Dissecting the key correlation between galaxies and their dark matter haloes: Mass or Structure?

Uluk. Rasulov
[1] University of Southampton, University Road, Southampton, SO17 1BJ, UK

23 April 2021

ABSTRACT

In this work we propose an original "Reverse Method" to assign galaxy stellar masses to a dark matter catalogue from galaxy size. We use the linear relation between the galaxy effective radius and halo virial radius $R_{1/2}$=AkR_{halo} to assign galaxy size. We vary the A_k normalisation and σ_k scatter until we match the mock data size function to the SDSS data size function. Stellar Mass is assigned from the observed mass-size distributions of SDSS. The original mapping of halo mass to stellar mass through galaxy size is compared to the traditional Abundance Matching. We find that to match the size function the applied scatter should be higher for the low stellar mass galaxies at 0.22 dex and should reduce to 0.14 dex for Ultra Massive Galaxies (UMG). This suggests that for UMGs the scatter in size can be explained by physical mechanisms in the SMHM relation, whereas for low mass galaxies the scatter is from unknown physical mechanisms. The Reverse Method is successful in reproducing the size function, the stellar mass function (SMF) and the two point projected correlation except for the high mass end (>11 [log10($M_\odot$)]). The failure of the clustering in the high mass end is explained by the large scatter applied to match the size function of SDSS. In comparison to this, Abundance Matching is able to reproduce the SMF and clustering for all masses. The ability of Abundance Matching to reproduce more of the observed data compared to the Reverse Method suggest that the Stellar Mass -Halo Mass relation is more fundamental than Galaxy Size-Halo size relation when it comes correlating haloes and galaxies.

Key words: projected correlation- Abundance Matching – galaxies: statistics – galaxies: groups: general – galaxies: haloes

1 INTRODUCTION

The observable universe has many large scale structures consisting of voids, filaments, mega-clusters, clusters, galaxy groups and galaxies Springel (2006). The structures of galaxies can take many different forms: elliptical, spiral, irregular , bulge dominated and many others Jeremy (2008). The formation and evolution of galaxies is of particular interest as it sheds light into the current understanding of the universe, galaxies and the subsequent distribution of galaxy structures.

The academically accepted method of galaxy formation in accordance with the Lambda CDM concordance model is the *hierarchical structure formation* Cole et al. (2000) Bower et al. (2006). In this model a complex heterogeneous Universe was formed from a homogeneous uniform universe due to tiny quantum fluctuations. Through gravitational attraction dark matter particles form dark matter haloes by continuous merging of smaller units. Subsequently, galaxies form within dark matter haloes by condensation of baryons within the gravitational potential of the dark matter haloes (Figure 1 demonstrates a simple diagram of halo-galaxy structure). White & Frenk (1991) White & Rees (1978). Over time, the haloes evolve through accretion and mergers, whereas galaxies form through cooling of Baryonic mass and evolve into various morphologies through complex physical mechanisms and mergers with other galaxies Huertas-Company et al. (2012) López-Sanjuan, C. et al. (2012).

Considering all the complex formation mechanisms, galaxies display a copious morphological diversity Dimauro et al. (2019), from disc-dominated "spiral" galaxies, such as our own Milky Way, to "lenticulars", discs with central stellar bulges and no spiral arms, and massive, red in colour and with little star formation, stellar spheroids ("ellipticals").

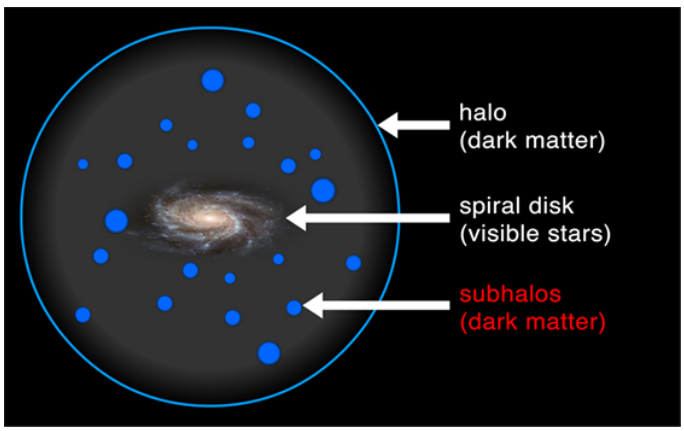

Figure 1. Example figure of a galaxy forming within a parent dark matter halo

Despite such a rich "zoology", all galaxies appear to have tight relations between various galaxy properties. These are called *scaling relations* which can allow one to probe into galaxy formation and evolution White (1997)Posti et al. (2019). One of these observed relations in the Universe and simulated data is the linear one-to-one scaling relation between galaxy size (half-light radius, $R_{1/2}$) and the

virial radius of the host haloes R_{vir} Kravtsov (2013) Shen et al. (2003) Huang et al. (2013) Lange et al. (2015)

In this paper the linear relation between the half-light radius, $R_{1/2}$ of a galaxy and its host halo virial radius (Kravtsov Relation) will be tested to see if the $R_{1/2}$ R_{vir} relation is more or less fundamental than the more widely adopted stellar mass M_* versus halo mass M_H relation (SMHM) Shankar et al. (2006). This will give insight on showing which property is more fundamental when it comes to the observed stellar mass of the galaxy: the size of the haloes or the mass of the parent haloes. The SMHM relation used to populate galaxies is based upon abundance matching Campbell et al. (2018), which is a widely used methodology proven to reproduce the spatial clustering of galaxies Norberg et al. (2001) Li et al. (2006) Norberg et al. (2002).To the best of our knowledge this paper proposes the "Reverse method" - a completely new methodology - to populate dark matter haloes with galaxies. The "Reverse method" produces a SMHM relation which is based on the linear $R_{1/2}$ R_{vir} relation and mass-size distribution of the SDSS data. The clustering of the galaxies is compared between the "Reverse Method" and the "Forward method" (Abundance Matching) to see which relation is more fundamental.

It is worth noting that in the hierarchical structure formation the haloes and subsequent galaxies can either be "Centrals" or "Satellites", with the Satellite galaxies orbiting a host Central galaxy. Various studies have shown that the occupation of the galaxy can affect the formation and the evolution Wetzel et al. (2013) Baldry et al. (2006). For the purposes of this paper only central galaxies will be taken into account for simplicity of calculations to test this new methodology. Satellite galaxies can be a focus of further research.

The testing of the Forward and Reverse methods will be based on the comparison of the two point projected correlation between the observed data of the Universe from the Sloan Digital Sky Survey Data Release 7 (SDSS) and the the two point projected correlation of a simulated Dark matter Universe.

The two-point projected correlation function $w_p(r_p)$ was shown to be effectively used to constrain connections between parent haloes and the subsequent galaxies Moster et al. (2010). In this paper the hydro-dynamical simulation of dark matter haloes will be populated with galaxies using abundance matching based on the Stellar Mass Halo Mass relation (SMHM) of Grylls 19 (Grylls et al. (2019)). Subsequently, both the SMHM and the Kravtsov relation will be tested by applying them to the simulated dark matter catalogue to generate a mock galaxy catalogue. Then the two relations will be tested by calculating the correlation function in order to see if they reproduce the correlation of observed data.

In this work CDM Planck cosmology is adopted Ω_m = 0.302 , Ω_λ =0.698 , h=0.677

2 OBSERVATIONAL AND SIMULATION DATA

2.1 SDSS Data

For the observable data we shall use Sloan Digital Sky Survey (SDSS) Data release 7 Abazajian et al. (2009). It is the largest database of the observable Universe. We supplement the catalogue with classification of morphologies of galaxies based on deep learning mechanism by Domínguez Sánchez et al. (2018). The catalogue is further supplemented by measurements of half-light radius $R_{1/2}$ provided by Meert et al. (2015a) Meert et al. (2015b).The Meert et al. catalogues are matched with the Yang et al. (2007) group catalogues. The central galaxies within a group of galaxies are identified as the most luminous, while the remaining objects in the group are considered to be satellites.

In the Meert et al. catalogues galaxies are fit with Sersic+Exponentialas as well as a Sersic profile. The galaxy stellar masses are computed adopting such light profiles and mass-to-light ratio M_{star}/L Mendel et al. (2014). The effective radius $R_{1/2}$ (half-light) is the truncated semi-major axis half-light radius of the full fit Fischer et al. (2017).

Overall the catalogue consists of 484,263 total galaxies, with 371,382 centrals and 112,881 satellites

2.2 MultiDark Simulation Data

The dark matter catalogue is the MDPL2 – Rockstar from the MultiDark Bolshoi simulation from the CosmoSim database Prada et al. (2012) Klypin et al. (2011) Riebe et al. (2013). The mock catalogue is made up of a hydro-dynamical simulation of dark matter from the beginning of the Universe up to the present time. The catalogue consists of Central and Satellite Dark matter haloes, with mass and positions in a X,Y,Z coordinates in units Mpc/h. The catalogue is a snapshot of the simulation at z = 0.093 It is a cube of axis 500/h Mpc. It is 1/8th of the original simulation consisting of total 16,708,203 haloes, 14,614,328 Centrals and 2.093,875 Satellites. The halo catalogue will be populated with galaxies using abundance matching and the Reverse method using the Kravtsov relation to generate a mock galaxy catalogues.

3 ABUNDANCE MATCHING

Abundance matching is a semi-empirical technique used to relate the stellar mass of a galaxy to its parent halo. It works by assuming that one parent halo contains one galaxy. The stellar mass and halo mass relation is linked by mapping the Halo Mass Function (HMF) to the Stellar Mass Function (SMF) to create a mean mapping of Stellar Mass To Halo Mass (SMHM), which reproduces the observed SMF in the SDSS Observed catalogue. In this paper we will use the Grylls et al. (2019) SMHM in parametric form based on the Moster et al. (2010) SMHM.

$$
\begin{aligned}
M_*(M_h,z) = 2M_h N(z)[(\frac{M_h}{M_n(z)})^{-\beta(z)} + (\frac{M_h}{M_n(z)})^{-\gamma(z)}]^{-1} \\
N(z) = N_0.1 + N_z(\frac{z-0.1}{z+1}) \\
M_n(z) = M_{n,0.1} + M_{n,z}(\frac{z-0.1}{z+1}) \\
\beta(z) = \beta_0.1 + \beta_z(\frac{z-0.1}{z+1}) \\
\gamma(z) = \gamma_0.1 + \gamma_z(\frac{z-0.1}{z+1})
\end{aligned}
\tag{1}
$$

For red shift z =0.1 the parameters M_n , N , β , γ , ϑ are selected such that SMF of the mock data matches the SMF of SDSS.

Only central galaxies are used to calculate the SMF for SDSS. The catalogue is supplemented with Vmax weights of measurements representing the accuracy of the values. To calculate the SMF for the observable data only galaxies above mass 10^9 M_{sun} are used due to the loss of accuracy of measurements at low masses.

The SMHM relation in the form of a broken power law used by Grylls et al takes into account all galaxies and uses an optimisation algorithm to fit the parameters of the SMHM to the SMF galaxies in SDSS by minimising the root mean square. In this paper the parameters are varied individually until the SMF of the mock data takes a reasonably accurate form of the SDSS SMF. This is an incomplete

[H]

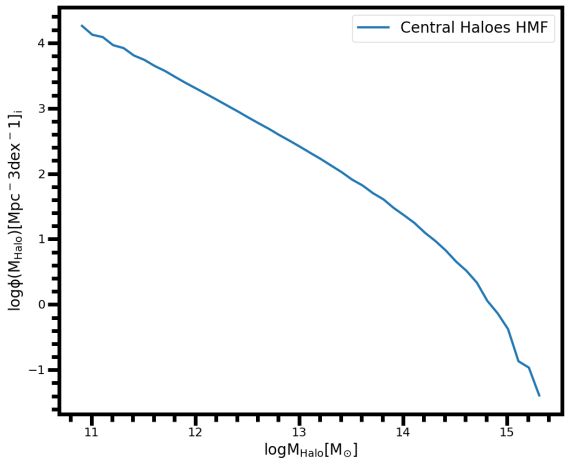

Figure 2. The Halo Mass Function of Central Galaxies of the mock dark matter halo catalogue

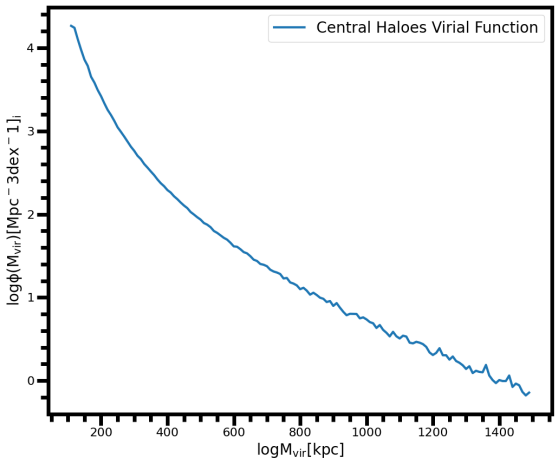

Figure 3. The Virial Radius Function of Central Galaxies of the mock dark matter halo catalogue

methodology and an optimisation algorithm such as the Markov Chain Monte Carlo (MCMC) should be used for full accuracy in future work.

The SMF of the mock is generated from the MultiDark HMF and the SMHM relation is from Moster et al. (2012) Buchan & Shankar (2016). The HMF is generated by the number density of MultiDark haloes for each parent mass bin [M_{cent}(z),M_{cent}(z) + dM_{cent}(z)]. Figure 2 shows the HMF used in this paper. The SMF is generated by multiplying the HMF densities by the SMHM relation and adding Gaussian distribution of (ϑ).

The parameters used for abundance matching are: M_n=11.72, N = 0.032 , β = 1.90 , γ = 0.49. A Gaussian scatter of θ = 0.1 is added to the stellar masses.These are different from the original parameters for Grylls since we are not including satellite haloes. The final SMF and the SMHM relation for the mock catalogue can be seen in Figure 11 and Figure 13.

4 THE REVERSE METHOD

It has been shown by Kravtsov (2013) that the galaxy effective radius is linearly proportional to the virial radius of parent haloes. The relationship found by the author is based purely on empirical findings. The author adopted an abundance matching technique to calculate M_{200} halo mass of galaxies from the stellar mass. The halo radii R_{200} was calculated from the halo mass and the author estimated the half-light radii $R_{\frac{1}{2}}$ of the observed galaxies (the radius which contains half of the stellar mass of the galaxies). The author found that for many morphologies of galaxies the size-virial relation could be well reproduced using a single normalisation:

$$R_{1/2} = AkR_{halo} \tag{2}$$

The author used a normalisation of Ak=0.015 with an added scatter 0.2dex.

It has been found by Zanisi et al. (2020) that the galaxy size function of SDSS could be reproduced using the Kravtsov relation applied to the halo catalogue and with an added scatter to sizes. The author found that in the Kravtsov model the scatter must decrease for massive galaxies, irrespective of galaxy morphology.

Similar to this approach in the Reverse method we shall assign stellar masses to galaxies through the following steps:

i) Matching galaxy size function of the mock catalogue to the SDSS central data in the log stellar mass bins:

9.0-9.5 | 9.5-10.0 | 10.0-10.5 | 10.5-11.0 | 11.0-11.5 | 11.5-12.0

The free parameters to assign galaxy sizes in this case are the Kravtsov normalisation A_k and the added scatter σ_k

ii) Once the galaxy sizes are assigned, we assign stellar mass from galaxy size using the mass-size distributions directly from SDSS data. This is done by binning the SDSS masses by galaxy size. If the mock galaxy size is within the SDSS bin window, we randomly assign masses such that the mass size distribution of SDSS is reproduced.

iii) Now that the full mock catalogue is complete with R_{vir} M_{halo} $R_{1/2}$ M_*. One can build a SMHM relation independently from abundance matching.

iv) The last step would be to calculate the clustering of mock galaxies by cutting the catalogue by mass and seeing how the clustering compares to the SDSS data.

The backbone of assigning masses from galaxy size can be seen in Figure 4.

Throughout the paper for central haloes we use R_{vir} as the virial radius, the value of R_{vir} measured at the time when virial theorem applies Coe (2010). It is defined as the radius at which the density is equal to the critical density ρ_c of the Universe at the redshift of the system, multiplied by an overdensity constant Δ_c Navarro et al. (1996). We use the colossus package Diemer (2018) to do the

calculation following the equation:

$$R_{vir} = [\frac{3 * M_{vir}}{4 * \pi * \Delta_{vir}(z) * \rho_m(z)}]^{\frac{1}{3}} \quad (3)$$

where for $\Delta_{vir}(z)$ the Bryan & Norman (1998) fitting function is used for the definition of the "Virial". The virial radius function used in this paper can be seen in Figure 3.

4.1 Matching The Size-Function

We use our SMHM relation to convert the stellar mass bin cuts into halo mass. Then from M_{halo} we calculate R_{vir} using the virial theorem. These R_{vir} values work as our boundaries for haloes when assigning $R_{1/2}$ using the Kravtsov relation. We vary the Ak normalisation and the σ_k scatter until we match the peaks and the distribution of the size-function of the observed catalogue. Due to our abundance matching SMHM relation not being fully optimised we slightly vary the halo boundaries to match the size functions. The used parameters can be seen in Table 1 with the size functions for cuts in stellar mass in Figure 5.

4.2 Effective Radius From Virial Radius

Kravtsov used a single normalisation of 0.015 working directly from observable data. In our model we work starting with mock dark matter catalogue to match the SDSS with two free parameters: the A_k normalisation and the log normal σ scatter (both of which are not known a priori). Our results agree with Zanisi et al. (2020) who found that a scatter of dex 0.2 reproduces the SDSS size function at low masses and the scatter needs to be reduced to dex 0.1 for the high mass end (The results of our paper is dex=0.22/dex=14 Low mass/High mass). The author also found that two separate normalisations are needed for Early Type and Late Galaxies to match the SDSS data. Higher normalisation values are needed for Late Type Galaxies (LTG) compared to the Early Type Galaxies. These results qualitatively agree with Huang et al. (2017) Lapi et al. (2018). In this work when assigning half-light radii to the mock catalogue there is no distinction between ETGs and LTGs. We work with all centrals, due to the limitation of the dark matter catalogue.

One can see in Figure 6 the total size function (all central galaxies) of the mock catalogue compared to the SDSS data only for central galaxies. The errors for SDSS are calculated from bootstrap re-sampling. Evidently using our model the size function of the mock catalogue can be used to reproduce the size function of the SDSS to a reasonable degree.

In figure 13 the SMHM for our abundance matching model shows a typical double power law. The distinct shape of the double power law and two separate gradients are believed to be caused by the efficiency of star formation. The star formation rate is highest below a certain halo mass and is suppressed above a certain halo mass. This is where Supernova and a combination of Active Galactic Nuclei (AGN) feedback and virial shocks, respectively, are believed to be most efficient (e.g. Shankar et al. (2006)Behroozi et al. (2013)Kravtsov et al. (2004).

One can see the the steepness of the power law change after the halo mass of 10^{12} $M\odot$. This is after the highest star formation rate has been reached. The gradient is significantly less, hence larger virial halo masses are needed for an increase in stellar mass. Therefore, when one would consider the virial radii distribution for different mass cuts, greater distribution of values are expected for larger stellar mass. This is a feature that is usually described in terms of the halo occupation distribution function Zheng et al. (2005). This is exactly what is observed in our mock catalogue in Figure 7.

The distribution of virial radii is greater for larger stellar mass. Hence if the size distribution could be reproduced directly with normalisation Ak and σ_k=0 it could be argued that the same galactic physical processes of AGN feedback, virial shocks etc are responsible for the size function. However, our model does not show this, therefore the model suggest that there are other physical mechanisms unrelated to the double power law of the SMHM relation that determine the broadness of the size function. On the other hand, for the Ultra Massive Galaxies the scatter decreases to 0.14 dex suggesting at this mass range (>11.2 log$M_{\odot}$)the scatter is dominated by the processes in the SMHM relation.

4.3 Assigning Stellar Mass From The Half-Light Radius

Abundance matching has become a stable workhorse to create a SMHM relation and used to assign galaxy stellar masses to dark matter haloes. Abundance Matching has been extensively proven to reproduce the clustering for mock catalogues which match the clustering of the SDSS data Yang et al. (2004) Vale & Ostriker (2004) Kravtsov et al. (2004). It has also been used in reverse (M_*->M_halo)to assign halo mass and virial radii to the observed data to derive the R_{vir}-$R_{1/2}$ relation used in this paper. The Kravtsov relation R_{vir}-$R_{1/2}$ has been further utilised to reproduce size-mass distribution of SDSS data when applied to mock catalogues Somerville et al. (2017). Furthermore, Kravtsov relation has been proven to generate a broken power law of the Halo Mass-Stellar Mass (HMSM) relation from abundance matching when applied to SDSS data ($R_{1/2}$->R_{vir}->M_{halo}) Mowla et al. (2019). However, to the best of our knowledge it has not been used to assign stellar mass to haloes and generate the SMHM relation, which is the originality of this paper.

With normalisation and scatter chosen so that the size function of the mock catalogue matches the SDSS data, we assign stellar mass from the half-light radius $R_{1/2}$. This is done by using the Mass-Size Distribution directly from the SDSS data. One can see in Figure 8 the Mass-Size Distribution of the SDSS data for central galaxies, where for different bins of $R_{1/2}$ the median stellar mass is calculated with the uncertainty being the one sigma standard deviation of the stellar mass. This can be seen as a rough approximation of how the mass is assigned from galaxy size. Instead of assigning stellar mass from a Gaussian normal distribution we assign a random stellar mass, such that the distribution of stellar masses within the radii bin matches directly the SDSS data. For example, consider a mock galaxy with size 0.82 log(kpc), we assign a random stellar mass such that the mock catalogue stellar mass distribution between the radii 0.8-0.9 log(kc) matches the SDSS. An example of mass being assigned can be seen in Figure 9.

4.4 The Stellar Mass Function and SMHM Relation From The Reverse Method

Once the stellar mass has been assigned we can do an initial test of the model and check the resultant SMF and the SMHM relation. These can be seen in Figure 11 and Figure 13. As shown, the SMF matches the SDSS data to a reasonable degree, while the SMHM relation has a similar shape to Abundance Matching but lacks the prominent double power law shape. It is interesting to note the very vast uncertainty of the SMHM relation from the Reverse model. Both uncertainties are one sigma standard deviation of the data within the halo bin, but the Reverse Method uncertainty is significantly larger.

[h]

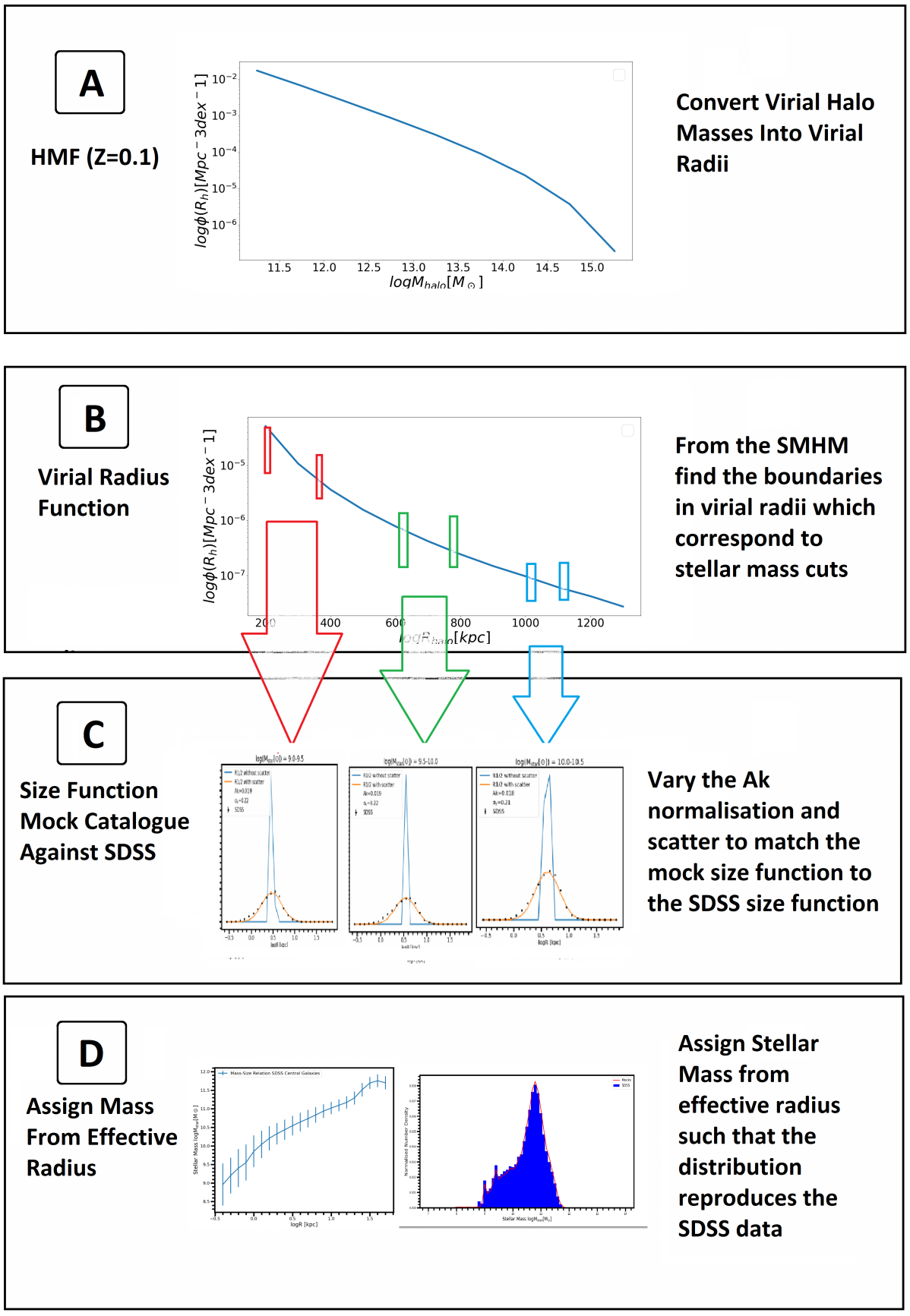

Figure 4. Figure to show the main steps in the Reverse Method. Stellar Mass is assigned from galaxy size rather than halo mass in traditional abundance matching

[t]

Stellar Mass log[$M\odot$]	9.0-9.5	9.5-10.0	10.0-10.5	10.5-11.0	11.0-11.5	11.5-12.0
Halo Mass log[$M\odot$]	11.25-11.45	11.45-11.66	11.66-12.03	12.03-12.94	12.94-14.04	>14.04
Virial Radius limits (kpc)	131-170	170-200	200-265	265-450	450-950	>950
Normalisation Ak	0.019	0.019	0.018	0.018	0.015	0.012
σ_k	0.22 dex	0.22 dex	0.21 dex	0.20 dex	0.17 dex	0.14 dex

Table 1. Table to show parameters used in the Reverse Method to match the size function of SDSS

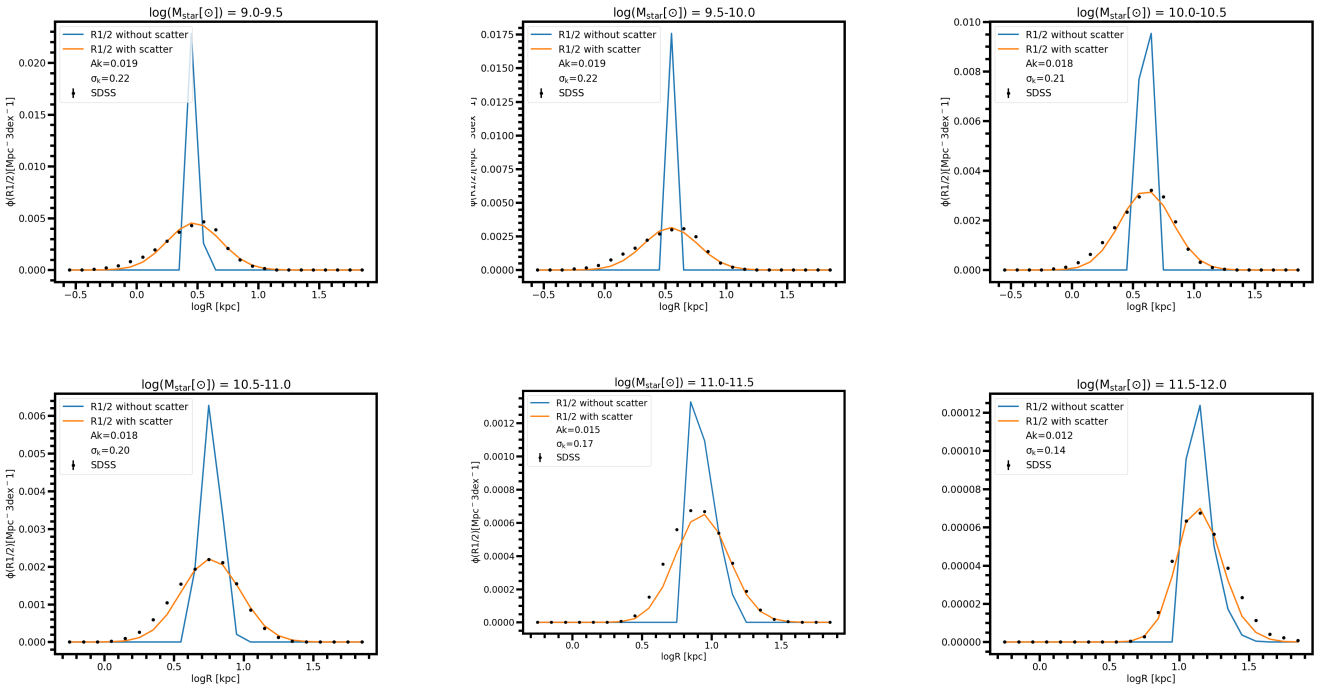

Figure 5. Size functions of the model for different cuts in Stellar Mass. The normalisation and scatter were chosen to match the peak and distribution of observed SDSS data. The errors in SDSS data are from bootstrap re sampling. Notice how scatter is reduced at high masses

The cause of this will be discussed in the next few sections and can be summed up that it arises from the larger scatter applied to the $R_{1/2}$. Overall, we have shown that from the Reverse Method we can reproduce the size function of SDSS and the SMF of the SDSS. The final test is if the Reverse Method can match the two point projected correlation of the SDSS data.

5 TWO POINT PROJECTED CORRELATION FUNCTION

One can not test galaxy formation models with experiments in the Universe, hence the next best method is done by comparing observable universe data to hydro-dynamical simulations of the Universe van Daalen et al. (2016). One powerful method is the two point correlation function, which can constrain empirical models of galaxy formation by comparing the correlation of observed and simulated data. This is done by selecting the data points influenced by empirical relations. For example, Abundance Matching links the Halo mass to the Stellar mass of the galaxies therefore, this relation can be tested and constrained using the correlation function by calculating the clustering of galaxies of specific galaxy stellar mass bins.

The two point correlation function gives a description of the distribution of galaxies in the Universe. $\varepsilon(r_p, \pi)$ is defined as the excess probability from random of finding a galaxy separated by that radial distance within the line of sight distance π_{max}.

For SDSS data we calculate w_p within the radial distance of 0.5-10.0 Mpc/h and π_{max} = 40. We use the correlcalc Rohin (2017) package to generate a random data-set three times the size of the input data, with the correct galaxy distributions. We use the SDSS data 7 mask file File (2004) which provides the survey geometry in .ply formats. We use mangle polygons Hamilton & Tegmark (2004) given by the galaxy surveys to create random points that fall within the survey geometry.

With the random data-set and input data with (Redshift, Right Ascension, Declination , Radial Weight) we calculate the number of pair points within the line of sight difference $\pi = (\vec{s} * \vec{l})/|\vec{l}|$ and transverse separation $r_p = \sqrt{s^2 - \pi^2}$. For two sample points $\vec{x}_1$ and $\vec{x}_2$ with $\vec{s}$ = $\vec{x}_2$ - $\vec{x}_1$ and $\vec{l}$ = 0.5($\vec{x}_2$ + $\vec{x}_1$). We use the Corrfunc package Sinha & Garrison (2020) to calculate correlation function using the Landy & Szalay Estimator Landy & Szalay (1993) DD-2DR + RR/RR where DD is galaxy-galaxy pair counts, DR is galaxy-

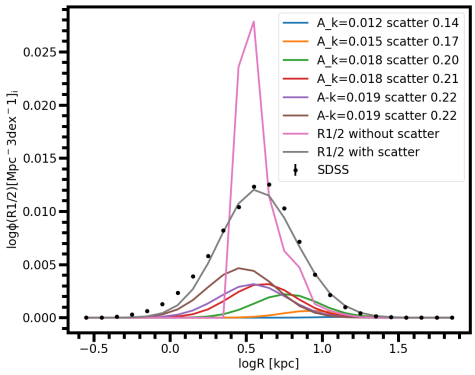

Figure 6. The total size function for the observed SDSS data and the mock catalogue. Note how the size function can be reproduced by the mock catalogue. The scatter needed is quite large and there is a large overlap of radii

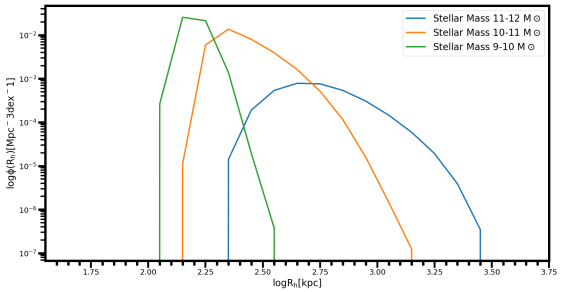

Figure 7. Virial radii distribution for different cuts of stellar mass. Shows the role of the SMHM relation in setting the scatter in halo size. If the σ_k=0 scatter reproduced SDSS data, then same mechanisms from SMHM relation take place in the Reverse Method

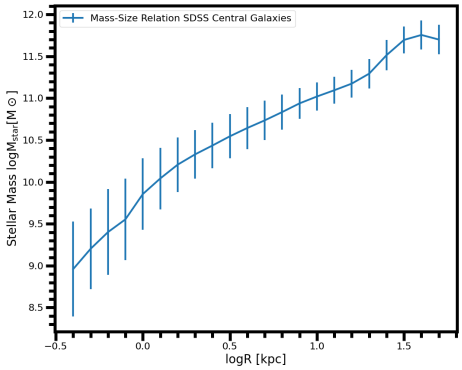

Figure 8. The median Mass-size Relation of SDSS data. The errors are one sigma standard deviation of data within the appropriate bin

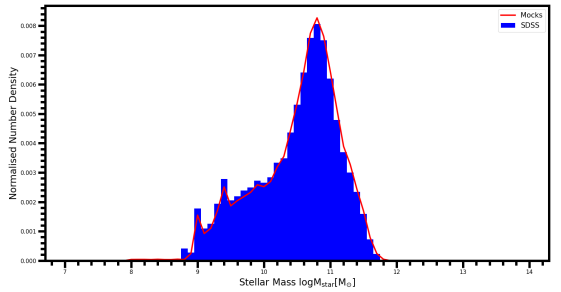

Figure 9. Stellar Mass distribution between 0.8-0.9 [kpc]. Figure shows how stellar mass is assigned to a halo from the $R_{1/2}$ radius such that the mock distribution reproduces the distribution of observed data

random pair counts and RR is random-random pair counts. The correlation function $\varepsilon(r_p, \pi)$ is then integrated over $|\pi|$ < 40 Mpc/h to calculate the projected correlation function w_p. The integral along the line of sight takes the influence of the red shift errors of the two point correlation function into account and gives a more accurate constraining method

$$\int_{\infty}^{-\infty} \varepsilon(r_p, \pi)\, d\pi = \int_{\infty}^{r_p} \varepsilon(r) \frac{r dr}{[r^2 - r_p^2]^{0.5}}\, dr$$

For mock data of our galaxies, with X,Y,Z coordinates, the projected correlation can be computed directly using Corrfunc.theory.wp function. Hence, the projected correlation function of simulated and observed data can be directly compared.

6 RESULTS OF CLUSTERING

The final test and comparison between the Reverse Method and Abundance Matching is the comparison of the clustering. We take four cuts of the mock catalogue by mass and compare them to the SDSS clustering. The cuts of data in the log10 basis are stellar masses 9.4-9.8, 10.0-10.3,11.0-11-3, 11.3-11.7. The projected correlation is calculated between radial bins 0-10 Mpc/h and π_{max} = 40Mpc/h. The results can be seen in Figure 10.

One can note that our abundance matching reproduces the projected clustering quite well especially at the high mass end. This is expected due to greater accuracy of results for high mass galaxies due to their increased luminosity. The errors for the SDSS data are from bootstrap re-sampling. Remember that the parameters for abundance matching were changed by hand without the full MCMC, hence it is expected that there is some tension in the clustering. The most interesting result is that the clustering is not reproduced by the Reverse Method for the high mass end above $logM_{star} > 11.0$. The low mass end is nearly identical to Abundance Matching and is in a reasonable agreement with the SDSS data. However, the high mass end is very vastly under-clustered. This suggests that through the Reverse Method lower mass haloes are being assigned to the larger galaxy stellar mass than expected, since it is expected that high mass haloes and galaxies have stronger clustering Yang et al. (2005). This is an interesting result since we have been able to reproduce the SMF of the SDSS catalogue, suggesting that the relative abundances of our galaxies are in agreement with the observed data. Furthermore, having a SMF matching the SDSS data has been proven to reproduce the clustering in Abundance Matching methods in previous work Contreras et al. (2020). One can conclude that using the Reverse Method allows a quite large distribution of masses assigned to haloes, which can explain the under clustering in the high mass end. The significance of the high mass end could be explained by the sensitivity of the data within the cut of data. There are considerably less galaxies at the high mass end, hence it would be very sensitive to the distribu-

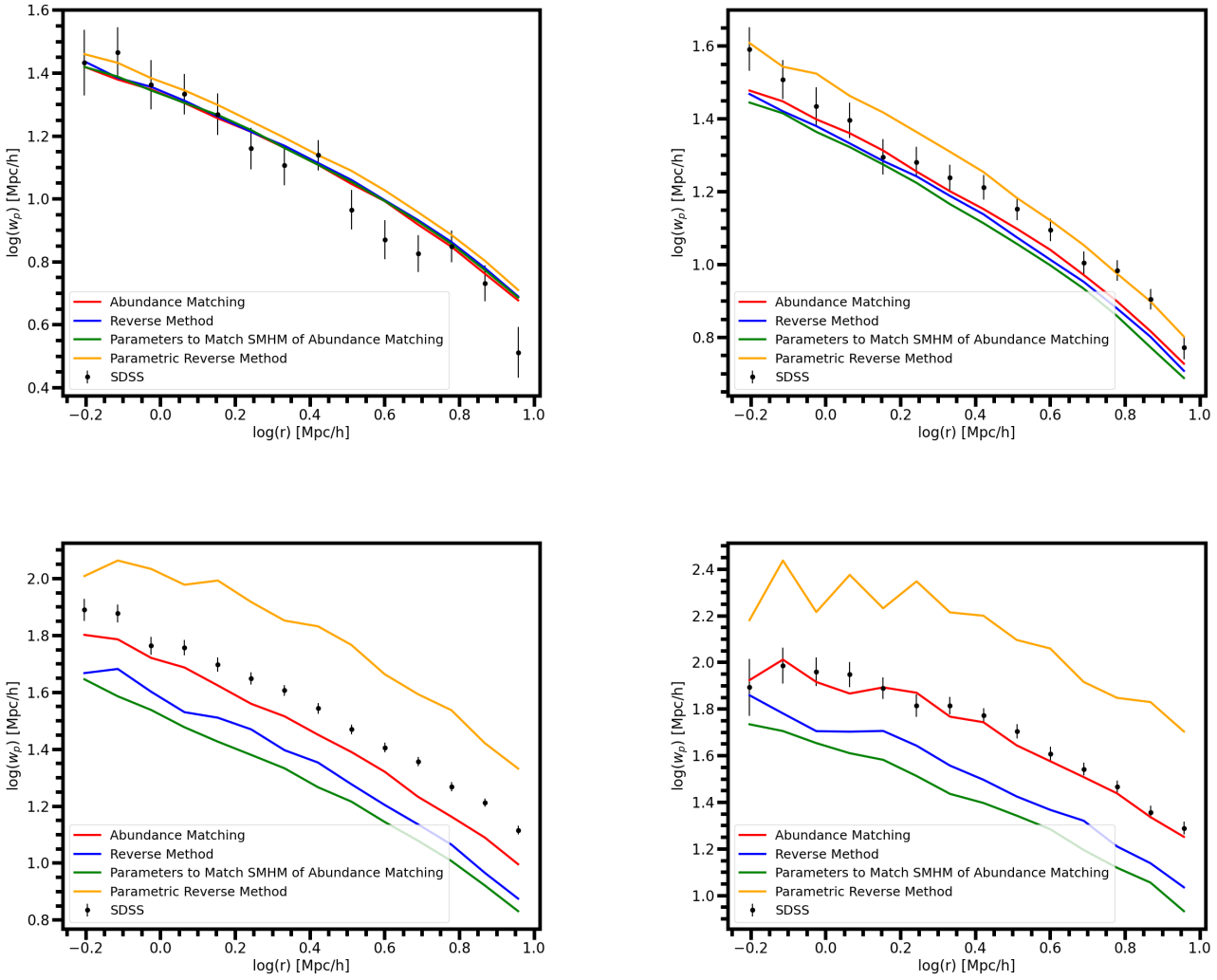

Figure 10. The two point projected correlation for different mass cuts compared to observed SDSS data. Notice how Abundance Matching is able to reproduce SDSS data to a reasonable degree. The Reverse Method reproduces clustering except the high mass end. The parametric form of the Reverse Method is vastly different from the original results. Changing the parameters of the Reverse Method to match the SMHM of Abundance Matching comes at a sacrifice of the clustering.

tion and scatter of the stellar mass. Conversely, through Abundance Matching the high mass end clustering is reproduced naturally due to the large haloes consistently being assigned to high stellar masses. Different possibilities of these results and the implications will be discussed in the next section.

7 DISCUSSION

7.1 Changing The Parameters and Effect of Scatter

One clear implication of these results is that there is break in the Kravtsov and clearly to create a model which can reproduce the SMF, size function, SMHM, and clustering the R_{vir}->$R_{1/2}$ relation must be complex and multi-functional. One can not reproduce the observed results with a single value when working from a mock catalogue.

Let us discuss the failure of the clustering in the high mass end. This is quite counter intuitive, since the high mass end clustering tends to occur naturally from Abundance Matching. A very important point is that our current model is very flexible. The two free parameters of A_k and σ_k can be varied to change our mock catalogue with a lot of freedom. Starting from the R_{vir} to M_* there are three instances which can affect final stellar mass of our individual halo in MultiDark.
i) The A_k normalisation
ii) The scatter applied σ_k
iii)The assignment of mass from the distribution of the SDSS
To understand this problem let us vary the parameters and see the implication. If one increases the A_k values this would shift the size function to the right. Hence there would be more high mass galaxies and the SMF for the high mass end would increase.

Consider changing the third Ak value (200-265 virial [kpc]) to 0.019 from 0.018 which would increase the number of galaxies within the 11.0-11.3 mass bin from 103,850 to 114,012. This would in turn

Stellar Mass [log($M_\odot$]	9.4-9.8	10.0-10.3	11.0-11.3	11.3-11.7
Abundance Matching	730373	481337	103154	29239
Reverse Method	754786	509711	103850	36416

Table 2. Table to show the number of galaxies within the various stellar mass cuts. Notice how the high mass end is made of fewest galaxies

[t]

Stellar Mass log[$M\odot$]	9.0-9.5	9.5-10.0	10.0-10.5	10.5-11.0	11.0-11.5	11.5-12.0
Halo Mass log[$M\odot$]	11.25-11.45	11.45-11.66	11.66-12.03	12.03-12.94	12.94-14.04	>14.04
Virial Radius limits (kpc)	131-170	170-200	200-265	265-450	450-950	>950
Normalisation Ak	0.020	0.026	0.027	0.027	0.023	0.024
σ_k	0.22 dex	0.22 dex	0.21 dex	0.20 dex	0.17 dex	0.14 dex

Table 3. Table to show parameters used in the Reverse Method to match the final SMHM relation to the SMHM from abundance matching

decrease the clustering, since the same haloes will be assigned to larger, more massive galaxies. Lower mass haloes cluster less, so the clustering becomes lower as well. This clustering problem is also further unique to the high mass end due to the low number of galaxies. In the reverse method when the scatter is assigned to match the size function and the SMF, the scatter has a very large effect on $R_{1/2}$ hence the high mass end gets washed out. Some very large haloes go to lower $R_{1/2}$and some smaller haloes get assigned high $R_{1/2}$, which assigns them high stellar masses. Furthermore, even if a halo is assigned a high $R_{1/2}$ there is further scatter from the SDSS distribution themselves.Table 2 and Figure 6 show how the high mass end of the mock catalogue from Abundance Matching and the Reverse Kravtsov mock make up a very small amount of the overall data. In addition, Figure 6 shows the individual scatter of every single A_k value. Note how vast the scatter causes an overlap in $R_{1/2}$

7.2 Matching the SMHM relation using the Reverse Model

In this work, we started from matching the size function and the SMF, where the SMHM relation was the final result. However, let us consider the scenario when we change our model parameters to reproduce our SMHM from Abundance Matching, since it has been proven to reproduce the clustering.

We change the values until our SMHM relation matches the Abundance Matching SMHM as seen in Figure 14. The parameters of the model can be seen in Table 3.

Note how we can change the SMHM function by increasing the normalisation values. This is mapping more lower mass haloes to higher galaxy stellar mass. This comes at a sacrifice of the SMF (Figure 11) and the size function (Figure 12). Furthermore, the clustering at the high mass end is still not to the acceptable level. This suggests that there are clearly fundamental differences between the two methodologies.

7.3 Fundamental Differences Between Methodologies

Our results so far show that the Reverse method is unable to reproduce the clustering at the high mass end. This could be explained by the large scatter applied to the haloes when assigning galaxy size to match the size function. This suggests that SMHM is the more fundamental property than size and structure when assigning stellar mass. However, one final test to outline the fundamental inconsistency between the methodologies is to work from the final SMHM arrived by the Reverse model and reapply it to our dark matter haloes. This is done by parameterizing the SMHM relation using linear interpolation of the median values and applying it to MultiDark with the same uncertainties. Figure 11 and Figure 13 show how the SMF and the clustering at the high mass end are very different, even though we are using the same SMHM relation from the Reverse Method. This suggests that there are fundamental inconsistencies between the Reverse and forward model, with the Abundance matching forward model reproducing a mock catalogue resembling the observed SDSS data to a greater degree.

There are several implications of this. Kravtsov explained the R_{vir}->$R_{1/2}$ relation by the Mo et al. (1998) model, where the galaxy sizes are determined by the angular momentum of Baryons and dark matter haloes. The angular momentum were acquired by tidal torques during cosmological collapse. Kravtsov used a single scatter of 0.2 dex which quantitatively agreed with Vitvitska et al. (2002), where the scatter was from the spin of parent halo. Our model based on this relation was able to reproduce the SMF, size function and the clustering except for the high mass end. Hence there is some interplay between the halo angular momentum, halo size and galaxy structure when determining how massive a galaxy is. However, due to vast scatters involved it suggests that there are other physical mechanisms not accounted in the model. The Mo et al. (1998) model is only for the formation of disk galaxies, hence not separating LTGs and ETGs could damage the accuracy of results, especially since ETGs play a more significant role in the high mass end Tamburri, S. et al. (2014). In addition, we found that we can match the SMHM by changing the components of our model, but this comes at a sacrifice of the SMF and the size function. Overall our result suggest that the Abundance Matching SMHM relation is more fundamental than the R_{vir}-> $R_{1/2}$ relation, which implies that it is the interplay between the halo mass and galaxy mass that determines how massive a galaxy is. Perhaps one can think of the interplay between the halo mass-galaxy mass in terms of the virial theorem $v^2 - \frac{GM^2}{R}$. The virial theorem and the physical attraction of halo masses is a feature of all stellar systems. The virial theorem is more general than angular momentum - Kravtsov's argument was drawn from Mo et al. (1998) which is only valid for disk galaxies (LTGs).

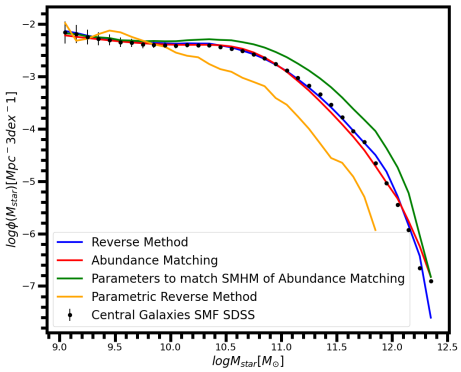

Figure 11. The Stellar Mass function mock data compared to the SDSS data

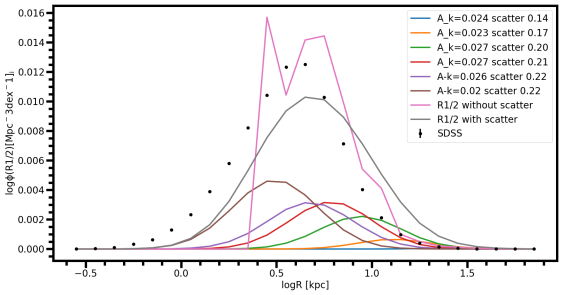

Figure 12. The total size function of mock catalogue when the parameters are changed to match the SMHM of abundance matching

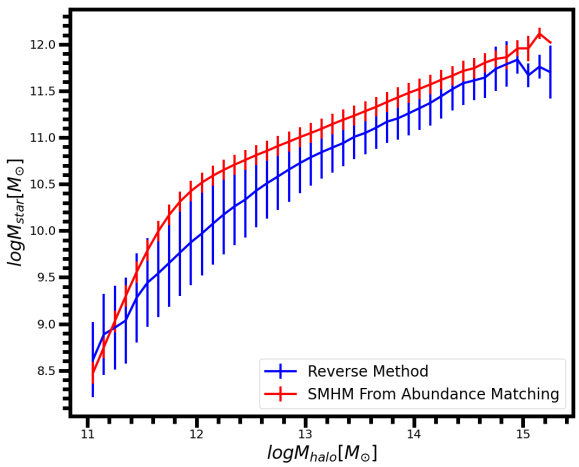

Figure 13. The final SMHM relation from the Reverse Method

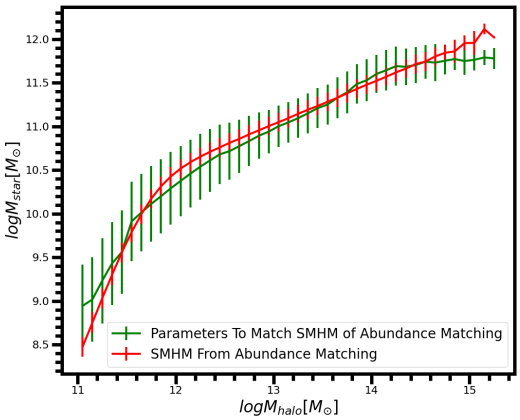

Figure 14. The SMHM relation of Reverse Method when paramters are changed to match the SMHM of Abundance Matching

8 LIMITATIONS

The results of the paper should be considered as an exploratory analysis, rather than a complete methodology. A significant section of this paper relies on the Abundance Matching parameters which were not chosen by an optimisation algorithm. Furthermore, the parameters selected for the Reverse Method were not optimised either. Hence it can be said that in the Reverse Method the true shape of the Kravtsov function is unknown. In principal, there should exist Kravtsov parameters to match the full spectrum of observed data (clustering at the high mass end), which would be a non trivial function of R_{halo}. A full MCMC should be used to fit the parameters to data. If indeed Kravtsov parameters that can fit all the data exist, it would need to be established whether there is a maximum or a minimum in $R_{1/2}/R_{halo}$ (such the one in M_{star}/M_{halo} SMHM). Such finding could shed light on physical mechanisms which occur in galaxy formation.

In addition to the previous limitation, only central galaxies were taken into account in this paper. Satellite galaxies should be included in future work.

9 CONCLUSION

In this work an original method to assign stellar mass to central dark matter haloes was explored, independent from abundance matching. The "Reverse Method" assigns stellar mass from the effective radius of galaxies ($R_{1/2}$ Half-Light Radius). The effective radii was assigned from the virial radii of haloes based on the linear R_{vir}->$R_{1/2}$ relation (Kravtsov), with the stellar masses assigned directly from the observed SDSS data. The Reverse Method has shown to be effective in reproducing the SMF, the size function and clustering except the high mass end $>10^{11}M_\odot$. Compared to this, Abundance Matching was able to reproduce the SMF and the clustering. This suggests that the SMHM relation is more fundamental than the halo size-galaxy size when assigning stellar mass.

Our results can be summarised as follows:

(i) Non trivial Kravtsov parameters are needed to reproduce size function of SDSS with decreasing scatter for Ultra massive galaxies. This implies that for massive galaxies the size distributions arise from the shape of the SMHM relation and hence the physical mechanisms behind it. Whereas for the low mass end a large scatter is needed to reproduce the size function which are caused by more complex unknown mechanisms.

(ii) The reverse method was able to reproduce the total SMF, the

total size function and the clustering up to the high mass end. The clustering fails at the high mass end due to the very large scatters needed to match the size function of the SDSS. The high mass end is particularly sensitive due to a low number of galaxies. The scatter applied could be explained by angular momentum of baryons and dark matter haloes Mo et al. (1998) model, where the scatter was from the spin of parent halo Vitvitska et al. (2002). The clustering at the high mass end was reproduced naturally by Abundance Matching.

iii) The Forward and Reverse methods are fundamentally different. If the same SMHM relation from the Reverse Method was applied to dark matter haloes, the resultant mock catalogue would be vastly different

iv) The final SMHM relation from the Reverse method was not identical to Abundance Matching. The parameters could be adjusted to match the SMHM, but this would come at the expense of the size function and the SMF

This work should be considered as a unique exploratory analysis in assigning stellar mass from galaxy size and structure.This sheds light onto a new approach on populating galaxies to dark matter halo catalogues. In addition to showing how mass is a more fundamental property in correlating haloes and galaxies. The work is limited due to the lack of use of optimisation algorithms when matching the mock data to the observed data. This paper identified ideas further work in including Satellites and using an optimisation algorithm to match parameters to data. Our team led by Prof. Shankar has applied for a further MNRAS grant to further continue this work over the summer and hopefully to publish the results.

REFERENCES

Abazajian K. N., et al., 2009, ApJS, 182, 543

Baldry I. K., Balogh M. L., Bower R. G., Glazebrook K., Nichol R. C., Bamford S. P., Budavari T., 2006, MNRAS, 373, 469

Behroozi P. S., Wechsler R. H., Conroy C., 2013, The Astrophysical Journal, 770, 57

Bower R. G., Benson A. J., Malbon R., Helly J. C., Frenk C. S., Baugh C. M., Cole S., Lacey C. G., 2006, Monthly Notices of the Royal Astronomical Society, 370, 645–655

Bryan G. L., Norman M. L., 1998, The Astrophysical Journal, 495, 80

Buchan S., Shankar F., 2016, Monthly Notices of the Royal Astronomical Society, 462, 2001

Campbell D., van den Bosch F. C., Padmanabhan N., Mao Y.-Y., Zentner A. R., Lange J. U., Jiang F., Villarreal A., 2018, MNRAS, 477, 359

Coe D., 2010, Dark Matter Halo Mass Profiles (arXiv:1005.0411)

Cole S., Lacey C. G., Baugh C. M., Frenk C. S., 2000, Monthly Notices of the Royal Astronomical Society, 319, 168

Contreras S., Angulo R., Zennaro M., 2020, A flexible subhalo abundance matching model for galaxy clustering in redshift space (arXiv:2012.06596)

Diemer B., 2018, ApJS, 239, 35

Dimauro P., et al., 2019, Monthly Notices of the Royal Astronomical Society, 489, 4135

Domínguez Sánchez H., Huertas-Company M., Bernardi M., Tuccillo D., Fischer J. L., 2018, MNRAS, 476, 3661

File S. D. R. . M., 2004

Fischer J.-L., Bernardi M., Meert A., 2017, Monthly Notices of the Royal Astronomical Society, 467, 490

Grylls P. J., Shankar F., Zanisi L., Bernardi M., 2019, MNRAS, 483, 2506

Hamilton A. J. S., Tegmark M., 2004, Monthly Notices of the Royal Astronomical Society, 349, 115

Huang S., Ho L. C., Peng C. Y., Li Z.-Y., Barth A. J., 2013, The Astrophysical Journal, 766, 47

Huang K.-H., et al., 2017, ApJ, 838, 6

Huertas-Company M., et al., 2012, Monthly Notices of the Royal Astronomical Society, 428, 1715

Jeremy B., 2008, The American Astronomical Society., 681, 225

Klypin A. A., Trujillo-Gomez S., Primack J., 2011, ApJ, 740, 102

Kravtsov A. V., 2013, The Astrophysical Journal, 764, L31

Kravtsov A. V., Berlind A. A., Wechsler R. H., Klypin A. A., Gottlober S., Allgood B., Primack J. R., 2004, The Astrophysical Journal, 609, 35

Landy S. D., Szalay A. S., 1993, ApJ, 412, 64

Lange R., et al., 2015, Monthly Notices of the Royal Astronomical Society, 447, 2603

Lapi A., Salucci P., Danese L., 2018, ApJ, 859, 2

Li C., Kauffmann G., Jing Y. P., White S. D. M., Börner G., Cheng F. Z., 2006, MNRAS, 368, 21

López-Sanjuan, C. et al., 2012, A&A, 548, A7

Meert A., Vikram V., Bernardi M., 2015a, MNRAS, 446, 3943

Meert A., Vikram V., Bernardi M., 2015b, Monthly Notices of the Royal Astronomical Society, 455, 2440

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

Moster B. P., Somerville R. S., Maulbetsch C., van den Bosch F. C., Macciò A. V., Naab T., Oser L., 2010, The Astrophysical Journal, 710, 903

Moster B. P., Naab T., White S. D. M., 2012, Monthly Notices of the Royal Astronomical Society, 428, 3121

Mowla L., Wel A. v. d., Dokkum P. v., Miller T. B., 2019, The Astrophysical Journal, 872, L13

Navarro J. F., Frenk C. S., White S. D. M., 1996, ApJ, 462, 563

Norberg P., et al., 2001, Monthly Notices of the Royal Astronomical Society, 328, 64

Norberg P., et al., 2002, Monthly Notices of the Royal Astronomical Society, 332, 827

Posti L., Marasco A., Fraternali F., Famaey B., 2019, Astronomy & Astrophysics, 629, A59

Prada F., Klypin A. A., Cuesta A. J., Betancort-Rijo J. E., Primack J., 2012, MNRAS, 423, 3018

Riebe K., et al., 2013, Astronomische Nachrichten, 334, 691

Rohin Y., 2017, correlcalc: A 'Generic' Recipe for Calculation of Two-point Correlation function (arXiv:1710.01723)

Shankar F., Lapi A., Salucci P., Zotti G. D., Danese L., 2006, The Astrophysical Journal, 643, 14

Shen S., Mo H. J., White S. D. M., Blanton M. R., Kauffmann G., Voges W., Brinkmann J., Csabai I., 2003, Monthly Notices of the Royal Astronomical Society, 343, 978

Sinha M., Garrison L. H., 2020, MNRAS, 491, 3022

Somerville R. S., et al., 2017, Monthly Notices of the Royal Astronomical Society, 473, 2714

Springel V., 2006, Nature, 440, pages1137–1144

Tamburri, S. Saracco, P. Longhetti, M. Gargiulo, A. Lonoce, I. Ciocca, F. 2014, A&A, 570, A102

Vale A., Ostriker J. P., 2004, Monthly Notices of the Royal Astronomical Society, 353, 189

Vitvitska M., Klypin A. A., Kravtsov A. V., Wechsler R. H., Primack J. R., Bullock J. S., 2002, ApJ, 581, 799

Wetzel A. R., Tinker J. L., Conroy C., van den Bosch F. C., 2013, Monthly Notices of the Royal Astronomical Society, 432, 336

White S. D. M., 1997, pp 3–14

White S. D. M., Frenk C. S., 1991, ApJ, 379, 52

White S. D. M., Rees M. J., 1978, MNRAS, 183, 341

Yang X., Mo H. J., Jing Y. P., Van Den Bosch F. C., Chu Y., 2004, Monthly Notices of the Royal Astronomical Society, 350, 1153

Yang X., Mo H. J., Van Den Bosch F. C., Jing Y. P., 2005, Monthly Notices of the Royal Astronomical Society, 357, 608

Yang X., Mo H. J., van den Bosch F. C., Pasquali A., Li C., Barden M., 2007, The Astrophysical Journal, 671, 153

Zanisi L., et al., 2020, MNRAS, 492, 1671

Zheng Z., et al., 2005, Astrophysical Journal, 633, 791

van Daalen M. P., Henriques B. M. B., Angulo R. E., White S. D. M., 2016, Monthly Notices of the Royal Astronomical Society, 458, 934

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