MNRAS **445,** 3382–3400 (2014) doi:10.1093/mnras/stu1972

Analytic model for the matter power spectrum, its covariance matrix and baryonic effects

Irshad Mohammed[1]★ and Uroš Seljak[2]

[1]*Physik-Institut, University of Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland*
[2]*Department of Physics, Department of Astronomy and Lawrence Berkeley National Laboratory, University of California, Berkeley, CA 94720, USA*

Accepted 2014 September 19. Received 2014 September 19; in original form 2014 June 30

ABSTRACT

We develop a model for the matter power spectrum as the sum of Zeldovich approximation and even powers of k, i.e. $A_0 - A_2k^2 + A_4k^4 - \ldots$, compensated at low k. With terms up to k^4, the model can predict the true power spectrum to a few per cent accuracy up to $k \sim 0.7\,h\,\mathrm{Mpc}^{-1}$, over a wide range of redshifts and models. The A_n coefficients contain information about cosmology, in particular amplitude of fluctuations. We write a simple form of the covariance matrix as a sum of Gaussian part and A_0 variance, which reproduces the simulations remarkably well. In contrast, we show that one needs an N-body simulation volume of more than 1000 $(\mathrm{Gpc}\,h^{-1})^3$ to converge to 1 per cent accuracy on covariance matrix. We investigate the supersample variance effect and show it can be modelled as an additional parameter that can be determined from the data. This allows a determination of σ_8 amplitude to about 0.2 per cent for a survey volume of $1(\mathrm{Gpc}\,h^{-1})^3$, compared to 0.4 per cent otherwise. We explore the sensitivity of these coefficients to baryonic effects using hydrodynamic simulations of van Daalen et al. We find that because of baryons redistributing matter inside haloes all the coefficients A_{2n} for $n > 0$ are strongly affected by baryonic effects, while A_0 remains almost unchanged, a consequence of halo mass conservation. Our results suggest that observations such as weak lensing power spectrum can be effectively marginalized over the baryonic effects, while still preserving the bulk of the cosmological information contained in A_0 and Zeldovich terms.

Key words: neutrinos – methods: analytical – galaxies: haloes – galaxies: statistics – cosmological parameters – large scale structures of Universe.

1 INTRODUCTION

The clustering of dark matter (DM) as a function of scale and redshift contains useful information about many cosmological parameters. For example, clustering as a function of redshift is very sensitive to the dark energy density and its equation of state. Clustering as a function of scale can reveal information about the primordial slope of the power spectrum and matter density, as well as about the presence of massive neutrinos. The best way to measure the DM clustering is via weak lensing (Bartelmann & Schneider 2001; Refregier 2003). In weak lensing light from distant galaxies, called sources, is being deflected by mass distribution along the line of sight, such as that the images are distorted. The primary distortion is shear, which changes ellipticity of the light of the source galaxy. By correlating these ellipticities between the source galaxies one can deduce the clustering strength of the matter along the line of sight. Over the past decade this recognition put weak lensing surveys at the forefront of cosmological probes, with several ground based and space based experiments proposed (Hoekstra et al. 2006; Massey et al. 2007; Fu et al. 2008; Schrabback et al. 2010). The primary statistic is the convergence power spectrum $C_l^{\kappa\kappa}$, which can be expressed as a weighted projection over the matter power spectrum $P(k)$ along the line of sight from the observer to the source. Future surveys will contain sources at many different redshifts, and by combining this information one can minimize the line-of-sight projection and measure a quantity close to the three-dimensional power spectrum, a procedure called weak lensing tomography. In this paper, we will focus on the three-dimensional power spectrum of matter $P(k)$.

★ E-mail: irshad@physik.uzh.ch

The procedure to extract information from the weak lensing measurements is in principle straightforward and, while experimentally challenging, its theoretical underpinnings have been known for a long time. What are the remaining theoretical challenges in this programme? The predictions of the dark matter only (DMO) clustering on small scales, where non-linear effects are important, was one of the uncertainties. For example, the widely used HALOFIT (Smith et al. 2003) is only accurate to 10 per cent, although the revised version (Takahashi et al. 2012) is argued to be 5 per cent accurate for $k < 1.0\,h\,\mathrm{Mpc}^{-1}$. Recent progress in N-body simulations suggests this problem will soon be solved. For example, the *The Coyote Universe* DMO power spectrum emulator (Heitmann et al. 2009, 2010; Lawrence et al. 2010) is accurate to nearly 1 per cent up to $k \sim 1\,h\,\mathrm{Mpc}^{-1}$ for the 38 cosmologies that have been simulated. The emulator provides an output power spectrum for any cosmological model, interpolated from the grid of 38 simulated models, with an error that can be as high as 5 per cent for some cosmological models. It seems likely that the precision will reach the required level in the near future as finer grids of simulations are developed, but it is also clear that by using better ways to interpolate between the models could improve the accuracy.

The second problem are the baryonic effects. Baryons differ from the DM in several aspects. First difference is that hot baryonic gas has pressure, which prevents clustering on small scales. These effects are particularly important inside the DM haloes, where gas temperature is high and pressure effects large. In addition, baryons cool and condense into stars, possibly bringing DM along in the process. However, baryons also form stars, which in turn lead to supernovae that can produce energy outflows. Even more dramatic effects can arise from the active galactic nuclei (AGN), which can also produce massive energy outflows. Recent studies with hydrodynamical simulations (van Daalen et al. 2011) have argued that these AGN feedback models are required to match the observations of X-ray groups and clusters, specially the temperature–luminosity relation in X-rays. The outflowing baryons can also redistribute the DM. Recent work (Semboloni et al. 2011; Semboloni, Hoekstra & Schaye 2013) shows that the baryonic correction in the matter power spectrum can be important above $k \sim 0.3\,h\,\mathrm{Mpc}^{-1}$ and if one does not take account for it, it will bias the cosmological constraints such as dark energy equation of state (Semboloni et al. 2011).

Third theoretical problem that remains unsolved is the issue of reliable covariance matrix for the observed power spectrum and optimal weighting of the data. The full covariance matrix consists of two parts: Gaussian and non-Gaussian. Both scale inversely with the volume of the survey. Gaussian contribution is very large at large scales (low wavemodes k) due to sample variance, i.e. finite number of long wavelength Fourier modes sampled in a finite volume. At higher k, the sampling variance becomes small, and Fisher matrix calculations based on Gaussian variance have predicted that most of the cosmological information in weak lensing comes from small scales. However, the non-Gaussian part becomes important on smaller scales and makes these predictions unreliable. There are two essential contributions to the covariance matrix: one arises from the Poisson fluctuations in the number of haloes relative to the average, and the second arises from the fluctuations on the scale of the survey, which induce curvature type effects that couple to all modes inside the survey (Baldauf et al. 2011; Takada & Hu 2013). For weak lensing applications these contributions become significant for $\sim\ell > 500$ (Yoo & Seljak 2012). So far the predictions have relied either on the halo model (Takada & Hu 2013) or on the simulations (Sato et al. 2009, 2011; Li, Hu & Takada 2014). It has been argued that large numbers of simulations are needed to converge for a single model (Sato et al. 2011; Blot et al. 2014). Without a reliable covariance matrix one cannot optimally combine the different power spectrum estimates, nor can one reliably estimate the errors, as emphasized in recent work (Percival et al. 2014; Taylor & Joachimi 2014).

In this paper, we propose a different approach to the DM power spectrum description that addresses all of the challenges above. We propose a novel form of the halo model for the DM power spectrum (Ma & Fry 2000; Peacock & Smith 2000; Seljak 2000; Cooray & Sheth 2002), in which we split the power spectrum into the quasi-linear 2-halo term, which we take to be the Zeldovich approximation (ZA), and the 1-halo term. Rather than relying on the analytic forms for the 1-halo term as in the original halo model (Seljak 2000; Ma & Fry 2000), we simply expand it into the series of even powers of k and fit each coefficient to the simulations. By doing so we obtain an accurate description of the DM power spectrum up to $k \sim 0.7\,h\,\mathrm{Mpc}^{-1}$. We then investigate the baryonic effects on these coefficients and address the question how to marginalize against these effects. Finally, the resulting solution we propose also simplifies the question of the covariance matrix calculations.

The outline of the paper is as follows: In Section 2, we review some important theoretical background, particularly the halo model (Section 2.1) and ZA (Section 2.2). We postulate the necessary modifications in the 1-halo term in Section 2.3 and calibrate the fitting functions on simulations in Section 3 and showing the comparison with the true matter power spectrum. In Section 4, we discuss the covariance matrix and cosmological information content of our model. We are also discussing supersample variance in Section 4.3. In Section 5, we describe the same method with baryons and the limits to which one can calculate the non-linear matter power spectrum and its full covariance matrix using this methodology. Finally in Section 6, we summarize and discuss the possibility of the future work.

2 THEORETICAL MODEL FOR DM POWER SPECTRUM

2.1 The halo model

There are several approaches to account for clustering of DM and its evolution in the Universe. One of the more successful frameworks is the halo model (McClelland & Silk 1977; Ma & Fry 2000; Peacock & Smith 2000; Seljak 2000; Cooray & Sheth 2002). We will first review the halo model as implemented in previous work before presenting a new version of the halo model that is more accurate. In the halo model approach, all the matter in the Universe is assumed to be in isolated haloes with mass defined by a threshold density as

$$M_{\triangle} = \frac{4}{3}\pi R_{\triangle}^3 \triangle \bar{\rho}_{\mathrm{m}}, \tag{1}$$

where M is the mass of the halo inside the radius $R_\triangle$ and the density of the halo is $\triangle$ times $\bar{\rho}_{\rm m}$, which is the mean matter density of the universe. We use $\triangle = 200$ throughout this paper unless stated otherwise. The power spectrum can be split into two parts:

$$P(k) = P_{\rm 1h}(k) + P_{\rm 2h}(k), \tag{2}$$

where the two terms in right are the 1-halo and 2-halo term, respectively. The 2-halo term gives the correlation between different haloes, also referred as halo–halo term, whereas the 1-halo term describe the correlation between DM particles within the halo, also referred to as Poisson term, and dominates at smaller scales. These two terms are given by

$$P_{\rm 1h}(k) = \int {\rm d}\nu\, f(\nu) \frac{M}{\bar{\rho}} |u(k|M)|^2, \tag{3}$$

$$P_{\rm 2h}(k) = \left[\int {\rm d}\nu\, f(\nu) b(\nu) u(k|M)\right]^2 P_{\rm L}(k), \tag{4}$$

where $P_{\rm L}(k)$ is the linear power spectrum. Throughout this paper, we use publicly available code CAMB (Lewis, Challinor & Lasenby 2000) to compute linear matter power spectrum, unless stated otherwise. We also used publicly available code CHOMP[1] to compute some functions like the halo mass function and density profiles. The Fourier transform of the density profile of the haloes $u(k|M)$ is normalized such that $u(k = 0|M) = 1$,

$$u(k|M) = \frac{4\pi}{M} \int_0^{R_{\rm vir}} {\rm d}r\; r^2\; \rho(r|M)\; \frac{\sin(kr)}{kr}. \tag{5}$$

One can see that upon expanding $\sin(kr)/kr$ only even powers of k will be present, as further developed below. The functions $f(\nu)$ and $b(\nu)$ are the mass function and halo bias, respectively. Both variables ν and M account for the scale and related as

$$\nu(M, z) = \left(\frac{\delta_c}{\sigma(M, z)}\right)^2, \tag{6}$$

where $\delta_c \sim 1.68$,

$$\sigma^2(M, z) = \sigma^2(M) D_+^2(z), \tag{7}$$

$$\sigma^2(M) = \frac{1}{2\pi^2} \int {\rm d}k\; k^2\; P_{\rm L}(k)\; |\bar{W}(kR)|^2, \tag{8}$$

with, $D_+(z)$ as the growth factor and $\bar{W}(x)$ as the Fourier transform of the top-hat function:

$$\bar{W}(x) = 3\frac{\sin(x) - x\cos(x)}{x^3}. \tag{9}$$

2.2 The new 2-halo term: ZA

The halo model is not sufficiently accurate for the 1 per cent precision required from the future surveys. The 2-halo term needs to be modified because in the halo model it is essentially given by the linear theory, and the non-linear effects such as the smearing of baryonic acoustic oscillations (BAO) are ignored. A useful improvement is the ZA (Zel'dovich 1970). In it we assume the particles stream along the initial trajectory, without being perturbed by the non-linear effects. Even though the ZA is in a sense linear, its effects on the density extend beyond linear effects, and ZA can even lead to caustics where the density is infinite. While ZA produces too little power to be a good approximation for the fully non-linear power spectrum, it smears the BAO in the amount that matches the simulations quite well (Taylor 1993; Matsubara 2008). As such it is a useful extension of the linear power spectrum. Here, we will consider ZA approximation for large scales, coupled to the 1-halo term for the small scales.

The Zeldovich power spectrum is given by (see e.g. Schneider & Bartelmann 1995)

$$(2\pi)^3 \delta^D(k) + P(k) = \int {\rm d}^3 q\; {\rm e}^{-{\rm i}\boldsymbol{q}\cdot\boldsymbol{k}} \exp\left[-\frac{1}{2} k_i k_j \mathbf{A}_{ij}(\boldsymbol{q})\right], \tag{10}$$

where

$$A_{ij}(\boldsymbol{q}) = X(q)\delta^K_{ij} + Y(q)\hat{q}_i\hat{q}_j, \tag{11}$$

and

$$X(q) = \int_0^\infty \frac{{\rm d}k}{2\pi^2} P_{\rm L}(k) \left[\frac{2}{3} - 2\frac{j_1(kq)}{kq}\right], \tag{12}$$

[1] http://code.google.com/p/chomp/

$$Y(q) = \int_0^\infty \frac{\mathrm{d}k}{2\pi^2} P_{\rm L}(k) \left[-2 j_0(kq) + 6 \frac{j_1(kq)}{kq} \right]. \tag{13}$$

Here, $P_{\rm L}(k)$ is the linear power spectrum and j_n is the spherical Bessel function of order n.

2.3 The 1-halo term expansion

In this section, we first motivate the 1-halo term expansion into even powers of k. In the next section, we analyse their dependence on the cosmological parameters and compare against the predictions of the halo model.

We begin by writing the ansatz for the 1-halo term,

$$P_{\rm 1h}(k) = (A_0 - A_2 k^2 + A_4 k^4 - \cdots) F(k). \tag{14}$$

To motivate the ansatz and calculate the coefficients A_n, we start with the Fourier transform of the normalized density profile, assuming for now $F(k) = 1$:

$$u(k|M) = \frac{4\pi}{M} \int_0^{R_{\rm vir}} \mathrm{d}r\, r^2\, \rho(r|M)\, \frac{\sin(kr)}{kr}. \tag{15}$$

The halo profile is spherically averaged and assumed to depend only on the mass of the halo. We can model the halo density profile in the NFW form (Navarro, Frenk & White 1997)

$$\rho(r|M) = \frac{\rho_{\rm s}}{(r/R_{\rm s})(1 + r/R_{\rm s})^2}. \tag{16}$$

This model assumes that the profile shape is universal in units of scale radius $R_{\rm s}$, while its characteristic density $\rho_{\rm s}$ at $R_{\rm s}$ or concentration $c = R_{\rm vir}/R_{\rm s}$ may depend on the halo mass M.

The function $\sin(kr)/kr$ can be expand as Taylor series with even powers of kr as

$$u(k|M) = \frac{4\pi}{M} \int_0^{R_{\rm vir}} \mathrm{d}r\, r^2\, \rho(r|M) \left[1 - \frac{k^2 r^2}{3!} + \frac{k^4 r^4}{5!} - \cdots \right]. \tag{17}$$

We can simplify this equation using function $\Im_n$ as

$$u(k|M) = \Im_0 k^0 - \Im_1 k^2 + \Im_2 k^4 - \cdots \equiv (-1)^n \sum_{n=0}^{\infty} \Im_n k^{2n}, \tag{18}$$

and

$$|u(k|M)|^2 = (-1)^{m+n} \sum_{(m,n)} \Im_n k^{2n} \Im_m k^{2m} = (-1)^{m+n} \sum_{(m,n)} \Im_m \Im_n k^{2(m+n)}, \tag{19}$$

where

$$\Im_n = \frac{4\pi}{(2n+1)! M} \int_0^{R_{\rm vir}} \mathrm{d}r\, r^{2(1+n)}\, \rho(r|M). \tag{20}$$

Note that the functions $\Im_n$ are the integrals over the density profiles and some power of r from 0 to $R_{\rm vir}$ and that $\Im_0 = 1$. However, there is nothing obviously special about truncating the integral there, and it can be changed to truncate the density profile at a different $R_{\rm max}$ than $R_{\rm vir}$, for example $2R_{\rm vir}$. This suggests that the halo model has some flexibility in its implementation and is not fully predictive. For this reason, we will just use it as a motivation and will not be doing the actual integrals over the halo profiles.

Next, we insert equation (19) into 1-halo term expression of equation (3) and group the terms in even powers of k,

$$P_{\rm 1h}(k) = \int \mathrm{d}\nu\, f(\nu) \frac{M}{\bar{\rho}} \sum_{(m,n)} \Im_m \Im_n k^{2(m+n)}, \tag{21}$$

$$P_{\rm 1h}(k) = \int \mathrm{d}\nu\, f(\nu) \frac{M}{\bar{\rho}} \left[\Im_0 \Im_0 k^0 - 2\Im_0 \Im_1 k^2 + (\Im_1 \Im_1 + 2\Im_0 \Im_2) k^4 - \cdots \right]. \tag{22}$$

Comparing equations (14) and (22), we obtain the coefficients and their variances as

$$\begin{aligned} A_0 &= \int \mathrm{d}\nu\, f(\nu) \frac{M}{\bar{\rho}}\, \Im_0 \Im_0 \\ A_2 &= \int \mathrm{d}\nu\, f(\nu) \frac{M}{\bar{\rho}}\, 2\Im_0 \Im_1 \\ A_4 &= \int \mathrm{d}\nu\, f(\nu) \frac{M}{\bar{\rho}}\, (\Im_1 \Im_1 + 2\Im_0 \Im_2) \end{aligned} \tag{23}$$

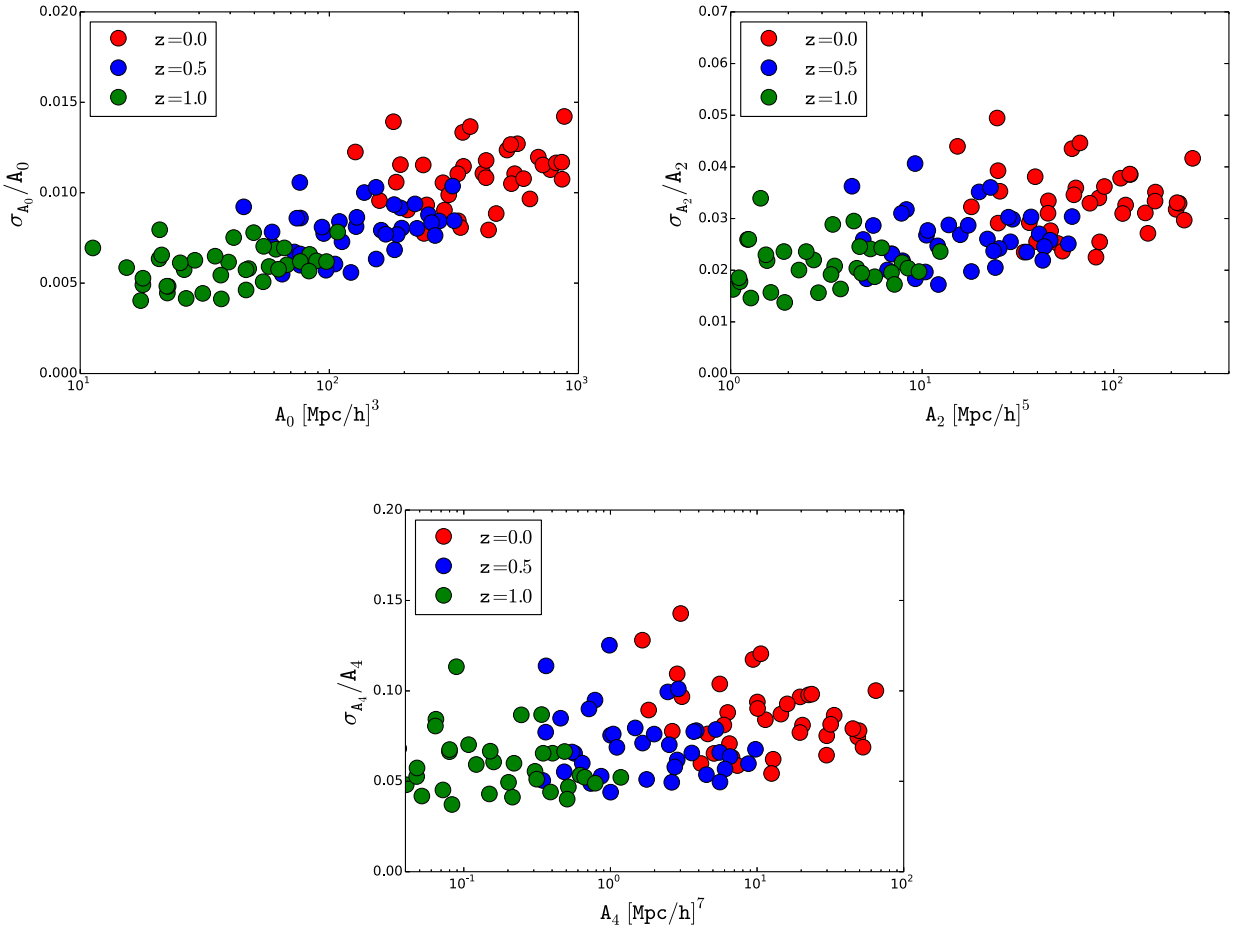

Figure 1. Relative variance $\Delta A_{2n}/A_{2n}$ versus A_n based on our model for A_0, A_2 and A_4 for three different redshift: 0.0 (red), 0.5 (blue) and 1.0 (green). Each circle bullet is one cosmological realization of the 38 cosmic emulator nodes.

with covariance,

$$\mathbf{Cov}(A_i A_j) = \begin{pmatrix} \int d\nu\, g(\nu)\,(\mathfrak{I}_0\mathfrak{I}_0)^2 & \int d\nu\, g(\nu)\,(\mathfrak{I}_0\mathfrak{I}_0)(2\mathfrak{I}_0\mathfrak{I}_1) & \int d\nu\, g(\nu)\,(\mathfrak{I}_0\mathfrak{I}_0)(\mathfrak{I}_1\mathfrak{I}_1+2\mathfrak{I}_0\mathfrak{I}_2) \\ \int d\nu\, g(\nu)\,(\mathfrak{I}_0\mathfrak{I}_0)(2\mathfrak{I}_0\mathfrak{I}_1) & \int d\nu\, g(\nu)\,(2\mathfrak{I}_0\mathfrak{I}_1)^2 & \int d\nu\, g(\nu)\,(2\mathfrak{I}_0\mathfrak{I}_1)(\mathfrak{I}_1\mathfrak{I}_1+2\mathfrak{I}_0\mathfrak{I}_2) \\ \int d\nu\, g(\nu)\,(\mathfrak{I}_0\mathfrak{I}_0)(\mathfrak{I}_1\mathfrak{I}_1+2\mathfrak{I}_0\mathfrak{I}_2) & \int d\nu\, g(\nu)\,(2\mathfrak{I}_0\mathfrak{I}_1)(\mathfrak{I}_1\mathfrak{I}_1+2\mathfrak{I}_0\mathfrak{I}_2) & \int d\nu\, g(\nu)\,(\mathfrak{I}_1\mathfrak{I}_1+2\mathfrak{I}_0\mathfrak{I}_2)^2 \end{pmatrix} \tag{24}$$

where $i, j = 0, 2, 4$ and

$$g(\nu) = \frac{1}{\text{Volume}} f(\nu) \left(\frac{M}{\bar{\rho}}\right)^3. \tag{25}$$

In this paper, we terminate this series after A_4 term. One can always go to higher order terms to get desired accuracy at higher k.

We will present the results of analytic calculations of A_{2n} in the next section. Calculating the variance of each of these coefficients is as straightforward as calculating the coefficient itself, performing the integrals over the halo mass function. We calculate the variance on these terms for a volume of 1 (Gpc h^{-1})^3 for different cosmological models (the 38 models explained in next section) at three different redshifts: 0.0, 0.5 and 1.0. Fig. 1 shows the relative variance of the three coefficients. We find for 1 (Gpc h^{-1})^3 the relative error σ_{A_0}/A_0 varies from 0.5 to 2 per cent, whereas on σ_{A_2}/A_2 and σ_{A_4}/A_4 vary from 1 to 7 per cent and from 2 to 20 per cent, respectively, depending on the cosmology and redshift. We see that the relative error on A_{2n} increases with n: this is a consequence of the fact that terms with higher n receiving a larger contribution from higher mass objects, since the mass scaling of the integrand for A_{2n} in the equations above is $M^{1+2n/3}$, while for the variance it is $M^{3+4n/3}$. Higher mass objects are rarer and their Poisson fluctuations are larger, hence the relative variance is increased. Below we will compute the sensitivity of these parameters to cosmology: we will show that A_0 contains most of the information on the amplitude σ_8. In this paper, we use halo mass function of Tinker et al. (2008).

So far we assumed $F(k) = 1$ without specifying its role. It was pointed out already in the original halo model (Seljak 2000) that the 1-halo term of the halo model fails to account for mass and momentum conservation at low k: the non-linear corrections to the power spectrum have to scale as k^4 or $-k^2P(k)$ at low k, while the leading order of the 1-halo term scales as k^0. At very low k such a term may even dominate over the linear term, which cannot be physical in the context of DM, even though it can happen in the context of galaxies Baldauf et al. (2013).

We will impose this constraint by simply fitting the residuals to the simulations at low k and apply the derived transfer function $F(k)$, which vanished at low k, to the model. We will show that the function $F(k)$ does not strongly depend on the cosmological model and we will thus ignore its dependence on cosmological parameters.

3 CALIBRATING THE MODEL WITH SIMULATIONS

We use cosmic emulator (Heitmann et al. 2009, 2010; Lawrence et al. 2010) to evaluate the power spectra for each of the 38 emulator simulations and assume in each case it gives the true non-linear matter power spectrum. These reference power spectra are correct to nearly 1 per cent up to $k \sim 1\,h\,\mathrm{Mpc}^{-1}$ at 38 different nodes (labelled as 0–37) in cosmological parameter space. This accuracy degrades to 5 per cent when computing the power spectrum away from the nodes. Node 0 cosmology is closest to the *Wilkinson Microwave Anisotropy Probe*-7 (*WMAP*-7) cosmology and we use it as a reference cosmology. We fit these simulation power spectra with our description – quasi-linear Zeldovich term plus modified 1-halo term as a sum of even powers of k, to determine coefficients A_{2n} as a function of cosmology.

To begin with, we fit the even power law (equation 14 with $F(k) = 1$) to the difference between matter power spectrum from emulator $P_{\rm Emu}$ and the Zeldovich term $P_{\rm Zel}$ for all 38 cosmologies and three redshifts: 0.0, 0.5, 1.0 between $k = 0.2$ and $0.8\,h\,\mathrm{Mpc}^{-1}$.

All the coefficients fitted, A_0, A_2 and A_4, are strongly correlated with σ_8, with A_0 having the least scatter. Fig. 2 shows the scaling of these coefficients with $\sigma_8(z)$ and $\sigma_{11.3}(z)$, where the latter was chosen to minimize the scatter in A_0. Each of these coefficients can be approximately fit as a power law irrespective of the redshift and cosmology, with $\sigma_8(z)$ scaling

$$A_0 \propto \sigma_8^{3.9},\ A_2 \propto \sigma_8^{3.0},\ A_4 \propto \sigma_8^{2.2}. \tag{26}$$

It is not straightforward to determine the errors since this is not a formal fit to a set of data points with individual errors. In Fig. 2, we also show results when the slope of A_0 is 4.0: we see this is also a good fit over the range.

Fig. 2 also shows the predictions of the halo model for these coefficients (in black crosses). While the halo mode predicts well A_0 at low redshifts, it fails for higher order coefficients. This can be improved if the virial radius is increased by roughly a factor of 2 at low redshifts, and more than that at higher redshifts (which needs to be taken to power $2n$ to evaluate the effect on A_{2n}), shown as red crosses in Fig. 2. The failure of the halo model to quantitatively predict these coefficients is not surprising: the haloes do not suddenly stop at the virial radius and the halo model has some flexibility in how it is implemented. Our goal here is not to understand the halo model, but to have accurate predictions. For this reason, we will just use the fits of A_{2n} coefficients to simulations in this paper.

The next step is to correct for the scatter around the best fit σ_8. A correlation is noticed between the residual of the coefficients with the effective slope $n_{\rm eff}$. This is shown in Fig. 3. Here, the residual means the difference between the diamond-bullets and best-fitting lines in Fig. 2 and effective slope $n_{\rm eff}$ is calculated as the slope of the linear matter power spectrum at $k \sim 0.2\,h\,\mathrm{Mpc}^{-1}$. The higher order coefficients have larger scatter and stronger correlation between this residual and effective slope. We tested the scalings for few different values of R in σ_R and found minimum scatter for $\sigma_{11.3}$, which can be seen in Figs 2 and 3. By using $\sigma_{11.3}$ instead of σ_8, one can remove the correlation with effective slope for A_0, so no $n_{\rm eff}$ correction is needed for A_0. However, A_2 and A_4 still need to be corrected for this correlation, although the correction is smaller in case of $\sigma_{11.3}$ than σ_8. Hence, the corrected expressions for these coefficients are

$$A_0 = 1529.87\sigma_8^{3.9} \times (1 + [-0.22 n_{\rm eff} - 0.4]),\ \text{ or } A_0 = 2167.39\sigma_{11.3}^{3.9}, \tag{27}$$

$$A_2 = 1299.75\sigma_8^{3.0} \times (1 + [-1.58 n_{\rm eff} - 2.8]),\ \text{ or } A_2 = 1724.16\sigma_{11.3}^{3.0} \times (1 + [-1.39 n_{\rm eff} - 2.5]), \tag{28}$$

$$A_4 = 758.31\sigma_8^{2.2} \times (1 + [-2.27 n_{\rm eff} - 4.2]),\ \text{ or } A_4 = 947.47\sigma_{11.3}^{2.3} \times (1 + [-2.12 n_{\rm eff} - 3.9]). \tag{29}$$

We still need to account for the mass conservation, which forces the 1-halo term to go to 0 at low k. In Fig. 4, we plot the ratio of the difference between $P_{\rm Emu}$ and $P_{\rm Zel}$ with $P_{\rm SimFit}$ which is given by $P_{\rm SimFit} = A_0 - A_2 k^2 + A_4 k^4$, where these coefficients are the best-fitting values to $(P_{\rm Emu} - P_{\rm Zel})$ for all 38 cosmologies (diamond bullets in Fig. 2). The top-left, top-right and bottom-left panel shows the same quantity at three different redshifts: 0.0, 0.5 and 1.0, respectively. All 38 curves in each panel are very close to 1 for k between 0.2 and $0.8\,h\,\mathrm{Mpc}^{-1}$, which is expected as these coefficients are fitted in that range in the first place. Outside this range the scatter increases. We took the average of all these 38 curves at all three redshifts and fit it to a 10th order polynomial, requiring to vanish at low k. The thick solid black line and dashed red curve represents the average and best fit to the average, respectively. It can be seen in bottom-right panel of Fig. 4 that these best fit to the average are very close to node 0 cosmology curve and also very close to each other for different redshifts for $k < 0.8\,h\,\mathrm{Mpc}^{-1}$. We average of these three best-fitting curves, at three different redshifts, to build the function $F(k)$ for the 1 halo term, which we model as

$$F(k) = \sum_{n=0}^{10} a_n k^n, \tag{30}$$

where the coefficients a_n are listed in Table 1. As expected by the mass conservation arguments, and seen in Fig. 4, this correction drops to zero for $k < 0.1\,h\,\mathrm{Mpc}^{-1}$. In principle we should force it to go to 0 as k^2, but we found this caused problems to the fit at higher k: the effects of $F(k)$ are very small in any case and in most instances below 1 per cent, since at low k the Zeldovich term dominates. For this reason, we will assume that this correction is independent of the cosmological model or redshift.

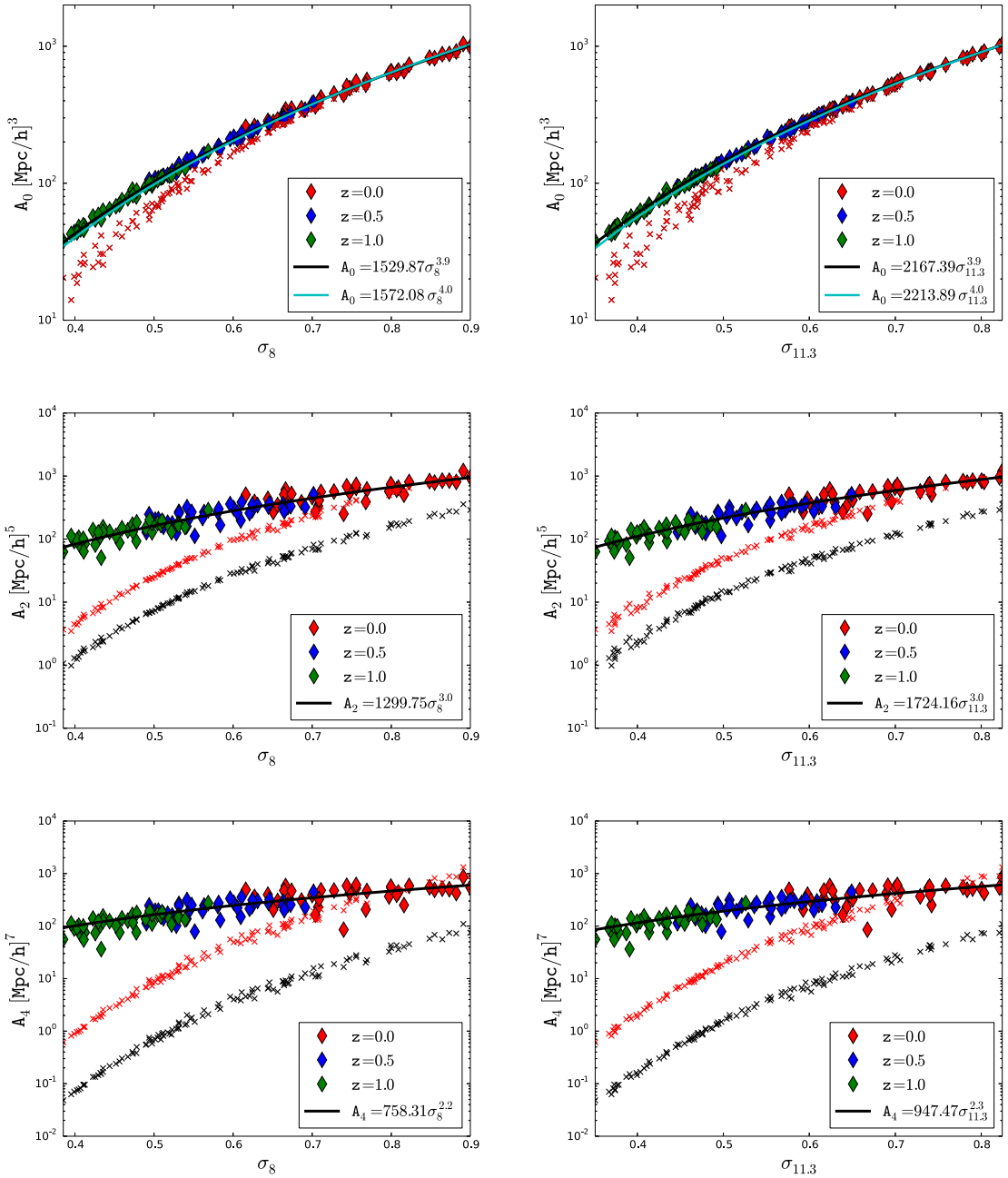

Figure 2. Fitted coefficients A_0, A_2 and A_4 versus σ_8 (left column) and $\sigma_{11.3}$ (right column). We see that $\sigma_{11.3}$ reduces the scatter relative to σ_8 for A_0. Solid black line is the best-fitting power law stated in the legend. The halo model prediction is shown in crosses, using the usual halo concentration parameter $R_s = R_{vir}/c$, with haloes extending to the virial radius R_{vir}, defined at the mean overdensity of 200 (black crosses), and doubling that to $2R_{vir}$ (red crosses). Halo model agrees well with simulations for A_0 at late redshifts, but not for A_2 and A_4 both in terms of amplitude and in terms of σ_8 or $\sigma_{11.3}$ scaling. Extending the halo profile to twice the virial radius improves the agreement.

We combine the above two terms to obtain the matter power spectrum as

$$P(k, z) = P_{\rm Zel}(k, z) + P_{\rm 1h}(k, z) \tag{31}$$

and,

$$P_{\rm 1h}(k, z) = (A_0 - A_2 k^2 + A_4 k^4) F(k), \tag{32}$$

where A_0, A_2 and A_4 are given by equations (27), (28) and (29), respectively, and $F(k)$ is given by equation (30).

We tested this expression against the matter power spectrum from emulator ($P_{\rm Emu}$) on 38 emulator nodes where the stated accuracy is 1 per cent. Fig. 5 shows the deviation of our predictions from the true matter power spectrum of Emulator at three different redshifts: 0.0, 0.5 and 1.0.

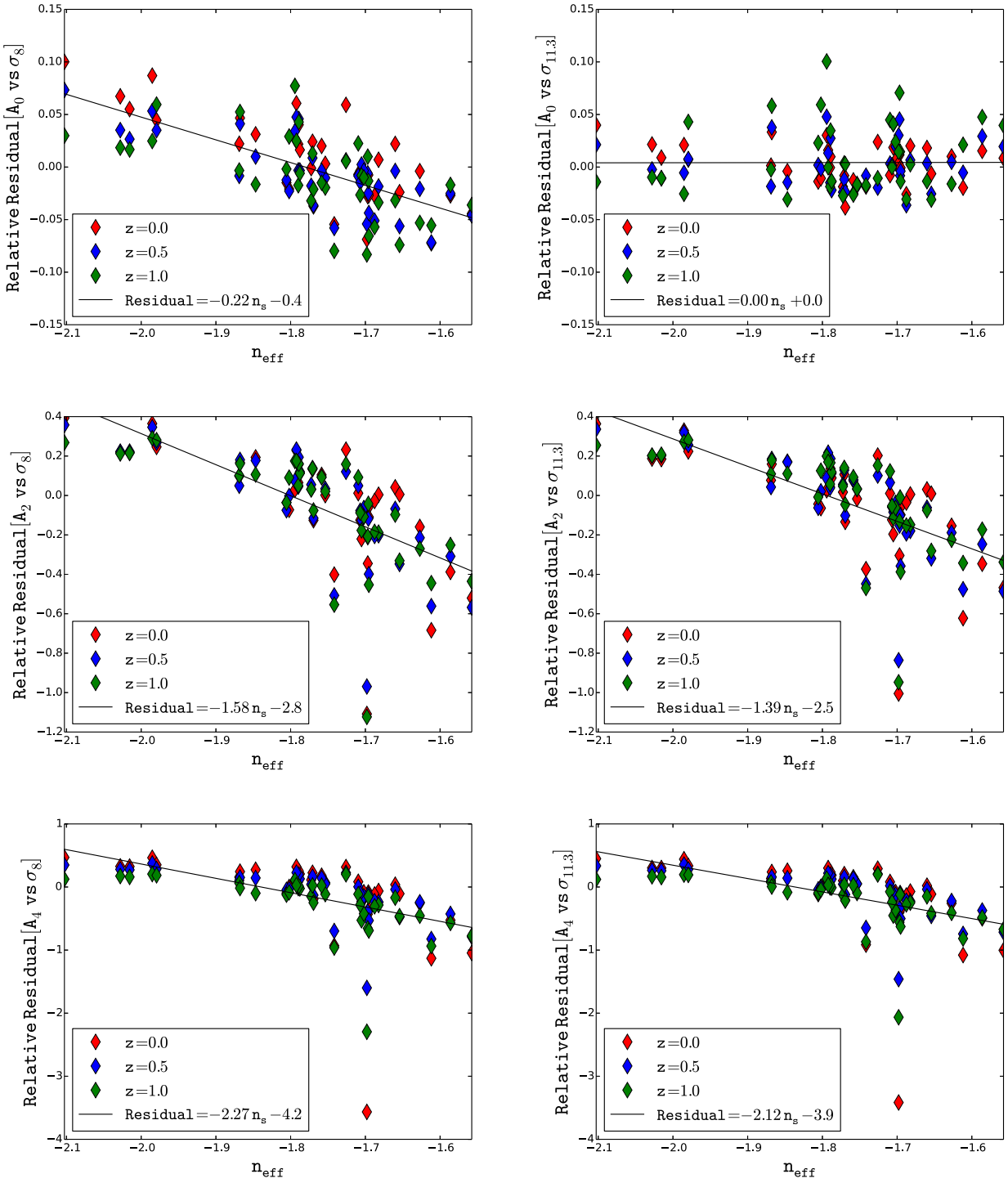

Figure 3. Correlation between effective slope ($n_{\rm eff}$) and residuals after σ_8 (left column) or $\sigma_{11.3}$ (right column) scaling is taken out and their respective best fit. Solid black line is the best linear fit as stated in the legend.

At redshift 0, we can predict the power spectra to a precision of 2–3 per cent up to $k \sim 0.5\,h\,{\rm Mpc}^{-1}$, except in some cosmologies which turn out to be unusual (typically equation of state very different from $w = -1$). At higher redshifts, this accuracy is even better for the same k, as expected since the non-linear effects are smaller. For most of the cosmological models, we can calculate these spectra to 5 per cent up to $k \sim 0.7\,h\,{\rm Mpc}^{-1}$ and much better for lower k.

In Fig. 6, we show the prediction of our model for *WMAP*-7 cosmology (node 0) with all its components plotted separately. Note that the $A_2k^2F(k)$ (in blue) term has a negative contribution while all other components have a positive contribution. The prediction of node 0 power spectrum is correct to about 2 per cent up to $k \sim 0.6\,h\,{\rm Mpc}^{-1}$ increasing to 4 per cent at $k \sim 0.7\,h\,{\rm Mpc}^{-1}$. This can also be seen in Fig. 5 where thick black line shows the ratio of the predicted and true matter power spectrum for node 0.

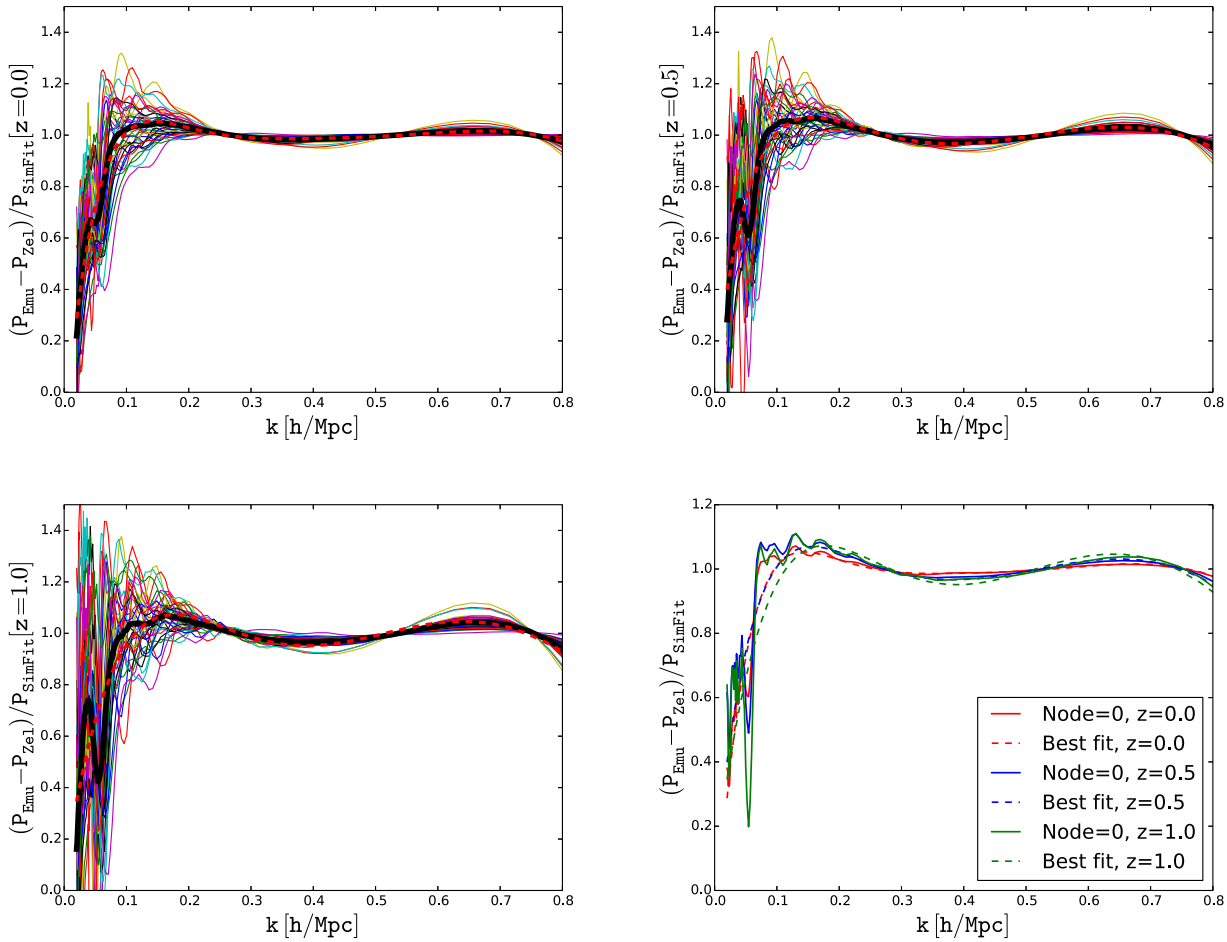

Figure 4. The first three panel (in reading order), shows the ratio of $(P_{\rm Emu} - P_{\rm Zel})$ and $P_{\rm SimFit} = A_0 - A_2k^2 + A_4k^4$, where the coefficients A_0, A_2, A_4 are the best-fitting coefficients to the emulator matter power spectrum for all 38 cosmological models (in different colours) at three different redshifts: 0.0 (top left), 0.5 (top right) and 1.0 (bottom left). Bottom-right panel shows same quantity for node 0 and the best fit to the average (of 38 coloured curves in first three panels) at three different redshifts.

Table 1. Coefficients to calculate the correction function, equation (30). The units of the coefficient a_n is $(\mathrm{Mpc}\,h^{-1})^n$.

a_n	a_0	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9	a_{10}
Value	0.0	21.814	−174.134	747.369	−2006.792	3588.808	−4316.241	3415.525	−1692.839	474.377	−57.228

We also explored how well can this expression predict the changes in the matter power spectrum when cosmological parameters are changed. We take emulator node 0 as the fiducial model and plot the relative difference with other nodes. The first three panel of Fig. 7 (in reading order) shows these derivatives for different components: linear term (in red), Zeldovich term (in green), emulator (in blue) and our predicted model (in thick black). Our predictions are matching very well with that of the true matter power spectrum from emulator, and certainly much better than pure linear theory or pure ZA. Note that we also get very good agreement of BAO smoothing, in contrast to linear theory predictions: this is because we are using ZA which smears out BAO. The broad-band effects of ZA are often anticorrelated with A_0: this is because an increase in σ_8 increases the non-linear smearing caused by the linear streaming of the displacement field, reducing the amplitude of the power spectrum in the ZA, while at the same time the amplitude of the A_0 is increased by the 1-halo term, generated by having more haloes at the same halo mass. The latter effect typically wins: the total power spectrum and the Zeldovich power spectrum are typically, but not always, on the opposite side relative to the linear power spectrum.

Of particular interest is the change in neutrino mass, also shown in Fig. 7. We compare the model predictions to the simulations of Bird, Viel & Haehnelt (2012). We see that our model predicts nearly perfectly the changes in the non-linear power spectrum induced by massive neutrinos. This shows that non-linear effects of massive neutrinos are no different than any other parameter: on large scales they follow linear theory, while on small scales the effects are dominated by the change in A_0. For $\sum m_\nu = 0.15\,\mathrm{eV}$ the change in σ_8 is about 3 per cent and the corresponding change in $A_0 \propto \sigma_8^{3.9}$ is 13 per cent, while ZA goes in the opposite direction, so the linear suppression of 7 per cent at $k \sim 0.2\,h\,\mathrm{Mpc}^{-1}$ is increased to 11 per cent at $k \sim 0.8\,h\,\mathrm{Mpc}^{-1}$, in perfect agreement with simulations.

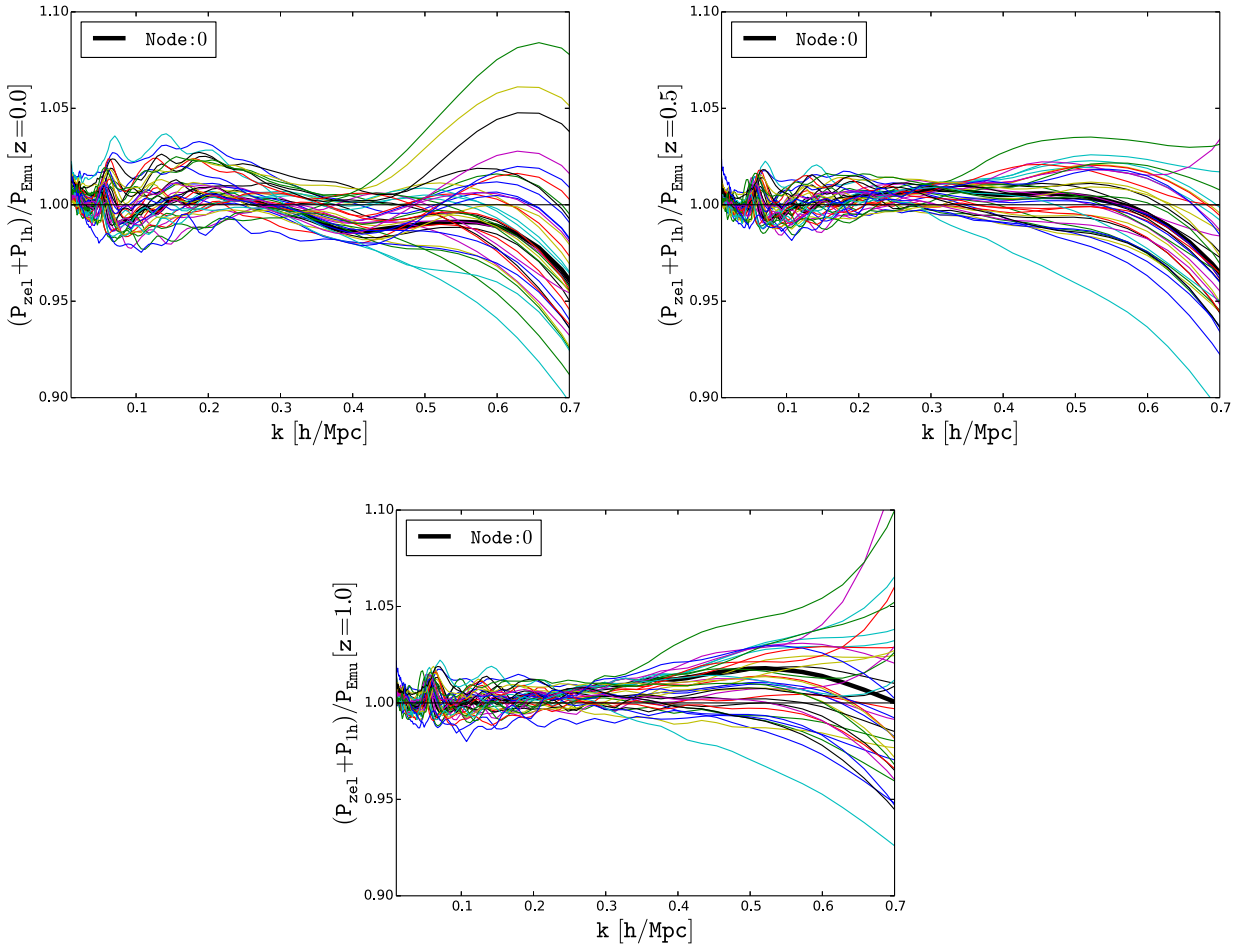

Figure 5. The residuals of our $P(k) = P_{\rm Zel}(k) + P_{\rm 1h}(k)$ expression against simulations for 38 different cosmological models (different colour curves in each panel) for three different redshifts.

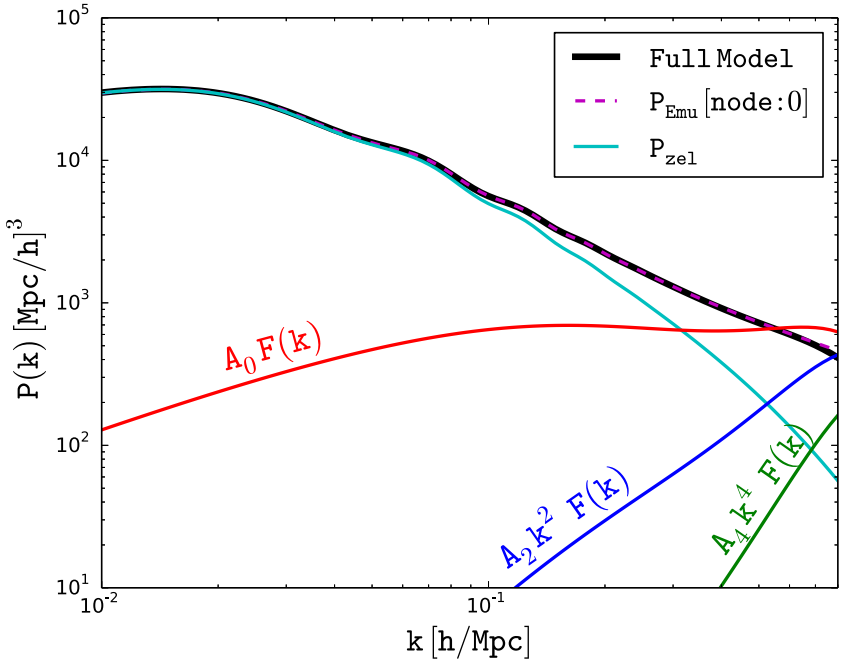

Figure 6. Matter power spectrum for *WMAP*-7 cosmology at redshift 0.0 from simulations (dashed magenta line) versus Zeldovich term (cyan line), $A_0F(k)$ term (red line), $A_2k^2F(k)$ term (blue line), $A_4k^4F(k)$ term (green line). Thick black line is the full predicted model from this work and is nearly indistinguishable from the simulations.

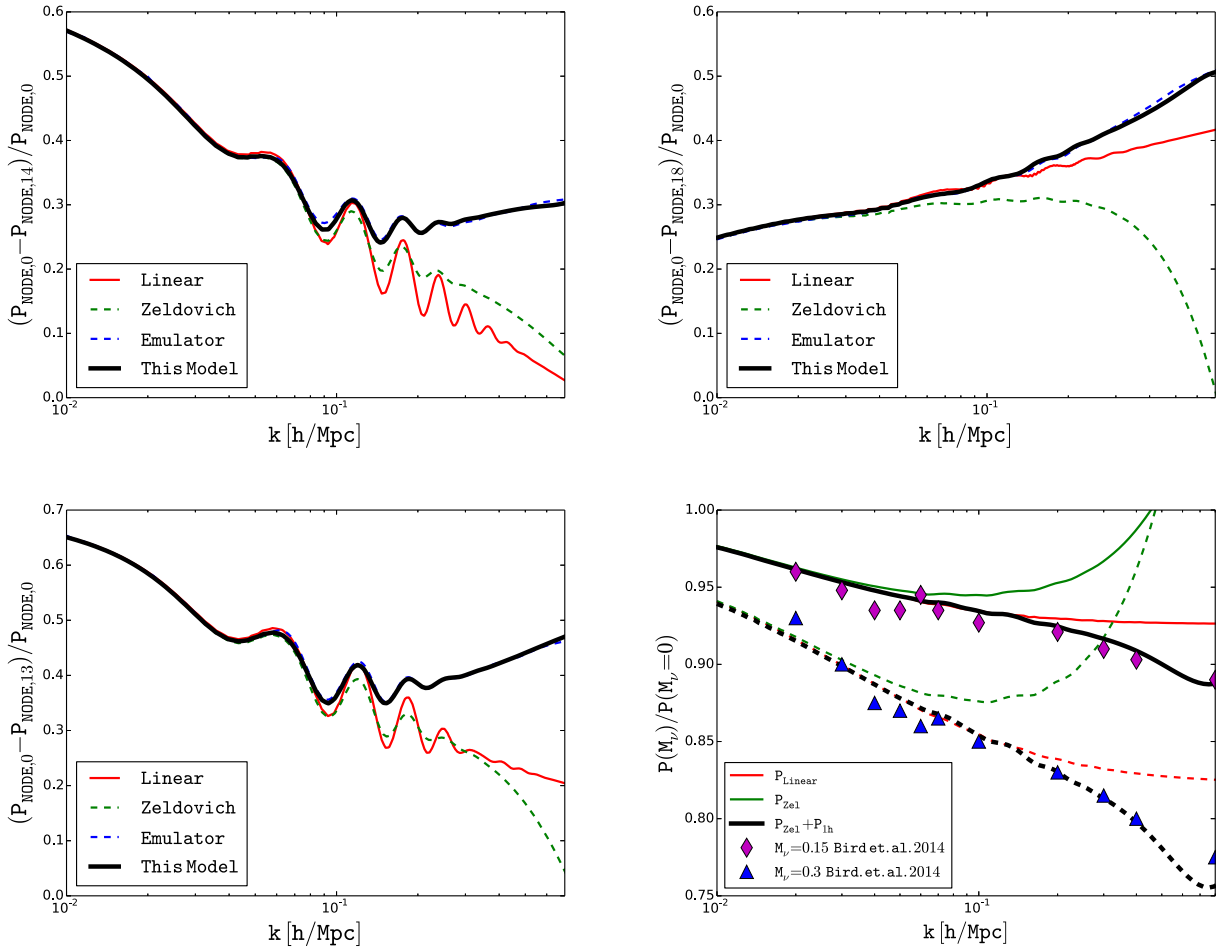

Figure 7. Relative difference in matter power spectrum between node 0 (Emulator) and node 14 (top-left), 18 (top-right), 13 (bottom-left). Showing the same quantity for linear term (in solid red), Zeldovich term (in dashed green), Emulator power spectrum (in dashed blue) and our prediction (in solid black). Bottom-right panel shows the ratio of the matter power spectrum with and without neutrino mass, for $\sum M_\nu$=0.15 (in solid lines) and 0.3 (in dashed lines) from Bird et al. (2012).

4 COVARIANCE MATRIX AND THE COSMOLOGICAL INFORMATION CONTENT OF $P(k)$

We next turn to the issue of covariance matrix. On large scales, low k, the covariance matrix is based on Gaussian approximation. As we move to higher k, the modes become correlated and the covariance matrix becomes non-Gaussian. In our model, the non-Gaussianity comes from two separate terms. First is the non-Gaussian nature of the Zeldovich term and second is the non-Gaussian nature of the 1-halo term. We will not analyse the non-Gaussian covariance matrix in ZA in this paper, as there are currently no analytic calculations available. We also do not have any analytic predictions for the correlation between the Zeldovich part and the 1-halo part. For the 1-halo term, we will focus on A_0 contribution, since as we will argue in next section we should marginalize over the higher order terms anyways. In our initial discussion, we will ignore the supersample variance contribution (Takada & Hu 2013; Li et al. 2014), which will be discussed separately below.

The halo model calculations in Fig. 1 suggest that the relative variance σ_{A_0}/A_0 should be around $0.01\sqrt{(\mathrm{Gpc}\,h^{-1})^3/\mathrm{Volume}}$, depending on the cosmological model and redshift. This calculation is given by

$$\left(\frac{\sigma_{A_0}}{A_0}\right)^2 = \frac{\int f(\nu)\mathrm{d}\nu M^3}{[\int f(\nu)\mathrm{d}\nu M]^2 \bar{\rho} V}, \tag{33}$$

and is determined by the fourth moment of mass integrated over the halo mass function and thus very sensitive to the halo mass function accuracy at the high-mass end. Just as in the case of the halo model predictions for the scalings of A_0, A_2 and A_4, we may not completely trust the halo model predictions. We will write the following ansatz to the covariance matrix $\mathrm{Cov}(P(k_i), P(k_j)) = \langle P(k_i)P(k_j)\rangle - \langle P(k_i)\rangle\langle P(k_j)\rangle$,

$$\mathbf{Cov}(P(k_i), P(k_j)) = P(k_i)P(k_j)\left(\frac{2}{N_i}\delta_{ij} + \left(\frac{\sigma_{A_0}}{A_0}\right)^2\right). \tag{34}$$

Here, N_i is the number of Fourier modes in the ith bin. Our model predicts that the scaling of the variance is

$$\frac{\sigma_{A_0}}{A_0} = \frac{\delta_{A_0}}{[(V/1h^{-1}\mathrm{Gpc})^3]^{1/2}},\ \delta_{A_0} = 0.0079(h^{-1}\mathrm{Gpc})^{3/2}, \tag{35}$$

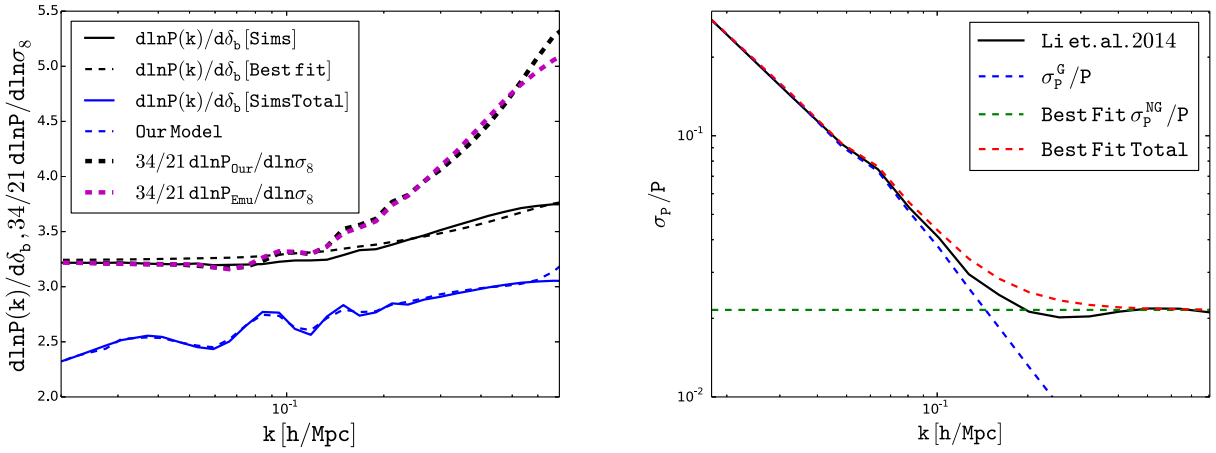

Figure 8. Left: derivative of the matter power spectrum with respect to the change in curvature (i.e. background density) from simulations (blue solid line) and our best-fitting model (blue dashed line). The same, but only for the growth effect without dilation, is shown with the corresponding black solid and black dashed lines. The thick dashed lines show the derivative with respect to the amplitude change, such that it is degenerate with the curvature change at low k: shown are the predictions from simulations (in magenta) and from our model (in black). We see that the degeneracy is broken at higher k even in the absence of the dilation effect. Right: relative variance in the matter power spectrum: $\sqrt{2/N}$ (blue dashed line) where N is the number of modes, best fit σ_p/P (green dashed line), and the total (red dashed line) as the norm of the two terms.

where V is the volume in units of $(h^{-1}\mathrm{Gpc})^3$, and the value of $\delta_{A_0} = 0.0079(h^{-1}\mathrm{Gpc})^{3/2}$ was obtained from a fit of the model to the diagonal part of the covariance matrix derived from *Planck* cosmology simulations in Li et al. (2014), shown in the right-hand panel of Fig. 8. This value is slightly lower than the predictions of the halo model in Fig. 1. Since the predictions are very sensitive to the massive end of the halo mass function, which is not well determined, we should not expect perfect agreement.

It is important to note that the covariance matrix depends on the simulated volume: if the volume changes the covariance matrix will change, and this means that comparing one set of covariance matrix results to another is not trivial. We can simplify the expression if we express the number of modes in terms of a fixed width of the k bin Δk, $N = 4\pi k^2 \Delta k V/(2\pi)^3$. One can see that both the Gaussian sampling variance term and the Poisson term scale with volume, so that

$$\mathbf{Cov}(P(k_i), P(k_j)) = P(k_i)P(k_j)V^{-1}\left(\frac{4\pi^2}{k_i^2\Delta k}\delta_{ij} + \delta_{A_0}^2\right). \tag{36}$$

The relative contribution of diagonal versus off-diagonal terms still depends on the width of the binning in k, but the overall volume scaling is the same.

Now that we have fixed the only free parameter of our model δ_{A_0}, we can apply it to another set of simulations to see the agreement. We have compared it to results in Blot et al. (2014), which used 12 288 boxes of size 656.25 h^{-1} Mpc to derive the full covariance matrix. In Fig. 9 (upper panels), we have compared our model to these simulations for both diagonal and off-diagonal parts of the covariance matrix. We show that the diagonal part of the covariance matrix (left-hand panel of Fig. 9) is an excellent fit, even better than comparison with Li et al. (2014), and this is without any free parameters. In the right-hand panel of Fig. 9, we show the off diagonal terms for six different k values. Our model predicts the off-diagonal correlation coefficients are simply a constant, except at the diagonal where there is an additional Gaussian contribution. Our prediction is in reasonable agreement with these simulations: we are able to reproduce simulation results for both diagonal and off-diagonal terms to within 10–20 per cent, which is remarkable given its simple form and no free parameters.

4.1 Variance of the covariance matrix

An interesting and important question is how big do the simulations need to be to converge. For the convergence of the power spectrum, the answer is given by $\sigma_{A_0}/A_0 = \delta_{A_0}/V^{1/2}$ and we can see that $V = 1\,(h^{-1}\,\mathrm{Gpc})^3$ is sufficient for 1 per cent accuracy. For the covariance matrix, this requirement becomes considerably stricter. One can write an expression for the relative variance of the covariance term as

$$\left(\frac{\sigma(\sigma_{A_0})}{\sigma_{A_0}}\right)^2 = \frac{\int f(\nu)\mathrm{d}\nu M^7}{[(\int f(\nu)\mathrm{d}\nu M^3]^2\bar{\rho}V}, \tag{37}$$

so we can see that this is given by the eighth moment of the mass averaged over the halo mass function. The results of this prediction are shown in Fig. 10. The rms variance for $V = 1\,(h^{-1}\,\mathrm{Gpc})^3$ is now about 10–30 per cent and the corresponding error on the covariance matrix (which goes as a square of σ_{A_0}) is thus 20–70 per cent. There is a large spread in the value because the calculation is so sensitive to the very high mass end of the halo mass function, which is poorly known, so the resulting values should only be taken as indicative and can probably vary by a factor of 2. This is simple to understand: occasionally there will be a large cluster formed which will significantly change the value of A_0, and consequently make its variance change considerably.

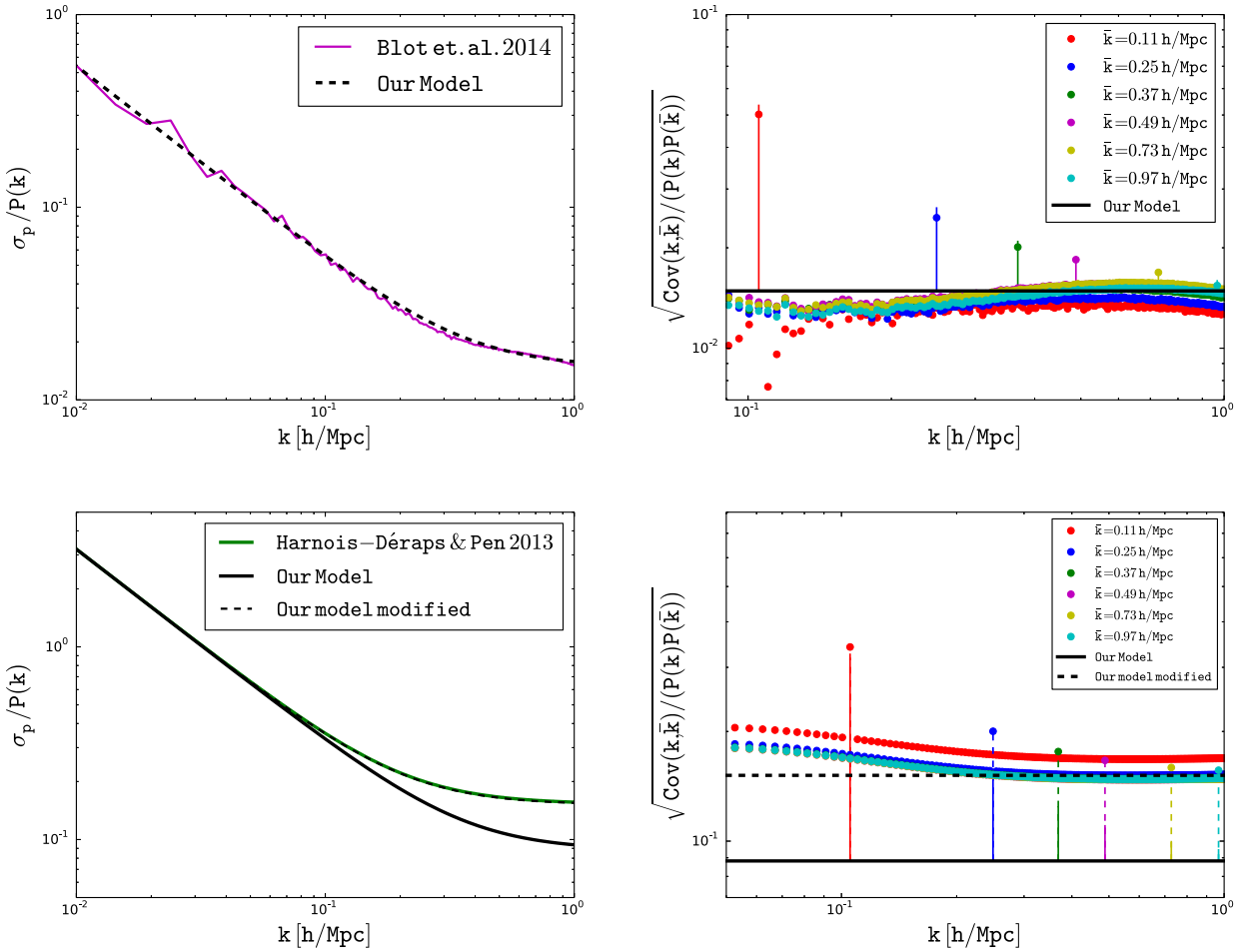

Figure 9. Comparison between our model prediction of covariance matrix with Blot et al. 2014 (upper panels) and Harnois-Déraps & Pen (2012, lower panels) for diagonal (left-hand panels) and off-diagonal elements (right-hand panels). Note that there are no free parameters in the top, while for the bottom panel we show both our best model without a free parameter as well as a modified model where we fit for the value of σ_{A_0}/A_0, which is a valid procedure for these simulations, as discussed in the text. Our covariance matrix model (equation 34) is very simple, yet it is able to reproduce the full covariance matrix from simulations to within 10–20 per cent.

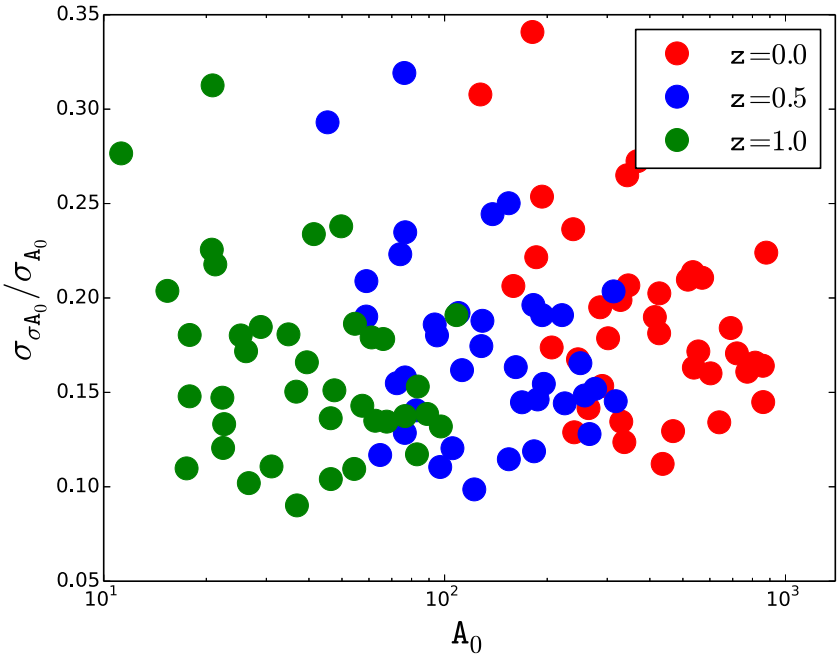

Figure 10. Relative variance $\sigma(\sigma_{A_0})/\sigma_{A_0}$ versus A_0 based on our model for three different redshifts: 0.0 (red), 0.5 (blue) and 1.0 (green). Each bullet is one cosmological realization of the 38 cosmic emulator nodes.

As an example, when we compare our model predictions of the covariance matrix to Harnois-Déraps & Pen (2012), we find that the agreement is not very good, in that our model predicts lower covariance matrix than measured, and the predicted value of σ_{A_0}/A_0 is about 40 per cent below the required for fit the simulations. However, Harnois-Déraps & Pen (2012) used a total simulated volume of 1.6 $(h^{-1}\,\mathrm{Gpc})^3$, suggesting that the value of σ_{A_0}/A_0 has only been determined to about 10–25 per cent. If we let the value of σ_{A_0}/A_0 to be free, we again find a remarkable agreement with the simulations.

To converge on the covariance matrix at 1 per cent one needs a simulated volume to be of the order of 500–5000 $(h^{-1}\,\mathrm{Gpc})^3$. This is an enormous volume: it explains why in recent work of Blot et al. (2014) they needed to simulate 12 288 simulations with a total volume of 3350 $(h^{-1}\,\mathrm{Gpc})^3$ to converge.

4.2 Information content

We can now combine the variance of A_0 with its scaling with σ_8, $A_0 \propto \sigma_8^{3.9}$, to derive the cosmology information content of the A_0 term,

$$\frac{\sigma_{\sigma_8}}{\sigma_8} = \frac{\sigma_{A_0}}{3.9A_0} = 0.002\sqrt{(h^{-1}\,\mathrm{Gpc})^3/\mathrm{Volume}}. \tag{38}$$

This is a remarkably small number, which suggests that much of the cosmological information on the rate of growth of structure, and consequently on the Figure of Merit for dark energy equation of state (Mortonson, Huterer & Hu 2010), resides in this term. To achieve a comparable precision on linear scales one would need about 5×10^5 modes, which for 1 $(h^{-1}\,\mathrm{Gpc})^3$ volume would correspond to $k_{\rm max} = 0.31\,h\,\mathrm{Mpc}^{-1}$. This is already well into the non-linear regime for $z < 1$ implying that we do not have this number of linear modes available, so the bulk of the cosmological information on the amplitude comes from A_0 term. However, since A_0 is mostly sensitive to amplitude (best correlation is with $\sigma_{11.3}$) and nothing else, this also suggests that information on other parameters that depend on the shape of $P(k)$ and not its amplitude will be less well determined.

While we do not have reliable variance predictions for A_2 and A_4 from simulations, Fig. 1 suggests that A_2 has variance three times larger than A_0 and A_4 has variance another three times larger than A_2. This is mostly caused by the fact that Poisson fluctuations get larger for higher order coefficients because of their mass weighting: for example, A_2 weighting is $M^{5/3}$ as opposed to M for A_0, giving more weight to higher mass haloes, which are rarer and therefore have larger Poisson fluctuations. This combined with less steep scaling of A_2 and A_4 with σ_8 compared to A_0 (equations 27, 28 and 29), suggests that there is little additional information in these two coefficients. Another argument for why information in A_2 and A_4 should be ignored, based on baryonic effects, will be presented below.

4.3 Supersample variance

Supersample variance (Hamilton, Rimes & Scoccimarro 2006) arises from the very long wavelength density modes that appear as constant on the scale of the survey. These can be viewed as a change of curvature inside the observed volume (Baldauf et al. 2011), and this couples to all the short wavelength modes. On large scales the effect can mimic a change in the amplitude of fluctuations, together with a rescaling of the length (Sherwin & Zaldarriaga 2012):

$$\delta \ln P(k) = \left(\frac{47}{21} - \frac{1}{3}\frac{\mathrm{d}\ln P}{\mathrm{d}\ln k}\right)\delta_b = \left(\frac{68}{21} - \frac{1}{3}\frac{\mathrm{d}\ln(k^3P)}{\mathrm{d}\ln k}\right)\delta_b, \tag{39}$$

where δ_b is the density perturbation on the scale of the survey volume. The first term is the effect of the curvature on the growth of small-scale modes, while the second term is the dilation due to the presence of local curvature. It is important to recognize that on large scales the growth effect is degenerate with a $(34/21)\delta_b$ change of amplitude σ_8, while the dilation effect of $-\delta_b/3$ is degenerate with a change in scale, i.e. with a change in the angular diameter distance that can arise from a change in cosmological parameters. We will assume that the change in scale cannot be used as an indicator of the supersample variance because of its degeneracy with these other parameters, so we will only focus on the change in growth rate. The rms fluctuations of $1(\mathrm{Gpc}\,h^{-1})^3$ volume are about 0.4 per cent (Takada & Hu 2013), which together with the 34/21 factor implies that at low k one cannot determine σ_8 to better than 0.6 per cent in the linear regime, which is a factor of 3 larger error than the error on σ_8 without the supersample variance in equation (38). It is therefore clear that without addressing this issue the supersample variance dominates the errors.

On smaller scales we expect the non-linear effects are no longer degenerate with a change in σ_8. Physically the reason for difference is in the curvature nature of the supersample variance: curvature effects grow with the growth rate, i.e. the growth of short wavelength mode $\delta_{\rm s}$ due to the coupling to the long wavelength mode scales as $\delta_{\rm s}(z)[1 + 34D(z)\delta_{b0}/21]$, where $D(z)$ is the linear growth rate and δ_{b0} is the long wavelength mode today, and thus this coupling only matters at low redshifts since $D(z) \ll 1$ for $z \gg 1$. This is different from a simple change in overall amplitude $\delta_{\rm s}(z)(1 + \delta\sigma_8)$, which has no redshift dependence.

To understand this more quantitatively we can compute the logarithmic derivative of A_0 (equation 23) with respect to the two parameters in the context of the universal halo mass function $f(\nu)$, where ν is given by equation 6 (Slosar et al. 2008). The Lagrangian bias is defined as $b_{\rm L} = \bar{n}^{-1}\partial n/\partial\delta_b$, which can be rewritten using $\nu = (\delta_c - \delta_b)^2/\sigma^2$ as $b_{\rm L} = (-2\nu/\delta_c)\partial\ln[\nu f(\nu)]/\partial\nu$. In addition we also have the mean

density increased by δ_b inside the patch. We are still dividing the density with the global mean density, so $\bar{\rho}$ does not change. Using this we find

$$\frac{\mathrm{d}\ln A_0}{\mathrm{d}\delta_b} = \frac{\int (1+b_{\rm L}(\nu))\nu f(\nu) M \mathrm{d}\ln\nu}{\int \nu f(\nu) M \mathrm{d}\ln\nu} = \langle (1+b_{\rm L})\rangle. \tag{40}$$

So the logarithmic slope of A_0 with respect to a long wavelength modulation is given by the appropriate average of the Eulerian bias $b_{\rm E} = 1 + b_{\rm L}$.

If instead one looks at the logarithmic growth of the amplitude with respect to amplitude $\sigma_8 \propto \sigma(M)$, then $\mathrm{d}\nu/\mathrm{dln}\,\sigma_8 = -2\nu$, and so

$$\frac{\mathrm{d}\ln A_0}{\mathrm{d}\ln\sigma_8} = \delta_c \frac{\int b_L(\nu)\nu f(\nu) M \mathrm{d}\ln\nu}{\int \nu f(\nu) M \mathrm{d}\ln\nu} = \delta_c \langle b_{\rm L}\rangle. \tag{41}$$

Since $\mathrm{dln}\,A_0/\mathrm{dln}\,\sigma_8 = 3.9$ we find $\mathrm{dln}\,A_0/\mathrm{d}\delta_b = 3.9/1.68 + 1 = 3.3$. The response to the long wavelength mode has thus a lower logarithmic slope of growth relative to σ_8 and is not much larger than the linear regime value 68/21. This should be contrasted against the response to the amplitude change, which goes from σ_8^2 in the linear regime to $\sigma_8^{3.9}$ in the non-linear regime. Note that this calculation is valid if the density is divided by the global background density, as appropriate for weak lensing observations, which are sensitive to the total density. Whenever the density perturbation is defined using local mean density these numbers should be reduced by 2.

Numerical results are shown in the left-hand panel of Fig. 8, where we show the non-linear response to δ_b from simulations of Li et al. (2014), and the corresponding response to a change in σ_8 that mimics δ_b at low k. We can still model a change in δ_b as a quasi-linear term and $(A_0 - A_2k^2 + A_4k^4)F(k)$. For the quasi-linear term, we adopt simply the ZA model multiplied with the corresponding linear factor of $68/21\delta_b$, and we fit for the other three parameters. The result is shown in Fig. 8 and provides a reasonable fit to the simulations. Note that we show results with and without $\mathrm{dln}\,P/\mathrm{dln}\,k$ term, against simulations with and without it (Li et al. 2014). We find that for $\delta_b = 0.02$ A_0 has changed by 7.4 per cent, while the quasi-linear term has changed by 6.4 per cent, so that $\mathrm{dln}\,P/\mathrm{dln}\,\delta_b = 3.2$ at low k and 3.7 around $k \sim 0.5\,h\,\mathrm{Mpc}^{-1}$ where A_0 dominates. This is in reasonable agreement with the analytic estimate of 3.3. For the σ_8 scaling a change of 6.4 per cent in the linear term corresponds to 13 per cent change in A_0. The contrast between the two effects is shown in Fig. 8. The supersample variance is thus not degenerate with σ_8, so if one can determine both the quasi-linear term and A_0 term with sufficient accuracy, one can break the degeneracy between the two effects.

How well can one determine σ_8 in the presence of supersample variance? If we only have information from A_0, then the analysis above suggests that one can determine σ_8 to about $(3.7/3.9)0.4 \sim 0.38$ per cent in 1 $(h^{-1}\,\mathrm{Gpc})^3$ volume, about a factor of 2 worse than without the supersample variance. If we have information both from linear regime and from A_0 dominated regime, then we can break the degeneracy between the supersample variance and σ_8. The extent to which this can be achieved depends on how well we can measure the amplitude in the linear regime: to reach 0.4 per cent accuracy we would need to measure all the modes up to $k \sim 0.2\,h\,\mathrm{Mpc}^{-1}$ in a 1 $(h^{-1}\,\mathrm{Gpc})^3$ volume, which seems possible to achieve. Moreover, we note that a change in curvature cannot be modelled well with just a change in linear term and A_0, higher order terms also change significantly. Even though we argue below that these effects are degenerate with baryonic effects, this degeneracy may be broken in this situation given how different these effects are and given that there is a lot of information present at high k. In summary, the amplitude of fluctuations in a 1 $(h^{-1}\,\mathrm{Gpc})^3$ volume can be determined to an accuracy of 0.4 per cent if the supersample variance cannot be determined, which can be reduced by a factor of 2 if the degeneracy between the supersample variance and σ_8 amplitude can be broken.

Instead of including the supersample variance effect in the covariance matrix, one can include it as an additional curvature parameter that one can marginalize over. The parameter is δ_b and its prior should be a Gaussian with a zero mean and rms variance σ_V determined by the survey window (see Takada & Hu 2013; Takada & Spergel 2014 for predictions for simple survey geometries). The response of the power spectrum to the long wavelength δ_b parameter should be

$$\delta P = \left(\frac{47}{21}P_{\rm Zel} - \frac{1}{3}\frac{\mathrm{d}P}{\mathrm{d}\ln k} + \left[3.7A_0 - 3A_2k^2 + 2.5A_4k^4\right]F(k)\right)\delta_b, \tag{42}$$

where $P_{\rm Zel}$, A_0, A_2 and A_4 are the values of the fiducial model around which we are exploring the supersample variance effect. For example, in a Markov-Chain Monte-Carlo (MCMC) chain this would be the model one is testing at a given chain position. We found that the fit to the simulations must include A_2 and A_4 terms and that the fit is only valid to $k \sim 0.7\,h\,\mathrm{Mpc}^{-1}$. Note that the change of A_2 and A_4 relative to A_0 is similar to that of amplitude change in equation (26).

5 EFFECTS OF BARYONS

Baryonic effects inside the DM haloes change the matter power spectrum relative to the DM alone and these effects must be incorporated into the analysis, otherwise they can lead to substantial bias in the cosmological parameter estimation (Semboloni et al. 2011, 2013). Baryonic effects can come in different forms. First is simply the fact that gas distribution inside DM haloes is distributed differently than the DM, because gas is hot and has significant pressure. As a result, gas has a core at the centre of the cluster, leading to reduced clustering strength on small scales. Second effect is baryon cooling, which causes gas to cool and condense into galaxies at the DM halo centres. This leads to an enhancement of the clustering relative to pure DM case. Baryons can also be pushed out of the halo centres by processes such as supernova and AGN feedback, which can in some cases push the gas quite far out. Furthermore, in all of these examples DM may also be redistributed

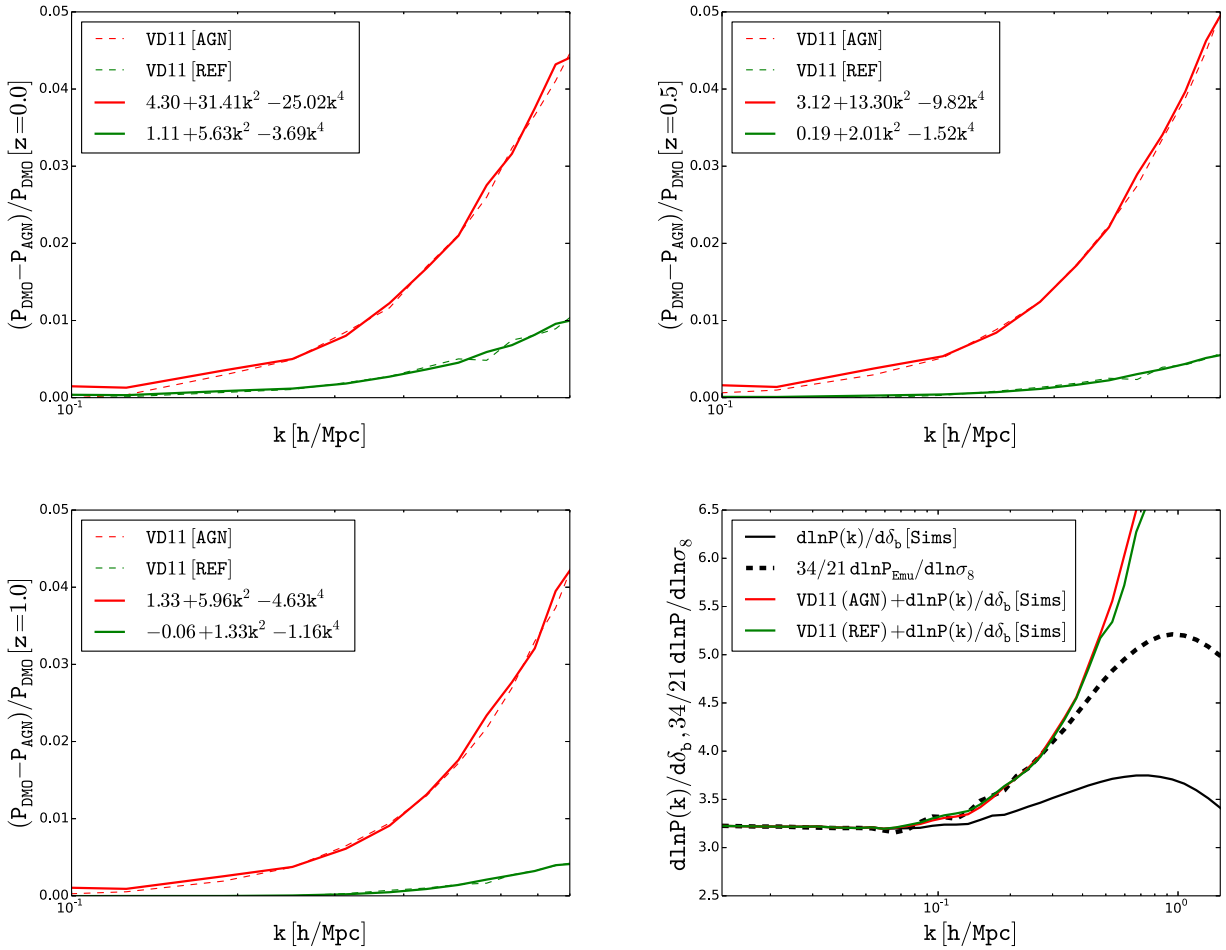

Figure 11. The first three panel (in reading order) are the relative difference between DMO model and AGN (red dashed line) and REF (green dashed line) from van Daalen et al. 2011 (VD11) at redshift 0.0 (top-left), 0.5 (top right) and 1.0 (bottom left). REF model contains the baryonic physics without any AGN feedback model. Solid lines (red and green) are the corresponding best fit $\delta A_0 - \delta A_2 k^2 + \delta A_4 k^4$ as explained in Section 5. Bottom-right panel shows the derivative of the matter power spectrum with respect to a change in background density (2 per cent) in solid-black from Li et al. 2014, and with respect to a change in amplitude using prediction from emulator (thick dashed black line).

as a consequence of the baryons either condensing on to the halo centres or being pushed out. For example, for baryonic cooling on to a galactic disc this process is known as adiabatic contraction (Blumenthal et al. 1984).

From the halo model point of view, the main effect of the baryons is the redistribution of the gas, and possibly DM, inside the haloes. This can be qualitatively described as the change in the scale radius R_s. The total mass of the halo M is unchanged, since these baryonic processes do not push the gas or the DM far out of the virial radius of the halo such that the halo mass would be affected. As a consequence, we expect that A_0 parameter is essentially unchanged, while A_2, A_4 etc. will change during the baryonic redistribution of matter.

To investigate this further we used simulation based matter power spectra from van Daalen et al. (2011) to compute the effects of baryons on the coefficients A_0, A_2 and A_4. In particular, we use the DM only and the supernova and AGN feedback models, corresponding to hydrodynamical simulations with supernova or AGN feedback model. It was argued that the latter is needed to reproduce cluster observations such as X-ray luminosity–temperature relation (McCarthy et al. 2010). We use the AGN model as the main model since it provides the largest effects, but we also explore reference supernova feedback model from van Daalen et al. (2011). Baryon corrections to the matter power spectrum from AGN feedback model exceed 1 per cent level for $k > 0.3\,h\,\mathrm{Mpc}^{-1}$ (van Daalen et al. 2011). We use the results at three different redshifts: 0.0, 0.5 and 1.0. In Fig. 11, we try to fit the difference between the pure DM and AGN model, $P_{\rm DMO} - P_{\rm AGN}$, or reference supernova feedback model $P_{\rm DMO} - P_{\rm REF}$, with the model $\delta A_0 + \delta A_2 k^2 + \delta A_4 k^4$, to estimate the changes in these coefficients due to baryons at each redshift (note that since the changes are only important at high k we can set $F(k) = 1$). We fit these models over the k range between 0.2 and $0.8\,h\,\mathrm{Mpc}^{-1}$. Fig. 11 shows the best-fitting models, which are a good fit to the simulations over this range. We also calculated these coefficients for the cosmology assumed in this paper using the results from Fig. 2.

We find that for the AGN model the relative change in A_0 is about 0.5–1 per cent, depending on the redshift, whereas the changes in A_2 and A_4 are about 4–7 and 4–8 per cent, respectively. If we assume no change in A_0 the fit is a bit worse and the change in A_2 and A_4 is larger. This confirms that the coefficient A_0 is quite indifferent to baryonic effects, while A_2 and A_4 are significantly more contaminated. The change is positive. This is expected since AGN feedback expands the gas and makes the scale radius R_s larger. It is less obvious why A_0 should increase

when gas is being pushed outwards, but the effect on A_0 is small and it could also be driven by the numerical fitting procedure. If we assume that the baryonic uncertainty is at the level suggested by these AGN models, then using equation (24) the corresponding uncertainty on σ_8 will be 0.5–1 per cent/3.9 $\sim$ 0.1–0.2 per cent from A_0, and about an order of magnitude larger from A_2 and A_4. Given that the difference between AGN and DM models is probably an overestimate of the error associated with the baryonic effects our analysis suggests that these effects can be effectively marginalized over without any loss of cosmological information from A_0. We also note that other baryonic feedback models from van Daalen et al. (2011), such as the reference model, while giving a lower amplitude of the effect, have very similar k-dependence, as can be seen from Fig. 11.

Above we argued that supersample variance effect should not be treated as a variance but as a separate parameter that can be determined from the data. Using linear theory and A_0 may not contain enough information to break the degeneracy between the amplitude σ_8 and supersample variance. Using higher k information may be more promising, since the two effects also have very distinctive signatures on A_2, A_4 etc. Since our model expansion to A_4 only works to $k \sim 0.7\,h\,\mathrm{Mpc}^{-1}$, we explore this question numerically. In Fig. 11 bottom-right panel, we plot the supersample effect and amplitude effect such that they are degenerate at low k, while also adding the baryonic effect such that it is degenerate with change of amplitude up to $k \sim 0.3\,h\,\mathrm{Mpc}^{-1}$. We see from the Fig. 11 that above $k \sim 0.3\,h\,\mathrm{Mpc}^{-1}$ the degeneracy is broken: the effects of σ_8 and δ_b are smaller compared to the effect of AGN feedback, which continuous to increase with k, because gas is being pushed out on small scales, suppressing all small-scale clustering. While this analysis is only restricted to a specific form of baryonic effects and is less robust than the other analyses in this paper, it suggests that one may be able to break the degeneracy between the baryonic effects, cosmological parameters such as amplitude, and supersample variance, using high k information.

6 DISCUSSION AND CONCLUSIONS

In this work, we propose a model of the matter power spectrum using the ZA power spectrum as the 2-halo term and even powers of k expansion of the 1-halo term, compensated on large scales to satisfy mass conservation, with coefficients calibrated on simulations. The leading order 1-halo term is k^0 term amplitude A_0, which in the halo model can be determined as a mass-dependent integral over the halo mass function. Simulations predict $A_0 \propto \sigma_8^{3.9}$, and the halo model is only able to reproduce this at low redshifts. The amplitude of A_0 is related to the cluster abundance method, where one counts clusters above a given mass, which also depends on the halo mass function and has a similarly steep dependence on σ_8. It is also related to Sunyaev-Zeldovich (SZ) power spectrum scaling, which is dominated by the 1-halo term and scales as σ_8^7 (Komatsu & Seljak 2002), because the SZ signal from individual clusters scales as $M^{5/3}$ rather than halo mass M and it is a projection over line of sight, leading to a steeper dependence on σ_8. Our analysis thus explicitly connects the cluster abundance method to the amplitude of the leading non-linear correction to the matter power spectrum, and shows the two use similar information. As a consequence, these two methods cannot be combined independently if the dominant errors are Poisson or large-scale structure fluctuations.

Using the first three coefficients of expansion, we accurately predict variations of basic cosmological parameters up to $k \sim 0.7\,h\,\mathrm{Mpc}^{-1}$, including amplitude σ_8, matter density Ω_m, Hubble parameter H_0, primordial slope n_s, equation of state w_0 and even neutrino mass $\sum m_\nu$. In all cases our model predicts well the BAO smoothing, a consequence of using the ZA rather than linear theory for the 2-halo term.

We present a very simple model for the covariance matrix of matter power spectrum (equation 34). We stress that the covariance matrix depends on the simulated volume both in linear and non-linear regimes, so a direct comparison between covariance matrices from different simulations needs to account for this. In this model the large-scale variance is dominated by the sampling variance, while on small scales where A_0 dominates the dominant term is the Poisson sampling of the haloes. Using the halo mass function of Tinker et al. (2008) to predict the latter gives about 20–30 per cent higher value than fitting with simulations of Li et al. (2014), which we consider good agreement given the inaccurate nature of halo mass function fits in the high mass regime. Using this value we show that our model gives remarkable agreement with the simulations of Blot et al. (2014), where 12 288 simulations of $656\,h^{-1}$ Mpc box size were run to construct a covariance matrix. We use our Poisson model to compute the convergence rate of the covariance matrix and find that simulated volumes of 500–5000 $(h^{-1}\,\mathrm{Gpc})^3$ are needed to converge at 1 per cent level. This explains why our model without any free parameters does not reproduce covariance matrix Harnois-Déraps & Pen (2012), because the total volume used in Harnois-Déraps & Pen (2012) was only 1.6 $(h^{-1}\,\mathrm{Gpc})^3$, and has thus not converged with high enough accuracy. Changing the parameter σ_{A_0}/A_0 from the predicted 0.09 to 0.15 we obtain perfect agreement.

Using this model we argue that most of the cosmological information about the amplitude is in A_0, which can determine the amplitude σ_8 to 0.2 per cent within 1 $(h^{-1}\,\mathrm{Gpc})^3$ volume. The higher order coefficients A_2, A_4 etc. are less sensitive to σ_8 and have a larger variance. We discuss the supersample variance and argue that due to its origin as a curvature effect it differs from the amplitude rescaling and so it should be treated as a separate cosmological parameter with a prior given by the rms variance on the scale of the survey volume. If its degeneracy with the amplitude is not broken, then it approximately doubles the errors, so that σ_8 can be determined to 0.4 per cent within 1 $(h^{-1}\,\mathrm{Gpc})^3$ volume. Note that both of these errors are a lot smaller than the currently available constraints, which at best are at 4 per cent (Kilbinger et al. 2013): observational and modelling errors dominate the error budget at the moment, but future data sets may be able to reach the levels where supersampling variance or Poisson error will dominate (Yoo & Seljak 2012).

We also investigate the baryonic effects on the matter power spectrum. We argue that these should not change A_0 much because of the mass conservation. Indeed, comparison of our model to simulations of baryonic effects in van Daalen et al. (2011) suggests that A_0 is almost unchanged, while higher order coefficients change significantly, because baryonic effects redistribute gas and DM inside the haloes without changing the overall halo mass. We advocate that marginalizing over higher order expansion coefficients should immunize against baryonic

effects without much loss of information. We explore the degeneracy between the amplitude, supersample variance and baryonic effects, finding that it can be broken using information above $k \sim 0.3\,h\,\mathrm{Mpc}^{-1}$.

Our results suggest that analytic modelling of DM clustering provides important insights even in the era of large simulations. It offers a promising venue not only for an accurate power spectrum description, but also for the covariance matrix modelling, for optimal extraction of information from the data, and for description of baryonic effects. We have shown that in the context of covariance matrix calculations our model is likely to be more reliable than simulations with insufficient total volume. However, more work remains to be done before it can be applied to the weak lensing observations. For example, in this paper we focused on the DM clustering description in terms of its power spectrum. If one wants to apply the method to the weak lensing observations one needs to perform the line-of-sight projections of the model on to the weak lensing power spectrum $C_l^{\kappa\kappa}$, where κ is the convergence which can be written as a projection of the density along the line of sight. Projecting powers of k simply gives the same powers of l, so if the projection kernels are narrow, as would be the case for weak lensing tomography, the analysis remains essentially unchanged, except for the fact that weak lensing probes matter density rather than density perturbations, so convergence is also multiplied by an overall mean matter density. If the projection kernels are broad and there are significant contributions from nearby structures for which $k > 0.7\,h\,\mathrm{Mpc}^{-1}$ projects to a low l, then one needs to assess these effects and improve the model to account better for the high k contributions. Similarly, one also needs to project baryonic effects and covariance matrix. This programme is feasible and if implemented it will give a completely analytical description of the weak lensing power spectrum and its covariance matrix without any need to use simulations.

ACKNOWLEDGEMENTS

We thank Z. Vlah for extensive discussions and for providing the Zeldovich power spectrum code, F. Schmidt and M. Takada for useful comments, Y. Li for electronic form of the plots of Li et al. (2014), M. van Daalen for electronic form of the plots of van Daalen et al. (2011), L. Blot for providing us data from their simulations, and J. Harnois-Deraps and U. Pen for interpretation of their covariance matrix. US is supported in part by the NASA ATP grant NNX12AG71G. IM would like to thank the hospitality of LBNL.

REFERENCES

Baldauf T., Seljak U., Senatore L., Zaldarriaga M., 2011, J. Cosmol. Astropart. Phys., 10, 31
Baldauf T., Seljak U., Smith R. E., Hamaus N., Desjacques V., 2013, Phys. Rev. D, 88, 083507
Bartelmann M., Schneider P., 2001, Phys. Rep., 340, 291
Bird S., Viel M., Haehnelt M. G., 2012, MNRAS, 420, 2551
Blot L., Corasaniti P. S., Alimi J.-M., Reverdy V., Rasera Y., 2014, preprint (arXiv:1406.2713)
Blumenthal G. R., Faber S. M., Primack J. R., Rees M. J., 1984, Nature, 311, 517
Cooray A., Sheth R., 2002, Phys. Rep., 372, 1
Fu L. et al., 2008, A&A, 479, 9
Hamilton A. J. S., Rimes C. D., Scoccimarro R., 2006, MNRAS, 371, 1188
Harnois-Déraps J., Pen U.-L., 2012, MNRAS, 423, 2288
Heitmann K., Higdon D., White M., Habib S., Williams B. J., Lawrence E., Wagner C., 2009, ApJ, 705, 156
Heitmann K., White M., Wagner C., Habib S., Higdon D., 2010, ApJ, 715, 104
Hoekstra H. et al., 2006, ApJ, 647, 116
Kilbinger M. et al., 2013, MNRAS, 430, 2200
Komatsu E., Seljak U., 2002, MNRAS, 336, 1256
Lawrence E., Heitmann K., White M., Higdon D., Wagner C., Habib S., Williams B., 2010, ApJ, 713, 1322
Lewis A., Challinor A., Lasenby A., 2000, ApJ, 538, 473
Li Y., Hu W., Takada M., 2014, Phys. Rev. D, 89, 083519
Ma C., Fry J. N., 2000, ApJ, 543, 503
McClelland J., Silk J., 1977, ApJ, 217, 331
McCarthy I. G. et al., 2010, MNRAS, 406, 822
Massey R. et al., 2007, MNRAS, 376, 13
Matsubara T., 2008, Phys. Rev. D, 77, 063530
Mortonson M. J., Huterer D., Hu W., 2010, Phys. Rev. D, 82, 063004
Navarro J. F., Frenk C. S., White S. D. M., 1997, ApJ, 490, 493
Peacock J. A., Smith R. E., 2000, MNRAS, 318, 1144
Percival W. J. et al., 2014, MNRAS, 439, 2531
Refregier A., 2003, ARA&A, 41, 645
Sato M., Hamana T., Takahashi R., Takada M., Yoshida N., Matsubara T., Sugiyama N., 2009, ApJ, 701, 945
Sato M., Takada M., Hamana T., Matsubara T., 2011, ApJ, 734, 76
Schneider P., Bartelmann M., 1995, MNRAS, 273, 475
Schrabback T. et al., 2010, A&A, 516, A63
Seljak U., 2000, MNRAS, 318, 203
Semboloni E., Hoekstra H., Schaye J., van Daalen M. P., McCarthy I. G., 2011, MNRAS, 417, 2020
Semboloni E., Hoekstra H., Schaye J., 2013, MNRAS, 434, 148
Sherwin B. D., Zaldarriaga M., 2012, Phys. Rev. D, 85, 103523
Slosar A., Hirata C., Seljak U., Ho S., Padmanabhan N., 2008, J. Cosmol. Astropart. Phys., 8, 31
Smith R. E. et al., 2003, MNRAS, 341, 1311

Takada M., Hu W., 2013, Phys. Rev. D, 87, 123504
Takada M., Spergel D. N., 2014, MNRAS, 441, 2456
Takahashi R., Sato M., Nishimichi T., Taruya A., Oguri M., 2012, ApJ, 761, 152
Taylor A. N., 1993, in Bouchet F., Lachieze-Rey M., eds, Cosmic Velocity Fields. Editions Frontieres, Gif-sur-Yvette, p. 585
Taylor A., Joachimi B., 2014, MNRAS, 442, 2728
Tinker J., Kravtsov A. V., Klypin A., Abazajian K., Warren M., Yepes G., Gottlöber S., Holz D. E., 2008, ApJ, 688, 709
van Daalen M. P., Schaye J., Booth C. M., Dalla Vecchia C., 2011, MNRAS, 415, 3649
Yoo J., Seljak U., 2012, Phys. Rev. D, 86, 083504
Zel'dovich Y. B., 1970, A&A, 5, 84

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