FERMILAB-PUB-06-083-A

Cosmology and the Bispectrum

Emiliano Sefusatti[1,2], Martín Crocce[2], Sebastián Pueblas[2], Román Scoccimarro[2]
[1] *Particle Astrophysics Center, Fermi National Accelerator Laboratory, Batavia, IL 60510-0500 and*
[2] *Center for Cosmology and Particle Physics, Department of Physics, New York University, New York, NY 10003*

The present spatial distribution of galaxies in the Universe is non-Gaussian, with 40% skewness in $50\,h^{-1}$ Mpc spheres, and remarkably little is known about the information encoded in it about cosmological parameters beyond the power spectrum. In this work we present an attempt to bridge this gap by studying the bispectrum, paying particular attention to a joint analysis with the power spectrum and their combination with CMB data. We address the covariance properties of the power spectrum and bispectrum including the effects of beat coupling that lead to interesting cross-correlations, and discuss how baryon acoustic oscillations break degeneracies. We show that the bispectrum has significant information on cosmological parameters well beyond its power in constraining galaxy bias, and when combined with the power spectrum is more complementary than combining power spectra of different samples of galaxies, since non-Gaussianity provides a somewhat different direction in parameter space. In the framework of flat cosmological models we show that most of the improvement of adding bispectrum information corresponds to parameters related to the amplitude and effective spectral index of perturbations, which can be improved by almost a factor of two. Moreover, we demonstrate that the expected statistical uncertainties in σ_8 of a few percent are robust to relaxing the dark energy beyond a cosmological constant.

I. INTRODUCTION

Several recent studies have stressed the importance of combining different observations to constrain cosmological parameters. A clear example is provided by the analysis of the galaxy power spectrum in the Sloan Digital Sky Survey (SDSS) [1], and in the 2dF Galaxy Redshift Survey [2], which have shown the central role played by the information contained in the large-scale galaxy distribution to break the degeneracies still present in the cosmic microwave background (CMB) data despite the precision of the WMAP satellite observations [3, 4].

One of the main challenges in extracting cosmological information from galaxy clustering is knowing how good tracers of the underlying mass distribution galaxies are. This is often bypassed altogether, for example in [1, 2] only infomation on the *shape* of the galaxy power spectrum was used, since its amplitude is degenerate with the linear bias parameter relating galaxy to dark matter fluctuations at large scales.

The determination of galaxy bias has been, so far, among the main reasons of interest in the galaxy higher-order statistics in general [5–15] and the bispectrum in particular [16–20]. At large scales, the dependence on triangle configuration of the bispectrum generated by gravitational instability allows to disentagle the gravitational contribution from the bispectrum generated by non-linear biasing and ultimately remove the degeneracy between the linear bias and the amplitude of the dark matter perturbations. In weak gravitational lensing at smaller scales, the bispectrum can similarly be used to break degeneracies between matter content and the amplitude of fluctuations and probe dark energy [21–24].

Observational applications of this method to the galaxy distribution in the past involved fixing the cosmological model [14, 17–19], thus the information of the bispectrum on cosmological parameters has not been properly taken advantage of. In this work we study the constraining power of the bispectrum as a tool in the determination of cosmological parameters and the nature of primordial fluctuations, going beyond the determination of galaxy bias alone. As shown in [59], higher-order correlation functions such as the bispectrum or the trispectrum in galaxy surveys show, when all measurable configurations are taken into account, a signal-to-noise ratio comparable or even exceeding the signal-to-noise of the power spectrum at mildly non-linear scales.

Figure 1 illustrates this point. Here we show constraints on flat cosmological models depending on nine parameters: the physical dark matter density $\omega_d = \Omega_d h^2$, the physical baryon density $\omega_b = \Omega_b h^2$, the dark energy density Ω_Λ, the amplitude of scalar fluctuations A_s, the scalar spectral index n_s, the dark energy equation of state parameter w, the optical depth to Thomson scattering τ, plus the linear (b_1) and quadratic (b_2) galaxy bias parameters. We also show "derived" parameters such as h, the Hubble constant in units of $100\,\mathrm{km\ s^{-1}\,Mpc^{-1}}$, the baryon density Ω_b and the amplitude of dark matter fluctuations at $8\,h^{-1}\,\mathrm{Mpc}$, σ_8.

These constraints are from a hypothetical analysis that combines the CMB data from WMAP (first year) with measurements in the North part of SDSS by the end of the survey in two cases: using the SDSS galaxy power spectrum (blue, dashed line) and *replacing* the SDSS power spectrum by the bispectrum (red, continuous line). Both cases include the covariance between different power spectrum bins or bispectrum configurations (see below for a full discussion). Figure 1 shows that when all triangle configurations are included down to wavenumber $k_{\max} = 0.3\,h\,\mathrm{Mpc}^{-1}$ the bispectrum even improves the power spectrum results.

In practice one would like to combine the information in the power spectrum and bispectrum, which requires a calculation of their covariance properties. This is the

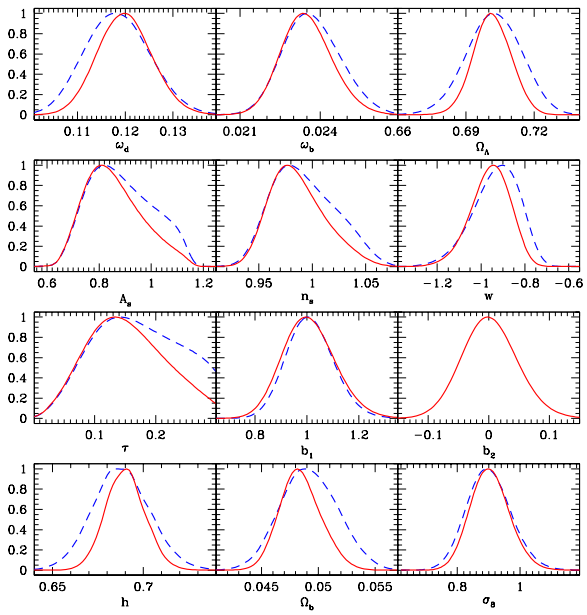

FIG. 1: An example showing the constraining power of the bispectrum compared to the power spectrum. The panels show marginalized likelihood functions corresponding to a hypothetical joint analysis of WMAP (first year) and SDSS North (by the end of the survey) where only the galaxy power spectrum is used (blue, dashed line), or only the galaxy bispectrum is used (red, solid line). Assumes a flat cosmology and scales up to $k_{\rm max} = 0.3\,h\,{\rm Mpc}^{-1}$.

main subject of the present work. The cross-covariance between the power spectrum and bispectrum turns out to have some non-trivial properties that help constraining cosmological parameters, in a way that is unexpected from a separate consideration of the covariance of each statistic by itself.

Although the main properties of the covariance matrices can be understood analytically, the details of the survey under consideration are important in practice, thus we compute covariance matrices from mock catalogs designed to reproduce the geometry of the SDSS survey by its completion. In particular, we consider the power spectrum and bispectrum of the north hemisphere main sample (MS) of galaxies. We also discuss how our constrains change as we include the power spectrum of the Luminous Red Galaxies (LRG) sample in the same geometry. As an example of what should be expected in combining large-scale structure (LSS) with the CMB we use the WMAP first year data; as this paper was being finished the third-year data became available [4]. Using the newer data set instead would of course improve the forecasted constraints.

This paper is organized as follows. In section II we review some basic results regarding the large-scale bispectrum of the galaxy distribution and discuss the main features of the covariance measured from our mock catalogues in section III. In section IV we present the likelihood functions both for the LSS and CMB correlations. In section V we present results on expected constraints on cosmological parameters and we conclude in section VI.

II. PREDICTIONS AND MOCK CATALOGS

We will assume primordial fluctuations to be Gaussian, so that every connected higher-order correlation function in the dark matter overdensity field δ results from gravitational instability. The dark matter bispectrum, *i.e.* the Fourier counterpart of the 3-point correlation function, is defined as

$$\langle \delta_{\mathbf{k}_1} \delta_{\mathbf{k}_2} \delta_{\mathbf{k}_3} \rangle \equiv \delta_D(\mathbf{k}_{123})\; B(k_1, k_2, k_3), \qquad (1)$$

with $\delta_{\mathbf{k}}$ the density contrast in Fourier space and $\mathbf{k}_{123} \equiv \mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3$. If the bispectrum can be reliably predicted by tree-level perturbation theory (PT), it follows that [25]

$$B(k_1, k_2, k_3) = 2F_2(\mathbf{k}_1, \mathbf{k}_2) P_1 P_2 + {\rm cyc.}, \qquad (2)$$

where $P_1 \equiv P(k_1)$ is the linear power spectrum while the kernel F_2 reads

$$F_2(\mathbf{k}_1, \mathbf{k}_2) = \frac{5}{7} + \frac{x}{2}\left(\frac{k_1}{k_2} + \frac{k_2}{k_1}\right) + \frac{2}{7}x^2, \qquad (3)$$

with $x = \hat{\mathbf{k}}_1 \cdot \hat{\mathbf{k}}_2$. In this work we consider scales up to $k \leq 0.3\,h\,{\rm Mpc}^{-1}$, for which the validity of Eq. (2) is only accurate to about 20% [50, 52]. A more accurate description of the bispectrum at these scales, particularly in redshift space, is given by second-order Lagrangian PT (2LPT), [52]. Therefore, we will only use tree-level PT to model *deviations* from a fiducial model calculated by using mock catalogs generated by `PTHalos` [26] and 2LPT simulations, which are similar at these scales, since halos in `PTHalos` are placed in the large-scale 2LPT density field. The advantage of using `PTHalos` is that a biased population of galaxies can be chosen by using appropriate prescriptions for their occupation inside halos. This method is therefore necessary for LRG galaxies, which are strongly biased tracers, whereas main sample galaxies are close enough to being unbiased that the difference between using 2LPT and `PTHalos` is not significant.

We therefore use the mock catalogs for the main sample of galaxies in SDSS generated by using 2LPT in [58], where the following cosmological parameters were used: dark matter density $\Omega_d = 0.225$, baryon density $\Omega_b = 0.045$, cosmological constant with density $\Omega_\Lambda = 0.73$, Hubble constant $h = 0.71$ and fluctuation amplitude $\sigma_8 = 0.82$ at the mean redshift of the survey of $z_{\rm mean} \simeq 0.1$. As discussed above, we assume these galaxies to be unbiased, and have included the detailed geometry of the expected final angular and radial selection functions. The redshift-space density field is weighted using the Feldman-Kaiser-Peacock (FKP) procedure [47, 52, 53] with $P_0 = 5000\ (h^{-1}\,{\rm Mpc})^3$. We

measure the power spectrum from $k_{\rm min} = 0.02\,h\,{\rm Mpc}^{-1}$ to $k_{\rm max} = 0.31\,h\,{\rm Mpc}^{-1}$, with a bin size given by $\Delta k = 0.015\,h\,{\rm Mpc}^{-1}$. We consider $N_k = 20$ k-bins for the power spectrum, corresponding to $N_T = 1015$ triangle bins, including all triangle shapes and orientations corresponding to 7.5×10^{10} elementary triangles. We use 6000 realizations of the survey [58, 59], such a large number is needed in order to estimate covariance matrices larger than $10^3 \times 10^3$ in size (see next section).

A second source of non-Gaussianity in the galaxy density field is given by non-linear galaxy bias. At scales much larger than the typical size of virialized structures the relation between the galaxy distribution and the underlying dark matter distribution is expected to be local [27–29], that is, in terms of the respective overdensities, $\delta_g(x) = f[\delta(x)]$. For small fluctuations we can Taylor-expand and describe such function in terms of few, constant, bias parameters [29]

$$\delta_g(x) = b_1\delta + \frac{1}{2}b_2\delta^2 + ... \tag{4}$$

The large-scale galaxy power spectrum will therefore be given $P_g(k) \simeq b_1^2 P(k)$ while the galaxy bispectrum B_g will be related to dark matter bispectrum B as

$$B_g(k_1,k_2,k_3) = b_1^3 B(k_1,k_2,k_3) + b_1^2 b_2 (P_1P_2 + {\rm cyc.}) \tag{5}$$

The different behaviour of the first and second terms on the right hand side of Eq. (5) as a function of triangle configuration given by the wavenumbers k_1, k_2 and k_3 allows a simultaneous measurement of the linear bias parameter b_1 and the quadratic bias parameter b_2 [5, 16]. This becomes obvious when Eq. (5) is rewritten in terms of the reduced bispectrum, defined, for the dark matter field as $Q \equiv B(k_1,k_2,k_3)/(P_1P_2 + {\rm cyc.})$ and analogously for the galaxy distribution, so that

$$Q_g(k_1,k_2,k_3) = \frac{1}{b_1}Q(k_1,k_2,k_3) + \frac{b_2}{b_1^2}. \tag{6}$$

While the first term on the left hand side depends on the specific triangle via the $F_2(\mathbf{k}_1,\mathbf{k}_2)$ kernel, the second just amounts to an overall additive constant.

If there is a scatter about the deterministic relationship given by Eq. (4), the bispectrum method recovers the *mean* relationship between galaxy and matter overdensities. This has been shown for models with significant scatter (see Fig. 1 in [52]) and galaxies populated using Halo Occupation Distributions (HOD) where the scatter is typically not very significant at the scales we consider here (see Fig. 6 in [30]).

For the mock catalogs of the LRG sample, we use 6000 mock catalogs constructed with `PTHalos` [26], using the following cosmological parameters: dark matter density $\Omega_d = 0.229$, baryon density $\Omega_b = 0.046$, cosmological constant with density $\Omega_\Lambda = 0.725$, Hubble constant $h = 0.71$ and fluctuation amplitude $\sigma_8 = 0.75$ at the mean redshift of the survey of $z_{\rm mean} \simeq 0.35$. In these mock catalogs the LRG galaxies populate dark matter halos according to an HOD prescription [31] for the mean number of galaxies in a halo of mass m

$$\langle N_{\rm gal}(m)\rangle = {\rm e}^{-m_{\rm min}/m}\left[1 + \left(\frac{m}{m_1}\right)^\alpha\right], \tag{7}$$

where the first contribution is that due to a central galaxy (with nearest integer scatter), the rest being satellite galaxies which are taken with a Poisson distributed scatter [32]. The parameters are chosen by a best fit procedure of the large-scale redshift-space correlation function given in [33] (including the survey covariance matrix) and the small-scale redshift-space correlation function given in [34]. The resulting parameters are $m_{\rm min} = 5 \times 10^{13}\,M_\odot h^{-1}$, $m_1 = 10^{15}\,M_\odot h^{-1}$ and $\alpha = 1.95$. Given Eq. (7), the large-scale bias parameters are given by,

$$b_i = \frac{1}{n_g}\int dm\; n(m)\;\langle N_{\rm gal}(m)\rangle\; b_i(m), \tag{8}$$

where $n(m)$ is the halo mass function (assumed to be that in [35]), $b_i(m)$ are the corresponding halo bias parameters [36] and the galaxy number density is given by

$$n_g = \int dm\; n(m)\;\langle N_{\rm gal}(m)\rangle. \tag{9}$$

For the parameters given above, $n_g = 8\times10^{-5}$, $b_1 = 2.11$, $b_2 = 1.1$ and $b_3 = -2.8$. In practice, the values in the mock catalogs are slightly different due to a somewhat different implementation, for example, $b_1 = 2.17$. This is what we use below as our fiducial value for the LRG large-scale linear bias. The mock catalogs have the radial selection function and angular selection expected by the end of the SDSS survey. The redshift-space density field in them is weighted according to the FKP method with $P_0 = 40,000\ (h^{-1}\,{\rm Mpc})^3$.

In section IV below we discuss in more detail how we take into account redshift distortions. In brief, we assume the 2LPT or `PTHalos` simulations give the correct answer (note that these *do not* assume perturbation theory for the real-to-redshift space mapping), and compute *deviations* from the fiducial model by tree-level perturbation theory. A full discussion about accurate theoretical predictions for statistics of galaxies in redshift-space and their possible systematics is beyond the scope of this paper, and will be presented elsewhere.

III. COVARIANCE MATRICES

In order to perform a joint likelihood analysis of the power spectrum and bispectrum, as detailed in section IV below, we need to compute their covariance properties. The *full* covariance matrix C_{ij} obtained from our mocks catalogs by measuring the power spectrum and bispectrum, is defined as

$$C_{ij} \equiv \langle \delta X_i \delta X_j\rangle \tag{10}$$

where $\delta X_i = X_i - \bar{X}_i$ and X_i equals the power spectrum P_i for $i = 1, ..., N_k$ with N_k the number of power spectrum bins, or the bispectrum for $i = N_k + 1, ..., N_k + N_T$, with N_T the number of bins in triangle space.

In what follows, we consider the three contributions to the general covariance matrix C_{ij}, that is $\langle \delta P_i \delta P_j \rangle$, $\langle \delta B_i \delta B_j \rangle$ and $\langle \delta P_i \delta B_j \rangle$ in turn, and compare the expected contributions to the values measured from the mock catalogues.

A. Power Spectrum Covariance

Our power spectrum estimator can be written as

$$\hat{P}(k) \equiv \frac{k_f^3}{V_P(k)} \int_k \mathrm{d}^3 q_1 \int_k \mathrm{d}^3 q_2 \; \delta_D(\mathbf{q}_{12}) \; \delta_{\mathbf{q}_1} \delta_{\mathbf{q}_2}, \qquad (11)$$

where

$$V_P(k) \equiv \int_k \mathrm{d}^3 q_1 \int_k \mathrm{d}^3 q_2 \; \delta_D(\mathbf{q}_{12}) \simeq 4\pi k^2 \Delta k, \qquad (12)$$

and the integral over the bin k of width Δk is given by

$$\int_k \mathrm{d}^3 q \equiv \int_{k-\Delta k/2}^{k+\Delta k/2} \mathrm{d}k \, k^2 \int \mathrm{d}\Omega. \qquad (13)$$

The bin width Δk does not necessarily coincide with the fundamental frequency k_f (in our analysis we will consider the case $\Delta k = 3k_f$). If the value of the power spectrum averaged over all the realizations is $P(k) = \langle \hat{P}(k) \rangle$, it is easy to see that the covariance between power spectrum bins can be expressed as [37, 38]

$$\begin{aligned} C_{ij}^P &\equiv \langle \delta P(k_i) \delta P(k_j) \rangle = \\ &\simeq \delta_{ij} \frac{2 \, k_f^3}{V_P(k_i)} P^2(k_i) \\ &+ \frac{k_f^3}{2} \int_{-1}^{+1} \mathrm{d}\cos\theta \; \widetilde{T}(k_i, k_j, \theta) \end{aligned} \qquad (14)$$

where the first diagonal term is the Gaussian contribution and in the second, non-Gaussian, term $\widetilde{T}(k_i, k_j, \theta) = T(\mathbf{k}_i, -\mathbf{k}_i, \mathbf{k}_j, -\mathbf{k}_j)$ is the trispectrum of the dark matter field and θ is the angle between the vectors $\mathbf{k}_i$ and $\mathbf{k}_j$. Note that here we are assuming that the survey window is effectively a delta function in Fourier space, for doing simple covariance estimates this suffices *except* for the case of the mixed power spectrum - bispectrum covariance as we discuss in section III C below. Of course, in analyzing our mock catalogs we do not make any such approximation as the survey geometry is properly taken into account.

Figure 2 shows the redshift-space power spectrum cross-correlation coefficients

$$r_{ij}^P \equiv \frac{C_{ij}^P}{\sqrt{C_{ii}^P C_{jj}^P}}, \qquad (15)$$

for the main (left) and LRG (right) sample power spectra measured from our mock catalogs. The values are ranging from 0 (white) to 1 (black). We used $N_k = 20$ bins for the main sample, $N_k^{LRG} = 27$ bins in the LRG sample case. The numerical value of the cross-correlation coefficients for the main sample power spectrum is given in Table I, whereas Table II presents the LRG power spectrum case. Note that in this case the maximum value for the wavenumber considered is $k_{\max}^{LRG} = 0.2\, h\,\mathrm{Mpc}^{-1}$, instead of $k_{\max}^{MS} = 0.3\, h\,\mathrm{Mpc}^{-1}$ for the main sample.

As evident from these tables and Fig. 2, the cross-correlation between different scales is stronger for the LRG power spectrum case. If LRG galaxies were simply a linearly biased population compared to the main galaxy sample (here assumed to be unbiased), then one would expect exactly the *opposite* given our choice of bin widths. This is so because the cross-correlation coefficient is independent of the volume of the sample (which appears in Eq. (14) only through k_f), proportional to the amount of non-Gaussianity (here given by the averaged trispectrum divided by the power spectrum squared), and proportional to the bin width Δk. The reason for this last

TABLE I: Power spectrum cross-correlation coefficients between different scales as measured from the main sample SDSS mock catalogues. For brevity, only even bins are displayed.

k	bin	2	4	6	8	10	12	14	16	18	20
0.031	2	1.00	0.13	0.22	0.2	0.25	0.21	0.22	0.2	0.18	0.17
0.063	4		1.00	0.31	0.34	0.37	0.36	0.34	0.34	0.32	0.30
0.094	6			1.00	0.42	0.49	0.48	0.48	0.47	0.45	0.43
0.126	8				1.00	0.54	0.55	0.55	0.54	0.53	0.51
0.157	10					1.00	0.61	0.62	0.62	0.61	0.60
0.188	12						1.00	0.69	0.7	0.71	0.70
0.220	14							1.00	0.73	0.74	0.74
0.251	16								1.00	0.78	0.78
0.283	18									1.00	0.82
0.314	20										1.00

TABLE II: Power spectrum cross-correlation coefficients between different scales as measured from the LRG sample SDSS mock catalogues. For brevity, only one bin every three is shown.

k	bin	2	5	8	11	14	17	20	23	26
0.015	2	1.00	0.18	0.24	0.26	0.24	0.26	0.24	0.25	0.23
0.038	5		1.00	0.46	0.51	0.52	0.51	0.52	0.5	0.5
0.060	8			1.00	0.65	0.66	0.66	0.66	0.64	0.66
0.083	11				1.00	0.77	0.77	0.77	0.76	0.78
0.106	14					1.00	0.83	0.83	0.82	0.83
0.123	17						1.00	0.85	0.84	0.84
0.151	20							1.00	0.86	0.87
0.173	23								1.00	0.87
0.196	26									1.00

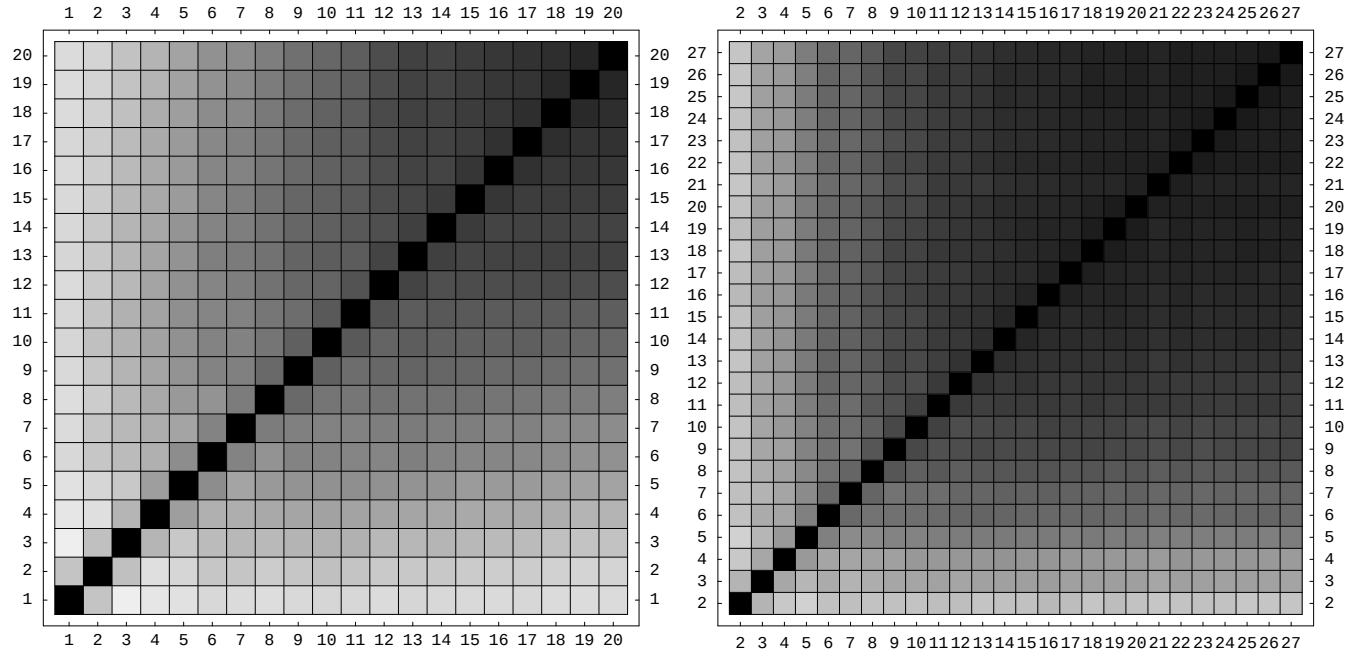

FIG. 2: Power spectrum cross-correlation coefficients r_{ij}^P between different scales for the main (left) and LRG sample (right) as measured from the SDSS mock catalogs. Black indicates maximum cross-correlation ($r_{ij}^P = 1$), white no cross-correlation ($r_{ij}^P = 0$). The wavenumbers are given in units of $\Delta k \simeq 0.015\, h\,\mathrm{Mpc}^{-1}$ for the main sample power spectrum, while $\Delta k \simeq 0.0075\, h\,\mathrm{Mpc}^{-1}$ for the LRG power spectrum. Note that bin 27 in the LRG case corresponds to bin 13 in the main sample case.

dependence is that the non-Gaussian noise term does not get beaten down by bin averaging, whereas the Gaussian one (which dominates in the denominator in Eq. (15)) does [37]. Since our bin size for LRG power is half that of the main sample, and linear bias does not alter non-Gaussianity (and redshift-distortions do so very slightly at large scales, and furthermore σ_8 is somewhat larger in the main sample), this would give that LRG power should be cross-corrrelated about half as much as main sample galaxy power.

The difference in behavior is thus a reflection of non-linearity in the LRG bias, which creates additional non-Gaussianity. In fact, this is *expected* in standard scenarios of galaxies, since Eq. (8) naturally predicts that for galaxies that populate high-mass halos where $b_1 \simeq 2$, b_2, b_3 should be at least of order unity. However, we caution that, unlike the case of linear bias $b_1(m)$ [39–41, 43, 44], the expressions for non-linear bias parameters for halos $b_2(m), b_3(m)$ given by the peak-background split [36, 42–44] (also assumed by `PTHalos`) have not been tested against current numerical simulations (see [45] for early work). This is an important issue since the prediction is that b_2, b_3 are strong functions of halo mass for the range relevant for LRG galaxies (see e.g. Fig. 8 in [36]), and small changes in the HOD parameters that leave the linear bias within observational bounds can change the non-linear bias parameters significantly. It is for this reason that we do not consider the bispectrum of LRG galaxies in this work, since its prediction has significant uncertainties. We are currently working on addressing these issues.

The nonlinearity in the bias relation for LRG galaxies implies that the power spectrum can only be reasonably well approximated by linear bias up to larger scales than in the main sample case. This is why we take $k_{\max} = 0.2\, h\,\mathrm{Mpc}^{-1}$ for LRG's, instead of $k_{\max} = 0.3\, h\,\mathrm{Mpc}^{-1}$ for the main sample. These are reasonable, though somewhat arbitrary values. In practice the allowed values of $k_{\max}$ can be empirically tested by looking at higher-order correlations and looking for scale-dependence in the derived bias parameters [17, 18].

B. Bispectrum Covariance

In an analogous way to the power spectrum case, we can define the estimator for the bispectrum, [50],

$$\hat{B}(k_1, k_2, k_3) \equiv \frac{k_f^3}{V_B} \int_{k_1} d^3q_1 \int_{k_2} d^3q_2 \int_{k_3} d^3q_3 \times \delta_D(\mathbf{q}_{123}) \delta_{\mathbf{q}_1} \delta_{\mathbf{q}_2} \delta_{\mathbf{q}_3}, \quad (16)$$

with

$$V_B \equiv \int_{k_1} d^3q_1 \int_{k_2} d^3q_2 \int_{k_3} d^3q_3\, \delta_D(\mathbf{q}_{123}) \simeq 8\pi^2\, k_1 k_2 k_3\, \Delta k^3, \quad (17)$$

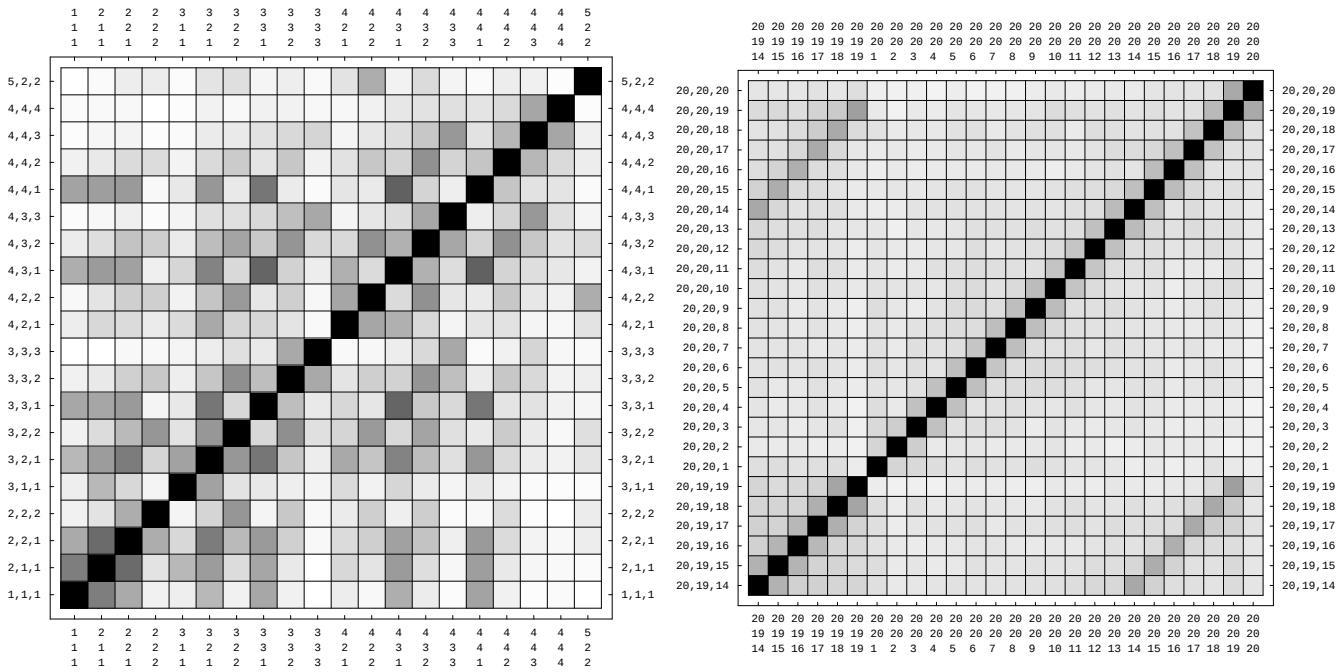

FIG. 3: Bispectrum cross-correlation coefficients r^B_{ij} among the first 20 triangle at large scales (left), and among the last 25 triangles at the smallest scales we consider (right). The triplets indicate the wavenumbers of the triangles sides in units of the k-bin width, $\Delta k \simeq 0.015\, h\,\mathrm{Mpc}^{-1}$.

then the covariance between two triangle configurations (where i and j represents the triangles while (i_1, i_2, i_3) and (j_1, j_2, j_3) are the corresponding k-vectors triplets) is,

$$C^B_{ij} \equiv \langle \delta B_i \delta B_j \rangle = +\delta_{i_1 j_1}\delta_{i_2 j_2}\delta_{i_3 j_3}\frac{k_f^3}{V_B(i)} P_{i_1} P_{i_2} P_{i_3} + \mathrm{cyc.}$$

TABLE III: Bispectrum cross-correlation coefficients for triangles at the largest scales. Each triangular configuration is given in terms of the three vectors k_1, k_2, k_3 in units of the k-bin width, $\Delta k \simeq 0.015\, h\,\mathrm{Mpc}^{-1}$.

triangle	1	2	2	2	3	3	3	3	3	3
	1	1	2	2	1	2	2	3	3	3
	1	1	1	2	1	1	2	1	2	3
1,1,1	1.00	0.51	0.33	0.06	0.07	0.28	0.06	0.35	0.06	0.00
2,1,1		1.00	0.58	0.11	0.28	0.39	0.14	0.35	0.09	0.00
2,2,1			1.00	0.32	0.15	0.52	0.27	0.4	0.19	0.03
2,2,2				1.00	0.04	0.17	0.41	0.04	0.22	0.03
3,1,1					1.00	0.37	0.11	0.09	0.07	0.04
3,2,1						1.00	0.41	0.54	0.23	0.07
3,2,2							1.00	0.15	0.44	0.12
3,3,1								1.00	0.26	0.09
3,3,2									1.00	0.34
3,3,3										1.00

$$+\delta_{i_1 j_1}\frac{k_f^3}{V_B(i)V_B(j)}\int_{k_{i_1}}\mathrm{d}^3q_1\ldots\int_{k_{i_3}}\mathrm{d}^3q_3\int_{k_{j_2}}\mathrm{d}^3p_2\int_{k_{j_3}}\mathrm{d}^3p_3 \times\delta_D(\mathbf{q}_{123})\delta_D(\mathbf{q}_1+\mathbf{p}_{23})B(\mathbf{q}_1,\mathbf{p}_2,\mathbf{p}_3)B(\mathbf{q}_1,\mathbf{q}_2,\mathbf{q}_3) + \mathrm{cyc.}$$

$$+\delta_{i_1 j_1}\frac{k_f^3}{V_B(i)V_B(j)}\int_{k_{i_1}}\mathrm{d}^3q_1\ldots\int_{k_{i_3}}\mathrm{d}^3q_3\int_{k_{j_2}}\mathrm{d}^3p_2\int_{k_{j_3}}\mathrm{d}^3p_3 \times\delta_D(\mathbf{q}_{123})\delta_D(\mathbf{p}_{23}-\mathbf{q}_1)P(q_1)T(\mathbf{q}_2,\mathbf{q}_3,\mathbf{p}_2,\mathbf{p}_3) + \mathrm{cyc.}$$

$$+\frac{k_f^3}{V_B(i)V_B(j)}\int_{k_{i_1}}\mathrm{d}^3q_1\ldots\int_{k_{i_3}}\mathrm{d}^3q_3\int_{k_{j_1}}\mathrm{d}^3p_1\ldots\int_{k_{j_3}}\mathrm{d}^3p_3 \times\delta_D(\mathbf{q}_{123})\delta_D(\mathbf{p}_{123})T_6(\mathbf{q}_1,\mathbf{q}_2,\mathbf{q}_3,\mathbf{p}_1,\mathbf{p}_2,\mathbf{p}_3), \quad (18)$$

with $T_6(\mathbf{k}_1, ..., \mathbf{k}_6)$ representing the 6-point connected correlation function in Fourier space. At large scales, the main contribution to the variance of the bispectrum is Gaussian and therefore

$$\Delta\hat{B}^2(k_1,k_2,k_3) \simeq s_B\frac{k_f^3}{V_B}\, P(k_1)P(k_2)P(k_3), \quad (19)$$

with $s_B = 6, 2, 1$ for equilateral, isosceles and general triangles, respectively.

From the expression for C^B_{ij} we see that the largest non-Gaussian contribution to the extra-diagonal elements of the bispectrum covariance matrix should arise in triangular configurations sharing two sides, with an extra factor when these are equal sides of isosceles triangles. Such large terms can be easily identified in Fig. 3 where we

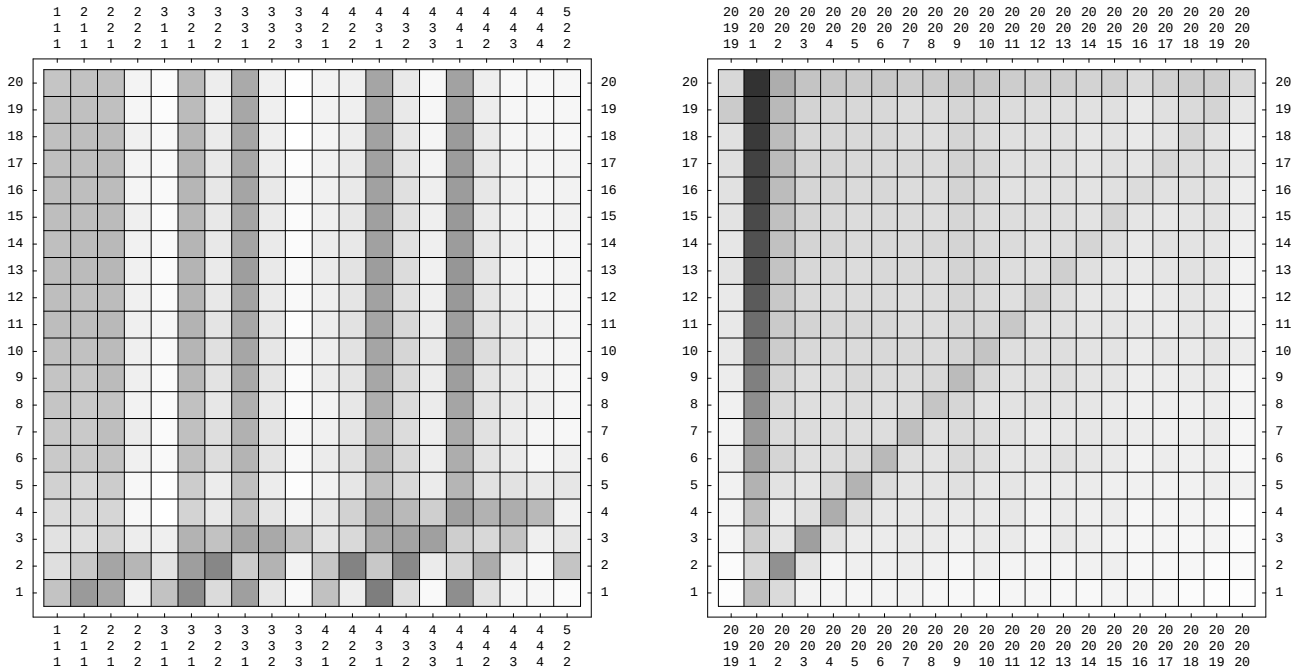

FIG. 4: Mixed cross-correlation coefficients r_{ij}^{PB} between main sample power spectrum and bispectrum. We show the 20 largest scale triangles (left) and the 21 smallest scale triangles (right) against all power spectrum bins. All numbers indicate wavenumbers in units of the k-bin width, $\Delta k \simeq 0.015\, h\,\mathrm{Mpc}^{-1}$. The vertical bands are due to beat-coupling.

show the bispectrum cross-correlation coefficients. The value of the cross-correlation coefficients at the largest scales is given in Table III. Note that, even at small scales, the bispectrum cross-correlation coefficients remain small, with values usually lower than 0.3, typically quite smaller than in the power spectrum case.

C. Mixed Covariance: Beat Coupling

Given the estimators for the power spectrum and bispectrum defined in Eqs. (11) and (16), the mixed terms in the general covariance matrix are

$$\begin{aligned} C_{ij}^{PB} &\equiv \langle \delta P_i \delta B_j \rangle = \\ &\simeq \delta_{ij_1} \frac{2k_f^3}{V_P(i)} P(k_i) B(k_{j_1}, k_{j_2}, k_{j_3}) + \text{cyc.} \\ &\quad + \frac{k_f^3}{V_P(i) V_B(j)} \int_{k_i} \mathrm{d}^3 q_1 \int_{k_{j_1}} \mathrm{d}^3 p_1 \ldots \int_{k_{j_3}} \mathrm{d}^3 p_3 \\ &\quad \times \delta_D(\mathbf{p}_{123}) T_5(\mathbf{q}_1, -\mathbf{q}_1, \mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3), \end{aligned} \tag{20}$$

where $T_5(\mathbf{k}_1, \ldots, \mathbf{k}_5)$ stands for the 5-point connected correlation function in Fourier space. At large scales, the first term in Eq. (20) dominates, and moreover, this is expected to be an important contribution. To see this, compare its magnitude to the expected signal

$$\frac{\langle \delta P_i \delta B_j \rangle}{P_i B_j} \simeq \frac{2 s_B k_f^3}{V_P(i)} \simeq \frac{s_B}{2\pi} \left(\frac{k_f}{k} \right)^2, \tag{21}$$

which is comparable to the same ratio for the diagonal covariance of the power spectrum,

$$\frac{\Delta P_i^2}{P_i^2} \simeq \frac{2k_f^3}{V_P(i)} \simeq \frac{1}{2\pi} \left(\frac{k_f}{k} \right)^2. \tag{22}$$

Figure 4 shows the cross-correlation coefficients between the first 20 (left) and last 21 (right) bispectrum configurations and all power spectrum bins. The terms just discussed correspond to the diagonal in the right panel, and a few of the elements in the bottom part of the left panel, where the power is calculated at one of the sides of the triangle.

However, it is evident that there are significant correlations beyond these, for triangles which include the smallest value of k as a side with *every* bin of the power spectrum. Indeed, Eq. (20) ignores important contributions that *dominate* the mixed covariance matrix. The reason is that so far we have ignored the effects of the window of the survey.

In a finite survey of size $\simeq L$, the uncertainty principle implies one cannot measure Fourier modes to a better accuracy than $\delta k \simeq \pi/L$, since two waves of frequency k and $k \pm \delta k$ differ only by half an oscillation from one end to the other of the survey, *i.e.* there is not enough room inside the survey to tell them apart. This implies that in reality the power spectrum estimator in Eq. (11) written in terms of the *observed* Fourier modes will necessarily contain, due to the survey window, cross-terms in the

underlying Fourier modes, written schematically as

$$\delta_{\mathbf{q}}\, \delta_{-\mathbf{q}+\epsilon}, \tag{23}$$

where $\epsilon \lesssim \delta k$, apart from "true power" contributions $\delta_{\mathbf{q}}\delta_{-\mathbf{q}} = |\delta_{\mathbf{q}}|^2$. Although such terms do not contribute to the expectation value of the power, they do correlate very well with appropriate bispectrum configurations. Indeed, due to quadratic nonlinearities two nearby Fourier modes couple to the beat mode between them,

$$\delta_{\mathbf{q}}\, \delta_{-\mathbf{q}+\epsilon} \sim \delta_{\mathbf{q}}\, F_2(-\mathbf{q}, \epsilon)\, \delta_{-\mathbf{q}}\, \delta_{\epsilon}, \tag{24}$$

which means that these terms dominate the fluctuation in power at high wavenumbers where $q \gg \epsilon$, giving the non-intuitive result that the errors of the power spectrum in the nonlinear regime are dominated by the large-scale power [61, 62]. From Eq. (24) it follows that such terms cross-correlate very well with the bispectrum of isosceles triangles with one small side of the order of the survey window $\simeq \epsilon$,

$$\langle \delta_k\, \delta_{-k+\epsilon}\, \delta_p\, \delta_{-p}\, \delta_{\epsilon} \rangle \sim P(k) P(p)\, P(\epsilon). \tag{25}$$

Therefore, *for all* k we expect power spectra to cross-correlate with bispectra of "narrow" isosceles triangles. These are the vertical features seen in Fig. 4.

Beat coupling implies that the whole power spectrum and the bispectrum of "narrow" isosceles triangles fluctuate together depending on the large-scale power. As we shall see in section V C this has interesting implications for the likelihood analysis.

IV. THE LIKELIHOOD FUNCTIONS

We now consider a hypothetical joint analysis of large-scale structure (LSS) and cosmic microwave background (CMB) anisotropies. In order to be specific and illustrate the amount of information that we expect to extract in the very near future, we consider the first year WMAP data, the power spectrum and bispectrum of the SDSS main sample of galaxies, and also include the SDSS power spectrum of the luminous red galaxies (LRG). The SDSS "data" is obtained from the mock catalogs described in section II and corresponds to the survey in its expected final form. In this section we describe the LSS and CMB likelihood functions that we use to derived the constraints discussed in the next sections.

TABLE IV: Mixed cross-correlation coefficients between SDSS main sample bispectrum and power spectrum at the largest scales. Each triangular configuration is given in terms of the three vectors k_1, k_2, k_3 in units of the k-bin width, $\Delta k \simeq 0.015\, h\, \mathrm{Mpc}^{-1}$.

	1	2	3	4	5	6	7	8	9	10
1,1,1	0.24	0.13	0.12	0.14	0.18	0.21	0.22	0.23	0.23	0.25
2,1,1	0.40	0.22	0.13	0.15	0.16	0.21	0.21	0.21	0.23	0.24
2,2,1	0.34	0.35	0.18	0.16	0.20	0.24	0.26	0.23	0.27	0.27
2,2,2	0.06	0.29	0.07	0.03	0.03	0.06	0.08	0.05	0.06	0.06
3,1,1	0.24	0.11	0.06	0.00	0.01	0.03	0.02	0.02	0.02	0.03
3,2,1	0.45	0.38	0.30	0.17	0.20	0.25	0.25	0.25	0.28	0.29
3,2,2	0.14	0.47	0.24	0.09	0.08	0.13	0.13	0.10	0.11	0.12
3,3,1	0.38	0.20	0.35	0.24	0.25	0.29	0.31	0.31	0.34	0.35
3,3,2	0.10	0.30	0.34	0.10	0.07	0.11	0.11	0.09	0.11	0.10
3,3,3	0.03	0.05	0.25	0.04	0.01	0.03	0.04	0.02	0.02	0.03

A. The LSS likelihood

For simplicity we assume that the power spectrum and bispectrum estimators are Gaussian distributed. This is certainly a good approximation near the maximum wavenumber we include, but becomes worse at large scales, where only a few modes (for the power spectrum) or triangles (for the bispectrum) contribute. The deviations from Gaussian likelihood can be included by calculating the likelihood from the Monte Carlo pool of mock catalogs [52]. Ignoring the non-Gaussianity of the likelihood can lead to a biased estimation of parameters [52, 54]. Since here we are only trying to understand the improvement brought by adding the bispectrum to the standard joint analysis of CMB and LSS, and most of the information added by the bispectrum is coming from scales small compared to the survey, our assumption should be safe. The combined power spectrum and bispectrum likelihood function is then

$$\ln \mathcal{L} = \ln \mathcal{L}_P + \ln \mathcal{L}_B + \ln \mathcal{L}_{PB}, \tag{26}$$

where

$$\ln \mathcal{L}_P = -\frac{1}{2} \sum_{i=1}^{N_k} \sum_{j=1}^{N_k} \delta P_i C_{ij}^{-1} \delta P_j, \tag{27}$$

$$\ln \mathcal{L}_B = -\frac{1}{2} \sum_{i=1}^{N_T} \sum_{j=1}^{N_T} \delta B_i C_{ij}^{-1} \delta B_j, \tag{28}$$

and

$$\ln \mathcal{L}_{PB} = -\sum_{i=1}^{N_k} \sum_{j=1}^{N_T} \delta P_i C_{ij}^{-1} \delta B_j \tag{29}$$

takes into account the mixed elements of the inverse covariance matrix C^{-1}. In Eqs. (27-29) the indices i and j run over the bins in k-space for the power spectra, N_k in all, as well as over the N_T configurations included in the bispectrum analysis. When we consider below results from the power spectrum or bispectrum individually the inverse matrix C^{-1} in Eq. (27) or (28) will be replaced by the inverse of the individual matrix $C_{ij}^P \equiv \langle \delta P_i \delta P_j \rangle$ or $C_{ij}^B \equiv \langle \delta B_i \delta B_j \rangle$. We will study as well the case of

combining the two statistics *without* taking into account their mixed covariance, in which case also only C^P and C^B will be needed.

The likelihood function for the LRG power spectrum is equivalent to Eq. (27) and the corresponding covariance matrix is independently determined from the LRG mock catalogues. Since the mean redshift of the LRG sample is $z \simeq 0.35$ compared to $z \simeq 0.1$ for the main sample, with little overlap, we assume the two samples are independent.

In what follows, starred quantities denote fiducial values, while $\hat{p}$ generically indicates particular values of the cosmological and bias parameters. The former in the most general case we consider can be written as

$$\mathbf{p} = (\tau, A_s, \omega_d, \omega_b, \Omega_\Lambda, n_s, w), \tag{30}$$

defined as the reionization optical depth, τ, the primordial amplitude of scalar fluctuations, A_s, the physical dark matter density, $\omega_d \equiv \Omega_d h^2$, the physical baryon density, $\omega_b \equiv \Omega_b h^2$, the dark energy density, Ω_Λ, the scalar spectral index, n_s and the dark energy equation of state parameter, $w = p_\Lambda/\rho_\Lambda$. The bias parameters include the main sample linear and quadratic bias, b_1 and b_2, and the LRG linear bias b_1^{LRG}.

In Eqs. (27-29), $\delta P \equiv P_s - P_s^*$ and $\delta B \equiv B_s - B_s^*$, where $P_s \equiv P_s(\hat{p}; k)$ and $B_s \equiv B_s(\hat{p}; k_1, k_2, k_3)$ are the redshift space galaxy power spectrum and bispectrum as a function of the parameters $\hat{p}$ while $P_s^* \equiv P_s(\hat{p}^*; k)$ and $B_s^* \equiv B_s(\hat{p}^*; k_1, k_2, k_3)$ are the redshift space galaxy power spectrum and bispectrum of the fiducial model (with parameters $\hat{p}^*$).

The covariance matrices are calculated at maximum likelihood from our mock catalogs, that is, we do not include a possible dependence on parameters to be estimated. A simple, approximate, check of such dependence on the bias parameters did not yield appreciable differences in the final results we present in section V.

1. Power Spectrum

Deviations from the fiducial *redshift space* power spectrum monopole P_s^*, as a function of the parameters $\hat{p}$, are modeled in the following way

$$P_s(\hat{p}; k) = b_1^2 \; f_s^P(\hat{p}) f(\hat{p}; k) \; P_s^*(k), \tag{31}$$

where P_s^* is the power spectrum measured from our mocks catalogs in redshift space and where

$$f(\hat{p}; k) \equiv \frac{A}{A^*} \left[\frac{T(\hat{p}; k)}{T^*(k)}\right]^2 \left[\frac{D(\hat{p})/\Omega_m}{D^*/\Omega_m^*}\right]^2 \times \left(\frac{k}{k_P}\right)^{n_s - n_s^*} \frac{h^{n_s}}{(h^*)^{n_s^*}}, \tag{32}$$

$T(\hat{p}; k)$ being the transfer function, $D_+(\hat{p})$ the growth factor and $k_P = 0.05\ Mpc^{-1}$ the pivot point corresponding to the scale whose power is unaffected by varying the spectral index, and

$$f_s^P(\hat{p}) \equiv \frac{a_0(\beta)}{a_0(\beta^*)} \tag{33}$$

is the redshift-space correction where

$$a_0 = 1 + \frac{2}{3}\beta + \frac{1}{5}\beta^2, \tag{34}$$

with $\beta \simeq \Omega_m^{5/9}/b_1$, corresponds to the power spectrum monopole [55]. Note that we use Eq. (34) only to model *deviations* from our fiducial cosmology assumed in the mock catalogs that include nonlinear effects from the redshift-space mapping. We assume a fiducial model that is unbiased, $b_1 = 1$ and $b_2 = 0$. The likelihood function for the LRG power spectrum is computed in the same way, except for the fiducial value of the LRG linear bias parameter $b_1^{LRG} = 2.17$.

Note that in this work we include only the monopole of the redshift-space power spectrum. By using the quadrupole as well one can better constrain β and further strengthen the constraints presented in section V below.

We calculate the transfer functions from CMBFAST [49], computing the value of $T(\hat{p}; k)$ for every value on a limited grid in parameter space, then interpolating over the final parameter grid for each value of the wavenumber k involved in the analysis.

For the growth factor $D(\Omega_m, w)$ we take advantage of the fitting formula provided in [60]. This is given by

$$D(a) = a \exp\left\{\int_0^a \mathrm{d}\ln a \left[\Omega_m(a)^\gamma - 1\right]\right\}, \tag{35}$$

with $\gamma = 0.55 + 0.05(1 + w)$ and where a is the cosmological scale factor.

2. Bispectrum

We describe the deviations of the bispectrum from the fiducial model of the mock catalogs by means of Eulerian perturbation theory. Since we are averaging the redshift space bispectrum over triangles with all possible orientations we can only need the *monopole* term in a Legendre polynomials expansion. We will consider the following approximation, see Eqs. (20-28) in [51]

$$B_s^{(0)} \simeq a_0^B(\beta) \frac{2}{b_1} F_2(\mathbf{k}_1, \mathbf{k}_2) P_g(k_1) P_g(k_2) + a_0^B(\beta) \frac{b_2}{b_1^2} P_g(k_1) P_g(k_2) + \text{cyc.}, \tag{36}$$

where $P_g(k) = b_1^2 P(k)$ is the galaxy power spectrum, and

$$a_0^B(\beta) \equiv 1 + \frac{2}{3}\beta + \frac{1}{9}\beta^2, \tag{37}$$

describes the bispectrum monopole redshift space correction, obtained from Eq.(24) and (28) in [51] by averaging

over the angle between $\mathbf{k}_1$ and $\mathbf{k}_2$ and dropping the dependence on the second-order velocity kernel and velocity dispersion which should partially cancel at large scales, to approximate the configuration dependence found in simulations for the redshift-space bispectrum in [51].

To compute the dependence of the bispectrum on cosmological parameters we will therefore use

$$B_s(\hat{p}) = f_s^B(\hat{p},)\, b_1^3\, \mathcal{B}\, \frac{B_s^*}{\mathcal{B}^*} + f_s^B(\hat{p})\, b_2\, b_1^2\, \mathcal{S}, \quad (38)$$

where B_s^* is the redshift-space bispectrum measured from the mock catalogs and

$$\mathcal{B} = 2F_2(\mathbf{k}_1,\mathbf{k}_2)P_s^*(k_1)f_1P_s^*(k_2)f_2 + \text{cyc.}, \quad (39)$$

$$\bar{\mathcal{S}} = P_s^*(k_1)f_1P_s^*(k_2)f_2 + \text{cyc.}, \quad (40)$$

while

$$f_s^B(\hat{p},\hat{p}^*) \equiv \frac{a_0^B(\beta)}{a_0^B(\beta^*)}, \quad (41)$$

with β^* as fiducial β and $f_1 \equiv f(\hat{p};k_1)$ as defined above.

3. *Inverting the Covariance Matrix*

The values of the entries of the complete covariance matrix $C_{ij} \equiv \langle X_iX_j\rangle$ with $X_i = P_i, B_i$ span several orders of magnitude and thus a direct computation of its inverse is susceptible to numerical instabilities. We therefore "normalize" the covariance matrix by factoring out in the $X = P, B$ vector the power spectrum and bispectrum predicted by linear theory and Eulerian PT, respectively. The resultant entries for this "normalized" covariance matrix are therefore all of order unity. Still, by performing a singular value decomposition (SVD) one can notice a poor determination of a few singular values, about 17 out of 1035 in the complete, $k_{\max}^{MS} = 0.3\,h\,\mathrm{Mpc}^{-1}$ case, indicating that 6000 mock catalogs is enough to determine most of the elements except for a small fraction. In the final analysis we compute the inverse by means of its SVD inverse by dropping these few singular values, assuming this might be a sign of a not optimal determination of the matrix C_{ij}. By doing this we make a conservative choice since the operation amounts to discard some of the potential information contained in the covariance matrix. Therefore our final error bars increase slightly. In computing the inverse of the individual power spectrum and bispectrum covariance matrices no such limitation is needed.

B. The CMB likelihood

To combine the results of the LSS likelihood analysis with CMB data as measured by WMAP, we need to compute the CMB anisotropies power spectrum and its corresponding likelihood for each model in our grid. This procedure is computationally expensive: version 4.5.1 of the CMBFAST code [49] takes about 30 seconds per model, and the WMAP likelihood code for the first year data release [46] takes $2-4$ seconds.

A possible approach to reduce the computing time is investigated in [56], where a polynomial approximation to the multidimensional log-likelihood is computed, allowing for an evaluation of the likelihood of each model in tenths of a second. However, their available code, CMBFIT 1.0, does not include the dark energy equation of state parameter, w. Although calculating such a polynomial fit still requires sampling the likelihood surface, it has to be done only once, thus reducing enormously the computational time needed for any ulterior likelihood analysis.

Motivated by that idea, we compute a polynomial fit to the CMB likelihood function based on the 7 parameters in Eq. (30). We use uniform priors in the following ranges: $0 \le \tau \le 0.3$, $0.018 \le \omega_b \le 0.028$,$-2 \le w \le 0.2$,$0.5 \le A_s \le 1.4$,$0.4 \le \Omega_\Lambda \le 0.9$,$0.08 \le \omega_d \le 0.22$,$0.8 \le n_s \le 1.1$.

We compute the likelihood on a homogeneous grid with 15 points per dimension for the parameters $(\omega_d, \omega_b, \Omega_\Lambda, n_s, w)$, and 30 points for (τ, A_s). To speed up the calculation, we divided the problem in two steps: firstly, we computed the likelihood on the grid approximating the dependence of the power spectrum on τ with the multiplicative factor $e^{-2\tau}$. These 7×10^8 approximate models allowed us to find the region of high likelihood. Secondly, we selected the subset of the $263,022$ most likely models and recalculated their likelihood for the exact τ dependence.

We fitted a 4th order polynomial to the log-likelihood surface spanned by our reduced dataset using a weighted least squares method. We weighted the fitting error of each model with its likelihood to counterbalance the fact that our grid was relatively coarse and there were many more low likelihood models than high likelihood ones. The covariance matrix of our 7-dimensional reduced set of models is given by $C_{ij} \equiv \langle p_ip_j\rangle - \langle p_i\rangle\langle p_j\rangle$, with $\langle p_i\rangle \equiv \sum p_i\,\mathcal{L}(\mathbf{p})$. In order to improve the numerical behavior of the fitting algorithm we first changed from $\mathbf{p}-$space to the variables $\mathbf{z}$ with zero mean and unit covariance defined as

$$\mathbf{z} \equiv \mathbf{E}(\mathbf{p} - \langle\mathbf{p}\rangle), \quad (42)$$

where the columns of $\mathbf{E}$ were defined as the eigenvectors of the covariance matrix $\mathbf{C}$ divided by the square root of their corresponding eigenvalues, *i.e.* such that $\mathbf{ECE}^t = \mathbf{I}$, and thus $\langle\mathbf{zz}^t\rangle = \mathbf{I}$, [56]. The 4th order polynomial, containing $M = 330$ terms, can be written in terms of the new variables as

$$y \equiv \ln\mathcal{L} = q_0 + \sum_{i_1=1}^{7} z_{i_1}\left\{ q_1^{i_1} + \sum_{i_2=i_1}^{7} z_{i_2}\left[q_2^{i_1 i_2} + \sum_{i_3=i_2}^{7} z_{i_3}\left(q_3^{i_1 i_2 i_3} + \sum_{i_4=i_3}^{7} q_4^{i_1 i_2 i_3 i_4} z_{i_4}\right)\right]\right\}. \quad (43)$$

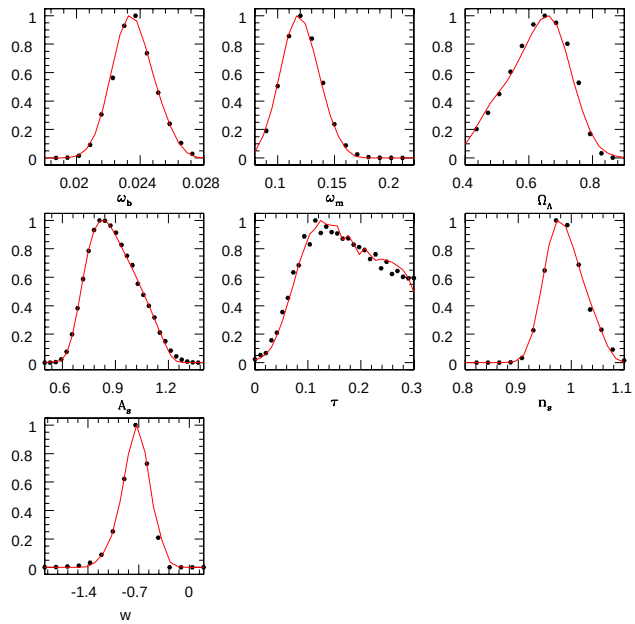

FIG. 5: WMAP TT+TE marginalized likelihoods for the 7 cosmological parameters in our wCDM model. Dots correspond to the marginalization over the reduced dataset in the grid (263,022 points) while the solid line was obtained using the *weighted* polynomial fit.

In order to make this expression compact we arranged all possible products of z_i's (up to the 4th power) into an M-dimensional array $\mathbf{x}$,

$$\mathbf{x} = \{1, z_1, \ldots, z_7, z_1 z_1, z_1 z_2, \ldots, z_7^4\}, \tag{44}$$

and the corresponding coefficients q into,

$$\mathbf{q} = \{q_0, q_1^1, \ldots, q_1^7, q_2^{11}, \ldots, q_4^{7777}\}, \tag{45}$$

so that Eq. (43) could be casted as

$$y = \mathbf{x} \cdot \mathbf{q}. \tag{46}$$

Next, we arranged our $N = 263,022$ datapoints, written in the $\mathbf{x}$ format, into an $(N \times M)$ matrix $\mathbf{X}$ and their corresponding likelihoods into an N-dimensional vector $\mathbf{y}$. Therefore, the weighted error of the polynomial fit was given by,

$$\epsilon \equiv \left[\frac{1}{M} \sum_{i=1}^{M} w_i (y_i - \sum_{j=1}^{M} X_{ij} q_j)^2 \right]^{1/2} \tag{47}$$

where $w_i = \mathcal{L}(\mathbf{p}_i)$. The coefficients q_i were then chosen such as to minimize ϵ,

$$\mathbf{q} = (\tilde{\mathbf{X}}^t \mathbf{X})^{-1} \tilde{\mathbf{X}}^t \mathbf{y}, \tag{48}$$

where we defined $\tilde{\mathrm{X}}_{ij} = w_i \mathrm{X}_{ij}$.

TABLE V: Fiducial values for the cosmological and bias parameters assumed.

Parameter		Fiducial value
ω_d	physical dark matter density	0.1222
ω_b	physical baryon density	0.0232
Ω_Λ	dark energy density	0.699
n_s	scalar spectral index	0.977
A_s	scalar fluctuation amplitude	0.81
w	dark energy equation of state	-1
τ	reionization optical depth	0.124
b_1	main sample linear galaxy bias	1
b_2	main sample quadratic galaxy bias	0
b_1^{LRG}	LRG linear galaxy bias	2.17
Derived parameters		**Fiducial value**
σ_8	galaxy fluctuation amplitude	0.917
Ω_m	matter density	0.301
Ω_b	baryon density	0.048
h	Hubble parameter	0.695

In order to avoid unphysical likelihood values due to polynomial artifacts in low confidence regions that were poorly sampled, we replaced the polynomial fit by a simple Gaussian for $z = |\mathbf{z}|$ outside the 2-sigma level, given that the likelihood distribution has a spherically symmetric tail for $z > 2.5$, well approximated by $\mathcal{L} \propto \mathrm{e}^{-z^2/2}$.

To test our weighted fit we compared it, for w $= -1$, against CMBFIT for a 6-parameter ΛCMD model $(\tau, A_s, \omega_d, \omega_b, \Omega_\Lambda, n_s)$, finding very good agreement. Another estimator of the goodness of the fit was the rms fitting error $\langle \Delta \ln \mathcal{L} \rangle$. We found that $\langle \Delta \ln \mathcal{L} \rangle = 0.29$ for our fitting dataset of $263,022$ models. Furthermore, we made a Monte Carlo Markov Chain test of 3000 models, which yielded $\langle \Delta \ln \mathcal{L} \rangle = 0.32$. These errors are similar to those reported in [56] for the 7-parameter models of CMBFIT.

Figure 5 shows the first year WMAP TT+TE marginalized likelihoods for the case of the 7-parameter wCDM models, Eq. (30), obtained by a grid marginalization over the reduced dataset (dots) versus the ones using the weighted polynomial fit (solid line). This shows the procedure described above is robust enough for our purposes.

V. RESULTS

In this section we present the results of the likelihood analysis in two classes of flat cosmological models. The first, section V A, corresponds to ΛCDM models depending on six cosmological plus three bias parameters: the density parameters ω_d, ω_b, Ω_Λ, the spectral index n_s, the fluctuations amplitude A_s, the reionization optical depth τ plus the linear and quadratic galaxy bias coefficients b_1 and b_2 for the main sample and the linear bias for

TABLE VI: ΛCDM models: expected marginalized errors (68% CL) for WMAP (temperature and polarization, W) combined with the SDSS main sample power spectrum (P) and bispectrum (B) and with the LRG power spectrum (P_L). The percentage in parenthesis indicates the improvement over the analysis including the main sample power spectrum alone (W+P), numbers in bold indicate errors down by at least 1.5. In brackets we quote the W+P+B errors obtained by ignoring the mixed power spectrum - bispectrum covariance.

	W+P	W+B	W+P+B	W+P+B (no mix. cov.)	W+P+B+P_L	W+P+P_L	W+P_L
	$k_{\max}^{MS} = 0.2\,h\,\mathrm{Mpc}^{-1}$					$k_{\max}^{LRG} = 0.2\,h\,\mathrm{Mpc}^{-1}$	
$\Delta\omega_d$	0.0035	0.0041 (−15%)	0.0031 (13%)	[0.0030 (17%)]	0.0025 (40%)	0.0026 (35%)	0.0029 (21%)
$\Delta\omega_b$	0.00093	0.00098 (−5%)	0.00078 (19%)	[0.00082 (13%)]	0.00074 (26%)	0.00081 (15%)	0.00087 (7%)
$\Delta\Omega_\Lambda$	0.0133	0.0113 (18%)	**0.0085** (56%)	[0.0085 (56%)]	**0.0063** (111%)	**0.0078** (70%)	0.0091 (46%)
Δn_s	0.022	0.024 (−8%)	0.0158 (39%)	[0.0176 (25%)]	**0.0140** (57%)	0.0171 (28%)	0.020 (10%)
ΔA_s	0.091	0.094 (−3%)	0.064 (42%)	[0.074 (23%)]	0.062 (47%)	0.078 (17%)	0.085 (7%)
$\Delta\tau$	0.052	0.053 (−2%)	0.039 (33%)	[0.044 (18%)]	0.038 (37%)	0.047 (11%)	0.049 (6%)
Δb_1	0.086	0.113 (−24%)	0.060 (43%)	[0.074 (16%)]	**0.054** (59%)	0.070 (23%)	-
Δb_2	-	0.069	0.054	[0.062]	0.051	-	-
Δb_1^{LRG}	-	-	-	-	0.099	0.125	0.137
$\Delta\sigma_8$	0.068	0.074 (−8%)	0.047 (45%)	[0.054 (26%)]	**0.043** (58%)	0.054 (26%)	0.061 (11%)
Δh	0.0152	0.0140 (9%)	**0.0101** (50%)	[0.0108 (41%)]	**0.0087** (75%)	0.0106 (43%)	0.0127 (19%)
$\Delta\Omega_b$	0.00151	0.00141 (7%)	0.00124 (22%)	[0.00124 (22%)]	0.00111 (36%)	0.00117 (29%)	0.00127 (19%)
	$k_{\max}^{MS} = 0.3\,h\,\mathrm{Mpc}^{-1}$					$k_{\max}^{LRG} = 0.2\,h\,\mathrm{Mpc}^{-1}$	
$\Delta\omega_d$	0.0033	0.0031 (6%)	0.0029 (14%)	[0.0026 (27%)]	0.0024 (37%)	0.0026 (27%)	
$\Delta\omega_b$	0.00090	0.00083 (8%)	0.00073 (23%)	[0.00072 (25%)]	0.00070 (28%)	0.00080 (12%)	
$\Delta\Omega_\Lambda$	0.0112	**0.0065** (72%)	**0.0063** (78%)	[0.0057 (96%)]	**0.0052** (115%)	**0.0073** (53%)	
Δn_s	0.021	0.018 (17%)	**0.014** (50%)	[0.013 (61%)]	**0.012** (75%)	0.016 (31%)	
ΔA_s	0.087	0.081 (7%)	**0.053** (64%)	[0.063 (38%)]	**0.052** (67%)	0.077 (13%)	
$\Delta\tau$	0.050	0.047 (6%)	**0.033** (52%)	[0.039 (28%)]	**0.033** (52%)	0.046 (9%)	
Δb_1	0.081	0.094 (−14%)	**0.051** (59%)	[0.060 (35%)]	**0.046** (76%)	0.068 (19%)	
Δb_2	-	0.045	0.041	[0.044]	0.040	-	
Δb_1^{LRG}	-	-	-	-	0.084	0.123	
$\Delta\sigma_8$	0.064	0.059 (8%)	**0.037** (73%)	[0.041 (56%)]	**0.034** (88%)	0.053 (21%)	
Δh	0.0132	0.0095 (39%)	**0.0082** (61%)	[0.0080 (65%)]	**0.0082** (61%)	0.0101 (31%)	
$\Delta\Omega_b$	0.00138	0.00112 (23%)	0.00111 (24%)	[0.00106 (30%)]	0.00104 (33%)	0.00115 (20%)	

the LRG sample, b_1^{LRG}. In the second class, denoted as wCDM models, section V E, we allow for a dark energy equation of state parametrized by the ratio of pressure to energy density w, assumed to be constant.

We include the temperature and polarization WMAP likelihood by means of the interpolation fit described in section IV B. We introduce here a flat prior on τ by limiting its values from zero to 0.3. The difference with the case of τ taking values up to 0.8 is negligible (tested for $w = -1$) and, most importantly, such a prior is more than justified by the three-year WMAP data which favors values of τ close to 0.1 [3, 57].

The fiducial values chosen for the present analysis are given in Table V. Note that they do not coincide with the maximum likelihood values obtained from the WMAP data alone, rather they correspond to those obtained for the WMAP+SDSS power spectrum 6 parameters case in [1]. These values are only relevant in the sense that they determine the point in parameter space about which we compute the errors. As long as this point is realistic, the results we present should be insensitive to their precise values.

A. ΛCDM models

We present now the expected errors on cosmological parameters from an analysis that considers different combinations of the main sample power spectrum and bispectrum and the power spectrum of the LRG sample with WMAP CMB data. We restrict here to the case of ΛCDM models, *i.e.* $w = -1$. The results for the 1-σ marginalized uncertainties are given in Table VI where we show, for clarity, the average between upper and lower limits. To see more clearly the benefits brought by using different statistics, in parenthesis we indicate the frac-

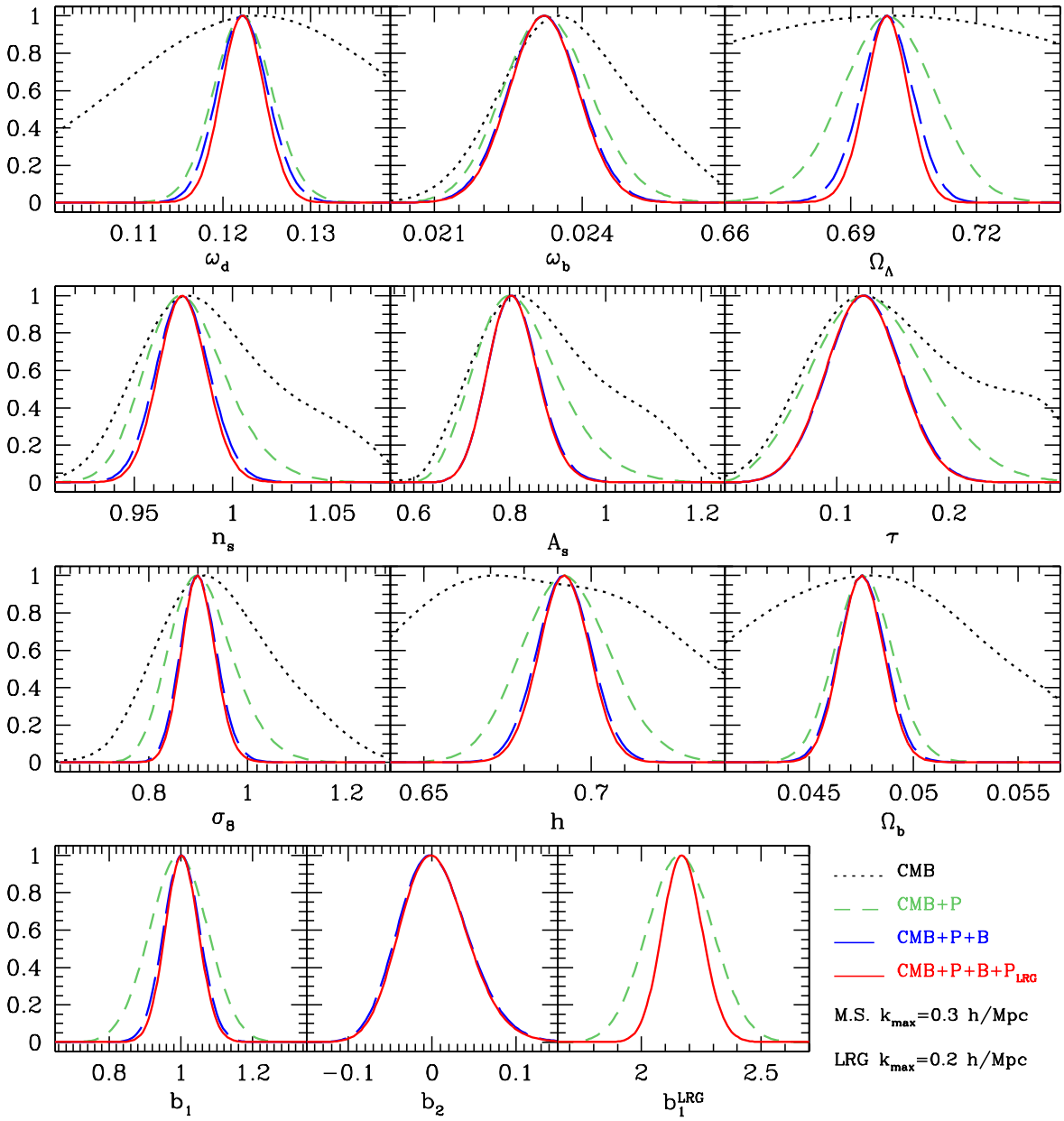

FIG. 6: Marginalized likelihood functions for the ΛCDM models assuming $k_{\rm max}^{MS} = 0.3\,h\,{\rm Mpc}^{-1}$. For the LRG linear bias parameter b_1^{LRG}, the dashed line denotes the WMAP plus LRG power spectrum likelihood.

tional improvement over the WMAP plus main sample power spectrum case (W+P), defined as

$$\text{improvement factor} = \frac{\Delta_{\rm W+P}}{\Delta} - 1, \qquad (49)$$

so a 50% (100%) improvement corresponds to reducing the errors by a factor of 1.5 (2).

The first two columns in Table VI show that analyzing the power spectrum and bispectrum separately can provide similar constraints (with the bispectrum determining an extra parameter, b_2). This can provide important consistency checks, as the Gaussian and non-Gaussian properties of galaxy clustering must yield consistent results.

As expected, the effectiveness of the bispectrum in constraining cosmology depends significantly on the smallest scale considered due to the fast rise in the number of triangles available. One can notice how, already

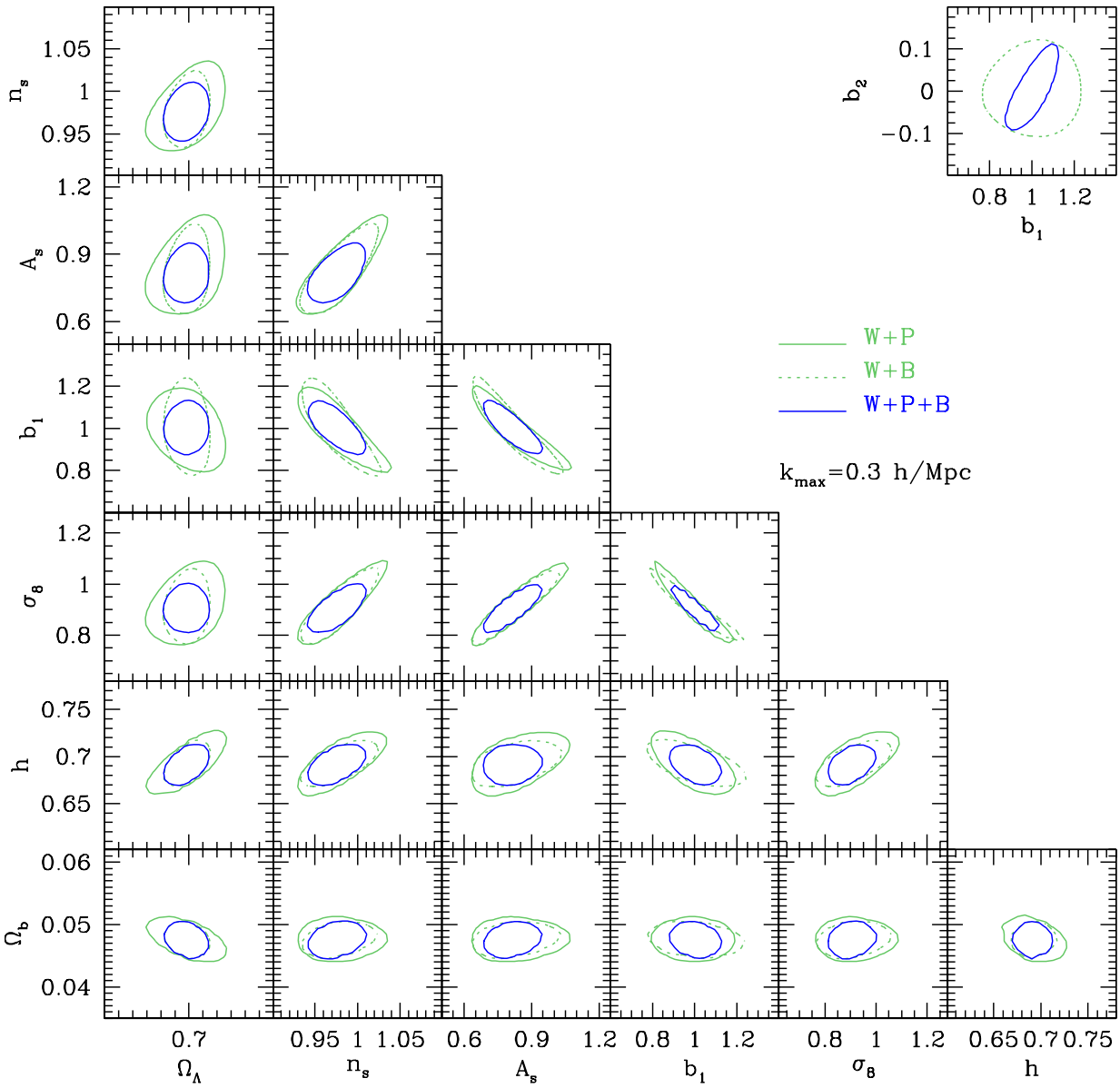

FIG. 7: Marginalized 95% contour plots for pairs of cosmological parameters in the ΛCDM models assuming $k_{\rm max}^{MS} = 0.3\,h\,{\rm Mpc}^{-1}$.

at $k_{\rm max}^{MS} = 0.2\,h\,{\rm Mpc}^{-1}$ when combined with the power spectrum, it can improve errors by a 13 to 56%. At $k_{\rm max}^{MS} = 0.3\,h\,{\rm Mpc}^{-1}$ when considered alone with CMB information the bispectrum can actually improve over the power spectrum by 6% to more than 70% for Ω_Λ, although at the expense of a poorer determination of the linear bias. A quick glance at Table VI shows that most of the improvement (numbers in bold) brought by the bispectrum are in parameters related to the overall amplitude of fluctuations and the effective spectral index. This is expected as the bispectrum breaks the degeneracy between bias and dark matter amplitude fluctuations [5, 16], and its configuration dependence is sensitive to the spectral index because of the anisotropy of tidal gravitational fields and velocity flows [64].

In Fig. 6 we compare the CMB power spectrum likelihoods to the combined power spectrum, bispectrum and LRG power spectrum likelihoods. From this and Table VI one sees that most of improvement over CMB alone is coming from pairs of statistics that involve the

TABLE VII: Comparison between W+P and W+P+B with and without galaxy bias assumptions. Includes W+P with a prior on b_1 to reproduce the error from W+P+B (left) and a case where the bias parameters are fixed (right). The percentage in parenthesis indicate the improvement over the respective W+P result. Assumes a ΛCDM cosmology and $k_{\rm max}^{MS} = 0.3\,h\,{\rm Mpc}^{-1}$.

	W+P	W+P+B	W+P+ b_1 prior	W+P+b_1 prior (Q)	W+P fixed bias	W+P+B fixed bias
$\Delta\omega_d$	0.0033	0.0029 (14%)	0.0032 (3%)	0.0031 (6%)	0.0031	0.0026 (19%)
$\Delta\omega_b$	0.00090	0.00073 (23%)	0.00078 (15%)	0.00073 (23%)	0.00068	0.00065 (5%)
$\Delta\Omega_\Lambda$	0.0112	0.0063 (78%)	0.0109 (3%)	0.0107 (5%)	0.0106	0.0062 (71%)
Δn_s	0.021	0.014 (50%)	0.016 (31%)	0.014 (50%)	0.0112	0.0096 (17%)
ΔA_s	0.087	0.053 (64%)	0.061 (43%)	0.049 (77%)	0.033	0.029 (14%)
$\Delta\tau$	0.050	0.033 (51%)	0.038 (32%)	0.032 (56%)	0.026	0.024 (8%)
Δb_1	0.081	0.051 (59%)	0.051 (59%)	0.036 (125%)	-	-
Δb_2	-	0.041	-	-	-	-
$\Delta\sigma_8$	0.064	0.036 (78%)	0.043 (49%)	0.032 (100%)	0.016	0.011 (45%)
Δh	0.0139	0.0088 (58%)	0.0116 (20%)	0.0110 (26%)	0.0103	0.0074 (39%)
$\Delta\Omega_b$	0.00139	0.00111 (25%)	0.00134 (4%)	0.00134 (4%)	0.00134	0.00104 (29%)

bispectrum (*i.e.* either P+B or P_L+B, not shown for clarity). This is so because the most significant improvements arise due to breaking of degeneracies present in the LSS or CMB [65]. This manifests itself in the entries in Table VI ($k_{\rm max}^{MS} = 0.2\,h\,{\rm Mpc}^{-1}$) in several ways : 1) W+P_L improves mildly over W+P, but W+P+P_L improves significantly over W+P (consistent with Table 2 in [63]) 2) W+P+B is better than W+P+P_L in most parameters (except those related to Ω_m: ω_d, Ω_b and Ω_Λ, since a better detection of the acoustic scale in the LRG sample gives a high quality constraint [63]). This holds even though the signal to noise in B (for $k_{\rm max}^{MS} = 0.2\,h\,{\rm Mpc}^{-1}$) is not as large as in P_L (e.g. compare W+B vs. W+P_L), because B is more complementary than P_L to W+P, *i.e.* using non-Gaussian information provides a substantially different direction in parameter space. When using information up to $k_{\rm max}^{MS} = 0.3\,h\,{\rm Mpc}^{-1}$, W+P+B constrains all parameters better than W+P+P_L except for ω_d. In this case, adding P_L to W+P+B still helps in improving parameters slightly (see Fig. 6), particularly for Ω_Λ (or $\Omega_m = 1 - \Omega_\Lambda$).

It is interesting to compare the results on bias parameters to those in a fixed cosmology, as assumed in past work [14, 17–19]. Performing an analysis of the bispectrum alone with fixed cosmology, one finds for linear and quadratic bias the errors ($k_{\rm max}^{MS} = 0.3\,h\,{\rm Mpc}^{-1}$)

$$\Delta b_1 = 0.015, \qquad \Delta b_2 = 0.045. \tag{50}$$

Comparing this to the corresponding entry (W+B) in Table VI we see that when cosmology is allowed to vary, the determination of b_1 suffers from the degeneracy with A_s and, when combined with CMB data, with the optical depth τ, while the result for b_2 is essentially unaffected. On the other hand, this is the price one pays for constraining cosmological parameters more accurately.

Figure 7 shows the marginalized 95% CL contour plots of pairs of parameters. The role played by the bispectrum in lifting the degeneracy between the galaxy bias parameter b_1 and the parameter A_s determining the amplitude of dark matter fluctuations is particularly evident. It is clear in particular from the b_1-A_s contours, that the combination of power spectrum and bispectrum, by narrowing the uncertainty on these two parameters, affects the determinations of all the others. The question then arises, are the improvements on cosmological parameters brought by using the bispectrum just a result of having constrained galaxy bias?

B. Not Just Galaxy Bias

In order to answer this question, we present in Table VII a couple of tests that compare W+P with W+P+B for $k_{\rm max}^{MS} = 0.3\,h\,{\rm Mpc}^{-1}$. The first two columns repeat the constraints shown before in Table VI, whereas the third column shows the W+P results when a prior on b_1 is added to mimic the W+P+B constraint on bias. Since the marginalized likelihood of b_1 is approximately Gaussian (see Fig. 6) we can add a Gaussian prior with width σ given by

$$\frac{1}{\sigma^2} = \frac{1}{\sigma_B^2} - \frac{1}{\sigma_P^2} \tag{51}$$

where σ_B is the error on b_1 from the W+P+B analysis and σ_P is that from W+P. We see from Table VII that this reproduces the W+P+B bias constraint closely enough. By comparing the rest of the entries in W+P+B against W+P+b_1 prior *it follows that the improvement on cosmological parameter determination from the bispectrum is not only due to constraining galaxy bias.*

The right side of Table VII presents another test, where the bias parameters are fixed ($b_1 = 1$, $b_2 = 0$). Comparing these last two columns we see a significant improvement from adding bispectrum information.

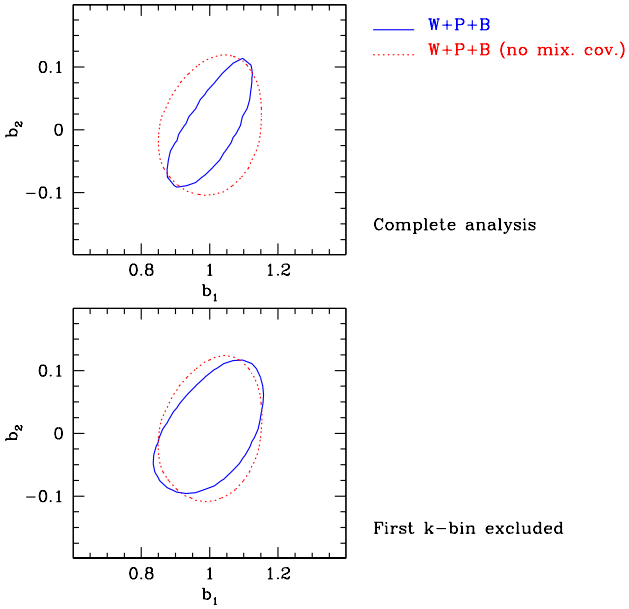

FIG. 8: Marginalized 95% contour plots for the linear bias parameter b_1 and the quadratic bias parameter b_2 in the W+P+B case with $k_{\rm max}^{MS} = 0.3\,h\,{\rm Mpc}^{-1}$. The top panel shows the full analysis with (blue, continuous line) and without (red, dashed line) the mixed covariance matrix, whereas in the bottom panel the two analyses are repeated excluding the lowest power spectrum bin, thus suppressing beat coupling.

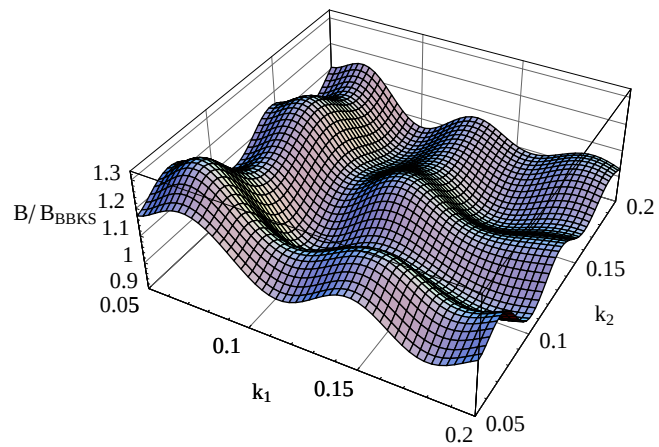

FIG. 9: Baryon acoustic oscillations imprinted in the bispectrum as a function of k_1 and k_2 for fixed angle given by $\hat{k}_1 \cdot \hat{k}_2 = -1/2$. The diagonal $k_1 = k_2$ corresponds to equilateral triangles. As the angle between the two vectors is varied the pattern of peaks moves accordingly.

The fourth column in Table VII shows the analysis of the W+P case with a prior on linear bias, $\Delta b_1 = 0.036$ [59], corresponding to the case where the bispectrum is analyzed through the hierarchical amplitude Q [see Eq. (6)] for a fixed cosmology. We see that in this case some of the constraints agree, but the error on bias is significantly underestimated, whereas the errors on Ω_Λ and h are significantly overestimated. Interestingly, the uncertainty in σ_8 is robust to this analysis (which is incorrect due to neglecting cross-correlations between Q and P and bias with cosmology).

C. The Effects of Beat Coupling

It is interesting to see what happens with the constraints on parameters if the mixed covariance between power spectrum and bispectrum is ignored, this is given in brackets in the fourth column of Table VI for the W+P+B case. Naively, one would expect that ignoring the mixed covariance should lead to better constraints, but as shown in Table VI this is incorrect for most parameters: this is due to the effects of beat coupling.

As discussed in section III C, beat coupling means that the structure of the mixed covariance matrix is dominated by up and down correlated fluctuations of the whole power spectrum and bispectrum of narrow isosceles triangles depending on the power of the largest mode in the survey, as shown in Eq. (25). Not allowing for such effect in the covariance matrix means that these fluctuations will be mistaken as a signature of larger errors in the parameters that characterize the amplitude of galaxy fluctuations. Indeed, as seen by comparing the third and fourth columns in Table VI, including the mixed covariance (thus allowing for beat coupling) reduces the errors mostly on A_s, τ, b_1, b_2 and thus σ_8.

In Fig. 8 we illustrate this point further by showing the marginalized 95% contour plots for the linear bias parameter b_1 and the quadratic bias parameter b_2 in the W+P+B case with $k_{\rm max}^{MS} = 0.3\,h\,{\rm Mpc}^{-1}$. In the top panel we show the full analysis, whereas the bottom panel drops the lowest k-bin in the power spectrum. The top panel shows a significant difference between including the mixed terms $\langle \delta P_i \delta B_j \rangle$ in the complete covariance matrix (solid) and dropping them (dashed). We see that including the mixed covariance gives clearly a tighter constraint on the two parameters together with a slight degeneracy.

In the lower panel the same contours are plotted but now the analysis excludes the smallest power spectrum bin, thus suppressing the effects of beat coupling. We see that in this case there is not much difference between including or not the mixed covariance, in fact the mild degeneracy induced by including the mixed covariance leads to slightly larger errors for b_1 and b_2. The same behavior is seen with all other parameters when the first k-bin is excluded, except for A_s and τ which still show a minor improvement when the mixed covariance is included. This is likely due to residual beat coupling, *e.g.* a careful look at the left panel in Fig. 4 shows that verti-

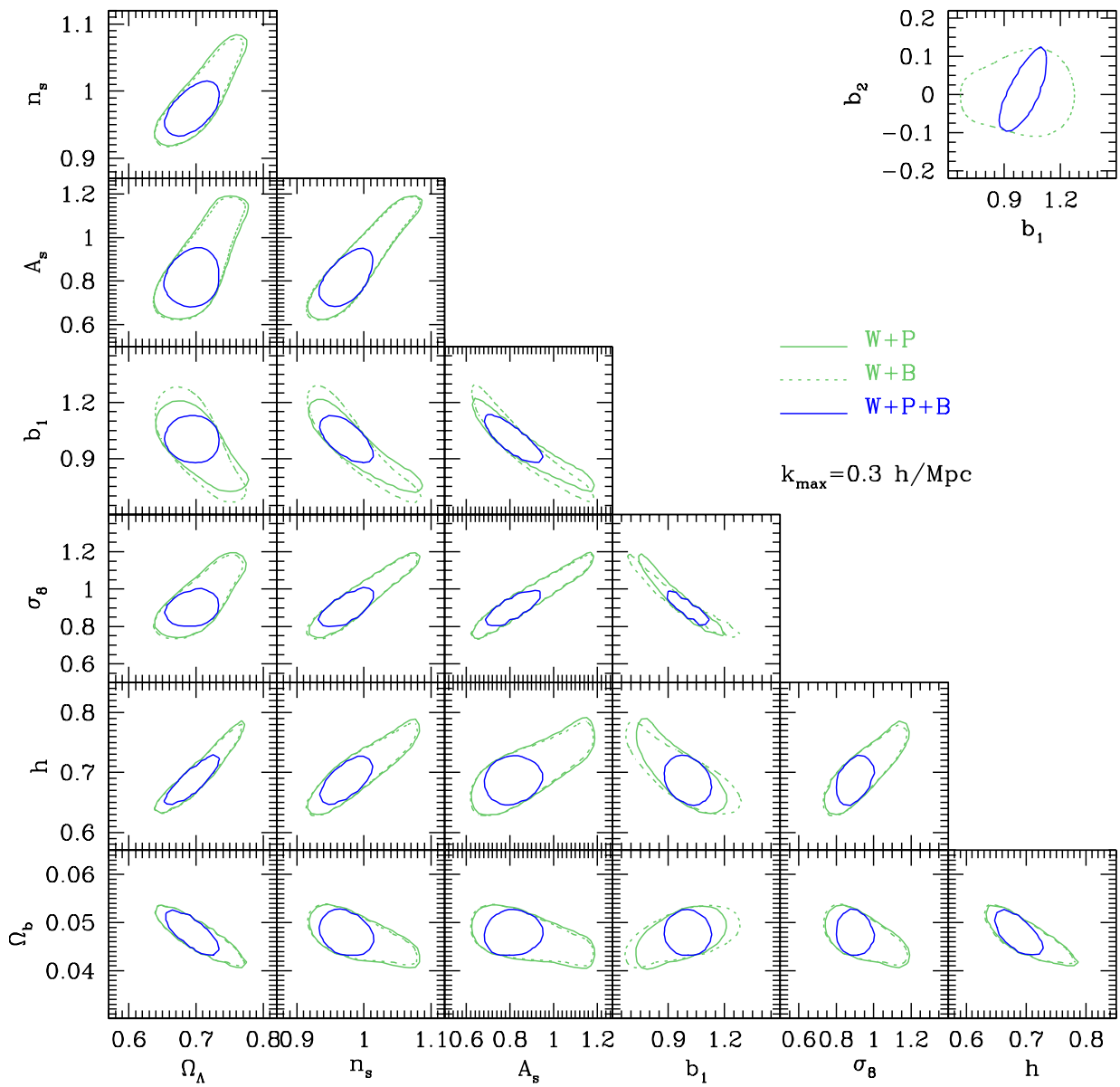

FIG. 10: Same as Fig. 7 but for a featureless (no baryon acoustic oscillations) transfer function.

cal features also exist for modes with $k = 2\Delta k$, although at a much lower amplitude.

D. Baryon Acoustic Oscillations

The same baryon acoustic oscillation features induced in the dark matter power spectrum [66, 67] and recently seen in galaxy surveys [63, 68] are expected to be present in the bispectrum [69] and can also be used to help in determining cosmological parameters. Figure 9 shows the ratio of the bispectrum to a featureless (no acoustic oscillation) bispectrum obtained from the BBKS fitting formula [48] by setting the shape parameter $\Gamma = 0.175$. Because the bispectrum scales as the square of the power spectrum, the 15% modulation in power leads to a 30% modulation in the bispectrum. At $k = 0.1\,h\,\mathrm{Mpc}^{-1}$ the signal to noise in the bispectrum is about twice smaller than for the power spectrum [59], so this scale roughly presents the limit after which the bis-

TABLE VIII: Same as Table VI but for wCDM models

	W+P	W+B	W+P+B	W+P+B (no mix. cov.)	W+P+B+P$_L$	W+P+P$_L$	W+P$_L$
	$k_{\rm max}^{MS} = 0.2\,h\,{\rm Mpc}^{-1}$					$k_{\rm max}^{LRG} = 0.2\,h\,{\rm Mpc}^{-1}$	
$\Delta\omega_d$	0.0070	0.0073 (-4%)	0.0061 (15%)	[0.0058 (21%)]	0.0047 (49%)	0.0050 (40%)	0.0056 (25%)
$\Delta\omega_b$	0.00109	0.00113 (-4%)	0.00094 (16%)	[0.00097 (12%)]	0.00084 (30%)	0.00094 (16%)	0.00102 (7%)
$\Delta\Omega_\Lambda$	0.0141	0.0121 (17%)	0.0102 (38%)	[0.0098 (44%)]	**0.0076** (86%)	**0.0089** (58%)	0.0101 (40%)
Δn_s	0.032	0.032 (0%)	0.026 (23%)	[0.027 (19%)]	0.022 (45%)	0.025 (28%)	0.029 (10%)
ΔA_s	0.131	0.129 (2%)	0.107 (22%)	[0.112 (17%)]	0.092 (42%)	0.109 (20%)	0.119 (10%)
Δw	0.107	0.103 (4%)	0.098 (9%)	[0.096 (11%)]	0.084 (27%)	0.089 (20%)	0.095 (13%)
$\Delta\tau$	0.077	0.075 (3%)	0.067 (15%)	[0.068 (13%)]	0.058 (33%)	0.066 (17%)	0.071 (8%)
Δb_1	0.099	0.132 (-25%)	**0.064** (55%)	[0.079 (25%)]	**0.055** (80%)	0.073 (36%)	-
Δb_2	-	0.072	0.055	[0.063]	0.051	-	-
Δb_1^{LRG}	-	-	-	-	0.101	0.129	0.14
$\Delta\sigma_8$	0.076	0.085 (-11%)	**0.049** (55%)	[0.057 (33%)]	**0.042** (81%)	0.056 (36%)	0.066 (15%)
Δh	0.0183	0.0164 (12%)	**0.0115** (59%)	[0.0123 (49%)]	**0.0096** (91%)	**0.0119** (54%)	0.014 (31%)
$\Delta\Omega_b$	0.0028	0.0025 (12%)	0.0022 (27%)	[0.0022 (27%)]	**0.0017** (65%)	0.0019 (47%)	0.0022 (27%)
	$k_{\rm max}^{MS} = 0.3\,h\,{\rm Mpc}^{-1}$					$k_{\rm max}^{LRG} = 0.2\,h\,{\rm Mpc}^{-1}$	
$\Delta\omega_d$	0.0068	0.0056 (21%)	0.0054 (26%)	[0.0053 (28%)]	**0.0044** (54%)	0.0050 (36%)	
$\Delta\omega_b$	0.00107	0.00094 (14%)	0.00087 (23%)	[0.00083 (29%)]	0.00080 (34%)	0.00092 (16%)	
$\Delta\Omega_\Lambda$	0.0120	0.0083 (45%)	0.0082 (46%)	[0.0074 (62%)]	**0.0066** (82%)	0.0084 (43%)	
Δn_s	0.031	0.026 (19%)	0.024 (29%)	[0.022 (41%)]	0.021 (48%)	0.025 (24%)	
ΔA_s	0.129	0.110 (17%)	0.096 (34%)	[0.094 (37%)]	**0.084** (54%)	0.108 (19%)	
Δw	0.105	0.090 (17%)	0.091 (15%	[0.085 (23%)]	0.080 (31%)	0.088 (19%)	
$\Delta\tau$	0.076	0.066 (15%)	0.061 (25%)	[0.059 (29%)]	0.053 (43%)	0.065 (17%)	
Δb_1	0.091	0.100 (-9%)	**0.052** (75%)	[0.062 (47%)]	**0.047** (94%)	0.071 (28%)	
Δb_2	-	0.046	0.042	[0.045]	0.040	-	
Δb_1^{LRG}	-	-	-	-	0.085	0.126	
$\Delta\sigma_8$	0.070	0.062 (13%)	**0.038** (84%)	[0.042 (67%)]	**0.035** (100%)	0.055 (27%)	
Δh	0.0160	0.0108 (48%)	**0.0091** (76%)	[0.0089 (80%)]	**0.0083** (93%)	0.0114 (40%)	
$\Delta\Omega_b$	0.0025	0.0018 (39%)	0.0018 (39%)	[0.0017 (47%)]	**0.0016** (56%)	0.0019 (32%)	

pectrum gives a better constraint on acoustic oscillations than the power spectrum. Unfortunately, at $z \simeq 0$ the acoustic oscillations are washed out by nonlinearities for $k \gtrsim 0.1\,h\,{\rm Mpc}^{-1}$ [70–73].

A fair assessment of the improvement on the detection of acoustic oscillations by using the bispectrum is beyond the scope of this paper and will be presented elsewhere. Here we note that our mock catalogs, although not exact in their nonlinear properties, do include the suppression of acoustic features, that is, we do not assume Eulerian second-order perturbation theory as done in Fig. 9 for illustrative purposes.

In order to assess the impact of acoustic features in our study we compute marginalized likelihoods using instead the BBKS fitting formula for the transfer function, the results are shown in Fig. 10 where we reproduce the same marginalized 95% CL contour plots given in Fig. 7. Since this transfer function depends exclusively on the shape parameter $\Gamma = \Omega_m h \exp[-\Omega_b(1+\sqrt{2h}/\Omega_m)]$ [74], we generically expect an enhanced degeneracy between Ω_m (or, equivalently here, Ω_Λ) and the Hubble parameter h. One can immediately notice in general a stronger degeneracy in all contour plots and, in particular, a more similar behavior of the bispectrum and power spectrum contours, especially for those involving Ω_Λ. This is the degeneracy that gets broken by acoustic features, as it is well known in the power spectrum case [65].

Focusing for instance on the Ω_Λ *vs.* h case in Fig. 10 and Fig. 7, we can see that the bispectrum, by virtue of its several different triangular configurations, is remarkably sensitive to features in the dark matter linear power spectrum such as the baryonic acoustic oscillations. The marginalized errors on individual parameters are, overall, larger when the BBKS transfer function is used. We notice, however, that the improvement provided by adding the bispectrum improves for parameters such as A_s, τ, b_1 and the spectral index n_s (about a factor of two better than the power spectrum alone) while it reduces for Ω_Λ.

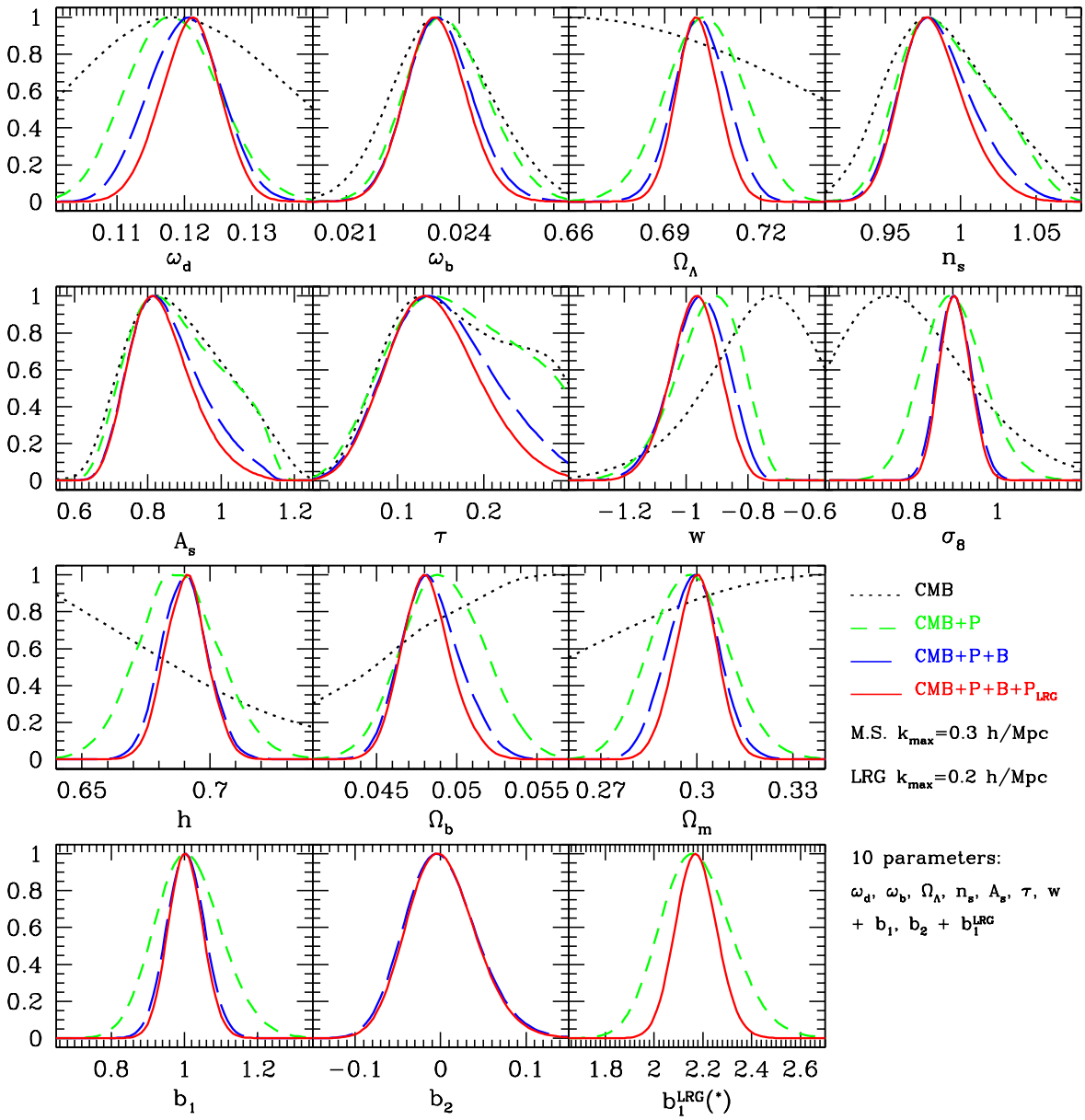

FIG. 11: wCDM models: marginalized likelihood functions for the six+four cosmological parameters assuming $k_{\rm max}^{MS} = 0.3\,h\,{\rm Mpc}^{-1}$. For the b_1^{LRG} parameter, the dotted line denotes the WMAP plus LRG power spectrum likelihood.

E. Dark Energy: wCDM models

We now extend the analysis performed above to include the determination of the dark energy equation of state parameter w under the assumption of an homogeneous dark energy component. We assume w to be constant. Introducing w, on which the growth function depends, leads to an increased degeneracy with the other parameters controlling the amplitude of the galaxy fluctuations such as A_s and b_1, which can be ameliorated by including bispectrum information.

Table VIII presents the expected errors on the various parameters for the two cases of $k_{\rm max}^{MS} = 0.2\,h\,{\rm Mpc}^{-1}$ and $0.3\,h\,{\rm Mpc}^{-1}$. We find that for $k_{\rm max}^{MS} = 0.2\,h\,{\rm Mpc}^{-1}$ the determination of the w parameter improves, over the W+P case, by 10% when the bispectrum is included, while by comparison a 20% if the LRG power spectrum is added instead. For $k_{\rm max}^{MS} = 0.3\,h\,{\rm Mpc}^{-1}$ adding the bis-

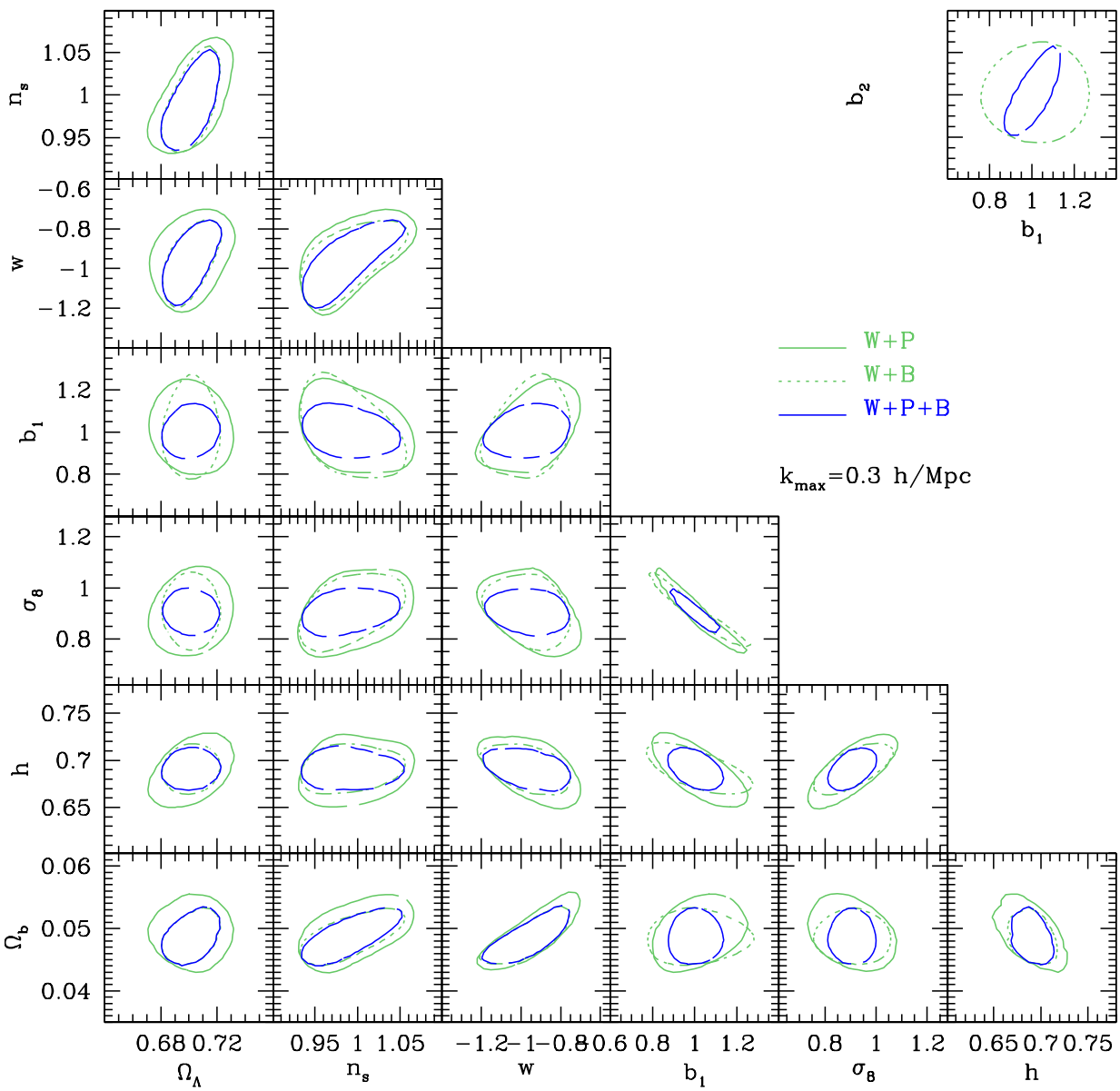

FIG. 12: wCDM models: marginalized 95% contour plots for a selection of the cosmological parameters assuming $k_{\rm max}^{MS} = 0.3\,h\,{\rm Mpc}^{-1}$.

pectrum improves the determination of w by 15%. These are mild improvements, but note that there are basically three different ways of getting below 10% errors on w by using the power spectrum of main and LRG samples, and the non-Gaussian information in the bispectrum and thus important consistency checks. Of course, the addition of extra information such as type IA supernovae or weak gravitational lensing (apart from the latest CMB data) will tighten the constraints.

In Fig. 11 we plot the marginalized likelihood functions for the case where the maximum wavenumber is $k_{\rm max}^{MS} = 0.3\,h\,{\rm Mpc}^{-1}$. Note that unlike the ΛCDM models previously considered, some of the maximum likelihood values of the first year WMAP data differ substantially from the chosen fiducial values for the LSS likelihood function. For instance, WMAP gives a marginalized probability distribution for w with a maximum close to -0.72, rather far from our fiducial value $w = -1$.

This implies that part of the constraing power of the LSS statistics is spent in shifting the maximum to the $w = -1$.

Finally, in Fig. 12 we show the contour plots for some selection of parameters. Comparing these results to the previous on ΛCDM models we see that the improvement brought by the bispectrum is increased in the case of b_1 and σ_8, and in fact the final error bars on these parameters (and b_2) are almost insensitive to including a more generic dark energy. This is good news as σ_8 is one of the least known parameters and subject to tension between different data sets [4, 75–78].

On the other hand, the constraints on Ω_Λ, A_s, τ and particularly ω_d and n_s are significantly worse than in the cosmological constant case. Regarding the behavior of the mixed covariance matrix, we see the same impact of beat coupling that we discussed before, the parameters responsible for the amplitude of galaxy fluctuations A_s, τ, σ_8 and bias parameters improve by the inclusion of the mixed covariance, though the behavior of the rest of the parameters is somewhat more complicated. Performing the same test as in Fig. 8 in this case we see that excluding the first k-bin completely erases the effects due to beat coupling.

VI. CONCLUSIONS

We have provided a first detailed study on the information about cosmological parameters contained in the bispectrum of the galaxy distribution at large scales, paying particular attention to the joint analysis with the power spectrum and their combination with CMB data. We have shown that the bispectrum has significant information on cosmological parameters and when combined with the power spectrum it is more complementary than combining power spectra of different samples of galaxies, since non-Gaussianity provides a somewhat different direction in parameter space. Moreover, replacing the power spectrum with the bispectrum gives similar constraints on cosmological parameters and can therefore serve as a consistency check.

In order to properly combine bispectrum with power spectrum information, we worked out their covariance properties. Due to the effects of beat coupling [61], the mixed terms in the covariance matrix are enhanced. We demonstrated that including this effect into the likelihood analysis provides a slight improvement on the error bars of cosmological parameters related to the amplitude of galaxy fluctuations.

In the framework of flat cosmological models we showed that most of the improvement of adding bispectrum information corresponds to parameters related to the amplitude and effective spectral index of perturbations, in particular Ω_m (or $\Omega_\Lambda = 1 - \Omega_m$) and σ_8, which can be improved by factors of 1.5 to 2 and, interestingly, are presently among the least well known. In particular, we showed that the uncertainties on σ_8 are robust to relaxing the equation of state parameter w beyond a cosmological constant. This is good news as σ_8 is subject to tension between different data sets [4, 75–78]. We also showed that the improvements are not directly a consequence of just constraining galaxy bias but of genuine information on cosmological parameters.

As far as future theoretical work is concerned, the single most pressing issue is that of systematic errors in the predictions, which were addressed in [52] for a previous generation of galaxy surveys. Those methods, based on second-order Lagrangian perturbation theory (which we have used here), are likely not enough given the expected statistical errors we derived in this work. Fortunately, powerful methods based on first principles have recently become available [79], and together with numerical simulations they should provide a sound theoretical basis. Work on this is in progress and will be reported soon.

Acknowledgments

We thank Josh Frieman for helpful comments and Mulin Ding for his technical assistance and his remarkable patience in maintaining our aging computer cluster Mafalda. We wish to thank Alan Sokal for giving us access to his group's computer cluster Guille, which was funded in part by NSF grant PHY–0424082. We thank NYU Information Technology Services for making its High Performance Computation Cluster available to us, and thank Joseph Hargitai for his technical assistance. E. S. is supported by the US Department of Energy and by NASA grant NAG5-10842 at Fermilab.

[1] M. Tegmark *et al.*, Phys. Rev. D **69**, 103501 (2004).
[2] A. G. Sanchez *et al.*, Mon. Not. R. Astron. Soc. **366**, 189 (2006).
[3] D. N. Spergel *et al.*, Astrophys. J. Suppl. **148**, 175 (2003).
[4] D. N. Spergel *et al.*, arXiv:astro-ph/0603449.
[5] J. A. Frieman and E. Gaztañaga, Astrophys. J. **425**, 392 (1994).
[6] J. A. Frieman and E. Gaztañaga, Astrophys. J. **521**, L83 (1999).
[7] I. Szapudi *et al.*, Astrophys. J. **570**, 75 (2002).
[8] E. Gaztañaga, Astrophys. J. **580**, 144 (2002).
[9] M. Takada and B. Jain, Mon. Not. R. Astron. Soc. **340**, 580 (2003).
[10] D. J. Croton *et al.*, Mon. Not. R. Astron. Soc. **352**, 1232 (2004).
[11] Y. P. Jing and G. Börner, Astrophys. J. **607**, 140 (2004).
[12] I. Kayo *et al.*, Pub. Astron. Soc. J. **56**, 413 (2004).
[13] Y. Wang *et al.*, Mon. Not. R. Astron. Soc. **353**, 300

(2004).
[14] E. Gaztañaga *et al.*, Mon. Not. R. Astron. Soc. **364**, 620 (2005).
[15] P. Fosalba, J. Pan and I. Szapudi, Astrophys. J. **632**, 29 (2005).
[16] J. N. Fry, Phys. Rev. Letters **73**, 215 (1994).
[17] R. Scoccimarro, H. A. Feldman, J. N. Fry and J. A. Frieman, Astrophys. J. **546**, 652 (2001).
[18] H. A. Feldman, J. A. Frieman, J. N. Fry and R. Scoccimarro, Phys. Rev. Lett. **86**, 1434 (2001).
[19] L. Verde *et al.*, Mon. Not. R. Astron. Soc. **335**, 432 (2002).
[20] R. E. Smith, P. I. R. Watts and R. K. Sheth, Mon. Not. R. Astron. Soc. **365**, 214 (2006).
[21] F. Bernardeau, L. van Waerbeke and Y. Mellier, Astron. Astrophys. **322**, 1 (1997).
[22] L. Hui, Astrophys. J. **519**, L9 (1999).
[23] A. Cooray and W. Hu, Astrophys. J. **548**, 7 (2001).
[24] M. Takada and B. Jain, Mon. Not. R. Astron. Soc. **348**, 897 (2004).
[25] J. N. Fry, Astrophys. J. **279**, 499 (1984).
[26] R. Scoccimarro and R. K. Sheth, Mon. Not. R. Astron. Soc. 329, 629 (2002).
[27] P. Coles, Mon. Not. R. Astron. Soc. **262**, 1065 (1993).
[28] R. J. Scherrer and D. .H. Weinberg, Astrophys. J. **504**, 607 (1998).
[29] J. N. Fry and E. Gaztañaga, Astrophys. J. **413**, 447 (1993).
[30] E. Gaztañaga and R. Scoccimarro, Mon. Not. R. Astron. Soc. **361**, 824 (2005).
[31] A. A. Berlind and D. H. Weinberg, Astrophys. J. **575**, 587 (2002).
[32] A. V. Kravtsov *et al.*, Astrophys. J. **609**, 35 (2004).
[33] D. Eisenstein *et al.*, Astrophys. J. **633**, 560 (2005).
[34] I. Zehavi *et al.*, Astrophys. J. **621**, 22 (2005).
[35] R. K. Sheth and G. Tormen, Mon. Not. R. Astron. Soc. **329**, 61 (2002).
[36] R. Scoccimarro, R. K. Sheth, L. Hui and B. Jain, Astrophys. J. **546**, 20 (2001).
[37] R. Scoccimarro, M. Zaldarriaga and L. Hui, Astrophys. J. **527**, 1 (1999).
[38] A. Meiksin and M. White, Mon. Not. R. Astron. Soc. **308**, 1179 (1999).
[39] Y. P. Jing, Astrophys. J. **503**, L9 (1998).
[40] U. Seljak and M. S. Warren, Mon. Not. R. Astron. Soc. **355**, 129 (2004).
[41] J. L. Tinker, D. H. Weinberg, Z. Zheng and I. Zehavi, Astrophys. J. **631**, 41 (2005).
[42] H. J. Mo and S. D. M. White, Mon. Not. R. Astron. Soc. **282**, 347 (1996).
[43] R. K. Sheth and G. Tormen, Mon. Not. R. Astron. Soc. **308**, 119 (1999).
[44] R. K. Sheth, H. J. Mo and G. Tormen, Mon. Not. R. Astron. Soc. **323**, 1 (2001).
[45] Y. P. Jing, H. J. Mo and S. D. M. White, Mon. Not. R. Astron. Soc. **284**, 189 (1997).
[46] L. Verde *et al.*, Astrophys. J. Suppl. **148**, 195 (2003).
[47] H. A. Feldman, N. Kaiser and J. A. Peacock, Astrophys. J. **426**, 23 (1994).
[48] J. M. Bardeen, J. R. Bond, N. Kaiser and A. S. Szalay, Astrophys. J. **304**, 15 (1986).
[49] U. Seljak and M. Zaldarriaga, Astrophys. J. **469**, 437 (1996).
[50] R. Scoccimarro, S. Colombi, J. N. Fry, J. A. Frieman, E. Hivon and A. Melott, Astrophys. J. **496**, 586 (1998).
[51] R. Scoccimarro, H. M. P. Couchman and J. A. Frieman, Astrophys. J. **517**, 531 (1999).
[52] R. Scoccimarro, Astrophys. J. **544**, 597 (2000).
[53] S. Matarrese, L. Verde and A.F. Heavens, Mon. Not. R. Astron. Soc. **290**, 651 (1997).
[54] I. Szapudi, S. Colombi, A. Jenkins and J. Colberg, Mon. Not. R. Astron. Soc. **313**, 725 (2000).
[55] N. Kaiser, Mon. Not. R. Astron. Soc. **227**, 1 (1987).
[56] H. B. Sandvik, M. Tegmark, X. M. Wang and M. Zaldarriaga, Phys. Rev. D **69**, 063005 (2004).
[57] L. Page *et al.*, arXiv:astro-ph/0603450.
[58] R. Scoccimarro, E. Sefusatti and M. Zaldarriaga, Phys. Rev. D **69**, 103513 (2004).
[59] E. Sefusatti and R. Scoccimarro, Phys. Rev. D **71**, 063001 (2005).
[60] E. V. Linder, Phys. Rev. D **72**, 043529 (2005).
[61] A. J. S. Hamilton, C. D. Rimes and R. Scoccimarro, arXiv:astro-ph/0511416.
[62] C. D. Rimes and A. J. S. Hamilton, arXiv:astro-ph/0511418.
[63] D. J. Eisenstein *et al.*, Astrophys. J. **633**, 560 (2005).
[64] R. Scoccimarro, Astrophys. J. **487**, 1 (1997).
[65] D. J. Eisenstein, W. Hu and M. Tegmark, Astrophys. J. **518**, 2 (1999).
[66] P. J. E. Peebles and J. T. Yu, Astrophys. J. **162**, 815 (1970).
[67] J. R. Bond and G. Efstathiou, Astrophys. J. **285**, L45 (1984).
[68] S. Cole *et al.*, Mon. Not. R. Astron. Soc. **362**, 505 (2005).
[69] J. A. Frieman and M. Joffre, unpublished.
[70] A. Meiksin, M. White and J. A. Peacock, Mon. Not. R. Astron. Soc. **304**, 851 (1999).
[71] D. J. Eisenstein, Astrophys. J. **633**, 575 (2005).
[72] V. Springel *et al.*, Nature **435**, 269 (2005).
[73] M. White , Astropart. Phys. **24**, 334 (2005).
[74] N. Sugiyama, Astrophys. J. **100** (Suppl.), 281 (1995).
[75] M. L. Brown *et al.*, Mon. Not. R. Astron. Soc. **341**, 100 (2003).
[76] H. Hoekstra *et al.*, arXiv:astro-ph/0511089.
[77] M. Viel, M. G. Haehnelt and A. Lewis, arXiv:astro-ph/0604310.
[78] U. Seljak, A. Slozar and P. McDonald, arXiv:astro-ph/0604335.
[79] M. Crocce and R. Scoccimarro, Phys. Rev. D **73**, 063519 (2006); *ibid*, Phys. Rev. D **73**, 063520 (2006).