

# 5

## *The contributions of matter inside and outside of haloes to the matter power spectrum*

Halo-based models have been very successful in predicting the clustering of matter. However, the validity of the postulate that the clustering is fully determined by matter in haloes remains largely untested, and it is not clear a priori whether non-virialised matter might contribute significantly to the non-linear clustering signal. Here, we investigate the contribution of haloes to the matter power spectrum as a function of both scale and halo mass by combining a set of cosmological N-body simulations to calculate the contributions of different spherical overdensity regions, Friends-of-Friends groups and matter outside haloes to the power spectrum. We find that on scales  $k < 2 h \text{ Mpc}^{-1}$ , matter inside spherical overdensity regions of size  $R_{200,\text{mean}}$  accounts for less than 85% of the power, regardless of the minimum halo mass. Its relative contribution increases with increasing Fourier scale, peaking at  $\sim 95\%$  around  $k = 20 h \text{ Mpc}^{-1}$  and on smaller scales remaining roughly constant. For  $2 \lesssim k \lesssim 10 h \text{ Mpc}^{-1}$ , haloes below  $\sim 10^{11} h^{-1} M_{\odot}$  provide a negligible contribution to the power spectrum, the dominant contribution on these scales being provided by haloes with masses  $M_{200} \gtrsim 10^{13.5} h^{-1} M_{\odot}$ , even though such haloes account for only  $\sim 13\%$  of the total mass. When haloes are taken to be regions of size  $R_{200,\text{crit}}$ , the amount of power unaccounted for is larger on all scales. Accounting also for matter inside FoF groups but outside  $R_{200,\text{mean}}$  increases the contribution of halo matter on all scales probed here by 5 – 15%. Matter inside FoF groups with  $M_{\text{FoF}} > 10^9 h^{-1} M_{\odot}$  accounts for essentially all power for  $3 < k < 100 h \text{ Mpc}^{-1}$ . We therefore expect a halo model based approach to overestimate the contribution of haloes of any mass to the power on small scales ( $k \gtrsim 1 h \text{ Mpc}^{-1}$ ), while ignoring the contribution of matter outside  $R_{200,\text{mean}}$ , unless one takes the halo to be a broader non-spherical region similar to the FoF group.

Marcel P. van Daalen and Joop Schaye  
*In preparation*

## 5.1 Introduction

---

The matter power spectrum, a measure of how matter clusters as a function of scale, is a key observable and a powerful tool in determining the other cosmological parameters of our Universe. As future weak lensing experiments which will measure this quantity to unprecedented accuracy, such as DES<sup>1</sup>, LSST<sup>2</sup>, Euclid<sup>3</sup> and WFIRST<sup>4</sup>, draw ever closer, the precision with which the theoretical matter power spectrum is being predicted steadily increases as well. Currently, some of the largest uncertainties on fully non-linear scales come from our incomplete understanding of galaxy formation (e.g. van Daalen et al., 2011), which causes large unwanted biases in the cosmological parameters derived from observations. We expect that we may be able to account for these using independent measurements of, for example, the large-scale gas distribution, and/or to marginalise over these uncertainties using a halo model based approach, although for the largest of these future surveys more effective and less model-dependent mitigation strategies than currently exist will be needed (e.g. Semboloni et al., 2011; Zentner et al., 2013).

But even assuming that we can somehow account for the effects of galaxy formation on the distribution of matter, significant challenges remain before we are able to predict the matter power spectrum to the sub-percent accuracy needed to fully exploit future measurements (Huterer & Takada, 2005; Hearin, Zentner & Ma, 2012). These include converging on the “true” simulation parameters in N-body codes, although these too can be marginalised over (Smith et al., 2014). However, with each such marginalisation one should expect the constraining power of observations to be reduced.

Direct simulations are not the only way to obtain theoretical predictions of the matter power spectrum, however. Other avenues, such as through the analytical halo model (e.g. Seljak 2000, Ma & Fry 2000; see Cooray & Sheth 2002 for a review), exist, and are widely used in clustering studies. The halo model is based on the assumption that all matter is partitioned over dark matter haloes, which finds its origin in the model proposed by Press & Schechter (1974, , hereafter PS), later extended by Bond et al. (1991). The PS formalism is based on the ansatz that the fraction of mass in haloes of mass  $M(R)$  is related to the fraction of the volume that contains matter fluctuations  $\delta_R > \delta_{\text{crit}}$ , where  $R$  is the smoothing scale and  $\delta_{\text{crit}}$  is the critical density assuming spherical collapse. If the initial field of matter fluctuations is known, a halo mass function can be derived from this ansatz, which together with a model for the bias  $b(M)$  (the clustering strength of a halo of mass  $M$  relative to the clustering of matter) and a description of halo density profiles fully determines the clustering of matter.

Much work has been done to improve the predictions of the halo model since its introduction. More accurate mass functions have been derived based on, for

---

<sup>1</sup><http://www.darkenergysurvey.org/>

<sup>2</sup><http://www.lsst.org/lsst>

<sup>3</sup><http://www.euclid-imaging.net/>

<sup>4</sup><http://wfirst.gsfc.nasa.gov/>

example, ellipsoidal collapse (Sheth, Mo & Tormen, 2001), fits to N-body simulations (e.g. Jenkins et al. 2001; Warren et al. 2006; Reed et al. 2007; Tinker et al. 2008; Bhattacharya et al. 2011; Angulo et al. 2012; Watson et al. 2013; see Murray, Power & Robotham 2013 for a comparison of different models) and simulations taking into account the effects of baryons (e.g. Stanek, Rudd & Evrard, 2009; Sawala et al., 2013; Martizzi et al., 2014; Cusworth et al., 2014; Khandai et al., 2014; Cui, Borgani & Murante, 2014; Velliscig et al., 2014). Similarly, much effort has gone into deriving more accurate (scale-dependent) bias functions (e.g. Sheth & Tormen, 1999; Seljak & Warren, 2004; Smith, Scoccimarro & Sheth, 2007; Reed et al., 2009; Grossi et al., 2009; Manera, Sheth & Scoccimarro, 2010; Pillepich, Porciani & Hahn, 2010; Tinker et al., 2010) and concentration-mass relations for halo profiles (e.g. Bullock et al., 2001; Eke, Navarro & Steinmetz, 2001; Neto et al., 2007; Duffy et al., 2008; Macciò, Dutton & van den Bosch, 2008; Prada et al., 2012; Ludlow et al., 2014). Current halo models may incorporate additional ingredients like triaxiality, substructure, halo exclusion, primordial non-Gaussianity and baryonic effects (e.g. Sheth & Jain, 2003; Smith & Watts, 2005; Giocoli et al., 2010; Smith, Desjacques & Marian, 2011; Gil-Marín, Jimenez & Verde, 2011; Fedeli, 2014), and fitting formulae based on the halo model have also been developed (e.g. Smith et al., 2003b; Takahashi et al., 2012).

However, the validity of the postulate that the clustering of matter is fully determined by matter in haloes remains relatively untested. Even though matter is known to occupy non-virialised regions such as filaments, their mass may simply be made up of very small haloes itself, although recent results indicate that part of the dark matter accreted onto haloes is genuinely smooth (Angulo & White, 2010b; Fakhouri & Ma, 2010; Genel et al., 2010; Wang et al., 2011). Either way, it is not clear a priori whether this non-virialised matter contributes significantly to the non-linear clustering signal.

Here, we examine the contributions of halo and non-halo mass to the matter power spectrum with the use of a set of N-body simulations. We will first investigate the contribution to the redshift zero matter power spectrum of haloes that are defined analogous to the typical halo model approach, also examining the contributions of matter in smaller overdensity regions and outside of haloes. Next, we expand the haloes to include all matter associated to Friends-of-Friends (FoF) groups. Finally, we make predictions for the contribution of halo matter to the power spectrum as a function of both scale and minimum halo mass, which can serve as a test for halo models aimed at reproducing the clustering of dark matter.

This chapter is organized as follows. In §5.2 we describe our simulations and the employed power spectrum estimator. We present and discuss our results in §5.3 and summarise our findings in §5.4.

Name	Box size [ $h^{-1}$ Mpc]	Particle number	$m_{\text{dm}}$ [ $h^{-1} M_{\odot}$ ]	$\epsilon_{\text{max}}$ [ $h^{-1}$ kpc]
<i>L400N1024</i>	400	$1024^3$	$4.50 \times 10^9$	4.00
<i>L200N1024</i>	200	$1024^3$	$5.62 \times 10^8$	2.00
<i>L050N512</i>	50	$512^3$	$7.03 \times 10^7$	1.00
<i>L025N512</i>	25	$512^3$	$8.79 \times 10^6$	0.50

TABLE 5.1: The different simulations employed in this chapter. From left to right, the columns list their name, box size, particle mass and maximum proper softening length. All simulations were run with only dark matter particles and a WMAP7 cosmology.

## 5.2 Method

---

### 5.2.1 Simulations

We base our analysis on a set of dark matter only runs that were run with a modified version of GADGET III, the smoothed-particle hydrodynamics (SPH) code last described in Springel (2005b). The cosmological parameters are derived from the Wilkinson Microwave Anisotropy Probe (WMAP) 7-year results (Komatsu et al., 2011), and given by  $\{\Omega_{\text{m}}, \Omega_{\text{b}}, \Omega_{\Lambda}, \sigma_8, n_{\text{s}}, h\} = \{0.272, 0.0455, 0.728, 0.81, 0.967, 0.704\}$ .

We generate initial conditions assuming the Eisenstein & Hu (1998) transfer function. Prior to imposing the linear input spectrum, the particles are set up in an initially glass-like state, as described in White (1994). The particles are then evolved to redshift  $z = 127$  using the Zel'dovich (1970) approximation.

The relevant parameters of the simulations we employ here are listed in Table 5.1. The simulation volumes range from  $25 h^{-1}$  Mpc to  $400 h^{-1}$  Mpc. The mass resolution improves by a factor of 8 with each step, corresponding to an improvement of the spatial resolution by a factor of 2, from the largest down to the smallest volume. The gravitational forces are softened on a comoving scale of  $1/25$  of the initial mean inter-particle spacing,  $L/N$ , but the softening length is limited to a maximum physical scale of  $2 h^{-1}$  kpc [ $L/(100 h^{-1}$  Mpc)] which is reached at  $z = 2.91$ . As we will demonstrate, by combining these simulations, we can accurately determine the matter power spectrum from linear scales up to  $k \sim 100 h \text{ Mpc}^{-1}$ .

### 5.2.2 Power spectrum calculation

The matter power spectrum is a measure of the amount of structure that has formed on a given Fourier scale  $k$ , related to a physical scale  $\lambda$  through  $k = 2\pi/\lambda$ . It is defined through the Fourier transform of the density contrast,  $\hat{\delta}_{\mathbf{k}}$ . We will present our results in terms of the dimensionless power spectrum, defined in the

usual way:

$$\Delta^2(k) = \frac{k^3}{2\pi^2} P(k) = \frac{k^3 V}{2\pi^2} \left\langle |\hat{\delta}_{\mathbf{k}}|^2 \right\rangle_{\mathbf{k}}, \quad (5.1)$$

with  $V$  the volume of the simulation under consideration. As all particles have the same mass, the shot noise is simply equal to  $\langle |\hat{\delta}_{\mathbf{k}}|^2 \rangle_{\mathbf{k}, \text{shot}} = 1/N_{\text{p}}$ , with  $N_{\text{p}}$  the number of particles in the simulation. All power spectra presented here have had shot noise subtracted to obtain more accurate results on small scales.

We calculate the matter power spectrum using the publicly available F90 package POWMES (Colombi et al., 2009). The advantages of POWMES stem from the use of the Fourier-Taylor transform, which allows analytical control of the biases introduced, and the use of foldings of the particle distribution, which allow the dynamic range to be extended to arbitrarily high wave numbers while keeping the statistical errors bounded. For a full description of these methods we refer to Colombi et al. (2009). As in van Daalen et al. (2011), we set the grid parameter to  $G = 256$  and use a folding parameter  $F = 7$  for the two smallest volumes. To calculate the power spectrum down to similar scales for the 200 and 400  $h^{-1}$  Mpc boxes, we set  $F$  equal to 8 and 9, respectively. Our results are insensitive to this choice of parameters.

Both box size and resolution effects lead to an underestimation of the power – at least on scales where a sufficient number of modes is available so that the effects of mode discreteness can be ignored ( $k \gtrsim 8\pi/L$ ) – while all simulations show excellent agreement on scales where they overlap (see §5.2). In order to cover the dynamic range from  $k = 0.01 h \text{ Mpc}^{-1}$  to  $100 h \text{ Mpc}^{-1}$ , we therefore combine the power spectra of different simulations. By always taking the largest value of  $\Delta^2(k)$  at each  $k$ . In the case of the full power spectrum, i.e. the power spectrum of all matter, we take the combined power spectrum to be the one predicted by linear theory up to  $k = 0.12 h \text{ Mpc}^{-1}$ , where the power starts to become non-linear. While the largest boxes show excellent agreement with the linear power spectrum on these scales, we wish to avoid box size effects as much as possible. For  $k > 0.12 h \text{ Mpc}^{-1}$  – or, in the case of power spectra of subsets, for the smallest  $k$ -value available – we individually average each power spectrum over each of 25 bins in Fourier space  $k_i$  and assign the combined power spectrum the largest  $\Delta^2(k_i)$  of all simulations derived in this way.

We combine the power spectra of selections of particles (e.g. all particles that reside in haloes above a certain mass) in a similar way, but without including the linear theory power spectrum.

### 5.2.3 Halo particle selection

In the halo model approach, haloes are commonly defined through a spherical overdensity criterion, usually relative to the mean density of the Universe. In order to investigate the contribution of such haloes to the matter power spectrum, we define our haloes consistently.

Overdense regions are identified in our simulations with a spherical overdensity finder, as implemented in the SUBFIND algorithm (Springel et al., 2001). We define a halo as a spherical region with an internal mass overdensity of  $200 \times \Omega_m \rho_{\text{crit}}$ , where  $\rho_{\text{crit}}$  is the critical density of the Universe. These haloes therefore have a mass equal to:

$$M_{200} = M_{200,\text{mean}} = 200 \times \frac{4\pi}{3} \Omega_m \rho_{\text{crit}} R_{200}^3, \quad (5.2)$$

where  $R_{200} = R_{200,\text{mean}}$  is the radius of the region. In the remainder of the chapter, we will define halo particles as any particle with a distance  $R < R_{200}$  from any halo centre. All other particles are treated as non-halo particles, irrespective of their possible FoF group membership, or having been identified as part of a bound subhalo by SUBFIND.

While we focus on halo matter as defined through  $R_{200}$ , we will also briefly discuss the contribution of halo matter to the power spectrum for other overdensity regions and halo definitions (i.e.  $R_{500}$ ,  $R_{2500}$ ,  $R_{200,\text{crit}}$  and Friends-of-Friends) during the course of the chapter.

## 5.3 Results

---

### 5.3.1 Fractional mass in haloes

We first examine the fraction of the mass that resides in haloes,  $f_h$ . As in each simulation there is a lower limit to the masses of haloes that we can reliably resolve, we compute  $f_h$  as a function of the minimum mass of the included haloes. Knowing the minimum resolved masses also allows us to estimate over which halo mass range we can probe the contribution of halo particles to the power spectrum in each simulation.

The results for  $f_h$  are shown in Figure 5.1. Different colours are used for each of our four different simulations, as indicated in the legend. Vertical dotted lines denote the masses corresponding to 100 particles. Below these limits the fraction of mass in haloes flattens off, indicating that such low-mass haloes are unresolved. A thick dashed line shows the result of combining the mass fractions of all four simulation for  $M_{\text{min}} > 10^9 h^{-1} M_\odot$ , through  $f_{h,\text{comb}} = \max(f_{h,i})$ . The bottom panel shows the ratio of  $f_h$  of each simulation to this combined fraction.

At the massive end, the high-resolution but low-volume *L025* and *L050* simulations significantly underestimate  $f_h$ . This is most clearly seen in the bottom panel: for *L025* the mass fraction in haloes becomes significantly underestimated for halo masses  $M_{200} \gtrsim 10^{11} h^{-1} M_\odot$ , while for the *L050* box this happens for  $M_{200} \gtrsim 10^{12} h^{-1} M_\odot$ . This is consistent with the points at which the halo mass functions are underestimated for these simulations (not shown). The fluctuations seen in the bottom panel for *L200* for  $M_{200} > 10^{14} h^{-1} M_\odot$  are due to the rarity of such massive haloes, but as the fraction of the mass residing in such haloes is  $< 10\%$  this does not impact our conclusions. All simulations in which haloes at a

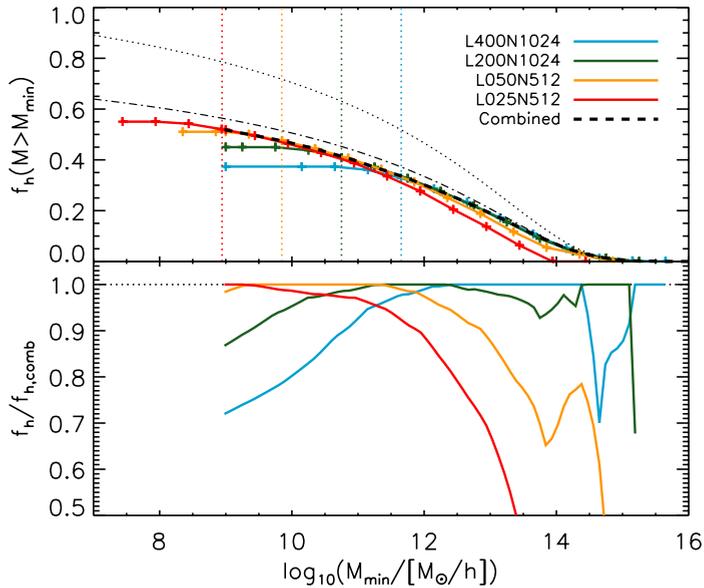


FIGURE 5.1: The cumulative fraction of mass inside haloes,  $f_h$ , as a function of minimum halo mass, for different collisionless simulations as indicated in the legend. The resolution limit, defined as the mass of haloes containing 100 particles, is shown as a vertical dotted line for each simulation. Below this limit, the fraction of mass in haloes is underestimated. For the two highest-resolution simulations these fractions are also significantly underestimated at high masses, as these haloes are under-represented in these small volumes. Between the limits imposed by resolution and box size effects, the simulations are in excellent agreement, and show that the fraction of mass in haloes is  $\sim 52\%$  for  $M_{200} > 10^9 h^{-1} M_\odot$ . A black dashed line shows the combined result, taking the maximum fraction of mass in haloes between the different simulations at every mass, while the bottom panel shows the fraction of this combined function predicted by each simulation. We also show predictions for the Tinker et al. (2008) mass function as a black dotted and dot-dashed line (see main text).

certain  $M_{\min}$  are both well-resolved and well-represented show excellent agreement in  $f_h(M > M_{\min})$ .

The fraction of mass in haloes increases with decreasing halo mass. Only  $\sim 19\%$  of matter is found in groups and clusters ( $M_{\min} > 10^{13} h^{-1} M_\odot$ ), which increases to  $\sim 30\%$  for Milky Way haloes and up ( $M_{\min} > 10^{12} h^{-1} M_\odot$ ). But even at the lowest resolved mass of roughly  $10^9 h^{-1} M_\odot$ , the fraction of mass in haloes is still barely more than 50%. We therefore expect a significant contribution from particles in haloes with  $M < 10^9 h^{-1} M_\odot$ , and possibly from dark matter particles that do not reside in haloes of any mass, to the matter power spectrum on large scales, which we calculate in the next section.

For comparison, the top panel of Figure 5.1 also shows predictions of the fraction of mass in haloes above a certain mass based on the Tinker et al. (2008)  $M_{200}$  halo mass function. Using the normalized halo mass function fit provided by these authors, we have calculated  $f_h(M > M_{\min})$  under the standard halo

model assumption that all mass resides in haloes. The results are shown by the black dotted line. Under this assumption, far more mass is predicted to reside within resolved haloes than we find for our simulations, at any mass. However, the Tinker et al. (2008) mass function converges to a mass density of only about  $0.72 \times \Omega_m \rho_{\text{crit}}$ , meaning that either the true mass function predicts far more mass in haloes  $M_{200} \lesssim 10^{11} h^{-1} M_{\odot}$  (roughly the lowest-mass haloes considered by Tinker et al. 2008), or that about 28% of the dark matter mass is genuinely smoothly distributed at  $z = 0$ , not residing in haloes of any mass. To compensate for this “missing mass”, we have also calculated the mass in haloes above some  $M_{\text{min}}$  predicted by the Tinker et al. (2008) mass function relative to the total mass in the Universe. This result is shown by the dot-dashed line, and shows much better agreement with our simulations.

Up to  $M_{\text{min}} \approx 10^{12}$  the relative difference between the Tinker et al. (2008) prediction and our combined result is constant at about 10% before decreasing at higher masses. One possible reason for this discrepancy is that we only count matter in regions where haloes overlap once, which is not taken into account when integrating the mass function. However, we have checked that the mass residing in overlap regions in our simulation is always  $\lesssim 1.7\%$ , with the largest overlap fraction being found for the most massive haloes. The  $\lesssim 10\%$  differences found for  $f_h$  are therefore likely due to the non-universality of the halo mass function at this level of precision (e.g. Tinker et al., 2008; Murray, Power & Robotham, 2013).

As it seems that  $f_h(M > M_{\text{min}})$  continues to rise on mass scales unresolved by our simulations, the total contribution of matter in haloes to the power spectrum will be underestimated in our simulations. However, as we will see in §5.3.2.1, this depends on the scale considered. A range in Fourier space exists where the fraction of power from halo particles is bounded below unity, and the contribution of haloes with masses  $M_{200} \lesssim 10^{11} h^{-1} M_{\odot}$  is negligible. Additionally, on scales where this does not hold we can still constrain the contribution from haloes above a certain mass.

In the remainder of the chapter, we will only consider particles residing in haloes with  $M_{200} > 10^9 h^{-1} M_{\odot}$  to be halo particles, as this corresponds roughly to the smallest haloes we can resolve.

### 5.3.2 Halo contribution to the power spectrum

We first show the full dimensionless matter power spectrum, i.e. using all particles, in Figure 5.2. Here each simulation is shown by a different colour, and it is immediately clear that no single one is converged over the full dynamic range up to  $k \sim 100 h \text{ Mpc}^{-1}$ . The linear theory power spectrum, as generated by the F90 package CAMB (Lewis, Challinor & Lasenby, 2000, , version January 2010), is shown as a long dashed purple line. *L400* and *L200* show good agreement with the linear power spectrum on scales where non-linear evolution is negligible ( $k \lesssim 0.12 h \text{ Mpc}^{-1}$ ) and a sufficient number of modes is available ( $k > 0.04$  and  $0.08 h \text{ Mpc}^{-1}$  respectively, roughly corresponding to  $\lambda = 0.4 L$ ), while *L050* and

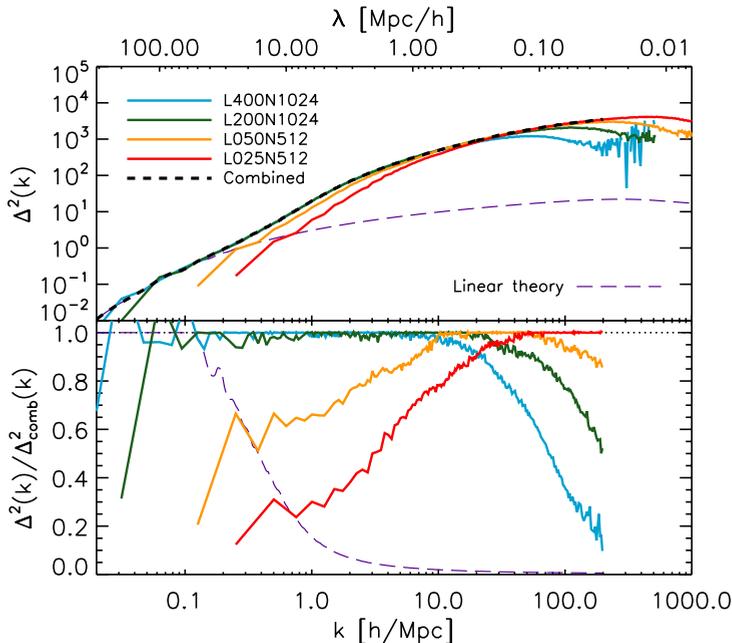


FIGURE 5.2: The dimensionless power spectrum derived from each simulation, along with the linear power spectrum (long-dashed purple line) and the combined power spectrum (dashed black line). While the *L025* and *L050* simulations significantly underestimate the power on large scales due to missing modes, their high resolution allows us to accurately extend the power spectrum of the larger volumes up to  $k \sim 100 h \text{ Mpc}^{-1}$ . The erratic behaviour seen for low-resolution simulations at large  $k$  is due to shot noise subtraction. The bottom panel shows for each simulation (as well as for the linear theory prediction) the fraction of power relative to the combined power spectrum. For  $k < 20 h \text{ Mpc}^{-1}$ , multiple simulations show the same results, indicating convergence on these scales.

*L025* show severe box size effects due to their lack of large-scale modes. These box size effects become negligible only for  $k > 10$  and  $k > 40 h \text{ Mpc}^{-1}$  respectively.

Due to their finite resolution, all simulations underestimate the power on sufficiently small scales. Note that all power spectra shown here have had shot noise subtracted, which explains the erratic behaviour of the power spectra on the smallest scales. The underestimation of small-scale power becomes significant already on scales corresponding to  $\sim 100$  softening lengths. However, for every  $k \lesssim 100 h \text{ Mpc}^{-1}$ , there is at least one simulation for which neither box size nor resolution leads to an underestimation of the power at the  $\gtrsim 1\%$  level. We therefore combine the different power spectra as described in §5.2.2 to obtain the combined power spectra, shown as a dashed black line.

The bottom panel of Figure 5.2 shows the fraction of power predicted by each simulation, as well as the fraction predicted by linear theory, relative to the combined power spectrum. By construction, this fraction is bounded to unity on

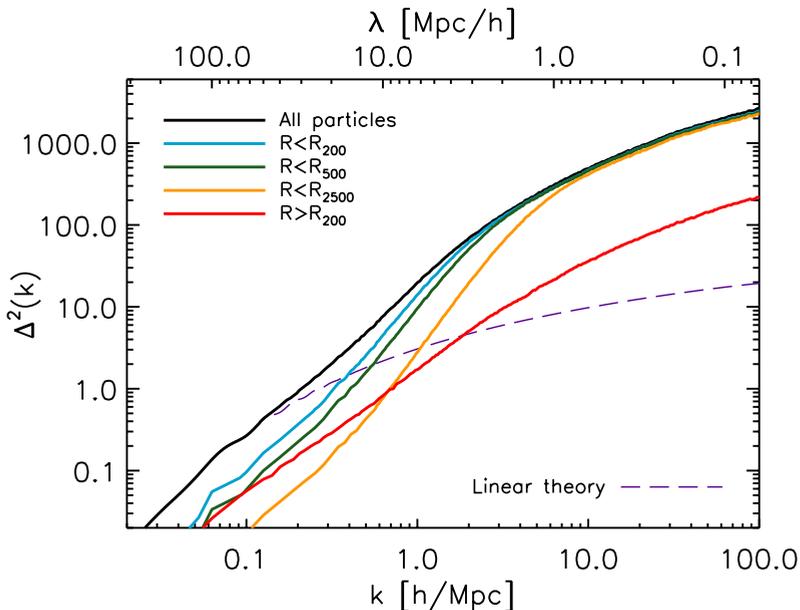


FIGURE 5.3: The combined power spectrum for different sets of particles:  $R < R_{200}$  (halo particles),  $R < R_{500}$ ,  $R < R_{2500}$  and  $R > R_{200}$  (non-halo particles). Only haloes with  $M > M_{\min} = 10^9 h^{-1} M_{\odot}$  were considered in the cuts made, which in total contain about 38% of all dark matter. The halo particles easily dominate the power on small scales; however, there is a significant range of non-linear scales ( $k < 1.4 h \text{ Mpc}^{-1}$ ) where this subset does not provide the most power. While the non-halo particles account for a larger fraction of the power for  $k < 0.6 h \text{ Mpc}^{-1}$ , it is the cross-terms between the halo and non-halo particles (not shown) which dominate in the mildly non-linear regime. Note that the horizontal range has been shortened relative to Figure 5.2.

non-linear scales. Note that on scales  $k \lesssim 20 h \text{ Mpc}^{-1}$ , the fractions of multiple simulations are within a few percent of unity, indicating convergence on these scales. For smaller scales, however, convergence is uncertain, although based on the results for larger scales we expect our combined power spectrum to be accurate to  $\sim 1\%$  up to  $k \sim 100 h \text{ Mpc}^{-1}$ .

Next, we repeat this procedure for halo and non-halo particles. We also consider particles within the  $R_{500}$  and  $R_{2500}$  overdensity regions, defined analogously to  $R_{200}$ , which probe the inner parts of haloes. As we cannot reliably resolve haloes with less than about 100 particles with our highest-resolution simulation, we only consider the contribution of haloes with masses  $M > M_{\min} = 10^9 h^{-1} M_{\odot}$  here, treating matter in lower-mass haloes as non-halo particles. The results are shown in Figure 5.3. Note that for clarity only the combined power spectra are shown, and that the horizontal range has been shortened with respect to Figure 5.2, only showing the range of scales for which we can reliably determine the power spectrum.

The contribution from halo particles strongly dominates the power on small scales. The halo contribution is in turn dominated by the very inner regions of haloes, at least on scales smaller than the size of these regions. However, towards larger scales this contribution diminishes, and for  $k < 0.4 h \text{ Mpc}^{-1}$  less than half of the total power is provided by matter in haloes alone. On large scales the significant fraction of the mass that occupies non-virialised regions becomes more important, increasing to about 20%, roughly half of the contribution of halo matter on the same scales. The remaining  $\sim 40\%$  of the total matter power on large scales is therefore contributed by the cross-terms of halo and non-halo matter (not shown here).

Note that on the scales shown here, only  $L400$  and  $L200$  contribute to the combined power spectrum of non-halo particles. However, as these two are in excellent agreement for  $k \gtrsim 0.4 h \text{ Mpc}^{-1}$  even though the mass resolution is eight times worse in  $L400$ , we do not have reason to believe that this component would significantly change on non-linear scales if lower-mass haloes were resolved. On linear scales, the contribution of halo matter is mostly determined by the fraction of mass in haloes, which does of course depend on the minimum halo mass resolved. We will return to this point in §5.3.2.1.

We investigate the contribution of halo matter in more detail in Figure 5.4, which shows the ratio of the power spectrum of matter within  $R_{200}$  of haloes with masses  $M_{200} > 10^9 h^{-1} M_{\odot}$  to the power spectrum of all matter. The black dashed line shows the ratio of the combined power spectra, obtained from the smoothed power spectra of all four simulations shown here as described in §5.2.2, relative to the combined total power spectrum (black line in Figure 5.3). The solid lines show the relative contribution of halo matter in each simulation separately.

The contribution of halo matter to the total power increases with decreasing physical scale. On large (linear) scales, the contribution from haloes seems to converge to  $\sim 30\%$ , in good agreement with  $f_h(M > 10^9 h^{-1} M_{\odot})^2 \approx 0.27$ . This is expected, as the contribution of any subset of matter to the power spectrum on linear scales should scale only with (the square of) the fraction of mass contained in such a subset. However, as the fraction of power in haloes on large scales is fully determined by  $L200$  and  $L400$ , with both predicting roughly the same fraction as can be seen in Figure 5.4, while the fraction of mass in haloes  $M > 10^9 h^{-1} M_{\odot}$  is only accurately measured for  $L025$ , this correspondence is actually surprising.

On non-linear scales the ratio rapidly increases down to physical scales of  $\lambda \sim 2 h^{-1} \text{ Mpc}$  ( $k \sim 3 h \text{ Mpc}^{-1}$ ), reaching at most 95%, before slowly levelling off towards smaller scales. Note that the combined results are fully determined by  $L050$  around  $k \approx 20 h \text{ Mpc}^{-1}$ , where we are unable to show convergence due to the too-low resolution of  $L200$  and too-small volume of  $L025$ . However, the results of §5.3.2.1 imply that little would change on these scales if higher-resolution simulations were available.

While  $L400$  and  $L200$  are in good agreement for  $0.2 \lesssim k \lesssim 10 h \text{ Mpc}^{-1}$ , on sub-Mpc scales the contribution of halo matter to the total matter power spectrum starts to show a strong dependence on resolution. On these scales fluctuations

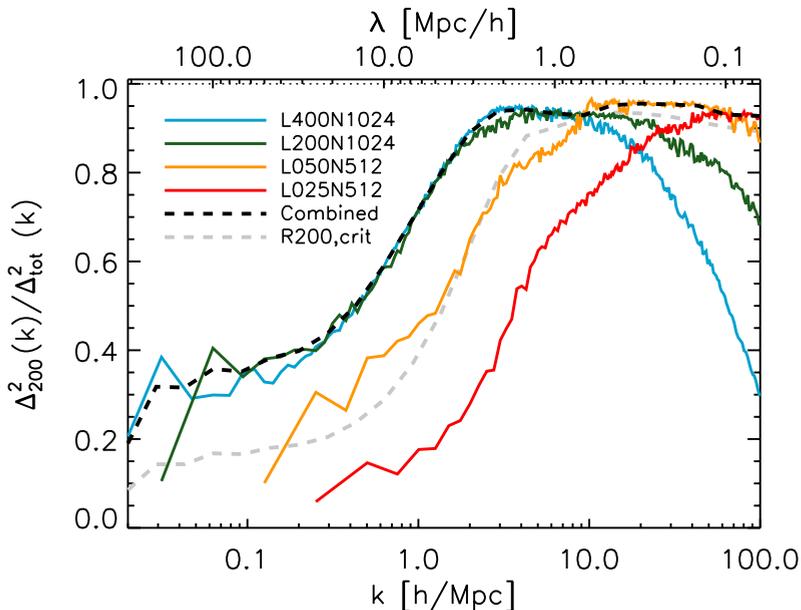


FIGURE 5.4: The fraction of power within haloes with masses  $M_{200} > 10^9 h^{-1} M_{\odot}$  as a function of scale. A dashed black line again shows the combined power spectrum derived from the smoothed power spectra of the four simulations employed in this chapter, each of which is shown as well. The halo contribution rapidly rises down to  $\lambda \sim 2 h^{-1} \text{Mpc}$ , peaking at  $\sim 95\%$  for  $k \approx 20 h \text{Mpc}^{-1}$  ( $\lambda \approx 300 h^{-1} \text{kpc}$ ) and remaining roughly constant for larger  $k$ . The power spectrum on smaller scales is dominated by increasingly smaller haloes, while the power spectrum on the largest scales depends mainly on the total mass fraction. The grey dashed line shows the result if  $R_{200,\text{crit}}$  is used instead of  $R_{200}$ .

within the same halo dominate the power spectrum (i.e. the 1-halo term in halo model terminology), so naturally the contribution to the power will be underestimated on scales  $\lambda \lesssim R_{200,\text{min}}$ , where  $R_{200,\text{min}}$  is the virial radius of a halo with the minimum resolved mass,  $M_{\text{min}}$ , in that particular simulation. In practice, the power is significantly underestimated already on larger scales, due to the gravitational softening employed in the simulation, which leads to an underestimation of the inner density of haloes. This leads to a significant power deficit on scales  $\lambda \lesssim 100 \epsilon_{\text{max}}$  (see Table 5.1). Fortunately, the combination of simulations chosen here still allows us to probe the contribution of halo matter up to  $k_{\text{max}} \sim 100 h \text{Mpc}^{-1}$ .

As the power on sub-Mpc scales is dominated by the 1-halo term, adding lower-mass haloes than those resolved here will have a negligible impact on the measured contribution of halo matter on the scales considered here, as  $2\pi/R_{200,\text{min}} > k_{\text{max}}$ . Therefore, 5 – 7% of small-scale power is unaccounted for by halo particles, regardless of resolution effects. Instead, it is the cross-term between halo matter and matter *just outside* the  $R_{200}$  regions that makes up the deficit.

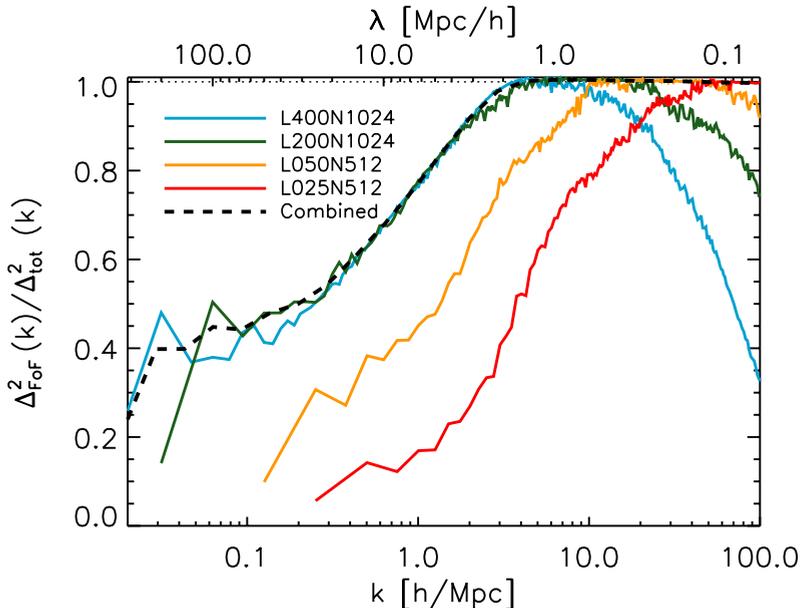


FIGURE 5.5: As Figure 5.4, but now for all mass inside FoF groups with  $M_{\text{FoF}} > 10^9 h^{-1} M_{\odot}$ . While the scale dependence is very similar (i.e. a rapid rise down to  $\lambda \sim 1 h^{-1} \text{Mpc}$  and roughly constant on smaller scales), the contribution to the power spectrum is higher than for the  $R_{200}$  overdensity regions on any scale. The contribution of halo matter power spectrum is increased by 5 – 10% on all scales relative to the results of Figure 5.4, and matter in FoF groups accounts for essentially all power on scales  $k > 3 h \text{Mpc}^{-1}$ . This implies that the  $R_{200}$  overdensity regions do not fully capture the halo.

To demonstrate that this is indeed the case, we calculate the contribution of matter in FoF groups (with a linking length of 0.2) to the total power spectrum, with a mass limit of  $M_{\text{FoF}} > 10^9 h^{-1} M_{\odot}$ . The results are shown in Figure 5.5. Here we see that, while the scale-dependence of the contribution is similar to that shown in Figure 5.4 for the  $R_{200}$  regions, the contribution is significantly larger on all scales, and is essentially 100% for  $k \gtrsim 3 h \text{Mpc}^{-1}$ . This Fourier scale corresponds well to the virial radius of the largest clusters in the simulation.

On scales  $k \lesssim 0.3 h \text{Mpc}^{-1}$ , the contribution of matter in FoF groups is consistently  $\sim 25\%$  higher than that of matter in  $R_{200}$  haloes. Correspondingly, the fraction of mass in FoF groups is also higher than the fraction of mass in  $R_{200}$  haloes at every mass, from 10% higher at  $M = 10^9 h^{-1} M_{\odot}$  to 15% higher at  $M = 10^{14} h^{-1} M_{\odot}$  (not shown).

Finally, we also show the results if  $R_{200,\text{crit}}$  is used instead (with  $M_{200,\text{crit}} > 10^9 h^{-1} M_{\odot}$ ), as a dashed grey line in Figure 5.4. As such an overdensity criterion picks out smaller regions than  $R_{200}$ , containing less mass, the contribution of halo matter to the power spectrum is also smaller, especially on large scales. On sub-Mpc scales, however, the differences are small, with the contribution to the power

spectrum of halo matter peaking at 94%.

We conclude that what region is chosen to represent a halo has a large impact on the contribution of haloes to the matter power spectrum, in a scale-dependent way. In what follows, we will continue to define haloes using the mean overdensity criterion, as this is typically used in the halo model approach.

### 5.3.2.1 Contribution as a function of mass

To see which haloes contribute most to the matter power spectrum as a function of scale, while simultaneously examining the dependence of our results on the mass of the lowest resolved halo, we turn to Figure 5.6. Each panel corresponds to a different simulation and each curve to a different minimum halo mass. The halo contributions are shown relative to the combined power spectrum of all matter (black line in Figure 5.3). The legend shows the minimum halo mass  $\log_{10}(M_{\min}/[M_{\odot}/h])$ . Note that the minimum masses differ for each simulation, because these are based on the particle masses of the simulations, in such a way that the first fractional contribution to the power shown includes matter in haloes of 100 particles or more, corresponding to our imposed resolution limit in mass (see Figure 5.1). The minimum halo mass increases by half a dex with each step. Grey regions indicate the approximate scales on which the full matter power spectrum of the simulation is not converged to  $\sim 1\%$  with respect to the combined one. While this gives an indication of which scales to trust, note that the relative contribution of each halo mass may to be converged for a different range of scales. Finally, the bottom half of each panel shows the difference between consecutive lines, i.e. the contribution added by decreasing the minimum halo mass by half a dex. Here  $f_{\Delta,i} \equiv \Delta_{200,i}^2/\Delta_{\text{all}}^2$ . As we will show shortly, while the relative contributions of haloes of a certain mass shown in the bottom halves of the panels can be compared between different simulations, the same does not hold for the absolute contributions, as box size effects play an important role on a large range of scales.

Several things can be learned from this figure. First, we can ask whether we are resolved with respect to the minimum mass, and on which scales. Low-mass haloes become increasingly important towards larger values of  $k$  on sub-Mpc scales. As we discussed before, we therefore do not expect to be converged on scales  $k \gtrsim 2\pi/R_{200,\min}$ , limiting us to  $k \lesssim 100 h \text{ Mpc}^{-1}$  when all simulations are combined. Interestingly, on scales  $k \sim 10 h \text{ Mpc}^{-1}$ , roughly where the contribution from halo matter plateaus, the *L200* simulation (top right panel) is just about converged with minimum halo mass. This can be seen most clearly by comparing the relative halo contribution on this scale at approximately fixed minimum halo mass between different simulations: as for example the *L050* simulation (bottom left panel) shows, well-resolved haloes below  $\sim 10^{11} h^{-1} M_{\odot}$  (about a factor of two above the halo resolution limit of *L200*) provide a negligible contribution ( $< 1\%$ ) at  $k \approx 10 h \text{ Mpc}^{-1}$ , indicating convergence on this scale. Indeed, both simulations agree that most of the power at  $k \sim 10 h \text{ Mpc}^{-1}$  comes from haloes with masses  $M_{200} \gtrsim 10^{13} h^{-1} M_{\odot}$ . These group and cluster-scale haloes remain

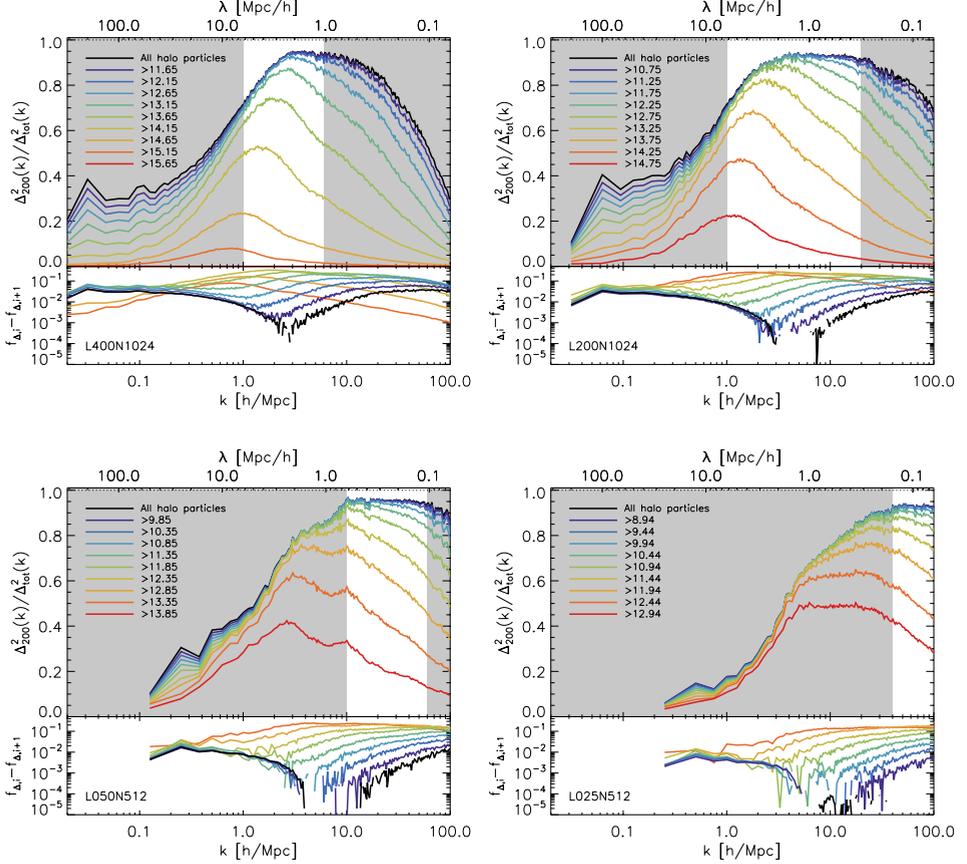


FIGURE 5.6: Comparison of the contribution of haloes above a certain mass to the matter power spectrum relative to the total combined power spectrum of all simulations, for *L400* (top left), *L200* (top right), *L050* (bottom left) and *L025* (bottom right). The legend shows the minimum halo mass  $\log_{10}(M_{\text{min}}/[M_{\odot}/h])$ . Note that lines of the same colour do not correspond to the same minimum halo mass in the four panels, as the binning is based on the minimum resolved halo mass (see text). The grey regions denote where box size or resolution effects are  $\gtrsim 1\%$  for the full power spectrum; the relative contribution may be converged on a different range, as can be seen by comparing the panels. The bottom half of each panel shows the difference between consecutive lines, i.e. the contribution added by decreasing the minimum halo mass by half a dex. Note that scales exist where we are converged with minimum halo mass: for example, around  $k = 4 h \text{ Mpc}^{-1}$  haloes with  $M_{200} < 10^{11} h^{-1} M_{\odot}$  contribute negligibly to the total matter power spectrum. For  $1 < k < 10 h \text{ Mpc}^{-1}$ , the power spectrum is dominated by the contributions of haloes with  $M_{200} > 10^{13} h^{-1} M_{\odot}$ , even though these account for only about 19% of the total mass. On large scales, haloes in the full mass range probed here contribute significantly to the power, and no convergence is obtained. Note that box size effects strongly influence the contributions of large haloes measured for the smallest two boxes, especially in the case of *L025*.

the dominant contributors on somewhat larger scales as well, their contribution peaking around  $k = 2 - 3 \text{ h Mpc}^{-1}$  before gradually falling off. Note that haloes with  $M_{200} > 10^{13} h^{-1} M_{\odot}$  account for only about 19% of the total mass (see Figure 5.1).

On larger scales, convergence is obtained for increasingly larger minimum halo masses: as the bottom halves of the panels show, around  $k = 4 \text{ h Mpc}^{-1}$  haloes with masses  $M_{200} < 10^{11} h^{-1} M_{\odot}$  do not make a significant contribution to the power spectrum. While the simulations shown here do not predict exactly the same contribution to the power spectrum for haloes with  $M_{200} > 10^{11} h^{-1} M_{\odot}$ , due to the limited box size of *L050* and *L025*, it is clear that the results for the relevant importance of such haloes are converged.

However, for  $k < 1 \text{ h Mpc}^{-1}$  the contributions of haloes which are not well resolved by *L400* and *L200* once again become important, and even *L050* may no be longer converged with minimum halo mass at the 1% level. This is because on these scales the matter power spectrum is dominated by the cross-correlation of matter in different haloes (i.e. the 2-halo term in halo model terminology). Therefore, while the bottom panels for each simulation do show that the relative contribution of haloes on small enough scales decreases with decreasing  $M_{200}$ , on scales  $k \lesssim 1 \text{ h Mpc}^{-1}$  our results can only provide a minimum contribution of halo matter to the total matter power spectrum.

As we noted before, the panels show that even when compared at roughly the same minimum halo mass, different simulations make different predictions for the contributions of haloes above a certain mass to the matter power spectrum. When the box size decreases, the contribution on large scales and that of high-mass haloes decreases as well. This is expected, as large-scale modes are missed in the smaller boxes and massive haloes are under-represented. However, the role of low-mass haloes is simultaneously overestimated in the smaller boxes, and the total contribution of halo matter at fixed minimum halo mass is therefore higher than it should be.

To demonstrate explicitly that this is the case, we show in Figure 5.7 again the results for *L200* (top-right panel in Figure 5.6), but with the results for *L100* superimposed as dashed lines. The *L100* simulation has  $512^3$  particles and therefore the same resolution as *L200*, but in an  $8\times$  smaller volume. Comparing the two simulations therefore shows the effects of box size at fixed resolution. On large scales and for high-mass haloes, the contribution of halo matter is underestimated in *L100*, relative to *L200*. Meanwhile, the contribution of low-mass haloes on small scales tends to be overestimated, even though the resolution is identical. Interestingly, there are mass and spatial scales where the simulations are in perfect agreement, such as for a minimum halo mass of  $10^{13.75} h^{-1} M_{\odot}$  where  $k > 3 \text{ h Mpc}^{-1}$ . Most important, however, is that the contributions at a certain halo mass shown in the bottom half of the panel are in perfect agreement over the entire range of scales, excepting the very highest mass bin (which is under-represented in *L100*) and the principal modes. This shows that we can still derive the correct contribution of haloes within a certain mass range, and investigate whether we are converged with

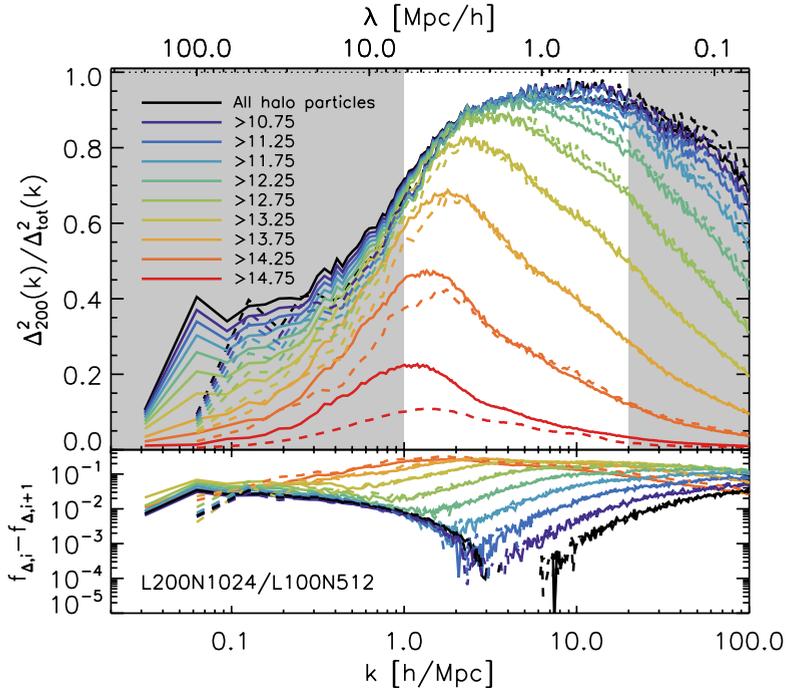


FIGURE 5.7: As the top-right panel of Figure 5.6, but with the results for  $L100$  added for the same minimum halo masses as dashed lines, showing the effects of box size at fixed resolution. Due to the missing large-scale modes in  $L100$ , the large-scale contribution is underestimated. Additionally, high-mass haloes are under-represented and the role of low-mass haloes on small scales is overestimated. However, the relative contributions of haloes of a certain mass shown in the bottom half of the figure are in excellent agreement for all but the highest mass bin.

mass on a certain scale, even when box size effects play a role.

## 5.4 Summary & conclusions

In this work we investigated the contribution of haloes to the matter power spectrum as a function of both scale and halo mass. This was motivated by the assumption typically made in halo-based models that all matter resides in spherical haloes of size  $R_{200}$ . To do so, we combined a set of cosmological N-body simulations to calculate the contributions of different spherical overdensity regions, FoF groups and matter outside haloes to the power spectrum.

Our findings can be summarised as follows:

- On scales  $k < 1 h \text{Mpc}^{-1}$ , haloes – defined as spherical regions with an enclosed overdensity of 200 times the mean matter density in the Universe – with masses  $M_{200} \lesssim 10^{9.5} h^{-1} M_{\odot}$ , which are not resolved here, may signif-

icantly contribute to the matter power spectrum. For  $2 < k < 60 h \text{ Mpc}^{-1}$ , our simulations suggest their contribution to be  $< 1\%$ .

- For  $k \gtrsim 2 h \text{ Mpc}^{-1}$ , the minimum mass of haloes that contribute significantly to the matter power spectrum decreases towards smaller scales, with more massive haloes becoming increasingly less important.
- On scales  $k < 2 h \text{ Mpc}^{-1}$ , matter in haloes accounts for less than 85% of the power in our simulations. Its relative contribution increases with increasing Fourier scale, peaking at  $\sim 95\%$  around  $k = 20 h \text{ Mpc}^{-1}$ . On smaller scales, its contribution is roughly constant. When  $R_{200,\text{crit}}$  is used to define haloes instead of the fiducial  $R_{200,\text{mean}}$ , the contribution of haloes to the power spectrum decreases significantly on all scales.
- For  $2 \lesssim k \lesssim 10 h \text{ Mpc}^{-1}$ , haloes below  $\sim 10^{11} h^{-1} M_{\odot}$  provide a negligible contribution to the power spectrum. The dominant contribution on these scales is provided by haloes with masses  $M_{200} \gtrsim 10^{13.5} h^{-1} M_{\odot}$ , even though such haloes account for only  $\sim 13\%$  of the total mass.
- Matter just outside the  $R_{200}$  overdensity regions, but identified as part of FoF groups, provides an important contribution to the power spectrum. Taken together, matter in FoF groups with  $M_{\text{FoF}} > 10^9 h^{-1} M_{\odot}$  accounts for essentially all power for  $3 < k < 100 h \text{ Mpc}^{-1}$ . Switching from  $R_{200}$  to FoF haloes increases the contribution of halo matter on any scale probed here by 5 – 15%.

As we have demonstrated, the halo model assumption that all matter resides in (spherical overdensity) haloes may have significant consequences for the predictions of the matter power spectrum. Specifically, we expect such an approach to overestimate the contribution of haloes to the power on small scales ( $k \gtrsim 1 h \text{ Mpc}^{-1}$ ), mainly because it ignores the contribution of matter just outside  $R_{200}$  to the power spectrum. While defining haloes to be larger regions similar to FoF groups mitigates the small-scale power deficits, the fact that such regions are often non-virialised and typically non-spherical may lead to other problems.

Clearly, the validity of the postulate that the clustering of matter is fully determined by matter in haloes is strongly dependent on the definition of a halo used – but it is hard to say what the “best” definition to use in this context is. For example, while haloes defined through  $R_{200,\text{crit}}$  will be more compact and therefore have a smaller overlap fraction than  $R_{200,\text{mean}}$  or FoF haloes, their contribution to the power spectrum will be smaller for the same minimum halo mass. And while FoF groups seem to contain nearly all mass important for clustering, the fact that they are not completely virialised, may be non-spherical and have boundaries which not necessarily correspond to some fixed overdensity (e.g. More et al., 2011) prohibit their use in traditional halo based models.

Some optimal choice of halo definition may exist, but whatever definition one uses, it remains difficult to say what the contribution to the matter power spec-

trum of haloes below the resolution limit is. Convergence with mass is extremely slow over a large range of scales, and as the results of §5.3.1 show, apparent convergence over a decade in mass is no guarantee, as successive decades in mass may contribute equally. This makes it extremely difficult to give a definitive answer on whether mass outside haloes significantly contributes to the power spectrum, or even if such mass exists: for example, we cannot exclude the possibility that matter outside  $R_{200}$  but inside FoF groups itself consists purely of very small  $R_{200}$  haloes. Therefore, any claims about the role of haloes or the mass contained in them needs to be quoted together with a minimum halo mass in order to have a meaningful interpretation.

## Acknowledgements

---

The authors thank Simon White for useful discussions. The simulations presented here were run on the Cosmology Machine at the Institute for Computational Cosmology in Durham (which is part of the DiRAC Facility jointly funded by STFC, the Large Facilities Capital Fund of BIS, and Durham University) as part of the Virgo Consortium research programme. We also gratefully acknowledge support from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013) / ERC Grant agreement 278594-GasAroundGalaxies.