The Astrophysical Journal, Submitted
Preprint typeset using LaTeX style emulateapj v. 04/21/05

# THE HALO MODEL

Andrew Zentner[1]

*The Astrophysical Journal, Submitted*

ABSTRACT

*Subject headings:* cosmology: theory, galaxies: formation, large-scale structure of universe

## 1. THE HALO MODEL

The halo model is a convenient formalism for predicting and interpreting the clustering statistics of dark matter, galaxies, and anything else associated with dark matter halos. The basic idea behind the halo model framework has a long history, initially in analytic models that described galaxy clustering as a superposition of randomly-distributed clusters with specified profiles and a range of cluster masses (Neyman & Scott 1952; McClelland & Silk 1977; Peebles 1974). The explosion of recent activity in this field has been partly fueled by the recognition that a combination of this approach with recently developed tools for predicting the spatial clustering of dark matter halos (e.g., Mo & White 1996; Sheth & Tormen 1999; Sheth et al. 2001c; Seljak & Warren 2004; Tinker et al. 2005b) provides a powerful formalism for analytic calculations of dark matter clustering, which can be extended naturally to biased galaxy populations. Various aspects of the halo model are discussed in a number of contributions to the literature (e.g., Scherrer & Bertschinger 1991; Ma & Fry 2000; Seljak 2000; Scoccimarro et al. 2001; Sheth et al. 2001a,b; Berlind & Weinberg 2002; Cooray & Sheth 2002). In this Section, we give a brief introduction to the halo model.

### 1.1. *The Dark Matter Correlation Function and Power Spectrum*

The halo model is predicated upon several fundamental assumptions. The first assumption is that all dark matter is bound in halos of some size. The gross properties and large-scale distribution of these halos are well-understood thanks to high-resolution, cosmological, $N$-body simulations. The distribution of dark matter can then be described by the mass function of dark matter halos $\mathrm{d}n(m)/\mathrm{d}m$ (e.g., Sheth & Tormen 1999; Jenkins et al. 2001), the bias of halos with respect to dark matter $b_{\rm h}(m)$ (e.g., Efstathiou et al. 1988; Mo & White 1996; Jing 1998; Sheth & Tormen 1999; Tinker et al. 2005a), and the spatial distribution of dark matter within these halos as a function of halo mass (Navarro et al. 1997; Bullock et al. 2001). These ingredients can be calculated from direct simulation or approximated analytically using the techniques described in the references above. The standard implementation of the halo model assumes that halo clustering is independent of all halo properties aside from halo mass though the halo model can be extended to include halo clustering as a function of other halo attributes.

The simplest application of the halo model is to calculate the dark matter two-point correlation function or, equivalently, the power spectrum. The calculation can be broken into two parts. First, one can compute the "one-halo" term due to distinct mass elements that lie within the same dark matter halo. The second component is the "two-halo" term which is due to mass elements in distinct pairs of halos. Thus the dark matter correlation function can be written as the sum

$$\xi_{\rm DM}(r) = \xi_{\rm 1h}(r) + \xi_{\rm 2h}(r). \quad (1)$$

The one-halo term dominates the correlation function on scales smaller than the virial radii of halos and the two-halo term dominates the correlation function on scales much larger than the virial radii of the largest halos.

The two terms can be computed by enumerating pairs of infinitesimal mass elements (e.g., Scherrer & Bertschinger 1991). The one-halo term is then

$$\xi_{\rm 1h}(r) = \frac{1}{\rho_{\rm M}^2}\int \mathrm{d}m\, m^2 \frac{\mathrm{d}n(m)}{\mathrm{d}m}\int \mathrm{d}^3x\, \lambda_m(\vec{x})\lambda_m(\vec{x}+\vec{r}), \quad (2)$$

where $\mathrm{d}n(m)/\mathrm{d}m$ is the halo mass function at mass $m$ as described above, $\rho_{\rm M}$ is the mean mass density of the universe, and $\lambda_m(\vec{x})$ is the mass density distribution within a halo of mass $m$ normalized so that the integral of the profile over the volume of the halo is unity. The distribution of mass in halos is most often taken to be of the form proposed by Navarro et al. (1997, NFW hereafter),

$$\lambda_m(r) \propto \left(c_{\rm vir}(M)\,\frac{r}{R_{\rm vir}}\right)^{-1}\left(1+c_{\rm vir}(M)\,\frac{r}{R_{\rm vir}}\right)^{-2}, \quad (3)$$

where $R_{\rm vir}$ is the virial radius of the halo and $c_{\rm vir}(M)$ is the mass-dependent concentration parameter that describes the radius of the transition between the two power laws. Simulation results for $c_{\rm vir}(M)$ are given by Bullock et al. (2001).

The two-halo term can be computed by counting mass elements in distinct halo pairs. The two-halo term is given by

$$\begin{aligned}\xi_{\rm 2h}(r) = &\frac{1}{\rho_{\rm M}^2}\int \mathrm{d}m_1 \int \mathrm{d}m_2\, m_1\frac{\mathrm{d}n(m_1)}{\mathrm{d}m_1}\, m_2\frac{\mathrm{d}n(m_2)}{\mathrm{d}m_2}\\ &\times\int \mathrm{d}^3x \int \mathrm{d}^3y\ \lambda_{m_1}(\vec{x})\lambda_{m_2}(\vec{y})\\ &\times \xi_{\rm hh}(\vec{x}-\vec{y}+\vec{r}|m_1,m_2), \end{aligned} \quad (4)$$

where $\xi_{\rm hh}(\vec{x}|m_1,m_2)$ is the cross-correlation function of halos of mass $m_1$ and $m_2$ and $r \equiv |\vec{r}|$. The most commonly utilized assumption is that $\xi_{\rm hh}(r|m_1,m_2) \simeq b_{\rm h}(m_1)b_{\rm h}(m_2)\xi_{\rm DM}^{\rm lin}(r)$, where $\xi_{\rm DM}^{\rm lin}$ is the dark matter power spectrum in linear perturbation theory, as in models of deterministic halo bias. The halo bias relation can

[1] KICP

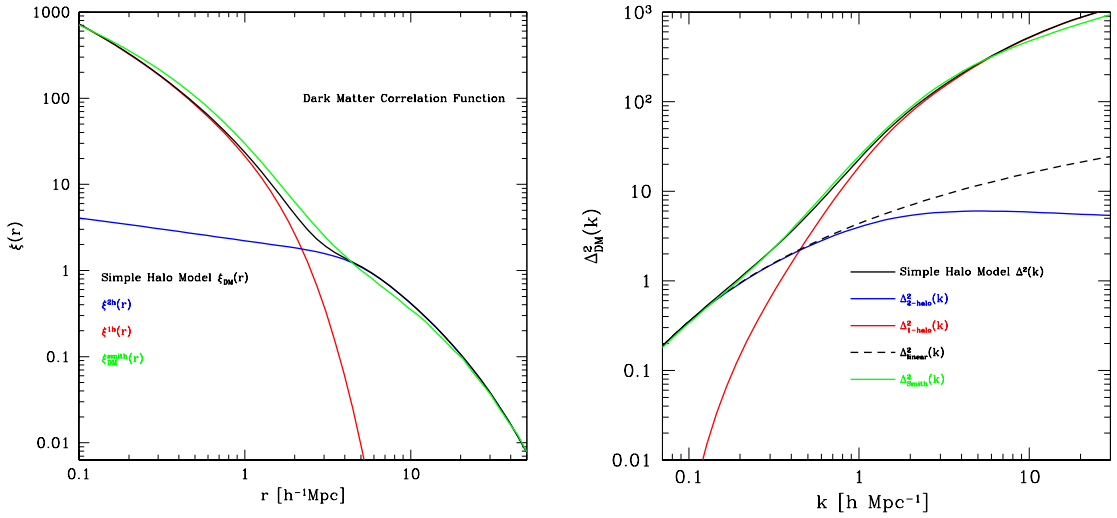


FIG. 1.— Dark matter correlation function and power spectrum computed according to a simple halo model. The left panel shows the dark matter correlation function while the right panel shows the dark matter power spectrum. The *solid, black* lines show the results from the full halo model, *red* lines show the one-halo contribution, *blue* lines show the two-halo contribution, *green* line show the halo model-based fitting formula provided by Smith et al. (2003), and the *dashed* line shows the dark matter power spectrum in linear theory.

subsequently be pulled out of the spatial integrals and the two-halo term can be written as

$$\begin{aligned}\xi_{\rm 2h}(r) = & \frac{1}{\rho_{\rm M}^2}\int {\rm d}m_1 \int {\rm d}m_2\, m_1 b_{\rm h}(m_1)\frac{{\rm d}n(m_1)}{{\rm d}m_1} \\ & \times \quad m_2 b_{\rm h}(m_2)\frac{{\rm d}n(m_2)}{{\rm d}m_2} \\ & \times \int {\rm d}^3x \int {\rm d}^3y\; \lambda_{m_1}(\vec{x})\lambda_{m_2}(\vec{y}) \\ & \times \xi_{\rm hh}(\vec{x}-\vec{y}+\vec{r}|m_1,m_2), \end{aligned} \quad (5)$$

In the limit of separations much larger than the sizes of the largest halos, the correlation function is determined by the two-halo term alone. On such large scales, the correlation functions vary little over the length scales of halos so that $\xi^{\rm lin}_{\rm DM}(\vec{x}-\vec{y}+\vec{r}) \simeq \xi^{\rm lin}_{\rm DM}(\vec{r})$ so that the spatial integrals in Eq. (5) can be replaced by $\xi^{\rm lin}_{\rm DM}(\vec{r})$. In the limit of very large separations, the two-halo term must give the linear dark matter correlation function so that the bias must obey the integral constraint

$$\int {\rm d}m\, \frac{{\rm d}n(m)}{{\rm d}m}\left(\frac{m}{\rho_{\rm M}}\right) b_{\rm h}(m) = 1. \quad (6)$$

This is the well-known normalization rule for the mass-dependent halo bias.

It is generally easier to work in Fourier space in terms of power spectra rather than correlation functions due to the convolution integrals in Eq. (2) and Eq. (4). The most common Fourier convention gives the Fourier transform and inverse Fourier transform as

$$f(\vec{k}) = \int {\rm d}^3x\, f(\vec{x})\exp(i\vec{k}\cdot\vec{x}) \quad (7)$$

and

$$f(\vec{x}) = \frac{1}{(2\pi)^3}\int {\rm d}^3k\, f(\vec{k})\exp(-i\vec{k}\cdot\vec{x}). \quad (8)$$

The convolution theorem then states that the Fourier transform of the convolution

$$H(\vec{y}) = \int {\rm d}^3\vec{x}\, f(\vec{y}-\vec{x})g(\vec{x}) \quad (9)$$

is

$$H(\vec{k}) = g(\vec{k})\, f(\vec{k}), \quad (10)$$

and similarly the Fourier transform of the double convolution

$$H(\vec{z}) = \int {\rm d}^3\vec{x}\, f(\vec{x}) \int {\rm d}^3\vec{y}\, g(\vec{y})\, j(\vec{x}-\vec{y}-\vec{z}) \quad (11)$$

is

$$H(\vec{k}) = g(\vec{k})\, f(\vec{k})\, j(\vec{k}). \quad (12)$$

The power spectrum is then a sum of one-halo and two-halo contributions $P(k) = P^{\rm 1h}(k) + P^{\rm 2h}(k)$. The one-halo contribution to the matter power spectrum is

$$P^{\rm 1h}(k) = \frac{1}{\rho_{\rm M}^2}\int {\rm d}m\, m^2 \frac{{\rm d}n}{{\rm d}m}\lambda_m^2(k). \quad (13)$$

Similarly, the two-halo power is

$$\begin{aligned} P^{\rm 2h}(k) = & \frac{P^{\rm lin}_{\rm DM}(k)}{\rho_{\rm M}^2}\left(\int {\rm d}m\, m \frac{{\rm d}n}{{\rm d}m}\lambda_m(k) b_{\rm h}(m)\right)^2 \quad (14) \\ \equiv & \frac{P^{\rm lin}_{\rm DM}(k)}{\rho_{\rm M}^2} I_m^2, \quad (15) \end{aligned}$$

where $P^{\rm lin}_{\rm DM}(k)$ is the linear power spectrum of the dark matter and $\lambda_m(k)$ is the Fourier transform of the mean mass density profile of halos of mass $m$. In addition to the power spectrum, which has units of volume, it is often convenient to work in terms of the dimensionless mass density variance per logarithmic interval in wavenumber

$$\Delta^2(k) \equiv k^3\frac{P(k)}{2\pi^2}. \quad (16)$$

The integral over mass in Eq. (14) converges very slowly, requiring integration over a very large mass range. It is more practical to recast this integral in terms of the variable $\nu(m) \equiv \delta_{\rm c}/\sigma(m)$ where $\delta_{\rm c}$ is the equivalent linear collapse overdensity (e.g. Lacey & Cole 1993) and $\sigma(m)$ is the *rms* mass fluctuation in the linear density

field smoothed on a scale containing mass $m$. The most common convention is to use a real-space tophat window to smooth the density field and compute $\sigma(m)$. Let $f(\nu)\mathrm{d}\nu$ denote the fraction of mass contained in halos in a range $\mathrm{d}\nu$ about $\nu$. Then the mass function is related to $f(\nu)$ by

$$\frac{\mathrm{d}n}{\mathrm{d}m} = 2\nu^2 f(\nu)\left(\frac{\rho_\mathrm{M}}{m}\right)\frac{\mathrm{d}\ln\nu}{\mathrm{d}m} \tag{17}$$

and the integral in Eq. (14) can be written as

$$I_m = 2\rho_\mathrm{M}\int \mathrm{d}\ln\nu\,\nu^2 f(\nu)\lambda_\nu(k)b_\mathrm{h}(\nu). \tag{18}$$

As a simple example, consider the matter power spectrum and correlation function. In Figure 1 I show the dark matter correlation function and power spectrum in the standard ΛCDM cosmology with $\Omega_\mathrm{M} = 0.3$, $\Omega_\Lambda = 0.7$, $\Omega_\mathrm{B}h^2 = 0.0221$, $h = 0.7$, and $\sigma_8 = 0.9$. I use the simple implementation of the halo model described above with dark matter halo profiles given by the NFW form with concentrations assigned according to the model of Bullock et al. (2001). In addition, I use a halo mass function and a mass-dependent halo bias given by Sheth & Tormen (1999), which explicitly satisfies the integral constraint of Eq. (6). The green lines in Fig. 1 show the halo model-based fitting formula for the non-linear clustering of dark matter given by Smith et al. (2003).

### 1.2. *Adapting the Halo Model to the Clustering of Other Populations*

In the context of the halo model, halos serve as surrogates that contain the population of interest and facilitate the counting of members of this population. In the previous example, that population was the dark matter. In this section we demonstrate the simple extension of the halo model to other populations by computing the power spectrum of baryonic gas.

The first additional ingredient we need to know is the average amount of gas in a halo of mass $m$. We can parameterize this in terms of the gas mass fraction, $f_\mathrm{gas}(m)$ so that the average gas mass in a halo of mass $m$ is $mf_\mathrm{gas}(m)$. The second additional ingredient is the spatial distribution of gas within halos, $\lambda_m^\mathrm{gas}(r)$. In this example, I take the adiabatic hydrodynamic simulations of Rudd (in preparation) to motivate particular choices. These simulations show that the gas distribution within halos can be well described by the profile of Burkert (1995),

$$\rho_\mathrm{B}(r) = \frac{\rho_\mathrm{B}}{(1+c_\mathrm{B}r/R_\mathrm{vir})(1+[c_\mathrm{B}r/R_\mathrm{vir}]^2)}. \tag{19}$$

The simulations provide the detailed information on the relation between the Burkert concentration $c_\mathrm{B}$ and halo mass[2]. The Burkert profile goes to a constant-density core at $r \ll R_\mathrm{vir}/c_\mathrm{B}$ and this results in a suppression in small-scale power at high wavenumbers relative to an NFW profile. This relative suppression can be seen in Figure 2.

[2] The gas fractions and concentrations used here can be viewed at http://logrus.uchicago.edu/group/index.php/BaryonPowerSpectrum

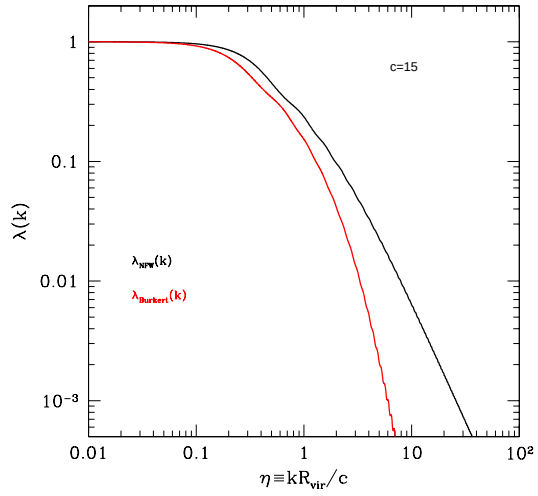


FIG. 2.— The Fourier transforms of Burkert and NFW profiles at a representative concentration of $c = c_\mathrm{vir} = c_\mathrm{B} = 15$.

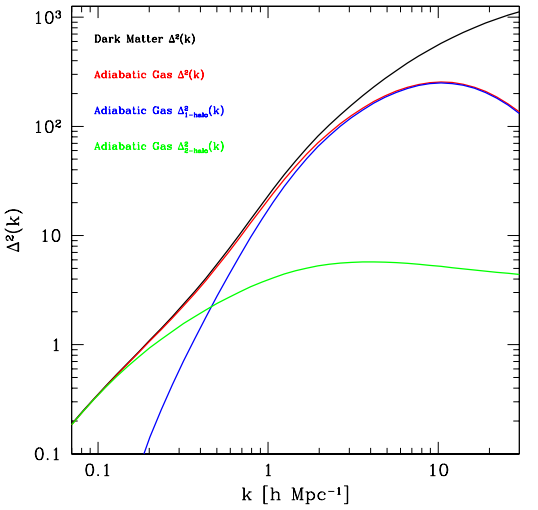


FIG. 3.— The baryonic gas power spectrum computed from a halo model that has been designed to mock up an adiabatic hydrodynamic simulation.

With these extra ingredients, the one- and two-halo contributions to the power spectrum can be written down immediately. The one-halo power is

$$P_\mathrm{gas}^\mathrm{1h}(k) = \frac{1}{\rho_\mathrm{gas}^2}\int \mathrm{d}m\, m^2 f_\mathrm{gas}^2(m)\frac{\mathrm{d}n}{\mathrm{d}m}[\lambda_m^\mathrm{gas}(k)]^2. \tag{20}$$

Similarly, the two-halo power for gas is

$$P_\mathrm{gas}^\mathrm{2h}(k) = \frac{P_\mathrm{DM}^\mathrm{lin}(k)}{\rho_\mathrm{gas}^2}\left(\int \mathrm{d}m\, m f_\mathrm{gas}(m)\frac{\mathrm{d}n}{\mathrm{d}m}[\lambda_m^\mathrm{gas}(k)]^2 b_\mathrm{h}(m)\right)^2. \tag{21}$$

The net results is the adiabatic gas power spectrum shown in Figure 3.

### 1.3. *Multi-component Models and Cross Correlations*

Of course, the halo model can be extended to model an arbitrary number of components using the approach of the previous two sections. In multi-component models

the cross correlations and cross power spectra between the different components can likewise be computed according to the same approach, one simply counts pairs of objects that are members of the two populations that are under consideration.

For concreteness, consider the dark matter-gas cross power spectrum using the quantities defined in the previous sections. The one-halo contribution to the cross power spectrum comes from gas elements and dark matter mass elements that reside within the same halo and is given by

$$P_{\rm gm}^{\rm 1h} = \frac{1}{\rho_{\rm DM}\rho_{\rm gas}} \int {\rm d}m \, \frac{{\rm d}n}{{\rm d}m} \, [m\, \lambda_m(k)] \, [m\, f_{\rm gas}(m)\, \lambda_m^{\rm gas}(k)]. \tag{22}$$

Similarly, the two-halo term is

$$P_{\rm gm}^{\rm 2h} = \frac{P_{\rm DM}^{\rm lin}(k)}{\rho_{\rm gas}\rho_{\rm DM}} \times \left( \int {\rm d}m\, m f_{\rm gas}(m) \frac{{\rm d}n}{{\rm d}m} \lambda_m^{\rm gas}(k) b_{\rm h}(m) \right) \tag{23}$$

$$\times \left( \int {\rm d}m\, m \frac{{\rm d}n}{{\rm d}m} \lambda_m(k) b_{\rm h}(m) \right). \tag{24}$$

The cross power spectrum is then $P_{\rm gm} = P_{\rm gm}^{\rm 1h} + P_{\rm gm}^{\rm 2h}$.

The remaining two-point statistic of interest is the total power spectrum for the population consisting of the sum of the various components. This can be computed from the aforementioned quantities according to

$$P_{\rm total}(k) = \bar{f}_{\rm DM}^2 P_{\rm DM}(k) + \bar{f}_{\rm gas}^2 P_{\rm gas}(k) + 2\bar{f}_{\rm gas}\bar{f}_{\rm DM} P_{\rm gm}(k), \tag{25}$$

where $\bar{f}_{\rm gas} \equiv \Omega_{\rm B}/\Omega_{\rm M}$ and $\bar{f}_{\rm DM} \equiv \Omega_{\rm DM}/\Omega_{\rm M}$, for the specific case of a two-component dark matter-gas model, and

$$P_{\rm total} = \sum_i \bar{f}_i^2 P_i(k) + 2 \sum_i \sum_{j\neq i} \bar{f}_i \bar{f}_j P_{\rm ij}(k) \tag{26}$$

for a more general many-component model where the $\bar{f}_i$ are the universal mass fractions in each component. For the dark matter-adiabatic gas two-componenet model, the total and cross power spectra are shown in Figure 4.

### 1.4. *Effects of Baryons on the Matter Power Spectrum: Some Particulars*

In order to mock up the effect of baryonic physics on the matter power spectrum, I have used the techniques and formulas set out in the previous sections. There are a number of specific details that need to be addressed in order to reproduce these numbers.

#### 1.4.1. *Mass Function and Halo Bias*

I used both the mass function and the halo bias of Sheth & Tormen (1999). This combination explicitly satisfies the normalization integral for bias. The drawback is that the halo bias is a bit high on intermediate mass scales leading to an overprediction of the power spectra near $k \sim 0.5 h Mpc^{-1}$.

For the cooling run I boost the by the factors shown on the wiki page [3]. I extrapolate away from the tables by assuming a constant boost at the level of the nearest bin.

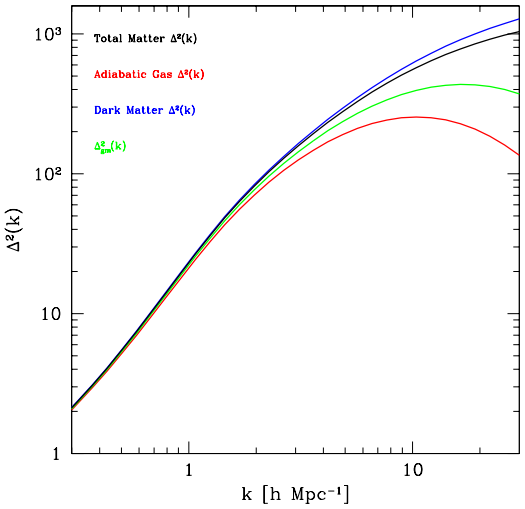


FIG. 4.— The dark matter-gas cross power spectra and the total matter power spectrum in a two-component halo model of dark matter and adiabatic gas.

#### 1.4.2. *Halo Concentrations*

I modeled the N-body data using a particular realization of the Bullock et al. (2001) concentration model in order to interpolate between and extrapolate from the N-body data for the L60B simulation box. L60B halos have a lower normalization and a shallower slope than the standard Bullock et al. (2001) model and so I have mocked up the L60B halo concentrations using a modified version of the model. I have chosen the parameters $F = 10^{-5}$ and $K = 1.75$.

In order to model the halos in the Adiabatic and CSF runs, I kept the assumption that the mass distributions are described by NFW profiles. However, I modified the concentrations of the profiles. I did this by taking the ratios $c_{\rm vir}^{\rm adiabatic}/c_{\rm vir}^{\rm N-body}$ and $c_{\rm vir}^{\rm CSF}/c_{\rm vir}^{\rm N-body}$ and using these to multiply the concentrations in the basic concentration model that I described in the previous paragraph. These ratios are shown on our wiki page.

I extrapolated outside of the bounds by taking the enhancements to be constant and set by the value of the enhancement at the closest bin. This is certainly crude and can easily be changed, but I do not know of a robust basis for changing this prescription.

#### 1.4.3. *Gas Profiles*

I use Burkert profiles to model the gas in the adiabatic run. I use the gas concentrations and gas fractions shown on the wiki page. I eliminate the two lowest mass bins for these variables and extrapolate to low mass as a constant gas fraction. I extrapolate to higher mass assuming the universal gas mass fraction. Notice that not all of the gas is contained in halos according to this prescription. Notice also that in Eq. (20) and Eq. (21) I have divided the power spectra by the universal gas fraction. This is tantamount to an assumption that the remaining gas is unclustered. This is the reason for the small anti-bias of gas relative to dark matter in the adiabatic plots.

## 2. THE HALO MODEL FOR GALAXY CLUSTERING

To compute galaxy clustering properties, it is necessary to take as an additional assumption that all galaxies re-

[3] http://logrus.uchicago.edu/group/index.php/BaryonPowerSpectrum

side within dark matter halos. The third assumption of the halo model is that the galaxy content of halos is statistically independent of halo environment and depends only upon halo mass. There is some evidence that this last assumption does not hold to high accuracy (e.g., Gao et al. 2005; Wechsler et al. 2005); however, the net effect of these environmental dependences is small and we neglect them for the purposes of this study.

In order to compute galaxy clustering statistics, it is necessary to specify the number and spatial distribution of galaxies within halos. The number of galaxies is specified by the probability for a halo of fixed virial mass $M$, to play host to $N_{\rm gal}$ galaxies, $P(N_{\rm gal}|M)$. We follow convention and refer to the probability distribution $P(N_{\rm gal}|M)$ as the *halo occupation distribution* (HOD). Let the average relative spatial distribution of galaxies within halos of mass $M$ be denoted by $\lambda(\vec{x}|M)$.

Given a halo population and an HOD, the galaxy autocorrelation function is simple to compute. The correlation function is conventionally written as the sum of two terms:

$$\xi_{\rm gg}(r) = \xi^{\rm 1h}_{\rm gg}(r) + \xi^{\rm 2h}_{\rm gg}(r). \tag{27}$$

The *one halo term* $\xi^{\rm 1h}_{\rm gg}(r)$, is due to pairs of galaxies at separation $r$ that lie within a common host halo and the *two halo term* $\xi^{\rm 2h}_{\rm gg}(r)$, is due to pairs of galaxies that reside in distinct halos. The two terms can be computed straightforwardly by counting pairs at fixed separation and are given by

$$\xi^{\rm 1h}_{\rm gg}(r) = \bar{n}_{\rm g}^{-2} \int {\rm d}M \, \frac{{\rm d}n}{{\rm d}M} \langle N_{\rm gal}(N_{\rm gal}-1)\rangle_M \times \int {\rm d}^3x \, \lambda(\vec{x}|M) \, \lambda(\vec{x}+\vec{r}|M) \tag{28}$$

and

$$\xi^{\rm 2h}_{\rm gg}(r) = \bar{n}_{\rm g}^{-2} \int {\rm d}M \int {\rm d}M' \, \frac{{\rm d}n}{{\rm d}M} \langle N_{\rm gal}(M)\rangle \, \frac{{\rm d}n}{{\rm d}M'} \langle N_{\rm gal}(M')\rangle \times \int {\rm d}^3x \int {\rm d}^3y \, \lambda(\vec{x}|M)\lambda(\vec{y}|M') \times \xi_{\rm hh}(\vec{x}-\vec{y}-\vec{r}|M,M') \, , \tag{29}$$

where $r \equiv |\vec{r}|$, $\bar{n}_{\rm g}$ is the mean number density of galaxies being considered, and $\xi_{\rm hh}(\vec{r}|M,M')$ is the cross correlation of halos of mass $M$ and $M'$. The one-halo term depends upon the HOD through the mean number of galaxy pairs per halo as a function of mass

$$\langle N_{\rm gal}(N_{\rm gal}-1)\rangle_M \equiv \sum_{N_{\rm gal}=1}^{\infty} N_{\rm gal}(N_{\rm gal}-1) P(N_{\rm gal}|M), \tag{30}$$

while the two-halo term is sensitive to the mean number of galaxies per halo

$$\langle N_{\rm gal}(M)\rangle \equiv \sum_{N_{\rm gal}=1}^{\infty} N_{\rm gal} P(N_{\rm gal}|M). \tag{31}$$

In principle, higher order correlators would require higher order moments of $P(N_{\rm gal}|M)$. For example, the three-halo term of the galaxy three-point function depends upon $\langle N_{\rm gal}(N_{\rm gal}-1)(N_{\rm gal}-2)\rangle$ and so on.

On scales much larger than the largest halos ($r \gg$ a few Mpc), the two-halo term dominates the correlation function as there are no halos large enough to host two galaxies at such a large separation. Supposing that the halo-halo correlation function $\xi_{\rm hh}(r|M,M')$ varies slowly over the scale of a single halo, the last two integrals in Eq. (29) can be approximated by $\xi_{\rm hh}(r|M,M')$. Making the standard assumption that we can express the halo-halo correlation function in terms of the halo bias $b_{\rm h}(M)$, such that $\xi_{\rm hh}(r|M,M') = b_{\rm h}(M) b_{\rm h}(M') \xi_{\rm mm}(r)$, where $\xi_{\rm mm}(r)$ is the mass autocorrelation function further simplifies Eq. (29) resulting in a simple relation for the large-scale bias of any particular galaxy population, namely

$$b_{\rm gal} \simeq \frac{1}{\bar{n}_{\rm g}} \int {\rm d}M \, \frac{{\rm d}n}{{\rm d}M} \langle N_{\rm gal}(M)\rangle b_{\rm h}(M). \tag{32}$$

In what follows, we follow the implementation of the halo model for galaxy clustering used by Tinker et al. (2005a) which is based upon extensive tests against mock galaxy populations distributed throughout large cosmological $N$-body simulations. In particular, we use the host halo mass function of Jenkins et al. (2001), the dark matter correlation function of Smith et al. (2003), and the scale-dependent halo bias relation of Tinker et al. (2005a).

### 2.1. *The Halo Occupation Distribution*

Given a cosmological model, the halo mass function, halo concentrations, and the halo bias are all specified so that the galaxy clustering statistics depend only upon the HOD. A convenient and conceptually fruitful approach to the HOD is to model the full distribution $P(N|M)$ as a sum of contributions from two distinct populations: central galaxies associated with the host dark matter halo and located at the host halo center, and satellite galaxies distributed throughout the halo (Guzik & Seljak 2002; Kravtsov et al. 2004). The motivation for this decomposition comes from the properties of galaxies in hydrodynamic simulations (Berlind et al. 2003), the distinction between halo and subhalo populations in dissipationless simulations (Kravtsov et al. 2004), and studies of central elliptical galaxies in groups and clusters which are often considered a separate population from the remainder of the cluster galaxies.

The central galaxy contribution can take on only two values, $N_{\rm cen} = 0$ or $N_{\rm cen} = 1$, depending upon whether or not there is a central galaxy associated with the host halo. The simplest approach is to model the central galaxy population as a step function where halos more massive than some threshold $M_{\rm min}$ contain a central galaxy, but as we demonstrate below, it is also possible to model a continuous transition between $\langle N_{\rm cen}\rangle = 0$ at low mass and $\langle N_{\rm cen}\rangle = 1$ at high mass by accounting for the expected scatter in halo and galaxy properties at fixed mass. Thus, the distribution $P(N_{\rm cen}|M)$ is that of a *nearest integer* distribution with the frequency of the values $N_{\rm cen} = 0$ and $N_{\rm cen} = 1$ fixed such that the mean is maintained at $\langle N_{\rm cen}\rangle$. For such a distribution, the second moment of the distribution is clearly $\langle N_{\rm cen}(N_{\rm cen}-1)\rangle = \sum_{N_{\rm cen}} N_{\rm cen}(N_{\rm cen}-1) P(N_{\rm cen}|M) = 0$.

The satellite galaxy population can then be modeled by a power law, $\langle N_{\rm sat}\rangle = (M/M_1)^{\alpha}$. In the regime where the host halo always contains a central galaxy, the HOD of satellite galaxies is simply related to the total HOD be-

cause $N_{\rm gal} = 1 + \langle N_{\rm sat}\rangle$. The most assumption that is often made is that the number of satellite galaxies at fixed mass $P(N_{\rm sat}|M)$, follows a Poisson distribution. This is the case for the satellite halos embedded in host halos in cosmological $N$-body simulations (e.g., Kravtsov et al. 2004; Zentner et al. 2005) and these satellite halos constitute natural sites for satellite galaxies. In the regime where the host halos seldom contain a satellite galaxy $\langle N_{\rm sat}\rangle \ll 1$, the second moment of the total HOD is $\langle N_{\rm gal}(N_{\rm gal}-1)\rangle = \langle N_{\rm sat}^2\rangle + \langle N_{\rm sat}\rangle \approx 2\,\langle N_{\rm sat}\rangle$ in the case of a Poisson distribution. As a result, the total HOD will have a distribution that is narrower than Poisson even if the satellite HOD is governed by a Poisson distribution or even a distribution that is significantly broader than Poisson. In the following, we discuss the evolution of the HOD as structure forms and its influence on the galaxy two-point function.

### 2.2. *Illustrative Examples*

Before proceeding, we discuss the dependence of the correlation function on the HOD using several toy models for the HOD. Our goal is to study well motivated models rather than to be comprehensive, therefore we limit the scope of this section to the consideration of models that can be decomposed into a step function for the central galaxy,

$$N_{\rm cen} = \begin{cases} 0 & \text{if } M < M_{\rm min} \\ 1 & \text{if } M \geq M_{\rm min} \end{cases} , \tag{33}$$

and a simple power law for the satellite galaxies,

$$\langle N_{\rm sat}(M)\rangle = \begin{cases} 0 & \text{if } M < M_{\rm min} \\ (M/M_1)^\alpha & \text{if } M \geq M_{\rm min} \end{cases} , \tag{34}$$

as described in the previous section, and models that are pure power laws above a minimum halo mass threshold,

$$\langle N_{\rm gal}(M)\rangle = \begin{cases} 0 & \text{if } M < M_{\rm min} \\ (M/M_{\rm min})^\alpha & \text{if } M \geq M_{\rm min} \end{cases} , \tag{35}$$

In models that decompose the HOD into central and satellite galaxy components the most common assumption that is made is that the central galaxies follow a nearest integer distribution and that the satellites follow a Poisson distribution, motivated largely by the results of $N$-body and hydrodynamic cosmological simulations (e.g., Berlind et al. 2003; Kravtsov et al. 2004; Zheng et al. 2005), while in models with pure power-law HODs the most commonly used assumption is that of a nearest integer distribution for $N_{\rm gal}$.

In Figure 5, we summarize the net effect of changes in the mean of the HOD on the correlation function. In each case, we have assumed that the galaxy spatial distribution is given by the NFW profile of Eq. (3) with concentration parameters given by the model of Bullock et al. (2001) and that the galaxy population in each halo can be decomposed into a single central galaxy with additional satellites that follow a Poisson distribution for $P(N_{\rm sat}|M)$. The top panel of Fig. 5 shows the effect of a change in the ratio $M_1/M_{\rm min}$ at fixed mean total number density $\bar{n}_{\rm g} = 10^{-2}h^3{\rm Mpc}^{-3}$. Decreasing $M_1/M_{\rm min}$ at fixed number density places relatively more galaxies in large halos. This results in an increase in the clustering amplitude at large scales simply because the high-mass halos that are the most highly biased are weighted more heavily in Eq. (6). Decreasing $M_1/M_{\rm min}$ also increases the relative fraction of galaxies that are satellites rather than central galaxies, leading to a boost in the one-halo contribution to the correlation function at small scales. An important scale to remember is that, roughly speaking, nearly power-law correlation functions tend to have $M_1/M_{\rm min} \sim 20$ (in accordance with Zehavi et al. 2004), while HODs with smaller $M_1/M_{\rm min}$ have relatively larger one-halo terms and correspondingly higher small-scale correlations and HODs with larger $M_1/M_{\rm min}$ tend to have relatively smaller one-halo terms and correspondingly weaker small-scale correlation functions.

The second row of panels in Figure 5 shows the simple effect of varying the number density with the power-law index fixed to $\alpha = 1$ and $M_1/M_{\rm min} = 20$. This is tantamount to varying $M_{\rm min}$ with $M_1$ set to a fixed scale relative to $M_{\rm min}$. Increasing the number density leads to larger correlations at both large and small scales. On large scales, this iresults from the fact that host halos become more strongly clustered with increasing halo mass and the overall bias follows from Eq. (6).

The bottom row of panels in Figure 5 shows the effect of varying the power-law index $\alpha$ at fixed number density and fixed $M_1/M_{\rm min}$. Again, the net effect on the large-scale correlation function is clear: increasing $\alpha$ places more galaxies in massive halos and weights these highly-biased halos relatively more in Eq. (6). The effect on small scales is to increase the correlation function with increasing $\alpha$ as more satellite galaxies contribute to a larger one-halo term. The figure also shows that the scale at which the one-halo term becomes important is strongly affected by changing the power-law index. A large value of $\alpha$ means that large halos (mass greater than the typical collapsing mass $M_* \sim 10^{13}h^{-1}M_\odot$) can contribute to the one-halo term, partially overcoming the rapid decrease of halo abundance with $M$. As such, the one-halo term is important out to several Mpc in the case of $\alpha = 1.3$. Conversely, low values of $\alpha$ coupled with the rapid decline of ${\rm d}n(M)/{\rm d}M$ with increasing $M$ deemphasize large host halos so that the one-halo term only becomes important at $\sim 600h^{-1}$kpc in the $\alpha = 0.7$ case.

## REFERENCES


Berlind, A. A. & Weinberg, D. H. 2002, ApJ, 575, 587

Berlind, A. A., Weinberg, D. H., Benson, A. J., Baugh, C. M., Cole, S., Davé, R., Frenk, C. S., Jenkins, A., Katz, N., & Lacey, C. G. 2003, ApJ, 593, 1

Bullock, J. S., Kolatt, T. S., Sigad, Y., Somerville, R. S., Kravtsov, A. V., Klypin, A. A., Primack, J. R., & Dekel, A. 2001, MNRAS, 321, 559

Burkert, A. A. 1995, ApJL, 447, L25

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

Efstathiou, G., Frenk, C. S., White, S. D. M., & Davis, M. 1988, MNRAS, 235, 715

Gao, L., Springel, V., & White, S. D. M. 2005, MNRAS, 363, L66

Guzik, J. & Seljak, U. 2002, MNRAS, 335, 311

Jenkins, A., Frenk, C. S., White, S. D. M., Colberg, J. M., Cole, S., Evrard, A. E., Couchman, H. M. P., & Yoshida, N. 2001, MNRAS, 321, 372

Jing, Y. P. 1998, ApJ, 503, L9+

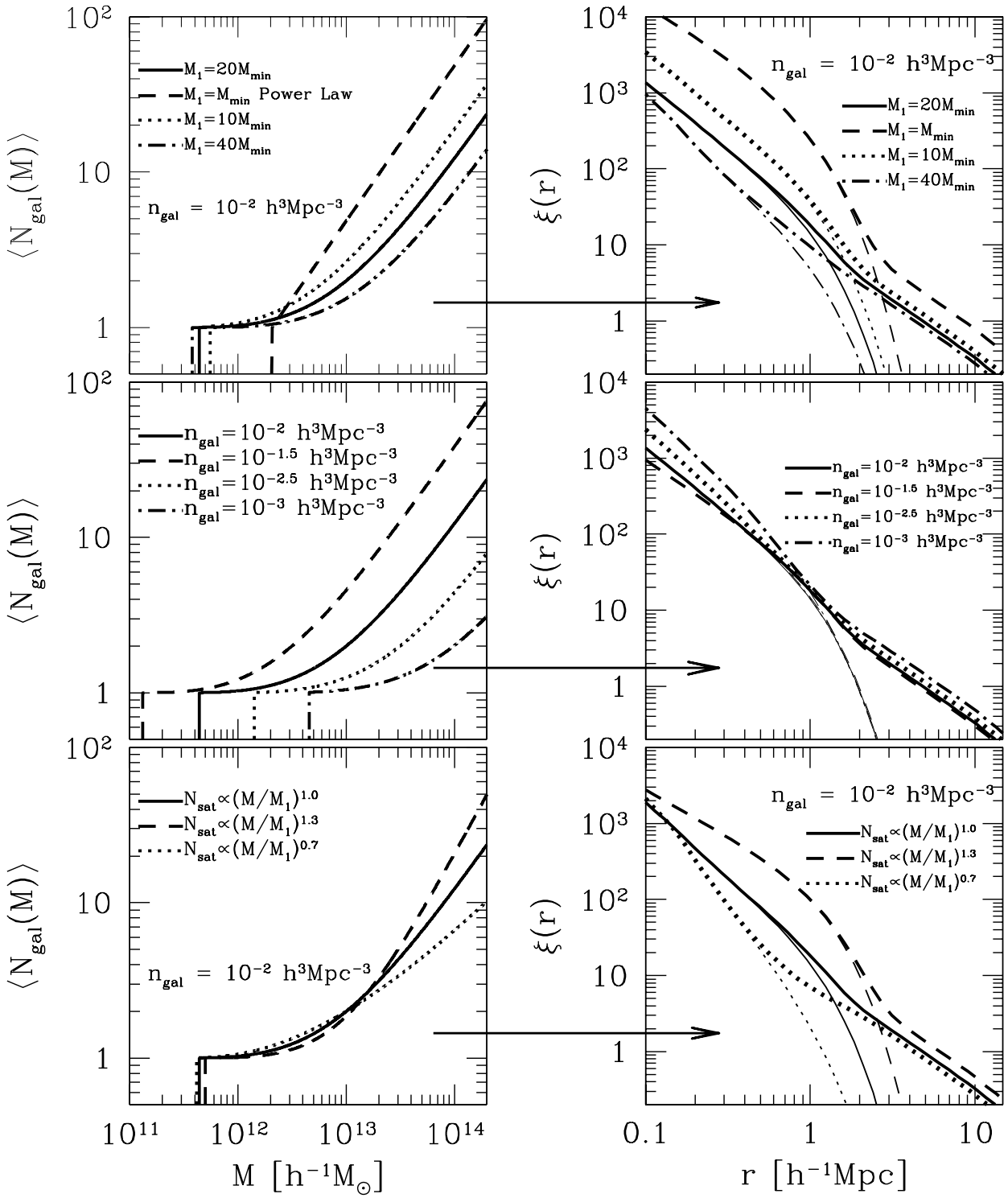


FIG. 5.— The dependence of the correlation function on several of the basic HOD parameters that will be relevant in the discussion that follows. The figure consists of three rows of two panels each. Each row illustrates the effect of a change in a particular HOD parameter on the resulting correlation function. The mean galaxy occupations are shown in the left panels. The right panels show the resulting correlation functions (*thick lines*) and the contribution of the one-halo terms to each of these correlation functions (*thin lines*). *Top:* The effect of the ratio of $M_1/M_{\rm min}$ on the correlation function at fixed mean number density $\bar{n}_{\rm g} = 10^{-2}h^3{\rm Mpc}^-3$ and fixed slope $\alpha = 1.0$. *Middle:* The effect of changing the number density with fixed *alpha* = 1.0 and $M_1/M_{\rm min} = 20$. Changing the number density with these other parameters fixed corresponds simply to shifting $M_{\rm min}$ as can be seen in the left panel of the middle row. *Bottom:* Changing the power-law index of the HOD with number density fixed to $\bar{n}_{\rm g} = 10^{-2}h^3{\rm Mpc}^{-3}$ and $M_1/M_{\rm min} = 20$.

Kravtsov, A. V., Berlind, A. A., Wechsler, R. H., Klypin, A. A., Gottlöber, S., Allgood, B., & Primack, J. R. 2004, ApJ, 609, 35
Lacey, C. & Cole, S. 1993, MNRAS, 262, 627
Ma, C.-P. & Fry, J. N. 2000, ApJ, 543, 503
McClelland, J. & Silk, J. 1977, ApJ, 217, 331
Mo, H. J. & White, S. D. M. 1996, MNRAS, 282, 347
Navarro, J. F., Frenk, C. S., & White, S. D. M. 1997, ApJ, 490, 493 (NFW)
Neyman, J. & Scott, E. L. 1952, ApJ, 116, 144
Peebles, P. J. E. 1974, ApJ, 189, L51
Scherrer, R. J. & Bertschinger, E. 1991, ApJ, 381, 349
Scoccimarro, R., Sheth, R. K., Hui, L., & Jain, B. 2001, ApJ, 546, 20
Seljak, U. 2000, MNRAS, 318, 203
Seljak, U. & Warren, M. S. 2004, MNRAS, 355, 129
Sheth, R. K., Diaferio, A., Hui, L., & Scoccimarro, R. 2001a, MNRAS, 326, 463
Sheth, R. K., Hui, L., Diaferio, A., & Scoccimarro, R. 2001b, MNRAS, 325, 1288
Sheth, R. K., Mo, H. J., & Tormen, G. 2001c, MNRAS, 323, 1
Sheth, R. K. & Tormen, G. 1999, MNRAS, 308, 119
Smith, R. E., Peacock, J. A., Jenkins, A., White, S. D. M., Frenk, C. S., Pearce, F. R., Thomas, P. A., Efstathiou, G., & Couchman, H. M. P. 2003, MNRAS, 341, 1311
Tinker, J. L., Weinberg, D. H., & Zheng, Z. 2005a, MNRAS, Submitted, (astro-ph/0501029)
Tinker, J. L., Weinberg, D. H., Zheng, Z., & Zehavi, I. 2005b, ApJ, 631, 41
Wechsler, R. H., Zentner, A. R., Bullock, J. S., & Kravtsov, A. V. 2005, ApJ, Submitted (astro-ph/0512416)
Zehavi, I., Weinberg, D. H., Zheng, Z., Berlind, A. A., Frieman, J. A., Scoccimarro, R., Sheth, R. K., Blanton, M. R., Tegmark, M., Mo, H. J., et al. 2004, ApJ, 608, 16
Zentner, A. R., Berlind, A. A., Bullock, J. S., Kravtsov, A. V., & Wechsler, R. H. 2005, ApJ, In Press (astro-ph/0411586), 624
Zheng, Z., Berlind, A. A., Weinberg, D. H., Benson, A. J., Baugh, C. M., Cole, S., Davé, R., Frenk, C. S., Katz, N., & Lacey, C. G. 2005, ApJ, 633, 791