. R., Bullock J. S., Kravtsov A. V., Allgood B., 2006, ApJ, 652, 71
Weinberger R. et al., 2017, MNRAS, 465, 3291
White S. D. M., Frenk C. S., 1991, ApJ, 379, 52
Xu X., Zheng Z., 2020, MNRAS, 492, 2739
York D. G. et al., 2000, AJ, 120, 1579
Zahid H. J., Geller M. J., Fabricant D. G., Hwang H. S., 2016, ApJ, 832, 203
Zahid H. J., Geller M. J., Damjanov I., Sohn J., 2019, ApJ, 878, 158
Zehavi I. et al., 2005, ApJ, 630, 1
Zehavi I., Contreras S., Padilla N., Smith N. J., Baugh C. M., Norberg P., 2018, ApJ, 853, 84
Zehavi I., Kerby S. E., Contreras S., Jiménez E., Padilla N., Baugh C. M., 2019, ApJ, 887, 17
Zentner A. R., Hearin A. P., van den Bosch F. C., 2014, MNRAS, 443, 3044
Zentner A. R., Hearin A., van den Bosch F. C., Lange J. U., Villarreal A., 2019, MNRAS, 485, 1196
Zhu G., Zheng Z., Lin W. P., Jing Y. P., Kang X., Gao L., 2006, ApJ, 639, L5
Zu Y., Mandelbaum R., 2016, MNRAS, 457, 4360
Zu Y., Mandelbaum R., Simet M., Rozo E., Rykoff E. S., 2017, MNRAS, 470, 551

This paper has been typeset from a TEX/LATEX file prepared by the author.