Galaxy sizes and the galaxy-halo connection - I. The remarkable tightness of the size distributions

Lorenzo Zanisi, Francesco Shankar, Andrea Lapi, Nicola Menci, Mariangela Bernardi, Christopher Duckworth, Marc Huertas-Company, Philip Grylls, Paolo Salucci

Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY

MNRAS **492,** 1671–1690 (2020) doi:10.1093/mnras/stz3516
Advance Access publication 2019 December 16

Galaxy sizes and the galaxy–halo connection – I. The remarkable tightness of the size distributions

Lorenzo Zanisi,[1,2]★ Francesco Shankar,[1]★ Andrea Lapi,[3,4] Nicola Menci,[5] Mariangela Bernardi,[6] Christopher Duckworth ,[7,8] Marc Huertas-Company,[9,10] Philip Grylls [1] and Paolo Salucci[3]

[1]*Department of Physics and Astronomy, University of Southampton, Highfield SO17 1BJ, UK*
[2]*DISCnet Center for Doctoral Training, University of Southampton, Highfield SO17 1BJ, UK*
[3]*SISSA, Via Bonomea 265, I-34136 Trieste, Italy*
[4]*IFPU – Institute for Fundamental Physics of the Universe, Via Beirut 2, I-34014 Trieste, Italy*
[5]*INAF – Osservatorio Astronomico di Roma, via Frascati 33, I-00078 Monteporzio, Italy*
[6]*Department of Physics and Astronomy, University of Pennsylvania, Philadelphia, PA 19104, USA*
[7]*School of Physics and Astronomy, University of St Andrews, North Haugh, St Andrews KY16 9SS, UK*
[8]*Center for Computational Astrophysics, Flatiron Institute, 162 Fifth Avenue, New York, NY 10010, USA*
[9]*Instituto de Astrofísica de Canarias (IAC), Departamento de Astrofísica, Universidad de La Laguna (ULL), E-38200 La Laguna, Spain*
[10]*LERMA, Observatoire de Paris, CNRS, PSL, Université Paris Diderot, F-75013, France*

Accepted 2019 December 4. Received 2019 November 26; in original form 2019 May 20

ABSTRACT

The mass and structural assembly of galaxies is a matter of intense debate. Current theoretical models predict the existence of a linear relationship between galaxy size (R_e) and the host dark matter halo virial radius (R_h). By making use of semi-empirical models compared to the size distributions of central galaxies from the Sloan Digital Sky Survey, we provide robust constraints on the normalization and scatter of the R_e-R_h relation. We explore the parameter space of models in which the R_e-R_h relation is mediated by either the spin parameter or the concentration of the host halo, or a simple constant the nature of which is in principle unknown. We find that the data require extremely tight relations for both early-type and late-type galaxies (ETGs, LTGs), especially for more massive galaxies. These constraints challenge models based solely on angular momentum conservation, which predict significantly wider distributions of galaxy sizes and no trend with stellar mass, if taken at face value. We discuss physically motivated alterations to the original models that bring the predictions into better agreement with the data. We argue that the measured tight size distributions of SDSS disc galaxies can be reproduced by semi-empirical models in which the R_e-R_h connection is mediated by the *stellar* specific angular momenta j_{star}. We find that current cosmological models of galaxy formation broadly agree with our constraints for LTGs, and justify the strong link between R_e and j_{star} that we propose, however the tightness of the R_e-R_h relation found in such ab initio theoretical models for ETGs is in tension with our semi-empirical findings.

Key words: galaxies: disc – galaxies: formation – galaxies: fundamental parameters – galaxies: kinematics and dynamics – galaxies: structure.

1 INTRODUCTION

In the local Universe it is observed that the physical properties of galaxies belonging to different morphological types (i.e. the Hubble sequence; Hubble 1926) define different scaling relations. This evidence is widely interpreted as the signature of the different physical processes that have shaped galaxies from their formation to our epoch. Of particular interest are the scaling relations that link galaxy structure and dynamics with galaxy stellar mass, the R_e-M_{star} relation (Shen et al. 2003; Bernardi et al. 2014; Lange et al. 2015) and the $j_{star}-M_{star}$ relation (Romanowsky & Fall 2012; Obreschkow & Glazebrook 2014), where R_e is the radius that encloses half of the light of the galaxy and j_{star} is the galaxy stellar angular momentum. These relations are believed to bear a significant trace of how galaxies stellar mass is assembled through cosmic time (e.g. Somerville & Davé 2015; Cappellari 2016).

★ E-mail: L.Zanisi@soton.ac.uk (LZ); f.shankar@soton.ac.uk (FS)

The standard scenario of the formation of galactic discs (first formulated in Fall 1983, Fall & Efstathiou 1980 and subsequently expanded by Mo, Mao & White 1998, MMW hereafter) envisages that LTGs form out of gas that cools and falls towards the centre of a host dark matter halo. If the gas shares the specific angular momentum of dark matter and retains a fraction f_j of it during the collapse, the model predicts that the scale length $R_{\rm d}$ of the newly formed disc should be

$$R_d \simeq \frac{\lambda}{\sqrt{2}} f_c f_R f_j R_h. \quad (1)$$

Here f_c and f_R are factors of order unity that provide minor corrections that account for adiabatic contraction (Blumenthal et al. 1986; Shankar et al. 2017) and disc self-gravity, and $R_{\rm h}$ is the dark matter halo radius

$$R_{\rm h} = \left(\frac{3M_{\rm h}}{4\pi\,\Delta\rho_\Delta}\right)^{\frac{1}{3}}, \quad (2)$$

where Δ is the virial overdensity with respect to the cosmological background density (Bryan & Norman 1998). The parameter λ in equation (1) is the spin parameter of dark matter defined by Peebles (1969) as

$$\lambda_{\rm P} = \frac{J \mid E \mid^{1/2}}{G M_{\rm h}^{5/2}}, \quad (3)$$

(J, E, and $M_{\rm h}$ are respectively the angular momentum, energy, and mass of the halo) or by Bullock et al. (2001) as

$$\lambda = \frac{J}{\sqrt{2} M_{\rm h} R_{\rm h}}. \quad (4)$$

The analysis of the Bolshoi–Planck and Multidark-Planck N-body numerical simulations (Klypin, Trujillo-Gomez & Primack 2011; Prada et al. 2012; Klypin et al. 2016) carried out in Rodríguez-Puebla et al. (2016) have revealed that when adopting the latter definition the distribution of the spin parameter is well fitted by a lognormal distribution, while the spin parameter in equation (3) follows a skewed distribution, closer to a 'Schechter-like' distribution (see Rodríguez-Puebla et al. 2016, for details). The dispersion in both cases has been found to be $\sigma_{\log\lambda} \sim 0.25$ dex. It is crucial to note that in equation (1) $\sigma_{\log\lambda}$ is the dominant source of scatter in disc scale length at fixed halo radius. In the following, for completeness, we will adopt both of the above definitions of the spin parameter in our semi-empirical models.

The success of the angular momentum conservation model put forward by MMW in predicting the structural properties of disc galaxies has been suggested in several works (Somerville et al. 2008; Kravtsov 2013; Huang et al. 2017; Straatman et al. 2017; Lapi, Salucci & Danese 2018b; Somerville et al. 2018). For instance, Kravtsov 2013 (K13 hereafter) found that a linear relationship between the sizes of galaxies and their haloes, as predicted by equation (1), indeed seems to exist. Somerville et al. (2018) used CANDELS (Gnedin et al. 2011; Koekemoer et al. 2011) and GAMA (Driver et al. 2011; Liske et al. 2015) observations to extend the results of K13 to $z \sim 3$, and claimed that the MMW model is able to explain the size distributions of all galaxies, irrespective of morphology. Lapi et al. (2018b) found that the normalization of the $R_{\rm e}$–$R_{\rm halo}$ relation of local discs is in good agreement with the predictions of the MMW model, although it is significantly offset with respect to that of the ETG-dominated sample of K13 (see also Huang et al. 2017). This is indeed reminiscent of the separation in angular momentum at fixed stellar mass reported by Romanowsky & Fall (2012). On the other hand, some studies question the validity of the MMW model based on the fact that the scatter that it would predict (i.e. at least 0.25 dex) overestimates the one found in observations (Desmond & Wechsler 2015a; Lapi et al. 2018b). For example, Jiang et al. (2019) have used two suites of hydrodynamical cosmological zoom-in simulations (VELA, Ceverino et al. 2014; Zolotov et al. 2015 and NIHAO, Wang et al. 2015) to study the connection between galaxy and halo size and found a weak link between galaxy and halo angular momenta, which would undermine the MMW model (similar conclusions were reported in Desmond et al. 2017 for the EAGLE simulation, Schaye et al. 2015). Fits to their simulations suggest instead an anticorrelation between $R_{\rm e}$ and the halo concentration c (Navarro, Frenk & White 1996) with $R_{\rm e} \propto c^{-0.7} R_{\rm h}$ and no correlation with the halo spin parameter. Moreover, Desmond (2017b) showed that introducing an anticorrelation between galaxy size and halo concentration results in an improved prediction of the radial dependence of the mass discrepancy–acceleration relation in abundance-matching-type models. Nevertheless, the physical motivation behind this empirical finding is yet to be found.

In this work we further investigate the role of dark matter in setting galaxy sizes and angular momenta from a semi-empirical point of view. In particular, we aim to directly test the dispersion predicted by the MMW model against a large photomorphological catalogue (Domínguez Sánchez et al. 2018). Our model builds on the assumption that a $R_{\rm e}$–$R_{\rm h}$ connection exists, to which we add an intrinsic scatter tuned to match observations of the size distributions. We then use the models proposed by Mo et al. (1998) and Jiang et al. (2019) to give empirical constraints on the galaxy–halo connection. We will show that in general observations require very small intrinsic scatters, a challenge for some of the assumed models. We will then discuss and interpret our results in the broader context of disc formation and angular momentum conservation The formation of ETGs in cosmological models will also be briefly discussed in relation to our semi-empirical constraints.

The paper is structured as follows. We first introduce the data set that we use to constrain the $R_{\rm e}$–$R_{\rm h}$ relation in Section 2. The details of the implementation of the models are given in Section 3. In Sections 4 and 5 respectively the results are presented and discussed. We discuss our results in the light of large-scale cosmological simulations in Section 6, and compare to other studies in Section 7. We will give our final remarks in Section 8. Caveats and additional discussion on the models are given in the Appendices.

In this work we adopt a standard flat ΛCDM cosmology with $\Omega_{\rm m} = 0.3$, $\Omega_\Lambda = 0.7$ and $h = 0.7$ and $\sigma_8 = 0.82$. We note that our results are largely independent on the exact choice of cosmology within the current constraints (Komatsu et al. 2011; Planck Collaboration et al. 2016).

2 DATA

In the following we will use the SDSS DR7 (Abazajian et al. 2009) spectroscopic sample as presented in Meert, Vikram & Bernardi (2015, 2016) (hereafter M15/16, available at http://alan-meert-website-aws.s3-website-us-east-1.amazonaws.com/fit_catalog/download/index.html.). The Meert et al. catalogues consist of 670722 objects the photometry of which benefits of substantial improvement both in background subtraction and fits to the light profiles. In the M15/M16 catalogues galaxies are fit with a `Sérsic + Exponential` as well as a `Sérsic` profile. In our work we only adopt the r-band best fit. The galaxy stellar masses are computed adopting such light profiles and the mass-to-light ratio $M_{\rm star}/L$ by Mendel et al. (2014), and the effective radius $R_{\rm e}$ is the truncated semimajor axis half-light radius of the full fit (e.g. Fischer, Bernardi & Meert 2017).

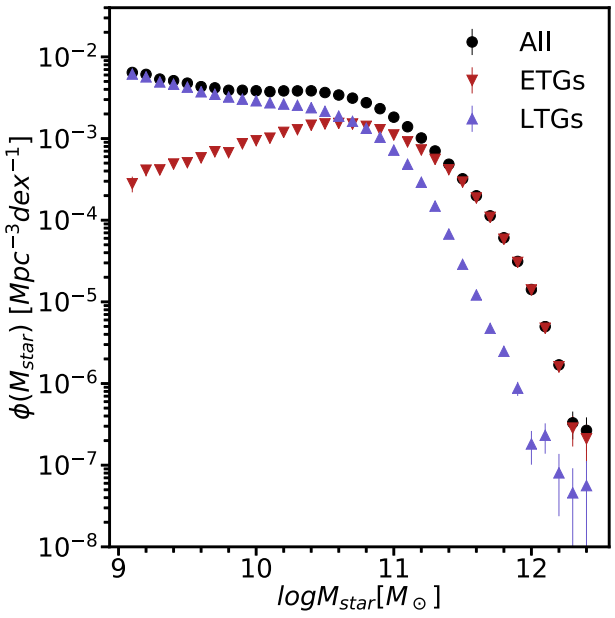

Figure 1. The V_{max} weighted morphological SMF from the M15/16 SDSS catalogues combined with the DS18 morphological catalogues.

Recently, Domínguez Sánchez et al. (2018) (hereafter DS18) have classified the morphology[1] of the objects in the Meert et al. catalogues by means of Convolutional Neural Networks. Previous works (Desmond & Wechsler 2015a; Lapi et al. 2018b) rely on catalogues that are orders of magnitude smaller than that used in this work and therefore quantities such as the morphological stellar mass function $\phi(M_{star})$ (SMF) or the morphological size function $\phi(R_e)$ are not available. One of the main aims of our work is exploring whether current models of disc formation are able to explain the size distribution of disc galaxies and the catalogues of morphology by DS18 offer a unique testbed of such models.

In this work we define LTGs and ETGs as having TType > 0 and TType ≤ 0, respectively. Note that we include S0 galaxies as part of the ETGs population. We further exclude from our selection Elliptical galaxies for which the `Sérsic + Exponential` fits provide a bulge-to-total ratio lower than 0.5. Indeed, visual inspection of a sample of these objects reveal crowded fields, close companions or classification errors.

We match the Meert et al. catalogues with the Yang et al. (2007) group catalogues. For each group we identify the central galaxy as the most luminous, while the remaining objects in that group are considered to be satellites. From the matched catalogues we compute the V_{max}-weighted stellar mass functions (SMF) of *central* galaxies for the full catalogues and for both ETGs and LTGs. Error bars are computed via jackknife resampling.[2] The inferred SMFs are reported in Fig. 1 – they agree with the results of Bernardi et al. (2017a) and they compare well to the morphological SMF of Bernardi et al. (2013). We then infer the fraction of late-type galaxies in each stellar mass bin, $f_L(M_{star}) = \phi(M_{star})^{LTGs}/\phi(M_{star})^{tot}$.

As for the sizes, we compute the V_{max}-weighted size functions $\phi(R_e)$ similarly to the SMF. Fig. 2 shows that $\phi(R_e)$ is only weakly bimodal (red downward triangles and blue upward triangles). At low and high masses the distributions are dominated by LTGs and

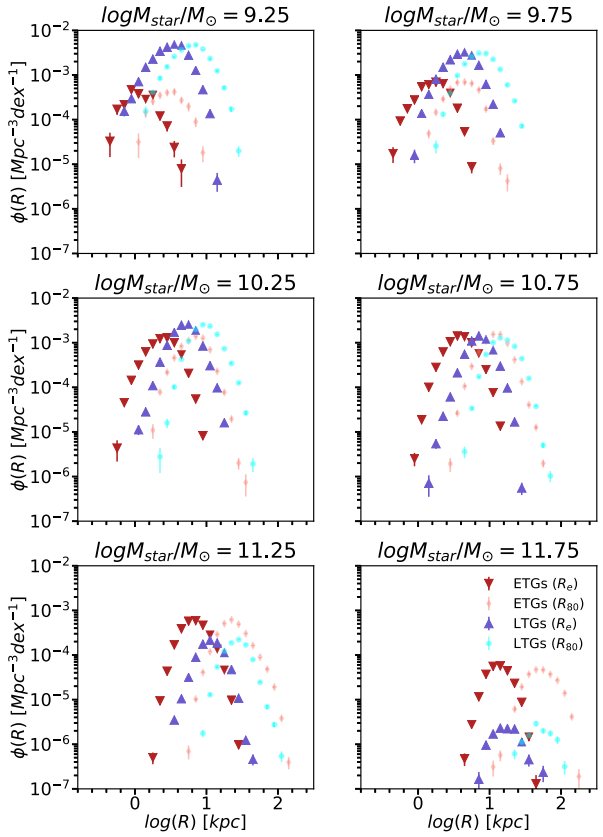

Figure 2. Size functions of ETGs and LTGs from the M15/16 SDSS catalogues combined with the DS18 morphological catalogues. Red downward triangles and blue upward triangles are for the R_e of ETGs and LTGs, respectively, while light pink diamonds and light cyan circles show the results for R_{80} for ETGs and LTGs.

ETGs, respectively, while the bimodality is most pronounced for $10^{10} \leq M_{star}/M_\odot \leq 10^{11}$. However, most strikingly, we see that the width of the size functions of ETGs and LTGs are comparable at all masses. It is also worth noticing that the total size function has a larger scatter than those of LTGs and ETGs taken singularly, at least for $M_{star} \leq 10^{11}\,M_\odot$.

We also notice that the mass dependence of the peak of the size function of ETGs is quite strong. Moreover, it is interesting to see that the size functions are somewhat skewed. While this feature was reported for LTGs also by van der Wel et al. (2014) for galaxies in CANDELS (Gnedin et al. 2011; Koekemoer et al. 2011), it is the first time that this is reported for ETGs. We have checked that using circularized sizes[3] leads to a reduced skewness in the size functions of ETGs (not so for LTGs). However we choose to use semimajor-axis sizes to enable a more direct comparison with LTGs, for which circularized sizes would be difficult to interpret physically as LTGs are intrinsically two-dimensional structures.

The aim of our theoretical work is to explore the mass dependence of the scatter and normalization of the input R_e–R_h relation to reproduce the measured SDSS size functions in different stellar mass bins.

To conclude, we also retrieve the size functions for the radius that encloses 80 per cent of the light, R_{80}, which is also shown in Fig. 2, and we comment on it in Section 7.3.

[1] More specifically, their TTypes (Nair & Abraham 2010).
[2] We adopt the publicly available library ASTROPY http://www.astropy.org/.
[3] Defined as $R_{e,circ} = R_{e,maj}\sqrt{b/a}$, where b and a are the semiminor and semimajor axis, respectively.

3 THE MODEL

The aim of this paper is to predict the fine details of the size functions of low-redshift galaxies. This issue has been explored in fully cosmological models only in a few instances. For example, Shankar et al. (2010) showed that such level of accuracy was not yet achievable in semi-analytic models. To the best of our knowledge, the size functions have never been explored in hydrodynamical simulations of galaxy evolution.

We here adopt the transparent and flexible approach offered by semi-empirical models. The latter are based on input observables, thus removing several degrees of freedom that characterize more standard approaches. Only a few parameters are required to match observables which are independent of the input. From the values of the model parameters it is then possible, we will show, to infer constraints on the processes that shape galaxy formation and evolution.

In brief, our semi-empirical model is structured as follows:[4]

(i) We extract catalogues of dark matter haloes from the Tinker et al. (2008) halo mass function.

(ii) We model the link between galaxies and dark matter via abundance matching (Section 3.1), and produce large mock catalogues of galaxies with moderate-to-high stellar masses.

(iii) We assign a half-light radius R_e to each galaxy according to diverse models of galaxy structure that exploit the galaxy–halo connection (Section 3.2).

We ultimately build a catalogue of dark matter haloes with mass M_h and size R_h and central galaxies with given stellar mass M_{star} and effective radius R_e. An accurate comparison to data will be able to set valuable constraints on the assumptions and related parameters in input to each of our adopted models.

For the remainder of the paper, we will consider dark matter haloes to follow a Navarro et al. (1996) density profile with scale radius R_s,

$$\rho(r) \propto \frac{1}{\frac{r}{R_s}\left[1+\frac{r}{R_s}\right]^2}, \tag{5}$$

and with $R_h = cR_s$ defining the *concentration* parameter c.

3.1 Abundance matching

The link between galaxies and their haloes is modelled following the popular *abundance matching* technique (AM) (e.g. Shankar et al. 2006; Vale & Ostriker 2006), which is essentially based on assigning a galaxy of stellar mass M_{star} to a host dark matter halo of mass M_h via rank ordering of the cumulative relative number densities,

$$n(>M_{star}) = n(>M_h) \tag{6}$$

to infer a (mean) stellar-mass-to-halo-mass relation (SMHM, Fig. B1). It has been shown that the simple ansatz of equation (6) is in excellent agreement with direct measurements of the SMHM with various techniques such as group finding algorithms (e.g. Yang et al. 2007), satellite kinematics (More et al. 2011), X-ray measurements of galaxy clusters (Kravtsov, Vikhlinin & Meshcheryakov 2018; Erfanianfar et al. 2019), and clustering analyses Shankar et al. 2006; Shankar et al. 2014b) as well as simulations (e.g. Guo et al. 2011; Matthee et al. 2017). The SMHM has been exploited by several authors for the most diverse purposes (only to mention a few, Diemer, More & Kravtsov 2013; Desmond & Wechsler 2017; Shankar et al. 2017; Shi et al. 2017; Posti et al. 2018a). The SMHM is modelled as a lognormal distribution with scatter σ_{SMHM}:

$$P(M_{star}|M_h) \equiv \mathrm{SMHM} = \frac{1}{\sqrt{2\pi\sigma^2_{SMHM}}} \times \exp\left[-\frac{(\log M_{star} - \langle(\log M_{star}\rangle)^2}{2\sigma^2_{SMHM}}\right]. \tag{7}$$

The overall scatter σ_{SMHM} stems from a convolution of observational errors σ_* and the intrinsic scatter σ_{int} that may be related to the stochasticity of the formation histories of galaxies within dark matter haloes (Tinker et al. 2017). We use the parametrization of the SMHM from Behroozi, Wechsler & Conroy (2013), which reads

$$\langle \log M_{star}\rangle = \log(\epsilon M_{10}) + g(x) - g(0), \tag{8}$$

where

$$g(x) = \delta\frac{\log(1+e^x)^\gamma}{1+e^{10^{-x}}} - \log(10^{\alpha x}+1) \tag{9}$$

and $x = \log(M_h/M_{10})$. We also assume that $\sigma_{SMHM} = 0.16$ dex,[5] as suggested by other studies at low redshift (e.g. Tinker et al. 2017), with no dependence on halo mass, which is a very good approximation especially at the high-mass end of the SMF (Shankar et al. 2014b; Rodríguez-Puebla et al. 2015). The parameters of the SMHM are $\vec{p} = (\epsilon, M_{10}, \delta, \alpha, \gamma)$.

According to AM, the SMF is retrieved from the SMHM by computing the integral

$$\phi(M_{star}) = \int \mathrm{SMHM}(M_h;\vec{p})\phi(M_h)\,\mathrm{d}M_h, \tag{10}$$

where $\phi(M_h)$ is the halo mass function. Hence, abundance matching reproduces the observed galaxy SMF by design, and it can therefore be used to produce realistic mock catalogues.

We run a Markov chain Monte Carlo[6] to fit the parameters of the SMHM to the SMF of central galaxies in SDSS adopting the Tinker et al. (2008) halo mass function for central haloes only by maximizing the likelihood $\mathcal{L} \propto \exp(-\chi^2)$. The parameters of our $z \sim 0.1$ SMHM are the following:

$$M_{10} = 11.632^{+0.008}_{-0.009} \tag{11}$$

$$\epsilon_0 = -1.785^{+0.010}_{-0.008} \tag{12}$$

$$\alpha_0 = -2.352^{+0.026}_{-0.021} \tag{13}$$

$$\delta_0 = 3.797^{+0.052}_{-0.052} \tag{14}$$

$$\gamma_0 = 0.600^{+0.100}_{-0.013} \tag{15}$$

$$\sigma_{SMHM} = 0.16 \quad \text{(fixed)}. \tag{16}$$

It is perhaps not surprising that the uncertainty on the inferred parameters is so low compared to other works, given the very small error bars on the SMF. Moreover, here the fit is performed at one redshift only, as opposed to, e.g. Rodríguez-Puebla et al. (2017).

[4]To build our model we extensively rely on the open-source PYTHON package COLOSSUS (Diemer 2018).

[5]We note that our choice for a relatively small scatter in the SMHM relation is a conservative one. Larger values of the scatter σ_{SMHM} would strengthen our main result for the need of a tight R_e–R_h relation.

[6]We use the publicly available PYTHON package EMCEE, (Foreman-Mackey et al. 2013).

While many studies include satellite galaxies in their models (Behroozi et al. 2013; Rodríguez-Puebla et al. 2015; Rodríguez-Puebla et al. 2017; Behroozi et al. 2019; Grylls et al. 2019), we choose to restrict our analysis to central galaxies only. Hearin et al. (2019) have shown that the sizes of satellite galaxies are linked to their halo mass at infall time, which is not straightforwardly available in the analytic halo catalogues that we are using.

We note that in principle LTGs and ETGs may occupy different loci in the $M_{\rm star}$–$M_{\rm halo}$ plane, as suggested by some studies (Dutton et al. 2010; More et al. 2011; Rodríguez-Puebla et al. 2015; Moster, Naab & White 2018). However, as pointed out in Wechsler & Tinker (2018), there is no agreement between different studies, which sometimes even reach opposite results (Rodríguez-Puebla et al. 2015; Behroozi et al. 2019; Moster et al. 2018). In what follows, we will therefore adopt the same SMHM for both LTGs and ETGs. We will show in Appendix B1 that our core results are largely independent of the choice of SMHM.

3.2 Galaxy sizes

We now proceed to assign to our mock galaxies a size. This is done according to theoretically or empirically justified models that link galaxy sizes to the size of their host dark matter halo, $R_{\rm h}$.

We adopt three models of galaxy sizes:

(i) The *MMW model* (or λ model). This model is inspired by the classical picture in which galaxies are born as discs out of cooling from the hot gas in the halo (MMW, see Section 1). We recast equation (1) as

$$R_{\rm e} = \frac{1.68}{\sqrt{2}} f_c f_j f_R \lambda R_{\rm h} = 1.68 A_\lambda \lambda R_{\rm h}, \tag{17}$$

where we define $A_\lambda = f_c f_j f_R/\sqrt{2}$, and the factor 1.68 comes from $R_{\rm e} \approx 1.68 R_{\rm d}$, appropriate for an exponential profile. We will take $A_\lambda = 1$ and rescale our results as needed to match the data. We discuss the implications of this assumption in Section 5. The spin parameter λ is defined either by Peebles (1969) (equation 3) or by Bullock et al. (2001) (equation 4), see Section 1. Note that the MMW model was devised to explain the formation of the *baryonic* size of galactic discs, while we will compare it to the *stellar* sizes. In particular, the factor f_j addresses the angular momentum retention of *baryons* and not stars. We will discuss the implications of this difference in Section 5.2.

(ii) The *K13 model*. This model is based on the empirical findings by Kravtsov (2013). The author adopted abundance matching techniques similar to the ones presented here[7] and found evidence that

$$R_{\rm e} = A_k R_{\rm h}. \tag{18}$$

Here A_k is the normalization which may vary with halo mass or galaxy stellar mass. We add to equation (18) an intrinsic lognormal scatter σ_K, which, as A_k, is a free parameter that can be tuned to match observations. The K13 model is hence purely empirical and will be applied to both LTGs and ETGs. Note that the physical meaning of both A_k and σ_K is not known a priori. However the K13 model reduces precisely to the MMW model when applied to LTGs in the case $\sigma_K = \sigma_{\log\lambda} \approx 0.25$ dex. The K13 model, in this respect, is more flexible. In fact, being empirically based, it can allow for any input scatter. As suggested by Kravtsov (2013), constraining the value of σ_K can be crucial to probe models of galaxy formation.

(iii) The *concentration model*. Recently, based both on observational and numerical studies, some groups have suggested that galaxy sizes should scale in a way that is inversionally proportional to halo concentration (Desmond et al. 2019; Jiang et al. 2019). Following Jiang et al. (2019), mathematically this model can be expressed as

$$R_{\rm e} = A_c \left(\frac{c}{10}\right)^{\gamma} R_{\rm h}, \tag{19}$$

with $\gamma < 0$. Similarly to what assumed in the other two models, we initially adopt $A_c = 0.012$ (Jiang et al. 2019) and then rescale our results to match data. We also adopt the concentration–mass relation by Dutton & Macciò (2014)

$$\log c = a + b \log M_{\rm h}[{\rm M}_\odot]/10^{12}/h \tag{20}$$

with $a(z) = 0.537 + (1.025–0.537)\exp(-0.718 z^{1.08})$ and $b(z) = -0.097 + 0.024z$. Dutton & Macciò (2014) report a lognormal scatter of about ∼0.11 dex at $z \sim 0$, which is independent on halo mass. Hydrodynamical simulations suggest that the intrinsic scatter $\sigma_{\rm CM}$ in this model is lower than in the K13 model (Jiang et al. 2019). Indeed, we will not include any other source of scatter in the concentration model other than the scatter in concentration (i.e. $\sigma_{\rm CM} = 0$, see discussion in Section 3.3).

We find that the concentration model may in fact be interpreted as a further generalization of the K13 model. Indeed, combining equations (19) and (20), and bearing in mind that $M_{\rm h} \propto R_{\rm h}^3$, yields

$$R_{\rm e} \propto R_{\rm h}^{3b\gamma+1}, \tag{21}$$

which reduces to the K13 model when $\gamma = 0$. The scatter implied by this version of the concentration model and the difference with that produced by equation (19) is discussed at the end of the next section.

Although, following the seminal approach by K13, we model the link between galaxies and their haloes in terms of the projected effective radius $R_{\rm e}$, such relation would be more physically motivated when expressed in terms of the 3D physical half-mass radii of galaxies $R_{\rm e,3D}$. However, the deprojection of galaxy shapes is a very hard task (Jiang et al., in preparation). In any event, as discussed in Appendix B2, projection effects tend to increase the variance in the measured effective radii, implying even tighter distributions in intrinsic sizes $R_{\rm e,3D}$. Accounting for deprojection effects would then further tighten the measured distribution of 3D galaxy sizes, which would constitute an even harder challenge for models. In Appendix B2 we give an estimate of the (small) biases induced by assuming that $R_{\rm e,2D} = R_{\rm e,3D}$ based on mock observations of galaxies from the Illustris TNG simulation (Nelson et al. 2019; Huertas-Company et al. 2019; Rodriguez-Gomez et al. 2019).

3.3 Sources of scatter in our models

At fixed bin in stellar mass, the width of the implied size distribution resulting from our three adopted models depends on a combination of different effects. In all models, there is always a contribution from the intrinsic scatter in the SMHM, as shown in Fig. 3. In fact, at fixed stellar mass there is a distribution of possible host

[7] We note that Kravtsov (2013) *backwards* models, that is in his work dark matter haloes are assigned to galaxies via the inverse of the SMHM relation, without accounting for its scatter. Somerville et al. (2018) have shown that doing so would result in severe biases in the estimate of the host halo masses. These authors stress that the *forward* modelling approach that we adopt here is instead more accurate (see also discussion in Shankar et al. 2017 and Rodríguez-Puebla et al. 2017).

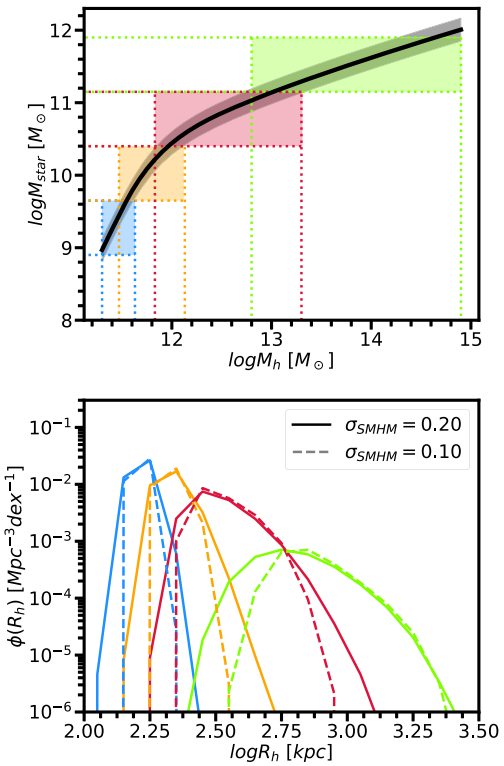

Figure 3. Role of the shape of the SMHM and its $\sigma_{\rm SMHM}$ in setting the scatter in halo size (and hence in galaxy size according to our models). *Upper panel.* The black line is the SMHM retrieved from MCMC fitting of the total SMF in SDSS (Section 3). Different cuts in stellar mass highlight different regions of the SMHM with different colours. Each coloured band corresponds to a stellar mass cut of the same width (0.75 dex). Their projections on to the x-axis select the halo mass range in which galaxies of a given stellar mass are expected to reside. *Lower panel.* The halo size functions resulting from the stellar mass cuts applied in the upper panel, with the same colour code. Dashed and solid lines indicate predictions for $\sigma_{\rm SMHM} = 0.10$ dex and $\sigma_{\rm SMHM} = 0.20$ dex. No additional scatter in size is added. Higher stellar mass cuts are naturally mapped in broader distributions. Larger values of $\sigma_{\rm SMHM}$ correspond to broader distributions with an effect that is larger the higher the stellar mass cut.

haloes, a feature that is usually described in terms of the *halo occupation distribution function* $P(M_{\rm h}|M_{\rm star})$,[8] which translates into a distribution in halo sizes $P(R_{\rm h}|M_{\rm star})$ (see equation 2), the main ingredient in all our models. The distributions get progressively broader for higher stellar mass cuts, given the shallow slope of the SMHM at high halo masses in combination with its intrinsic scatter. As this feature is mainly driven by the double power-law shape of the SMHM, it would be present even in the case of $\sigma_{\rm SMHM} = 0$.

The origin of the double power-law shape of the SMHM is thought to be linked to the efficiency of star formation, which is suppressed below and above a certain halo mass, where Supernova and a combination of AGN feedback and virial shocks, respectively, are believed to be most efficient (e.g. Shankar et al. 2006; Faucher-Giguère, Kereš & Ma 2011; Pillepich et al. 2018a). If zero intrinsic scatter σ_K in the K13 model was required to match observations, it could be argued that the same physics that shapes the SMHM is responsible for the width of the observed size distributions. On the other hand, wherever $\sigma_K > 0$ is needed to match the data, there must be some physical processes unrelated to the build-up of the shape of the SMHM at play in determining the broadness of the observed size functions.

[8]Which is *different* from the inverse of $P(M_{\rm star}|M_{\rm h})$ (Shankar et al. 2014b; Somerville et al. 2018).

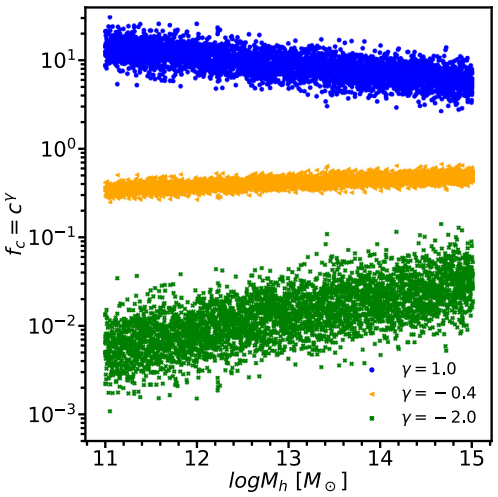

Figure 4. Scatter induced by different choices of γ in the factor $f(c) = c^\gamma$ as a function of halo mass. Blue dots, orange triangles and green crosses are for $\gamma = 1$, -0.4, and -2.0, respectively. Concentrations are from Dutton & Macciò (2014).

In the MMW and concentration models, the scatter is due to both the halo occupation distribution (described above) and the internal properties of the dark matter hosts. In fact, most of the scatter of the *MMW model* derives from the distribution of the spin parameter λ, with a typical dispersion of $\sigma_{\log\lambda} \approx 0.25$ dex and very weak dependence on halo mass. Interestingly, we find that for the concentration model one additional source of scatter derives from the factor c^γ in equation (19). As shown in Fig. 4, the (quite tight) distribution in concentration at fixed halo mass (blue dots) is modified for different values of γ. As we will see in Section 4.2, adopting lower values of γ will result in broader distributions. Such effect is degenerate with the intrinsic scatter in the concentration model $\sigma_{\rm CM}$. We set $\sigma_{\rm CM} = 0$ in this work, noting that having $\sigma_{\rm CM} > 0$ would require higher values of γ to match the observed size functions. Therefore our constraints are *lower limits* to γ.

As a final note, we recall that the concentration model may be seen as a further generalization of the K13 model (see Section 3.2). It must however be noted that at fixed $R_{\rm h}$ the concentrations follow a lognormal distribution, while the expression in equation (21) has been derived assuming only the *mean* of equations (19) and (20). Hence, the arguments about the scatter in the concentration model presented in this section would not apply straightforwardly to equation (21). However, studying this issue in more detail is outside of the scope of our paper.

4 RESULTS

We now proceed to a careful comparison of our three models to the size functions extracted from the SDSS photomorphological catalogues (see Section 2).

In our models we do not differentiate between ETGs and LTGs, and so the size function from the model should be compared to the *total* observed distribution. In each bin of stellar mass we retrieve the size function from our model and rescale it to match the observed

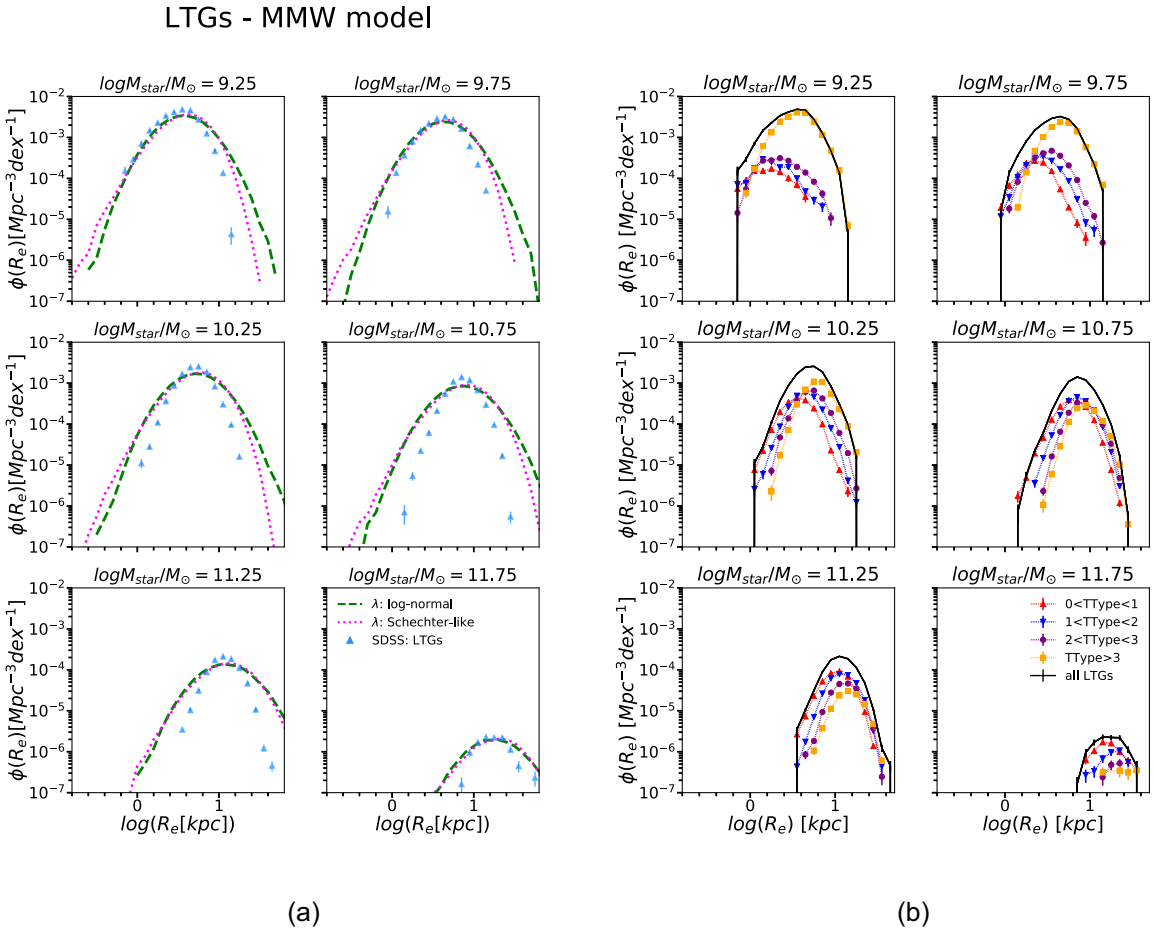

Figure 5. *Left*: Size functions from the MMW model (λ model, equation 17). The spin parameter λ is retrieved either from the lognormal (pink dotted lines) or Schechter-like (purple dashed lines) fits from Rodríguez-Puebla et al. (2016). Data points are LTGs from the photo + morphological SDSS catalogues described in Section 2. *Right*: Size functions for LTGs divided in bins of TType. The total distribution is shown with solid black lines, the distributions for $0 < \mathrm{TType} < 1$, $1 < \mathrm{TType} < 2$, $2 < \mathrm{TType} < 3$, and $\mathrm{TType} > 3$ are instead shown with red upward triangles, blue downward triangles, purple circles, and yellow squares plus dotted lines, respectively.

ones for different morphological types

$$\phi(R_\mathrm{e}|M_\mathrm{star})^{\mathrm{LTGs}}_{\mathrm{obs}} = f_\mathrm{L}(M_\mathrm{star})\phi(R_\mathrm{e}|M_\mathrm{star})^{\mathrm{tot}}_{\mathrm{model}} \quad (22)$$

$$\phi(R_\mathrm{e}|M_\mathrm{star})^{\mathrm{ETGs}}_{\mathrm{obs}} = (1 - f_\mathrm{L}(M_\mathrm{star}))\phi(R_\mathrm{e}|M_\mathrm{star})^{\mathrm{tot}}_{\mathrm{model}}, \quad (23)$$

where $f_\mathrm{L}(M_\mathrm{star})$ is the fraction of late-type galaxies as defined in Section 2. Note that here we are implicitly assuming that ETGs and LTGs at fixed stellar mass live on average in the same dark matter haloes.

Figs 5(a), 6(a), and (b) show a comparison between the observed size functions $\phi(R_\mathrm{e})$ of LTGs and our models (the MMW, K13, and concentration models, respectively). We bin both our model galaxies and data in bins of 0.5 dex in stellar mass. In all the figures the model size functions are shifted to match the location of the peaks of the observed distributions. The normalizations of the different models (A_λ, A_K, A_c) in each stellar mass bin are reported in Table 1. Results for ETGs are given in Appendix A.

4.1 The MMW model

In Fig. 5(a) it can be seen that the classical λ-disc model by MMW does not provide a good fit to data, irrespective of the definition of spin parameter adopted (lognormal or Schechter-like, see Section 1). This effect becomes gradually more severe as more massive populations of LTGs are considered. As for the normalization A_λ we note that the values listed in Table 1, recalling equation (17) and that $R_\mathrm{e} \approx 1.68 R_\mathrm{d}$ for LTGs, are consistent with $R_\mathrm{d} \approx 0.3\lambda R_\mathrm{h}$, in agreement with the study by Lapi et al. (2018b). Notably, given that $A_\lambda = f_j f_\mathrm{R} f_\mathrm{c}/\sqrt{2}$, this is fully consistent with the MMW model with an angular momentum retaining factor f_j of about 0.5–0.7 (see also Desmond & Wechsler 2015b).

4.1.1 The case of bulgeless galaxies

To select LTGs from the catalogues by DS18 we applied the cut $\mathrm{TType} > 0$. We note that this cut might still include galaxies with prominent bulges, which may have a non-negligible contribution in determining the half-light radius of the whole galaxy, especially at high masses (Kormendy 2016). On the contrary, the MMW model is expected to work for pure disc galaxies only and therefore comparing the MMW model with LTGs selected as above may not be entirely accurate.

In Fig. 5(b) we show the size functions of LTGs in our SDSS photo + morphological catalogues divided by TType. We find

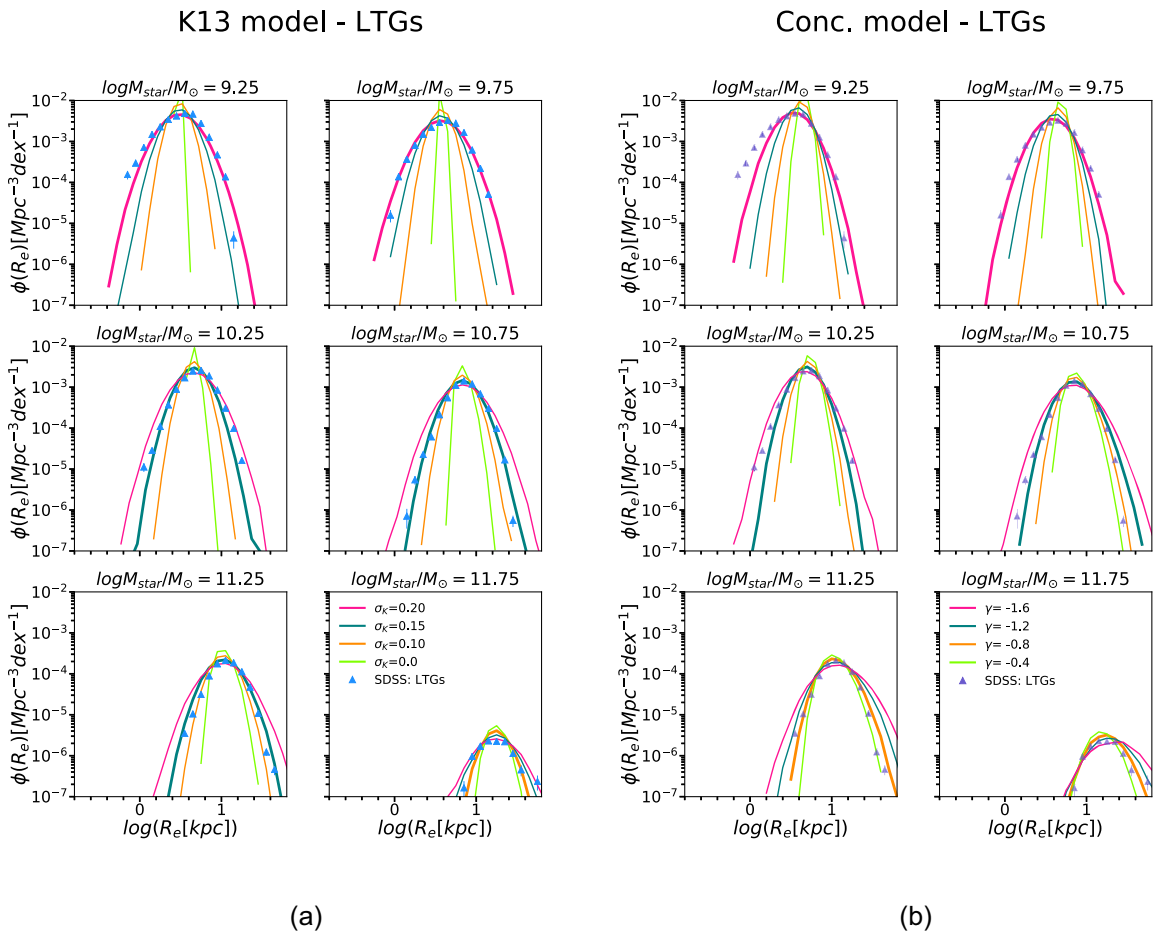

Figure 6. *Left*: Size functions from the K13 model (equation 18) for values of $\sigma_K = 0.00, 0.10, 0.15, 0.20$. *Right*: Size functions from the concentration model (equation 19) for values of $\gamma = -1.6, -1.2, -0.8, -0.4$. Models that work best for a given stellar mass bin are highlighted in each panel by a thicker line. Data points are LTGs from the photo + morphological SDSS catalogues described in Section 2.

Table 1. Values of A_k, A_c, and A_λ in different bins of $M_{\rm star}$, for LTGs. Compare to Table A1.

$M_{\rm star}$	9.25	9.75	10.25	10.75	11.25	11.75
A_k	0.018	0.019	0.019	0.019	0.024	0.024
A_c	0.034	0.030	0.027	0.026	0.021	0.015
A_λ	0.60	0.60	0.60	0.60	0.60	0.47

that for $M_{\rm star} \lesssim 10^{10.5}\,{\rm M}_\odot$ the population is entirely dominated by galaxies with TType > 3, which represent the disc-dominated Sb–Sc–Sd galaxies according to the Nair & Abraham (2010) classification against which the CNN in DS18 was trained. At higher masses earlier types start dominating, with the peaks of their size functions being located at lower $R_{\rm e}$ due to the progressively important contribution of the bulge. Interestingly, LTGs with TType > 3 display an even tighter size distribution than that of the overall population, which dismisses our concerns.[9] It might however be argued that the comparison between model and data may not be ideally set up since not all Sb–Sc–Sd can be fitted by a pure exponential disc. To check for the latter effect, we further restricted our analysis to LTGs with TType > 3 and $B/T < 0.2$ and still did not find significant changes in the width of the size distributions.

[9]Note that the skewness of the size function is partially explained by the morphological mix of LTGs, but that for the later types the skewness still persists.

4.2 The K13 and concentration models

The size distributions from the K13 and concentration models are reported in Figs 6(a) and (b). The free parameters in these models are (A_k, σ_K) and (A_c, γ), respectively. The values of A_k and A_c are reported in Table 1. Although we do not aim to give exact fits for γ and σ_K, we show how the models depend on these parameters by plotting results for models with different values of σ_K and γ as labelled. As it can be seen in Figs 6(a) and (b), varying σ_K and γ leads to quite drastic differences in the model distributions. In each panel of the figures we highlight with a thicker line the parameter that seems to best reproduce observations. For the K13 model an intrinsic scatter larger than ∼0.20 dex would be strongly disfavoured by current data. Likewise, K13 models with $\sigma_K \lesssim 0.1$ dex provide a poor agreement with data. However, our model suggests a trend in which σ_K decreases as higher stellar mass bins are considered, with $\sigma_K \sim 0.20$ dex for the lowest masses and $\sigma_K \sim 0.10$ dex for the most massive galaxies.

Turning to the *concentration model*, at lower stellar masses lower values of γ are preferred, while for more massive galaxies $\gamma \sim -0.8$ gives a better match to data. Adopting $\gamma \gtrsim -0.4$ or $\gamma \lesssim -1.6$ would

produce distributions that are too tight or too wide, respectively, compared to the observed ones.

It is worth pointing out here that the same identical considerations about γ and σ_K can be applied to ETGs, as shown in Appendix A. In Table A1 we report the values of A_c and A_K, which instead are significantly lower than those of LTGs (compare to Table 1). Thus, ETGs and LTGs define two separate relations in the $R_e - R_h$ plane, qualitatively in agreement with the findings of Huang et al. (2017) (see Section 1). However, we recall that in our framework $R_e \equiv R_{e,2D}$ is the 2D projection on the sky of the intrinsic galaxy shape. While in this work we do not model deprojection explicitly, we will show that cosmological models where the intrinsic galaxy shape is available still produce a rather marked dichotomy in the $R_{e,3D} - R_h$ relation (see Section 6), qualitatively in agreement with our empirical findings for $R_{e,2D}$.

5 WHAT DRIVES THE TIGHTNESS OF THE OBSERVED SIZE DISTRIBUTIONS?

5.1 Implications for ultra-massive galaxies

We recall that part of the scatter σ_K originates from the shape of the SMHM (i.e. the halo occupation distribution, see Fig. 3). The latter contributes very little to the observed size functions at low masses (see Figs 6a and A1a) and further scatter is needed to obtain a good match to data. On the other hand for UMGs (ultra-massive galaxies, for which $M_\odot > 10^{11.5}\,M_\odot$ – and which are mostly ETGs, see Fig. 2) the contribution of the halo occupation distribution is the most relevant source of scatter. Therefore only a very small intrinsic additional scatter $\sigma_K \lesssim 0.1$ is necessary to match observations. Thus, essentially, the K13 model predicts that the width of the galaxy size distribution at the high mass end can be entirely interpreted in terms of their halo occupation distribution, that is, *the information about the broadness of the size distribution of UMGs is already contained in the SMHM*. We also note that adopting a flatter high mass end slope in the SMHM, as proposed by other groups (e.g. Behroozi et al. 2013; Moster, Naab & White 2013), would result in an even larger source of scatter, perhaps in tension with the width of the observed size function of UMGs.

5.2 On the validity of the 'MMW framework'

Although the MMW model strictly predicts galaxy sizes only, it sets a framework in which also galaxy angular momenta can be predicted. Thus, in the following we will refer to the general notion of angular momentum conservation, which works for both galaxy angular momentum and sizes, as the *MMW framework*.

The prediction for galaxy angular momenta in the MMW framework is straightforward. In the context of a biased collapse scenario (e.g. Kassin et al. 2012; Romanowsky & Fall 2012; Lapi et al. 2018a; Posti et al. 2018a)

$$j_{\rm inf} = f_{\rm inf}\, j_{\rm h} = f_{\rm inf}\sqrt{2}\lambda V_{\rm h} R_{\rm h}, \tag{24}$$

where $f_{\rm inf}$ is the fraction of gas that is able to cool efficiently (Shi et al. 2017). If angular momentum was strictly conserved during gas collapse, the distribution of $j_{\rm inf}$ should be such that $\sigma_{j_{\rm inf}} \approx \sigma_{\log\lambda} \approx 0.25$ dex. Studies have constrained $f_{\rm inf} \approx 1$ for LTGs (e.g. Shi et al. 2017), so that the gas that cools has the same specific angular momentum of the host halo. The factor f_j that appears in equation (17) corresponds exactly to $f_{\rm inf}$ in the biased collapse scenario.

5.2.1 *Is it safe to compare the MMW model to stellar sizes?*

The MMW model gives a clear prediction for the *baryonic* sizes of LTGs, but here we compare to the stellar effective radius instead. It could therefore be argued that from this comparison it is not possible to draw conclusions about the MMW model. We now show that instead $R_{\rm e,star}$ is a good proxy of $R_{\rm e,bar}$.

Kravtsov (2013) have shown that in a sample of local LTGs from Leroy et al. (2008), the gas and stellar mass surface densities are well described by exponential profiles with $R_{d,\rm gas} \approx 2.5 R_{d,\rm star}$. Using this information, it can be shown that at fixed radius, $\Sigma_{\rm bar}$ and $\Sigma_{\rm star}$ differ by only $\lesssim$10 per cent, and hence $R_{d,\rm bar} \approx R_{d,\rm star}$. However, this is not sufficient to confirm that observations of $R_{d,\rm star}$ can be compared to the predictions of the MMW model for $R_{d,\rm bar}$. The reason for this is that the factor f_j that appears in equation (17) strictly refers to the angular momentum retained by all *baryons*, which might well be different from that retained by stars $f_{j,\rm star}$, since gas is so spread out in the outskirts of LTGs with substantially high velocities traced by HI emission. On the other hand, using constraints from chemical abundances and star formation efficiency, Shi et al. (2017) have shown that $f_{\rm inf} \approx f_{j,\rm star}$ for LTGs, and we can therefore conclude that the MMW model can be also extended to the stellar component as well. More details on this subject can be found in Section 7.1, where we compare our Semi-Empirical Model to a state-of-the-art semi-analytic model.

5.2.2 *Is the MMW model consistent with observed LTGs scaling relations?*

We now show that the observed proportionality between R_e and R_h, as well as its scatter, is fully consistent with observations of galaxy angular momenta and a high angular momentum retention factor $f_j \sim 0.5$ in the MMW framework (Romanowsky & Fall 2012; Posti et al. 2018a).

We recall the mathematical form of the MMW model

$$R_{\rm d} \simeq \frac{\lambda}{\sqrt{2}} f_{\rm c} f_{\rm j} f_{\rm R} R_{\rm h}. \tag{25}$$

We now note that $f_{\rm j} = j_{\rm star}/j_{\rm h}$ with $j_{\rm h} = \sqrt{2}\lambda V_{\rm h} R_{\rm h} \propto \lambda M_{\rm h}^{2/3}$ (see e.g. Romanowsky & Fall 2012). With this in mind, the equation above reads

$$R_{\rm d} \simeq B \frac{j_{\rm star}}{M_{\rm h}^{2/3}} R_{\rm h}, \tag{26}$$

where, critically, λ disappears, which will be an important point in the discussion that follows. Here B is a factor that encloses all the dependencies not relevant for our discussion. Posti et al. (2018a) have shown that in the mass range $9 < \log M_{\rm star}/M_\odot < 11$ the ratio $\tilde{f}_j \equiv B j_{star}/M_h^{2/3}$ depends very weakly on stellar mass.[10,11] The only dependence left on stellar mass is in the factor $R_h \propto M_h^{1/3} \propto M_{star}^{1/6}$ (Dutton et al. 2010), for LTGs with $M_{\rm star} < 10^{11}\,M_\odot$.

With all this in mind, equation (1) reads

$$R_{\rm d} \propto \tilde{f}_j M_{\rm star}^{1/6}. \tag{27}$$

[10] Actually Posti et al. (2018a) constrain $f_j = j_{\rm star}/j_{\rm h} \approx 0.5$, but since λ is mass independent the same applies to $\tilde{f}_j$.

[11] A close look at their fig. 5 for the Dutton et al. (2010) SMHM reveals that at most $f_j \sim M_{\rm star}^{0.1}$. Moreover, the factor $f_c \propto c^{-0.2}$ (see Mo et al. (1998) and Jiang et al. (2019)) depends very weakly on halo mass ($c \propto M_h^{-0.1}$, Dutton & Macciò 2014) and therefore on stellar mass.

A slope of 1/6 is consistent with measurements of the slope of the $R_e - M_{star}$ relation of LTGs (see Shen et al. 2003; Bernardi et al. 2014), plus minor corrections mainly due to the factor $\tilde{f}_j$. We now note that the normalization A_λ is remarkably constant over the whole mass range studied here (see Table 1). Furthermore, the scatter in this relation is entirely governed by j_{star}, as for the mass range under consideration the halo occupation distribution is not critical (Fig. 3) and therefore for this purpose $\sigma_{\tilde{f}_j} \approx \sigma_{j\,star} \approx 0.20$ dex (Posti et al. 2018b). Notably, this is consistent with the scatter of the $R_e - R_h$ relation that we calibrate with the K13 model $\sigma_K \lesssim 0.2$ dex. Moreover, the scatter that would come from f_c is negligible (see Fig. 4).

We note that to compute j_{star} some authors adopt the simple scaling

$$j_{star} \approx R_e V_c, \tag{28}$$

where V_c is the circular velocity of a galaxy assuming a flat rotation curve (Romanowsky & Fall 2012). In this case the observed scatter in R_e would drive the one in j_{star}, making the argument above circular. However, the constraints on the scatter in j_{star} by Posti et al. (2018b) quoted above, are found by direct integration of the observed rotation curves in the SPARC sample (Lelli, McGaugh & Schombert 2016). It is also intriguing that, to first order, equation (26) is consistent with the empirical finding of equation (28). Indeed, assuming for simplicity an isothermal profile for dark matter haloes, for which the circular velocity V_h is proportional to the halo radius R_h, $V_h \propto R_h$, equation (28) can be easily inferred from equation (26) assuming that $V_c \approx V_h$ (Lapi et al. 2018b).

Our conclusions above further corroborate the theoretical link between galaxy sizes and their angular momentum. It is interesting to investigate whether the origin of such a connection lies in the MMW framework. Indeed, Cervantes-Sodi et al. (2013) and Burkert et al. (2016) have observationally constrained the quantity λf_j and have found that its dispersion is $\lesssim$0.2 dex. In the light of the discussion above, where we have shown that λf_j does not actually depend on λ, we argue that what these authors have found is in fact the scatter of the distribution of $\tilde{f}_j$, $\sigma_{\tilde{f}_j}$, which essentially boils down to the distribution of galaxy stellar angular momenta, $\sigma_{j\,star}$. As a caveat, it should be noted that actually in both studies it is the gas kinematics that is probed, which may differ from the stellar kinematics. Nevertheless, in a recent study Aquino-Ortíz et al. (2018) have shown that for a sample of local LTGs from the CALIFA survey (Sánchez et al. 2012) gas and stellar kinematics show similar scaling relations. In the discussion above, we have tentatively assumed that this is also true for the sample of high-redshift galaxies used in Burkert et al. (2016).

To summarize, our work suggests that the MMW taken at face value is able to recover the median values of the observables, but it fails at reproducing the width of their distributions. Conversely, observations of galaxy angular momenta combined with the MMW model recover our semi-empirically determined constraints on the $R_e - R_h$ relation $\sigma_K \lesssim 0.20$ dex. Moreover, we have analytically shown that the MMW model naturally gives the slope of the $R_e - M_{star}$ relation. Our discussion confirms and extends to greater detail the results of Obreschkow & Glazebrook (2014), who have shown that if angular momentum and mass are known for LTGs, then the size–mass relation, as well as other observable LTG scaling relations, are automatically reproduced.

5.3 Reconciling the MMW framework and observations

We have seen in the previous section that both galaxy sizes and angular momenta predicted in the MMW framework suffer from the same shortcomings. In particular, although a mean $\langle\lambda\rangle \approx 0.035$ seems to work well in predicting the mean of the observables, the width of the predicted distributions is $\approx$0.25 dex, whereas independent observations systematically find evidence for a width of $\lesssim$0.10–0.20 dex. We thus conclude that either the MMW model is an oversimplification of an underlying more complex problem (Section 5.3.1), or we must introduce additional physical processes that limit the acceptable values of λ (Section 5.3.2).

5.3.1 Insights from hydrodynamic cosmological zoom-in simulations

Using hydrodynamical cosmological zoom-in simulations, Danovich et al. (2015) have traced the build-up of galaxy angular momentum in four phases that are linked to different spatial scales, from the cosmic web ($R \approx 2R_h$) to the innermost part of the halo where $R \lesssim 0.1R_h$. The region where $0.1 \lesssim R/R_h \lesssim 0.3$, termed as the 'messy region' (Ceverino, Dekel & Bournaud 2010), is particularly interesting. This is the zone where the cold streams coming from 3–5 different independent directions start to interact. These streams have had their angular momentum set at $R \approx 2R_h$, which does not significantly vary during its transport down to the 'messy region'. In this region substantial angular momentum exchange and torquing occurs, which eventually drive the baryons down to $R \lesssim 0.1R_h$. The resulting dynamics is such that the stellar spin parameter of the disc stars, defined as $\lambda_{star} = j_{star}/(\sqrt{2}R_h V_h)$, is well described by a lognormal distribution with $\langle\lambda_{star}\rangle = 0.019$ and dispersion of 0.2 dex.

The results by Danovich et al. (2015) have several implications for our work. First of all, they agree with the estimate that $f_j = j_{star}/j_h \equiv \lambda_{star}/\lambda \approx 0.5$. Secondly, they provide a tighter distribution of angular momenta than that predicted from the MMW model, qualitatively in agreement with observations (Cervantes-Sodi et al. 2013 and Burkert et al. 2016) and with our empirical constraints on the scatter of the size functions of LTGs (although possibly still too large for very massive LTGs and indeed to wide if only pure discs are considered, see Section 4.1.1). Third, recall that our results for the MMW model imply that $R_d \approx 0.3\lambda R_h$. Here the factor $0.3R_h$, is reminiscent of the outer boundary of the 'messy region' seen in Ceverino et al. (2010) and Danovich et al. (2015), where, effectively, the final angular momentum of baryons that will settle in a disc at $R \lesssim 0.1R_h$ originates. On the other hand, in the MMW model the factor $0.3R_h$ boils down to $f_j R_h/\sqrt{2}$ with $f_j \approx 0.5$ constrained in various ways, which is related to the angular momentum 'conserved during the gas collapse'. Thus, the longstanding success of the MMW framework in predicting the normalization of LTG scaling relations (Somerville et al. 2008; Straatman et al. 2017; Lapi et al. 2018b; Marasco et al. 2019 among many others) might be attributed only to the fact that the relevant physics is set at $0.3R_h$, which also regulates the normalization of the MMW model. Note that this is not a matter of semantic, but of the underlying physics. The scenario envisaged in the MMW model is that of a rather smooth formation history. The gas is assumed to be tight to the *overall* spin parameter of dark matter, and to slowly accrete on to the protogalaxy at the centre of the halo. Conversely, the simulations described in Danovich et al. (2015) reveal a quite more violent scenario where the gas is funnelled towards the inner halo in only a few streams with an angular momentum higher than that of dark matter, which is

then lowered by gravitational torques in the 'messy region'. Indeed, the value of $f_j \approx 0.5$ can be understood in the light of these torques. Notably, in the MMW framework $f_j \approx 0.5$ provides a good fit to the mean observed size and angular momentum distributions, but it is not possible to predict it from first principles.

5.3.2 *Shrinking the predicted size distributions: a toy model*

A narrower observed size distribution for LTGs could still be reconciled with the MMW model if we consider that not all the values of the spin parameter are physically acceptable.

Low values of λ are for example disfavoured by the standard disc instability scenario (e.g. Efstathiou, Lake & Negroponte 1982) according to which a disc becomes unstable to its own self-gravity if

$$V_{\rm disc}^2 < \epsilon^2 \frac{GM_{\rm d}}{R_{\rm d}}. \tag{29}$$

Here we will consider discs that are dominated by stars, for which $\epsilon \sim 0.9$ (e.g. Christodoulou, Shlosman & Tohline 1995), so that $M_{\rm d} \approx M_{\rm star}$. Using the definition $M_{\rm star}/f_{\rm b}M_{\rm h} = f_*$[12] and equation (1), the condition above reads

$$\lambda < \lambda_{\rm DI} \equiv \sqrt{2}\frac{\epsilon^2}{f_j}\frac{f_{\rm b}}{f_v} f_* f_{\rm R} f_{\rm c}, \tag{30}$$

where $f_v \approx 1.07 V_{\rm h}/V(3R_d)$ (Lapi et al. 2018b) and $V_{\rm h}$ is the circular velocity of the halo.

On the other hand, it could be envisaged that the high-spin tail of the λ distribution provides a substantial centrifugal barrier that prevents the gas from collapsing and forming stars. In such a scenario, the gas would set in a rotationally supported disc at $R_{\rm rot} > R_{d,\rm gas}$, where $R_{\rm rot}$ is given by solving the following expression:

$$\frac{j_{\rm gas}^2}{R_{\rm rot}^3} = \frac{[GM_{\rm gas}(< R_{\rm rot}) + M_{\rm h}(< R_{\rm rot})]}{R_{\rm rot}^2}. \tag{31}$$

Here $R_{\rm rot}$ is the size of the galaxy that cannot form stars, while $R_{d,\rm gas}$ the size that would be predicted by the MMW model, if the gas could collapse beyond the centrifugal barrier. In the biased collapse scenario, $M_{gas} \equiv M_{\rm inf} = f_{\rm inf} f_b M_{\rm h}$ and $M_{\rm h}(< R_{\rm rot}) \approx 10 M_{\rm h} \left(R_{\rm rot}/R_{\rm h}\right)^2$ (Lapi et al. 2018a) and $j_{\rm gas} = f_{\rm inf} j_{\rm h}$. The solution to the equation above is

$$R_{\rm rot} = 2\lambda^2 \frac{f_{\rm inf}}{f_{\rm b}} R_{\rm h}, \tag{32}$$

and therefore the condition $R_{\rm rot} > R_{d,gas}$ reads

$$\lambda > \lambda_{\rm CB} = \frac{f_{\rm b}}{2\sqrt{2}} \frac{1}{f_{\rm R} f_{\rm c} f_{\rm v}}. \tag{33}$$

Fig. 7 shows the results of this toy model, compared to the observed size function of bulgeless galaxies with $M_{\rm star} < 10^{11}\,{\rm M}_\odot$. In this case, the values of A_λ in Table 1 would need to be increased by 0.1 dex, which could be interpreted as a higher f_j for pure discs, which agrees with Romanowsky & Fall (2012) and Obreschkow & Glazebrook (2014). Overall, it can be seen that our simple framework improves tremendously on the shape of the size function that would be predicted from the MMW model. Indeed, the cuts that we apply are able to tighten the predicted size function in a mass-dependent fashion, as suggested by the mass dependence of σ_K, with disc instabilities being important only for more massive galaxies (De Lucia et al. 2011, however see also Romeo & Mogotsi 2018; Romeo 2020). Moreover, these cuts behave very differently at the low- and high-spin ends respectively, in a way that almost fully recovers the skewness of the observed size functions (see also Section 2).

The power of this very simple toy model is that it is able to shrink not only the predicted size functions, but also the angular momentum distributions, since $j_{\rm gas} \propto \lambda$. The resulting angular momentum distribution would retain the same skewness seen in the size distributions, and future data would be a powerful test for this model.

5.4 Comments on the concentration model

As regards to the concentration model, we have seen (cf. Section 4) that lower values of γ produce wider distributions, and that γ may be tuned to match the size functions without adding any intrinsic scatter $\sigma_{\rm CM}$ in the concentration model. There are two factors at play here. The first is that halo concentration has already an intrinsic scatter that amounts to $\sigma_{logc} \sim 0.11$ dex. The second is that, as it turns out, adopting different values of γ also ends up contributing to the total scatter at different levels (see Section 3.3). We note that in principle some degeneracy may be expected in the contribution to the total scatter from γ and $\sigma_{\rm CM}$. Inspired by the results of Jiang et al. (2019), we discuss what the consequences of having a mass independent $\gamma \sim -0.7$ would be. It is clear from Figs 6(b) and A1(b) that such a value of γ would account for some of the observed width of the size functions. In fact, it can be seen that the scatters produced by $\gamma = -1.6$, -1.2, -0.8 and -0.4 are roughly equivalent to those given by $\sigma_K = 0.20$, 0.15, 0.10, and 0.0. A constant value of $\gamma = -0.7$ from Jiang et al. (2019) would be able to account for ≈ 13 per cent, ≈ 25 per cent and all of the scatter observed for galaxies with $M_{\rm star} < 10^{10}\,{\rm M}_\odot$, $10^{10} < M_{\rm star}/{\rm M}_\odot < 10^{11}$, and $10^{11} < M_{\rm star}/{\rm M}_\odot < 10^{12}$, respectively. Note that if γ is independent of $M_{\rm star}$, then a mass-dependent $\sigma_{\rm CM}$ is expected due to the mass dependence of the width of the size functions. Overall, the concentration model could be favoured due to its lower intrinsic scatter, however its explanation from a theoretical standpoint remains a challenge.

6 THE K13 MODEL IN STATE-OF-THE-ART SIMULATIONS AND THE FORMATION OF ETGS

We now proceed to test whether current cosmological models of galaxy formation are consistent with the semi-empirical constraints outlined in the previous sections, i.e. the existence of a tight relationship $R_{\rm e}$–$R_{\rm h}$, between galaxy size and host halo radius, and a lower normalization in the relationship $R_{\rm e}$–$R_{\rm h}$ relation for ETGs compared to LTG. To this purpose, we will use the Rome semi-analytic model (the Rome SAM hereafter) and the Illustris TNG simulation.

6.1 The Rome SAM

We updated the Rome SAM (described in detail in Menci et al. 2005; Menci et al. 2008; Menci et al. 2014) with a standard prescription for galaxy sizes (Cole et al. 2000) as adopted in many other semi-analytic models (e.g. Guo et al. 2011; Shankar et al. 2013). We refer the reader to the original papers for the details of the SAM. Here we just recall that in the Rome SAM galaxies are initialized as discs following the MMW model, with $\langle\lambda\rangle = 0.035$ and $\sigma_{log\lambda} =$

[12] $f_b \approx 0.16$ is the cosmological baryon fraction, Planck Collaboration XIII (2016).

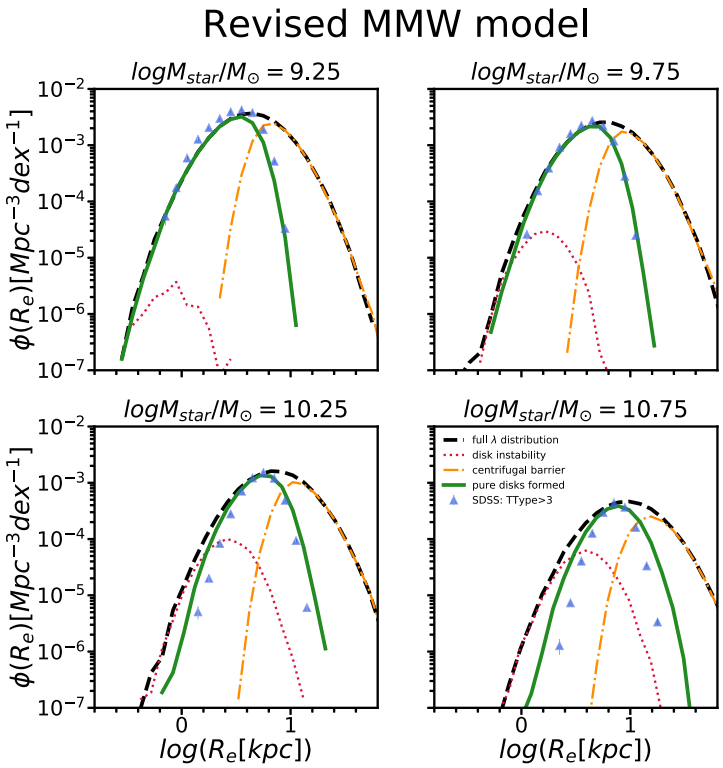

Figure 7. Result of the toy model outlined in Section 5.3.2. The full distribution of λ is indicate with dashed black lines. The contributions to the total size function that would come from haloes with $\lambda < \lambda_{DI}$ and $\lambda > \lambda CB$ are shown in red dotted and orange dash–dotted lines. Only galaxies within the range $\lambda_{DI} < \lambda < \lambda CB$ are able to form (green solid lines).

0.25 dex from dark matter only simulations (e.g. Rodríguez-Puebla et al. 2016). Although in principle both internal torques (i.e. disc instabilities) and galaxy mergers may contribute to the evolution of galaxy sizes, in our SAM we deliberately choose to ignore the former to isolate the sole role of mergers in setting galaxy sizes. In our SAM after a merger the size of the remnant is computed from energy conservation and the virial theorem,

$$\frac{M_{\rm fin}^2}{R_{\rm fin}} = \frac{M_1^2}{R_1} + \frac{M_2^2}{R_2}, \tag{34}$$

where $M_{\rm fin}$, M_1, M_2 and $R_{\rm fin}$, R_1, R_2 are the masses and half-mass radii of the remnant and the merging partners. Here we neglect the term of gravitational interaction between the merging galaxies which corresponds to assuming that all mergers occur on parabolic orbits. We have checked that at this level of the modelling, including the gravitational interaction term mostly impacts the relative normalization of galaxy sizes, but not their distribution, which is the main aim of this work. Major mergers ($M_1/M_2 > 0.3$) are assumed to completely destroy the stellar disc, while minor mergers leave the disc intact and grow the galactic bulge only. The total size of a galaxy is computed as the mass-weighted mean of the disc and bulge radii. During a major merger substantial energy dissipation may occur (Covington et al. 2011), which will modify the size of the remnant as (Shankar et al. 2013)

$$R(\text{dissipation}) = \frac{R(\text{dissipationless})}{1 + f_{\rm gas}/0.2}, \tag{35}$$

where f_{gas} is the gas fraction of the merging pair. We run the Rome SAM with and without the implementation of such process, and we will show that it is not crucial to our conclusions.

In the Rome SAM we classify galaxies according to their B/T ratio, with 'pure discs' being the galaxies with $B/T < 0.3$, and 'pure bulges' those with $B/T > 0.7$. To enable a closer comparison to observations, the sizes of our semi-analytic galaxies are convolved with a measurement error of ∼0.1 dex.

6.2 Illustris TNG

The Illustris TNG project (described in detail in Nelson et al. 2019; Pillepich et al. 2018a, b) is a suite of cosmological simulations run with the same parameters in three boxes with side 50, 100, and 300 h^{-1} Mpc. The simulation builds and improves on the previous Illustris project (Genel et al. 2014; Vogelsberger et al. 2014; for a review of the differences between Illustris and Illustris TNG see Pillepich et al. 2018b), and has proven to achieve a good agreement with observed galaxy sizes (Genel et al. 2018), SMF (Pillepich et al. 2018a), and morphologies (Huertas-Company et al. 2019). Here we use the box of 100 h^{-1} Mpc a side, which is publicly available[13] and described in Nelson et al. (2019). For a complete description of the numerical implementation of Illustris TNG we refer the reader to Springel (2010), Marinacci et al. (2018), Naiman et al. (2018), Springel et al. (2018), Pillepich et al. (2018b), Weinberger et al. (2017), and Nelson et al. (2019).

For IllustrisTNG100, we make use of the catalogue of optical morphologies presented in Huertas-Company et al. (2019) based on STATMORPH (Rodriguez-Gomez et al. 2019), a PYTHON package for calculating non-parametric morphologies of galaxy images as

[13] www.tng-project.org/data/

well as fitting 2D Sérsic profiles.[14] Briefly, Huertas-Company et al. (2019) have selected galaxies with $M_{\rm star} > 10^{9.5}\,{\rm M}_{\odot}$ in the snapshot 95 at $z \sim 0.045$ and processed their images with the radiative transfer code SKIRT (Baes et al. 2011; Camps & Baes 2015). The mock images are then observed in the SDSS *r*-band filter and further realism is added using `RealSim`[15] (Bottrell et al. 2017a, b, 2019). The full procedure is outlined in Rodriguez-Gomez et al. (2019) and Huertas-Company et al. (2019). Morphological information is also available for the mock galaxies. Using the Nair & Abraham (2010) catalogue as training set, Huertas-Company et al. (2019) have trained an ensemble of Convolutional Neural Networks in binary classification mode to distinguish LTGs from ETGs; a finer within-class classification is also available. The mock images from Illustris TNG galaxies are then classified as LTGs or ETGs using the same neural networks. Huertas-Company et al. (2019) have found that the morphologically classified Illustris TNG galaxies follow the same scaling relations of SDSS galaxies almost everywhere. The issue of how exactly the morphologies of simulated galaxies resemble observations is the subject of a forthcoming paper (Zanisi et al., in preparation). We match the catalogue with the SubFind catalogue of Illustris TNG and select central galaxies only. This leaves us with a total of 7222 galaxies.

For each of the IllustrisTNG100 galaxies, we compute the specific angular momentum of the stellar particles:

$$j_{\rm star} = \frac{1}{\sum_n m^{(n)}} \sum_n m^{(n)} \boldsymbol{x}^{(n)} \times \boldsymbol{v}^{(n)}, \tag{36}$$

where $\boldsymbol{v}^{(n)}$ is the velocity of the *n*th particle relative to the centre of mass for the galaxy. $\boldsymbol{x}^{(n)}$ is the position of a given particle with respect to the position of the most gravitationally bound particle in the galaxy. We choose this definition since the centre of mass velocity can be biased by structure at large radii and hence may not represent the true rotational centre. We compute the angular momentum relative to the centre of mass since the rest frame as defined by the most bound particle is often noise dominated. See Duckworth, Tojeiro & Kraljic (2019) for more details.

As for the sizes of IllustrisTNG100 galaxies, we adopt four of the available STATMORPH estimates. First, we use the semimajor axis size of the best Sérsic fit, $R_{\rm e,maj}$, ensuring that the flag `flag_sersic` is equal to zero to include only good photometric fits. We will also show results for R_{50} and R_{80}, the radii of a circular area that encloses 50 per cent and 80 per cent of the light contained in 1.5 times the Petrosian radius, where no prior assumption on the light profile is made. Finally, from the SubFind (Springel et al. 2001) catalogue we extract the physical size $R_{\rm e,3D}$. The correlation between $R_{\rm e,maj}$ and $R_{\rm e,3D}$ is shown in Appendix B2.

6.3 Results

We produce $R_{\rm e}-R_{\rm h}$ relations and their scatter at $z = 0$ for central galaxies only from our SAM (Fig. 8) and the mock observations of IllustrisTNG100 (Fig. 9). At a first glance, we do not see much difference between the $R_{\rm e}-R_{\rm h}$ relation found in TNG100 and in the Rome SAM. It is indeed pleasing that both models predict that ETGs and LTGs lie on two separate relations, in agreement with our semi-empirical constraints. We now discuss the outcomes of the two models in more detail.

In the SAM, the relation of ETGs is offset by ∼0.3 and ∼0.4 dex with respect to that of LTGs in the dissipationless and dissipative scenarios, respectively. The two left-hand panels of Fig. 8 show that using the distribution of spin parameters taken from dark matter only simulations result in a scatter $\sigma_{\rm K} > 0.2$ for both ETGs and LTGs, which is larger than that found by our semi-empirical approach. In the two right-hand panels, instead, we have assumed that the distribution of spin parameters from which LTGs can form is $\sigma_{log\,\lambda} = 0.15$ which, once convolved with measurement uncertainty, is consistent with the upper limits to σ_K that we give in Section 4 (Fig. 6a). In this case, the scatter in the $R_{\rm e}-R_{\rm h}$ relation of ETGs is somewhat reduced, and it becomes consistent with our semi-empirical findings, especially at high values of $R_{\rm h}$. We also note that dissipation does not affect the scatter in either case.

In IllustrisTNG100 we can see that using the semimajor axis size $R_{\rm e,maj}$ gives a scatter that is somewhat larger than the one we find with our semi-empirical model, while the size R_{50} of mock-observed LTGs follows more closely our constraints on the scatter of the $R_{\rm e} - R_{\rm h}$ relation. Indeed, it can be seen that for intermediate values of $R_{\rm h}$ the scatter is just about 0.2 dex, declining with increasing $R_{\rm h}$. However, it seems that for ETGs the scatter is larger than 0.2 dex in both cases. The right-hand panel of Fig. 9 shows that the distribution of physical sizes at fixed halo radius is indeed already of the order of 0.2 dex for ETGs even before the mock observations are performed. On the positive side, we observe that such scatter decreases as $R_{\rm h}$ increases also for ETGs in all cases. In passing, we also note that the relationship between the physical sizes of very large central galaxies and that of their dark matter haloes is extremely tight, of the order of 0.05 dex, in agreement with our constraints (see Fig. A1a).

6.4 Discussion

The difficulty of maintaining a tight scatter in the observed structural scaling relations of ETGs implied by a pure merger scenario has been discussed in, e.g. Nipoti et al. (2009, 2012) (see also discussion in Shankar et al. 2014a). Using our SAM we find that a pure merger scenario, where ETGs are only formed as a consequence of merging of LTGs, requires a very tight distribution for the sizes of LTGs, of the order of ∼0.15 dex, which is supported by the estimate of the intrinsic scatter of the size distributions of star- forming galaxies at $z > 0$ provided in van der Wel et al. (2014). We discuss a comparison with another SAM in Section 7.1.

On the other hand, in hydrodynamical simulations internal torques and mergers arise naturally from the local and global gravitational fields respectively. The implementation of IllustrisTNG100 achieves naturally a tight relation between $R_{\rm e}$ and $R_{\rm h}$ for LTGs but not so for ETGs. We now speculate on the possible reasons behind the success of the simulation in reproducing the semi-empirical trends for LTGs. In Section 5.2.2 we have shown that the MMW model may be consistent with the observed scaling relations of LTGs if the stellar angular momentum, rather than the halo spin parameter, is used. Being consistent with our determination on the $R_{\rm e}-R_{\rm h}$ relation for LTGs, Illustris TNG offers an ideal testbed for this hypothesis. We recall that empirically, and using the MMW model, our argument would predict that at fixed stellar mass and halo radius the scatter in $R_{\rm e}$ should be completely driven by that in $j_{\rm star}$. Therefore, a variance of about 0.2 dex in the distribution of $j_{\rm star}$ at fixed $M_{\rm star}$ and $R_{\rm h}$ would support our argument.

In Fig. 10 we show the relationship between the stellar angular momentum $j_{\rm star}$ and stellar mass in IllustrisTNG100 in bins of $R_{\rm h}$.

[14] Available at: https://statmorph.readthedocs.io.
[15] Available at: https://github.com/cbottrell/RealSim.

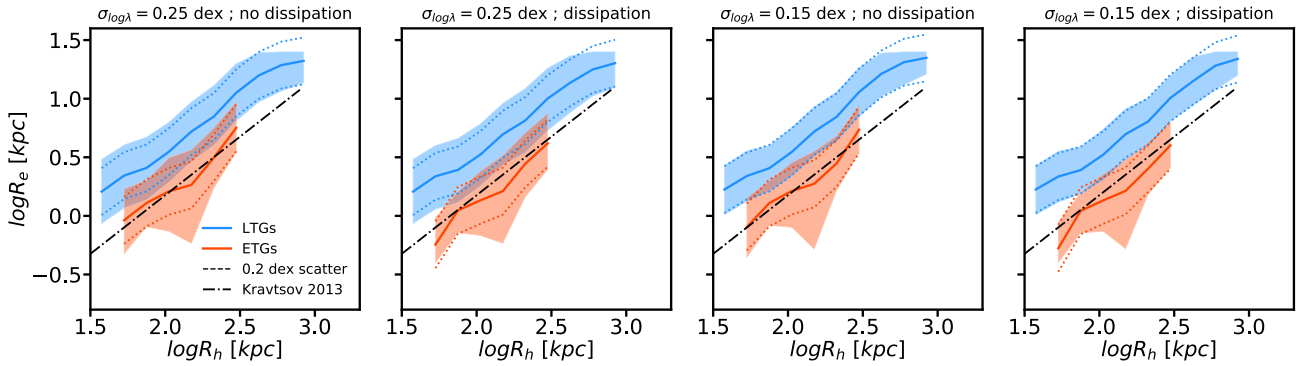

Figure 8. $R_e - R_h$ relation in the Rome SAM. Each panel represents a run of the model where $\sigma_{log\lambda}$ is varied or dissipation is included, as labelled. The red and blue lines are for LTGs and ETGs respectively, while the cyan and salmon shaded areas indicate the 16th and 84th percentiles of the distributions at fixed R_h. Dashed lines show a scatter of 0.2 dex from the mean, consistent with the upper limit provided our semi-empirical model. The relation by Kravtsov (2013) is shown as dot–dashed lines for comparison. The predicted R_e are convolved with an observational scatter of 0.1 dex.

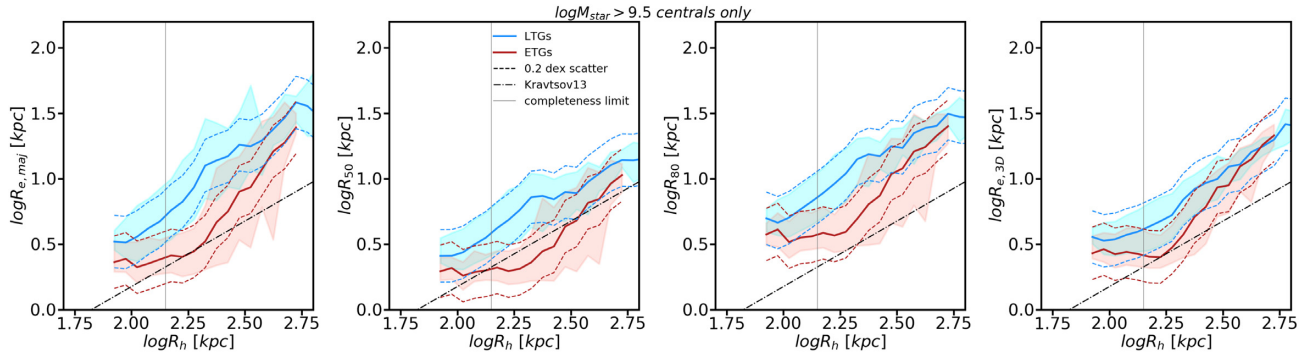

Figure 9. *First panel:* STATMORPH Sérsic semimajor axis sizes of the mock observed Illustris TNG galaxies as a function of R_h. *Second panel:* STATMORPH estimates for R_{50} of the mock observed Illustris TNG galaxies as a function of R_h. *Third panel:* STATMORPH estimates for R_{80} of the mock observed Illustris TNG galaxies as a function of R_h. *Fourth panel:* Physical 3D radius $R_{e,3D}$ of the same Illutris TNG galaxies as a function of R_h. Red and blue lines are for LTGs and ETGs, respectively, while the cyan and salmon shaded areas indicate the 16th and 84th percentiles of the distributions at fixed R_h. Dashed lines show a scatter of 0.2 dex from the mean, consistent with the upper limit provided our semi-empirical model. The relation by Kravtsov (2013) is shown as dot–dashed lines for comparison. The completeness limit on R_h induced by the stellar mass cut is shown as a vertical grey line. The difference between the left-hand and right-hand panels may be understood in the light of Fig. B2.

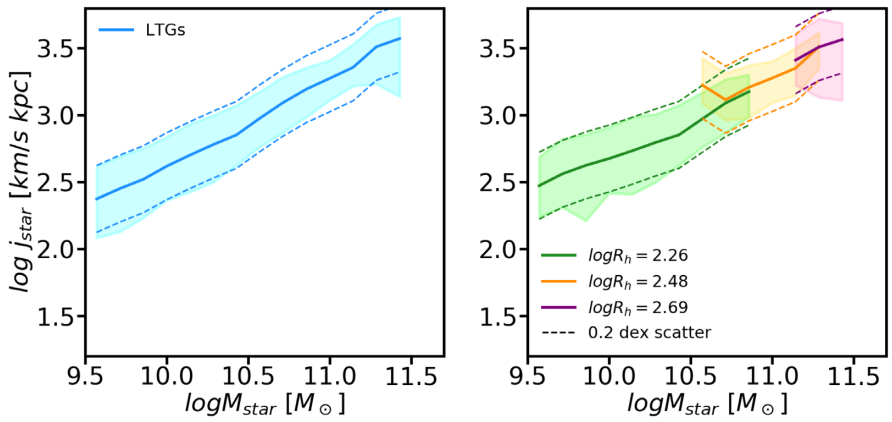

Figure 10. The relationship between stellar angular momentum and stellar mass for Illustris TNG LTGs (left), binned in three ranges of halo radius (right).

We also show the relation for all LTGs since two highest bins in R_h suffer from low number statistics. It can be seen that the predicted scatter is about 0.2 dex and decreasing with increasing stellar mass and bin of R_h. This is consistent with our argument, and also with the decrease in scatter in the $R_e - R_h$ relation at high halo radii. We test more directly the connection between galaxy size and stellar angular momentum in Fig. 11, where we show the size functions of IllustrisTNG100 LTGs in narrow bins of j_{star}.

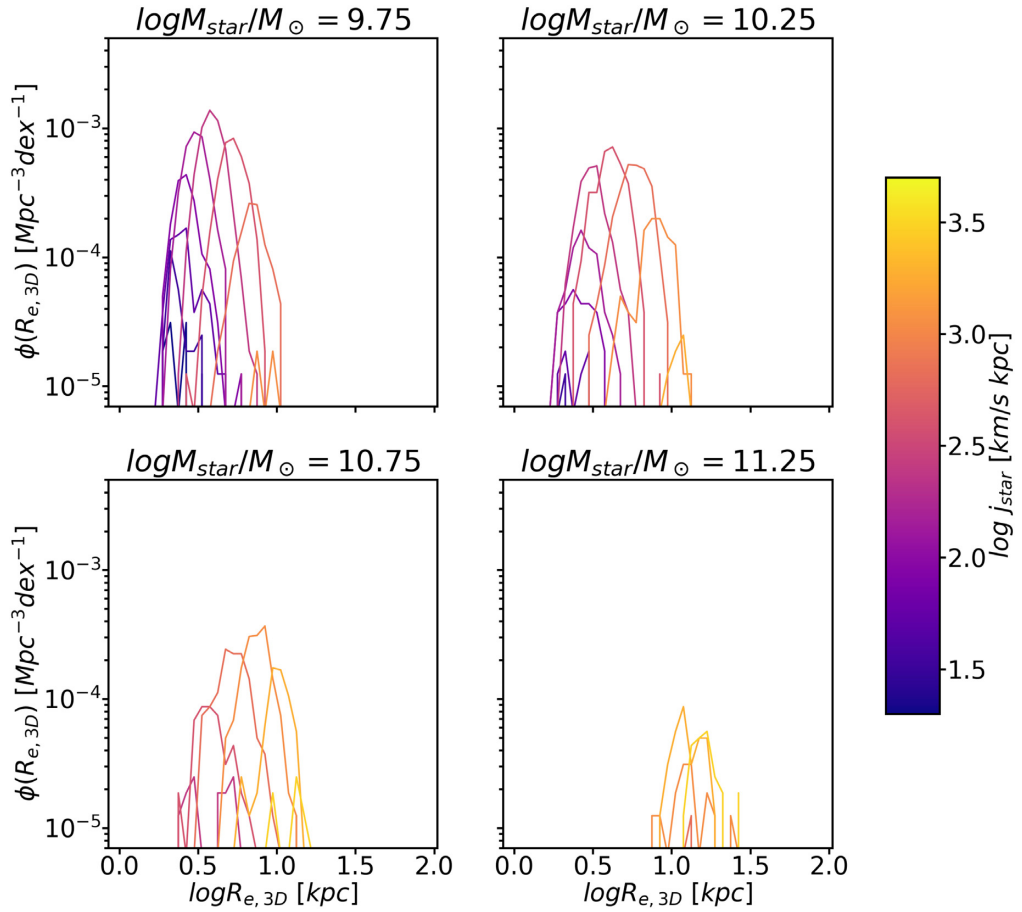

Figure 11. Size functions of IllustrisTNG100 galaxies in bins of stellar mass and colour coded by the value of specific stellar angular momentum. The tightness of the size functions at fixed $j_{\rm star}$ is remarkable.

The first striking feature of Fig. 11 is that in a given bin of $M_{\rm star}$ larger galaxies have a larger specific stellar angular momentum. Even more remarkable is the fact that the tightness of the size functions[16] is extraordinarily narrow at fixed $j_{\rm star}$, with a scatter of the order of $\lesssim$0.1 dex. These findings suggest that the link between galaxy sizes and their stellar angular momentum is extremely tight. We therefore advocate that an empirically motivated model where the relationship between $R_{\rm e}$ and $R_{\rm h}$ is mediated by stellar angular momentum seems to be supported by our analysis of IllustrisTNG100.

To conclude, we briefly note that mechanical feedback from the AGN may also 'puff-up' the galaxy structure (Fan et al. 2008; Fan et al. 2010; Ragone-Figueroa & Granato 2011), which may be critical to decrease the scatter in the ETGs scaling relations (Lapi et al. 2018a). This is not included in the Rome SAM, but it is modelled in Illustris TNG (Weinberger et al. 2017; Pillepich et al. 2018b). It seems however that both the purely hierarchical scenario adopted in the version of the Rome SAM where $\sigma_{log\lambda} = 0.15$ dex and IllustrisTNG100 follow our semi-empirical constraints, which makes it difficult to disentangle the effects of mergers and AGN puffing-up here.

[16]Here we use $R_{\rm e,3D}$ since we want to investigate the intrinsic relationship between size and angular momentum.

7 DISCUSSION AND COMPARISON TO PREVIOUS WORK

7.1 Comparison to other semi-analytic models

The constraints that we give on the $R_{\rm e}-R_{\rm h}$ connection with our semi-empirical model stem from single-epoch abundance matching. As such, we are unable to follow galaxies during their formation history in an evolutionary context. In particular, the MMW model is directly applied to dark matter haloes at $z \sim 0$, while disc galaxies are likely to have grown their stellar mass steadily in time during the last 10 Gyr or so (e.g. Patel et al. 2013). The implementation of the Rome SAM is such that whenever gas cooling occurs, the galaxy is assigned a size according the MMW model. Admittedly, this model is not able to capture the inside-out growth of galaxy discs (e.g. Sánchez et al. 2018). Recently, Zoldan et al. (2018, 2019) have presented the result of a semi-analytic model (based on De Lucia et al. 2014; Hirschmann, De Lucia & Fontanot 2016; Xie et al. 2017) in which stellar angular momentum and galaxy sizes are evolved self-consistently. In their model it is assumed that the stellar angular momentum is built up gradually as star formation proceeds and depletes the gas discs. The size of the stellar disc is then computed at each step assuming that a close analogue of equation (28) holds. The Zoldan et al. model is built on the assumption that angular momentum is conserved during both star formation and disc instabilities, which are also

included, contrary to the Rome SAM. Their studies point out that mergers are the primary drivers of the observed LTGs–ETGs dichotomy in the $j_{\rm star}-M_{\rm star}$ relation. This qualitatively agrees with our result that bulge growth via mergers may lead to the same bimodality but in the $R_{\rm e}-R_{\rm h}$ plane. These authors also obtain a tight scatter in the $R_{\rm e}-R_{\rm h}$ relation for ETGs, as shown in their fig. 7. Whether such scatter is consistent with our constraints remains to be seen.

7.2 Comparison to other semi-empirical models

Using a semi-empirical technique similar to ours, Somerville et al. (2018) found that the *total* size distributions observed in GAMA and CANDELS are in agreement with the MMW model. These findings are suggestive that both the population of ETGs and LTGs may be described in the MMW framework. For example, ETGs could be formed in dark matter haloes with preferentially lower λ, which would account for the fact that the distribution of ETGs is peaked at lower $R_{\rm e}$ than that of LTGs. However, this model would not be able to explain the observed angular momenta of ETGs, as shown by Romanowsky & Fall (2012) and Posti et al. (2018a). Alternatively, one could note that the normalization of the MMW model bears the dependence on the fraction of the halo angular momentum f_j that was retained by the collapsing gas, since $A_\lambda \propto f_j$ (see Section 3). In principle ETGs and LTGs could then be two populations that retained preferentially lower and higher f_j respectively but that did not form in haloes with different value of λ. Such a scenario may also be able explain why ETGs always have smaller sizes than LTGs. However, although ETGs are sometimes envisioned to form mostly in situ (e.g. Shi et al. 2017; Lapi et al. 2018a), it is often suggested that they have likely undergone merger events, which may have led to lower f_j on average (Romanowsky & Fall 2012). We would thus be cautious in interpreting f_j for ETGs in the context of the MMW model, at least at low redshift. Indeed, we have shown that a purely hierarchical model is able to produce smaller sizes for ETGs, while preserving the linearity of the MMW model. As a side note, we recall that the total size function shown in Fig. 2 is wider than those of ETGs and LTGs taken individually and therefore it might well be that the agreement between the MMW model and the total size function found by Somerville et al. (2018) occurs only by chance.

Another possible explanation for the difference in the normalization of the $R_{\rm e}-R_{\rm h}$ relation is that the size of a galaxy is more tightly bound to that of its halo at the redshift of formation than to the size of the halo at the time of observation. In particular, given the older ages of ETGs (e.g. Bernardi et al. 2010), they must have formed at high redshift where haloes were smaller (see equation 2). The late evolution of ETGs, which seems to be dominated by minor dry mergers (e.g. Oser et al. 2010; Shankar et al. 2013), will however modify the $R_{\rm e}-R_{\rm h}$ relation on to which ETGs formed. Unfortunately, with the semi-empirical model used in this work we are able to constrain only the present-day relation, and therefore we cannot directly infer any information about the formation of ETGs.

7.3 Using R_{80} instead of $R_{\rm e}$

In recent work (Miller et al. 2019; Mowla et al. 2019) it has been proposed to use as proxy for galaxy size R_{80}, the size that encloses 80 per cent of the light, rather than the half-light radius $R_{\rm e}$. This suggestion has been made on the grounds that: (i) the sizes of passive and star-forming galaxies tend to collapse on the same size–stellar mass relation in the case where R_{80} is used (Miller et al. 2019); (ii) R_{80} is more closely linked to the size of the host dark matter halo $R_{\rm e}$ (Mowla et al. 2019).

Fig. 2 shows a comparison between the size functions computed for $R_{\rm e}$ and R_{80} for both ETGs and LTGs. We observe that the difference in the size functions of ETGs and LTGs computed using $R_{\rm e}$ is only slightly reduced when using R_{80}. While such difference appears to be somewhat more pronounced at $M_{\rm star} > 10^{11}\,{\rm M}_\odot$, the bimodality of the size functions $\phi(R_{\rm e})$ seems to be substantially conserved also for $\phi(R_{80})$ at lower masses. It is also noteworthy that the scatter of the individual size functions is not affected by the choice of the definition of galaxy size. Therefore, adopting R_{80} rather than $R_{\rm e}$ in our work would only require an overall higher normalization for the $R_{\rm e}-R_{\rm h}$ relations studied here, but the results for the implied scatters remain robust. In particular, such result would not undermine our empirical model where galaxy and halo sizes are mediated by stellar angular momenta.

The discussion above is in agreement with the fact that the STATMORPH estimate of R_{80} for the mock-observed IllustrisTNG100 galaxies entails a similar scatter in the galaxy size-halo size relation of LTGs and ETGs compared to that of R_{50} (see central panel of Fig. 9). Moreover, the relations for the two morphological classes keep being separated also in the $R_{80}-R_{\rm h}$ plane also in the case of IllustrisTNG100.

We stress that our analysis of SDSS makes use of a mix of `Sérsic` and `Sérsic + Exponential` fits, contrary to Mowla et al. (2019) and Miller et al. (2019) where only `Sérsic` profiles are assumed, while in IllustrisTNG100 R_{80} is the size of a region that contains 80 per cent of the light inside an area of 1.5 the Petrosian radius. We refer the reader to Bernardi et al. (2013, 2014, 2017b) for a detailed discussion of the implications of using different fits to photometric light profiles.

8 CONCLUSIONS

In this work we have used a semi-empirical approach to study three models of galaxy sizes, where the sizes of galaxies are linked to that of their haloes by means of the dynamical (the MMW model, equation 1) or structural (the concentration model, equation 19) properties of the dark matter halo in which they are hosted, or by a simple constant (the K13 model, equation 18) the origin of which is a priori unknown.

Our results can be summarized as follows:

(i) The scatter in the K13 model must decrease for more massive galaxies, irrespective of galaxy morphology. This implies that most of the information on the size distributions of the most massive galaxies is fully dependent on the shape of the SMHM and hence on the physical processes that determine it.

(ii) In the concentration model we find that γ is degenerate with the model intrinsic scatter $\sigma_{\rm CM}$. This suggests that a lower $\sigma_{\rm CM}$ may be needed to account for the width of the size functions, and that γ must be low for massive galaxies. A lower $\sigma_{\rm CM}$ might make the concentration model more fundamental than any other model studied here, however its physical origin remains unclear.

(iii) Similarly to other studies (Huang et al. 2017; Lapi et al. 2018b) we find that the normalization of both the K13 model and concentration model must be different for ETGs and LTGs.

(iv) The classical disc model by MMW taken at face value overestimates the tails of the size and angular momentum distributions of disc galaxies, but is able to predict the correct normalizations of

the scaling relations for LTGs. We discuss two scenarios that bring the model in better agreement with data:

(a) We outline a model where only some values of λ are physically acceptable. This model reproduces well the skewness and tightness of the size functions of LTGs.

(b) Based on our constraints from the K13 model, we discuss a scenario where the link between the sizes of LTGs and their dark matter haloes is mediated by the stellar angular momentum, and where the halo spin parameter may not play any major role.

We also investigate whether our empirical constraints are reproduced in current cosmological models of galaxy formation and evolution.

(i) In the Rome SAM, which implements a purely hierarchical scenario where the MMW model is taken at face value, we find that mergers of LTGs alone are able to reproduce the dichotomy of the R_e-R_h relation, but overestimate its scatter. We show that with a tighter scatter in the LTGs R_e-R_h relation it could be possible to lower the inferred scatter in the sizes of ETGs at fixed halo radius to meet our semi-empirical constraints.

(ii) In IllustrisTNG100, where both mergers and internal torques are at work, the morphological segregation in the R_e-R_h plane is also present, with a scatter which is within the empirical constraints given in this work for LTGs, and somewhat higher for ETGs.

(iii) We exploit the information about the dynamics available from IllustrisTNG100 to show that the scatter of the galaxy size–halo size connection of LTGs is consistent with being driven by the stellar specific angular momentum, which corroborates our empirical model based on the MMW model and the scatter of the K13 model.

ACKNOWLEDGEMENTS

We thank the referee for a constructive report that significantly improved this paper. L.Z. acknowledges support from the Science and Technologies Facilities Council and from the Royal Astronomical Society. F.S. acknowledges partial support from a Leverhulme Trust Research Fellowship and from the European Union's Horizon 2020 programme under the AHEAD project (grant agreement 654215). A.L. acknowledges PRIN MIUR 2017 prot. 20173ML3WW002, Opening the ALMA window on the cosmic evolution of gas, stars, and supermassive black holes. L.Z. wishes to thank Avishai Dekel, Mauro Giavalisco, Helena Domínguez-Sanchez, Fangzhou Jiang, Gabriella De Lucia, Anna Zoldan, Vladimir Avila-Reese, and Gigi Danese for helpful discussions and Dylan Nelson for guidance with the Illustris TNG simulation database.

REFERENCES

Abazajian K. N. et al., 2009, ApJS, 182, 543
Aquino-Ortíz E. et al., 2018, MNRAS, 479, 2133
Baes M., Verstappen J., De Looze I., Fritz J., Saftly W., Vidal Pérez E., Stalevski M., Valcke S., 2011, ApJS, 196, 22
Behroozi P. S., Wechsler R. H., Conroy C., 2013, ApJ, 770, 57
Behroozi P., Wechsler R., Hearin A., Conroy C., 2019, MNRAS, 488, 3143
Bernardi M., Shankar F., Hyde J. B., Mei S., Marulli F., Sheth R. K., 2010, MNRAS, 404, 2087
Bernardi M., Meert A., Sheth R. K., Vikram V., Huertas-Company M., Mei S., Shankar F., 2013, MNRAS, 436, 697
Bernardi M., Meert A., Vikram V., Huertas-Company M., Mei S., Shankar F., Sheth R. K., 2014, MNRAS, 443, 874
Bernardi M., Meert A., Sheth R. K., Fischer J. L., Huertas-Company M., Maraston C., Shankar F., Vikram V., 2017a, MNRAS, 467, 2217
Bernardi M., Fischer J. L., Sheth R. K., Meert A., Huertas-Company M., Shankar F., Vikram V., 2017b, MNRAS, 468, 2569
Blumenthal G. R., Faber S. M., Flores R., Primack J. R., 1986, ApJ, 301, 27
Bottrell C., Torrey P., Simard L., Ellison S. L., 2017a, MNRAS, 467, 1033
Bottrell C., Torrey P., Simard L., Ellison S. L., 2017b, MNRAS, 467, 2879
Bottrell C. et al., 2019, MNRAS, 490, 5390
Bryan G. L., Norman M. L., 1998, ApJ, 495, 80
Bullock J. S., Dekel A., Kolatt T. S., Kravtsov A. V., Klypin A. A., Porciani C., Primack J. R., 2001, ApJ, 555, 240
Burkert A. et al., 2016, ApJ, 826, 214
Camps P., Baes M., 2015, Astron. Comput., 9, 20
Cappellari M., 2016, ARA&A, 54, 597
Cervantes-Sodi B., Li C., Park C., Wang L., 2013, ApJ, 775, 19
Ceverino D., Dekel A., Bournaud F., 2010, MNRAS, 404, 2151
Ceverino D., Klypin A., Klimek E. S., Trujillo-Gomez S., Churchill C. W., Primack J., Dekel A., 2014, MNRAS, 442, 1545
Christodoulou D. M., Shlosman I., Tohline J. E., 1995, ApJ, 443, 551
Cole S., Lacey C. G., Baugh C. M., Frenk C. S., 2000, MNRAS, 319, 168
Covington M. D., Primack J. R., Porter L. A., Croton D. J., Somerville R. S., Dekel A., 2011, MNRAS, 415, 3135
Danovich M., Dekel A., Hahn O., Ceverino D., Primack J., 2015, MNRAS, 449, 2087
De Lucia G., Fontanot F., Wilman D., Monaco P., 2011, MNRAS, 414, 1439
De Lucia G., Tornatore L., Frenk C. S., Helmi A., Navarro J. F., White S. D. M., 2014, MNRAS, 445, 970
Desmond H., 2017a, MNRAS, 464, 4160
Desmond H., 2017b, MNRAS, 464, 4160
Desmond H., Wechsler R. H., 2015a, MNRAS, 454, 322
Desmond H., Wechsler R. H., 2015b, MNRAS, 454, 322
Desmond H., Wechsler R. H., 2017, MNRAS, 465, 820
Desmond H., Mao Y.-Y., Wechsler R. H., Crain R. A., Schaye J., 2017, MNRAS, 471, L11
Desmond H., Katz H., Lelli F., McGaugh S., 2019, MNRAS, 484, 239
Diemer B., 2018, ApJS, 239, 35
Diemer B., More S., Kravtsov A. V., 2013, ApJ, 766, 25
Domínguez Sánchez H., Huertas-Company M., Bernardi M., Tuccillo D., Fischer J. L., 2018, MNRAS, 476, 3661
Driver S. P. et al., 2011, MNRAS, 413, 971
Duckworth C., Tojeiro R., Kraljic K., 2019, preprint (arXiv:1910.10744)
Dutton A. A., Macciò A. V., 2014, MNRAS, 441, 3359
Dutton A. A., Conroy C., van den Bosch F. C., Prada F., More S., 2010, MNRAS, 407, 2
Efstathiou G., Lake G., Negroponte J., 1982, MNRAS, 199, 1069
Erfanianfar G. et al., 2019, A&A, 631, A175
Fall S. M., 1983, in Athanassoula E., ed., Proc. IAU Symp. 100, Internal Kinematics and Dynamics of Galaxies. Reidel, Dordrecht, p. 391
Fall S. M., Efstathiou G., 1980, MNRAS, 193, 189
Fan L., Lapi A., De Zotti G., Danese L., 2008, ApJ, 689, L101
Fan L., Lapi A., Bressan A., Bernardi M., De Zotti G., Danese L., 2010, ApJ, 718, 1460
Faucher-Giguère C.-A., Kereš D., Ma C.-P., 2011, MNRAS, 417, 2982
Fischer J. L., Bernardi M., Meert A., 2017, MNRAS, 467, 490
Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, PASP, 125, 306
Genel S. et al., 2014, MNRAS, 445, 175
Genel S. et al., 2018, MNRAS, 474, 3976
Gnedin O. Y., Ceverino D., Gnedin N. Y., Klypin A. A., Kravtsov A. V., Levine R., Nagai D., Yepes G., 2011, preprint (arXiv:1108.5736)
Grylls P. J., Shankar F., Zanisi L., Bernardi M., 2019, MNRAS, 483, 2506
Guo Q. et al., 2011, MNRAS, 413, 101
Hearin A., Behroozi P., Kravtsov A., Moster B., 2019, MNRAS, 489, 1805
Hearin A. P. et al., 2017, AJ, 154, 190
Hirschmann M., De Lucia G., Fontanot F., 2016, MNRAS, 461, 1760
Huang K.-H. et al., 2017, ApJ, 838, 6
Hubble E. P., 1926, ApJ, 64
Huertas-Company M. et al., 2019, MNRAS, 489, 1859
Jiang F. et al., 2019, MNRAS, 488, 401

Kassin S. A., Devriendt J., Fall S. M., de Jong R. S., Allgood B., Primack J. R., 2012, MNRAS, 424, 502
Klypin A. A., Trujillo-Gomez S., Primack J., 2011, ApJ, 740, 102
Klypin A., Yepes G., Gottlöber S., Prada F., Heß S., 2016, MNRAS, 457, 4340
Koekemoer A. M. et al., 2011, ApJS, 197, 36
Komatsu E. et al., 2011, ApJS, 192, 18
Kormendy J., 2016, in Laurikainen E., Peletier R., Gadotti D., eds, Astrophysics and Space Science Library, Vol. 418, Galactic Bulges. Springer International Publishing, Switzerland, p. 431
Kravtsov A. V., 2013, ApJ, 764, L31 (K13)
Kravtsov A. V., Vikhlinin A. A., Meshcheryakov A. V., 2018, Astron. Lett., 44, 8
Lange R. et al., 2015, MNRAS, 447, 2603
Lapi A. et al., 2018a, ApJ, 857, 22
Lapi A., Salucci P., Danese L., 2018b, ApJ, 859, 2
Lelli F., McGaugh S. S., Schombert J. M., 2016, AJ, 152, 157
Leroy A. K., Walter F., Brinks E., Bigiel F., de Blok W. J. G., Madore B., Thornley M. D., 2008, AJ, 136, 2782
Liske J. et al., 2015, MNRAS, 452, 2087
Marasco A., Fraternali F., Posti L., Ijtsma M., Di Teodoro E. M., Oosterloo T., 2019, A&A, 621, L6
Marinacci F. et al., 2018, MNRAS, 480, 5113
Matthee J., Schaye J., Crain R. A., Schaller M., Bower R., Theuns T., 2017, MNRAS, 465, 2381
Meert A., Vikram V., Bernardi M., 2015, MNRAS, 446, 3943
Meert A., Vikram V., Bernardi M., 2016, MNRAS, 455, 2440
Menci N., Fontana A., Giallongo E., Salimbeni S., 2005, ApJ, 632, 49
Menci N., Fiore F., Puccetti S., Cavaliere A., 2008, ApJ, 686, 219
Menci N., Gatti M., Fiore F., Lamastra A., 2014, A&A, 569, A37
Mendel J. T., Simard L., Palmer M., Ellison S. L., Patton D. R., 2014, ApJS, 210, 3
Mo H. J., Mao S., White S. D. M., 1998, MNRAS, 295, 319
Miller T. B., van Dokkum P., Mowla L., van der Wel A., 2019, ApJ, 872, L14
More S., van den Bosch F. C., Cacciato M., Skibba R., Mo H. J., Yang X., 2011, MNRAS, 410, 210
Moster B. P., Naab T., White S. D. M., 2013, MNRAS, 428, 3121
Moster B. P., Naab T., White S. D. M., 2018, MNRAS, 477, 1822
Mowla L., van der Wel A., van Dokkum P., Miller T. B., 2019, ApJ, 872, L13
Naiman J. P. et al., 2018, MNRAS, 477, 1206
Nair P. B., Abraham R. G., 2010, ApJS, 186, 427
Navarro J. F., Frenk C. S., White S. D. M., 1996, ApJ, 462, 563
Nelson D. et al., 2019, ComAC, 6, 2
Nipoti C., Treu T., Auger M. W., Bolton A. S., 2009, ApJ, 706, L86
Nipoti C., Treu T., Leauthaud A., Bundy K., Newman A. B., Auger M. W., 2012, MNRAS, 422, 1714
Obreschkow D., Glazebrook K., 2014, ApJ, 784, 26
Oser L., Ostriker J. P., Naab T., Johansson P. H., Burkert A., 2010, ApJ, 725, 2312
Patel S. G. et al., 2013, ApJ, 778, 115
Peebles P. J. E., 1969, ApJ, 155, 393
Pillepich A. et al., 2018a, MNRAS, 473, 4077
Pillepich A. et al., 2018b, MNRAS, 473, 4077
Planck Collaboration XIII, 2016, A&A, 594, A13
Posti L., Pezzulli G., Fraternali F., Di Teodoro E. M., 2018a, MNRAS, 475, 232
Posti L., Fraternali F., Di Teodoro E. M., Pezzulli G., 2018b, A&A, 612, L6
Prada F., Klypin A. A., Cuesta A. J., Betancort-Rijo J. E., Primack J., 2012, MNRAS, 423, 3018
Ragone-Figueroa C., Granato G. L., 2011, MNRAS, 414, 3690
Rodriguez-Gomez V. et al., 2019, MNRAS, 483, 4140
Rodríguez-Puebla A., Avila-Reese V., Yang X., Foucaud S., Drory N., Jing Y. P., 2015, ApJ, 799, 130
Rodríguez-Puebla A., Behroozi P., Primack J., Klypin A., Lee C., Hellinger D., 2016, MNRAS, 462, 893
Rodríguez-Puebla A., Primack J. R., Avila-Reese V., Faber S. M., 2017, MNRAS, 470, 651
Romanowsky A. J., Fall S. M., 2012, ApJS, 203, 17
Romeo A. B., 2020, MNRAS, 491, 4843
Romeo A. B., Mogotsi K. M., 2018, MNRAS, 480, L23
Sánchez S. F. et al., 2012, A&A, 538, A8
Sánchez S. F. et al., 2018, RMxAA, 54, 217
Schaye J. et al., 2015, MNRAS, 446, 521
Shankar F., Lapi A., Salucci P., De Zotti G., Danese L., 2006, ApJ, 643, 14
Shankar F., Marulli F., Bernardi M., Boylan-Kolchin M., Dai X., Khochfar S., 2010, MNRAS, 405, 948
Shankar F., Marulli F., Bernardi M., Mei S., Meert A., Vikram V., 2013, MNRAS, 428, 109
Shankar F. et al., 2014a, MNRAS, 439, 3189
Shankar F. et al., 2014b, ApJ, 797, L27
Shankar F. et al., 2017, ApJ, 840, 34
Shen S., Mo H. J., White S. D. M., Blanton M. R., Kauffmann G., Voges W., Brinkmann J., Csabai I., 2003, MNRAS, 343, 978
Shi J., Lapi A., Mancuso C., Wang H., Danese L., 2017, ApJ, 843, 105
Somerville R. S., Davé R., 2015, ARA&A, 53, 51
Somerville R. S. et al., 2008, ApJ, 672, 776
Somerville R. S. et al., 2018, MNRAS, 473, 2714
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
Straatman C. M. S. et al., 2017, ApJ, 839, 57
Tinker J., Kravtsov A. V., Klypin A., Abazajian K., Warren M., Yepes G., Gottlöber S., Holz D. E., 2008, ApJ, 688, 709
Tinker J. L. et al., 2017, ApJ, 839, 121
Vale A., Ostriker J. P., 2006, MNRAS, 371, 1173
van der Wel A. et al., 2014, ApJ, 788, 28
Vogelsberger M. et al., 2014, MNRAS, 444, 1518
Wang L., Dutton A. A., Stinson G. S., Macciò A. V., Penzo C., Kang X., Keller B. W., Wadsley J., 2015, MNRAS, 454, 83
Wechsler R. H., Tinker J. L., 2018, ARA&A, 56, 435
Weinberger R. et al., 2017, MNRAS, 465, 3291
Wojtak R., Mamon G. A., 2013, MNRAS, 428, 2407
Xie L., De Lucia G., Hirschmann M., Fontanot F., Zoldan A., 2017, MNRAS, 469, 968
Yang X., Mo H. J., van den Bosch F. C., Pasquali A., Li C., Barden M., 2007, ApJ, 671, 153
Zoldan A., De Lucia G., Xie L., Fontanot F., Hirschmann M., 2018, MNRAS, 481, 1376
Zoldan A., De Lucia G., Xie L., Fontanot F., Hirschmann M., 2019, MNRAS, 487, 5649
Zolotov A. et al., 2015, MNRAS, 450, 2327

APPENDIX A: EARLY-TYPE GALAXIES

Figs A1(a) and (b) show a comparison between data and the size functions $\phi(R_e)$ from our models (K13 and concentration models, respectively). We bin both our model galaxies and data in bins of 0.5 dex in stellar mass. In all the figures the model size functions are shifted to match the peaks of the observed distributions. The normalization of the different models in each stellar mass bin is reported in Table 1. The values in Table A1 are not meant to be best fits, rather they only indicate that there is trend with stellar mass.

Similarly to what we did for LTGs, we report models for $\sigma_K = 0.10, 0.15, 0.20$ and $\gamma = -1.6, -1.2, -0.8, -0.4$. In each panel of Figs A1(b) and (a) we highlight with a thicker line the parameter that seems to best reproduce observations, keeping in mind that this should not be considered as a fit to data. Qualitatively, we find that ETGs obey relations that are very similar to those of LTGs in terms of σ_K and γ, with higher (lower) σ_K (γ) for lower (higher) stellar masses.

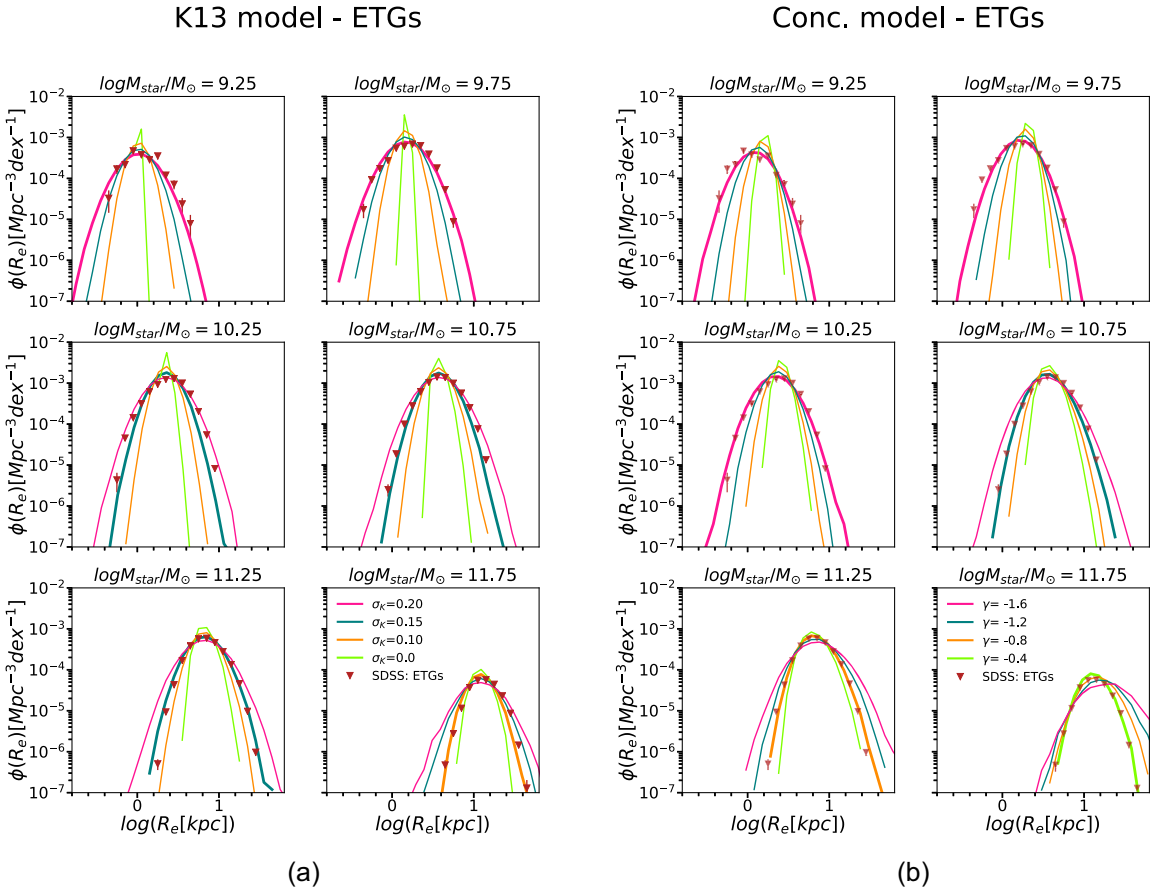

Figure A1. *Left*: Size functions from the K13 model (equation 18) for values of $\sigma_K = 0.00, 0.10, 0.15, 0.20$. *Right*: Size functions from the concentration model (equation 19) for values of $\gamma = -1.6, -1.2, -0.8, -0.4$. Models that work best for a given stellar mass bin are highlighted in each panel by a thicker line. Data points are ETGs from the photo + morphological SDSS catalogues described in Section 2.

Table A1. Values of A_k and A_c in different bins of $M_{\rm star}$, for ETGs. Compare to Table 1.

$M_{\rm star}$	9.25	9.75	10.25	10.75	11.25	11.75
A_k	0.006	0.007	0.010	0.011	0.015	0.016
A_c	0.012	0.013	0.013	0.013	0.013	0.013

APPENDIX B: CAVEATS

B1 Model assumptions

The backbone of our work is the SMHM and all our results depend on it. In Fig. B1 we compare our SMHM fitted to reproduce the SDSS M15/16 SMF to that of Rodríguez-Puebla et al. (2017) and of Dutton et al. (2010) for LTGs at $z \sim 0.1$. Notably, the high mass slope of our SMHM is much steeper than that of Rodríguez-Puebla et al. (2017). Our estimate of the SMHM agrees with other studies where an improved photometry was used (Shankar et al. 2017; Grylls et al. 2019). With a flatter high mass end slope in the SMHM the halo occupation distribution of massive galaxies would be wider (see Section 3.3).

It can also be seen that the difference between our SMHM and that of Dutton et al. (2010) for LTGs is not critical. In the Dutton et al. (2010) SMHM LTGs tend to live in haloes ≈0.2 dex less massive

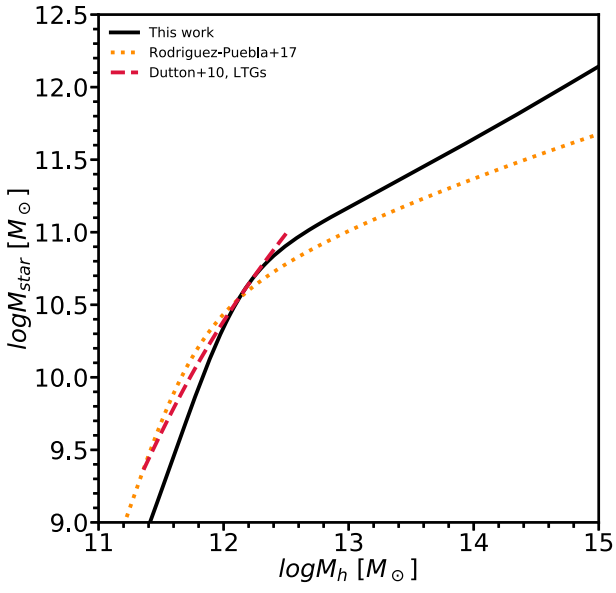

Figure B1. The SMHM fit to reproduce the SDSS SMF adopted in our work compared to that by Rodríguez-Puebla et al. (2017) and Dutton et al. (2010) for LTGs at $z \sim 0.1$.

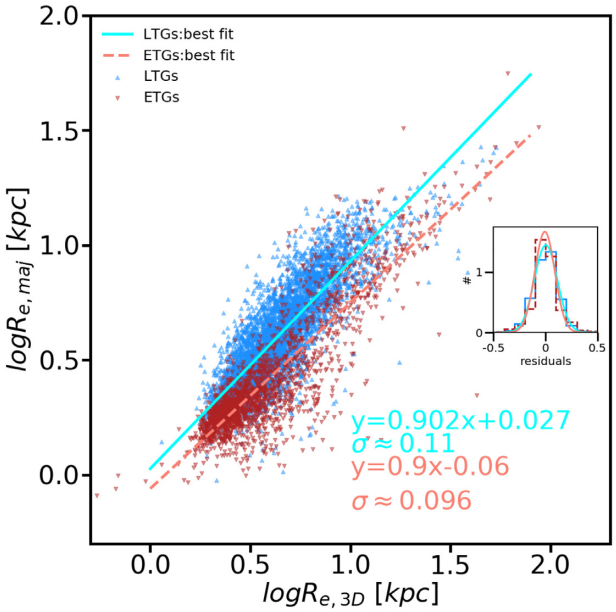

Figure B2. Correlation between 3D physical size and the semimajor axis sizes from `statmorph` (Huertas-Company et al. 2019; Rodriguez-Gomez et al. 2019) of galaxies in Illustris TNG morphologically classified as ETGs and LTGs using the threshold $P(\mathrm{Late}) = 0.5$. The flag `flag_sersic` has been enforced to ensure that only good photometric Sérsic fits are used. Red downward triangles and blue upward triangles indicate ETGs and LTGs, respectively, while the solid cyan and salmon dashed lines are the best linear fit to the relations. The inset shows the distribution of residuals around the best fit for each relation, where the best-fitting Gaussian to the residuals has been superimposed in both cases.

than in our determination of the SMHM, which would correspond to ≈5 per cent difference in halo size, which provides only a very minor corrections to our results.

We also point out that some studies have reported the intrinsic scatter in the SMHM to be larger at lower halo masses (e.g. Rodríguez-Puebla et al. 2015; Moster et al. 2018), but we have adopted a constant $\sigma_{\mathrm{SMHM}} = 0.16$. Since larger σ_{SMHM} would mean a larger scatter in the derived size distribution (Fig. 3), this would exacerbate the tension between data and the MMW model, even after accounting for only the physically acceptable values of λ.

We note that the self-gravity of baryons may lead to important modifications in the structure of the halo, as suggested by several works (Blumenthal et al. 1986; Gnedin et al. 2011; Desmond & Wechsler 2015a; Desmond & Wechsler 2017; Shankar et al. 2017). For example, Jiang et al. (2018) have shown that in a dark matter only simulation matched to a complementary hydrodynamical simulation the correlation of galaxy and halo spin is much less strong than in the hydrodynamical simulation in itself, proving that the effect of baryons in the inner regions of the halo may be crucial. The lack of this kind of information in our approach could potentially affect our conclusions.

It is also important to note that some authors suggest that LTGs live preferentially in haloes with lower concentration (e.g. Wojtak & Mamon 2013; Desmond 2017a). This would mean that the distribution of concentration at fixed halo mass for LTGs is tighter than the full $c{-}M_{\mathrm{h}}$ relation, which would account for less scatter in the observed size functions. In such case, if $\gamma \sim -0.7$ is adopted, a larger intrinsic scatter in the concentration model would be needed. Whether this would still be lower than that in the K13 model remains to be seen, perhaps in the context of conditional abundance matching (e.g. Hearin et al. 2017).

B2 The role of projection effects

In our model we link the halo size directly to the observable 2D effective radius $R_{\mathrm{e}} \equiv R_{\mathrm{e,2D}}$, which is a projection of the true galaxy shape on the sky. However it is the physical half-light radius $R_{\mathrm{e,3D}}$ the quantity that should be physically linked to R_{h}, which may be different from R_{e}. Indeed, Jiang et al. (in preparation) study the relation between R_{e} and intrinsic 3D sizes, and find that a considerable scatter in the relation is present at fixed $R_{\mathrm{e,3D}}$ and intrinsic shape, depending on the line of sight. Therefore the intrinsic scatter in the size distributions would be even tighter and would constitute a further challenge for models of galaxy formation and evolution.

To explore how projection effects affect our analysis of Illustris TNG, we have used the catalogue of optical morphologies and photometric mock observations of IllustrisTNG100 presented in Huertas-Company et al. (2019) and briefly introduced in Section 6 and we plot $R_{\mathrm{e,maj}}$ from the mock observations against the intrinsic 3D size $R_{\mathrm{e,3D}}$ in Fig. B2. It can be seen that the measured $R_{\mathrm{e,maj}}$ for LTGs (ETGs) are only about ∼0.03 (0.06) dex higher (lower) than their physical size, while the slope of the correlation is close to 1 in both cases. Interestingly, the dispersion of the residuals of both the $R_{\mathrm{e,maj}}{-}R_{\mathrm{e,3D}}$ relations are quite small, of the order of ∼0.1 dex. This is even more striking in the light of the fact that the estimate of $R_{\mathrm{e,maj}}$ is prone to both projection effects (as galaxies are mock observed along random lines of sight) and photometric errors. Based on the analysis above and on the fact that galaxy morphologies in IllustrisTNG100 are reasonably well reproduced (Huertas-Company et al. 2019) we conclude that projection effects may not strongly bias the comparison between observations and models that predict the 3D sizes of galaxies.

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