Mon. Not. R. Astron. Soc. **000**, 000–000 (0000) Printed 15 September 2017 (MN LaTeX style file v2.2)

The Emergent Simplicity of Galaxy Size

Andrew Hearin[1], Peter Behroozi[2], Andrey Kravtsov[3], Benjamin Moster[4]
[1] *Argonne National Laboratory, Argonne, IL, USA 60439, USA*
[2] *Department of Physics, University of Arizona, 1118 E 4th St, Tucson, AZ 85721 USA*
[3] *Department of Astronomy & Astrophysics, The University of Chicago, Chicago, IL 60637 USA*
[4] *Universitäts-Sternwarte, Ludwig-Maximilians-Universität München, Scheinerstr. 1, 81679 München, Germany*

15 September 2017

ABSTRACT
We derive empirical modeling constraints on the connection between dark matter halos and the half-light radius $R_{1/2}$ of galaxies. Using forward-modeling techniques based on `Halotools`, we confirm previous results in Kravtsov (2013) that $R_{1/2}$ is well-described by a linear scaling relation with halo virial radius. Novel to this work, we use new SDSS measurements of the $R_{1/2}-$dependence of galaxy clustering to test this modeling assumption. With no changes to the parameters, the model accurately predicts the observed two-point clustering on small- and large-scales over a wide range of stellar mass. This quantitative success is remarkable since the Kravtsov (2013) parameters were fit to the observed $\langle R_{1/2}|M_*\rangle$, and the $R_{1/2}-$dependence of SDSS galaxy clustering has heretofore never been measured. Moreover, this success is non-trivial, as we demonstrate that galaxy clustering is highly sensitive to the assembly history-dependent physics that shapes the relative size of centrals and satellites. Our results can be treated as a boundary condition for more complex and fine-grained models of galaxy size, and provide a simple means for cosmological surveys to generate synthetic galaxy populations with realistic sizes across the cosmic web.

1 INTRODUCTION

Some introduction goes here.

2 DATA AND SIMULATIONS

Our galaxy sample comes from the catalog of SDSS galaxy profile decompositions provided by Meert et al. (2015). This catalog is based on Data Release 10 of the Sloan Digital Sky Survey (SDSS, Ahn et al. 2014), with improvements to the photometry pipeline and light profile fitting methods (Vikram et al. 2010; Bernardi et al. 2013, 2014; Meert et al. 2013). In the version of this catalog that we use, two-dimensional $r-$band profiles were fit with a two-component de Vaucouleurs + exponential profile to determine the half-light radius $R_{1/2}$. We apply the Bell et al. (2003) mass-to-light ratio to the $r-$band flux and $g-r$ colors in this catalog to obtain an estimate for the total stellar mass M_* of every galaxy.

We calculate two-point clustering $w_{\rm p}$ of our SDSS galaxy sample using line-of-sight projection of $\pi_{\rm max}=$ 20Mpc using the `correl` program in `UniverseMachine`. Our results in § 4 will give special focus on the dependence of $w_{\rm p}$ upon $R_{1/2}$. We will quantify this dependence in terms of *clustering ratios* of "large" vs. "small" galaxies, defined according to whether composite galaxy size is above or below $\langle R_{1/2}|M_*\rangle$, computed as the median of a sliding stellar mass window with a width of $N_{\rm gal}=1000$.

As the bedrock of our modeling, we use the catalog of `Rockstar` subhalos identified at $z=0$ in the Bolshoi-Planck simulation (Klypin et al. 2011; Behroozi et al. 2013,?; Riebe et al. 2013; Rodríguez-Puebla et al. 2016). the particular version of the catalog we use is made publicly available through `Halotools` (Hearin et al. 2016), with `version_name` = 'halotools_v0p4'. For mock galaxies, to compute galaxy clustering we employ the distant observer approximation by treating the simulation $z-$axis as the line-of-sight. We compute $w_{\rm p}$ using the `mock_observables.wp` function in `Halotools`, which is a python implementation of the algorithm in the `Corrfunc` C library (Sinha & Garrison 2017).

All numerical values of $R_{1/2}$ will be quoted in kpc, and all values of M_* and $M_{\rm halo}$ in $M_{\odot}$, assuming $H_0=67.8$ km/s $\equiv 100h$ km/s, the best-fit value from Planck Collaboration et al. (2016). To scale stellar masses to "$h=1$ units" (Croton 2013), our numerically quoted values for M_* should be multiplied by a factor of h^2, while our halo masses and distances should be multiplied by a factor of h.

3 GALAXY-HALO MODEL

We map M_* onto subhalos with the best-fit stellar-to-halo mass relation from Moster et al. (2013):

$$\langle M_*/M_{\rm halo}\rangle = 2N\left[(M_{\rm halo}/M_1)^{-\beta} + (M_{\rm halo}/M_1)^{\gamma}\right]^{-1}.$$

For halo mass $M_{\rm halo}$ we use $M_{\rm peak}$, the largest value of $M_{\rm vir}$ ever attained along the main progenitor branch of the subhalo. The values of the best-fit parameters in Moster et al. (2013) were fit to a stellar mass function (SMF) with values $M_*^{\rm MPA-JHU}$ based on the MPA-JHU catalog (Kauffmann et al. 2003; Brinchmann et al. 2004), which differs from the SMF in our galaxy sample (see, e.g., Bernardi et al. 2014). We account for this difference by manually tabulating the median value $\langle M_*^{\rm Meert+15}|M_*^{\rm MPA-JHU}\rangle$ in logarithmic bins spanning $9 < \log_{10} M_*^{\rm MPA-JHU}/M_\odot < 12$, and applying the median correction to the Monte Carlo realization of the mock galaxy sample. This results in a typical boost of ~ 0.25 dex at $M_*^{\rm MPA-JHU} \approx 10^{9.75} M_\odot$, and ~ 0.4 dex at $M_*^{\rm MPA-JHU} \approx 10^{11.5} M_\odot$.

In Kravtsov (2013), it was found that if a stellar-to-halo mass relation is inverted to map halo mass estimates $M_{\rm halo}$ onto SDSS galaxies, and then the $M_{\rm halo} - R_{\rm vir}$ relation is applied to map values of $R_{\rm vir}$ onto the galaxies, then the resulting $R_{1/2} - R_{\rm vir}$ relation of SDSS galaxies exhibits the following linear scaling across a wide range of stellar mass:

$$R_{1/2} = 0.0125 R_{\rm vir} \tag{1}$$

Motivated by the simplicity of this scaling relation, we transform the Kravtsov (2013) into a forward model using `Halotools`. For the virial radius of halos and subhalos, we use $R_{\rm M_{peak}}$, the value of $R_{\rm vir}$ in physical units of kpc measured at the time of peak subhalo mass. When generating Monte Carlo realizations of our model galaxy sizes, we add log-normal scatter $\sigma_{\rm R_{1/2}} = 0.15$ dex.

4 RESULTS

In Figure 1 we show the scaling of galaxy size $R_{1/2}$ with M_*.

In this section we present new measurements of the $R_{1/2}-$dependence of projected galaxy clustering, $w_{\rm p}(r_{\rm p})$. Galaxy clustering has complex, simultaneous dependence upon M_*, sSFR, $R_{1/2}$, and B/T. Since the purpose of this work is to test models of composite galaxy size $R_{1/2}$, we design a measurement that is specifically tailored to isolate the influence of $R_{1/2}$ on the two-point function, while minimizing the dependence upon the other correlated variables.

To achieve this, we compute projected clustering separately for M_*-complete samples of galaxies defined by simultaneous cuts on B/T and $R_{1/2}$, described as follows. First, we construct subsamples of disk-dominated and bulge-dominated galaxies using cuts of B/T < 0.25 and B/T > 0.75, respectively, as in § **??** above. Separately for each B/T−selected subsample, we calculate $\langle R_{1/2}|M_*; {\rm B/T}\rangle$, the median composite size of the sample as a function of total galaxy stellar mass. We split each B/T−selected sample in half according to whether the galaxy has a size $R_{1/2}$ that is above or below this median value. For any M_*-complete threshold, this gives four subsamples of large and small, disk- and bulge-dominated

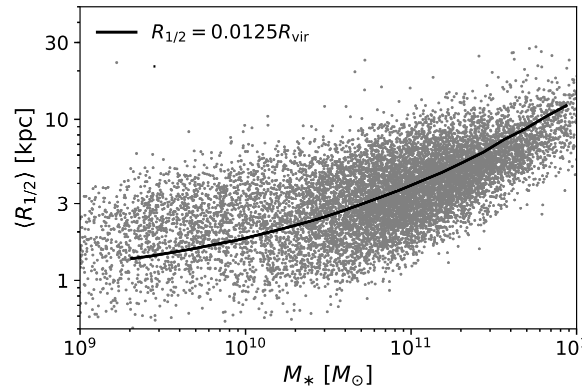

Figure 1. One-point data used to fit the fiducial model. Scattered points show the $R_{1/2} - M_*$ relation for SDSS galaxies as measured in Meert et al. (2015). The black curve in each panel shows the $R_{1/2} - M_*$ relation implied by the model. This figure confirms the findings in Kravtsov (2013) that a linear relationship between $R_{\rm vir}$ and $R_{1/2}$, convolved against the nonlinear relationships between $R_{\rm vir}$, $M_{\rm halo}$ and M_*, correctly predicts the characteristic curvature in the relation $\langle R_{1/2}|M_*\rangle$ over a wide range in stellar mass.

galaxies. For each B/T−selected sample, we calculate the difference $w_{\rm p}^{\rm large} - w_{\rm p}^{\rm small}$, scaled by $w_{\rm p}^{\rm all}$, the clustering of the B/T−selected sample prior to splitting on $R_{1/2}$, and refer to this quantity as *the* $R_{1/2}$ *clustering ratio.*

The y-axis of each panel in Figure 2 shows these $R_{1/2}$ clustering ratios, with left and right panels showing measurements for $M_* > 10^{10} M_\odot$ and $M_* > 10^{10.5} M_\odot$ thresholds in total stellar mass, and top and bottom panels showing results for disk- and bulge-dominated samples, respectively. Points with error bars show SDSS measurements, solid curves show the clustering ratios of model galaxies. Before unpacking the information contained in these clustering measurements, we stress that the good agreement shown between model and data in Figure 2 is a genuine model prediction, since we only tuned parameters to fit the one-point measurements in Fig. 1.

The salient feature of the clustering ratio measurements is that they are negative: for either disk- or bulge-dominated samples, small galaxies cluster more strongly than large galaxies of the same stellar mass threshold, a new result. This feature also holds true for model galaxies, for which both $R_{1/2}^{\rm bulge}$ and $R_{1/2}^{\rm disk}$ scale in power law proportion to $R_{\rm vir}$. Since halo mass $M_{\rm vir} \propto R_{\rm vir}^3$, and clustering is a strong function of $M_{\rm vir}$, the result may seem surprising.

The resolution to this puzzle is illustrated in Figure **??**, which shows the PDF of B/T for halos of the same mass $M_{\rm halo}$, defined as $M_{\rm peak}$, the peak mass ever attained through the history of the (sub)halo, so that host halos and subhalos can be treated on equal footing. In the model, central galaxies are diskier than satellites, a feature that is inherited from the method we use to map B/T to model galaxies.

In light of this explanation, it is natural to ask whether this simple feature *alone* is all that is needed for any model of $R_{1/2}$ to achieve this level of success at predicting galaxy clustering on small and large scales.

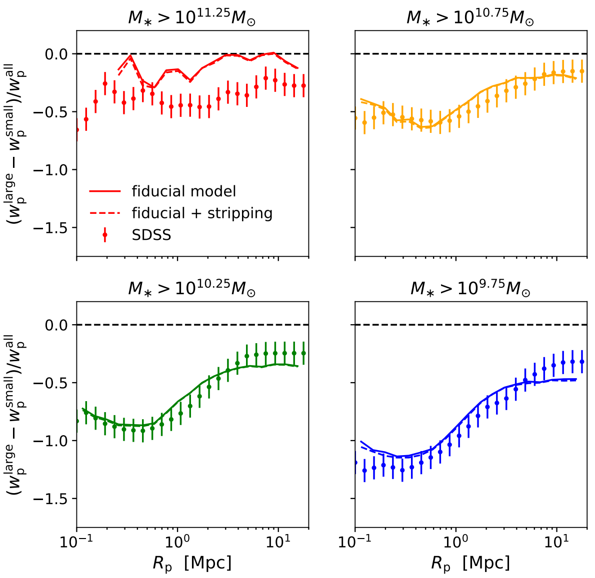

Figure 2. Two-point clustering predictions of the fiducial model. Points with error bars show new SDSS measurements of the $R_{1/2}-$dependence of projected galaxy clustering, w_p, compared to predictions by the model tuned to the measurements shown in Fig. 1. We define a disk or bulge as "large" or "small" according to whether it is above or below the median size for its stellar mass. The y-axis shows clustering strength ratios, so that, for example, a y-axis value of -0.5 corresponds to small galaxies being 50% more strongly clustered than large galaxies of comparable stellar mass. We show results separately for disk-dominated galaxies (*top* panels) and bulge-dominated galaxies (*bottom* panels), and different thresholds in total stellar mass in the *left* and *right* panels. The successful prediction shown here is remarkable because the model was not fit to these data, and because two-point clustering is highly sensitive to the physics that shapes satellite galaxy profiles (see Fig. 3).

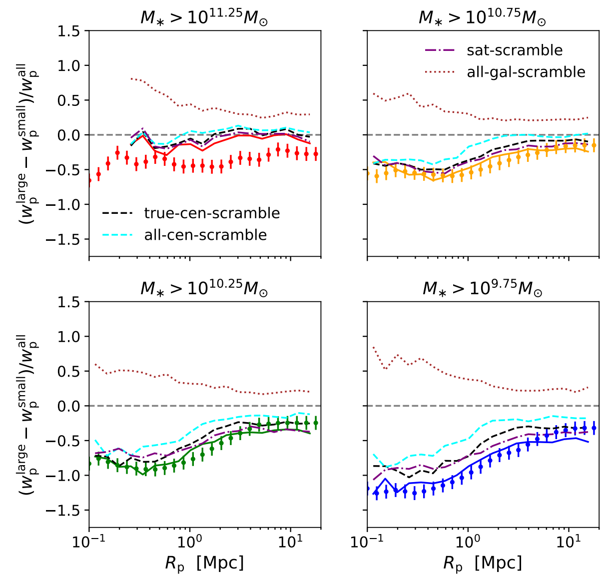

Figure 3. Clustering tightly constrains satellite-specific size evolution. Here we compare our fiducial model, in which satellite galaxy size is set by $R_{\rm vir}$ at the time of infall, to an alternative model analogous to Watson et al. (2012) in which satellite sizes contract in proportion to $(M_{\rm vir}/M_{\rm acc})^{1/3}$. The large differences between solid and dashed curves in the top panels show that the $R_{1/2}-$dependence of galaxy clustering ratios is highly sensitive to the post-infall evolution of satellite galaxy profiles. The successful prediction of our fiducial model, in which satellite galaxies neither contract nor puff up after infall, places tight constraints on satellite-specific physical processes, which must be either negligible or conspiratorially produce little-to-no size change after accretion.

We address this question in § 4.1 below, finding that the answer is no: the reasonably correct magnitude and scale-dependence of the clustering ratios predicted by our model is a non-trivial result that places tight constraints on the post-infall physics of satellite galaxies.

4.1 Sensitivity of Clustering to Satellite Profile Evolution

In this section we explore how post-infall changes in satellite $R_{1/2}$ manifest in $w_p(r_p)$. To do so, we construct an extension our fiducial model with a simple additional ingredient for post-infall size evolution of satellites.

In the fiducial model described in § 3, recall that in the power law a scaling relations (Eq. 1) the halo radius $R_{\rm vir}$ used for satellite galaxies is taken to be the value at the time of infall. The simple interpretation of this assumption is that, in a statistical sense, the size of satellite galaxies is determined by its history as a central galaxy, and that post-infall physics leaves no distinct imprint on satellite $R_{1/2}$.

Here we suppose that this assumption is violated, and that $R_{1/2}$ decreases after infall according to $(M_{\rm vir}/M_{\rm acc})^{1/3}$. That is, in this alternative formulation, rather than using $R_{\rm vir}$ at the time of infall for satellites in the power law scaling relations, we instead use $R_{\rm vir}(M_{\rm vir}/M_{\rm acc})^{1/3}$. This simple toy model attempts supposes that the same physical processes leading to halo mass loss are also responsible for a post-infall decrease in satellite size.

Figure 3 compares the clustering ratios of our fiducial model (solid curves) to the clustering ratios of this alternative model (dashed curves). The difference between the small-scale clustering of the two models is stark: the assumption that $R_{1/2}$ decreases after infall leaves a strong imprint on the $R_{1/2}-$dependence of $w_p(r_p)$, such that small galaxies cluster *much* more strongly relative to large galaxies, particularly for disk-dominated systems with $M_* \approx 10^{10} M_\odot$.

The large differences between the solid and dashed curves in Fig. 3 establishes that the largely successful prediction for the clustering ratios fiducial model is a non-trivial: $w_p(r_p)$ is indeed providing good constraining power on the assumptions underlying our profile modeling and not simply our morphology modeling, c.f., §**??** and §3. We refer the reader to §5.2 for further discussion of the physical implications of this result.

5 DISCUSSION

5.1 Progression from Backwards to Forwards Modeling

5.2 Implications for Satellite Mass Loss

5.3 Future Directions for Empirical Modeling of Morphology

6 CONCLUSIONS

6.1 Summary

ACKNOWLEDGMENTS

APH thanks John Baker for the *Toejam & Earl* soundtrack.

REFERENCES

Ahn C. P., Alexandroff R., Allende Prieto C., Anders F., Anderson S. F., Anderton T., Andrews B. H., Aubourg É., Bailey S., Bastien F. A., et al. 2014, ApJS , 211, 17
Behroozi P. S., Wechsler R. H., Wu H.-Y., 2013, ApJ , 762, 109
Behroozi P. S., Wechsler R. H., Wu H.-Y., Busha M. T., Klypin A. A., Primack J. R., 2013, ApJ , 763, 18
Bell E. F., McIntosh D. H., Katz N., Weinberg M. D., 2003, ApJS , 149, 289
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
Brinchmann J., Charlot S., White S. D. M., Tremonti C., Kauffmann G., Heckman T., Brinkmann J., 2004, MNRAS , 351, 1151
Croton D. J., 2013, PASA, 30, e052
Hearin A., Campbell D., Tollerud E., et al., 2016, ArXiv e-prints
Kauffmann G., Heckman T. M., White S. D. M., et al., 2003, MNRAS , 341, 33
Klypin A. A., Trujillo-Gomez S., Primack J., 2011, ApJ , 740, 102
Kravtsov A. V., 2013, ApJL , 764, L31
Meert A., Vikram V., Bernardi M., 2013, MNRAS , 433, 1344
Meert A., Vikram V., Bernardi M., 2015, MNRAS , 446, 3943
Moster B. P., Naab T., White S. D. M., 2013, MNRAS , 428, 3121
Planck Collaboration Ade P. A. R., Aghanim N., Arnaud M., Ashdown M., Aumont J., Baccigalupi C., Banday A. J., Barreiro R. B., Bartlett J. G., et al. 2016, AAP, 594, A13
Riebe K., Partl A. M., Enke H., Forero-Romero J., Gottlöber S., Klypin A., Lemson G., Prada F., Primack J. R., Steinmetz M., Turchaninov V., 2013, Astronomische Nachrichten, 334, 691
Rodríguez-Puebla A., Behroozi P., Primack J., Klypin A., Lee C., Hellinger D., 2016, MNRAS , 462, 893
Sinha M., Garrison L., , 2017, Corrfunc: Blazing fast correlation functions on the CPU, Astrophysics Source Code Library
Vikram V., Wadadekar Y., Kembhavi A. K., Vijayagovindan G. V., 2010, MNRAS , 409, 1379
Watson D. F., Berlind A. A., Zentner A. R., 2012, ApJ , 754, 90