Halo Definition and Environmental Effects

Antonio Villarreal[1]⋆, Andrew R. Zentner[1]†, Christopher W. Purcell[2]‡

[1] *Department of Physics and Astronomy &*
Pittsburgh Particle Physics, Astrophysics, and Cosmology Center (Pitt-PACC),
University of Pittsburgh, Pittsburgh, PA
[2] *Department of Physics and Astronomy,*
West Virginia University, Morgantown, WV

In preparation

ABSTRACT
Recent work has shown the importance of environment to the properties of dark matter halos. This brings conflict to standard implementations of the halo model and excursion set theory which assume that the properties of a population within the halo is determined by the mass of the halo alone. We seek to find a definition of the size of a halo that allows us to circumvent these environmental effects. We analyze the dependence on environment of our properties using the method of marked correlation functions for several different halo definitions, utilizing the Consuelo and Diemer et al simulations. We find that environmental dependencies are dramatically different as we vary the definition of the halo radius in terms of the overdensity parameter Δ. Further, we analyze the mass dependence on these effects and how they factor into the ultimate applicability of our methodology.

Key words: dark matter – galaxies: halos – galaxies: formation – large-scale structure of universe – methods: numerical

1 INTRODUCTION

ARZ:We will need to work on the introduction considerably as we get a better handle on the final results. The first two paragraphs can probably be combined into a single shorter paragraph. I also like to end the first paragraph by telling the reader what it is that we aim to do in the paper.

In the current concordance cosmology, the creation of observed galaxies and clusters is often seen as arising from the hierarchical mergers of dark matter halos.(White & Rees 1978). Being able to model the properties of dark matter halos and the galaxies within gives us a potential probe for the physical processes that go into galaxy formation. The excursion-set formalism of galaxy clustering (Bond et al. 1991; Lacey & Cole 1993; Somerville & Kolatt 1999; Zentner 2007) and the standard halo model of galaxy clustering (Seljak 2000; Peacock & Smith 2000; Scoccimarro et al. 2001; Berlind & Weinberg 2002; Bullock et al. 2002; Cooray & Sheth 2002) both can help us in this task, but rely on underlying asumptions. The first is that the statistics of the objects within a dark matter halo is a function of the mass alone.

The second is that the clustering of dark matter halos is a function of mass as well. In this paper we propose a simple redefinition of halo size that will help correct for inaccuracies in these two assumptions.

It has previously been demonstrated that the clustering of halos is dependent on not only dependent on the mass, but also on the formation time of the halo (Sheth & Tormen 2004; Gao et al. 2005; Croton et al. 2007). Furthermore, it has been shown that the clustering of a given halo is dependent on the halo concentration (Wechsler et al. 2006). This necessitates corrections be made to the standard implementations to account for this. More complicated methodologies have been made to extend to using merger histories directly from simulation (Dvorkin & Rephaeli 2011), as well as to account for the dependence due to concentration (Gil-Marín et al. 2011). The relationship that clustering has to the properties of the halo can be described as environmental effects.

Our method of halo redefinition is motivated by the size of a halo being an ill-defined quantity. What if often referred to as the "virial radius" will not contain all gravitational bound dark matter particles in the halo (Kazantzidis et al. 2006). Rather, it is a matter of convention that does not have a common definition. We choose to define a halo radius in a way that encompasses the nearby environmental effects. These may be due to large scale structure or driven by baryonic physics. Using a

⋆ E-mail: asv13@pitt.edu
† E-mail: zentner@pitt.edu
‡ E-mail: cwpurcell@mail.wvu.edu

simulated box, we can then test how the redefinition of the halo size affects the relationship between the clustering and the properties of the halo. In the case that the properties of the halo become independent of the clustering, it is possible to utilize standard implementations of the halo model without necessitating more complicated theory.

In § 2 of this paper, we consider the simulated cosmological boxes that we utilize for our statistics and our method of halo size redefinition. In § 3, we discuss the statistics that we have used in order to test for environmental effects and the removal of known mass scaling from these statistics. In § 4, we present our results on our tests and draw conclusions as to the natures of the encountered environmental effects. In § 5, we discuss the significance of redefining environmental effects through a redefinition of halo environments and discuss possible applications of our methodology.

2   SIMULATION DATA AND HALO FINDING

In order to study the effects of halo redefinition, we make extensive use of two simulation suites. The Diemer et al simulations each utilize a Planck best fit cosmology with $\Omega_M = 0.32$, $\Omega_\Lambda = 0.68$, and $h_o = 0.67$. We use three simulation boxes with comoving sizes of 125, 250, and 500 $h^{-1}\text{Mpc}$ respectively. The particle mass of $1.6 \times 10^8 \ h^{-1}M_\odot$, with a total of 1024^3 particles in each simulation. This set of simulations allows us to probe the resolution effects on halo finding, as each box has a factor of eight difference in effective resolution. We also analyze the Consuelo simulation from the Large Suite of Dark Matter Simulations (LasDamas) (McBride et al, in prep? I literally cannot find this paper despite it being in prep since 2012!) :which ran with a cosmology with $\Omega_M = 0.25$, $\Omega_\Lambda = 0.75$, and $h_o = 0.7$. This simulation box has a comoving size of 420 $h^{-1}\text{Mpc}$ and a particle mass of $1.87 \times 10^9 \ h^{-1}M_\odot$, with a total of 1400^3 particles. The difference in cosmology between the two sets of boxes will allow us to determine if this has drastic implications on our method.

For halo finding we use the ROCKSTAR halo finder, which works on the phase space algorithm described in Behroozi, Wechsler & Wu (2013). In short, ROCKSTAR determines the initial groupings using a FOF algorithm in phase space, before then applying the spherical overdensity case in order to determine the properties of interest. Unbound particles are removed prior to the calculation of halo mass and other properties of interest. In addition, we take interest of the shape of the halo, which is determined through the sorted eigenvalues of the inertia tensor, with principal ellipsoid axes defined such that $a > b > c$. Some of these parameters are determined after fitting the particles to the Navarro-Frenk-White profile (Navarro, Frenk & White 1997).

Given the nature of the halo size as a hard to define statistic, we choose to redefine the halo in terms of the overdensity parameter, Δ, by defining the halo radius, r_Δ, as follows:
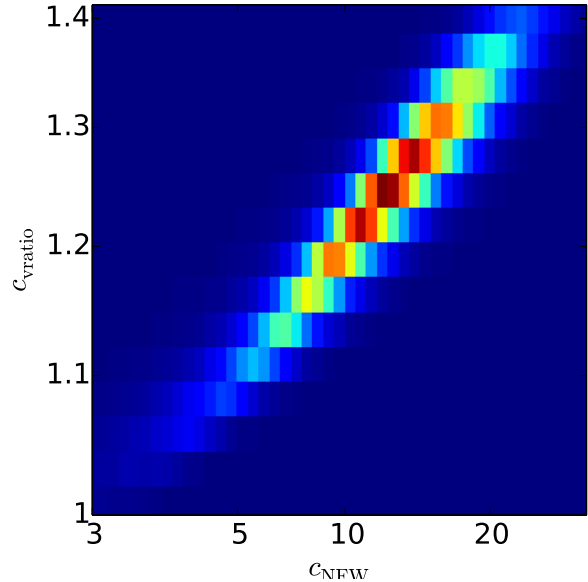
Figure 1. The relationship between the two different marks of concentration, using halos in the Consuelo simulation box.

$$\rho(r_\Delta) = \Delta\rho_b \tag{1}$$

Here, ρ_b is the background mass density of the simulation box, while ρ is the density measured within at a radius under the assumption that it uniformly distributed in a sphere of radius r_Δ. It should be noted that the use of the critical density or the mean density is interchanged within the literature, though the change is relatively mild. We allow the overdensity parameter Δ to vary over the range of the fiducial definition of $\Delta = 200$ down to values as extreme as $\Delta = 10$. As discussed above, this change is motivated due to the fact that the definition of halo size is primarily a matter of convention.

With the combination of this analysis and the ROCKSTAR data, we define the following properties of interest:

$$c_{\text{vratio}} = \frac{v_{\text{max}}}{v_\Delta}; \quad c_{\text{NFW}} = \frac{r_\Delta}{r_s}; \quad s = \frac{c}{a}; \Lambda = \frac{J\sqrt{|E|}}{GM_{vir}^{2.5}} \tag{2}$$

The first two marks listed above are different proxies for the concentration of the halo. The relationships between these marks are shown in Figure 1 and it can be seen that they can be related to each other with some scatter. The ROCKSTAR catalogs give us the max circular velocity of a hypothetical particle, v_{max}, the scale radius, r_s, the halo radius, r_Δ, both ellipsoidal axis parameters c and a, and the spin parameter, Λ as introduced by (?) in terms of the halo angular momentrum J and the total energy of the halo E. The circular velocity of a particle at the halo radius is calculated by $v_\Delta = \sqrt{\frac{GM(<r_\Delta)}{r_\Delta}}$.
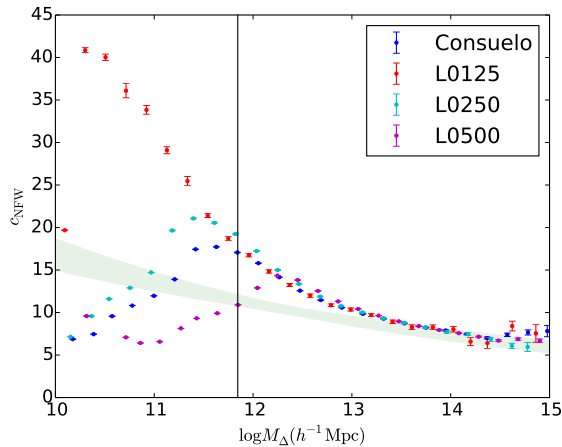
Figure 2. An example of the chosen lower limit on halo mass for our sample is marked as a black line, for the $\Delta = 200$ case. The shaded blue region depicts the spread on the power in the concentration-mass relationship proposed by Duffy et al. (2008). The shaded region illustrates the implied two-sigma band of Duffy et al. (2008) mass-concentration relation. At lower mass, halos are ill defined due to resolution limits.

3   METHODOLOGY

There are several well understood effects in our simulation that must be accounted for before we can draw any conclusions. The first is the matter of simulation resolution. In each of our generated halo catalogs, we can anticipate finding artificial halos. While existing under our constraint of having only a set overdensity, they will have properties that conflict with well known trends that have been shown in prior works, such as concentration decreasing as a function of halo mass (Wechsler et al. 2006). As shown in Figure 2, we can roughly identify the region in which these artificial halos become significant by looking for the turnover in the plot. In order to avoid our halo finding statistics being thrown off by these artificial halos, we set a lower mass threshold, limiting the sample to only those halos that are physically significant. We choose to use the higher mass constraint between the Max and Consuelo simulation boxes for consistency.

Another well known effect is the scaling of our properties as a function of halo mass, as well demonstrated within the literature (Duffy et al. 2008). We are interested in the clustering behavior beyond this well known effect, so we seek to remove this scaling. We take all host halos of interest and sort them by their halo mass. Each set of halo properties are then placed in bins of equal population in order to assure that enough data points exist for robust statistics. We then normalize each mark in a given bin to the mean value. The end result is halo properties that do not change as a function of mass, allowing us to more accurately analyze clustering behavior within our simulation.

To normalize the satellite number we follow the prescription of Wechsler et al. (2006). In addition to the mass cutoffs on our data, we eliminate ill-resolved satellites by choosing a cutoff in $v_{\max}$ for the host halo. We then attempt to match subhalos to this mass halo,

making a secondary cutoff in the case that the value of $v_{\max,\mathrm{sub}}/v_{\max,\mathrm{host}}$ is not above a threshold value. We set these two parameters such that there are no isolated or poorly resolved subhalos in our data sample.

In order to test for environmental effects, we choose to utilize two primary methods. The first is to look at the two-point correlation functions of the host halo catalog and compare it against the correlation functions of the top and bottom 20% of most concentrated host halos. Assuming that each halo was statistically similar, we would expect that there would be no discernible difference between these correlation functions in the case where there are no environmental effects. We utilize a jackknife method in order to estimate the errors on these correlation functions.

The second method is to use a weighted correlation function, often referred to in the literature as the marked correlation function (MCF). Our marks in this case are the log values of the halo concentration proxies described previously, the halo shape, and the halo satellite number. The use of this tool in measuring clustering has been shown previously in Wechsler et al. (2006) or Harker et al. (2006). We compare our MCFs to an error bar representing the approximate 2-sigma confidence region generated by randomizing the marks among our halos 400 times and taking the 10th lowest and 390th highest values of the mark among the randomizations. In the case of no environmental effects upon the property of interest, one would expect that the MCF would fall within the region bounded by these limits.

4   RESULTS AND DISCUSSION

Ultimately, as a result of the limited nature of simulation data, we must make some lower mass cutoff in our study of halo properties. Table 1 describes where we have chosen to place these cuts in a variety of different measures. The "hosts" set of cuts has been chosen for our primary form of analysis, as it it performs well across all four halo boxes. Figure 2 shows the example of this cutoff in the $\Delta = 200$ case. While there may be some concern about the number of low mass halos included in the L0500 simulation box, results seem to demonstrate that the effect is somewhat minimal compared to other factors. This is quite likely related to our methodology by which we remove the inherent mass dependence, as provided that the ranking of the halos in some mass bin does not significantly change, it should result in the same general trends. The remaining two cuts have been chosen to explore any resolution related effects. The "l0125" cuts explore the lowest possible mass cutoff that excludes ill resolved halos in the Diemer et. all simulations, while the "l0500" set of cuts represent the mass cutoff chosen to avoid contamination due to resolution effects in our largest simulation box.

Figure 3 demonstrates the results of our method as we change the value of Δ. It can be seen that for between the values of $\Delta = 75$ and $\Delta = 50$ that both the high and low concentration cuts of the data can be brought into close agreement in the Diemer et. al boxes. This same feature is seen in the Consuelo box, at a higher value of
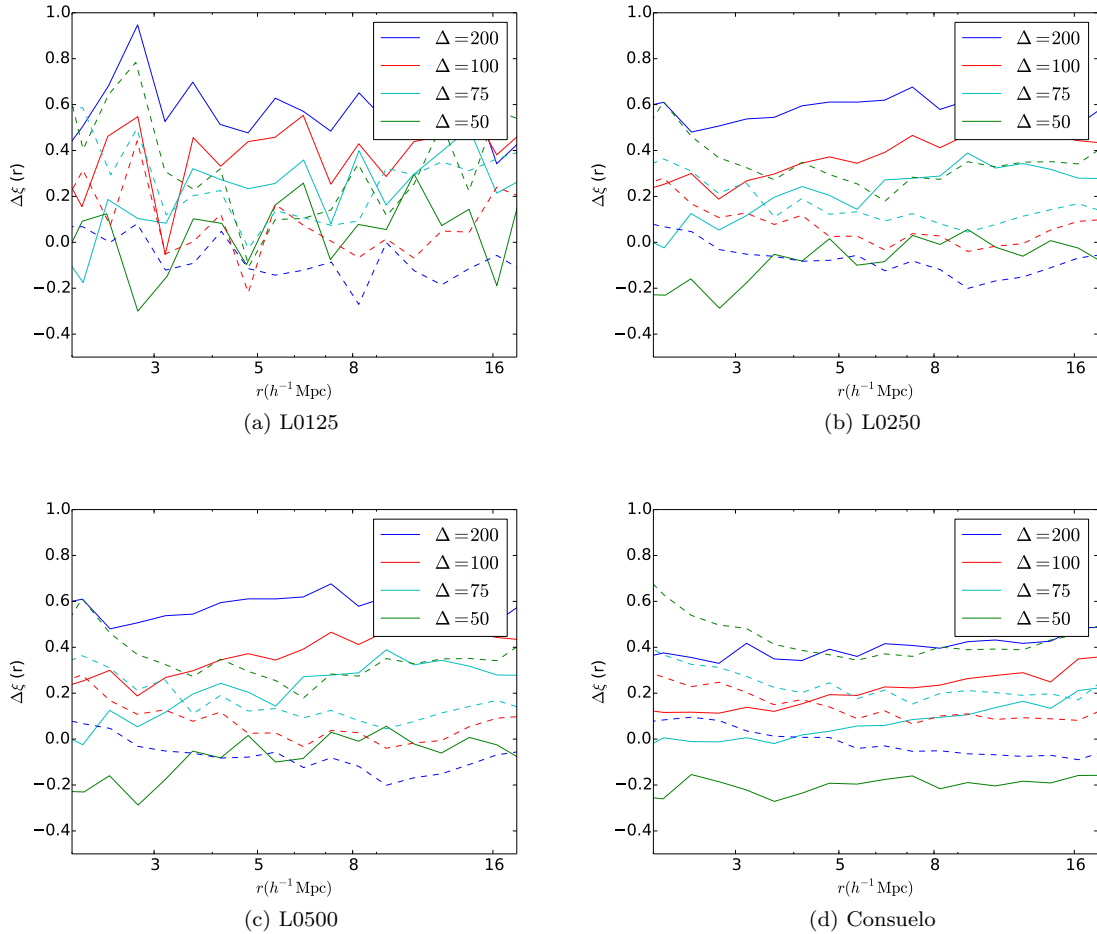
Figure 3. The difference of the correlation function of all hosts compared to a cut of the 20 % highest or lowest NFW defined concentration halos. The solid lines represent the difference from the highest concentration cut, the dashed lines represent the difference from the lowest concentration cut.

Table 1. Chosen mass cuts on the simulation boxes, in units of $h^{-1}M_{\odot}$. The values of the halo overdensity parameter Δ chosen are listed in the first row.

	200	100	75	50
"hosts"	7E11	8E11	9E11	1.5E12
"l0125cut"	7E10	8E10	9E10	1E11
"l0500cut"	4E12	5E12	6E12	7E12

Δ between 100 and 75. It should be noted that both the high and low concentration cuts move closer toward a state with environmental dependencies being thoroughly removed within the margin of error from opposite directions in terms of clustering. Changing the value of Δ beyond the observed "sweet spot" appears to reintroduce the problem we are attempting to remove.

The NFW defined concentration MCF is shown in Figure 4. It can been that as we decrease the value of the overdensity parameter Δ to lower values down to $\Delta = 50$. It can be seen that at scales of $r > 10\ h^{-1}\mathrm{Mpc}$, environmental effects are removed between $\Delta = 75$ and

$\Delta = 50$ in the Diemer et al. boxes and $\Delta = 100$ and $\Delta = 75$ in the Consuelo box.

The velocity ratio defined concentration MCF is shown in Figure 5. It can be immediately seen that many of the same features from the previous MCF can be seen in the different concentration proxy. We see that at large scales, the same Δ values are preferred in both the Diemer et al. and Consuelo boxes as in the previous figure. Combined with our previous result, this allows for this halo redefinition to prove useful for studies at this scale with regard to concentration.

Figure 6 demonstrates a case in which the halo property does not have environmental dependence removed under our simple method of halo redefinition. In particular, with decreasing values of Δ, we only serve to increase the strength of this environmental dependence. This result seems to indicate a process occurring to drive the shape of the halo is on scales that are significantly larger than the size of the halo, such that our method will not be able to encompass these effects. A similar result is located in both Figure 7 and Figure 8, in which the spin parameter mark and satellite number mark experience the exact same effect.
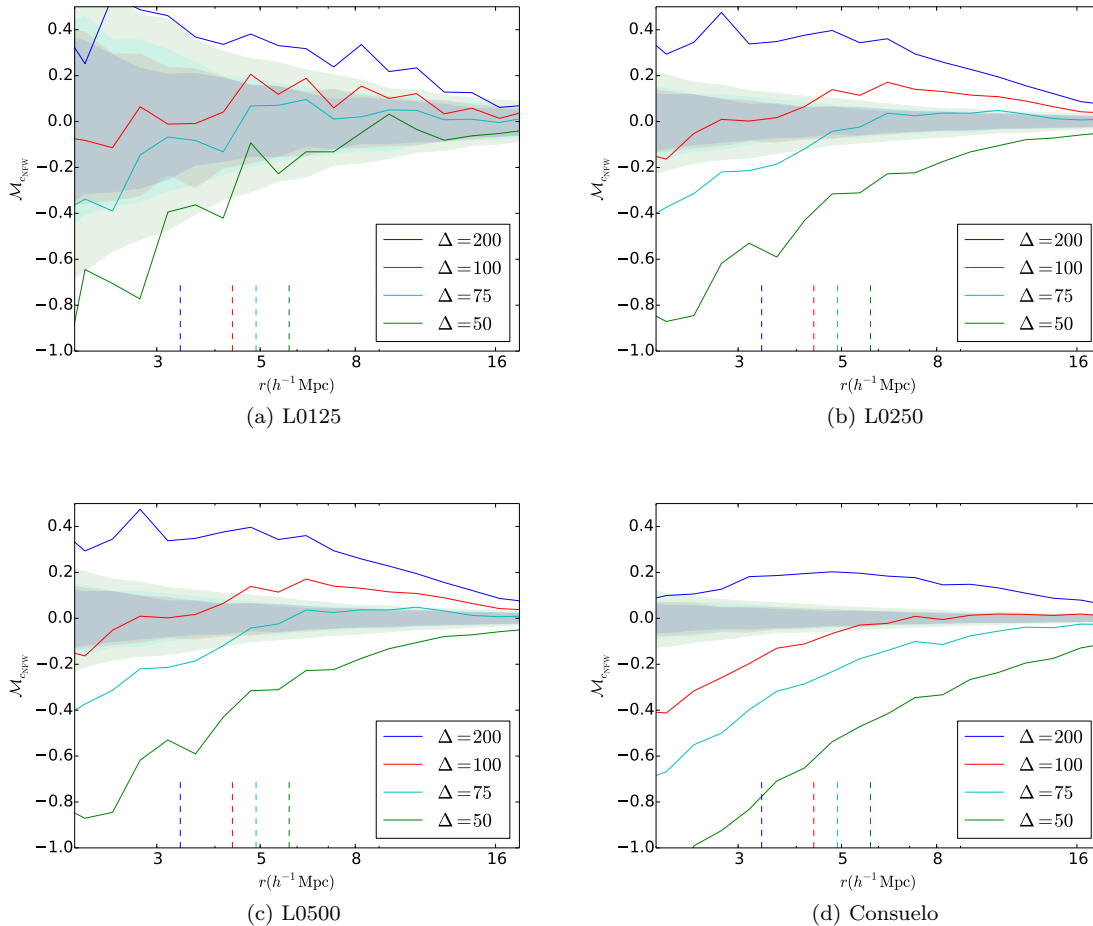
Figure 4. The marked correlation function for the concentration defined according to the NFW profile. The shaded bands represent 2-sigma confidence regions generated by randomization of the marks. The dashed line denotes the largest halo radius for a given value of the overdensity parameter.

Overall, while some statistics of the halo can be made to become independent of environmental effects, there exist several that cannot. These environmental effects are not negligible when considering applications of the halo model. In contrast, if one is only interested in the concentration of the halo, it is feasible to take advantage of our methodology in order to remove environmental effects. In contrast, it is not possible to apply the halo model directly to statistics such as the number of satellite halos or the shape without accounting for these environmental effects. We do not yet have a full understanding of what leads to the visible effects that we see on some of these marks, but we present a series of cases in which an intuitive understanding may be gained as to the root causes.

In the case that the halo can be described well by an NFW profile, one expects a direct relationship between the NFW defined halo concentration and the velocity ratio defined halo concentration. While some variation can be expected due to halos not perfectly being fit by an NFW profile, we do see that the features in one concentration proxy are mirrored in the other. This allows for the two concentration markers to support each other well

with regards to our ability to remove environmental effects on large scales.

Halo shape and satellite number are statistics that do not end up having their environmental effects removed and can even be made more prominent by our methodology. One intuitive way to consider the former statistic is in the context of the cosmic web. Studies have shown a statistically significant alignment between filaments and satellite galaxy position (Tempel et al. 2015; Velliscig et al. 2015). Our method then expands the halo radius and subsumes material that was previously outside of the halo. A simple graphic illustrates this potential effect in Figure 8. As there is a preferential distribution of these satellites that are being subsumed into the halo, this would serve to induce a shape to the halo that would then be determined by Rockstar. In addition, as our satellite number is chosen by those halos within the halo radius, we anticipate that the most clustered regions would see the largest increase in satellite count and thus see an increase in the satellite number mark.

The "sweet spot" behavior of this method is also of interest to us. The halo redefinition process that we use serves to decrease halo clustering for the most con-
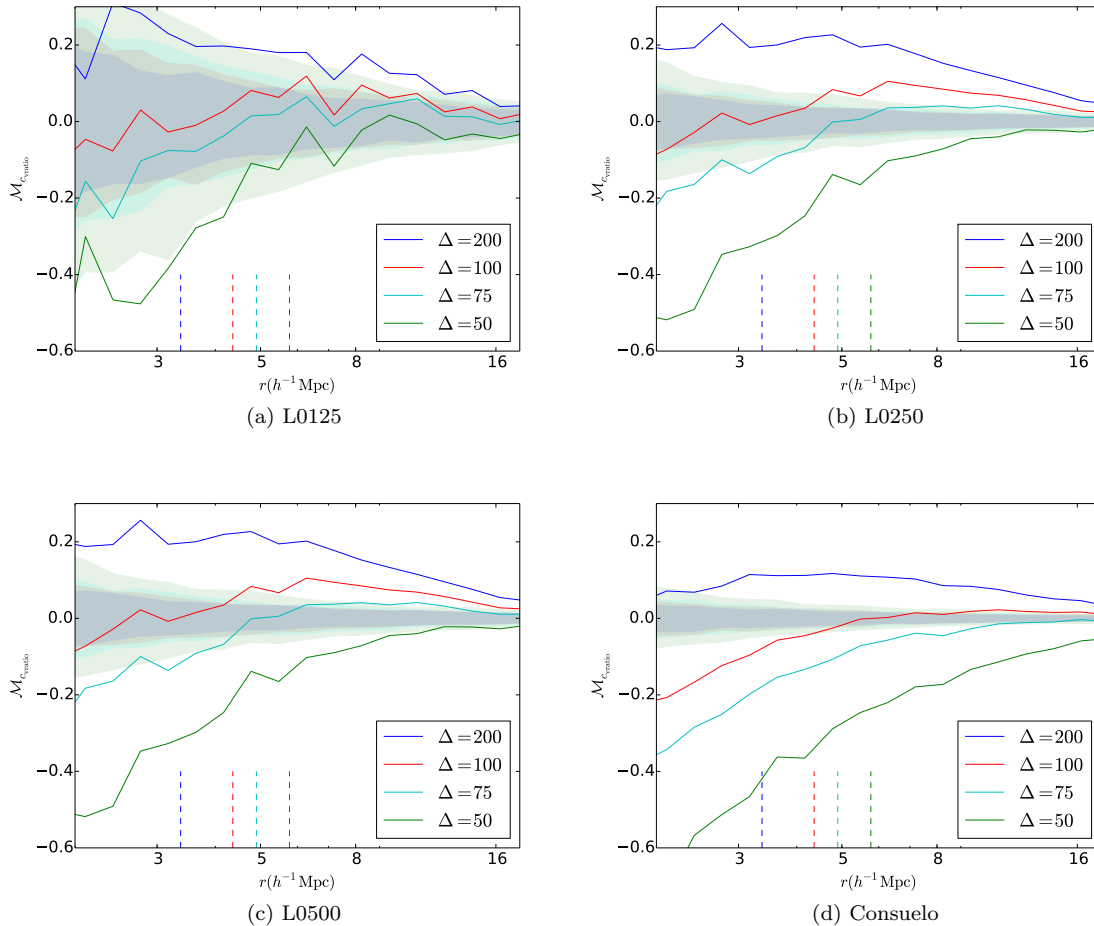
Figure 5. The marked correlation function for the concentration defined according to the velocity ratio. The shaded bands represent 2-sigma confidence regions generated by randomization of the marks. The dashed line denotes the largest halo radius for a given value of the overdensity parameter.

centrated halos and increase halo clustering for the least concentrated halos. In the case of the high concentration cut, the reduced clustering can be seen as a result of halo exclusion. As we exchange collections of tightly clustered smaller halos for a single larger halo, then the most prominent contribution of the two-point correlation function will be reduced at all scales.

Given the differences between the simulation results and with the data available to us with the Diemer et al. simulations, we then explore the effect of the choice in mass cut. The low mass cut effects are shown in Figures 10 through 15. This cut explores the area in which we are including ill-resolved halos potentially into the simulation. Namely, utilizing this mass cut on the L0250 simulation box data includes halos that do not fit the form of the expected monotonic halo mass-concentration relation. We can determine several facts from this exercise in resolution testing. The first is that the general behavior of the marks seems to be the same regardless of the influence of the resolution effects. Decreasing the value of Δ still moves our marked correlation functions in the same directions as the previous mass cut, although often from a very different starting location. A second interest-

ing fact is that despite the L0250 simulation potentially having ill-resolved halos included in the sample, the effect on the overall marked correlation function result is relatively minor. Both samples seem to converge to the same answers within simulation variance. Finally, it should be noted that an unreasonable amount of reduction in Δ is necessary to bring the lowest mass halos into a region where they will work with this method. We will further explore this effect in the highest mass cuts.

5 CONCLUSIONS

We have looked at how to use CFs and MCFs in order to analyze the environmental effects upon the properties of the halo. We have suggested a method of removing the mass dependence that is not subject to the small number statistics at large halo masses. Taking our various tests, we then apply a change to the threshold density Δ in an attempt to remove the effect that environment has upon these properties. We come to the following conclusions from our simulation data.
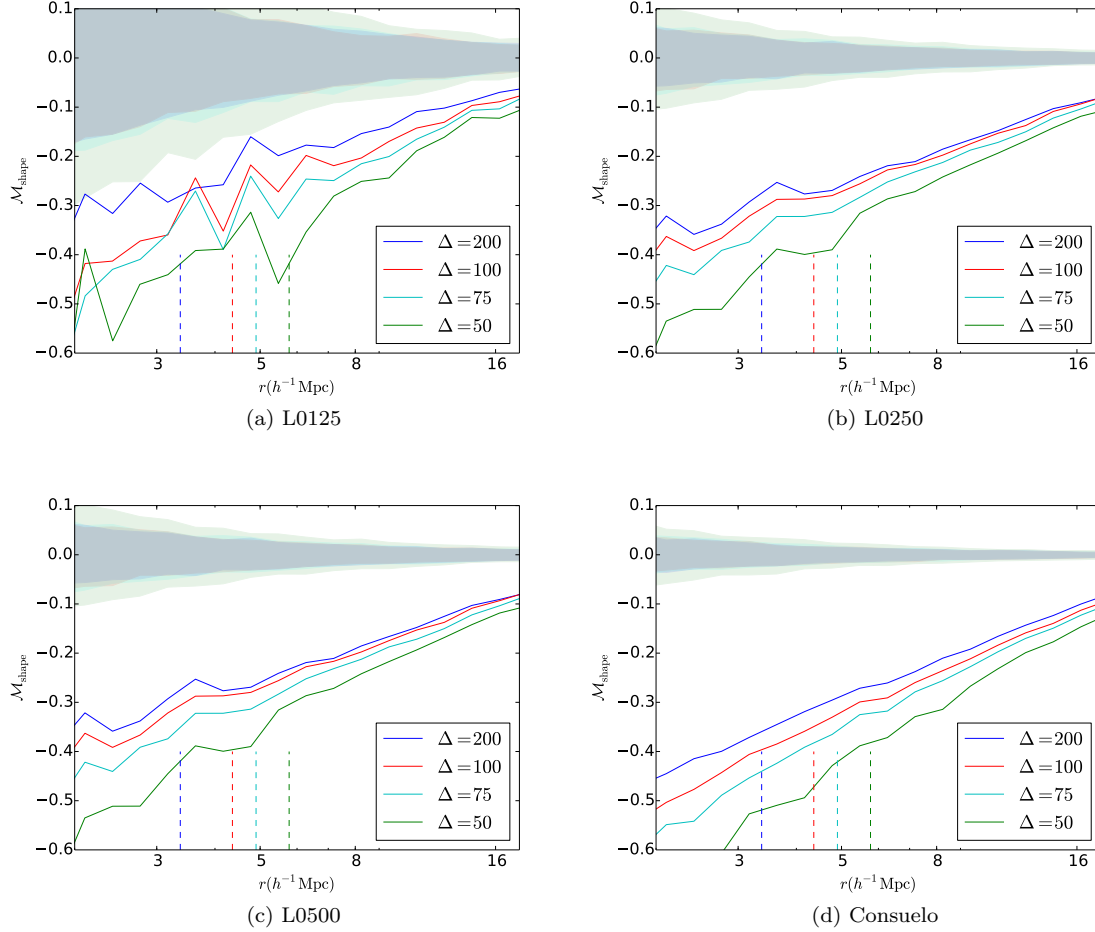
Figure 6. The marked correlation function for the shape of the halo. The shaded bands represent 2-sigma confidence regions generated by randomization of the marks. The dashed line denotes the largest halo radius for a given value of the overdensity parameter.

• Our halo redefinition method does not cause any substantial breakdown in the ROCKSTAR halo finding algorithm, though this may not be the case for every halo finding methodology. This is something that should be considered prior to utilization of this method, unless working directly from particle data. As our initial halo sizes and locations are determined through spherical over-densities, it cannot be assumed that starting from a FoF grouping and then determining values through particle data directly will produce identical results. Similarly, different cosmologies may remove environmental effects at different scales.

• When looking at the two-point correlation function, there appears to be a "sweet spot" that appears to remove environmental effects the most efficiently. Going beyond that seems to reintroduce environmental effects, possibly as an extreme side effect of halo exclusion.

• For our marked correlation functions we see that both proxies of concentration that we use as marks show significant removal of environmental effects at large scales for similar values of the overdensity parameter Δ. In cases where one is only interested in the concentration of dark matter halos and large scales (or correspondingly small values of k), this method will allow you to compensate for bias that environment could introduce to calculations dependent upon the halo model. This may prove valuable for calculations such as that of the shear power spectrum calculated through weak lensing.

• The environmental effects on the shape of the host halo and the satellite number of the host halo cannot be removed regardless of the chosen redefinition of Δ. We propose that this may be intrinsically tied to the nature of the filaments, whose effects cannot be removed by a simple redefinition of the halo radius.

This methodology, while certainly not perfect in removing environmental effects, may be of significance when applied to galaxy formation models. Provided that the properties of interest in a given model behave well under our redefinition, it will allow us to create better mock galaxy catalogs while avoiding the trouble that baryonic physics represent. This may allow us to better study the properties of galaxies within halos without having to substantially increase computation time.
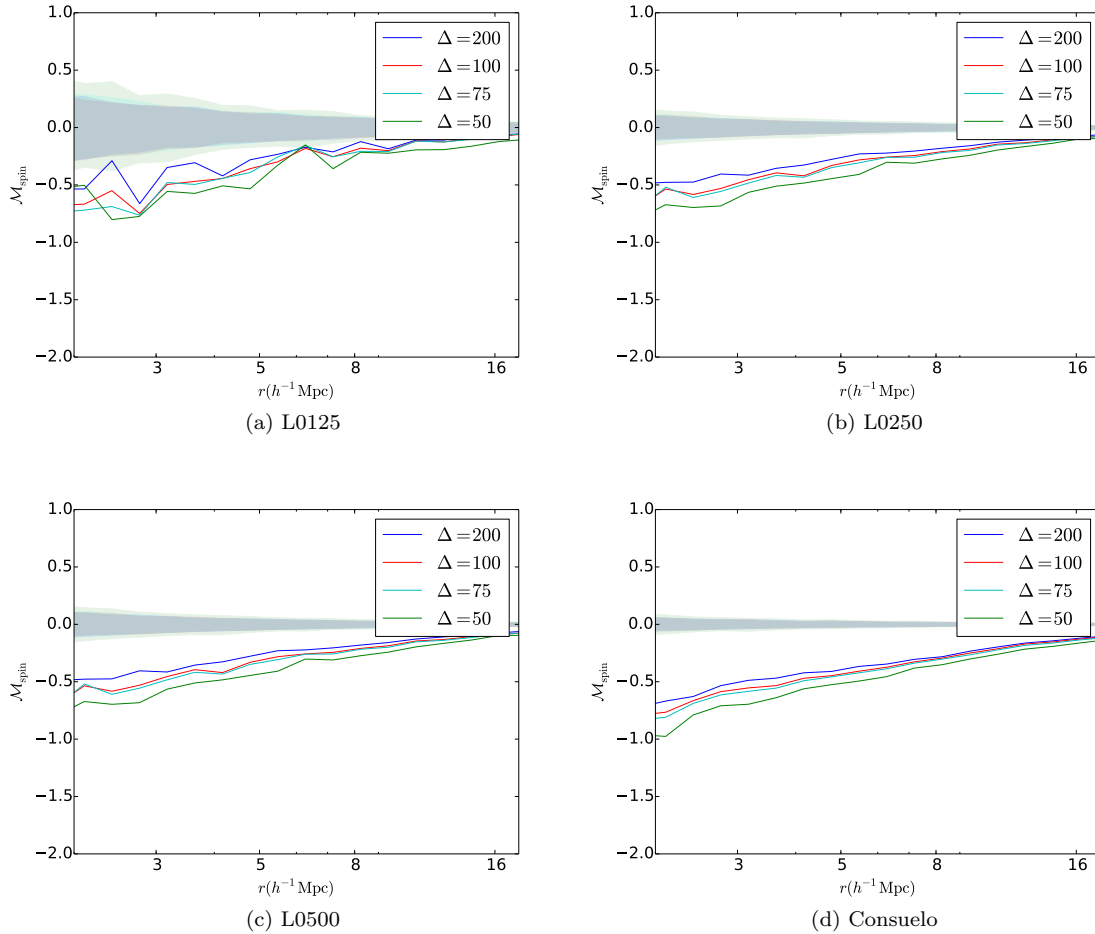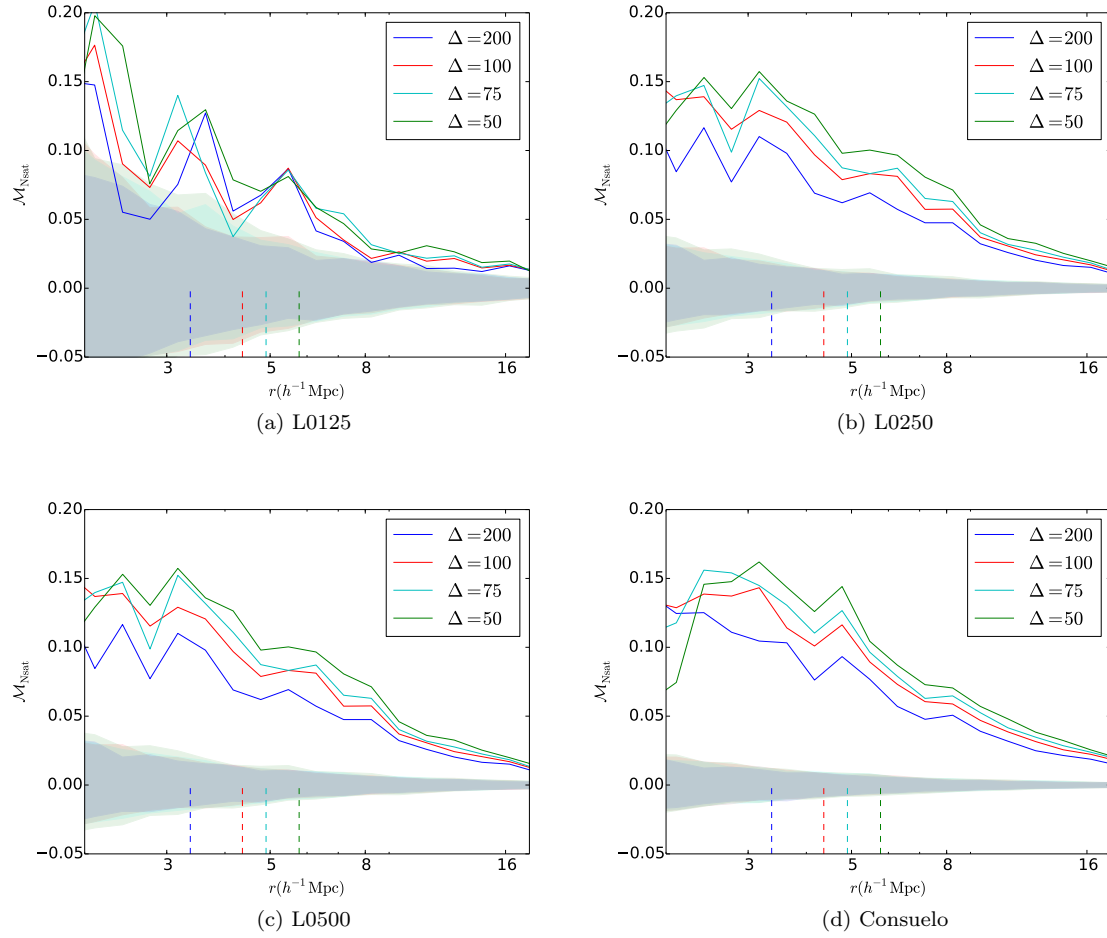
Figure 7. The marked correlation function for the spin parameter of the halo. The shaded bands represent 2-sigma confidence regions generated by randomization of the marks. The dashed line denotes the largest halo radius for a given value of the overdensity parameter.

REFERENCES

Behroozi P. S., Wechsler R. H., Wu H.-Y., 2013, Astrophys. J. , 762, 109

Berlind A. A., Weinberg D. H., 2002, Astrophys. J. , 575, 587

Bond J. R., Cole S., Efstathiou G., Kaiser N., 1991, Astrophys. J. , 379, 440

Bullock J. S., Wechsler R. H., Somerville R. S., 2002, MNRAS , 329, 246

Cooray A., Sheth R., 2002, Phys. Rep., 372, 1

Croton D. J., Gao L., White S. D. M., 2007, MNRAS , 374, 1303

Duffy A. R., Schaye J., Kay S. T., Dalla Vecchia C., 2008, MNRAS , 390, L64

Dvorkin I., Rephaeli Y., 2011, MNRAS , 412, 665

Gao L., White S. D. M., Jenkins A., Frenk C. S., Springel V., 2005, MNRAS , 363, 379

Gil-Marín H., Jimenez R., Verde L., 2011, MNRAS , 414, 1207

Harker G., Cole S., Helly J., Frenk C., Jenkins A., 2006, MNRAS , 367, 1039

Kazantzidis S., Zentner A. R., Kravtsov A. V., 2006, Astrophys. J. , 641, 647

Lacey C., Cole S., 1993, MNRAS , 262, 627

Navarro J. F., Frenk C. S., White S. D. M., 1997, Astrophys. J. , 490, 493

Peacock J. A., Smith R. E., 2000, MNRAS , 318, 1144

Scoccimarro R., Sheth R. K., Hui L., Jain B., 2001, Astrophys. J. , 546, 20

Seljak U., 2000, MNRAS , 318, 203

Sheth R. K., Tormen G., 2004, MNRAS , 350, 1385

Somerville R. S., Kolatt T. S., 1999, MNRAS , 305, 1

Tempel E., Guo Q., Kipper R., Libeskind N. I., 2015, ArXiv e-prints

Velliscig M., Cacciato M., Schaye J., Bower R. G., Crain R. A., van Daalen M. P., Dalla Vecchia C., Frenk C. S., Furlong M., McCarthy I. G., Schaller M., Theuns T., 2015, ArXiv e-prints

Wechsler R. H., Zentner A. R., Bullock J. S., Kravtsov A. V., Allgood B., 2006, Astrophys. J. , 652, 71

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

Zentner A. R., 2007, International Journal of Modern Physics D, 16, 763

Figure 8. The marked correlation function for the satellite number. The shaded bands represent 2-sigma confidence regions generated by randomization of the marks. The dashed line denotes the largest halo radius for a given value of the overdensity parameter.
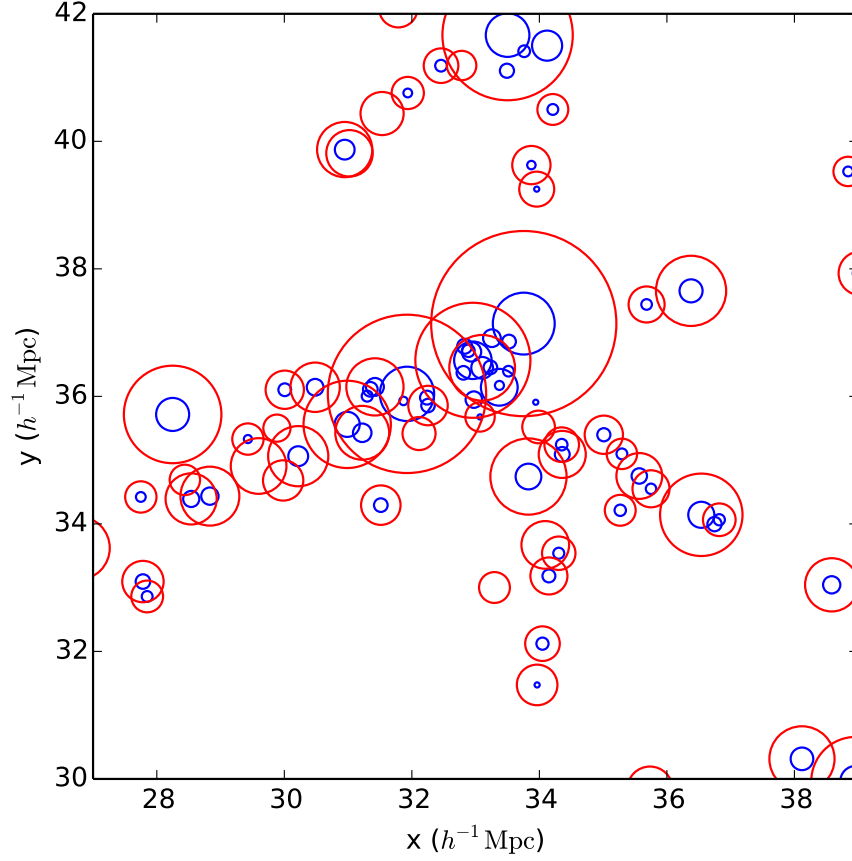
Figure 9. A 5 h^{-1}Mpc deep cut of the Consuelo simulation box along the z-axis. This zoom-in demonstrates the process that decreases the shape parameter as a function of clustering. The size of each circle represents the projection of a spherical dark matter halo with a given halo radius onto the x-y plane. Blue circles use the $\Delta = 200$ catalog and red circles use the $\Delta = 10$ catalog in order to make the effect more visible.
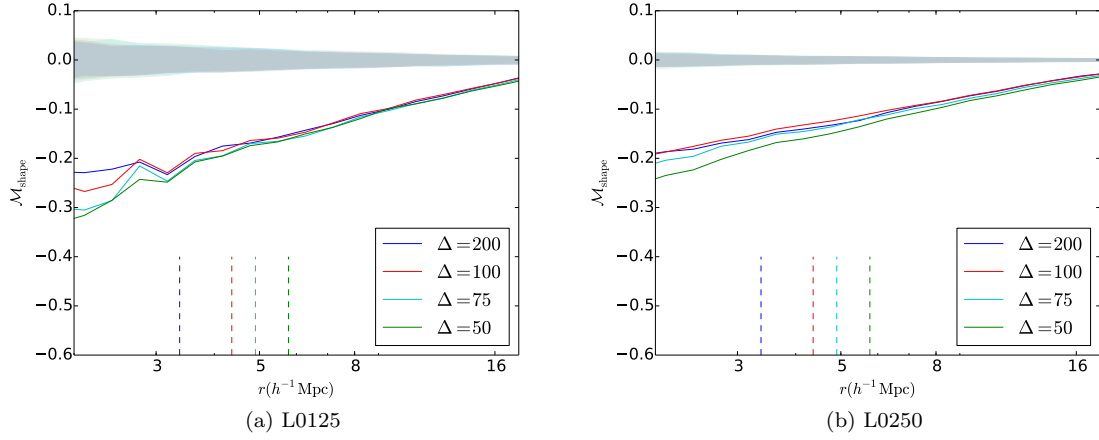

(a) L0125

(b) L0250
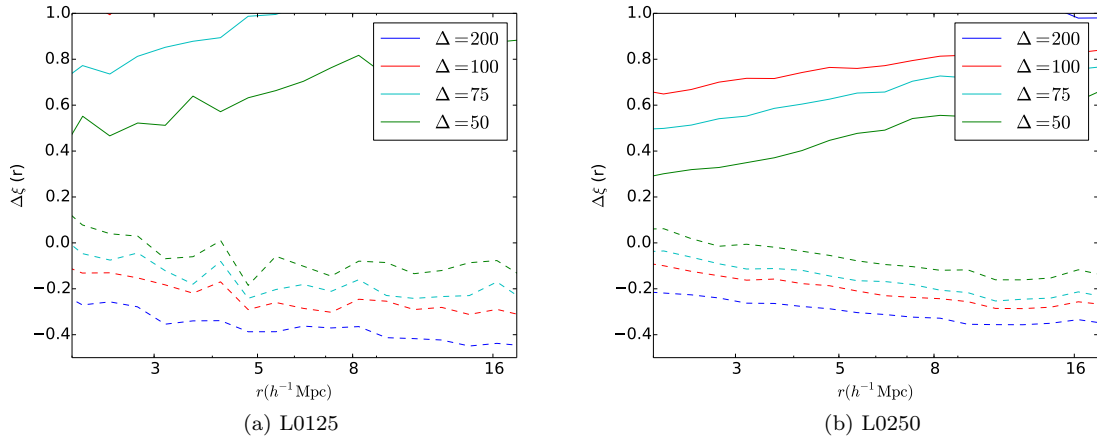
Figure 10. The difference of the correlation function of all hosts compared to a cut of the 20 % highest or lowest NFW defined concentration halos utilizing the 'l0125cut' mass cutoffs. The solid lines represent the difference from the highest concentration cut, the dashed lines represent the difference from the lowest concentration cut.
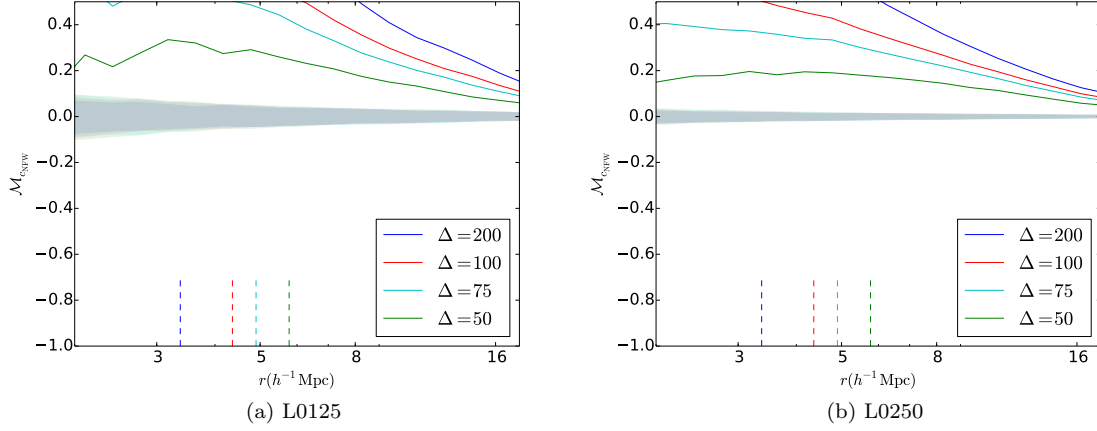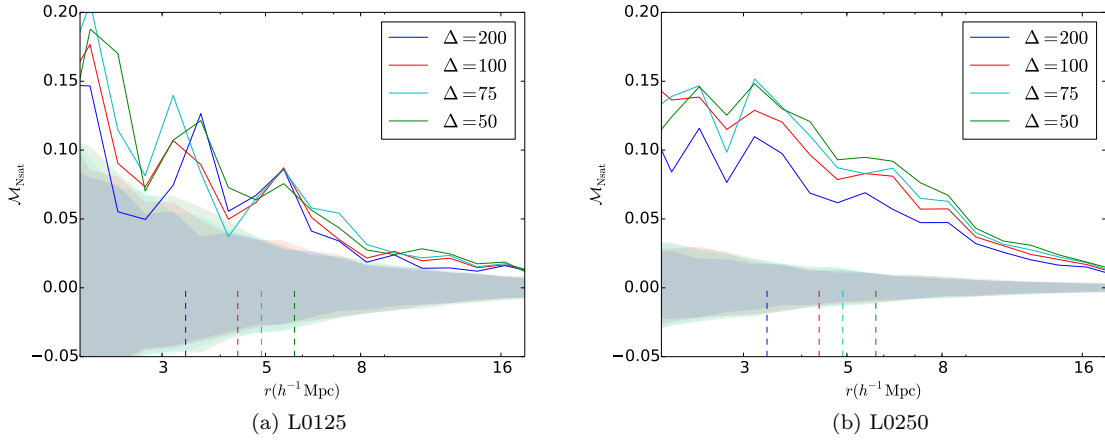
Figure 11. The marked correlation function for the concentration defined according to the NFW profile utilizing the 'l0125cut' mass cutoffs. The shaded bands represent 2-sigma confidence regions generated by randomization of the marks. The dashed line denotes the largest halo radius for a given value of the overdensity parameter.
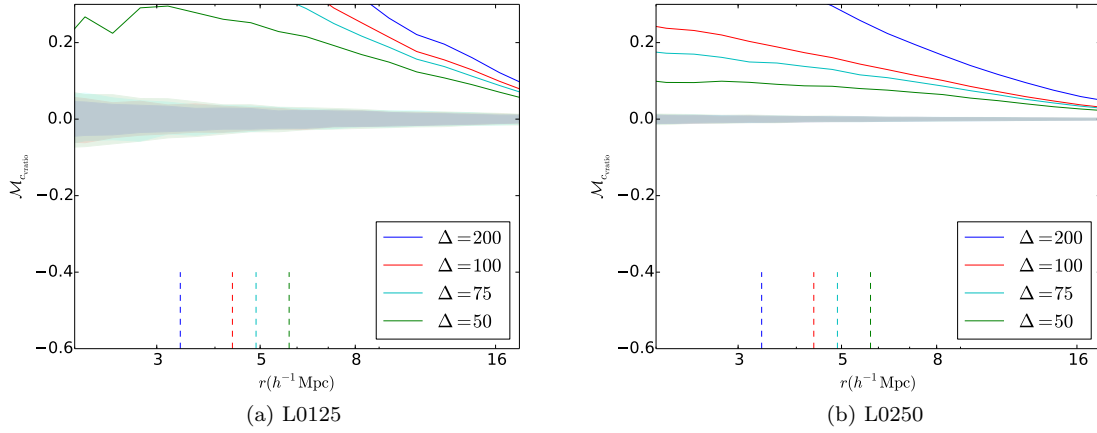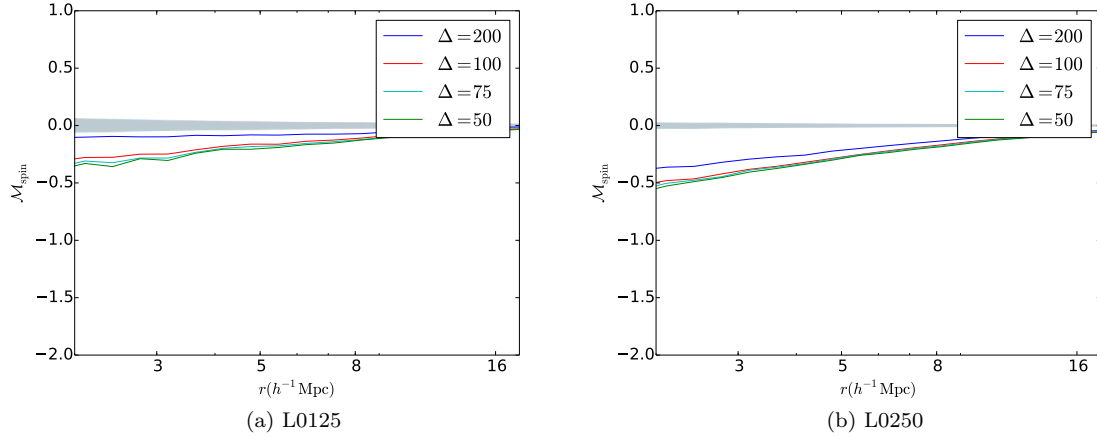

Figure 12. The marked correlation function for the concentration defined according to the velocity ratio utilizing the 'l0125cut' mass cutoffs. The shaded bands represent 2-sigma confidence regions generated by randomization of the marks. The dashed line denotes the largest halo radius for a given value of the overdensity parameter.

(a) L0125

(b) L0250

Figure 13. The marked correlation function for the shape of the halo utilizing the 'l0125cut' mass cutoffs. The shaded bands represent 2-sigma confidence regions generated by randomization of the marks. The dashed line denotes the largest halo radius for a given value of the overdensity parameter.

(a) L0125

(b) L0250

Figure 14. The marked correlation function for the satellite number utilizing the 'l0125cut' mass cutoffs. The shaded bands represent 2-sigma confidence regions generated by randomization of the marks. The dashed line denotes the largest halo radius for a given value of the overdensity parameter.

Figure 15. The marked correlation function for the spin of the halo utilizing the 'l0125cut' mass cutoffs. The shaded bands represent 2-sigma confidence regions generated by randomization of the marks. The dashed line denotes the largest halo radius for a given value of the overdensity parameter.