

Combining clustering and abundances of galaxy clusters to test cosmology and primordial non-Gaussianity

Annalisa Mana^{1,2}, Tommaso Giannantonio^{1,2}, Jochen Weller^{1,2,3}, Ben Hoyle^{1,2}, Gert Hütsi^{4,5} and Barbara Sartoris^{1,2,6,7}

¹ *Universitäts-Sternwarte München, Ludwig-Maximilians Universität München, Scheinerstr. 1, D-81679, München, Germany*

² *Excellence Cluster Universe, Technical University Munich, Boltzmannstr. 2, D-85748 Garching, Germany*

³ *Max-Planck-Institut für Extraterrestrische Physik, Giessenbachstr., D-85748 Garching, Germany*

⁴ *Max-Planck-Institut für Astrophysik, Karl-Schwarzschild-Str. 1, D-85748 Garching, Germany*

⁵ *Tartu Observatory, EST-61602 Tõravere, Estonia*

⁶ *Dipartimento di Fisica, Sezione di Astronomia, Università di Trieste, Via Tiepolo 11, I-34143 Trieste, Italy*

⁷ *INAF-Osservatorio Astronomico di Trieste, Via Tiepolo 11, I-34143 Trieste, Italy*

5 March 2013

ABSTRACT

We present the clustering of galaxy clusters as a useful addition to the common set of cosmological observables. The clustering of clusters probes the large-scale structure of the Universe, extending galaxy clustering analysis to the high-peak, high-bias regime. Clustering of galaxy clusters complements the traditional cluster number counts and observable-mass relation analyses, significantly improving their constraining power by breaking existing calibration degeneracies. We use the maxBCG galaxy clusters catalogue to constrain cosmological parameters and cross-calibrate the mass-observable relation, using cluster abundances in richness bins and weak-lensing mass estimates. We then add the redshift-space power spectrum of the sample, including an effective modelling of the weakly non-linear contribution and allowing for an arbitrary photometric redshift smoothing. The inclusion of the power spectrum data allows for an improved self-calibration of the scaling relation. We find that the inclusion of the power spectrum typically brings a $\sim 50\%$ improvement in the errors on the fluctuation amplitude σ_8 and the matter density Ω_m . Finally, we apply this method to constrain models of the early universe through the amount of primordial non-Gaussianity of the local type, using both the variation in the halo mass function and in the cluster bias. We find a constraint on the amount of skewness $f_{\text{NL}} = 12 \pm 157$ (1σ) from the cluster data alone.

Key words: galaxies: clusters: general – cosmology: cosmological parameters – methods: statistical

1 INTRODUCTION

Galaxy clusters are the most massive bound systems in the Universe which trace the evolution of the large-scale structure (LSS) (see the recent review by Allen et al. 2011). The initial density perturbations are thought to have formed in the early universe from inflationary physics (Lyth & Liddle 2009). In the simplest scenario, the perturbations can be modelled as Gaussian random fields (Bardeen et al. 1986), which evolve gravitationally. This leads to the formation of bound dark matter structures – the haloes, whose abundance is described by the halo mass function. The simplest infall formation model (Press & Schechter 1974) is complicated by dynamical effects, meaning that accurate modelling of the mass function requires partial calibration (Sheth & Tormen 1999; Maggiore & Riotto 2010; Corasaniti & Achitouv 2011) or full fitting (Jenkins et al. 2001; Tinker et al. 2008) to N -body simulations. Halo mergers and internal processes such as galaxy formation further complicate the picture at small scales (e.g., see

Borgani & Kravtsov 2009, for a review). Nonetheless galaxy clusters form at a comoving scale of $\sim 10 h^{-1} \text{Mpc}$, allowing for a simpler theoretical description than is possible for smaller structures such as galaxies. Due to their scale, clusters reside in the tail of the halo mass function and thus their numbers are exponentially sensitive to variations in cosmology (see e.g., Evrard 1989; Frenk et al. 1990; Bahcall et al. 1997).

Clusters are detected across multiple wavelengths with varying degrees of success. A few dozens have been found in the millimeter by the Atacama Cosmology Telescope (ACT) (Menanteau et al. 2012), hundreds with the South Pole Telescope (SPT) (Reichardt et al. 2012) and the Planck satellite (Planck Collaboration et al. 2011); also a few hundreds of clusters have been found in the X-ray (REFLEX, BCS, eBCS catalogues by Böhringer et al. 2004; Ebeling et al. 1998, 2000) using the the ROSAT satellite All-Sky Survey (Voges et al. 1999), the Chandra Cluster Cosmology Project (Burenin et al. 2007; Vikhlinin et al.

2009a), and by Mehrtens et al. (2012) using X-ray Multi-Mirror Mission Newton (Fassbender 2008). Many tens of thousands have been found in the optical using Sloan Digital Sky Survey (SDSS) data to construct the maxBCG (Koester et al. 2007b) and GM-BCG (Hao et al. 2010) catalogues, based on the selection of brightest cluster galaxies (BCG) to identify the clusters' centres, and by Gladders & Yee (2005) using the Canada-France-Hawai'i Telescope and Cerro Tololo Inter-American Observatory.

While detectable in large numbers, the main obstacle with using optical clusters as probes of cosmology is the difficult choice of a low-scatter mass proxy. Efforts on this front have been led by Rozo et al. (2010) and Zu et al. (2012), who derived cosmological constraints from the maxBCG cluster sample. The tightest scaling relation between observable and cluster mass comes from X-ray data ($< 10\%$ scatter, Allen et al. 2008). Constraints on dark energy with $\sim 20\%$ uncertainty were obtained from X-ray cluster samples studied by Mantz et al. (2008, 2010) and Vikhlinin et al. (2009b). Data on the cluster masses obtained from weak-lensing analyses of background galaxies have also been combined with the number counts to improve the constraining power of the cluster mass function (Johnston et al. 2007; Sheldon et al. 2009; Mahdavi et al. 2007). The statistics of rare events in the high-peak, high-mass limit has also been used by Hotchkiss (2011), Hoyle et al. (2012) to test cosmology.

The uncertainty in the scaling relation is one of the biggest obstacles in using galaxy clusters as cosmological probes, as pointed out by Haiman et al. (2001), Battye & Weller (2003). Majumdar & Mohr (2003) suggested to use the clustering of clusters as a complementary probe. So far, only limited efforts have been dedicated to the measurement of the clustering properties of galaxy clusters: Hütsi (2010) measured the power spectrum of maxBCG clusters resulting in weak detection of baryon acoustic oscillations (BAO), Estrada et al. (2009) measured the correlation function for the same, and Hong et al. (2012) measured the correlation function of extended versions of maxBCG. Finally Balaguera-Antolínez et al. (2011) measured the power spectrum of the REFLEX X-ray cluster catalogue. The goal of this paper is to fully include the clustering information in the cosmological analysis of cluster data: we show that its inclusion significantly improves the cosmological constraints, and also reduces the degeneracies between the scaling relation nuisance parameters. We present the improved cosmological results obtained in this way from the maxBCG data.

As an interesting application, we present the constraining power of these data on the amount of primordial non-Gaussianity (PNG) of the initial density perturbations, which is expected to be produced in some models of the early universe. Briefly, while the simplest single-field slow-roll inflation produces nearly-Gaussian initial conditions (Maldacena 2003; Acquaviva et al. 2003), there exist alternatives, such as multi-field models, which can produce large non-Gaussianities (see e.g. the recent review by Chen 2010). These would have multiple observable consequences, of which we here consider two: the halo mass function changes as a function of the non-zero skewness (Matarrese et al. 2000; LoVerde et al. 2008; Pillepich et al. 2010; Achitouv & Corasaniti 2012a,b), and in the local and orthogonal cases the halo bias becomes strongly scale-dependent due to the coupling of long- and short-wavelength modes (Dalal et al. 2008; Matarrese & Verde 2008; Slosar et al. 2008; Afshordi & Tolley 2008; Desjacques et al. 2009; Giannantonio & Porciani 2010; Desjacques & Seljak 2010; Schmidt & Kamionkowski 2010; Desjacques et al. 2011). Measurements of PNG can potentially rule out entire classes of infla-

tionary models (Bartolo et al. 2004; Suyama et al. 2010). The latest constraints on the local PNG parameter f_{NL} from the bispectrum of the cosmic microwave background (CMB) as measured by the WMAP satellite are $-3 < f_{\text{NL}} < 77$ at 95% c.l. (Hinshaw et al. 2012; Bennett et al. 2012); comparable bounds have been obtained from the LSS using multiple galaxy catalogues (Slosar et al. 2008; Afshordi & Tolley 2008; Xia et al. 2010a,b, 2011; Sefusatti et al. 2012; Ross et al. 2013; Giannantonio et al. 2013); future galaxy surveys such as Euclid are expected to reach an accuracy of $\Delta f_{\text{NL}} \sim 3$ (Giannantonio et al. 2012). Oguri (2009) suggested that measuring the variance of cluster counts can yield significant constraints on PNG, while Sartoris et al. (2010) showed in principle how such constraints can be improved by using the cluster power spectrum. Here we constrain PNG with the combined cluster data.

The paper is organised as follows. In Section 2 we describe measurements of the cluster abundances, weak lensing mass estimates and power spectrum of the maxBCG catalogue. In Section 3 we introduce the theoretical framework including number counts and total mass determination from the mass function, mass-observable relation, bias and power spectrum definitions, and the effects of non-Gaussian initial conditions. Section 4 presents our Monte Carlo Markov Chain (MCMC) analysis, the cosmological constraints and the relevant degeneracies. We draw our conclusions in Section 5. We assume that the Universe is spatially flat on large scales throughout.

2 DATA

2.1 The maxBCG cluster catalog

The maxBCG catalogue (Koester et al. 2007a) is a sample of 13,823 galaxy clusters compiled from SDSS photometric data. The catalogue is assembled by selecting the brightest cluster galaxy and applying a red-sequence method to identify cluster members in its neighbourhood. In this way clusters with richness (number of member galaxies) ranging from 10 to 190 are selected. The low mass limit of this sample is $M_{\text{lim}} \sim 7 \cdot 10^{13} h^{-1} M_{\odot}$, which evolves weakly with redshift. This relatively low mass limit results in a sample that is significantly larger than other current galaxy cluster catalogues. The clusters are chosen in an approximately volume-limited way from a 500 Mpc^3 region, covering $\sim 7500 \text{ deg}^2$ of sky with a photometric redshift (photo- z) range of $0.1 \leq z \leq 0.3$. The photo- z errors are small and of the order of $\Delta z = 0.01$. An analysis of mock samples shows that the maxBCG algorithm results in more than 90% purity and more than 85% completeness, for clusters with masses $M \geq 10^{14} M_{\odot}$ (Koester et al. 2007a).

We define the richness N_{gal} as the number of red galaxies within the radius R_{200} from the cluster centre. R_{200} is the radius within which the average overdensity is 200 times the mean density of the Universe. The catalogue is divided into nine richness bins in the range $11 \leq N_{\text{gal}} \leq 120$, which approximately corresponds to $7 \cdot 10^{13} h^{-1} M_{\odot} \leq M \leq 1.2 \cdot 10^{15} h^{-1} M_{\odot}$ (Roza et al. 2010). We found that adding the five remaining high-mass clusters of the maxBCG with richness $N_{\text{gal}} > 120$ has a negligible impact on the cosmological analysis, so we do not include them. We also use an additional bin at $9 \leq N_{\text{gal}} \leq 11$ (Roza, private communication), although we checked the results are not affected by this. For the cosmological analysis we include Poisson errors and sample variance due to large-scale structure (Hu & Kravtsov 2003). Furthermore, we assume 100% purity and completeness, including a 5% uncertainty (Roza et al. 2010), which we add in quadrature. We found that the

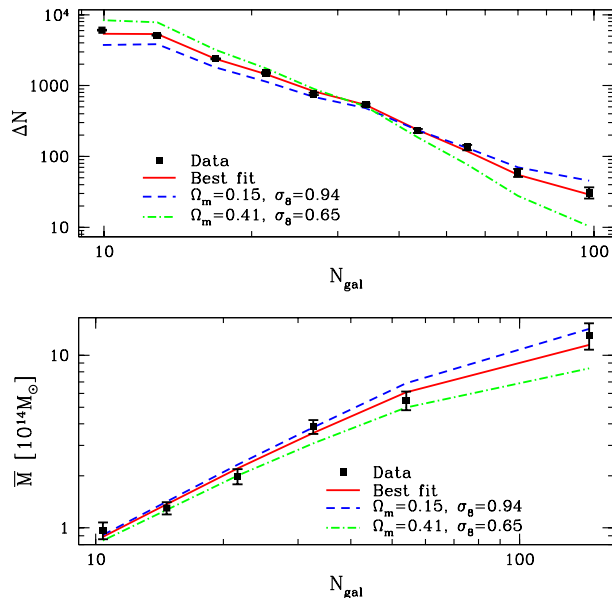


Figure 1. *Top panel:* MaxBCG cluster counts data (black points) and theoretical predictions according to the prescriptions of Section 3 for a choice of different cosmologies (without primordial non-Gaussianity). The red line represents the best-fit model to our full data set (counts, total masses and power spectrum). *Bottom panel:* Mean masses of maxBCG clusters (black points) and theoretical predictions for different cosmologies, as above.

photo- z errors have a negligible impact on the number counts analysis presented here, so we neglect their effect on the number counts covariance matrix.

We show in the top panel of Fig. 1 the counts data, together with the predicted counts for a selection of different cosmologies, modelled as described in Section 3.

2.2 Galaxy cluster masses from weak lensing observations

Sheldon et al. (2009) measured the weak lensing (WL) effect from clusters in the maxBCG catalogue. By stacking the clusters, mean cluster surface density profiles were created for different luminosity and richness bins. The stacking of clusters in a given richness bin improves the signal-to-noise considerably compared to the measurement of the profile of an individual cluster. Johnston et al. (2007) used these profiles and reconstructed mean 3D cluster density and mass profiles, which allows one to estimate the mass (and concentration) of clusters in a given redshift bin. For this reconstruction a Navarro-Frenk-White (NFW) profile (Navarro et al. 1997) for the cluster density was assumed. Johnston et al. (2007) were then able to construct a mean mass-richness relation, finding for the whole sample of groups and clusters

$$M_{200}(N_{\text{gal}}) \approx 8.8 \times 10^{13} h^{-1} M_{\odot} (N_{\text{gal}}/20)^{1.28}, \quad (1)$$

where M_{200} is the mass contained within the radius R_{200} . Due to photometric redshift bias these masses are corrected upwards by a factor of 1.18 as described in Mandelbaum et al. (2008b), Rozo et al. (2010). For the cosmological analysis presented here we follow Rozo et al. (2010) and fit simultaneously for the mass-richness relation using the Johnston et al. (2007) data and their errors. We use five richness bins for this, in the range $12 \leq N_{\text{gal}} \leq$

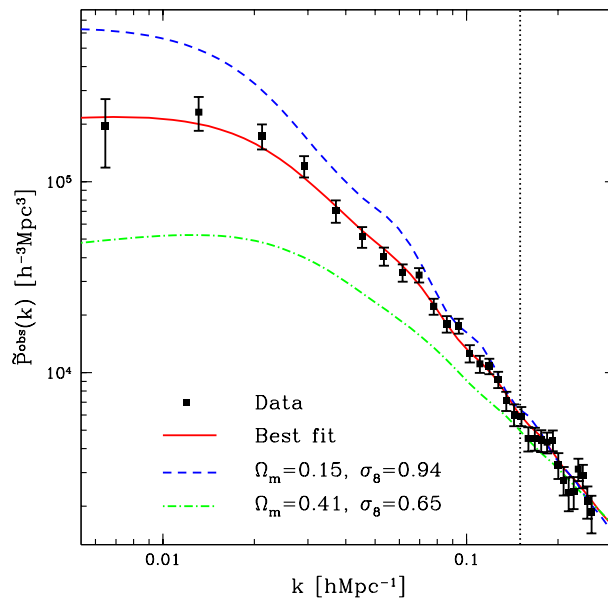


Figure 2. The observed power spectrum of maxBCG clusters (black points) compared with the full theoretical modelling \bar{P}^{obs} described in Section 3 for our best-fit model (red solid line), and for two other models, assuming no primordial non-Gaussianity. The dotted line at $k = 0.15 h \text{ Mpc}^{-1}$ represents our choice of k_{max} , which is the smallest scale we use in the analysis.

300, plus another extra bin at $9 \leq N_{\text{gal}} \leq 12$ (Roza, private communication).

An independent weak lensing analysis of the maxBCG sample was performed by Mandelbaum et al. (2008a), who found a mean mass difference of approximately 6% with respect to Johnston et al. (2007). We follow Rozo et al. (2010) and include this discrepancy by introducing an offset factor β with a suitable chosen prior as described in Section 3.3.

The bottom panel of Fig. 1 shows the mean WL mass estimates data, together with the theoretical mean masses for a selection of different cosmologies, modelled as described in Section 3.

2.3 MaxBCG power spectrum

We consider the redshift-space power spectrum of the maxBCG sample, as measured by Hütsi (2010). For the full details of the power spectrum measurement, along with systematics tests, we refer the reader to Hütsi (2006a,b, 2010). The direct Fourier method by Feldman et al. (1994) (FKP) was used, with the difference that fast Fourier transforms (FFTs) were used instead of direct summation. This method actually yields the pseudo-spectrum, i.e. the measurement products are convolved with the window function of the survey. We take this into account when modelling the theoretical spectra in our analysis. To implement the modified FKP method, the following steps were followed:

- (i) The survey selection function (footprint) was represented using a random (unclustered) catalogue with 100 times more points than maxBCG sample;
- (ii) The overdensity field was calculated on a regular grid using the triangular shaped cloud method (Hockney & Eastwood 1988) mass assignment scheme — it was checked that the aliasing effects due to the finite grid size were negligible for the measurements,

and were nonetheless corrected with the iterative method by Jing (2005);

(iii) The gridded overdensity field was transformed to Fourier space using the FFT;

(iv) The raw 3D power spectrum was estimated by taking the modulus squared of the FFT;

(v) The shot noise contribution was subtracted.

The uncertainties on the power spectrum measurements were estimated with three different methods: with the original FKP theoretical prescription, which assumes Gaussian errors from cosmic variance plus a shot noise contribution; with a jackknife method, implemented by dividing the survey into a total of 75 regions; and with a Monte Carlo (MC) method, based on the fiducial Λ CDM cosmology, in which 1000 mock realisations of the maxBCG survey were generated, including redshift-space distortions and photo- z errors. These three methods were shown to be comparable; in this work we use the MC covariance matrix.

The power spectrum measurements are shown in Fig. 2. To take into account data in the quasi-linear regime only, we restrict ourselves to scales larger than (wavenumbers smaller than) $k_{\max} = 0.15 h \text{ Mpc}^{-1}$.

2.4 The counts-clustering off-diagonal covariance

The diagonal blocks of the data covariance matrix for counts, masses and the power spectrum have been described above. We estimate the off-diagonal terms of the covariance matrix between the clustering and the binned number distributions ΔN of the maxBCG clusters. For simplicity, instead of the power spectrum we use here, as a clustering estimator, the projected correlation function $w(\theta)$, which is defined as

$$w(\theta) \equiv \langle \delta_h(\hat{\mathbf{n}}) \delta_h(\hat{\mathbf{n}}') \rangle, \quad (2)$$

where $\delta_h(\hat{\mathbf{n}})$ is the halo (cluster) projected overdensity in a direction $\hat{\mathbf{n}}$, and the average is carried over all pairs at an angular distance θ .

We use the jackknife technique as follows: we split the maxBCG footprint into 100 equal-area jackknife regions using HEALPIX¹ (Górski et al. 2005) and populate the full footprint with 50 random points for each maxBCG cluster, to reduce shot noise. We use the correlation estimator by Landy & Szalay (1993) to calculate the correlation function, and bin the number of clusters within six equal-width bins in \log_{10} space. We iteratively remove and replace each jackknife region and calculate the number histogram and correlation function at each iteration. The covariance matrix C_{JK} between measured statistics $x = x(\alpha)$ and $y = y(\beta)$ can be estimated from N jackknives using (see e.g. Efron 1982):

$$\left[C_{\text{JK}}(x_i, y_j) \right]_{\alpha\beta} = \frac{N-1}{N} \sum_{k=1}^N (x_{-i}^k - \bar{x}_i)_{\alpha} (y_{-j}^k - \bar{y}_j)_{\beta}, \quad (3)$$

where x_{-i} (y_{-j}) is the statistic with jackknife region i (j) removed, and \bar{x} (\bar{y}) is the average value of all x_{-i} (y_{-j}). We note that typically, but not necessarily, x and y are the same statistic.

We compare the square root of the diagonal elements of the covariance matrix $C_{\text{JK}}[w(\theta), w(\theta)]$ with the error expected from Poisson counting statistics and find agreement with the theoretical expectations (as described by e.g., Ross et al. 2009), and also find that the diagonal elements of $C_{\text{JK}}(\Delta N, \Delta N)$ are approximately Poissonian, independently of the number of jackknives used. In Fig. 3

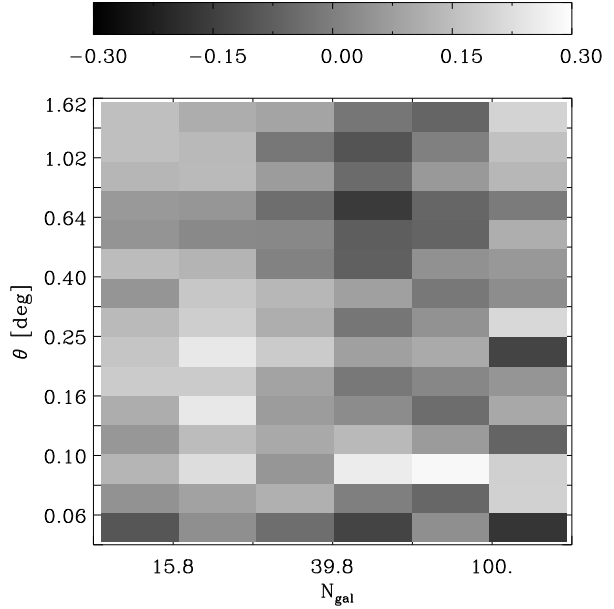


Figure 3. The off-diagonal elements of the normalised covariance matrix of the correlation function $w(\theta)$ and the histogram distribution of ΔN , as calculated using the jackknife technique.

we show the off-diagonal terms of the normalised $C_{\text{JK}}[\Delta N, w(\theta)]$, and note that the average value and 1σ error of the off-diagonal terms are -0.06 ± 0.10 , which is consistent with zero. We observe that as the number of jackknives increases, the mean and error of the average value of the off-diagonal terms approaches and fluctuates around zero.

We compare the magnitude of the off-diagonal terms obtained from the maxBCG clusters with simulated clusters from the Millennium Simulation (Springel et al. 2005; Lemson & Virgo Consortium 2006). Specifically, we join the lightcone table of `Henriques2012a.wmap1.BC03.AllSky_00` (Henriques et al. 2012; Guo et al. 2011) with the halo-tree table `MPAHaloTrees.MHalo`. We apply the same redshift and survey footprint constraints to mimic the maxBCG sample and calculate the correlation function and histogram distribution of ΔN . We find the data and simulations agree closely: e.g., for 100 jackknives the mean and 1σ error of the off-diagonal terms are 0.05 ± 0.12 from the simulations.

From these tests we conclude that ignoring the off-diagonal covariance matrix between clustering and number counts is a reasonable approximation.

3 THEORETICAL MODELLING OF CLUSTER STATISTICS

3.1 The cluster mass function

Press & Schechter (1974) first calculated the expected number of dark matter haloes of a given mass and redshift. This was better described by the excursion set approach (Bond et al. 1991) and generalised to non-spherical model by Sheth & Tormen (1999), who calibrated their mass function with N -body simulations. Even more accurate estimations are achieved with a full fitting to N -body simulation (i.e. Jenkins et al. 2001). The current state-of-the-art halo

¹ <http://healpix.jpl.nasa.gov/>

mass function has been estimated by Tinker et al. (2008, 2010): this mass function is valid over wide redshift and mass ranges.

The expected number density of virialised dark matter haloes as a function of mass and redshift can be expressed as

$$\frac{dn(M, z)}{d \ln M} = \bar{\rho}_m \left| \frac{d \ln \sigma^{-1}}{dM} \right| f(\nu), \quad (4)$$

where $\bar{\rho}_m$ is the mean matter density of the Universe, $\nu \equiv \delta_c / \sigma(M, z)$, $\delta_c = 1.686$ is the threshold linear overdensity for spherical collapse in a matter dominated Universe, and $\sigma^2(M, z)$ is the variance of the linear matter density field at $M = 4\pi R^3 \bar{\rho}_m / 3$. In this work we use the mass function given by Tinker et al. (2010) for cluster mass at R_{200} , with an overdensity of $\Delta = 200$ in units of the mean mass density of the Universe:

$$f_T(\nu) = 0.368 \left[1 + (\hat{\beta}\nu)^{-2\hat{\phi}} \right] \nu^{2\hat{\eta}} e^{-\hat{\gamma}\nu^2/2}, \quad (5)$$

where the parameters evolve in redshift as

$$\begin{aligned} \hat{\beta} &= 0.589 (1+z)^{0.20}, & \hat{\phi} &= -0.729 (1+z)^{-0.08}, \\ \hat{\eta} &= -0.243 (1+z)^{0.27}, & \hat{\gamma} &= 0.864 (1+z)^{-0.01}. \end{aligned} \quad (6)$$

3.2 The mass-richness scaling relation

In order to perform a cosmological analysis, we need to make some assumptions on the scaling relation between the true mass of a cluster and its observed richness $N_{\text{gal}}^{\text{obs}}$. We first consider the probability of observing $N_{\text{gal}}^{\text{obs}}$ member galaxies at R_{200} for a given true mass M of the cluster. We can write this probability as the product of the probability of a cluster having true richness N_{gal} given the mass M and the probability of observing $N_{\text{gal}}^{\text{obs}}$ member galaxies given the true richness N_{gal} :

$$p(N_{\text{gal}}^{\text{obs}}|M; z) = p(N_{\text{gal}}^{\text{obs}}|N_{\text{gal}}) p(N_{\text{gal}}|M). \quad (7)$$

Here $p(N_{\text{gal}}^{\text{obs}}|N_{\text{gal}})$ is assumed to follow a lognormal distribution as suggested by Lima & Hu (2005)

$$p(N_{\text{gal}}^{\text{obs}}|N_{\text{gal}}) = \frac{1}{\sqrt{2\pi\sigma^2 \ln N_{\text{gal}}}} \exp[-x^2(N_{\text{gal}}^{\text{obs}})], \quad (8)$$

where

$$x(N_{\text{gal}}^{\text{obs}}) = \frac{\ln N_{\text{gal}}^{\text{obs}} - \ln N_{\text{gal}}}{\sqrt{2\sigma^2 \ln N_{\text{gal}}}} \quad (9)$$

and $\sigma_{\ln N_{\text{gal}}}$ is the scatter in the mass-richness relation (Lima & Hu 2005; Battye & Weller 2003). Following Johnston et al. (2007) and Rozo et al. (2010), we assume the scaling relation to be a power law in mass, i.e.

$$\ln M = \ln \hat{M}_{200|20} + \alpha_N \ln(N_{\text{gal}}/20), \quad (10)$$

with $\hat{M}_{200|20}$ the mass of a cluster with 20 member galaxies within a radius of R_{200} and α_N the slope of the scaling relation. We fit this relation by fixing two pivot points in mass $M_1 = 1.3 \cdot 10^{14} M_{\odot}$ and $M_2 = 1.3 \cdot 10^{15} M_{\odot}$, while the corresponding richness values $\ln N_1 \equiv \ln N_{\text{gal}}|M_1$ and $\ln N_2 \equiv \ln N_{\text{gal}}|M_2$ are kept as free parameters.

The statistical scatter around this relation is assumed to be constant with redshift and mass for individual clusters. To get an estimate of this quantity is not trivial; however Rozo et al. (2009) used weak lensing and X-ray observations together with the maxBCG richness to have three different mass proxies. By demanding consistency between the X-ray and weak lensing measurements, Rozo et al. (2009) found $\sigma_{\ln M|N_{\text{gal}}} = 0.45_{-0.18}^{+0.20}$, which is

the scatter in mass given the richness. For our cosmological analysis described in Section 4 we need to place a prior on the converse scatter, $\sigma_{\ln N_{\text{gal}}|M}$. The two quantities can be readily related to each other by invoking the scaling relation of Eq. (10), which results in $\sigma_{\ln M|N_{\text{gal}}} = \alpha_N \cdot \sigma_{\ln N_{\text{gal}}|M}$.

3.3 Modelling galaxy cluster counts and total masses

In order to predict the number of observed galaxy clusters for an observed richness $N_{\text{gal}}^{\text{obs}}$, we can use the probability distribution and scaling relation defined in the previous section. The cluster average number density within a richness bin $[N_{\text{gal},i}^{\text{obs}}, N_{\text{gal},i+1}^{\text{obs}}]$ is given by

$$\begin{aligned} n_i &= \int_{N_{\text{gal},i}^{\text{obs}}}^{N_{\text{gal},i+1}^{\text{obs}}} d \ln N_{\text{gal}}^{\text{obs}} \int d \ln N_{\text{gal}} \frac{dn}{d \ln N_{\text{gal}}} p(N_{\text{gal}}^{\text{obs}}|N_{\text{gal}}) = \\ &= \int d \ln N_{\text{gal}} \frac{dn}{d \ln N_{\text{gal}}} \frac{1}{2} [\text{erfc}(x_i) - \text{erfc}(x_{i+1})], \end{aligned} \quad (11)$$

where $x_i \equiv x(N_{\text{gal},i}^{\text{obs}})$,

$$\frac{dn}{d \ln N_{\text{gal}}} = \frac{dn}{d \ln M} \cdot \frac{d \ln M}{d \ln N_{\text{gal}}} = \alpha_N \frac{dn}{d \ln M}, \quad (12)$$

and we have employed the scaling relation of Eq. (10). The total number of predicted galaxy clusters within a richness bin can be calculated as

$$\Delta N_i = \Delta \Omega \int_{z_{\text{min}}}^{z_{\text{max}}} dz \frac{d^2 V}{dz d\Omega} n_i, \quad (13)$$

where $\Delta \Omega$ is the survey sky coverage and $d^2 V/dz/d\Omega$ the volume element. The cosmology dependence is driven by the mass function and by the comoving volume element.

We can write similar expressions to Eqs. (11) and (13) for the total mass of clusters. The average total mass $(nm)_i$ contained within the same richness bin can be obtained as in Eq. (11) by weighting the integrand by the mass, estimated via the mass-observable relation. The total mass of clusters within a richness bin is then

$$(\Delta N \bar{M})_i = \beta \Delta \Omega \int_{z_{\text{min}}}^{z_{\text{max}}} dz \frac{d^2 V}{dz d\Omega} (nm)_i, \quad (14)$$

where β is an additional nuisance parameter introduced to account for possible mismatch with the WL masses, as discussed above and in Rozo et al. (2010).

3.4 Clustering of clusters

Galaxy clusters can be studied as tracers of the LSS (Mo et al. 1996), corresponding to the highest-density regions of the dark matter overdensity field $\delta(\mathbf{x}, z)$. If we assume linear theory and a local deterministic halo bias (Fry & Gaztanaga 1993), then the dark-matter haloes overdensity is $\delta_h(\mathbf{x}, M, z) = b_0 + b_h(M, z) \delta(\mathbf{x}, z)$. The local bias assumption breaks down in the case of primordial non-Gaussianity (see Section 3.5). As the effect of baryons is negligible for the clustering properties of the clusters, in the following we use the naming ‘cluster’ and ‘halo’ interchangeably.

Halo bias

The halo bias can be derived from a theory of the mass function via the peak-background split formalism (Cole & Kaiser 1989;

Mo & White 1996). This method gives a prediction for the halo bias in Lagrangian space $b^L(M, z)$, which can be evolved into the observable Eulerian space as $b = 1 + b^L$ (Mo & White 1996), considering linear perturbation only, spherical collapse and no large-scale velocity bias. We assume the bias is scale-independent (except for the modifications in the presence of PNG). At linear order, the Lagrangian bias is

$$b^L(M, z) = -\frac{f'(\nu)}{f(\nu)\sigma}, \quad (15)$$

where the derivative of the mass function is taken with respect to ν , and the mass and redshift dependences of ν and σ are implicit. When using the Tinker mass function and keeping the leading order terms, the Eulerian bias is (Tinker et al. 2010):

$$b_T(M, z) \simeq 1 + \frac{\hat{\gamma}\nu^2 - (1 + 2\hat{\eta})}{\delta_c} + \frac{2\hat{\phi}/\delta_c}{1 + [\hat{\beta}\nu]^{2\hat{\phi}}}, \quad (16)$$

where all parameters are defined as in Eq. (6). We show in Fig. 4 the Tinker halo bias $b_T(M)$ as a function of halo mass M at $\bar{z} = 0.2$, which is the mean redshift of the maxBCG clusters, compared with the Press & Schechter (1974) (PS) case $b_{PS} = 1 - 1/\delta_c + \delta_c/\sigma^2$, and with the scale-independent part of the bias in the presence of PNG described below, for our combined best-fit model.

We obtain the average cluster bias \bar{b} over the mass range of our data by weighting with the mass function (Lima & Hu 2005):

$$\bar{b}(z) = \frac{1}{\Delta N} \int_{M_{\min}}^{M_{\max}} d \ln M \frac{dn(M, z)}{d \ln M} b_T(M, z), \quad (17)$$

where the normalisation factor ΔN is the full integral of the mass function in the observed range.

Power spectra

We then define the observable clustering statistics in Fourier space (denoted by a tilde). As we consider linear scales only, the halo-halo power spectrum P_{hh} can be related to the linear matter power spectrum P_{lin} as

$$P_{hh}(k, M, z) = b^2(M, z) P_{\text{lin}}(k, z) = b^2(M, z) D^2(z) P_{\text{lin}}(k, 0), \quad (18)$$

where $D(z)$ is the linear growth function. We integrate the mass dependence by weighting the bias as described in Eq. (17) and we compute all quantities at the mean redshift of our cluster sample, $\bar{z} \simeq 0.2$. This is further justified by observing that the growth of $\bar{b}(z)$ is compensated by a similar drop in $D(z)$; we have checked that for our fiducial cosmology, in the observed range $0.1 \leq z \leq 0.3$ the variation of $\bar{b}(z)D(z)$ is at the percent level.

Before fitting models to the data, the following four effects have to be taken into account, following the description by Hütsi (2010): the photo- z errors, which are responsible for a damping of the spectrum on small scales; the convolution with the survey window, which suppresses the power on large scales; the non-linearities which add power on small scales, and the redshift-space distortions. The total observed power spectrum P^{obs} is modelled as

$$P^{\text{obs}}(k) = \int d \ln \kappa \kappa^3 P_{\text{NL}}(\kappa) K(\kappa, k), \quad (19)$$

where $K(\kappa, k)$ is the kernel accounting for the effect of the finite survey area, given in Eqs. (9-11) of Hütsi (2010), and P_{NL} contains the remaining corrections and the effect of non-linearities. In our analysis we only use data up to $k_{\text{max}} = 0.15 h \text{ Mpc}^{-1}$ and we follow Hütsi (2010), modelling the effect of residual weak non-linearities

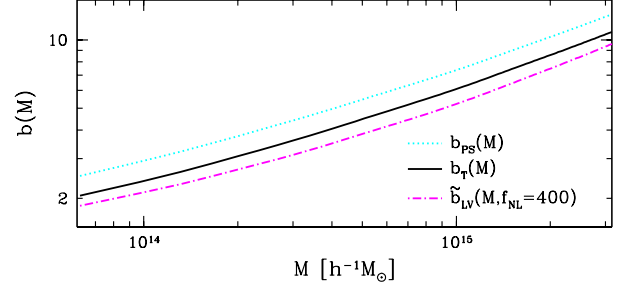


Figure 4. Mass dependence of the linear halo bias at $\bar{z} = 0.2$ for three mass functions: Press-Schechter (cyan dotted), Tinker (black solid) and modified LoVerde mass function in the presence of PNG (magenta dot-dashed), with $f_{\text{NL}} = 400$. Cosmology is fixed to our combined best-fit model.

with a simple effective fitting function with one free parameter q_{NL} . All these contributions lead to:

$$P_{\text{NL}}(k) = (b^{\text{obs}})^2 (1 + q_{\text{NL}} k^{3/2}) s(k) P_{\text{lin}}(k) \left[1 + \frac{2}{3} \beta_z + \frac{1}{5} \beta_z^2 \right]. \quad (20)$$

Here the bias is rescaled as $b^{\text{obs}} = \bar{b} \cdot B$, where we include a nuisance parameter B to represent the uncertainty on the bias derived from the mass function. We model the photo- z smoothing with a corrective factor

$$s(k) = \left(\frac{\sqrt{\pi}}{2 \sigma_z k} \right) \text{erf}(\sigma_z k), \quad (21)$$

assuming that photo- z errors follow a Gaussian distribution with dispersion δz and corresponding spatial smoothing scale $\sigma_z = \delta z \cdot c/H_0$. The last term of Eq. (20) is the correction due to redshift space distortions (RSD), for which we assume $\beta_z(\bar{z}) \simeq \Omega_m^{0.55}(\bar{z})/b^{\text{obs}}(\bar{z})$ (Kaiser 1987). We have checked that the RSD correction changes at most at the percent level if we calculate it at the limits of our redshift range. We finally take into account the Alcock-Paczynski effect (Alcock & Paczynski 1979): we rescale the full theoretical power spectrum with respect to the cosmology used to convert redshifts to distances in the measurements (denoted by the superscript ‘fid’), assuming a single isotropic dilation applies (Eisenstein et al. 2005; Hütsi 2006c), i.e.

$$\tilde{P}^{\text{obs}}(k) = \frac{1}{c_{\text{isotr}}^3} P^{\text{obs}}\left(\frac{k}{c_{\text{isotr}}}\right). \quad (22)$$

Here $c_{\text{isotr}} = (c_{\parallel} c_{\perp}^2)^{1/3}$, $c_{\parallel} = H^{\text{fid}}/H$, $c_{\perp} = D_A/D_A^{\text{fid}}$, D_A is the angular diameter distance, where all quantities are calculated at the mean redshift \bar{z} .

We show in Fig. 2 the full cluster power spectrum $\tilde{P}^{\text{obs}}(k)$. The different lines correspond to the theory curves for our combined best-fit cosmology (red solid) and for two other models (blue and green) chosen to be at the 2σ limit of the marginalised $\Omega_m - \sigma_8$ contour, compared with the data.

3.5 Primordial non-Gaussianity

We extend our model to constrain PNG from both bias and abundances of the maxBCG clusters. Briefly, following e.g. the notation of Giannantonio & Porciani (2010), we introduce the f_{NL} parameter to quantify the amount of PNG in the simplest local, scale-independent case as

$$\Phi(\mathbf{x}, z_*) = \varphi(\mathbf{x}, z_*) + f_{\text{NL}} \left[\varphi^2(\mathbf{x}, z_*) - \langle \varphi^2 \rangle(z_*) \right], \quad (23)$$

where Φ is the Bardeen's potential at a primordial redshift z_* and φ is an auxiliary Gaussian potential. Throughout this paper, we define f_{NL} by writing the previous equation at early times (i.e. $z_* \approx 1100$). The potential power spectrum can be approximated by its Gaussian part, $P_\Phi(k) \approx P_\varphi(k)$, at leading order in f_{NL} and neglecting trispectrum corrections. The matter perturbations are related to the primordial potential by the Poisson equation:

$$\tilde{\delta}(\mathbf{k}, z) = \alpha(\mathbf{k}, z) \tilde{\Phi}(\mathbf{k}, z_*), \quad (24)$$

with

$$\alpha(k, z) = \frac{2c^2 k^2 T(k) D(z) g(0)}{3\Omega_m H_0^2 g(z_*)}. \quad (25)$$

Here $T(k)$ is the transfer function and $g(z) \propto (1+z)D(z)$ is the growth function of the potential. We can then write for the matter power spectrum P

$$P(k, z) = \alpha^2(k, z) P_\Phi(k, z_*) \simeq \alpha^2(k, z) P_\varphi(k, z_*); \quad (26)$$

we consider linear theory only, so we assume $P = P_{\text{lin}}$.

The halo mass function is modified in the presence of PNG as it gains a dependence on the skewness. We use the LoVerde et al. (2008) mass function (LV), which was obtained by using the Edgeworth expansion, and is given by

$$f_{\text{LV}}(\nu) = \sqrt{\frac{2}{\pi}} e^{-\frac{\nu^2}{2}} \left[\nu + S_3 \frac{\sigma}{6} (\nu^4 - 2\nu^2 - 1) + \frac{dS_3}{d \ln \sigma} \frac{\sigma}{6} (\nu^2 - 1) \right], \quad (27)$$

where S_3 is the skewness of the matter density field defined as in Desjacques et al. (2009) (the mass dependence is implicit). To improve the agreement with N -body simulations, and for consistency with the rest of our analysis, we replace its Gaussian limit from the PS to the Tinker form, so that we use the rescaled form defined as:

$$\tilde{f}_{\text{LV}}(\nu) \equiv \frac{f_{\text{T}}(\nu)}{f_{\text{PS}}(\nu)} f_{\text{LV}}(\nu), \quad (28)$$

where $f_{\text{PS}} = \sqrt{2/\pi} \nu \exp(-\nu^2/2)$ is the PS mass function.

We apply the peak-background split formalism and analytically derive the Lagrangian linear halo bias associated to the LoVerde mass function using Eq. (15) as

$$b_{\text{LV}}^L(\nu) = \frac{\delta_c}{\sigma^2} - \frac{1}{\sigma} \frac{6 + S_3 \sigma (4\nu^3 - 4\nu) + 2 \frac{dS_3}{d \ln \sigma} \sigma \nu}{6\nu + S_3 \sigma (\nu^4 - 2\nu^2 - 1) + \frac{dS_3}{d \ln \sigma} \sigma (\nu^2 - 1)}, \quad (29)$$

while the Lagrangian bias associated to the rescaled mass function of Eq. (28) is

$$\tilde{b}_{\text{LV}}^L(\nu) = b_{\text{LV}}^L(\nu) + b_{\text{T}}^L(\nu) - b_{\text{PS}}^L(\nu), \quad (30)$$

which is the bias we use in the following.

In the presence of PNG, the halo density perturbations depend not only on the dark matter perturbations δ , but also on the potential φ . The latter can then be related back to the density in Fourier space by using the Poisson equation, so that the effective Eulerian bias can be written at the mean redshift $\bar{z} \approx 0.2$ as

$$b_{\text{eff}}(M, k, f_{\text{NL}}) = b(M, f_{\text{NL}}) + \Delta b(M, k, f_{\text{NL}}), \quad (31)$$

where the bias contains implicitly a scale-independent correction $\delta b(f_{\text{NL}}) \equiv b(M, f_{\text{NL}}) - b(M, 0)$ with respect to the Gaussian case, following from the difference in the mass function, and the scale-dependent part is

$$\Delta b(M, k, f_{\text{NL}}) = \frac{2 f_{\text{NL}} \delta_c b^L(M, f_{\text{NL}})}{\alpha(k, \bar{z})}. \quad (32)$$

As in the Gaussian case, we average the bias over the masses

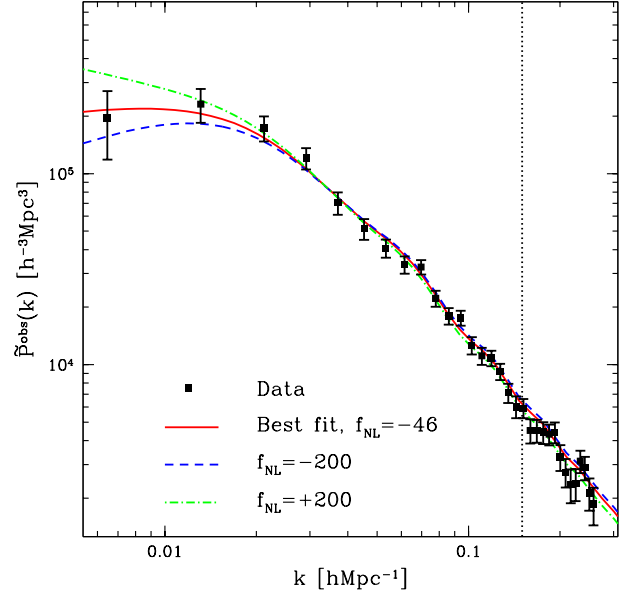


Figure 5. The effect of PNG on the cluster power spectrum. We compare the data (black points) with the predictions for the best-fit model to our data with $f_{\text{NL}} = -46$ (red solid) and for two cases with $f_{\text{NL}} = -200$ (blue dashed) and $f_{\text{NL}} = 200$ (green dot-dashed). The dotted line at $k_{\text{max}} = 0.15 h \text{ Mpc}^{-1}$ represents the smallest scale we use in the analysis.

in our catalogue following Eq. (17). In order to take into account the uncertainty on our assumption of a mass function, we also introduce a nuisance parameter B as in Section 3.4, which rescales the bias as $b^{\text{obs}} = \bar{b} \cdot B$.

The scale-independent correction $\delta b(f_{\text{NL}})$ is small, easily confused with other normalisation effects, and relies on the assumed form of the mass function and the peak-background split method. For these reasons, it is worth ensuring that the results do not depend on this contribution. We make sure this happens in our case because any constant rescaling of the bias can be equally explained by either a change in the nuisance parameter B or a change in f_{NL} . But since a model with $f_{\text{NL}} \neq 0$ also predicts the scale-dependent bias, it will be favoured only in case such a feature is indeed observed in the data, otherwise the $B \neq 1$ model will be assigned a better likelihood. In practice, we impose some Gaussian priors centred on $B = 1$, but we have checked that the results on f_{NL} do not depend significantly on this choice.

We show in Fig. 5 the full power spectrum $\tilde{P}^{\text{obs}}(k)$ in the presence of PNG for a choice of f_{NL} values, compared with the data. The scale-dependent bias induced by PNG is visible on large scales (small k), while the smaller scale-independent contribution can be seen on small scales (large k). Note that the survey window convolution of Eq. (19) partially suppresses the effect of PNG on the largest scales, which become comparable with the survey volume.

4 LIKELIHOOD ANALYSIS AND RESULTS

We use cluster counts, WL masses and the cluster power spectrum to fit the richness-mass relation and constrain cosmology simultaneously. In particular, our observables are:

- (1) Cluster counts ΔN , divided into 10 richness bins;

Type	Symbol	Definition	Prior without CMB	Prior with CMB
Cosmology	h	Dimensionless Hubble parameter	0.7	[0.4, 0.9]
	n_s	Scalar spectral index	0.96	[0.5, 1.5]
	Ω_b	Baryon energy density	0.04397	[0.01, 0.2]
	Ω_c	Cold dark matter energy density	[0.1, 0.9]	[0.1, 0.9]
	$\log(10^{10} A_s)$	Amplitude of primordial perturbations	[0.1, 6.0]	[0.1, 6.0]
	τ	Optical depth	0.09	[0.01, 0.125]
	f_{NL}	Primordial non-Gaussianity amplitude	[-900, 900]	[-900, 900]
Scaling relation	$\ln N_1 \equiv \ln N_{\text{gal}} M_1$	Richness at $M_1 = 1.3 \cdot 10^{14} M_{\odot}$	[1.0, 4.0]	[1.0, 4.0]
	$\ln N_2 \equiv \ln N_{\text{gal}} M_2$	Richness at $M_2 = 1.3 \cdot 10^{15} M_{\odot}$	[1.0, 4.0]	[1.0, 4.0]
	$\sigma_{\ln M N_{\text{gal}}}$	Scatter	0.45 ± 0.1	0.45 ± 0.1
Nuisance	β	Weak lensing mass measurements bias	1.0 ± 0.06	1.0 ± 0.06
	B	Scatter on bias derived from mass function	1.0 ± 0.15	1.0 ± 0.15
	q_{NL}	Non-linear correction to power spectrum	[0.0, 50.0]	[0.0, 50.0]
	σ_z	Photo-Z errors	[0, 120]	[0, 120]
	A_{SZ}	Amplitude of CMB SZ template	1	[0, 2]
Derived	Ω_m	Total matter energy density	—	—
	σ_8	Amplitude of density perturbations	—	—

Table 1. Parameters used in the analysis and their assumed priors. In the prior columns a single number n stands for a fixed value, $[a, b]$ stands for a flat prior, $\mu \pm \sigma$ means a Gaussian prior of mean μ and standard deviation σ .

- (2) Total mass of clusters $\Delta N \bar{M}$, divided into 6 richness bins;
- (3) Cluster power spectrum \bar{P}^{obs} , divided into 18 k bins.

The covariance matrix we use for the cosmological analysis is composed by the parts discussed in Section 2. In addition to the cluster data we also use the CMB power spectra from WMAP7 (Larson et al. 2011), in the cases specified below.

We assume a flat Λ CDM cosmological model. When using cluster data alone we fix the Hubble parameter $h = 0.7$, primordial spectral index $n_s = 0.96$ and baryon density $\Omega_b = 0.044$, as these parameters are not easily constrained in this case; we relax these assumptions when adding external CMB data. Note that we need to fix the spectral index of scalar density perturbations because of the small range in scale which our mass range corresponds to.

We then perform Bayesian parameter estimation by running Monte Carlo Markov Chains (MCMCs), using Metropolis sampling with a modified version of the COSMOMC code (Lewis & Bridle 2002). In Table 1 we list all the parameters of the analysis, including their assumed priors. We estimate the posterior probability distributions in the following cases:

- (i) **Counts only:** 6 free parameters [Ω_c , $\log(10^{10} A_s)$, $\ln N_1$, $\ln N_2$, $\sigma_{\ln M|N_{\text{gal}}}$, β], without the cluster power spectrum;
- (ii) **Counts with f_{NL} :** 7 free parameters [Ω_c , $\log(10^{10} A_s)$, $\ln N_1$, $\ln N_2$, $\sigma_{\ln M|N_{\text{gal}}}$, β , f_{NL}], without the cluster power spectrum;
- (iii) **Counts+ $P(k)$:** 9 free parameters [Ω_c , $\log(10^{10} A_s)$, $\ln N_1$, $\ln N_2$, $\sigma_{\ln M|N_{\text{gal}}}$, β , q_{NL} , σ_z , B], with the cluster power spectrum;
- (iv) **Counts+ $P(k)$ with f_{NL} :** 10 free parameters [Ω_c , $\log(10^{10} A_s)$, $\ln N_1$, $\ln N_2$, $\sigma_{\ln M|N_{\text{gal}}}$, β , q_{NL} , σ_z , B , f_{NL}], with the cluster power spectrum;
- (v) **CMB only:** 7 free parameters [Ω_b , h , τ , n_s , A_{SZ} , Ω_c , $\log(10^{10} A_s)$], with CMB data only;
- (vi) **CMB+clusters:** 14 free parameters [Ω_b , h , τ , n_s , A_{SZ} , Ω_c , $\log(10^{10} A_s)$, $\ln N_1$, $\ln N_2$, $\sigma_{\ln M|N_{\text{gal}}}$, β , q_{NL} , σ_z , B], with CMB and all cluster data;
- (vii) **CMB+clusters with f_{NL} :** 15 free parameters [Ω_b , h , τ , n_s , A_{SZ} , Ω_c , $\log(10^{10} A_s)$, $\ln N_1$, $\ln N_2$, $\sigma_{\ln M|N_{\text{gal}}}$, β , q_{NL} , σ_z , B , f_{NL}], with CMB and all cluster data.

4.1 Results

We summarise our results in Table 2, and we show in Figs. 6, 7, 8 and 9 the 2D 68% and 95% marginalised confidence regions for different pairs of parameters in our analysis. The colour scheme is the same for all figures: blue contours refer to runs with counts and WL mean masses data only, green contours include in addition the cluster power spectrum data, while orange contours also include CMB data from WMAP7.

The joint constraint in the $\Omega_m - \sigma_8$ plane in Fig. 6 displays the typical degeneracy from cluster counts: the counts increase with increasing Ω_m and σ_8 values, hence any increase in Ω_m must be balanced by a decrease in σ_8 (and viceversa), to keep the abundances at the observed values. The constraints on individual parameters with counts and masses only are consistent with Rozo et al. (2010), and we find $\Omega_m = 0.25 \pm 0.06$, $\sigma_8 = 0.80 \pm 0.06$ (1σ errors throughout), while the errors are improved by a factor between 1.5 and 3, depending on the parameter, when adding the maxBCG power spectrum: in this case we obtain $\Omega_m = 0.215 \pm 0.022$, $\sigma_8 = 0.84 \pm 0.04$. Combining then these results with the CMB data, the constraints shrink to $\Omega_m = 0.255 \pm 0.014$ and $\sigma_8 = 0.790 \pm 0.016$: the contribution of the CMB tightens the errors by a further factor of two.

In Fig. 7 we show the marginalised posterior probability contours of the scaling relation parameters $\ln N_1$ and $\ln N_2$. Constraints on individual parameters using counts and masses only are again compatible with Rozo et al. (2010) ($\ln N_1 = 2.44 \pm 0.11$, $\ln N_2 = 4.16 \pm 0.15$), while errors are reduced when adding the power spectrum, even if less significantly ($\ln N_1 = 2.49 \pm 0.09$, $\ln N_2 = 4.13 \pm 0.13$). Combining these results with the CMB data, the constraints are almost identical, as the CMB is not sensitive to the clusters scaling relation. Our constraints on the scaling relation scatter $\sigma_{\ln M}$ are in agreement with Rozo et al. (2010), and they are not improved by the addition of power spectrum and CMB data.

Figs. 8, 9 show the constraints on f_{NL} and its degeneracies with Ω_m and σ_8 . First we can see that, when only counts and masses are used, the constraints on f_{NL} are weak as expected. The situation improves when adding the cluster power spectrum: in this case, the constraints are tighter, and we observe a positive correlation between f_{NL} and Ω_m and an anti-correlation with σ_8 . To understand this behaviour, we notice that if we increase (decrease) Ω_m ,

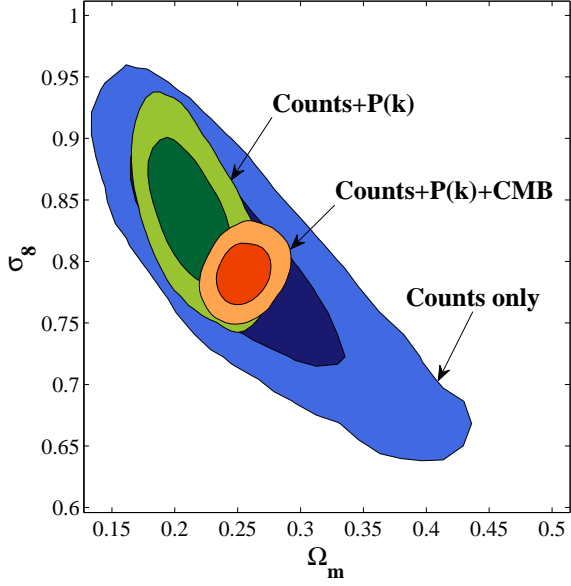


Figure 6. Marginalised posterior probability distributions on the parameters $\Omega_m - \sigma_8$ for the runs using **Counts only** (blue), **Counts+ $P(k)$** (green) and **Counts+ $P(k)$ +CMB** (orange), at 68% and 95% confidence levels.

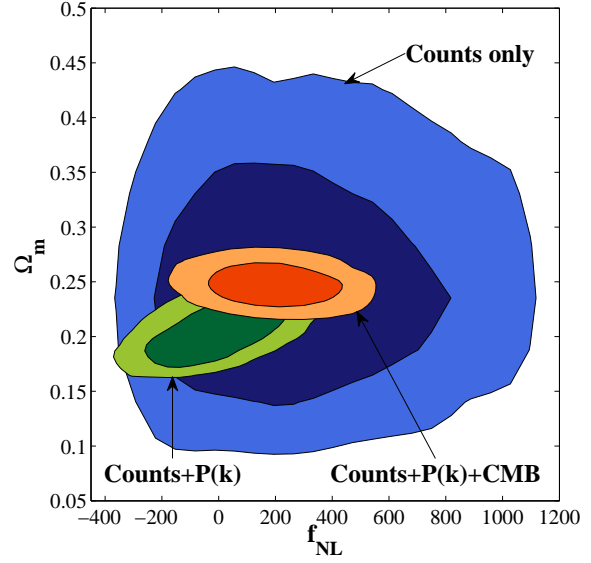


Figure 8. Marginalised posterior probability distributions on the parameters $f_{\text{NL}} - \Omega_m$ for the runs using **Counts with f_{NL}** (blue), **Counts+ $P(k)$ with f_{NL}** (green) and **CMB+clusters with f_{NL}** (orange), at 68% and 95% confidence levels.

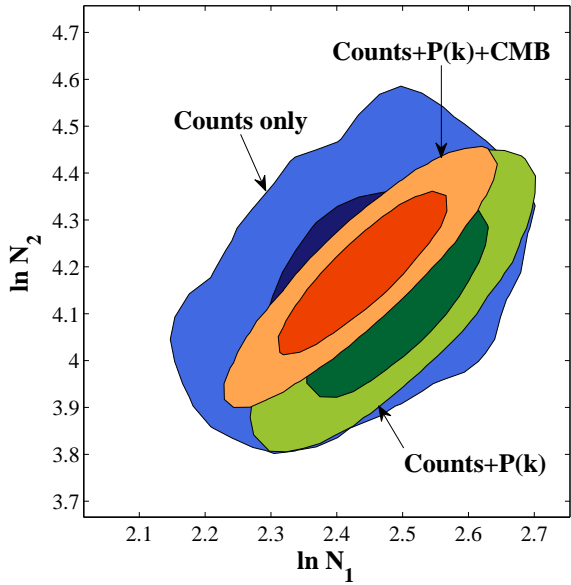


Figure 7. Constraints on the scaling relation parameters for the runs using **Counts only** (blue), **Counts+ $P(k)$** (green) and **Counts+ $P(k)$ +CMB** (orange), at 68% and 95% confidence levels. Notice that $\ln N_1 \equiv \ln N_{\text{gal}}|M_1$ and $\ln N_2 \equiv \ln N_{\text{gal}}|M_2$, where $M_1 = 1.3 \cdot 10^{14} M_{\odot}$ and $M_2 = 1.3 \cdot 10^{15} M_{\odot}$.

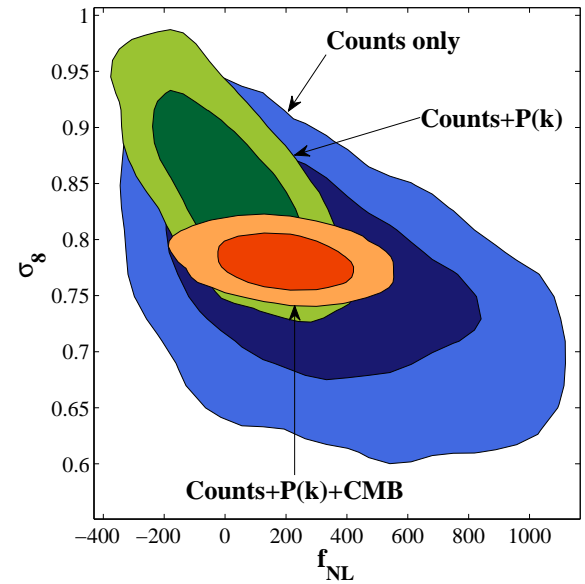


Figure 9. Marginalised posterior probability distributions on the $f_{\text{NL}} - \sigma_8$ plane for the runs including **Counts with f_{NL}** (blue), **Counts+ $P(k)$ with f_{NL}** (green) and **CMB+clusters with f_{NL}** (orange), at 68% and 95% confidence levels.

the peak of the power spectrum decreases (increases) while also being shifted to higher (lower) values of k , while σ_8 simply changes the overall normalisation. As described above, an increase in f_{NL} causes a boost in the power spectrum on large scales (small k), so that σ_8 needs to decrease to compensate a higher f_{NL} : this is exactly what is shown in Fig. 9. In addition to this Ω_m should increase to compensate a higher f_{NL} : this can be seen in Fig. 8. We also see that the addition of the CMB power spectrum data improves the constraints on Ω_m and σ_8 and only indirectly reduces the bounds

on f_{NL} , since PNG simply affects the higher-order statistics of the CMB.

Our constraints on PNG are $f_{\text{NL}} = 12 \pm 157$ (1σ) (without CMB) and $f_{\text{NL}} = 194 \pm 128$ (with CMB), which are statistically compatible with zero and with each other. The shift in the mean between the two results is clear by looking at Figs. 8, 9: the addition of the CMB favours lower values of σ_8 (and higher values of Ω_m), thus shifting the favoured f_{NL} values in the process. While not competitive with results from the CMB bispectrum or from combined

Params	Counts only		Counts+ $P(k)$		Clusters+CMB	
	no f_{NL}	+ f_{NL}	no f_{NL}	+ f_{NL}	no f_{NL}	+ f_{NL}
Ω_m	0.25 ± 0.06	0.25 ± 0.06	0.215 ± 0.022	0.209 ± 0.022	0.255 ± 0.014	0.248 ± 0.013
σ_8	0.80 ± 0.06	0.77 ± 0.07	0.84 ± 0.04	0.85 ± 0.05	0.790 ± 0.016	0.780 ± 0.016
$\ln N_1$	2.44 ± 0.11	2.44 ± 0.11	2.49 ± 0.09	2.49 ± 0.08	2.44 ± 0.08	2.43 ± 0.08
$\ln N_2$	4.16 ± 0.15	4.15 ± 0.15	4.13 ± 0.13	4.11 ± 0.12	4.19 ± 0.11	4.15 ± 0.11
$\sigma_{\ln M}$	0.38 ± 0.06	0.38 ± 0.06	0.36 ± 0.06	0.37 ± 0.06	0.378 ± 0.059	0.38 ± 0.06
β	1.00 ± 0.06	1.01 ± 0.06	1.01 ± 0.06	1.01 ± 0.06	1.01 ± 0.06	1.00 ± 0.06
q_{NL}	-	-	26 ± 10	27 ± 10	14 ± 6	16 ± 7
σ_z	-	-	46 ± 12	42 ± 8	43 ± 10	31 ± 5
B	-	-	1.07 ± 0.13	1.01 ± 0.15	1.19 ± 0.11	1.00 ± 0.14
f_{NL}	-	282 ± 317	-	12 ± 157	-	194 ± 128

Table 2. Marginalised mean values and 1σ errors on the cosmological parameters, for the runs **Counts only**, **Counts with f_{NL}** , **Counts+ $P(k)$** , **Counts+ $P(k)$ with f_{NL}** , **CMB+clusters** and **CMB+clusters with f_{NL}** . Note that Ω_m and σ_8 are derived parameters in our analysis.

analyses of multiple galaxy surveys, it is interesting to find such constraints independently and for the first time with the clustering of galaxy clusters.

Since we restrict our analysis to nearly-linear scales, by imposing the data cut at $k_{\text{max}} = 0.15 h \text{ Mpc}^{-1}$, we are not expecting strong constraints on q_{NL} . The constraints we found are indeed broad and in agreement within the errors with the results by Hütsi (2010), who found $q_{\text{NL}} = 14.2 \pm 2.8$ when marginalising over three parameters only: we obtain $q_{\text{NL}} = 26 \pm 10$ and $q_{\text{NL}} = 14 \pm 6$ when also using CMB data.

It is also worth mentioning the results on the B parameter, which was introduced to take into account the uncertainty in the bias expression derived from the mass function. As this parameter allows an arbitrary constant rescaling of the bias, it also has the desirable property of cancelling the effect of the scale-independent bias correction $\delta b(f_{\text{NL}})$, as described in Section 3.5. To check that the Gaussian prior we are imposing $B = 1.0 \pm 0.15$ is large enough for both purposes, we made an additional run replacing it with a flat prior $B \in [0.0001, 5]$. In this way, we obtain nearly unchanged results on f_{NL} .

5 CONCLUSIONS

In this work we have investigated the cosmological implications of the optically-selected SDSS maxBCG galaxy cluster data, obtaining extended cosmological constraints with respect to previous works. We considered the number counts of clusters in richness bins and the weak lensing mass estimations, including the respective covariances, for a cross-calibration of the scaling relation. We then combined such data for the first time with a measurement of the redshift-space power spectrum of the same clusters. In the modelling we included an effective treatment of the non-linear contribution, photo- z smoothing, redshift space distortions and Alcock-Paczynski effect. We only considered quasi-linear scales at $k < k_{\text{max}} = 0.15 h \text{ Mpc}^{-1}$ to be conservative. We estimated the off-diagonal terms of the counts-clustering covariance matrix with a jackknife method applied on both data and N -body simulations, and found consistently that such contributions are negligible.

We then performed a full MCMC analysis of the posterior probability distribution of cosmological parameters given the full data set. By thus combining the one- and two-point statistics, we achieved a factor 1.5 to 3 improvement on the errors on the cosmological parameters, if compared with previous analyses using

number counts and masses only (Roza et al. 2010), obtaining e.g. for the fluctuation amplitude $\sigma_8 = 0.84 \pm 0.04$ (1σ) and for the matter content $\Omega_m = 0.215 \pm 0.022$ (1σ). These are further tightened by a factor of 2 by the addition of the CMB data. On the other hand, we found that the errors on the scaling relation parameters are consistent with previous works, but not significantly improved by the addition of the cluster power spectrum.

As an interesting application, we also tested primordial non-Gaussianity, which is constrained through the non-Gaussian halo mass function and the scale-dependent cluster bias. Assuming deviations from Gaussianity at the three-point (skewness) level of the local type, we obtained $f_{\text{NL}} = 12 \pm 157$ (1σ) from our combined data set, which shifts to $f_{\text{NL}} = 194 \pm 128$ (1σ) when including the WMAP7 CMB data. While not competitive with the CMB bispectrum and with results from combined galaxy clustering data sets, this result is consistent with them and was obtained using the maxBCG cluster data alone. Our results can be seen as a proof of concept towards a full joint analysis of the LSS, consistently including both galaxies and clusters as dark matter tracers, to achieve the full potential of the upcoming galaxy surveys such as the Dark Energy Survey and the Euclid mission.

ACKNOWLEDGMENTS

We are grateful to Ixandra Achitouv, Richard Battye and Eduardo Roza for useful discussions. We acknowledge support from the Trans-Regional Collaborative Research Centre TRR 33 – ‘‘The Dark Universe’’.

REFERENCES

- Achitouv I. E., Corasaniti P. S., 2012a, JCAP, 2, 2
- Achitouv I. E., Corasaniti P. S., 2012b, Phys. Rev. D, 86, 8, 083011
- Acquaviva V., Bartolo N., Matarrese S., Riotto A., 2003, Nuclear Physics B, 667, 119
- Afshordi N., Tolley A. J., 2008, Phys. Rev. D, 78, 12, 123507
- Alcock C., Paczynski B., 1979, Nature, 281, 358
- Allen S. W., Evrard A. E., Mantz A. B., 2011, ARA&A, 49, 409
- Allen S. W., Rapetti D. A., Schmidt R. W., Ebeling H., Morris R. G., Fabian A. C., 2008, MNRAS, 383, 879
- Bahcall N. A., Fan X., Cen R., 1997, ApJ, 485, L53

- Balaguera-Antolínez A., Sánchez A. G., Böhringer H., Collins C., Guzzo L., Phleps S., 2011, *MNRAS*, 413, 386
- Bardeen J. M., Bond J. R., Kaiser N., Szalay A. S., 1986, *ApJ*, 304, 15
- Bartolo N., Komatsu E., Matarrese S., Riotto A., 2004, *Phys. Rep.*, 402, 103
- Battye R. A., Weller J., 2003, *Phys. Rev. D*, 68, 8, 083506
- Bennett C. L., et al., 2012, arXiv:1212.5225
- Böhringer H., Schuecker P., Guzzo L., et al., 2004, *A&A*, 425, 367
- Bond J. R., Cole S., Efstathiou G., Kaiser N., 1991, *ApJ*, 379, 440
- Borgani S., Kravtsov A., 2009, arXiv:0906.4370
- Burenin R. A., Vikhlinin A., Hornstrup A., Ebeling H., Quintana H., Mescheryakov A., 2007, *ApJS*, 172, 561
- Chen X., 2010, *Advances in Astronomy*, 2010
- Cole S., Kaiser N., 1989, *MNRAS*, 237, 1127
- Corasaniti P. S., Achitouv I., 2011, *Physical Review Letters*, 106, 24, 241302
- Dalal N., Doré O., Huterer D., Shirokov A., 2008, *Phys. Rev. D*, 77, 12, 123514
- Desjacques V., Jeong D., Schmidt F., 2011, *Phys. Rev. D*, 84, 6, 061301
- Desjacques V., Seljak U., 2010, *Advances in Astronomy*, 2010
- Desjacques V., Seljak U., Iliev I. T., 2009, *MNRAS*, 396, 85
- Ebeling H., Edge A. C., Allen S. W., Crawford C. S., Fabian A. C., Huchra J. P., 2000, *MNRAS*, 318, 333
- Ebeling H., Edge A. C., Böhringer H., et al., 1998, *MNRAS*, 301, 881
- Efron B., 1982, *CBMS-NSF conference series in applied mathematics, Soc. Ind. and App. Math., Philadelphia*
- Eisenstein D. J., et al., 2005, *ApJ*, 633, 560
- Estrada J., Sefusatti E., Frieman J. A., 2009, *ApJ*, 692, 265
- Evrard A. E., 1989, *ApJ*, 341, L71
- Fassbender R., 2008, arXiv:0806.0861
- Feldman H. A., Kaiser N., Peacock J. A., 1994, *ApJ*, 426, 23
- Frenk C. S., White S. D. M., Efstathiou G., Davis M., 1990, *ApJ*, 351, 10
- Fry J. N., Gaztanaga E., 1993, *ApJ*, 413, 447
- Giannantonio T., et al., 2013, In prep.
- Giannantonio T., Porciani C., 2010, *Phys. Rev. D*, 81, 6, 063530
- Giannantonio T., Porciani C., Carron J., Amara A., Pillepich A., 2012, *MNRAS*, 422, 2854
- Gladders M. D., Yee H. K. C., 2005, *ApJS*, 157, 1
- Górski K. M., Hivon E., Banday A. J., et al., 2005, *ApJ*, 622, 759
- Guo Q., White S., Boylan-Kolchin M., et al., 2011, *MNRAS*, 413, 101
- Haiman Z., Mohr J. J., Holder G. P., 2001, *ApJ*, 553, 545
- Hao J., McKay T. A., Koester B. P., et al., 2010, *ApJS*, 191, 254
- Henriques B. M. B., White S. D. M., Lemson G., et al., 2012, *MNRAS*, 421, 2904
- Hinshaw G., et al., 2012, arXiv:1212.5226
- Hockney R. W., Eastwood J. W., 1988, *Bristol, Hilger*, 1988
- Hong T., Han J. L., Wen Z. L., Sun L., Zhan H., 2012, *The Astrophysical Journal*, 749, 1, 81
- Hotchkiss S., 2011, *JCAP*, 7, 4
- Hoyle B., Jimenez R., Verde L., Hotchkiss S., 2012, *JCAP*, 2, 9
- Hu W., Kravtsov A. V., 2003, *ApJ*, 584, 702
- Hütsi G., 2006a, *A&A*, 449, 891
- Hütsi G., 2006b, *A&A*, 446, 43
- Hütsi G., 2006c, *A&A*, 459, 375
- Hütsi G., 2010, *MNRAS*, 401, 2477
- Jenkins A., Frenk C. S., White S. D. M., et al., 2001, *MNRAS*, 321, 372
- Jing Y. P., 2005, *ApJ*, 620, 559
- Johnston D. E., Sheldon E. S., Wechsler R. H., et al., 2007, arXiv:0709.1159
- Kaiser N., 1987, *MNRAS*, 227, 1
- Koester B. P., McKay T. A., Annis J., et al., 2007a, *ApJ*, 660, 239
- Koester B. P., McKay T. A., Annis J., et al., 2007b, *ApJ*, 660, 221
- Landy S. D., Szalay A. S., 1993, *ApJ*, 412, 64
- Larson D., et al., 2011, *ApJS*, 192, 16
- Lemson G., Virgo Consortium t., 2006, arXiv:astro-ph/0608019
- Lewis A., Bridle S., 2002, *Phys. Rev. D*, 66, 10, 103511
- Lima M., Hu W., 2005, *Phys. Rev. D*, 72, 4, 043006
- LoVerde M., Miller A., Shandera S., Verde L., 2008, *JCAP*, 4, 14
- Lyth D. H., Liddle A. R., 2009, *The Primordial Density Perturbation*
- Maggiore M., Riotto A., 2010, *ApJ*, 711, 907
- Mahdavi A., Hoekstra H., Babul A., Sievers J., Myers S. T., Henry J. P., 2007, *ApJ*, 664, 162
- Majumdar S., Mohr J. J., 2003, *ApJ*, 585, 603
- Maldacena J., 2003, *Journal of High Energy Physics*, 5, 13
- Mandelbaum R., Seljak U., Hirata C. M., 2008a, *JCAP*, 8, 6
- Mandelbaum R., Seljak U., Hirata C. M., et al., 2008b, *MNRAS*, 386, 781
- Mantz A., Allen S. W., Ebeling H., Rapetti D., 2008, *MNRAS*, 387, 1179
- Mantz A., Allen S. W., Rapetti D., Ebeling H., 2010, *MNRAS*, 406, 1759
- Matarrese S., Verde L., 2008, *ApJ*, 677, L77
- Matarrese S., Verde L., Jimenez R., 2000, *ApJ*, 541, 10
- Mehrtens N., Romer A. K., Hilton M., et al., 2012, *MNRAS*, 423, 1024
- Menanteau F., Sifón C., Barrientos L. F., et al., 2012, arXiv:1210.4048
- Mo H. J., Jing Y. P., White S. D. M., 1996, *MNRAS*, 282, 1096
- Mo H. J., White S. D. M., 1996, *MNRAS*, 282, 347
- Navarro J. F., Frenk C. S., White S. D. M., 1997, *ApJ*, 490, 493
- Oguri M., 2009, arXiv:0905.0920, 102, 21, 211301
- Pillepich A., Porciani C., Hahn O., 2010, *MNRAS*, 402, 191
- Planck Collaboration, Ade P. A. R., Aghanim N., et al., 2011, *A&A*, 536, A8
- Press W. H., Schechter P., 1974, *ApJ*, 187, 425
- Reichardt C. L., Stalder B., Bleem L. E., et al., 2012, arXiv:1203.5775
- Ross A. J., et al., 2013, *MNRAS*, 428, 1116
- Ross N. P., Shen Y., Strauss M. A., et al., 2009, *ApJ*, 697, 1634
- Rozo E., Rykoff E. S., Evrard A., et al., 2009, *ApJ*, 699, 768
- Rozo E., Wechsler R. H., Rykoff E. S., et al., 2010, *ApJ*, 708, 645
- Sartoris B., Borgani S., Fedeli C., et al., 2010, *MNRAS*, 407, 2339
- Schmidt F., Kamionkowski M., 2010, *Phys. Rev. D*, 82, 10, 103002
- Sefusatti E., Crocce M., Desjacques V., 2012, *MNRAS*, 425, 2903
- Sheldon E. S., Johnston D. E., Scranton R., et al., 2009, *ApJ*, 703, 2217
- Sheth R. K., Tormen G., 1999, *MNRAS*, 308, 119
- Slosar A., Hirata C., Seljak U., Ho S., Padmanabhan N., 2008, *JCAP*, 8, 31
- Springel V., White S. D. M., Jenkins A., et al., 2005, *Nature*, 435, 629
- Suyama T., Takahashi T., Yamaguchi M., Yokoyama S., 2010, *JCAP*, 12, 30
- Tinker J., Kravtsov A. V., Klypin A., et al., 2008, *ApJ*, 688, 709

- Tinker J. L., Robertson B. E., Kravtsov A. V., et al., 2010, *ApJ*, 724, 878
- Vikhlinin A., Burenin R. A., Ebeling H., et al., 2009a, *ApJ*, 692, 1033
- Vikhlinin A., Kravtsov A. V., Burenin R. A., et al., 2009b, *ApJ*, 692, 1060
- Voges W., Aschenbach B., Boller T., et al., 1999, *A&A*, 349, 389
- Xia J.-Q., Baccigalupi C., Matarrese S., Verde L., Viel M., 2011, *JCAP*, 8, 33
- Xia J.-Q., Bonaldi A., Baccigalupi C., et al., 2010a, *JCAP*, 8, 13
- Xia J.-Q., Viel M., Baccigalupi C., De Zotti G., Matarrese S., Verde L., 2010b, *ApJ*, 717, L17
- Zu Y., Weinberg D. H., Rozo E., Sheldon E. S., Tinker J. L., Becker M. R., 2012, arXiv:1207.3794