

# Improved primordial non-Gaussianity constraints from measurements of galaxy clustering and the integrated Sachs-Wolfe effect

Tommaso Giannantonio,<sup>1,2,\*</sup> Ashley J. Ross,<sup>3</sup> Will J. Percival,<sup>3</sup> Robert Crittenden,<sup>3</sup> David Bacher,<sup>1,2</sup> Martin Kilbinger,<sup>4</sup> Robert Nichol,<sup>3</sup> and Jochen Weller<sup>1,2,5</sup>

<sup>1</sup>Ludwig-Maximilians-Universität München, Universitäts-Sternwarte München, Scheinerstr. 1, D-81679 München, Germany

<sup>2</sup>Excellence Cluster Universe, Technical University Munich, Boltzmannstr. 2, D-85748 Garching, Germany

<sup>3</sup>Institute of Cosmology and Gravitation, University of Portsmouth, Portsmouth, PO1 3FX, UK

<sup>4</sup>Laboratoire AIM, CEA Saclay, Irfu/Sap (Service d'Astrophysique), Orme des Merisiers, Bât 709, F-91191 Gif-sur-Yvette, France

<sup>5</sup>Max-Planck-Institut für Extraterrestrische Physik, Giessenbachstr., D-85748 Garching, Germany  
(Received 8 March 2013; published 16 January 2014)

We present the strongest robust constraints on primordial non-Gaussianity (PNG) from currently available galaxy surveys, combining large-scale clustering measurements and their cross correlations with the cosmic microwave background. We update the data sets used by Giannantonio *et al.* (2012), and broaden that analysis to include the full set of two-point correlation functions between all surveys. In order to obtain the most reliable constraints on PNG, we advocate the use of the cross correlations between the catalogs as a robust estimator and we perform an extended analysis of the possible systematics to reduce their impact on the results. To minimize the impact of stellar contamination in our luminous red galaxy sample, we use the recent Baryon Oscillations Spectroscopic Survey catalog of Ross *et al.* (2011). We also find evidence for a new systematic in the NVSS radio galaxy survey similar to, but smaller than, the known declination-dependent issue; this is difficult to remove without affecting the inferred PNG signal, and thus we do not include the NVSS autocorrelation function in our analyses. We find no evidence of primordial non-Gaussianity; for the local-type configuration we obtain for the skewness parameter  $-36 < f_{\text{NL}} < 45$  at 95% C.L. ( $5 \pm 21$  at  $1\sigma$ ) when using the most conservative part of our data set, improving previous results; we also find no evidence for significant kurtosis, parameterized by  $g_{\text{NL}}$ . In addition to PNG, we simultaneously constrain dark energy and find that it is required with a form consistent with a cosmological constant.

DOI: 10.1103/PhysRevD.89.023511

PACS numbers: 98.80.Es, 98.65.Dx

## I. INTRODUCTION

It is widely assumed that the density perturbations that seeded the cosmic microwave background (CMB) anisotropies and later induced the formation of cosmic structure were produced by primordial quantum fluctuations in the early universe [1]. An early inflationary phase is then necessary to stretch such fluctuations beyond the horizon, and to address a number of other issues related to the initial conditions of the big bang. While the general paradigm of inflation is widely accepted, the precise details of the physical mechanism that produced an inflationary phase in the early universe remain largely unconstrained (see Ref. [2] for a recent review). However, any specific inflationary model is testable, as it leaves observable traces in the later evolution. For example, the simplest single-field slow-roll inflation model predicts a nearly flat universe and near scale-invariance of the primordial perturbations, both of which have been verified e.g. by the WMAP satellite

[3–6]. Additionally, the primordial perturbations are predicted to be purely adiabatic, which is also indicated (see Ref. [7] for recent constraints on isocurvature). Finally, the distribution of primordial perturbations should be nearly Gaussian [8,9]: testing this last prediction is our present focus.

Determining the amount of primordial non-Gaussianity (PNG) is instrumental in distinguishing between models of the early universe [10]. This has been typically constrained from the higher-order statistics (such as the bispectrum) of the CMB [4,11]; the latest CMB measurements from the nine-year analysis of data from the WMAP satellite constrain local PNG, as quantified by the skewness parameter  $f_{\text{NL}}$  (defined in Section II) to be  $-3 < f_{\text{NL}} < 77$  at 95% confidence [5,6] [12]. Recent CMB constraints on the kurtosis parameter  $g_{\text{NL}}$  are  $-7.4 \times 10^5 < g_{\text{NL}} < 8.2 \times 10^5$  [13]. Higher-order statistics of the large-scale structure (LSS) may also be used [14,15], but the difficulty of this approach is in disentangling PNG from late-time non-Gaussianity produced by nonlinear structure

\*tommaso.giannantonio@usm.lmu.de

formation. This is also an issue if using the effect on the nonlinear contributions to the matter power spectrum [16], probed e.g. by weak gravitational lensing [17]. The abundance of massive clusters is also a sensitive, complementary probe of PNG [18–21], and is less sensitive to later nonlinear evolution.

In addition to these established methods, a new technique has been discovered and developed in the recent years: it has been shown that PNG of the local and orthogonal types alters the biasing of dark matter halos and galaxies with respect to the underlying density field, making the bias strongly scale-dependent; thus by measuring the galaxy bias it is possible to constrain PNG [22–32]. This has been exploited by some authors [33–38], finding constraints which are competitive with CMB bispectrum analyses: Ref. [33] found  $-29 < f_{\text{NL}} < 70$  at 95%, while Ref. [37] reported hints of detection at  $8 < f_{\text{NL}} < 88$  (95%) for the local model. A measurement of  $g_{\text{NL}}$  was obtained from LSS data by Ref. [39] assuming  $f_{\text{NL}} = 0$  by rescaling the  $f_{\text{NL}}$  constraints by Ref. [33]:  $-3.5 \times 10^5 < g_{\text{NL}} < 8.2 \times 10^5$  (95%). The bias of galaxy clusters is similarly affected by PNG, and has been used in combination with cluster abundances to constrain  $f_{\text{NL}}$  [40]. The scale-dependent bias technique will remain competitive with the arrival of future galaxy and cluster surveys, and is potentially more powerful than the cosmic-variance limit of the CMB [41–47].

The large-scale divergence in galaxy bias produced by PNG can also be observed through cross correlations of galaxy surveys with the CMB, provided some of the CMB anisotropies are created locally. This is conveniently provided by the integrated Sachs-Wolfe (ISW) effect [48], where small secondary CMB anisotropies are induced by the decay of the gravitational potentials when the Universe undergoes a transition from a dark matter to a dark energy or curvature dominated phase. While the nature of dark energy is still unknown, a broad range of observations have confirmed its presence [49], including observations of the ISW effect [see Ref. [50], and references therein]. In most realistic models the ISW effect is too small to be directly observed in the primary CMB temperature power spectrum, but it can be detected by cross-correlating the CMB temperature anisotropies with tracers of the potentials, such as galaxy catalogs [51]. The latest compilations which examine multiple catalogs are those by Refs. [52] and [53] (G08 herein); the latter was recently updated in Ref. [50] (G12 herein) and we refer to this paper for a more detailed introduction and a longer description of the ISW effect and the data. These cross correlations are sensitive to the galaxy bias, and therefore also depend on PNG as above.

In this paper we use an updated compilation of LSS and ISW data to make robust and accurate PNG measurements. We first review the galaxy samples that we consider, revising those used by G12. In particular, we include the

latest available luminous red galaxy (LRG) photometric redshift sample [54], also used for ISW measurements by Ref. [55], which were calibrated using data from the Baryon Oscillation Spectroscopic Survey (BOSS) [56] CMASS sample. Refs. [54,57,58] have shown that properly addressing the possibility of spurious fluctuations in the observed number of galaxies is crucial in order to measure and interpret LSS measurements. References [38,59,60] have shown that these issues are particularly important when measuring PNG. We address these concerns in all of our analyses and study the impact of some well-known systematic effects present in the radio-galaxies of the NVSS survey.

We then explore the cosmological consequences of our data, using a nested-sampling Monte Carlo Markov chain (MCMC) method and marginalizing over extra nuisance parameters to account for a range of possible uncertainties and systematics. We derive cosmological constraints for both PNG and dark energy, extending the data set to include both the galaxy-CMB and all the galaxy-galaxy correlations, and include their full covariance. Indeed, we demonstrate the importance of the cross correlations between different galaxy catalogs as a robust method to estimate the amount of PNG in the data.

The plan of this paper is as follows: we briefly review the theoretical basis of the LSS in the presence of PNG and the ISW effect in Section II. We then describe our data set and the most important systematics in Section III, while Section IV clarifies how we model the observations. In Section V we discuss the cosmological interpretation of our data, before concluding in Section VI.

## II. THEORY

### A. Primordial non-Gaussianity and the LSS

We consider here deviations from Gaussianity at the three- and four-point level (skewness and kurtosis). For PNG of the local type, the Bardeen potential at early times ( $z_*$ ) is assumed to be written in terms of an auxiliary Gaussian potential  $\varphi$  as

$$\Phi(\mathbf{x}, z_*) = \varphi(\mathbf{x}, z_*) + f_{\text{NL}}[\varphi^2(\mathbf{x}, z_*) - \langle \varphi^2 \rangle(z_*)] + g_{\text{NL}}[\varphi^3(\mathbf{x}, z_*) - 3\langle \varphi^2 \rangle(z_*)\varphi(\mathbf{x}, z_*)], \quad (1)$$

where  $f_{\text{NL}}$  and  $g_{\text{NL}}$  quantify the amount of PNG. In this work we will also study simple deviations from the local case, as described below.

It was originally shown by Ref. [22] that primordial non-Gaussianity induces a strongly scale-dependent bias, whose features appear on the largest scales for both dark-matter halos and galaxies, and their analysis was later clarified and improved [24,28,33,61–65]. The scale dependence originates through the coupling of long- and short-wavelength modes in the density perturbations which arise in the presence of PNG by applying the peak-background split

technique to Eq. (1). In Fourier space, we can use the linear Poisson equation to show that the late-time density can be related to the initial potential field by

$$\tilde{\delta}(k, z) \simeq \alpha(k, z) \tilde{\phi}(k, z_\star), \quad (2)$$

where the tilde denotes Fourier transformation. Here,

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

where  $T(k)$  is the density transfer function,  $D(z)$  is the linear growth function,  $g(z) \propto (1+z)D(z)$  is the potential growth function, for which  $g(z_\star)/g(0) \simeq 1.306$ .

In the presence of PNG, the Gaussian bias  $b_{\text{Gauss}}$  is altered by two corrective terms, giving an effective bias

$$b_{\text{eff}}(k, f_{\text{NL}}, g_{\text{NL}}) = b_{\text{Gauss}} + \delta b(f_{\text{NL}}, g_{\text{NL}}) + \Delta b(k, f_{\text{NL}}, g_{\text{NL}}). \quad (4)$$

The first, scale-independent correction  $\delta b(f_{\text{NL}}, g_{\text{NL}})$  is produced by the effect of PNG on the halo mass function and it is typically negligible compared with the scale-dependent part  $\Delta b(k, f_{\text{NL}}, g_{\text{NL}})$ ; in the following we will absorb it into the constant part of the bias, which we define as  $b_1(f_{\text{NL}}, g_{\text{NL}}) \equiv b_{\text{Gauss}} + \delta b(f_{\text{NL}}, g_{\text{NL}})$ . For local-type non-Gaussianity, the scale-dependent term is (at any given redshift) [65]

$$\Delta b^{\text{loc}}(k, f_{\text{NL}}, g_{\text{NL}}) = \frac{\beta_f f_{\text{NL}} + \beta_g g_{\text{NL}}}{\alpha(k)}. \quad (5)$$

The coefficients  $\beta_f$  and  $\beta_g$  will in general depend on redshift and mass (or bias). Here the  $f_{\text{NL}}$  part is

$$\beta_f = 2\delta_c b_L \quad (6)$$

where  $\delta_c = 1.686$  is the spherical collapse threshold and  $b_L \equiv b_1 - 1$  is the Lagrangian bias.

For the  $g_{\text{NL}}$  contribution  $\beta_g$ , we use the fit to  $N$ -body simulations of the Edgeworth expansion introduced by Ref. [65]:

$$\beta_g \simeq \kappa_3^{(1)}(M) [-0.7 + 1.4(\tilde{\nu} - 1)^2 + 0.6(\tilde{\nu} - 1)^3] - \frac{d\kappa_3^{(1)}(M)}{d \ln \sigma^{-1}} \left( \frac{\tilde{\nu} - \tilde{\nu}^{-1}}{2} \right), \quad (7)$$

with  $\tilde{\nu} \equiv \sqrt{\delta_c b_L + 1}$ , where  $\tilde{\delta}_c = 1.42$ , which was found by Ref. [65] to improve the agreement between the halo collapse model and  $N$ -body simulations with respect to the Press-Schechter value. We adopt this value in the  $g_{\text{NL}}$  fitting formulae to be consistent with the derivation of Ref. [65] (although we keep the standard value when fitting  $f_{\text{NL}}$ , as the two parts of the model are separable).  $\kappa_3^{(1)}(M)$  is the

skewness of the density field for the case  $f_{\text{NL}} = 1$ , smoothed at a mass  $M$ , and  $\sigma$  the r.m.s. of the same field at the same scale. For the skewness and its derivative we use the fitting functions given in Eqs. (46)–(47) of Ref. [65]. Though their redshift dependence is different, both  $f_{\text{NL}}$  and  $g_{\text{NL}}$  cause the bias to become scale-dependent with the same power of  $k$ , albeit with different amplitudes; this makes distinguishing between the two parameters challenging [66].

The fitting formula of Eq. (7) comes with an important caveat, as it was obtained for a narrow range in halo mass and redshift [65]; furthermore, it was shown that the effect of  $g_{\text{NL}}$  changes depending on the halo occupation distribution of the used tracers, not only on the Gaussian bias, and this effect becomes more severe for less biased samples, such as in our case. This formula is, however, the best currently available, and the form of the constraints obtained using it will be indicative of those that would be available using a more accurate fit. We therefore adopt the formula, testing for specific non-Gaussian deviations of this form, making sure that when inferences are drawn from the results we remember these might be different from true  $g_{\text{NL}}$ -like deviations.

Any residual theoretical uncertainties on the local  $f_{\text{NL}}$  contribution to the bias are less significant, as the linear model of Eq. (6) has been tested both theoretically and numerically with  $N$ -body simulations by multiple authors [24,28,33,61–65]. Nonlinear effects have also been extensively studied by some of these authors, and they were shown to be negligible on the scales of interest; finally, while general-relativistic effects are present on the largest scales, they are small and do not affect the PNG constraints from the bias [67]. Even if small residual inaccuracies were present, they would primarily affect the interpretation of a detection, while the probability of  $f_{\text{NL}}$  being nonzero should not significantly change.

Scale-dependent bias was first used to constrain  $f_{\text{NL}}$  from the LSS by Refs. [33,62] with varying results, possibly caused by different assumptions for the bias evolution with redshift. Ref. [39] obtained bounds on  $g_{\text{NL}}$  from the same data consistent with Gaussian initial conditions. Some later work [35–37] found hints of detection of a positive  $f_{\text{NL}}$ , while more recent analyses that take into account possible systematic errors in the large-scale clustering measurements have found measurements consistent with no PNG [38].

Similar expressions have been derived for other varieties of PNG, such as orthogonal and equilateral [63,64], featuring weaker or no scale dependence. Furthermore, some models also predict a scale dependence of  $f_{\text{NL}}$  [68,69], often parameterized with a spectral index  $n_{f_{\text{NL}}}$ : recently, the first constraint on this quantity has been found  $n_{f_{\text{NL}}} = 0.30^{+1.9}_{-1.2}$  at 95% C.L. from CMB data [70]. Given the difficulty in distinguishing between different models with current bias measurements, and to be as general as

possible, one can adopt a simpler phenomenological approach by allowing the scale-dependent term to have an arbitrary amplitude and scaling index. To keep the interpretation of the parameter constraints closer to the usual local case, we adopt a parameterization

$$\Delta b^{\text{gen}}(k, f_{\text{NL}}) = \frac{2f_{\text{NL}}^{\text{gen}}\delta_c b_L}{\alpha'(k, a_{\text{NL}})}, \quad (8)$$

where we have only one extra parameter  $a_{\text{NL}}$ , changing the slope of the bias correction:

$$\alpha'(k, a_{\text{NL}}, z) = \frac{2k^{a_{\text{NL}}}T(k)D(z)g(0)}{3\Omega_m H_0^2 g(z_*)}. \quad (9)$$

The local, scale-independent case is recovered for  $a_{\text{NL}} = 2$ . Notice also that this can be related to the local spectral index as

$$k^{a_{\text{NL}}} = k^2 \left( \frac{k_{\text{piv}}}{k} \right)^{n_{f_{\text{NL}}}}, \quad (10)$$

where  $k_{\text{piv}}$  is the pivot scale. We will use  $k_{\text{piv}} = 0.002h \text{ Mpc}^{-1}$  in the following as in Ref. [43].

### B. The ISW effect and large-scale structure

For a recent, detailed summary of the ISW theory and its observational status, we refer to Ref. [50]. Briefly, if the gravitational potentials  $\Phi, \Psi$  are evolving in time, they will imprint secondary temperature anisotropies  $\Theta \equiv \delta T/T$  in the CMB in a direction  $\hat{\mathbf{n}}$  given by

$$\Theta(\hat{\mathbf{n}}) = - \int e^{-\tau(\eta)} (\dot{\Phi} + \dot{\Psi}) [\eta, \hat{\mathbf{n}}(\eta_0 - \eta)] d\eta, \quad (11)$$

where  $\eta$  is conformal time, the dots are derivatives with respect to it, and  $\tau$  is the optical depth of CMB photons. Similarly, for a galaxy catalog  $i$  with redshift distribution  $\varphi_i(z) \equiv dN_i/dz(z)$ , its expected overdensity is

$$\delta_{g_i}(\hat{\mathbf{n}}) = \int b_{g_i}(z)\varphi_i(z)\delta(\hat{\mathbf{n}}, z)dz, \quad (12)$$

where  $b_{g_i}(z)$  is the galactic bias and  $\delta(\hat{\mathbf{n}}, z)$  the dark matter overdensity.

From these relations, the auto- and cross-spectra in harmonic space can be constructed assuming linear theory as

$$\begin{aligned} C_l^{g_i g_j} &= \frac{2}{\pi} \int dk k^2 P(k) W_l^{g_i}(k) W_l^{g_j}(k) \\ C_l^{T g_i} &= \frac{2}{\pi} \int dk k^2 P(k) W_l^T(k) W_l^{g_i}(k); \end{aligned} \quad (13)$$

here  $P(k)$  is the linear matter power spectrum at  $z = 0$ , and the sources for galaxies and the ISW temperature anisotropies are (assuming  $\Phi = \Psi$ )

$$\begin{aligned} W_l^T(k) &= -\frac{3\Omega_m H_0^2}{k^2} \int dz e^{-\tau(z)} g'(z) j_l[k\chi(z)] \\ W_l^{g_i}(k) &= \int dz b_{g_i}(k, z) \phi_i(z) D(z) j_l[k\chi(z)], \end{aligned} \quad (14)$$

where  $\chi$  is the comoving distance and  $j_l$  the spherical Bessel function. These spectra are related to the auto- and cross-correlation functions (ACF and CCF), defined as

$$\begin{aligned} w^{g_i g_j}(\vartheta) &\equiv \langle \delta_{g_i}(\hat{\mathbf{n}}) \delta_{g_j}(\hat{\mathbf{n}}') \rangle \\ w^{T g_i}(\vartheta) &\equiv \langle \delta_{g_i}(\hat{\mathbf{n}}) \Theta(\hat{\mathbf{n}}') \rangle, \end{aligned} \quad (15)$$

where the averages are carried over all pairs at a distance  $\vartheta = |\hat{\mathbf{n}} - \hat{\mathbf{n}}'|$ . They are derived from the spectra by Legendre transformation:

$$w(\vartheta) = \frac{1}{4\pi} \sum_{l=0}^{\infty} (2l+1) C_l P_l[\cos(\vartheta)], \quad (16)$$

where  $P_l$  are the Legendre polynomials.

### III. DATA AND SYSTEMATICS

It is well known [38,59,60] that the large-scale clustering of the LSS is easily contaminated by systematic distortions in the galaxy samples caused by observational issues, which may significantly bias the PNG results. Indeed, any spurious long-wavelength contaminant, produced in the Galaxy (extinction-causing dust, stars, synchrotron emission) or even in the atmosphere (seeing, air mass) will generally introduce extra large-scale power, which may naively be attributed to PNG. This issue only has a limited impact when studying the ISW correlation with the CMB, as there is no particular reason why such systematics should correlate with the primordial temperature fluctuations; however it has the potential to substantially bias any result using the large-scale angular autocorrelation of these galaxies. However, as in the case of the galaxy-CMB correlations, cross correlations between independent galaxy catalogs will be in general less subject to such systematics, and can therefore provide estimates of the large-angle clustering that are more robust than the auto correlations.

To minimize these effects, and to obtain robust measurements of PNG, we study a number of data sets previously considered in G12, including the main galaxy, LRG and quasar samples from the Sloan Digital Sky Survey (SDSS) [71], as well as independent infrared, radio and x-ray surveys. We examine a number of these in greater depth in order to address the impact of potential observational systematics on the large-scale clustering measurements.

### A. Data and general considerations

Below we discuss potential contaminants to our LRG, radio and quasar samples. The other data sets are unchanged from G12: The 2MASS infrared catalog [72] contains 415,459 galaxies after masking, has a median redshift  $\bar{z} = 0.09$  and a bias  $b = 1.3$  if it is assumed constant. The final imaging SDSS Data Release 8 (DR8) main galaxy sample [73] contains 30 million galaxies at  $\bar{z} = 0.3$  and  $b = 1.2$ , while the HEAO survey of the x-ray background [74] has a broad redshift distribution of estimated  $\bar{z} = 0.9$  and  $b = 1.0$ . See G12 for further details.

We pixellate all maps using the HEALPIX [75] scheme at low resolution  $N_{\text{side}} = 64$ , corresponding to a pixel size of  $\sim 50$  arcmin, since we are primarily interested in large scales. We measure the angular two-point correlation functions between two pixellated maps  $a, b$  of weighted masks  $f^a, f^b$  using the estimator

$$\hat{w}^{ab}(\vartheta) = \frac{1}{N_\vartheta} \sum_{i,j=1}^{N_{\text{pix}}} \left( \frac{n_i^a}{\bar{n}^a} - 1 \right) \left( \frac{n_j^b}{\bar{n}^b} - 1 \right) f_i^a f_j^b, \quad (17)$$

where  $N_\vartheta = \sum_{i,j} f_i^a f_j^b$  is the weighted number of pixels at separation  $\vartheta$ ,  $\bar{n}$  indicates the mean expectation value in each pixel and  $N_{\text{pix}}$  is the total number of pixels in a map. This generalizes Eq. (10) of G12.

We account for the pixelization effects by smoothing the theoretical correlation functions with the HEALPIX window function in harmonic space. We checked that a higher resolution ( $N_{\text{side}} = 128$ ) leaves the results unchanged.

Note that the overlap between catalogs is significant, both in redshift and in footprint: this leads to significant cross correlations and covariances in most cases, as also discussed in G08, G12.

### B. The BOSS CMASS galaxies

The MegaZ LRGs [76–78] used in G08 and G12 have an excess of power on large scales with respect to the  $\Lambda$ CDM model [78]. It was shown by Ref. [54] that the likely cause for this excess is a systematic contribution to the clustering from stellar contamination. Here, we have decided to instead use an updated catalog of LRGs at redshifts  $z \sim 0.55$ , constructed from the final imaging DR8 of the SDSS [54], where the systematics are better understood. We use the DR8 photometric redshift (photoz) sample rather than the data release nine (DR9) [79] spectroscopic redshift sample recently used for multiple cosmological analyses [38,80–86], because of the larger cosmological volume covered. Even though it relies on photometric rather than spectroscopic redshifts, it covers a usable area of  $\sim 9,000 \text{ deg}^2$ , rather than  $3,000 \text{ deg}^2$  for the spectroscopic sample, giving an increase in volume that allows better large-scale clustering measurements.

### 1. The catalog and its correlations

The DR8 LRG sample was selected using the same color and magnitude cuts as the SDSS-III BOSS CMASS (constant mass) sample, defined in Ref. [56], using imaging data from the entire DR8 footprint. Here we will denote it as the CMASS sample. It used more than 100,000 BOSS spectroscopic redshifts for training to estimate photometric redshifts and evaluating the probability that an object is a galaxy. (3% of objects that are selected are stars.) A detailed analysis of possible systematic uncertainties was performed by Ref. [54], including seeing effects, sky brightness and possible stellar contamination. The most serious systematic problem found with present photometric redshift catalogs from SDSS imaging was the loss of faint galaxies around stars (even relatively faint stars to  $r \approx 20$ ). Reference [54] corrected this issue, which clearly illustrates the need to be diligent about stellar contamination, particularly when cross-correlating with other data sets which may also include some contamination from galactic sources.

We use CMASS LRGs with  $0.45 < z_{\text{phot}} < 0.65$  and color term (see Eq. (13) in Ref. [54]),  $c_{\parallel} > 1.6$ . This is an identical sample to that used in Refs. [57,87,88], except that we do not further divide the sample by photometric redshift. Our map exploits the “ $A_{\text{star}}$ ” method (see Section 4.1 of Ref. [54]) to correct for the systematic effect of stars. We also correct for the offset between SDSS photometry in the north and south galactic caps [89] (NGC/SGC from here on) using the method applied to obtain the “ $\Delta$ South” results of Ref. [54]. The north/south offset for the CMASS LRGs was applied by Ref. [54] and it was recommended that it should be done for all analyses with the catalog. However, it mainly affects the measured clustering for  $z_{\text{phot}} > 0.6$ , and it is a small effect; as it can be seen in Fig. 14 of Ref. [54], this makes a small difference in the  $0.6 < z_{\text{phot}} < 0.65$  bin only, so that it has a negligible effect in our case, where we do not cut in redshift bins. We use the same mask for this sample as constructed in Ref. [57], but use a marginally more conservative Galactic extinction cut of  $A_r < 0.18$  (compared to the fiducial  $A_r < 0.20$ ), for consistency with our other data sets and previous analyses. This yields a sample with 801,226 LRGs occupying a footprint of  $8,891 \text{ deg}^2$ . The redshift distribution of the sample is centered around  $\bar{z} \approx 0.5$ , as shown in Fig. 1. By fitting the ACF, we have confirmed the bias  $b \approx 2.1$ , found by earlier work [54].

We calculate the auto- and cross-correlation functions of these LRGs and the WMAP7 internal linear combination (ILC) map using the estimator of Eq. (17), binning the results in the range  $0 \leq \vartheta \leq 12$  degrees. We also estimate the full covariance matrix using the Monte Carlo method MC2 of G08. We see the results in Fig. 2, where we also compare them with the correlations of the MegaZ data. While the ACFs are comparable at small scales, the excess large-scale power is significantly reduced in the CMASS data. Furthermore, the correlation of the CMASS data with

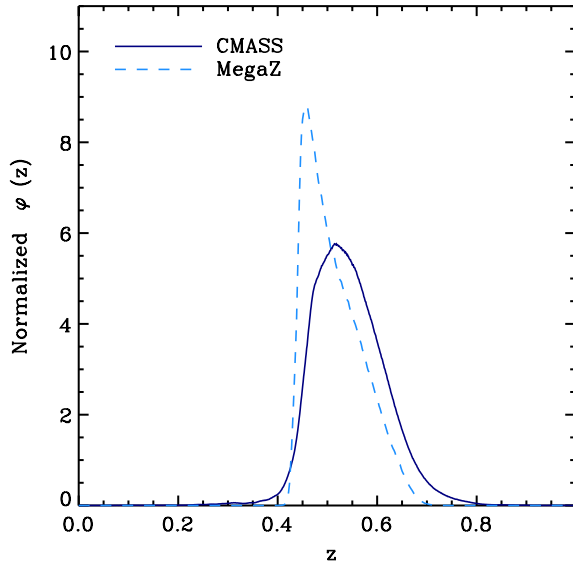


FIG. 1 (color online). Redshift distributions of the different LRG samples. The CMASS DR8 data are less peaked, but the difference in the distribution is small for the ISW purposes.

the CMB is much lower than for the MegaZ, and on most scales comparable to the theoretical prediction for a WMAP7 best-fit  $\Lambda$ CDM cosmology (assuming a bias  $b = 2.1$ , and no PNG), remaining well within the (highly correlated)  $1\text{-}\sigma$  error bars. Any significant deviations could reveal either residual systematics, or PNG.

## 2. Discussion

The differences between the CMASS and MegaZ CCFs seen in Fig. 2 are large compared to their errors, and it is interesting to investigate the reasons for this discrepancy in more detail.

*North-south divide.*—The footprint of the CMASS data is larger than that of the MegaZ data (8,891  $\text{deg}^2$  compared to 7,400  $\text{deg}^2$  after masking), with most of the additional area resulting from imaging data in the SGC [73]. As a first check, we tested whether the difference in the CCF signal is produced by the different sky coverage of the data. For this purpose, we measured the CCF of the CMASS data in the DR7 region only. We found that the signal increases significantly, as shown in Fig. 2 (red line), although it still does not match the MegaZ result, which used DR7 (i.e., mostly the NGC data).

We next examined whether the NGC and SGC footprints of DR8 give compatible ISW signals. As we can see in Fig. 3, this may not be the case, as the NGC of the DR8 data gives results which are in better agreement with the MegaZ catalog, but the SGC shows little or negative cross-correlation signal, and thus when added to the NGC data, lowers the overall CCF. However, given the relatively small coverage of DR8 in the SGC (1,300  $\text{deg}^2$ ), we should be wary of over-interpreting the results as the cosmic variance

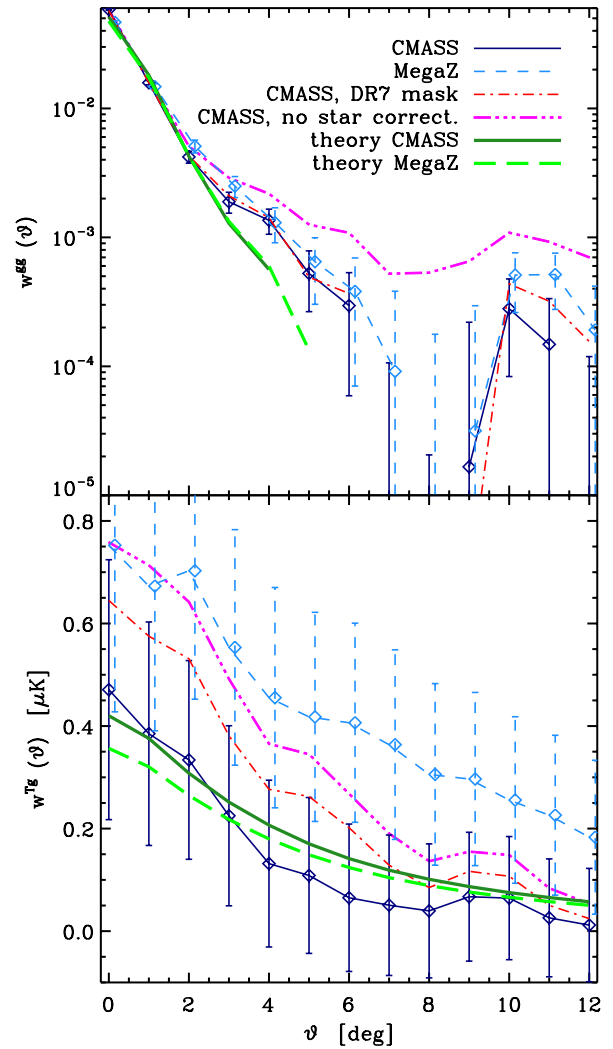


FIG. 2 (color online). Autocorrelation functions and cross correlations with WMAP (top and bottom panels, respectively) for the LRG data sets. The  $1\text{-}\sigma$  error bars are derived from Monte Carlo and are highly correlated; the data points have been slightly displaced in  $\vartheta$  for clarity. Dark and light blue data show the CMASS and MegaZ LRGs, and the relative theoretical expectations from the WMAP7 best-fit cosmology (without PNG) and constant bias of 2.1 and 1.7 are drawn in dark and light green (ACFs negative at  $\vartheta > 5$  deg). The red lines show the measured correlation functions for the CMASS data when the DR7 footprint mask is applied, while for the magenta line we have also removed the systematic correction for the effect of faint stars [54]. A part of the differences between the LRG samples can be accounted for by these two effects.

error on the ISW measurement is large. The difference between the NGC and the full DR8 data is more significant, but well within the expected statistical fluctuations given the size of the highly-correlated error bars in Fig. 3.

*Stellar correction.*—We have also checked whether the new correction for the systematic effect of faint stars introduced by Ref. [54] can better explain the differences between MegaZ and CMASS. We can see from the

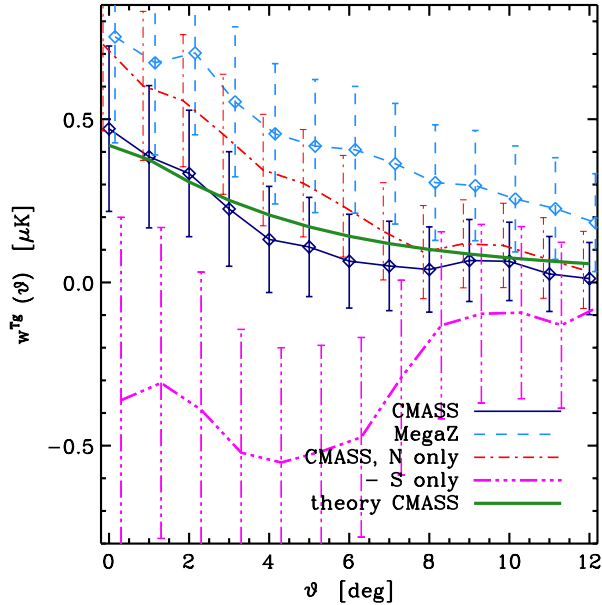


FIG. 3 (color online). North-south divide in the CMASS cross-correlation data. The results for the south galactic cap are contrary to the expected ISW signal, but the cosmic variance for this small area is large and it is potentially more subject to extinction systematics.

magenta line in Fig. 2 that indeed, if this correction is not applied to the CMASS data, the resulting CCF becomes even higher, close to the MegaZ result. However, even this effect is not enough to fully reproduce the earlier result on all angular scales.

*Frequency dependence.*—An important property of the ISW signal is that it is expected to be independent of CMB frequency, making any frequency dependence an indication of possible systematics. To quantify the level of contamination of the two LRG catalogs, we investigated in detail the frequency dependences of their CCFs, as shown in Fig. 4. Here we can see that the MegaZ result is more frequency dependent, while the CCF from the CMASS dataset is remarkably constant for all WMAP frequency bands, including the K and Ka WMAP bands which are most affected by residual galactic emission. This robustness, unmatched in any of the other galaxy catalogs from G12, is fully consistent with the ISW interpretation of the observed cross-correlation signal and appears to confirm that the stellar contamination in the CMASS data is negligible. From this test, we conclude that the CMASS catalog is the most robust data set available for our purposes.

*Extinction.*—As discussed in G08 and G12, dust extinction can be a major source of systematic uncertainty in ISW measurements. This is due to reddening from our Galaxy altering the inferred large-scale distribution of galaxies, which then introduces spurious correlations with the CMB, which may also have residual galactic emission. For this reason, we have carefully checked the effect of extinction on the measured CCF.

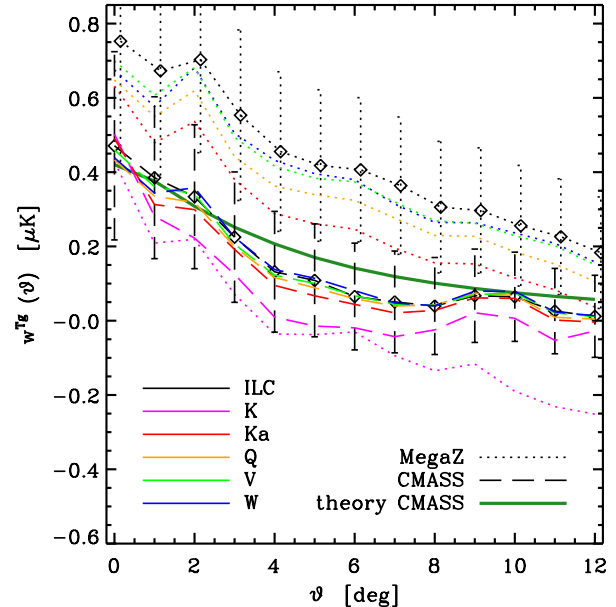


FIG. 4 (color online). Frequency dependence of the CCFs, for the MegaZ and CMASS data. We can see that the latter data are more stable across the whole frequency range.

Extinction can be corrected by excluding the most affected regions from the analysis. In Ref. [54], areas with a reddening  $A_r > 0.20$  magnitudes were excluded, while the cut in G12 was marginally stricter ( $A_r > 0.18$ ) and we have applied the latter in the present analysis. Fig. 5 shows how the CCFs vary when lowering the amount of reddening which is tolerated. When this cut is made stricter, the measured signal increases both for the full DR8 and NGC-only areas. The average extinction is higher in the SGC; the average in the whole footprint, calculated before applying the extinction cut itself, is  $\bar{A}_r = 0.12$ ; this reduces to  $\bar{A}_r = 0.10$  in the NGC and increases to  $\bar{A}_r = 0.18$  in the SGC.

The sky coverage decreases significantly for higher extinction cuts, going from  $f_{\text{sky}} = 0.22$  for  $A_r > 0.18$  to  $f_{\text{sky}} = 0.11$  for  $A_r > 0.08$ , of which only  $f_{\text{sky}} = 0.01$  remains in the south. This means that the error bars, which scale as  $\sqrt{1/f_{\text{sky}}}$  due to sample variance, are significantly higher. As we can see in Fig. 5, this increase in the errors makes the results for the different extinction cuts statistically consistent.

### 3. Conclusion

From the analyses performed in this section, we conclude that the CMASS data set is the most robust available for large-scale structure and cross-correlation measurements, as it appears to have low stellar contamination [54]. We also observe that the lower CCF from the CMASS data is partly due to a lower CCF in the SGC portion of the DR8 data; while consistent with cosmic variance, this could also be potentially connected to higher extinction effects in the SGC region.

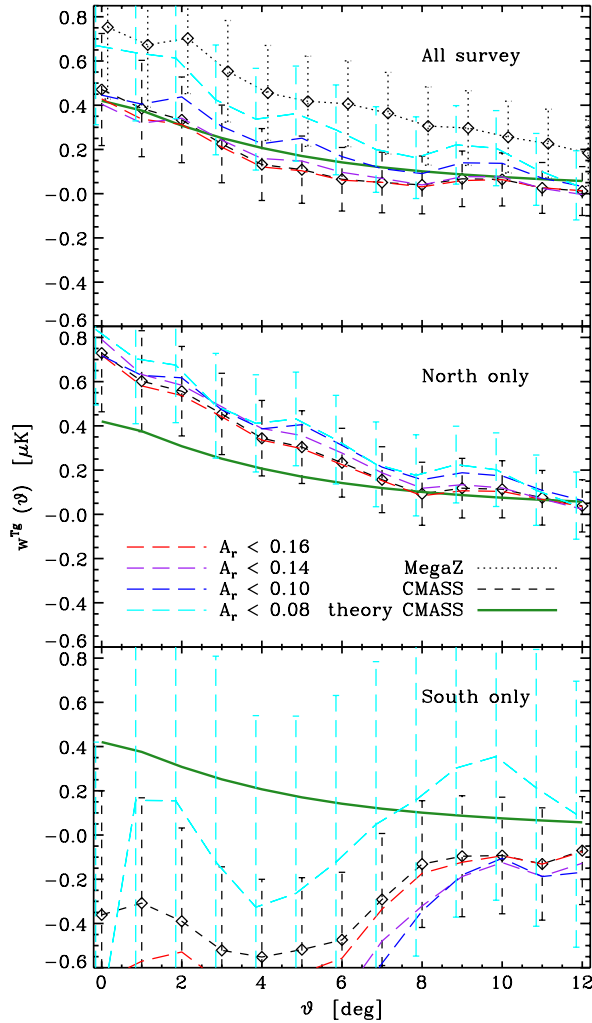


FIG. 5 (color online). Extinction dependence of the CCFs for the CMASS LRGs. The top panel shows the result for the whole surveyed area, while the middle and bottom panels present results from the NGC and SGC, respectively. The black lines refers to our baseline cut at  $A_r < 0.18$ . The greater variation in the SGC reflects both its smaller area and the higher average reddening in this region, and it is consistent within the errors.

### C. NVSS systematics

The NRAO VLA Sky Survey (NVSS) [90] produced a catalog of radio-galaxies covering the full sky at declinations  $\delta > -40$  deg. The clustering properties of these sources have been measured by several authors [91–94], who found excess power on the largest scales compared with the  $\Lambda$ CDM predictions and the analyses of other data sets. Its combination of large sky coverage and redshift depth makes the NVSS well-suited for the measurement of ISW cross correlations and this has been one of the most studied data sets in this context [95–102]. Furthermore, these data have been used for other studies, such as to detect CMB lensing [103] and to constrain inhomogeneous models [104].

Unfortunately this catalog is affected by some well-known issues. There is a large uncertainty in the redshift distribution of the sources, which can make its cosmological interpretation ambiguous. For many applications, this is not a serious issue; for example, it was shown in G12 that different assumptions for the distributions do not significantly affect the ISW measurements, given the size of the expected errors.

More worryingly, the sample contains clear variations in the density of sources with declination that are significantly greater than those expected for a statistically isotropic universe [91,95]. In particular, fewer sources are detected at  $\delta < -15$  deg, as can be seen in the upper panel of Fig. 6. In addition, we have found a similar albeit smaller issue in right ascension: the number density has a roughly linear decrease, reaching a minimum at right ascension (r.a.)  $> 250$  deg; this can be seen in the lower panel of Fig. 6. The data points at r.a. = 0/360 match as the distinction is purely artificial.

The main effect of such systematics, and of any long-wavelength noise in the data, will be to add spurious signal on the large-scale power spectrum (or ACF). Indeed the presence of spurious large-scale clustering power above an assumed cosmological model is often taken to be a clear indicator of uncorrected systematics in the data; however, we must be careful not to do so here as we are trying to measure, or limit, a large-scale PNG signal from the data.

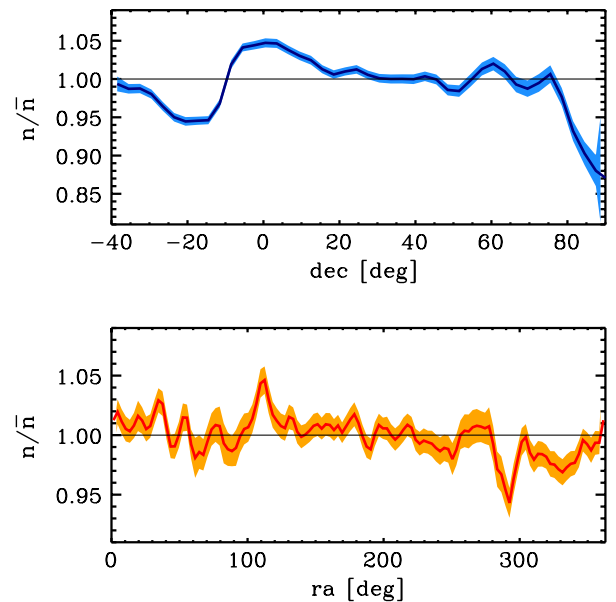


FIG. 6 (color online). Systematic density fluctuations in NVSS. The upper panel shows the well-known variation as a function of declination, while the lower panel presents a weaker deviation we have found as a function of right ascension. The data have been smoothed with a Gaussian beam of r.m.s.  $\sigma = 5$  deg. The colored bands show the Poisson  $1 - \sigma$  regions, which are largest near the poles.



In previous analyses, a number of methods have been used to correct the declination systematics: in some real-space analyses [95,96,104], including G08 and G12, the data were subdivided in declination bands, and expected densities were rescaled to match each. Other authors [91,93,36] found that this problem was minimized by imposing a strict flux cut of  $F > 10$  mJy; this however discards a large fraction of sources. Finally, others [103,105] working in harmonic space addressed the issue by assigning infinite variance to the  $m = 0$  modes of the data, effectively discarding any azimuthally invariant mode.

The method of fitting sub-bands in declination effectively assumes that any declination offsets are smooth and fit by the particular distribution of bands chosen. Going beyond the earlier analyses in G08 and G12, we have found that the signal fluctuates depending on the size chosen for these bands, leading to an uncertain autocorrelation. This suggests that we need a more robust technique for correcting for these effects if we are to use this sample to robustly measure PNG. However, the cross correlation with WMAP is fairly stable to such corrections: we show below in Section VA that the G12 results on the significance of ISW detection do not change significantly.

We have compared these existing methods for mitigating the NVSS systematics, and have also tested a new method, modelling the mean source density with an arbitrary mask that is assumed to be a separable function of the r.a. and dec coordinates. Any systematics or true fluctuations in the associated modes are absorbed into the inferred mask, so the inferred power is suppressed. We tested these methods with mock data to which we applied the observed r.a. and dec systematics, and found that this method does indeed allow us to remove their effects, and in particular the resulting large-angle power that is produced in the ACF. However, such corrections are problematic when constraining PNG, as they suppress the true power as well; when we applied these methods to non-Gaussian mocks, we found that the large-angle correlations produced by  $f_{\text{NL}}$  are also removed to a significant degree by this systematic mitigation procedure in ways that are difficult to model in real space. Since removing this type of systematic in real space without also removing any primordial signal is difficult, we decided to adopt the most conservative approach and not use the NVSS ACF to constrain cosmology nor PNG. The cross correlations with other density tracers and the CMB are instead expected to be reliable, as long as these other data sets are not affected by correlated systematics, and we therefore retain these measurements in our analysis.

If we include r.a. and dec systematic fluctuations in our mocks at a level consistent with the NVSS observations, then the clustering seen in the mocks is a good match to that in the observational data, as shown below in Fig. 10. We use these contaminated mocks with the density fluctuations

to estimate the covariance matrix, thus estimating the data errors using a set of mock ACFs with amplitude in agreement with the actual raw NVSS ACF.

#### D. Quasar systematics

Quasar samples are promising for future PNG measurements, as they are easily observed to high redshifts and are highly biased tracers of the matter distribution. The DR6 photometric quasar catalog has been used previously to constrain PNG, and Ref. [37] found that these data alone yield  $f_{\text{NL}} = 62 \pm 26$  ( $1\sigma$ ). However, this result is in conflict with Ref. [33], who found  $f_{\text{NL}} = 8_{-37}^{+26}$  (for their fiducial ‘QSO’ case): they used an earlier subset of the DR6 data from the SDSS DR3 [106]. Reference [60] (and more recently Ref. [105]) thoroughly tested the DR6 quasar sample against potential contaminants, and found significant effects associated with stellar density, the stellar color locus, airmass, and seeing. They concluded that the sample was not fit to use for  $f_{\text{NL}}$  measurements. Here we further investigate this issue, and argue that the quasar data should only be included through cross correlations with other surveys. Future spectroscopic quasar samples will be immune to many of these issues.

The fiducial photometric quasar sample we use is drawn from the SDSS data release six (DR6) [107]. The quasar catalog [108] contains more than one million objects selected via a nonparametric Bayes classifier–kernel density estimate (NBC-KDE) method, including both UV-excess (UVX) and high-redshift samples, up to a magnitude limit of  $i = 21.3$ . The total footprint is  $8417 \text{ deg}^2$ . Nevertheless, to reduce contamination as much as possible, we restrict the analysis to the UVX sample, as in G08 and G12, and we reduce the footprint with a stricter extinction cut, discarding areas with  $A_r > 0.14$ , so that we actually use 502,565 sources covering  $6,912 \text{ deg}^2$ .

Photometric quasars are particularly prone to stellar contamination, as their faint and compact nature matches those of stars. If we assume a stellar contamination fraction given by  $\kappa$ , the expected autocorrelation function will be given by

$$w^{\text{obs}}(\vartheta) = (1 - \kappa)^2 w^{\text{qso}}(\vartheta) + \kappa^2 w^{\text{star}}(\vartheta), \quad (18)$$

where we estimate the stellar ACF  $w^{\text{star}}(\vartheta)$  by measuring the clustering of stars from the SDSS survey. The distribution of stars over the DR6 footprint is distinctly anisotropic since they trace the structure of the Galaxy, and the autocorrelation of stars is therefore significant at large scales. A residual contamination of the order of  $\sim 2.5\%$  could explain the observed plateau in the ACF as shown in Fig. 7, as could  $f_{\text{NL}} \sim 50$ . We found from both the local relationship between quasar density and stellar density and also by directly measuring the cross-correlation function between the quasars and a catalog of SDSS stars that the actual contamination fraction is only  $\sim 1\%$  in these data.

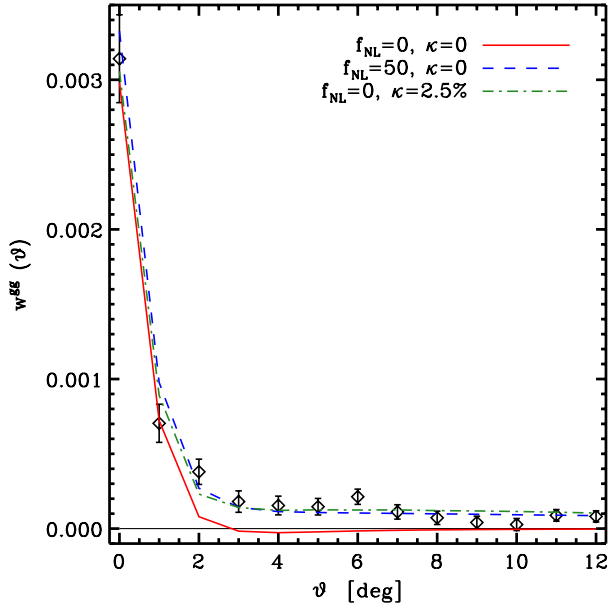


FIG. 7 (color online). Large-scale clustering of the quasars and stellar contamination. We plot in red (solid) the prediction for a model with no stellar contamination and no PNG. The green (dot-dashed) and blue (dashed) lines correspond to models with stellar contamination and with PNG, which are both able to explain the observed excess power (black data points). Constant linear bias is assumed.

The quasar ACF data thus suggest there are either further spurious fluctuations in the quasar density field, or  $f_{\text{NL}}$  is nonzero (as suggested by Ref. [37]).

It is reasonable to assume that subsamples at fainter magnitude should be less reliable and more severely affected by systematics. We therefore measured the ACF with cuts on  $i$ -band magnitude and we display the results in the top panel of Fig. 8. Indeed, the ACF of the faintest data ( $i > 20.5$ ; black triangles) displays much larger clustering amplitudes than the brighter subsamples (and the difference is inconsistent with the change in linear bias and a nonzero  $f_{\text{NL}}$ ). However, we are unable to isolate a particular magnitude at which we find that the systematic relationships become negligible and we do not find that the  $f_{\text{NL}}$  constraints from the DR6 quasar ACF become stable for samples brighter than any given  $i$ -band magnitude.

The bottom panel of Fig. 8 displays the ACF when we split the quasar sample by photometric redshift. Both the  $z > 2.2$  (black triangles) and  $z < 1.0$  (green stars) samples display significant large-scale clustering, that we can attribute to, in large part, significant stellar contamination in these subsamples (confirmed both via cross-correlation and local relationships with stellar density). However, we find that cutting by photometric redshift can introduce systematic fluctuations, as would be the case if there were any relationship between the estimated redshift and a potential systematic. This is illustrated by the fact that

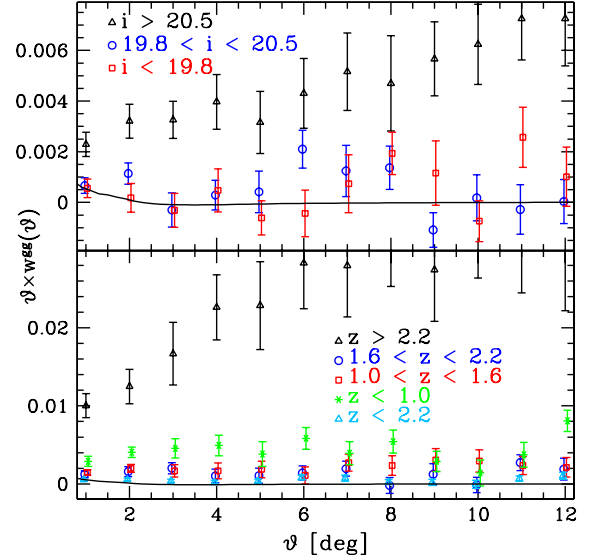


FIG. 8 (color online). Top panel: The angular autocorrelation function, ACF, multiplied by the angular scale  $\vartheta$ , when the DR6 quasar sample is subdivided based on  $i$ -band magnitudes. The solid black line is the prediction for the standard  $\Lambda$ CDM model. Bottom panel: The measured ACF when the DR6 quasar sample is split based on photometric redshift,  $z$ .

the ACF of both the  $1.6 < z < 2.2$  (blue circles) and  $1.0 < z < 1.6$  (red squares) samples has larger amplitudes than the  $z < 2.2$  (cyan triangles) ACF, despite the fact that the  $z < 2.2$  sample contains quasars with  $z < 1.0$  and thus has significant stellar contamination. Thus, while the  $1.0 < z < 2.2$  sample removes data with stellar contamination, the cuts on redshift appear to induce other systematic fluctuations. Some examples are shown in Fig. 9: e.g. we find a significant relationship with Galactic extinction  $A_r$ , when we cut the sample to  $1.0 < z < 2.2$ , but find no such relationship when no redshift cut is applied. We further find significant correlations with air mass, sky brightness, and especially at high redshift, with the surface density of stars.

We have not studied the DR6 quasar sample as thoroughly as Ref. [60], but we have, independently, reached a similar conclusion: the DR6 quasar ACF should not be used to obtain  $f_{\text{NL}}$  constraints. We have found that significant correlations exist with potential systematics, such as stellar density, Galactic extinction, and airmass, exist in excess of those expected. Also, these relationships depend nontrivially on the redshift and magnitude of the quasar sample that is selected. The effect of  $f_{\text{NL}}$  on the quasar density field will also depend on redshift and the bias of the sample, and we are therefore unable to address systematic concerns using the methods outlined in Refs. [54,57].

However, we do not expect these issues to be correlated with other samples, and should be able to trust correlations between the quasars and other data sets. In particular, the quasars have a large overlap in redshift with the NVSS data. Potential SDSS systematics, such as airmass and seeing, are

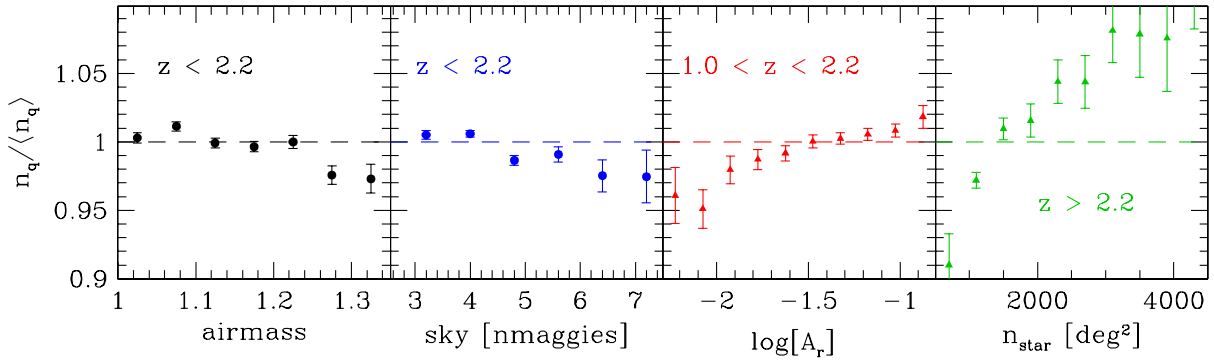


FIG. 9 (color online). Observed relationship between the number density of DR6 quasars  $n_q$  with respect to the average  $\langle n_q \rangle$  and a choice of systematics: airmass, sky brightness, galactic extinction in the  $r$  band, and stellar density, in different redshift ranges.

survey-specific and should thus have no correlation with NVSS data. In addition, we find no correlation with NVSS data and potential systematics (Galactic extinction, stellar density, synchrotron emission) that trace the structure of the Galaxy. Further, we trust correlations between the quasars and the LRGs, as the LRG sample has already proven to be robust to systematic fluctuations. Thus, while we do not consider the quasar ACF as a reliable probe of PNG, we will exploit the external correlations between the quasars and the other data sets. Also in this case, this includes the cross correlation with the CMB, which for the same reasons should be relatively free from contamination, as also confirmed by its frequency independence shown in G12.

## IV. MODELING THE DATA

### A. Data considered

We have discussed six different large-scale structure data sets, which yield six autocorrelations, fifteen cross correlations and six correlations with the WMAP CMB temperature. Our final data set is shown in Fig. 10, including the galaxy-CMB cross correlations and the galaxy-galaxy correlations between all the catalogs. As any of the corrections to the NVSS r.a. and dec fluctuations that we have tested can potentially remove the PNG signal, we are conservatively not using the NVSS ACF. We use the raw NVSS data without any correction, to measure external cross correlations as described below, since the other data sets should not correlate with these systematic fluctuations. While random correlations between the NVSS systematics and the other data may in principle occur by chance, we have checked that our results remain fully consistent and are not biased if instead we apply a systematic correction procedure to NVSS.

In all cases below, we discard the first five angular bins of the 2MASS-CMB CCF, which is believed to be contaminated by the SZ effect. We also discard the zero-lag data points in the density-density correlations based on contamination from nonlinear structure formation, which we are not modeling, and from possibly inaccurate

shot-noise subtraction in the correlations between different catalogs, due to the unknown fraction of common sources between the catalogs.

Given the potential impact of systematic contaminants, it is useful to define the following three subsets of data (we exclude in all cases the 0-lag bins of the density-density correlations and the first five data points in the 2MASS ISW measurement as described above):

- (i) The “naïve” data set is the full combination. As we are conservatively using the raw NVSS data, without any correction for the r.a. and dec systematics, we can not use the NVSS ACF for any cosmological constraints. So this data set now includes 26 correlation functions.
- (ii) The “fair” sample contains 25 correlations, excluding the autocorrelations of NVSS and QSO, which are expected to be overall the least reliable parts of our data set, due to the high risk or residual systematics (as discussed above).
- (iii) The “conservative” set includes only what we view as the most reliable correlations. We exclude all the autocorrelations except for the LRG case, as autocorrelations are expected to be more prone to systematics than the cross correlations. Further, we exclude any data from 2MASS and from the main SDSS galaxies; as they are not expected to contribute significantly to the PNG measurements due to their low bias, we feel it is better to exclude them rather than risk any systematic contamination that they might bring. Finally, we also exclude the LRG-quasar correlation: as both samples are derived from SDSS, there is a limited risk of residual correlated systematics, particularly related to stellar contamination. The resulting set is consists of ten correlations only.

We will discuss the results from these samples in the following section. Finally, it is worth noting that we discard samples by effectively setting the covariance for these data to a high value. The likelihood estimation from the inverse of the covariance matrix will then simply ignore the discarded points.

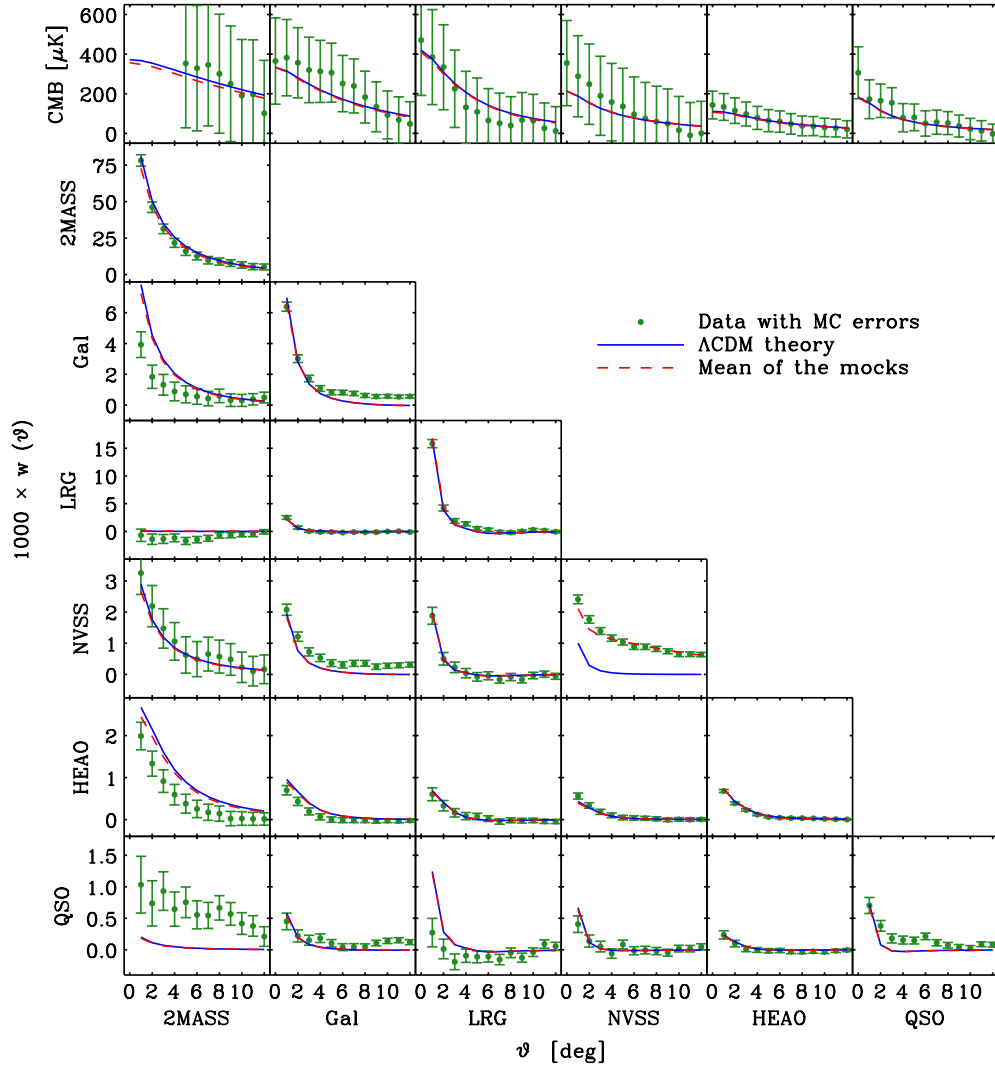


FIG. 10 (color online). Complete set of the two-point functions we use. The top row shows the CMB-galaxy correlation functions, while the remaining panels are the galaxy-galaxy correlations. Error bars are from 10,000 Monte Carlos, whose means are the red dashed lines, and the blue line is the standard  $\Lambda$ CDM cosmology from WMAP7, with constant biases (not a fit to these data). The ACF of the raw NVSS data presents a significant excess power with respect to the  $\Lambda$ CDM expectations, which is modelled by adding to the mocks the r.a. and dec density fluctuations observed in the data.

In addition to our data, we will consider also the CMB data from WMAP7 [3] and in some cases the luminosity distances to the Union2 compilation of Type Ia Supernovae [109].

We calculate the theoretical correlation functions using a modified version of CAMB [110] (validated with a fully independent code for accuracy). The covariance matrix is generated using the Monte Carlo method described in G12, now extended to the whole data set (we have  $n = 351$  data points, so this is the dimension of the full covariance). We base our Monte Carlo realizations on a fiducial flat  $\Lambda$ CDM model without PNG, which is appropriate since we do not find posterior evidence for PNG; if such evidence were found, we would need to re-calculate our likelihoods with a new covariance matrix, based on a non-Gaussian model. To

calculate the covariance, we use the linear bias parameters fitted from the autocorrelation functions and assumed constant in redshift. In the case of NVSS, we fitted the bias to the ACF when the systematics are subtracted, which is the only step for which we use the corrected NVSS data; the obtained bias value agrees with the external correlations of NVSS, and furthermore the fact that the mock samples reproduce the measured clustering also for the ACF implies the bias we used is reasonable.

## B. Gaussian bias and other nuisance parameters

In addition to the standard cosmological parameters, we must include other nuisance parameters to account for the systematics and other uncertainties, over which we marginalize. These describe our assumptions about the intrinsic

bias (in the absence of PNG effects), uncertainties in the redshift distributions of the surveys (and particularly their overlaps) and other effects.

For most catalogs  $i$ , we assume an intrinsic (Gaussian) bias evolution of the form:

$$b_1^i(z) = 1 + \frac{b_0^i - 1}{D^{\gamma_i}(z)}; \quad (19)$$

we assume  $\gamma_i = 2$  in most cases [111]; this bias approaches unity in the future, and is therefore appropriate for flux-limited samples. However, for the quasars we instead assume

$$b_1^{\text{QSO}}(z) = \frac{b_0^{\text{QSO}}}{D^{\gamma_{\text{QSO}}}(z)}, \quad (20)$$

which instead approaches zero in the future. This is consistent with observations of quasars [112,113], that suggest that the preferred halo mass of quasars is  $\sim 10^{12} M_\odot$ , independent of redshift. Setting  $\gamma_{\text{QSO}} = 1.6$  approximately satisfies this constant-mass condition for  $0.43 < z < 2.4$ , which spans the median redshifts of Refs. [112,113], and is the span of redshifts at which we may reasonably expect to obtain meaningful information from the DR6 quasar sample. We are relatively uncertain of the bias evolution of the NVSS sample, which has a broad redshift distribution but also the potential to provide strong PNG constraints due to its large sky coverage. We

therefore have tested allowing a free slope  $\gamma_{\text{NVSS}}$  for this sample and found that this results in negligible changes in our constraints. This suggests that our results are not strongly affected by uncertainties in the bias evolution relationship.

In addition we allow the possibility to add 15 extra parameters  $\beta_{ij}$  to account for uncertainties in the redshift distributions  $\varphi_i(z)$ . A factor arises for each pair of catalogs, and they are introduced in the density-density cross correlations so that:

$$w_{\text{obs}}^{g_i g_j}(\vartheta) = \beta_{ij} w^{g_i g_j}(\vartheta). \quad (21)$$

Effectively, the uncertainty in  $\varphi_i(z)$  becomes more important when the overlap between two surveys is smaller. For this reason, we introduce a Gaussian prior, with mean  $\mu_\beta = \langle \beta_{ij} \rangle = 1$  and a variance  $\sigma_\beta^2$  which depends on the fiducial survey overlap, such that the r.m.s. is  $\sigma_\beta = 0$  when the overlap is total (the ACFs), and  $\sigma_\beta = 2$  when the overlap is expected to be small. We linearly interpolate the prior between these two values. We also checked that the results on PNG do not change significantly when doubling the r.m.s. of these priors.

We also add a free stellar contamination fraction  $\kappa_i$  for the three samples derived from SDSS, used as described for the quasars in Eq. (18). We finally add one extra parameter for the HEAO catalog: an amplitude describing the clustering-independent autocorrelation due to source shot noise spread over the large PSF of the instrument  $\alpha_{\text{HEAO}}$  [114].

TABLE I. Summary of the parameters used in the likelihood runs, and their prior ranges. We checked that our posterior results are never affected by the edges of the priors. We specify in the last column the parameters that we are using for our final  $f_{\text{NL}}$  runs (“conservative” and “fair”).

	Parameter	Description	Prior range	In final runs
Cosmology	$\omega_b \equiv \Omega_b h^2$	baryon energy density	[0.015, 0.035]	Yes
	$\omega_c \equiv \Omega_c h^2$	dark matter energy density	[0.05, 0.2]	Yes
	$\theta$	sound horizon at the last-scattering surface	[0.8, 1.2]	Yes
	$\tau$	optical depth	[0.01, 0.25]	Yes
Our focus	$\log(10^{10} A_s)$	amplitude of primordial perturbations at $k_{\text{piv}} = 0.002 h/\text{Mpc}$	[2.9, 3.5]	Yes
	$f_{\text{NL}}$	amplitude of skewness-type primordial non-Gaussianity	[-200, 200]	Yes
	$g_{\text{NL}}$	amplitude of kurtosis-type primordial non-Gaussianity	$[-5 \times 10^6, 5 \times 10^6]$	No
	$a_{\text{NL}}$	scale-dependence of bias	[0, 4]	No
	$w$	dark energy equation of state	$[-2.5, -1/3]$	No
Nuisance	$A_{\text{SZ}}$	amplitude of the SZ CMB template	[0, 2]	Yes
	$b_0^i$	6 bias nuisance parameters for each catalog	[0, 3]	Yes
	$\gamma_i$	slope of bias evolution for each catalog	[0, 3]	No
	$\kappa_i$	3 stellar contamination nuisance parameters for each SDSS catalog	[0, 0.1]	Yes
	$\beta_{ij}$	15 $dN/dz$ nuisance parameters for each pair of catalogs	Gaussian priors	No
	$\alpha_{\text{HEAO}}$	1 PSF nuisance parameters for the HEAO catalog	[0.0001, 0.001]	Yes
Derived	$\Omega_\Lambda$	dark energy energy density	—	—
	$T_0$	age of the Universe	—	—
	$\Omega_m$	total matter energy density	—	—
	$\sigma_8$	LSS amplitude of density fluctuations	—	—
	$z_{\text{re}}$	redshift of reionization	—	—
	$H_0$	Hubble constant	—	—

## V. RESULTS

We examine a range of models; starting from a simple flat  $\Lambda$ CDM model, we add primordial non-Gaussianity of different types, or the dark energy equation of state  $w$ . The parameters which we use and their prior ranges are shown in Table I.

We study the cosmological consequences of these data via a Monte Carlo Markov chain (MCMC) method using the publicly available COSMOMC code [115] and its extension to nested sampling MULTINEST [116], which replaces the Metropolis-Hastings with a nested sampling algorithm, achieving in the process higher accuracy, greater speed, and an estimation of the Bayesian evidence for each model. We ensured to always use high enough accuracy within MULTINEST and CAMB to obtain numerical stability, also enforcing a sufficiently broad  $k$  range in the integrations.

### A. $\Lambda$ CDM

#### 1. ISW significance

We first consider a  $\Lambda$ CDM template with a single amplitude of the ISW  $A$ ; as described in G08 and G12 this can be done by assuming the CCFs in each bin  $i$  are written as  $w_i^{Tg} = Ag_i$ , where  $g_i$  is the template prediction from the fiducial model. Then the best-fit  $A$  can be calculated from the covariance matrix  $\mathcal{C}$  and from the observed CCFs  $\hat{w}_j^{Tg}$  by analytically maximizing the likelihood:

$$A = \frac{\sum_{i,j=1}^p C_{ij}^{-1} g_i \hat{w}_j^{Tg}}{\sum_{i,j=1}^p C_{ij}^{-1} g_i g_j}, \quad (22)$$

where in our case  $p = 13$  if considering a single catalog, and  $p = 13 \times 6 = 78$  for the full combined analysis. The variance is given by

$$\sigma_A^2 = \left( \sum_{i,j=1}^p C_{ij}^{-1} g_i g_j \right)^{-1}. \quad (23)$$

In this way we recover the significances reported in Table II, which are consistent with our earlier work, and with the expected signal-to-noise ratio within the statistical errors, calculated according to Eq. (6) in G12. The significance of the ISW detection in the CMASS LRGs is at  $1.8\sigma$ , down from  $2.5\sigma$  which was found with the MegaZ data, as the measured CCF has decreased, in agreement with the results by Ref. [55]. The significance of the NVSS ISW with raw data is increased to  $3.7\sigma$ , as the signal has increased. Given the highly correlated error bars and the different shape of the template to the data points, the best-fit curve happens to be higher than the data points. However the strong systematics that are present in NVSS

will also increase the errors, and thus the actual significance is almost certainly lower than the one derived from the statistical error. The NVSS ISW significance is marginally lower at the  $3.6\sigma$  level if we correct the r.a. and dec systematics, and the results using different treatments are anyway fully compatible within the errors. The overall significance has remained nearly unchanged from G12 at  $4.6\sigma$ , to which we add a  $0.4\sigma$  estimate of the systematic error as in G12.

Catalog	$A \pm \sigma_A$	S/N	expected
2MASS cut	$1.32 \pm 2.02$	0.7	0.5
SDSS gal DR8	$1.29 \pm 0.59$	2.2	1.6
SDSS LRG MegaZ	$2.10 \pm 0.84$	2.5	1.2
SDSS LRG CMASS	$1.10 \pm 0.61$	1.8	1.6
NVSS G12	$1.21 \pm 0.43$	2.8	2.6
NVSS corrected	$1.47 \pm 0.41$	3.6	2.6
NVSS raw	$1.88 \pm 0.51$	3.7	2.6
HEAO	$1.37 \pm 0.57$	2.4	2.0
SDSS QSO DR6	$1.46 \pm 0.60$	2.4	1.7
<b>TOTAL G12</b>	$1.38 \pm 0.32$	$4.4 \pm 0.4$	$\sim 3.1$
<b>TOTAL, corr. NVSS</b>	$1.39 \pm 0.31$	$4.5 \pm 0.4$	$\sim 3.1$
<b>TOTAL, raw NVSS</b>	$1.48 \pm 0.33$	$4.6 \pm 0.4$	$\sim 3.1$

will also increase the errors, and thus the actual significance is almost certainly lower than the one derived from the statistical error. The NVSS ISW significance is marginally lower at the  $3.6\sigma$  level if we correct the r.a. and dec systematics, and the results using different treatments are anyway fully compatible within the errors. The overall significance has remained nearly unchanged from G12 at  $4.6\sigma$ , to which we add a  $0.4\sigma$  estimate of the systematic error as in G12.

#### 2. Full MCMC

We then fully test a simple flat  $\Lambda$ CDM model using MULTINEST; we can see in the first row of Fig. 11 the resulting 1D marginalized posteriors on a choice of cosmological parameters. Adding our “fair” data to the WMAP7 CMB power spectrum yields results which agree within the errors, even though our data prefer a lower matter density  $\Omega_m$ ; we find  $0.21 < \Omega_m < 0.28$  (95% C.L.) when using our data and the CMB. We show in Fig. 12 the marginalized constraints on the most important nuisance parameters. We can see that about 5% stellar contamination seems present in the main galaxy sample, which is reduced to 1% in the quasars, while no evidence for residual contamination appears for the LRGs, in agreement with our positive assessment of the purity of this sample.

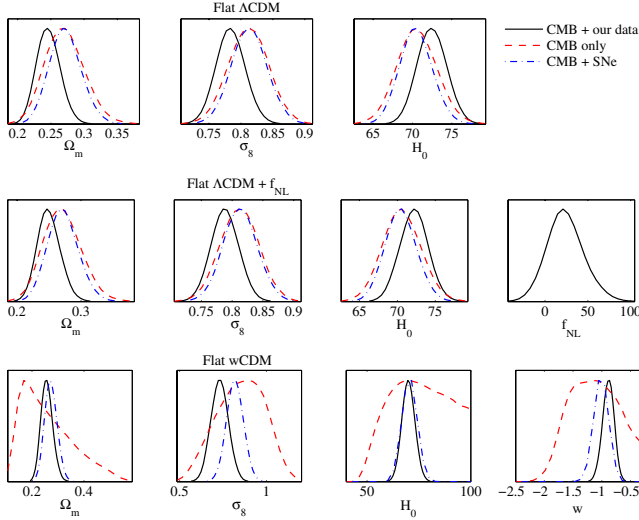


FIG. 11 (color online). Marginalized posterior probability distributions on a choice of cosmological parameters for different models. First row: The flat  $\Lambda$ CDM case. The results are in agreement with the concordance model, but we can see that our data (the “fair” data set is used) shift the peak of the distribution towards models with more dark energy. Second row: Same, with the addition of local  $f_{\text{NL}}$ . (The CMB temperature power spectrum can not constrain PNG and is ignored.) The standard cosmological parameters do not change, and  $f_{\text{NL}}$  is consistent with zero. Third row: Same, for the case of a wCDM model. Here the CMB can not constrain these parameters simultaneously, hence the broad posteriors. The degeneracy between  $\Omega_m - w$  is broken by adding either Type Ia SNe or our data. The HST  $H_0$  prior was added to the  $w$  run using our data.

## B. Primordial non-Gaussianity

### 1. The $f_{\text{NL}}$ model

We next consider the local  $f_{\text{NL}}$  model, which contains one extra parameter. When estimating the posterior likelihood of  $f_{\text{NL}}$  from the whole “naïve” data set, including all autocorrelations, we find a strong positive detection:

$31 < f_{\text{NL}} < 64\%$  at 95% C.L., as shown in two dimensions by the red contours of Fig. 13. This result is consistent with previous measurements that used similar data sets [37], but our uncertainty on  $f_{\text{NL}}$  is reduced due to the fact that we (i) effectively include all multipoles  $l$  (for SDSS data, we include the  $\kappa$  parameters to marginalize over the known low- $l$  issues), (ii) include fainter NVSS data (Ref. [37] applied a 10 mJy cut), (iii) include 33% more area for the LRGs, and (iv) include additional data sets.

To assess whether this result is robust, we study how it is changed when less reliable data are excluded. When we repeat this test with the “conservative” data, we find no detection:  $-36 < f_{\text{NL}} < 45$  at 95% C.L., corresponding to the green contours of Fig. 13. When we examine the “fair” data, we still see no hint for PNG (blue contours in Fig. 13). By looking at both panels of Fig. 13, we can see no evidence of strong degeneracies between  $f_{\text{NL}}$  and  $\Omega_m$  or  $\sigma_8$ . The results are in agreement with the simplest Gaussian  $\Lambda$ CDM model: the marginalized 95% C.L. interval from our “fair” sample is  $-15 < f_{\text{NL}} < 68$ .

We can see the corresponding marginalized 1D distributions in the second row of Fig. 11; by comparing with the first row, we can also notice that best-fit regions of a selection of other cosmological parameters do not change when adding  $f_{\text{NL}}$ : this indicates overall low degeneracy between these parameters and  $f_{\text{NL}}$ .

We further tested for possible degeneracies between  $f_{\text{NL}}$  and other parameters by examining the marginalized 2D likelihood contours, particularly between  $f_{\text{NL}}$ , the stellar contamination  $\kappa$  and the bias  $b_0$  for each sample. We found that the only parameter that presents some degeneracy with  $f_{\text{NL}}$  is the quasar bias  $b_0^{\text{QSO}}$ . When considering the ‘naïve’ data set we found that, perhaps surprisingly, there is no significant degeneracy with the quasars’  $\kappa$ , while the degeneracy with  $b_0$  is more pronounced: by raising it to  $b_0 \approx 1$ , values as low as  $f_{\text{NL}} \approx 30$  are allowed. When using the quasars’ ACF however, the Gaussian limit is always excluded at  $> 2\sigma$ . We found that a stronger degeneracy between  $\kappa$  and  $f_{\text{NL}}$  is present only when using the quasar

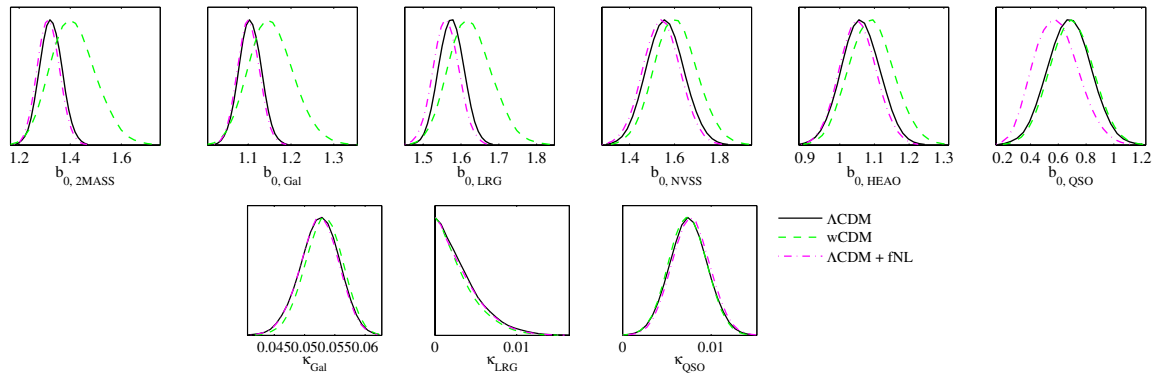


FIG. 12 (color online). Marginalized posterior probability distributions on a choice of nuisance parameters for different models. First row: The bias parameters  $b_0^i$  for each catalog. Second row: Same, for the stellar contamination parameters  $\kappa_i$ , which are introduced for the SDSS data only. The wCDM case includes the HST  $H_0$  prior.

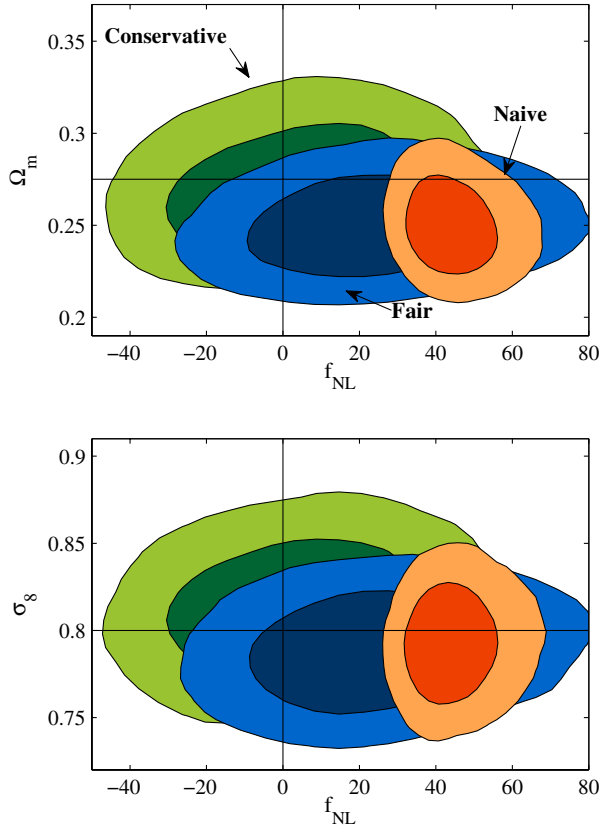


FIG. 13 (color online). Marginalized posterior probability distributions for  $f_{\text{NL}}$  versus  $\Omega_m, \sigma_8$  for different data sets. The naïve result obtained using the complete data set (red contours, at 68% and 95% C.L.), which would suggest the presence of significant PNG, is not stable. When using only the most reliable parts of our compilation, we obtain the conservative result (green), which remains stable when adding back most of the data (blue), except the quasar and NVSS ACFs. As these are the least reliable data, as we discussed in our systematic section, we decide to discard them.

ACF alone. When using the “conservative” data set, we found that the  $f_{\text{NL}} - b_0^{\text{QSO}}$  degeneracy is significantly moderated.

We summarize the constraints on  $f_{\text{NL}}$  in Table III and in Fig. 14 for clarity. Here we compare the marginalized results obtained when using the most constraining parts of our data set. We can see once again that most results agree with Gaussian initial conditions, and with each other. When considering single autocorrelation functions, we marginalize over cosmology including the WMAP CMB likelihood, and over one bias parameter and one stellar contamination fraction (for the data derived from SDSS). To better interpret the cross correlations on their own, we have assigned Gaussian priors on the relevant bias and stellar contamination parameters equal to the posteriors on these parameters obtained from the “fair” data. Applying these priors allows us to accurately portray the relative importance of each cross correlation to our bottom-line results.

TABLE III. Summary of the  $f_{\text{NL}}$  constraints. In the top section we show the results from the combined analyses. The two runs with the nuisance parameters  $\beta_{ij}$  have two different choices for their Gaussian priors. The runs with a single ACF below include marginalizations over one bias parameter and (for the SDSS catalogs) one stellar contamination parameter.

Data	$f_{\text{NL}}$ 95% interval
Conservative	$-36 < f_{\text{NL}} < 45$
Fair	$-15 < f_{\text{NL}} < 68$
Naïve	$31 < f_{\text{NL}} < 64$
Conservative; with $\beta_{ij}, \sigma_\beta^{\text{max}} = \ln 2$	$-36 < f_{\text{NL}} < 42$
Conservative; with $\beta_{ij}, \sigma_\beta^{\text{max}} = \ln 4$	$-36 < f_{\text{NL}} < 42$
LRG-LRG	$-116 < f_{\text{NL}} < 91$
NVSS-NVSS	$140 < f_{\text{NL}} < 245$
QSO-QSO	$-13 < f_{\text{NL}} < 91$

Furthermore, we found that applying the bias prior to the autocorrelations would increase the precision of their  $f_{\text{NL}}$  constraints by a factor of two. Accounting for this factor, the LRG autocorrelation is the best-constrained measurement that enters the “conservative” data set. When using the LRG ACF only we recover a result consistent with the recent analysis by Ref. [38], who found  $-45 < f_{\text{NL}} < 195$  at 95% using the spectroscopic sample of the CMASS LRGs, which contains  $\sim 1/3$  of the photz sample we use.

Notice that the factor  $(b_1 - 1)$  within the bias correction  $\Delta b$  is the leading contribution that determines the size of the  $f_{\text{NL}}$  error bars. For this reason, the low-bias data from 2MASS, the SDSS main galaxies, and HEAO bring little information on  $f_{\text{NL}}$ . Also the external correlations of the quasars bring less contribution than it may be expected, since the quasar bias at low redshift is also low. This explains why the strongest constraints come from NVSS, the LRGs and their external correlations. For this reason, we have also checked the effect of the assumed NVSS bias evolution with one additional run where the evolution parameter  $\gamma_{\text{NVSS}}$  is let free, and we found no significant changes in the results.

## 2. The $a_{\text{NL}}$ model

We then extend our model to generalized PNG defined in Eq. (8): in addition to  $f_{\text{NL}}$ , we thus allow for scale dependence of the bias of any slope  $a_{\text{NL}}$ , which reduces to  $a_{\text{NL}} = 2$  in the local, scale-independent case. We show our marginalized posterior likelihood distribution in the top panel of Fig. 15, where we can see that, in line with the lack of evidence for  $f_{\text{NL}}$ , there is no evidence for  $a_{\text{NL}}$  either. The full marginalized upper limit we find is  $a_{\text{NL}} < 1.8$  at 95%, but it must be born in mind that there is an infinite degeneracy along the direction  $f_{\text{NL}} = 0$  by construction: thus, this result is strongly dependent on our adopted priors, rather than being a “stand-alone measurement”. The correspondent bound on  $n_{f_{\text{NL}}}$  can be found using Eq. (10).



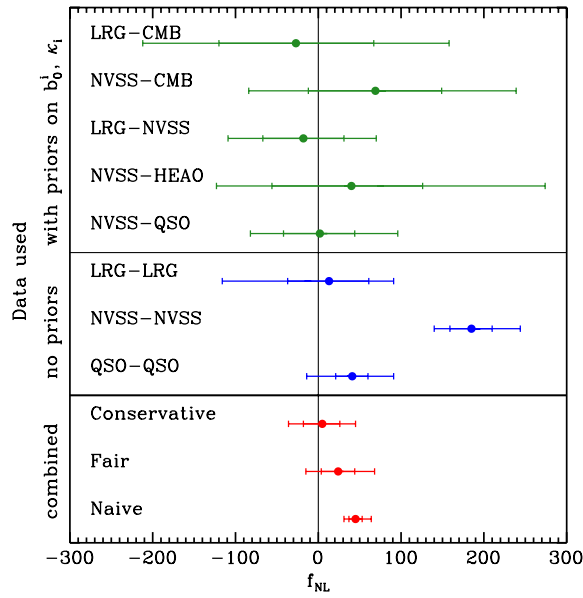


FIG. 14 (color online). Comparison of the marginalized posterior probability distribution on  $f_{\text{NL}}$  using the parts of our data set giving the strongest contributions. We show the results from single cross-correlation functions (top, green), autocorrelations (center, blue), and from combined subsamples of the whole data set (bottom, red). The lines correspond to 68% and 95% ranges, have been marginalized over the cosmological parameters, and include the WMAP7 CMB priors. The points represent the mean values of the posterior likelihoods. The results from single autocorrelation functions have also been marginalized over one bias parameter and one stellar contamination fraction (for the SDSS samples). The NVSS ACF result is inconsistent with the rest, but is discarded due to the high level of systematics. To best present the relative constraining power of the cross-correlation measurements, we have placed priors on the bias and stellar contamination parameters, which significantly overstate the constraints these cross correlation allow on their own. See the main text for more details.

### 3. The $g_{\text{NL}}$ model

We finally consider the  $g_{\text{NL}}$  model. We shall here make the optimistic assumption that the fitting formula of Eq. (7) is a reasonable approximation to the effect of  $g_{\text{NL}}$ , keeping in mind that this may not be accurate in all cases due to the low bias of our catalogs. Under this assumption we find  $-5.6 \times 10^5 < g_{\text{NL}} < 5.1 \times 10^5$  (95%) if assuming  $f_{\text{NL}} = 0$ . However as shown by Refs. [65,66], and as clear from Eq. (1), there is a degeneracy between  $f_{\text{NL}}$  and  $g_{\text{NL}}$ , as both parameters produce a scale dependence of the same order  $\sim k^{-2}$ ; the degeneracy is alleviated by the different redshift dependences. This is indeed what happens when we consider the complete model where both parameters are left free: we can see in the bottom panel of Fig. 15 that the marginalized posterior partially presents this degeneracy, as demonstrated with  $N$ -body simulations by Ref. [66]. Also in this case the Gaussian model remains well within the 95% region: the marginalized constraints on

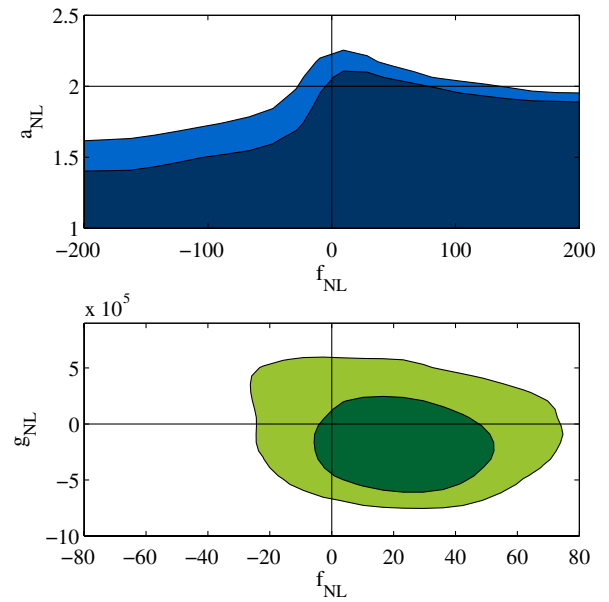


FIG. 15 (color online). Marginalized posterior probability distributions for extended PNG models. Top panel: The parameter  $a_{\text{NL}}$  defines the scale-dependence of the bias and it is  $a_{\text{NL}} = 2$  in the local, scale-independent model. The two shaded contours represent the 95% and 99% confidence regions. There is a vertical infinite degeneracy along the  $f_{\text{NL}} = 2$  direction, which is only partially visible due to finite sampling resolution and smoothing. Bottom panel: Marginalized constraints on the  $f_{\text{NL}} - g_{\text{NL}}$  plane (68% and 95% regions). As both parameters produce the same scale dependence of the bias, they are degenerate, but only partially, as the redshift dependence is different. Note that the  $g_{\text{NL}}$  constraints are optimistic, as they assume the validity of the fitting formula by Ref. [65] for our data.

the two parameters are marginally degraded to  $-17 < f_{\text{NL}} < 63$  and  $-5.9 \times 10^5 < g_{\text{NL}} < 4.7 \times 10^5$  at 95%, respectively, when they are both set free.

### C. Dark energy

Without the ISW effect from the presence of dark energy, we would not be able to use the CMB cross correlations to constrain PNG. In the above analysis, we jointly fitted the dark energy density and PNG to test the sensitivity of the PNG measurements to our assumptions about DE.

#### 1. The $w$ CDM Model

We next also consider models with Gaussian initial conditions, and use the ISW data to constrain the simplest dynamical dark energy of equation of state  $w$ . We present in the third row of Fig. 11 the marginalized 1D posterior probability distributions when considering this  $w$ CDM model. Here we can see that, as it is well known, the CMB temperature power spectrum alone can not break the  $\Omega_m - w$  degeneracy, which is broken instead by either Type Ia Supernovae or our data. This can be seen more clearly in the 2D plot of Fig. 16. Here we show that the results from

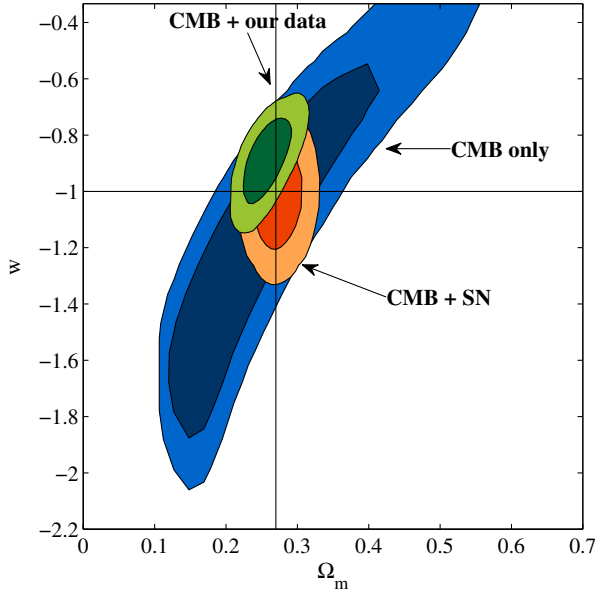


FIG. 16 (color online). Marginalized posterior probability distributions for  $\Omega_m$  vs  $w$  for different data sets: CMB only (blue), CMB + Type Ia Supernovae (red), and CMB + our data (green). The results are in agreement with the concordance  $\Lambda$ CDM model. The HST  $H_0$  prior was added to the run using our data.

our data and the CMB power spectrum reduce the parameter space significantly around the concordance values. Notice that we have added the HST prior [117] on the Hubble constant  $H_0 = 74.2 \pm 3.6$ , since we found that otherwise significant degeneracies appear between  $w$ , the bias parameters, and  $H_0$ , due to compensating effects in the growth function  $D(z)$ . When marginalizing over all other parameters, we find from our data  $-1.09 < w < -0.70$  at 95%. This confirms the overall agreement with the standard  $\Lambda$ CDM paradigm of the ISW measurements.

#### D. Model selection

Nested sampling allows model selection in addition to parameter estimation, thanks to the calculation of the

Bayesian evidence factor for each model. Briefly, if we assume a model  $M$  of parameters  $\Theta$ , and we compare it with data  $D$ , Bayes' theorem states that the posterior probability distribution  $\mathcal{P}$  on the parameters is given by

$$\mathcal{P}(\Theta) = \frac{\mathcal{L}(\Theta)\Pi(\Theta)}{\mathcal{Z}(M)}, \quad (24)$$

where the prior is  $\Pi(\Theta) = P(\Theta|M)$ , the likelihood is  $\mathcal{L}(\Theta) = P(D|\Theta, M)$ , and the Bayesian evidence is  $\mathcal{Z}(M) = P(D|M)$ . When interested in parameter estimation only, it is common to neglect the evidence  $\mathcal{Z}$  and simply study the posteriors of arbitrary normalization for a given model. However here we will also compare different models (e.g. with and without PNG), and for this we will use the evidence. Model selection between two models  $M, N$  can be performed by calculating the ratio of their evidences, also called Bayes factor:  $B \equiv \mathcal{Z}_N/\mathcal{Z}_M$ , so that  $\ln B = \Delta \ln \mathcal{Z}$ . This factor, which incorporates Occam's razor by penalizing models with unnecessary extra parameters, expresses the odds between the models given the data, and can be qualitatively interpreted with the heuristic Jeffrey's scale, which states that the model selection is inconclusive if  $|\ln B| < 1$ , and that the evidence for one model is weak, moderate, or strong for  $\ln B > 1$ ,  $\ln B > 2.5$ ,  $\ln B > 5$ , respectively [118].

To test non-Gaussianity, we can compare the Bayesian evidence  $\mathcal{Z}$  of the models with and without PNG. As shown in Table IV, the Bayes factor between these models is  $\Delta \ln \mathcal{Z} = -1.17$  in the case of  $f_{\text{NL}}$  and  $\Delta \ln \mathcal{Z} = -1.16$  in the case of  $g_{\text{NL}}$ , which are both interpreted as weak evidence against these models according to Jeffrey's scale, as the two models have odds  $\sim 1:3$  compared with  $\Lambda$ CDM given our data [118]. The situation is more defined for the extended model with both  $f_{\text{NL}}$  and  $g_{\text{NL}}$ : in this case we find  $\Delta \ln \mathcal{Z} = -2.34$ , corresponding to odds of 1:10.

We also performed a Bayesian model selection for dark energy: as shown in Table IV, when using the CMB and our data this test is also weakly disfavoring the  $w$ CDM model:  $\Delta \ln \mathcal{Z} = -1.46$ , which means the odds are 1:4 when

TABLE IV. Bayesian model selection. For each data set we compare models with extra parameters with the baseline  $\Lambda$ CDM. We can see that our "fair" data in combination with the CMB provide weak evidence against models with one extra PNG parameter; the odds become worse against a more complex model with both  $f_{\text{NL}}, g_{\text{NL}}$  present.

Parameters	Data	$\ln(\mathcal{Z})$	$\ln B = \Delta \ln(\mathcal{Z})$	Odds	Interpretation
$\Lambda$ CDM	CMB + our	$-3983.27 \pm 0.15$	0	—	—
$+f_{\text{NL}}$	("fair")	$-3984.44 \pm 0.15$	$-1.17 \pm 0.21$	1:3	Weak evidence against $f_{\text{NL}}$
$+g_{\text{NL}}$		$-3984.43 \pm 0.15$	$-1.16 \pm 0.21$	1:3	Weak evidence against $g_{\text{NL}}$
$+f_{\text{NL}} + g_{\text{NL}}$		$-3985.61 \pm 0.13$	$-2.34 \pm 0.20$	1:10	Weak evidence against $f_{\text{NL}} + g_{\text{NL}}$
$w$ CDM		$-3984.73 \pm 0.15$	$-1.46 \pm 0.21$	1:4	Weak evidence against $w$
$\Lambda$ CDM	CMB alone	$-3751.37 \pm 0.08$	0	—	—
$w$ CDM		$-3751.84 \pm 0.09$	$-0.47 \pm 0.12$	5:8	Inconclusive
$\Lambda$ CDM	CMB + SN	$-4017.15 \pm 0.09$	0	—	—
$w$ CDM		$-4019.10 \pm 0.09$	$-1.95 \pm 0.13$	1:7	Weak evidence against $w$

comparing  $\Lambda$ CDM with  $w$ CDM. The result becomes more defined when combining CMB and data from Type Ia Supernovae: in this case  $\Delta \ln \mathcal{Z} = -1.95$ .

In general it has to be kept in mind that the Bayesian selection is affected by the assumed choice of priors, as the prior normalization implies the evidence is decreased if the prior range is broader. We have however chosen a relatively narrow prior on  $f_{\text{NL}}$  and  $g_{\text{NL}}$  compared with existing constraints, meaning we are not unnecessarily penalizing the evidence for these PNG models.

## VI. CONCLUSIONS

Scale-dependent bias has arisen as a key means of detecting primordial non-Gaussianity, but its detection can be severely compromised by a number of potential systematics. These, if not accounted for, can introduce spurious large-scale structure and bias the inferred constraints on PNG. We have therefore conducted an analysis that carefully accounts for these systematic concerns and maximizes the information obtained from the most robust sources, in order to obtain the most precise to-date PNG measurements.

In particular, large differences are observed when comparing the MegaZ and CMASS LRG samples. As shown in Ref. [54], these appear to arise mainly from systematic relationships with foreground stars which, if unaccounted for, cause spurious large-scale power. Similarly, the NVSS autocorrelation is affected by severe systematic fluctuations in the source number density as a function of r.a. and declination. While correcting for these systematics is possible, such corrections are difficult to model in real space, and suppress the effect of PNG. We have chosen here the most conservative approach of not applying any correction and discarding the NVSS ACF. We have tested that that our results remain fully compatible if instead we apply a correction for the NVSS systematics. Finally, while quasar samples are potentially ideal for searching for primordial non-Gaussianity, we have confirmed earlier work [60] showing that the quasar autocorrelation measurements are compromised by systematic errors. In particular, there are correlations between the mean quasar density and a number of systematics including stellar density, galactic reddening and sky brightness.

While the existence of such systematics can be discouraging, they can often be corrected when they are understood. In addition, cross correlations between different data sets, which should generally contain different systematics, offer a much more robust means of addressing the question of primordial non-Gaussianity.

One example of this can be seen in the analysis of cross correlations with the CMB, previously analyzed in G12. Comparing the MegaZ and CMASS measurements, the CMB cross correlation was seen to be much less affected than the autocorrelations. While there were differences, these are partially accounted for by the larger sky coverage

of CMASS; the extra area has associated cosmic variance and is potentially more subject to extinction systematics. The signal decreased with respect to the G12 result, but the error bars are also smaller due to the larger DR8 sky coverage: the total signal-to-noise of this ISW detection alone is at the  $1.8\sigma$  level, in agreement with other recent measurements [55]. Similarly, using the raw NVSS data without subtracting the r.a. and declination dependent systematics changes its cross correlation somewhat but not dramatically, actually raising the inferred signal by 20%. However, the total ISW significance from the combined data is in the end nearly unchanged with respect to G12, as the increase in NVSS compensates the decrease in the LRGs. Also in this case, the results are consistent if we instead apply a correction to the NVSS systematics.

While measurements of cross correlations are more robust, they require more information for their interpretation, in particular understanding how the surveys overlap in redshift. We treat these uncertainties by marginalizing over nuisance parameters associated with the strength of the cross-correlation, so that it is its shape, rather than its amplitude, that constrains the PNG. Effectively, the overlap is determined by the small scale cross correlations, allowing us to use the large-scale measurements to constrain PNG. Fundamentally, it is the lack of large-scale excess power in any of the cross correlations which is responsible for our strong constraints. The precise strength of the constraints will depend on our assumptions about the bias evolution, particularly for the deepest surveys; however, the bias models which we use are consistent with our physical understanding of the tracers and other observations. Further, we have applied a general bias model, one that effectively assumes only that the bias smoothly tends towards unity as the redshift decreases, to the NVSS sample and found no significant change in our results.

We have measured that the allowed amount of PNG is  $-36 < f_{\text{NL}} < 45$  at 95%. This bound was determined from our conservative analysis, for which we excluded all autocorrelations apart from the CMASS LRGs, and instead relied on the the cross correlations between large-scale structure data sets and also CMB cross correlations which are sensitive to scale-dependent bias through the ISW effect. We found the data consistent with the simplest  $\Lambda$ CDM model in all cases. This is significantly different than the positive signal detected if all possible autocorrelations are included, as the signal is driven up by the large-scale clustering in the quasar catalog.

Our measurements have also allowed us to constrain more general PNG models, as well as models of dark energy. As yet, there are no indications of deviations from Gaussianity, or dynamic dark energy.

The new Planck bispectrum constraints on PNG, released while this paper was being revised [119], independently confirmed our findings. Future large-scale structure measurements, such as will arise from DES, LSST and

Euclid, may well determine whether the simplest inflationary models are correct. However, the more sensitive the results, the more prone they are to potential systematics, making a careful comparison and correlation between a range of different measurements indispensable.

## APPENDIX: FAST CALCULATION OF PNG SPECTRA

Our analysis requires scanning over a large number of possible model parameters, which is potentially very computationally expensive if a CMB code such as CAMB [110] must be rerun for every parameter choice. Here we describe how we can make the calculation of the PNG effects more efficient, so that a range of bias choices and  $f_{\text{NL}}$  values can be calculated using a single run of CAMB.

### 1. Gaussian case

In the Gaussian case, number density perturbations at a position  $\mathbf{x}$  and redshift  $z$  are defined at linear level as

$$\delta_g(\mathbf{x}, z) = b_g(z)\delta(\mathbf{x}, z), \quad (\text{A1})$$

where  $b_g(z)$  is the galaxy bias and  $\delta(\mathbf{x}, z)$  are the underlying dark matter perturbations. Projection onto the celestial sphere is achieved by summing over the past light cone, so that the 2D number density perturbation in a direction  $\hat{\mathbf{n}}$  is

$$\delta_g(\hat{\mathbf{n}}) = \int dz \varphi(z) b(z) \delta(\mathbf{x}, 0) D(z), \quad (\text{A2})$$

where  $D(z)$  is the linear growth function normalized to unity today and  $\varphi(z)$  is the normalized redshift distribution of the sources.

We can expand as usual these perturbations into harmonic space by means of a Fourier transform and an expansion in spherical harmonics  $Y_{lm}$

$$\begin{aligned} \delta_g(\mathbf{x}, z) &= \sum_{\mathbf{k}} \tilde{\delta}_g(k, z) e^{i\mathbf{k}\cdot\mathbf{x}} \\ &= \sum_{\mathbf{k}} \tilde{\delta}_g(k, z) \sum_{l,m} 4\pi i^l j_l(kr) Y_{lm}(\Omega_{\hat{\mathbf{k}}}) Y_{lm}^*(\Omega_{\hat{\mathbf{k}}}), \end{aligned} \quad (\text{A3})$$

where  $j_l$  are the spherical Bessel functions. We thus get the harmonic coefficients of the perturbations

$$a_{lm} = \sum_{\mathbf{k}} 4\pi i^l \tilde{\delta}(\mathbf{k}, z) Y_{lm}^*(\Omega_{\hat{\mathbf{k}}}) \int dz \varphi(z) b_g(z) D(z) j_l(kr).$$

We then introduce for simplicity the quantity

$$f_l(k) \equiv \int dz \varphi(z) b(z) D(z) j_l(kr), \quad (\text{A4})$$

so finally the projected angular power spectrum of the galaxy density perturbations can be written in a compact form in function of these quantities as

$$C_l^{gg} \equiv \langle a_{lm} a_{lm}^* \rangle = \frac{2}{\pi} \int dk k^2 P(k) f_l^2(k). \quad (\text{A5})$$

### 2. Non-Gaussian case

When local non-Gaussianity of the  $f_{\text{NL}}$  type is introduced, we know that at linear level the galactic bias will be corrected by

$$\Delta b(k, z) = \frac{2f_{\text{NL}}\delta_c}{\alpha(k, z)}, \quad (\text{A6})$$

where we have used the definition of  $\alpha$  of Eq. (3). While redshift evolution of the bias is difficult to model, a general template can be written as a function of three parameters  $b_0$ ,  $b_{-1}$ ,  $\gamma$  [111] as

$$b_1(z) = b_{-1} + \frac{b_0 - b_{-1}}{D^\gamma(z)}. \quad (\text{A7})$$

We will here consider three simplified cases, introducing different templates for it, such as simply a constant

$$b_1(z) = b_0, \quad (\text{A8})$$

or

$$b_1(z) = 1 + \frac{b_0 - 1}{D^\gamma(z)}, \quad (\text{A9})$$

which satisfies the property of approaching unity in the future, and is therefore appropriate for flux-limited samples, and

$$b_1(z) = \frac{b_0}{D^\gamma(z)}, \quad (\text{A10})$$

which instead tends to zero in the future, consistent with observations of quasars [112,113].

We can now generalize the 2D spectrum of Eq. (A5) between two galaxy catalogs  $i, j$  by noticing how the bias alterations due to PNG can be simply absorbed into the source function  $f_i(k)$ , which can be split in two:

$$\begin{aligned} C_l^{g_i g_j} &= \frac{2}{\pi} \int dk k^2 P(k) \\ &\times \left[ f_i^i(k) + \frac{g_l^i(k)}{\alpha(k, 0)} \right] \left[ f_i^j(k) + \frac{g_l^j(k)}{\alpha(k, 0)} \right], \end{aligned} \quad (\text{A11})$$

where we have absorbed the redshift dependence of  $\alpha$  (a factor of  $D$ ) into the integrand of  $g_l^i$ , and

$$\begin{aligned}
f_l^i(k) &\equiv \int dz \phi_i(z) b_1^i(z) D(z) j_l(kr) \\
g_l^i(k) &\equiv \int dz \phi_i(z) \{2f_{\text{NL}} \delta_c [b_1^i(z) - 1]\} j_l(kr). \quad (\text{A12})
\end{aligned}$$

In the integrand of  $g_l^i$  we have simplified the factor of  $D(z)$  from  $\alpha$ , so that its original  $D(z)$  has disappeared.

### 3. Fast decomposition of the spectra

It is useful to decompose the power spectra in a way which allows us to extract the bias and the  $f_{\text{NL}}$  parameters outside of the integrals, so that the calculation is faster.

#### a. Bias parameterization $b_1(z) = 1 + \frac{b_0 - 1}{D^{\gamma_i}(z)}$

To fix the ideas, let us assume for the redshift evolution of the bias that for a catalog  $i$  it holds

$$b_1^i(z) - 1 = \frac{b_0^i - 1}{D^{\gamma_i}(z)} \equiv \frac{b_{L0}^i}{D^{\gamma_i}(z)}. \quad (\text{A13})$$

We want now to express the power spectra using the parameter  $b_{L0}^i$ . The sources can be written explicitly as

$$\begin{aligned}
f_l^i(k) &= \int dz \phi_i(z) D(z) j_l(kr) \\
&\quad + b_{L0}^i \int dz \frac{\phi_i(z)}{D^{\gamma_i-1}(z)} j_l(kr) \\
g_l^i(k) &= 2b_{L0}^i f_{\text{NL}} \delta_c \int dz \frac{\phi_i(z)}{D^{\gamma_i}(z)} j_l(kr). \quad (\text{A14})
\end{aligned}$$

We can introduce here three source terms

$$\begin{aligned}
f_l^{i,0} &\equiv \int dz \phi_i(z) D(z) j_l(kr) \\
f_l^{i,1} &\equiv \int dz \frac{\phi_i(z)}{D^{\gamma_i-1}(z)} j_l(kr) \\
g_l^{i,1} &\equiv 2\delta_c \int dz \frac{\phi_i(z)}{D^{\gamma_i}(z)} j_l(kr), \quad (\text{A15})
\end{aligned}$$

which are defined inside CAMB in the file EQUATIONS.F90, subroutine OUTPUT. The previous expressions can be simply rewritten as

$$\begin{aligned}
f_l^i(k) &= f_l^{i,0}(k) + b_{L0}^i f_l^{i,1}(k) \\
g_l^i(k) &= f_{\text{NL}} b_{L0}^i g_l^{i,1}(k), \quad (\text{A16})
\end{aligned}$$

which can be done in CMBMAIN.F90, subroutine CALCSCALCLS. In this way the dependencies on both parameters  $b_{L0}^i$  and  $f_{\text{NL}}$  are factored out; the power spectra can then be easily calculated using Eq. (A11).

#### b. Bias parameterization $b_1(z) = \frac{b_0}{D^{\gamma_i}(z)}$

In this case,  $b_1^i(z) - 1 = \frac{b_0^i}{D^{\gamma_i}(z)} - 1 = \frac{b_{L0}^i + 1}{D^{\gamma_i}(z)} - 1$ , so that we can write (notice here we use  $b_0$  as nuisance parameter):

$$\begin{aligned}
f_l^i(k) &= b_0^i \int dz \frac{\phi_i(z)}{D^{\gamma_i-1}(z)} j_l(kr) \\
g_l^i(k) &= -2f_{\text{NL}} \delta_c \int dz \phi_i(z) j_l(kr) \\
&\quad + 2b_0^i f_{\text{NL}} \delta_c \int dz \frac{\phi_i(z)}{D^{\gamma_i}(z)} j_l(kr). \quad (\text{A17})
\end{aligned}$$

We can define then the source terms

$$\begin{aligned}
f_l^{i,1} &\equiv \int dz \frac{\phi_i(z)}{D^{\gamma_i-1}(z)} j_l(kr) \\
g_l^{i,0} &\equiv -\delta_c \int dz \phi_i(z) j_l(kr) \\
g_l^{i,1} &\equiv 2\delta_c \int dz \frac{\phi_i(z)}{D^{\gamma_i}(z)} j_l(kr), \quad (\text{A18})
\end{aligned}$$

so the previous expressions can be rewritten as

$$\begin{aligned}
f_l^i(k) &= b_0^i f_l^{i,1}(k) \\
g_l^i(k) &= f_{\text{NL}} g_l^{i,0}(k) + f_{\text{NL}} b_0^i g_l^{i,1}(k). \quad (\text{A19})
\end{aligned}$$

The constant bias case  $b_1(z) = b_0$  can be also derived from these expressions.

### ACKNOWLEDGMENTS

We thank Christian T. Byrnes, Marilena LoVerde, Sabino Matarrese, Hiranya Peiris, Cristiano Porciani and Jussi Väliiviita for useful discussions. T. G. acknowledges support from the Alexander von Humboldt Foundation, and from the Trans-Regional Collaborative Research Center TRR 33—“The Dark Universe” of the Deutsche Forschungsgemeinschaft (DFG). T. G. gratefully acknowledges the ICG Portsmouth for hospitality and for computing time on the Sciamia facility. R. C., R. N., W. P. and A. R. gratefully acknowledge financial support from the STFC via the rolling Grant No. ST/I001204/1. W. P. acknowledges funding from the European Research Council under the European Union’s Seventh Framework Programme (FP/2007-2013) / ERC Grant Agreement No. 202686—MDEPUGS. A. R. acknowledges funding from The University of Portsmouth Research Infrastructure Funding Project. Funding for SDSS-III has been provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Science Foundation and the US Department of Energy Office of Science. The SDSS-III web site is <http://www.sdss3.org/>. SDSS-III is managed by

the Astrophysical Research Consortium for the Participating Institutions of the SDSS-III Collaboration, including the University of Arizona, the Brazilian Participation Group, Brookhaven National Laboratory, University of Cambridge, Carnegie Mellon University, University of Florida, the French Participation Group, the German Participation Group, Harvard University, the Instituto de Astrofísica de Canarias, the Michigan State/Notre Dame/JINA Participation Group, The Johns Hopkins

University, Lawrence Berkeley National Laboratory, Max Planck Institute for Astrophysics, Max Planck Institute for Extraterrestrial Physics, New Mexico State University, New York University, Ohio State University, Pennsylvania State University, University of Portsmouth, Princeton University, the Spanish Participation Group, University of Tokyo, University of Utah, Vanderbilt University, University of Virginia, University of Washington and Yale University.

- 
- [1] D.H. Lyth and A.R. Liddle, *The primordial density perturbation: Cosmology, inflation and the origin of structure* (Cambridge University Press, Cambridge, England, 2009).
- [2] C. T. Byrnes and K.-Y. Choi, *Adv. Astron.* **2010**, 724525 (2010).
- [3] D. Larson *et al.*, *Astrophys. J. Suppl. Ser.* **192**, 16 (2011).
- [4] E. Komatsu *et al.*, *Astrophys. J. Suppl. Ser.* **192**, 18 (2011).
- [5] G. Hinshaw, *Astrophys. J. Suppl. Ser.* **208**, 19 (2013).
- [6] C.L. Bennett *et al.*, *Astrophys. J. Suppl. Ser.* **208**, 20 (2013).
- [7] J. Väliväita and T. Giannantonio, *Phys. Rev. D* **80**, 123516 (2009).
- [8] J. Maldacena, *J. High Energy Phys.* **05** (2003) 013.
- [9] V. Acquaviva, N. Bartolo, S. Matarrese, and A. Riotto, *Nucl. Phys.* **B667**, 119 (2003).
- [10] T. Suyama, T. Takahashi, M. Yamaguchi, and S. Yokoyama, *J. Cosmol. Astropart. Phys.* **12** (2010) 030.
- [11] E. Komatsu, *Classical Quantum Gravity* **27**, 124010 (2010).
- [12] Since this paper was submitted, Planck results have been released, showing no significant PNG detections at a level of  $-8.9 < f_{\text{NL}} < 14.3$  (95% C.L.) [119].
- [13] J. Smidt, A. Amblard, C. T. Byrnes, A. Cooray, A. Heavens, and D. Munshi, *Phys. Rev. D* **81**, 123007 (2010).
- [14] M. Liguori, E. Sefusatti, J.R. Fergusson, and E. P. S. Shellard, *Adv. Astron.* **2010**, 980523 (2010).
- [15] E. Sefusatti, M. Crocce, and V. Desjacques, *Mon. Not. R. Astron. Soc.* **425**, 2903 (2012).
- [16] A. Taruya, K. Koyama, and T. Matsubara, *Phys. Rev. D* **78**, 123534 (2008).
- [17] S. Hilbert, L. Marian, R. E. Smith, and V. Desjacques, *Mon. Not. R. Astron. Soc.* **426**, 2870 (2012).
- [18] S. Matarrese, L. Verde, and R. Jimenez, *Astrophys. J.* **541**, 10 (2000).
- [19] M. LoVerde, A. Miller, S. Shandera, and L. Verde, *J. Cosmol. Astropart. Phys.* **04** (2008) 014.
- [20] A. Pillepich, C. Porciani, and O. Hahn, *Mon. Not. R. Astron. Soc.* **402**, 191 (2010).
- [21] I. E. Achitouv and P. S. Corasaniti, *J. Cosmol. Astropart. Phys.* **02** (2012) 002.
- [22] N. Dalal, O. Doré, D. Huterer, and A. Shirokov, *Phys. Rev. D* **77**, 123514 (2008).
- [23] S. Matarrese and L. Verde, *Astrophys. J. Lett.* **677**, L77 (2008).
- [24] P. Valageas, *Astron. Astrophys.* **514**, A46 (2010).
- [25] V. Desjacques, U. Seljak, and I. T. Iliev, *Mon. Not. R. Astron. Soc.* **396**, 85 (2009).
- [26] V. Desjacques and U. Seljak, *Classical Quantum Gravity* **27**, 124011 (2010).
- [27] V. Desjacques and U. Seljak, *Adv. Astron.* **2010**, 908640 (2010).
- [28] T. Giannantonio and C. Porciani, *Phys. Rev. D* **81**, 063530 (2010).
- [29] V. Desjacques, M. Crocce, R. Scoccimarro, and R. K. Sheth, *Phys. Rev. D* **82**, 103529 (2010).
- [30] V. Desjacques, D. Jeong, and F. Schmidt, *Phys. Rev. D* **84**, 063512 (2011).
- [31] V. Desjacques, D. Jeong, and F. Schmidt, *Phys. Rev. D* **84**, 061301 (2011).
- [32] N. Hamaus, U. Seljak, and V. Desjacques, *Phys. Rev. D* **84**, 083509 (2011).
- [33] A. Slosar, C. Hirata, U. Seljak, S. Ho, and N. Padmanabhan, *J. Cosmol. Astropart. Phys.* **08** (2008) 031.
- [34] N. Afshordi and A. J. Tolley, *Phys. Rev. D* **78**, 123507 (2008).
- [35] J.-Q. Xia, A. Bonaldi, C. Baccigalupi, G. De Zotti, S. Matarrese, L. Verde, and M. Viel, *J. Cosmol. Astropart. Phys.* **08** (2010) 013.
- [36] J.-Q. Xia, M. Viel, C. Baccigalupi, G. De Zotti, S. Matarrese, and L. Verde, *Astrophys. J. Lett.* **717**, L17 (2010).
- [37] J.-Q. Xia, C. Baccigalupi, S. Matarrese, L. Verde, and M. Viel, *J. Cosmol. Astropart. Phys.* **08** (2011) 033.
- [38] A. J. Ross *et al.*, *Mon. Not. R. Astron. Soc.* **428**, 1116 (2013).
- [39] V. Desjacques and U. Seljak, *Phys. Rev. D* **81**, 023006 (2010).
- [40] A. Mana, T. Giannantonio, J. Weller, B. Hoyle, G. Huetsi, and B. Sartoris, *arXiv:1303.0287*.
- [41] C. Fedeli, L. Moscardini, and S. Matarrese, *Mon. Not. R. Astron. Soc.* **397**, 1125 (2009).
- [42] C. Carbone, O. Mena, and L. Verde, *J. Cosmol. Astropart. Phys.* **07** (2010) 020.
- [43] T. Giannantonio, C. Porciani, J. Carron, A. Amara, and A. Pillepich, *Mon. Not. R. Astron. Soc.* **422**, 2854 (2012).

- [44] Z. Huang, L. Verde, and F. Vernizzi, *J. Cosmol. Astropart. Phys.* **04** (2012) 005.
- [45] A. Pillepich, C. Porciani, and T. H. Reiprich, *Mon. Not. R. Astron. Soc.* **422**, 44 (2012).
- [46] C. Fedeli, C. Carbone, L. Moscardini, and A. Cimatti, *Mon. Not. R. Astron. Soc.* **414**, 1545 (2011).
- [47] C. Cunha, D. Huterer, and O. Doré, *Phys. Rev. D* **82**, 023004 (2010).
- [48] R. K. Sachs and A. M. Wolfe, *Astrophys. J.* **147**, 73 (1967).
- [49] J. A. Frieman, M. S. Turner, and D. Huterer, *Annu. Rev. Astron. Astrophys.* **46**, 385 (2008).
- [50] T. Giannantonio, R. Crittenden, R. Nichol, and A. J. Ross, *Mon. Not. R. Astron. Soc.* **426**, 2581 (2012).
- [51] R. G. Crittenden and N. Turok, *Phys. Rev. Lett.* **76**, 575 (1996).
- [52] S. Ho, C. Hirata, N. Padmanabhan, U. Seljak, and N. Bahcall, *Phys. Rev. D* **78**, 043519 (2008).
- [53] T. Giannantonio, R. Scranton, R. G. Crittenden, R. C. Nichol, S. P. Boughn, A. D. Myers, and G. T. Richards, *Phys. Rev. D* **77**, 123520 (2008).
- [54] A. J. Ross *et al.*, *Mon. Not. R. Astron. Soc.* **417**, 1350 (2011).
- [55] C. Hernández-Monteagudo *et al.* (to be published).
- [56] K. S. Dawson *et al.*, *Astron. J.* **145**, 10 (2013).
- [57] S. Ho *et al.*, *Astrophys. J.* **761**, 14 (2012).
- [58] A. J. Ross *et al.*, *Mon. Not. R. Astron. Soc.* **424**, 564 (2012).
- [59] D. Huterer, C. E. Cunha, and W. Fang, *Mon. Not. R. Astron. Soc.* **432**, 2945 (2013).
- [60] A. R. Pullen and C. M. Hirata, *Publ. Astron. Soc. Pac.* **125**, 705 (2013).
- [61] S. Matarrese and L. Verde, *Astrophys. J. Lett.* **677**, L77 (2008).
- [62] N. Afshordi and A. J. Tolley, *Phys. Rev. D* **78**, 123507 (2008).
- [63] F. Schmidt and M. Kamionkowski, *Phys. Rev. D* **82**, 103002 (2010).
- [64] V. Desjacques, D. Jeong, and F. Schmidt, *Phys. Rev. D* **84**, 061301 (2011).
- [65] K. M. Smith, S. Ferraro, and M. LoVerde, *J. Cosmol. Astropart. Phys.* **03** (2012) 032.
- [66] N. Roth and C. Porciani, *Mon. Not. R. Astron. Soc.* **425**, L81 (2012).
- [67] J. Yoo, N. Hamaus, U. Seljak, and M. Zaldarriaga, *Phys. Rev. D* **86**, 063514 (2012).
- [68] X. Chen, *Phys. Rev. D* **72**, 123518 (2005).
- [69] C. T. Byrnes, S. Nurmi, G. Tasinato, and D. Wands, *J. Cosmol. Astropart. Phys.* **02** (2010) 034.
- [70] A. Becker and D. Huterer, *Phys. Rev. Lett.* **109**, 121302 (2012).
- [71] D. G. York *et al.*, *Astron. J.* **120**, 1579 (2000).
- [72] M. F. Skrutskie *et al.*, *Astron. J.* **131**, 1163 (2006).
- [73] H. Aihara *et al.*, *Astrophys. J. Suppl. Ser.* **193**, 29 (2011).
- [74] E. Boldt, *Phys. Rep.* **146**, 215 (1987).
- [75] K. M. Górski, E. Hivon, A. J. Banday, B. D. Wandelt, F. K. Hansen, M. Reinecke, and M. Bartelmann, *Astrophys. J.* **622**, 759 (2005).
- [76] A. Collister *et al.*, *Mon. Not. R. Astron. Soc.* **375**, 68 (2007).
- [77] S. A. Thomas, F. B. Abdalla, and O. Lahav, arXiv:1011.2448.
- [78] S. A. Thomas, F. B. Abdalla, and O. Lahav, *Phys. Rev. Lett.* **106**, 241301 (2011).
- [79] C. P. Ahn *et al.*, *Astrophys. J. Suppl. Ser.* **203**, 21 (2012).
- [80] L. Anderson *et al.*, *Mon. Not. R. Astron. Soc.* **427**, 3435 (2012).
- [81] B. A. Reid *et al.*, *Mon. Not. R. Astron. Soc.* **426**, 2719 (2012).
- [82] L. Samushia, *et al.*, *Mon. Not. R. Astron. Soc.* **429**, 1514 (2013).
- [83] A. G. Sánchez *et al.*, *Mon. Not. R. Astron. Soc.* **425**, 415 (2012).
- [84] C. G. o. Scóccola, arXiv:1209.1394.
- [85] R. Tojeiro *et al.*, *Mon. Not. R. Astron. Soc.* **424**, 2339 (2012).
- [86] G.-B. Zhao *et al.*, arXiv:1211.3741.
- [87] H.-J. Seo *et al.*, *Astrophys. J.* **761**, 13 (2012).
- [88] R. de Putter *et al.*, *Astrophys. J.* **761**, 12 (2012).
- [89] E. F. Schlafly and D. P. Finkbeiner, *Astrophys. J.* **737**, 103 (2011).
- [90] J. J. Condon, W. D. Cotton, E. W. Greisen, Q. F. Yin, R. A. Perley, G. B. Taylor, and J. J. Broderick, *Astron. J.* **115**, 1693 (1998).
- [91] C. Blake and J. Wall, *Mon. Not. R. Astron. Soc.* **329**, L37 (2002).
- [92] C. Blake and J. Wall, *Mon. Not. R. Astron. Soc.* **337**, 993 (2002).
- [93] C. Blake, P. G. Ferreira, and J. Borrill, *Mon. Not. R. Astron. Soc.* **351**, 923 (2004).
- [94] M. Negrello, M. Magliocchetti, and G. De Zotti, *Mon. Not. R. Astron. Soc.* **368**, 935 (2006).
- [95] S. P. Boughn and R. G. Crittenden, *Phys. Rev. Lett.* **88**, 021302 (2001).
- [96] S. Boughn and R. Crittenden, *Nature (London)* **427**, 45 (2004).
- [97] M. R. Nolta *et al.*, *Astrophys. J.* **608**, 10 (2004).
- [98] D. Pietrobon, A. Balbi, and D. Marinucci, *Phys. Rev. D* **74**, 043524 (2006).
- [99] J. D. McEwen, Y. Wiaux, M. P. Hobson, P. Vanderghaynst, and A. N. Lasenby, *Mon. Not. R. Astron. Soc.* **384**, 1289 (2008).
- [100] A. Raccanelli, A. Bonaldi, M. Negrello, S. Matarrese, G. Tormen, and G. de Zotti, *Mon. Not. R. Astron. Soc.* **386**, 2161 (2008).
- [101] C. Hernández-Monteagudo, *Astron. Astrophys.* **520**, A101 (2010).
- [102] F. Schiavon, F. Finelli, A. Gruppuso, A. Marcos-Caballero, P. Vielva, R. G. Crittenden, R. B. Barreiro, and E. Martínez-González, arXiv:1203.3277.
- [103] K. M. Smith, O. Zahn, and O. Doré, *Phys. Rev. D* **76**, 043510 (2007).
- [104] K. M. Smith and D. Huterer, *Mon. Not. R. Astron. Soc.* **403**, 2 (2010).
- [105] B. Leistedt, H. V. Peiris, D. J. Mortlock, A. Benoit-Lévy, and A. Pontzen, *Mon. Not. R. Astron. Soc.* **435**, 1857 (2013).
- [106] K. Abazajian *et al.*, *Astron. J.* **129**, 1755 (2005).
- [107] J. K. Adelman-McCarthy *et al.*, *Astrophys. J. Suppl. Ser.* **175**, 297 (2008).

- [108] G. T. Richards, A. D. Myers, A. G. Gray, R. N. Riegel, R. C. Nichol, R. J. Brunner, A. S. Szalay, D. P. Schneider, and S. F. Anderson, *Astrophys. J. Suppl. Ser.* **180**, 67 (2009).
- [109] R. Amanullah *et al.*, *Astrophys. J.* **716**, 712 (2010).
- [110] A. Lewis, A. Challinor, and A. Lasenby, *Astrophys. J.* **538**, 473 (2000).
- [111] L. Moscardini, P. Coles, F. Lucchin, and S. Matarrese, *Mon. Not. R. Astron. Soc.* **299**, 95 (1998).
- [112] N. Padmanabhan, M. White, P. Norberg, and C. Porciani, *Mon. Not. R. Astron. Soc.* **397**, 1862 (2009).
- [113] M. White *et al.*, *Mon. Not. R. Astron. Soc.* **424**, 933 (2012).
- [114] S. P. Boughn, R. G. Crittenden, and G. P. Koehrsen, *Astrophys. J.* **580**, 672 (2002).
- [115] A. Lewis and S. Bridle, *Phys. Rev. D* **66**, 103511 (2002).
- [116] F. Feroz, M. P. Hobson, and M. Bridges, *Mon. Not. R. Astron. Soc.* **398**, 1601 (2009).
- [117] A. G. Riess *et al.*, *Astrophys. J.* **699**, 539 (2009).
- [118] R. Trotta, *Contemp. Phys.* **49**, 71 (2008).
- [119] Planck Collaboration, arXiv:1303.5084.