The large-scale 3-point correlation function of the SDSS BOSS DR12 CMASS galaxies

Zachary Slepian[1]⋆, Daniel J. Eisenstein[1]†, Florian Beutler[2], Antonio J. Cuesta[3],
Jian Ge[4], Héctor Gil-Marín[5,6,7], Shirley Ho[8], Francisco-Shu Kitaura[9],
Cameron K. McBride[1], Robert C. Nichol[7], Will J. Percival[7],
Sergio Rodríguez-Torres[10,11,12], Ashley J. Ross[13], Román Scoccimarro[14], Hee-Jong Seo[15],
Jeremy Tinker[14], Rita Tojeiro[16], & Mariana Vargas-Magaña[17]

[1] *Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, MA 02138, USA*
[2] *Lawrence Berkeley National Lab, 1 Cyclotron Rd, Berkeley CA 94720, USA*
[3] *Institut de Ciències del Cosmos (ICCUB), Universitat de Barcelona (IEEC-UB), Martí i Franquès 1, E08028 Barcelona, Spain*
[4] *Astronomy Department, University of Florida, 211 Bryant Space Science Center, Gainesville, FL 32611, USA*
[5] *Sorbonne Universités, Institut Lagrange de Paris (ILP), 98 bis Boulevard Arago, 75014 Paris, France*
[6] *Laboratoire de Physique Nucléaire et de Hautes Energies, Université Pierre et Marie Curie, 4 Place Jussieu, 75005 Paris, France*
[7] *Institute of Cosmology & Gravitation, University of Portsmouth, Dennis Sciama Building, Portsmouth PO1 3FX, UK*
[8] *Bruce and Astrid McWilliams Center for Cosmology, Department of Physics, Carnegie Mellon University,*
5000 Forbes Ave, Pittsburgh, PA 15213, USA
[9] *Leibniz-Institut für Astrophysik Potsdam (AIP), An der Sternwarte 16, D-14482 Potsdam, Germany*
[10] *Instituto de Física Teórica, (UAM/CSIC), Universidad Autónoma de Madrid, Cantoblanco, E-28049 Madrid, Spain*
[11] *Campus of International Excellence UAM+CSIC, Cantoblanco, E-28049 Madrid, Spain*
[12] *Departamento de Física Teórica M8, Universidad Autonoma de Madrid (UAM), Cantoblanco, E-28049 Madrid, Spain*
[13] *Center for Cosmology and Astroparticle Physics, Department of Physics, The Ohio State University, OH 43210, USA*
[14] *Center for Cosmology and Particle Physics, New York University, 4 Washington Place, NY 1003, New York, USA*
[15] *Department of Physics and Astronomy, Ohio University, Clippinger Labs, Athens, OH 45701*
[16] *University of St Andrews, North Haugh, St Andrews Fife, KY16 9SS, UK*
[17] *Instituto de Física, Universidad Nacional Autónoma de México, Apdo. Postal 20-364, México*

9 December 2015

ABSTRACT

We report a measurement of the large-scale 3-point correlation function of galaxies using the largest dataset for this purpose to date, $777,202$ Luminous Red Galaxies in the Sloan Digital Sky Survey Baryon Acoustic Oscillation Spectroscopic Survey (SDSS BOSS) DR12 CMASS sample. This work exploits the novel algorithm of Slepian & Eisenstein (2015b) to compute the multipole moments of the 3PCF in $\mathcal{O}(N^2)$ time, with N the number of galaxies. Leading-order perturbation theory models the data well in a compressed basis where one triangle side is integrated out. We also present an accurate and computationally efficient means of estimating the covariance matrix. With these techniques the redshift-space linear and non-linear bias are measured, with 2.6% precision on the former if σ_8 is fixed. The data also indicates a 2.8σ preference for the BAO, confirming the presence of BAO in the 3-point function.

1 INTRODUCTION

The spatial distribution of galaxies on large scales contains a wealth of information on both the Universe's contents and the laws under which these contents evolved. The 3-point correlation function (3PCF) is an important tool for quantifying this distribution (Groth & Peebles 1977; Peebles 1980; Bernardeau et al. 2002). It measures the excess probability over random of finding three galaxies whose separations form a given triangle. In the consensus picture of structure formation, density perturbations were set up at the end of inflation as a nearly Gaussian random field (Starobinsky 1980;

Bardeen, Steinhardt & Turner 1983). These primordial fluctuations are thus expected to be nearly fully characterised by their 2-point correlation function (2PCF), with a negligible 3PCF (Planck Paper XVII, 2015). However, subsequent evolution generates a 3PCF as well as imprinting additional information in the 2PCF. Further, the small level of primordial non-Gaussianity (PNG) generically predicted at the end of inflation produces its own unique signature in the 3PCF (Desjacques & Seljak 2008, for a review; Sefusatti & Komatsu 2007 for forecasts; for recent observational work using large-scale structure, see Xia et al. 2010, 2011; Ross et al. 2013;

Giannantonio et al. 2014; for Cosmic Microwave Background constraints, see Planck Paper XVII, 2015).

This picture applies within the context of perturbation theory (PT) for the cosmic density field, i.e for small perturbations around a homogeneous background. The galaxies we observe are tracers of this density field, but do so with some bias because galaxy formation is a complicated process (Kaiser 1984; Dekel & Rees 1987). In addition to information on the initial conditions and on high redshift physics, the 2PCF and 3PCF therefore illuminate galaxy formation, revealing how faithfully galaxies trace the underlying matter density.

Since the 1970s, both 3PCF and 2PCF have been measured in galaxy redshift surveys of increasingly large volume and number of objects. In the last decade a particular focus of 2PCF measurements has been to use the Baryon Acoustic Oscillation (BAO) method, which exploits a sharp bump in the 2PCF, imprinted by sound waves in the early Universe, to constrain the Universe's expansion history and thus the cosmological parameters (Eisenstein, Hu & Tegmark 1998; Blake & Glazebrook 2003; Hu & Haiman 2003; Linder 2003; Seo & Eisenstein 2003). The 3PCF is intrinsically weaker than the 2PCF, and consequently has not been a powerful additional lever on the BAO scale. Moreover, BAO analyses now routinely use reconstruction, a technique to move galaxies backwards along their inferred velocity vectors to try to undo some of the low-redshift non-linear evolution. Reconstruction reduces the amplitude of the 3PCF further and puts some of the distance information in the 3PCF back into the 2PCF (Eisenstein et al. 2007; Padmanabhan, White & Cohn 2009; Burden et al. 2014). Schmittfull et al. (2015) explicitly shows how 3- and 4-point statistics of the unreconstructed density field enter the reconstructed 2PCF.

Measurements of the 3PCF have further been complicated by its scaling with the number of objects. In contrast to the 2PCF, which requires pair counting and hence scales as $N(nV_{\mathrm{max}})$, the 3PCF involves triplets of galaxies and so scales as $N(nV_{\mathrm{max}})^2$, where N and n are the number and number density of galaxies and V_{max} is the volume within a sphere of radius R_{max}, the maximal scale at which the correlations are measured. While nV_{max} is much smaller than N for large-volume surveys, this scaling has nonetheless mostly inhibited exploitation of the 3PCF on the acoustic scale of $r_{\mathrm{s}} \sim 100\ \mathrm{Mpc}/h$. It is not only measuring the 3PCF of the data that is time-consuming; in most methods, the number of points must also include a set of random points, much larger than the actual galaxy set, in order to model the survey geometry. Further, computing covariance matrices to fit parameters requires using potentially thousands of mock catalogs generated by N-body simulations combined with some prescription for matching galaxies to dark matter halos. There is also a problem of visualization: the 3PCF depends on three independent variables (three triangle sides, or two triangle sides and an angle), so it has not been clear how to show the results of an analysis. In practice what is often done is to show the angle dependence of the amplitude for certain fixed triangle side lengths or ratios.

For these reasons, 3PCF measurements have been limited in ability to exploit the full data sets used by the 2PCF. They often have focused on only certain triangle configurations (two-to-one, isosceles) and worked on scales smaller than the BAO scale. Even so these measurements have been computationally challenging.

Historically the 3PCF traces back to the pioneering work of Groth & Peebles (1977); our list of works here focuses on the recent and is not intended to be exhaustive.[1] The largest sample used for the 3PCF to date is 220,000 galaxies from SDSS DR6 in McBride et al. (2011a,b), with 2-to-1 triangles up to 27 Mpc/h. There have been a number of other works also using SDSS data. Nichol et al. (2006) use 36,738 LRGs for 2-to-1, 3-to-1, and 4-to-1 triangles on scales up to $\sim 40\ \mathrm{Mpc}/h$; Kulkarni et al. (2007) use 50,967 SDSS DR3 LRGs for isosceles, 2-to-1, and 3-to-1 triangles on scales up to 30 Mpc/h; Marín (2011) uses 105,831 Luminous Red Galaxies (LRGs) from SDSS DR7 for 2-to-1 triangles on scales up to 90 Mpc/h; Guo et al. (2014) use DR7 LRGs to measure isosceles, 2-to-1, and 3-to-1 configurations up to 40 Mpc/h. Guo et al. (2015) use LRGs in the DR11 CMASS sample to measure 2-to-1 configurations up to 40 Mpc/h.

Data from the Two-degree-Field Galaxy Redshift Survey (2dFGRS) has also been used for 3PCF measurements: Jing & Börner (2004) and Wang et al. (2004) use respectively $\sim 60,000$ and $\sim 250,000$ galaxies from this survey on scales up to respectively 40 Mpc/h and $\sim 20\ \mathrm{Mpc}/h$. The 3PCF has also been measured using the WiggleZ Dark Energy Survey: Marín et al. (2013) use isosceles, 2-to-1, and 3-to-1 triangles for scales up to 90 Mpc/h.

By concentrating only on particular triangle configurations, these measurements leave information in the 3PCF unexploited. Importantly, they also do not probe the BAO scale, which in principle offers cosmological distance information.

The only work to probe the BAO scale in the 3PCF thus far has been Gaztañaga et al. (2009) using a sample of $\sim 40,000$ LRGs from SDSS DR7. They find a $2 - 3\sigma$ detection of the BAO using all opening angles of a single triangle configuration with side lengths $r_1 = 33\ \mathrm{Mpc}/h$ and $r_2 = 88\ \mathrm{Mpc}/h$. They did not fit for the BAO scale to extract distance information.

Similar datasets to those discussed above have been used for measuring the bispectrum, the Fourier space analog of the 3PCF. While in principle the signal-to-noise is the same for the bispectrum and the 3PCF, in practice we believe Fourier space has several disadvantages relative to configuration space in this context. First, the bispectrum requires gridding the galaxies to a regular mesh to allow an $N_{\mathrm{g}} \log N_{\mathrm{g}}$ Fourier Transform (FT) to be computed, with N_{g} the number of grid cells. Consequently, it only samples a discrete set of wavenumbers, so some spatial information is lost. Second, in practice there is a truncation of the transform at high wavenumber, set by the Nyquist frequency given by the grid spacing. Third, edge correction is not straightforward in the bispectrum, as even simple survey boundaries in real space lead to complicated, infinite-support ringing in Fourier space. Fourth, the BAO features in the 3PCF are localized but in the bispectrum are dispersed, much as the sharp bump in the 2PCF corresponds to extended wiggles in the power spectrum. This localization is likely to make it

---

[1] Different authors use different parametrizations; to standardize we quote the largest triangle side possible in each work.

easier in practice to fit for BAO information in the 3PCF than in the bispectrum.

Nonetheless, a major advantage of the bispectrum is that it is obvious how to estimate it quickly. Even the most naive approach only requires comparison of pairs of wavevectors rather than galaxy triplets. Additional speed can come from using templates to directly extract bias parameters within a given bispectrum model if the angular and wave-vector-length dependences are separable (Schmittfull, Baldauf & Seljak 2015); there are also accelerations available in the absence of separability provided the bispectrum model can be expanded in a finite number of separable basis functions (Schmittfull, Regan & Shellard 2013). Further, it has long been known that the opening-angle-averaged bispectrum can be computed using FTs (Scoccimarro 2000), a result recently extended to the full, unaveraged bispectrum (Scoccimarro 2015). Previous work on the bispectrum has been Feldman et al. (2001), Scoccimarro et al. (2001), and Verde et al. (2002), with the most recent measurement, using 690,827 LRGs from the SDSS DR11 CMASS sample, by Gil-Marín et al. (2015). This measurement uses isosceles and 2-to-1 triangles in Fourier space up to scales of $k = 0.03\ h/\mathrm{Mpc} \sim 30\ \mathrm{Mpc}/h$. Gil-Marín et al. (2015) use only diagonal elements of the covariance matrix for parameter fitting since the large number of matrix elements relative to the number of mock catalogs available means the full covariance cannot be accurately computed empirically.

In this paper, we present the first measurements of the large-scale 3PCF for the CMASS sample of $777,202$ LRGs within SDSS DR12, working on scales from $75 - 180\ \mathrm{Mpc}/h$. We use the novel algorithm of Slepian & Eisenstein (2015b), hereafter SE15b, which computes the 3PCF's multipole moments in $\mathcal{O}(N(nV_{\mathrm{max}}))$ time using spherical harmonic decompositions; the covariance matrix also turns out to be tractable in the multipole basis (SE15b). The main outcomes of this work are:

1) A highly accurate covariance matrix derived analytically but matching mock catalogs well.

2) Demonstrating that perturbation theory is in agreement with the data on these scales.

3) Measurement of the redshift-space linear and non-linear bias with 2.6% precision on the former.

4) A 2.8σ detection of the BAO feature in the 3PCF.

The paper is laid out as follows. In §2 we briefly introduce the data set, the random catalogs used for edge correction, and the mock catalogs used to verify our analysis pipeline and covariance matrix. §3 presents our multipole basis for the 3PCF and reviews the algorithm used to do the measurement. §4 recapitulates a compression scheme developed in previous work that we apply to the 3PCF here. §5 details our perturbation theory model, while §6 describes our approach to obtaining the covariance matrix. §7 shows the results of our analysis on both mock catalogs and the CMASS data. §8 concludes.

2 DATA, RANDOMS, AND MOCKS

Here we summarize the data set, random catalogs, and mock catalogs used. The data are $777,202$ LRGs within the CMASS sample of DR12 (Alam et al. 2015) of the Baryon Oscillation Spectroscopic Survey (BOSS; Eisenstein et al. 2011; Dawson et al. 2013) in the redshift range 0.43 to 0.7, with $568,776$ in the North Galactic Cap over $6,934$ square degrees and $208,426$ in the South Galactic Cap over $2,560$ square degrees, for a total of $9,493$ square degrees (Reid et al. 2016, Table 2). The sample is color-selected to have roughly constant stellar mass and observations were completed in 2014. Target selection and the algorithms used for creating the galaxy catalogues and the associated random catalogues are presented in Reid et al. (2016), with the treatment for observational systematic biases described in Ross et al. (2012) and Ross et al. (2015). The 2PCF of this dataset is used to constrain the cosmic distance scale via the BAO method in Cuesta et al. (2015).

In total, the SDSS (York et al. 2000), comprising SDSS I, II (Abazajian et al. 2009) and SDSS III (Eisenstein et al. 2011) imaged more than one-third of the sky (14,555 square degrees) with a drift-scanning mosaic CCD camera (Gunn et al. 1998) in five photometric bandpasses (Fukugita et al. 1996; Smith et al. 2002; Doi et al. 2010) on the 2.5-m Sloan Telescope (Gunn et al. 2006) at Apache Point Observatory in New Mexico. The data underwent astrometric calibration (Pier et al. 2003), photometric reduction (Lupton et al. 2001), photometric calibration (Padmanabhan et al. 2008), and reprocessing for Data Release 8 (Aihara et al. 2011). Within BOSS, targets were assigned to tiles with an adaptive algorithm (Blanton et al. 2003), spectra obtained using double-armed spectrographs (Smee et al. 2013), and redshifts derived as described in Bolton et al. (2012).

We also ran our pipeline on 299 mock catalogs developed for DR12 known as the MULTIDARK-PATCHY BOSS DR12 mocks (Kitaura, Yepes & Prada 2014; Kitaura et al. 2015a,b). These catalogs used second-order Lagrangian perturbation theory (2LPT) combined with a spherical collapse model on small scales (Kitaura & Heß 2013). They were calibrated on accurate N-body-based reference catalogs using halo abundance matching to reproduce the number density, clustering bias, selection function, and survey geometry of the BOSS data (Rodríguez-Torres et al. 2015).

3 BASIS AND ALGORITHM

We parameterize the 3PCF using two triangle sides r_1 and r_2 and the cosine of the angle between them, $\cos\theta_{12} = \hat{r}_1 \cdot \hat{r}_2$. We then project the angular dependence onto the basis of Legendre polynomials P_l of $\cos\theta_{12}$, so that the 3PCF is written as

$$\zeta(r_1, r_2; \hat{r}_1 \cdot \hat{r}_2) = \sum_l \zeta_l(r_1, r_2) P_l(\hat{r}_1 \cdot \hat{r}_2). \qquad (1)$$

The multipole coefficients ζ_l capture the 3PCF's dependence on side lengths at each multipole. We bin in the two side lengths but our measurement uses the exact angle between the triangle sides when projecting onto the multipoles. The Legendre polynomials of even or odd l involve a falling sequence of even or odd powers of $\cos\theta_{12}$, with l giving the maximal power. This expansion was first proposed by Szapudi (2004) but has up to now not been used save for the monopole term. Since $P_0 = 1$, the monopole corresponds to measuring the angle-averaged 3PCF (Pan & Szapudi 2005). The motivation for this expansion is that at leading order

in perturbation theory, prior to cyclically summing over all triangle vertices, the 3PCF has only $l = 0, 1$, and 2 moments (the perturbation theory (PT) is reviewed in Bernardeau et al. 2002). Cyclically summing around the triangle vertices then introduces higher moments but they do not contain substantial new information, tending to be dominated by the same population of squeezed triangles. These points are further discussed in Slepian & Eisenstein (2015a,b), hereafter SE15a and SE15b.

The Legendre polynomials and radial coefficients of equation (1) form a complete, orthogonal basis for any functions depending only on the relative angle $\hat{r}_1 \cdot \hat{r}_2$. This basis builds in isotropy, which is in reality not true due to redshift-space distortions (RSD; see Hamilton 1998 for a review). By using this basis, we spherically average over all orientations to the line of sight, so the effects of RSD get integrated over much as is the case in the spherically averaged, isotropic 2PCF. We discuss RSD further in §5.2.

We directly measure the 3PCF's multipole coefficients ζ_l using an algorithm presented in SE15b that substantially accelerates the computation. We construct the counts over triples of galaxies or randoms as follows. For each galaxy in the sample successively, we bin the surrounding galaxies into concentric spherical shells set by our radial binning around that galaxy and store the unit vectors for each of the separations. On each spherical shell, we compute the multipole moments of the unit vectors and then convert these to the spherical harmonic decomposition $a_{\ell m}$ of the galaxy field. We then form all of the cross-powers between the shells as a function of the order ℓ, summing over the azimuthal parameter m. We finally sum the resulting cross-powers over all central galaxies. The spherical harmonic computation around each galaxy is $\mathcal{O}(nV_{\max})$ and we must do this around each of the N galaxies in the survey, leading to a 3PCF measurement that scales as $N(nV_{\max})$. Further details such as implementation, tests and runtimes are given in SE15b.

We now briefly recapitulate the edge-correction framework of SE15b. We measure the projection onto multipoles of the Szapudi-Szalay (1998) estimator for the 3PCF,

$$\zeta = \frac{NNN}{RRR} = \sum \zeta_l(r_1, r_2) P_l(\hat{r}_1 \cdot \hat{r}_2) \qquad (2)$$

where $N \equiv D - R$, D denotes a data point, and R denotes a random point. It can be shown that if the NNN and RRR counts are given by the multipole series

$$NNN = \sum_j \mathcal{N}_j(r_1, r_2) P_j(\hat{r}_1 \cdot \hat{r}_2) \qquad (3)$$

and

$$RRR = \sum_{l'} \mathcal{R}_{l'}(r_1, r_2) P_{l'}(\hat{r}_1 \cdot \hat{r}_2) \qquad (4)$$

then

$$\frac{\mathcal{N}_k}{\mathcal{R}_0} = \zeta_k + \sum_l \zeta_l (2k+1) \sum_{l'>0} \left(\begin{array}{ccc} l & l' & k \\ 0 & 0 & 0 \end{array} \right)^2 f_{l'}. \qquad (5)$$

The ζ_l are the desired multipole coefficients of the 3PCF and the $f_{l'} \equiv \mathcal{R}_{l'}/\mathcal{R}_0$ are the ratios of higher multipoles of the randoms' triple count to its monopole. The Wigner 3j-symbol represents angular momentum coupling; it satisfies the Triangle Inequality $|l - l'| \leqslant k \leqslant |l + l'|$ and $l + l' + k$ must be even by parity.

We see that a vector of measured $\mathcal{N}_k/\mathcal{R}_0$ is simply a linear transformation of the vector of ζ_k we desire, and given the matrix $\mathbf{I} + \mathbf{M}$ where $\mathbf{I}$ is the identity matrix and $\mathbf{M}$ has elements

$$M_{kl} = (2k+1) \sum_{l'>0} \left(\begin{array}{ccc} l & l' & k \\ 0 & 0 & 0 \end{array} \right)^2 f_{l'}, \qquad (6)$$

the edge-correction equation becomes

$$\vec{\mathcal{N}}/\mathcal{R}_0 = (\boldsymbol{I} + \boldsymbol{M})\vec{\zeta} \equiv \boldsymbol{A}\vec{\zeta}, \qquad (7)$$

where $\vec{\mathcal{N}} = (\mathcal{N}_0, \mathcal{N}_1, \cdots, \mathcal{N}_{l_{\max}})$, and analogously for $\vec{\zeta}$, with $l_{\max}$ the maximum multipole to which we measure. This edge-correction equation can then be solved for $\vec{\zeta}$ by matrix inversion. In practice the $f_{l'}$ fall rapidly with l', meaning we can measure them to $l' = 10$ and compute edge-corrected ζ_l up to $l = 9$, as discussed in SE15b.

It is necessary to use a large number of random points so that the Poisson error from them is much smaller than the Poisson error from the galaxy data. We do this by computing NNN with 1.5 times more randoms than data, weighting the randoms so that the sum of their weights is equal to that of the data. The calculation of NNN is repeated with 32 different sets of randoms, and the results then averaged to get our final NNN; this is equivalent to using $1.5 \times 32 = 48$-fold more randoms than data when computing the data-random pairs. As described in SE15b, this process optimizes the minimization of the Poisson error for a fixed amount of computational time, by avoiding an expensive excess of random-random pairs. When computing the NNN of mock catalogs, we use only 10 sets of randoms. This incurs slightly more Poisson error in the resulting covariance matrix, but is 3-fold faster computationally. We estimate RRR and the resulting $\mathcal{R}_0$ and $f_{l'}$ factors using simply one set of randoms, 1.5 times more numerous than the data. This is sufficient for accurate answers as these terms do not involve the difference of two nearly canceling ingredients as N does.

4   COMPRESSION

Using an idea first advanced in SE15a, we compress the radial coefficients $\zeta_l(r_1, r_2)$ by integrating over r_2 in a region fixed by r_1. This has two motivations.

First, the integration region for r_2 can be set so as to avoid the squeezed limit where $r_1 = r_2$ and the third triangle side, r_3, can become zero. In this squeezed limit, as $\theta_{12} \to 0$ the two galaxies separated by r_3 begin to interact with each other strongly via gravity and so perturbation theory will likely be invalid. Further, the 3PCF becomes large when one triangle side is small, so these squeezed, third-side-zero triangles actually dominate the isosceles $r_1 = r_2$ signal. By choosing an integration region in r_2 that constrains r_2 to be greater than some minimum and also well different from r_1, we can avoid this squeezed limit. The compression scheme is shown in Figure 1.

Second, compression lowers the dimension of the problem, reducing the dimension of the covariance matrix, easing parameter fitting, and permitting simpler visualization of the data, as a 1-D line plot versus r_1 rather than a 2-D color map of $\zeta_l(r_1, r_2)$ amplitudes versus r_1 and r_2. This framework enables more robust interpretation and analysis of the 3PCF.
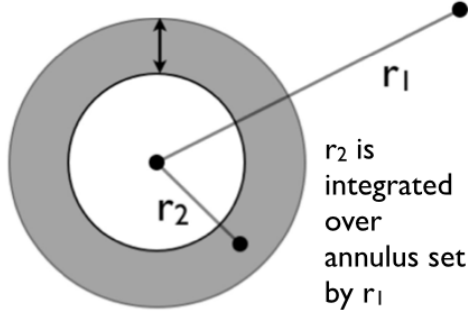
Figure 1. A visualization of the compression scheme; dots represent galaxies, and r_1 and r_2 are the triangle sides. The grey annulus contains allowed r_2 values given a particular r_1. In the present work, the inner radius of the annulus is fixed to be 30 Mpc/h and the outer edge $r_1 - 30$ Mpc/h.

However, we believe that the compression is not the optimal choice for extraction of BAO information. As SE15a shows, the full 3PCF in the multipole basis has structure where $r_3 = r_1 + r_2$ equals the BAO scale of $100\,\mathrm{Mpc}/h$. This structure gets averaged out in this compression. In the present work we present the compressed approach as a first step, in that it can validate the covariance matrix and the basic broadband agreement with perturbation theory. Even despite the non-optimality of the compression for this purpose, we find a BAO signal, and we will extend our fitting to the full two triangle side and multipole space soon.

We now present the explicit form of our compression scheme. The radial binning means the original compression scheme of SE15a, which took r_1 and r_2 as continuous variables, required modifications, described in SE15b. The compression used in that work and here is

$$\zeta_l^c (r_1) = \frac{\sum_{r_2 \in S(r_1)} \zeta_l(r_1, r_2) \Delta V(r_2)}{\sum_{r_2 \in S(r_1)} \Delta V(r_2)} \qquad (8)$$

where superscript c denotes compressed, $\Delta V(r_2)$ is the volume of bin r_2, and $S(r_1)$ is the set of all bins in r_2 where r_2 is greater than $3\Delta r$ and less than $r_1 - 3\Delta r$, with Δr the bin width. For our bin width of $\Delta r = 10\,\mathrm{Mpc}/h$, this prescription guarantees that $r_2 > 30\,\mathrm{Mpc}/h$ and $|r_2 - r_1| > 30\,\mathrm{Mpc}/h$, meaning by the Triangle Inequality that $r_3 > 30\,\mathrm{Mpc}/h$. As an example, for $r_1 = 100\,\mathrm{Mpc}/h$, the r_2 range integrated over is $30\,\mathrm{Mpc}/h - 70\,\mathrm{Mpc}/h$, or four r_2 bins given our Δr. This compression scheme means that for $r_1 < 70\,\mathrm{Mpc}/h$, there are no allowed r_2 bins, so it places a minimum on r_1 as well.

Overall, the compression guarantees that all the galaxies in the triangle are well-separated so PT should describe them well. The framework of compression and its sensitivity to linear, non-linear, and baryon-dark matter relative velocity bias are further discussed in SE15a. SE15b shows how this compression scheme translates to the covariance matrix, which we will also briefly summarize when we turn to the covariance matrix in §6.

5 MODEL

5.1 Standard perturbation theory

We use standard perturbation theory (SPT) to compute the 3PCF at leading (fourth) order in the linear density field δ. Summarizing the calculation performed in SE15a, we begin with the real-space local (Fry & Gaztañaga 1993) Eulerian galaxy bias model

$$\delta_{\mathrm{g}}(\vec{r}) = b_1 \delta_{\mathrm{m}}(\vec{r}) + b_2 \left[\delta_{\mathrm{m}}^2(\vec{r}) - \langle \delta_{\mathrm{m}}^2 \rangle \right] \qquad (9)$$

where δ_{g} is the galaxy overdensity at $\vec{r}$, b_1 is the linear bias, and b_2 is the non-linear bias. We do not include a bias term coupling to the tidal tensor here. δ_{m} is the matter density, given in terms of the linear density field as

$$\delta_{\mathrm{m}}(\vec{r}) = \delta(\vec{r}) + \delta^{(2)}(\vec{r}) \qquad (10)$$

with $\delta^{(2)}$ the usual second-order density field, most often given in Fourier space as

$$\tilde{\delta}^{(2)}\left(\vec{k}\right) =$$
$$\int d^3\vec{k}_1 d^3\vec{k}_2 \; \delta_{\mathrm{D}}^{[3]}(\vec{k}_1 + \vec{k}_2 - \vec{k}) \tilde{F}^{(2)}(\vec{k}_1, \vec{k}_2) \tilde{\delta}(\vec{k}_1) \tilde{\delta}(\vec{k}_2) \quad (11)$$

with $\delta_{\mathrm{D}}^{[3]}$ a 3-D Dirac delta function and $\tilde{F}^{(2)}$ the second-order kernel

$$\tilde{F}^{(2)}(\vec{k}_1, \vec{k}_2)$$
$$= \frac{17}{21} P_0(\hat{k}_1 \cdot \hat{k}_2) + \frac{1}{2}\left(\frac{k_1}{k_2} + \frac{k_2}{k_1}\right) P_1(\hat{k}_1 \cdot \hat{k}_2) + \frac{4}{21} P_2(\hat{k}_1 \cdot \hat{k}_2)$$
$$(12)$$

(Bernardeau et al. 2002 equation (45)).

We calculate the 3PCF to fourth order using the bias model (9) and focus on the pre-cyclic terms. This means that we choose one vertex of the triangle of galaxies at which to define θ_{12} and the two sides enclosing it, r_1 and r_2. Note that in the product of three copies of δ_g needed to form the 3PCF, each of the three galaxies can contribute a δ^2 and a $\delta^{(2)}$. The pre-cyclic term is written by choosing one galaxy to contribute each of these (it may be the same one). We have chosen the third galaxy to contribute these more complicated terms. Since we can then take this galaxy to be at the origin, this approach simplifies the calculation. However, to connect with observations, where there is no "preferred" vertex (galaxy), we eventually must sum cyclically, giving each galaxy in the survey the chance to contribute δ^2 and $\delta^{(2)}$.

Denoting "pre-cyclic" with a subscript pc, we find

$$\zeta_{\mathrm{pc}}(r_1, r_2; \hat{r}_1 \cdot \hat{r}_2) = \sum_{l=0}^{2} \zeta_{\mathrm{pc}l}(r_1, r_2) P_l(\hat{r}_1 \cdot \hat{r}_2), \qquad (13)$$

with the coefficients as

$$\zeta_{\mathrm{pc}0}(r_1, r_2) = \left[2b_1^2 b_2 + \frac{34}{21}b_1^3\right]\xi_1 \xi_2, \qquad (14)$$

$$\zeta_{\mathrm{pc}1}(r_1, r_2) = -b_1^3 \left[\xi_1^{[1-]}\xi_2^{[1+]} + \xi_2^{[1-]}\xi_1^{[1+]}\right] \qquad (15)$$

and

$$\zeta_{\mathrm{pc}2}(r_1, r_2) = \frac{8}{21}b_1^3 \xi_1^{[2]}\xi_2^{[2]}. \qquad (16)$$

Subscripts indicate sides r_1 or r_2 as a function's argument,

and ξ is the linear theory matter correlation function; $\xi^{[2]}$ and $\xi^{[1\pm]}$ are defined as:

$$\xi^{[1\pm]}(r) = \int_0^\infty \frac{dk}{2\pi^2} k^2 j_1(kr) P(k) k^{\pm 1} \qquad (17)$$

and

$$\xi^{[2]}(r) = \int_0^\infty \frac{dk}{2\pi^2} k^2 j_2(kr) P(k), \qquad (18)$$

where $P(k)$ is the matter power spectrum.

We will consider two different matter power spectra: the physical power spectrum, denoted $P_{\rm phys}$, and a "no-wiggle" power spectrum without BAO, denoted $P_{\rm nw}$. This no-wiggle power spectrum is computed using Eisenstein & Hu (1998)'s fitting formula for the no-wiggle transfer function, while the physical power spectrum is computed, following Eisenstein, Seo & White (2007) and Anderson et al. (2012), as

$$P_{\rm phys}(k) = [P(k) - P_{\rm nw}(k)] \exp[-k^2 \Sigma_{\rm nl}^2/2] + P_{\rm nw}(k), \quad (19)$$

where $\Sigma_{\rm nl} = 8$ Mpc/h is a non-linear smoothing scale to represent non-linear structure formation and RSD.

We pause to note an important feature of equation (17) regarding the BAO. The additional factor of k in $\xi^{[1+]}$ amplifies the sharp, localized real-space BAO bump encoded in the power spectrum (SE15a). This factor of k stems from the $\ell = 1$ term in the $\tilde{F}^{(2)}$ kernel (12), which in turn is generated primarily ($\sim 70\%$) by density gradients parallel to the velocity (SE15b; Bernardeau et al. 2002). The density Green's function (response to an initial Dirac-delta function overdensity at high redshift) has a sharp BAO bump (Slepian & Eisenstein 2015d), and so its divergence goes from positive to negative with a zero crossing at the BAO scale. Multiplying by the velocity amplifies structure at and beyond the BAO scale relative to that on much smaller scales because the velocity is larger at and beyond the BAO scale (SE15a). The remaining 30% of the $\ell = 1$ term in $\tilde{F}^{(2)}$, which also enters $\xi^{[1+]}$, stems from gradients of the velocity divergence parallel to the velocity, which also contain BAO information. Mathematically, the weighting by k in equations (12) and (17) occurs because the gradient operator's Fourier Transform scales as k. As §7 will show, these points lead to an order-unity BAO feature in the weighted, compressed 3PCF's dipole ($\ell = 1$).

After cyclically summing, we re-project onto the basis of Legendre polynomials to find the full radial coefficients in our multipole expansion (1) of the 3PCF as

$$\zeta_l(r_1, r_2) = \frac{2l+1}{2} \int_{-1}^1 d\mu_{12} \Big[\zeta_{\rm pc}(r_1, r_2, \mu_{12}) \qquad (20)$$
$$+ \zeta_{\rm pc}(r_2, r_3, \mu_{23}) + \zeta_{\rm pc}(r_3, r_1, \mu_{31}) \Big] P_l(\mu_{12}),$$

with $\mu_{12} \equiv \cos\theta_{12}$ and analogously for μ_{23} and μ_{31}. Note that r_3, μ_{23} and μ_{31} are all functions of r_1, r_2, and μ_{12}, easily found using the law of cosines. The factor of $(2l+1)/2$ is necessary because $\int_{-1}^1 P_m(\mu) P_n(\mu) d\mu = 2/(2n+1) \delta_{mn}^{\rm K}$; that is, the Legendre polynomials are an orthogonal but not orthonormal basis. Where in the pre-cyclic terms, we only found terms up to $l = 2$, cyclically summing introduces higher orders. These coefficients are shown versus r_1 and r_2 in SE15a Figure 9 split out by bias coefficient; since we set $b_v \equiv 0$, only the first two columns of this three-column figure are relevant to the present analysis. Also note that this Figure is in Mpc whereas the present work uses Mpc/h.

The power spectrum used in these predictions is computed from the matter transfer function $T_{\rm m}$ given by CAMB (Lewis 2000). We have $P = Ak^{n_s} T_{\rm m}^2$, with A a normalization set by matching σ_8 and n_s the scalar spectral tilt. The transfer function is on a grid equally spaced in $\log k$ with 5000 divisions per decade from $k = 3.46 \times 10^{-6}$ to 1.87 h/Mpc. To avoid ringing in the transforms of equations (17) and (18) we multiply by a Gaussian smoothing in Fourier space $\exp[-k^2]$, which suppresses structure below 1 Mpc/h. The predictions are not sensitive to the exact value of the smoothing scale.

We use a geometrically flat ΛCDM cosmology with parameters matching those used for the MULTIDARK-PATCHY mock catalogs (Kitaura et al. 2015): $\Omega_{\rm b} = 0.048$, $\Omega_{\rm m} = 0.307115$, $h \equiv H_0/(100$ km/s/Mpc$) = 0.6777$, $n_s = 0.9611$, $\sigma_8(z = 0) = 0.8288$, $T_{\rm CMB} = 2.7255$ K. These parameters do not differ substantially from the Planck values (Planck Paper XIII, 2015). We rescale σ_8 by the ratio of the linear growth factor at the survey redshift to the linear growth factor at redshift zero. We take the survey redshift to be $z_{\rm survey} = 0.565$ by simply averaging 0.43 and 0.7; since the redshift distribution of objects is roughly symmetric about the middle of the redshift interval this is a good approximation.

Broadly, it is worth emphasizing that the PT model we use is fourth order in the linear density field: this is the lowest order at which the 3PCF is non-zero. It does not incorporate any modeling of spherical collapse and virialization. However, on the large scales ($\gtrsim 30$ Mpc/h) we present here, it is likely that such effects are not important, especially because our compression scheme (§4) prevents any triangle side from becoming smaller than 30 Mpc/h.

5.2 Redshift-space distortions

Our model does not include any treatment of the effects of Redshift-Space Distortions (RSD). In principle, RSD can generate additional angular and side-length-dependent structure in the 3PCF. In practice, however, we believe that in our compressed basis the dominant effect of RSD is simply to rescale the multipole moments by some overall factor that depends only mildly on side length and multipole (SE15b). Consequently the no-RSD PT model presented above can fit the data well, but the bias parameters appearing in it are actually redshift-space quantities rather than the usual real-space biases.

In SE15b, we measured the real-space and redshift-space compressed 3PCF's multipole moments on scales from $45-90$ Mpc/h for the LasDamas SDSS DR7 mock catalogs. SE15b Figure 12 shows that RSD rescale the 3PCF by a factor of $\sim 1.5 - 1.8$ roughly independent of side length and multipole. Previous analytic work on the bispectrum by Scoccimarro, Couchman & Frieman (1999; hereafter SCF99) has shown that the rescaling is mostly set by $\beta = f/b_1$, where $f = d\ln D/d\ln a \approx \Omega_{\rm m}^{0.55}$ is the logarithmic derivative of the linear growth rate D with respect to scale factor a and b_1 is the real-space linear bias. For the LasDamas mocks we found $\beta = 0.33$; using the real-space linear bias $b_1 = 2.086$ from the CMASS power spectrum (Gil-Marín et al. 2015), $\beta = 0.37$ for the present work. Given the similarity of the CMASS and LasDamas values of β, we expect that the impact of RSD on the CMASS compressed 3PCF is comparable to that shown in SE15b Figure 12.

Importantly, since at fixed $\gamma \equiv 2b_2/b_1$ the 3PCF is proportional to b_1^3, any overall rescaling of the 3PCF due to RSD will only enter the measured redshift-space b_1 as a cube root.[2] Thus rescaling factors in the range $1.5 - 2$ will only change the implied real-space b_1 by $14 - 26\%$.

In detail, the rescaling due to RSD does also depend explicitly on b_1. The real-space linear bias $b_1 = 1.90$ for LasDamas, comparable to the value $b_1 = 2.086$ Gil-Marín et al. (2015) find for CMASS, so the b_1-dependence is unlikely to make the effect of RSD different between the two. The rescaling also depends on γ; for LasDamas we found $\gamma = 0.99$, while using the Gil-Marín et al. (2015) value for b_2 from the power spectrum we have $\gamma = 0.43$ for CMASS. While these values differ, the dependence on γ is weak, $\sim 3\%$ in $\ell = 0$ and $\ll 1\%$ for $\ell = 1$ and 2 for a factor of 2 change in b_2. Thus the empirical result that RSD rescale the LasDamas compressed 3PCF roughly independently of side length and multipole is likely to hold for the CMASS compressed 3PCF as well.

We now offer physical intuition as to why this rescaling is roughly a constant. One can show analytically that the rescaling is independent of side length at $\mathcal{O}(\beta)$ and nearly so at $\mathcal{O}(\beta^2)$. Over a realistic range of b_1 and b_2 the variation with multipole is also modest. Given that $\beta^3 \sim 3\%$, higher-order corrections might be expected to be unimportant.

In more detail, at $\mathcal{O}(\beta)$, simplifying SCF99 we find the following rescaling factors:

$$\frac{\zeta_{0,\mathrm{s}}}{\zeta_0} = 1 + \frac{94 + 42(b_1 + \gamma)}{102 + 63\gamma}\beta,$$

$$\frac{\zeta_{1,\mathrm{s}}}{\zeta_1} = 1 + \left(1 + \frac{b_1}{3}\right)\beta,$$

$$\frac{\zeta_{2,\mathrm{s}}}{\zeta_2} = 1 + \frac{4}{3}\beta. \tag{21}$$

Subscript s denotes redshift-space. At $\mathcal{O}(\beta^2)$, simple analytic formulae are not available but numerically taking the multipole moments of the SCF99 expression for the bispectrum (averaged over all orientations with respect to the line of sight) shows that dependence on the magnitudes of k_1 and k_2 is negligible. We have used these $\mathcal{O}(\beta^2)$ results to estimate the rescaling expected for the LasDamas compressed 3PCF. For $\ell = 0, 1$, and 2, the empirical rescalings from SE15b are 1.50, 1.71 and 1.72. Our $\mathcal{O}(\beta^2)$ results using SCF99 are 1.52, 1.70 and 1.80, which differ from the SE15b values by respectively 1.2%, -0.59%, and 4.7%. This agreement is encouraging and suggests that our approach could be used to relate redshift-space and real-space measurements to reasonable accuracy. For the CMASS sample, we estimate the rescalings are 1.74, 1.87, and 1.99 for $\ell = 0$, 1 and 2. While we expect these factors are accurate at the $\sim 5\%$ level, we defer detailed modeling of RSD to future work. Here we do not use these estimated rescalings to recover real-space biases from our redshift-space fits, but simply report the redshift-space best fit biases for the model equation (9).

We now summarize the approaches to RSD of other recent 3PCF and bispectrum works. For the bispectrum we focus on Gil-Marín et al. (2015), the most recent work on

the bispectrum and the one using the largest sample to date. This work uses the full bispectrum model of SCF99 with an additional overall damping factor to describe fingers-of-God (Jackson 1972) from velocity dispersions inside virialized structures. The SPT kernels of SCF99 are also modified to "effective" kernels, though this is to better describe the smaller-scale, weakly non-linear regime rather than to address RSD. Gil-Marín et al. (2015) also add a non-local bias term, which effectively rescales the P_2 and P_0 terms in the $\tilde{F}^{(2)}$ kernel.

For the 3PCF, we focus on Gaztañaga et al. (2009), which measures the 3PCF on BAO scales, and on McBride et al. (2011a,b), the largest sample used for the 3PCF previous to the present work. Gaztañaga et al. (2009) measure the reduced 3PCF, which is the 3PCF divided by the hierarchical ansatz that $\zeta(r_1, r_2, r_3) \sim \xi(r_1)\xi(r_2) + \mathrm{cyc.}$, with ξ the linear correlation function (Groth & Peebles 1977). They note that on large scales RSD cancel out of the reduced 3PCF, as shown in their Figure 2, so they do not incorporate additional modeling for RSD. The bias model of McBride et al. (2011a,b) has the reduced 3PCF of galaxies as a constant rescaling of the dark matter-only 3PCF. To remove the effects of RSD from their bias measurement, McBride et al. (2011a,b) apply a shift to the dark matter particles in the Hubble Volume simulations used for the dark matter-only 3PCF. As McBride et al. (2011a) Figure 8 shows, this is accurate save on the smallest scales they measure ($\lesssim 9\,\mathrm{Mpc}/h$).

Finally, the MULTIDARK-PATCHY mocks used for parts of the present analysis have been tuned to match small-scale quasi-virialized motions and the concomitant fingers-of-God, but for triangle opening angles $\theta = 0$ and π do not reproduce the observed 3PCF well (Kitaura et al. 2015; Figure 6). However this mismatch is for a triangle with r_1, $r_2 = 10$, $20\,\mathrm{Mpc}/h$, and on the larger scales in the present analysis we expect that fingers-of-God will not be an important effect. In particular, the compression scheme guarantees that all triangle sides are greater than $30\,\mathrm{Mpc}/h$, and further averages two triangle sides (r_2 and r_3) over rather large bins, additionally suppressing any finger-of-God effect.

5.3 Integral Constraint Model

We consider whether the randoms used for edge-correction may have been incorrectly normalized and hence fail to satisfy the integral constraint, that the mean number density of randoms matches the true number density of galaxies. Such a misnormalization can occur because we only know the measured number density of galaxies, which may be different from the true number density were one to perform the survey over an infinite volume.

We incorporate a possible failure of the integral constraint into the PT model fit to the data by adding a small fractional shift c to the randoms, taking our 3PCF estimator from

$$\zeta = \frac{(D - R)^3}{RRR} \rightarrow \frac{[D - R(1 + c)]^3}{RRR(1 + c)^3}. \tag{22}$$

Expanding this out and simplifying we find

$$\zeta^{\mathrm{M}} = \frac{\zeta^{\mathrm{T}} - c\xi_{\mathrm{cyc}} - c^3}{(1 + c)^3}, \tag{23}$$

---

[2] The term in $b_1^2 b_2$ of equation (14) can be rewritten as $b_1^3\gamma$. γ is defined following SCF99, but their real-space b_2 is defined such that it is twice the real-space b_2 of the present work.

where $\zeta^{\rm M}$ is the measured 3PCF and $\zeta^{\rm T}$ is the true 3PCF. $\xi_{\rm cyc}$ is the cyclic sum of the linear theory 2PCF around the three triangle sides,

$$\xi_{\rm cyc}(r_1, r_2, r_3) \equiv \xi(r_1) + \xi(r_2) + \xi(r_3). \qquad (24)$$

The 2PCF fits have often included similar nuisance polynomials, including a constant in ξ (e.g. Anderson et al. 2012). Including this constant in the BAO fits helps make them robust to observational systematics (Osumi et al. 2015; Ross et al. 2012; Ross et al. 2015). This term also is a good approximation of anomalous systematic power at small wavenumber.

6   COVARIANCE MATRIX

6.1   A Hybrid Approach

The covariance matrix $\mathbf{C}$ measures the independence of different multipole and radial bin combinations from each other. In the limit of fully independent measurements $\mathbf{C}$ is diagonal. The χ^2 for a parameter fit is calculated via $\vec{v}\mathbf{C}^{-1}\vec{v}^T$, where $\vec{v}$ is a vector of the data points minus the model being fit. Minimizing the χ^2 according to this prescription is guaranteed to give the optimal parameter fit if the likelihood is Gaussian.

However, this approach requires that the covariance matrix be invertible. Consequently if the covariance matrix is determined empirically by measuring the 3PCF of mock catalogs, a large number of mocks for each matrix element is desirable (Percival et al. 2014). Given the large dimension of the covariance matrix this is computationally costly, and so we have adopted an alternative approach. We compute the covariance matrix analytically and treat the effective survey volume and number density as free parameters to be fit by matching the empirical covariance derived from mock catalogs. This empirical covariance matrix, while likely too noisy to invert, is sufficient to constrain the volume and shot noise and reveal whether our analytic calculation describes the survey well.

Assuming that the density field is purely Gaussian and that the survey is boundary-free, the covariance of our multipole decomposition of the 3PCF can be analytically computed as (SE15b)

$$C_{\rm GRF, ll'}(r_1, r_2; r_1', r_2') = \frac{4\pi}{V}(2l+1)(2l'+1)(-1)^{l+l'}$$

$$\times \int r^2 dr \sum_{l_2}(2l_2+1)\begin{pmatrix} l & l' & l_2 \\ 0 & 0 & 0 \end{pmatrix}^2$$

$$\times \left\{ (-1)^{l_2}\xi(r)\Big[f_{l_2 ll'}(r; r_1, r_1')f_{l_2 ll'}(r; r_2, r_2') \right.$$

$$+ f_{l_2 ll'}(r; r_2, r_1')f_{l_2 ll'}(r; r_1, r_2')\Big] + (-1)^{(l+l'+l_2)/2}$$

$$\times \Big[f_{ll}(r; r_1)f_{l'l'}(r; r_1')f_{l_2 ll'}(r; r_2, r_2')$$

$$+ f_{ll}(r; r_1)f_{l'l'}(r; r_2')f_{l_2 ll'}(r; r_2, r_1')$$

$$+ f_{ll}(r; r_2)f_{l'l'}(r; r_1')f_{l_2 ll'}(r; r_1, r_2')$$

$$\left. + f_{ll}(r; r_2)f_{l'l'}(r; r_2')f_{l_2 ll'}(r; r_1, r_1')\Big] \right\}, \qquad (25)$$

with

$$f_{ll}(r; r_1) = \int \frac{k^2 dk}{2\pi^2}\left[P(k) + \frac{1}{n} \right]j_l(kr_1)j_l(kr) \qquad (26)$$

and

$$f_{l_2 ll'}(r; r_1, r_1') = \int \frac{k^2 dk}{2\pi^2}\left[P(k) + \frac{1}{n} \right]j_l(kr_1)j_{l'}(kr_1')j_{l_2}(kr), \qquad (27)$$

where V is the effective survey volume, and n is the effective number density of the survey. The $1/n$ term added to the power spectrum above characterizes the shot noise contribution to the covariance matrix. The covariance as above can be easily binned radially and compressed by integrating out r_2 and r_2', as detailed in SE15b.

Note that linear perturbation theory is insufficient to obtain the 3PCF for a purely Gaussian Random Field, because as an odd moment the 3PCF vanishes. However, linear theory does give the covariance because it is a 6-point function and hence an even moment. An analogy here is to a distribution with zero mean (the expected signal, the 3PCF), but non-zero width (the covariance).

To fit for volume and shot noise, we compute the empirical covariance matrix with elements

$$C_{\rm mock, ll'}(r_1, r_2; r_1', r_2') =$$

$$\frac{1}{N_{\rm mock}-1}\sum_{i=1}^{N_{\rm mock}}\left[\zeta_{i,l}(r_1, r_2) - \langle\zeta_l(r_1, r_2)\rangle \right]$$

$$\times \left[\zeta_{i,l'}(r_1', r_2') - \langle\zeta_{l'}(r_1', r_2')\rangle \right] \qquad (28)$$

where $N_{\rm mock}$ is the number of mock catalogs, i denotes the mock catalog number, and angle brackets represent the mean over the mocks.

The metric used for fitting the volume and shot noise of our analytic calculation to the empirical covariance is given by Xu et al. (2012) equation (58). For a fixed shot noise n, the likelihood-maximizing volume $V_{\rm best}$ can be found analytically as

$$V_{\rm best} = \frac{d}{\langle\chi_1^2\rangle} \qquad (29)$$

where d is the dimension of the covariance matrix. $\langle\chi_1^2\rangle$ is the mean χ^2 of all the mock covariances with respect to the analytic covariance matrix with the given shot noise and with the volume set to unity. Explicitly,

$$\langle\chi_1^2\rangle = \sum_{i=0}^{N_{\rm mock}}\vec{x}\,\mathbf{C}_{\rm 1,GRF}^{-1}\,\vec{x}^T \qquad (30)$$

where $\mathbf{C}_{\rm 1,GRF}$ is the analytic covariance matrix with $V=1$ and a given shot noise and $\vec{x}$ is a vector of the differences between the mock multipoles $\zeta_l(r_1, r_2)$ and their mean taken over all mock catalogs. We find a best-fit volume of 1.92 $\rm Gpc^3/h^3$ and a best-fit number density of $n = 1.80 \times 10^{-4}\ h^3/\rm Gpc^3$. Figure 2 shows the reduced covariance matrix used in our parameter fitting, i.e. that given by equation (25) after compression and with the values of V and n quoted above. Away from the diagonal, the covariance is strongest for low multipoles, indicated by the higher amplitude in the lower left corner.
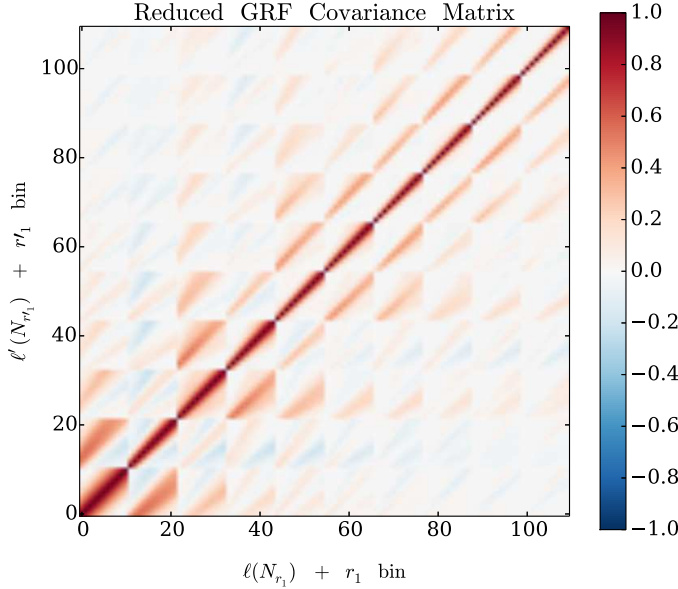
Figure 2. The reduced covariance matrix from our analytic calculation, $\mathbf{C}_{\mathrm{GRF},ll'}(r_1, r_1')/\sqrt{\mathbf{C}_{\mathrm{GRF},ll}(r_1, r_1)\mathbf{C}_{\mathrm{GRF},l'l'}(r_1', r_1')}$. We have mapped the 4-D compressed covariance tensor, which depends on r_1, r_1', l and l', to 2-D as indicated by the axis labels. Each small 11×11 tile corresponds to fixed l and l'; $N_{r_1} = N_{r_1'} = 11$ is the number of bins in side-length. The tile beginning at the origin represents the covariance of all radial bins at $l = l' = 0$. The two tiles neighboring the $l = l' = 0$ tile represent $l = 1$, $l' = 0$ and $l' = 1$, $l = 0$. The strong diagonal band within many of the tiles represents the covariance of radial bins where $r_1 = r_1'$.

6.2 Covariance Results and Tests

To assess how well our analytic covariance calculation fits the mock covariance, we compare the eigenvalues of the two matrices. We also compare the ratio of the eigenvectors at each eigenvalue. We find good agreement in both these tests. The agreement worsens somewhat as the eigenvalues become smaller, as we might expect because these will be more sensitive to the Poisson noise of estimating the covariance matrix using 299 mock catalogs rather than a larger number.

As a third test, we compute $\mathbf{C}_{\mathrm{GRF}}^{-1/2}\mathbf{C}_{\mathrm{mock}}\mathbf{C}_{\mathrm{GRF}}^{-1/2} - \mathbf{I}$, where $\mathbf{I}$ is the identity matrix. Note we apply a half-inverse on each side of the mock matrix purely to ensure our resulting test is symmetric. If the analytic and mock covariances are identical, this should be zero; given the noise of using a finite number of mocks in reality we expect the result to look like noise. This test is shown in Figure 3. The mean is 0.6% and the root mean square is 6%. This scatter about the (roughly zero) mean is attributable to having used a finite number of mocks, as with 110 degrees of freedom and 299 mocks we expect rms of $110/\sqrt{299} \approx 6\%$.

Figure 3 shows that our analytic calculation captures the true covariance as estimated from the mocks fairly well. We have windowed off the upper half of the test because the symmetry about the diagonal, present by construction, can misleadingly produce the appearance of more pattern by eye than is real. Note that we have mapped the 4-D compressed covariance matrix $\mathbf{C}_{ll'}(r_1, r_1')$ into a 2-D array using row-major ordering along the multipoles; $l = 0 = l'$ is in the
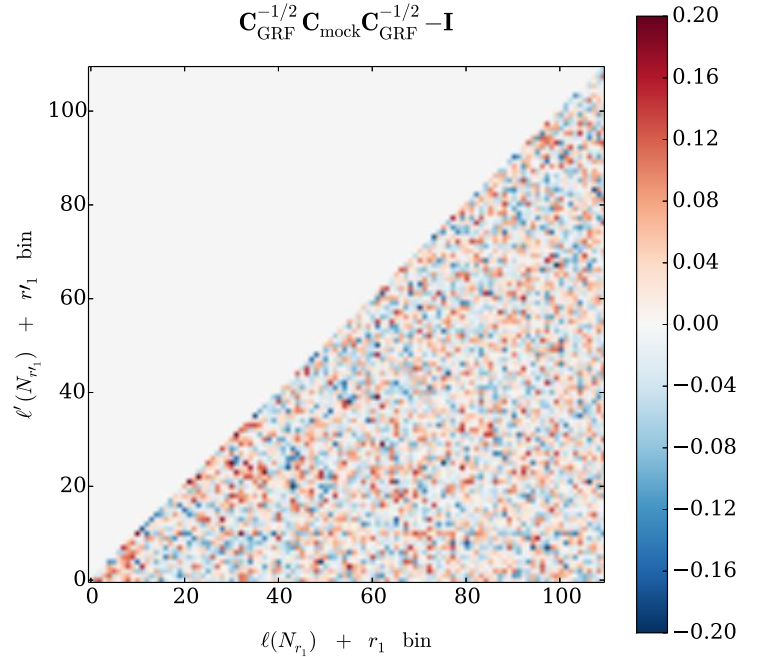
Figure 3. A test of the analytic covariance $\mathbf{C}_{\mathrm{GRF}}$ against that estimated from the mocks, $\mathbf{C}_{\mathrm{mocks}}$. The mean of this plot is 0.6% with root mean square 6%. The fact that the mean is close to zero shows that our analytic covariance matrix describes the true covariance well; 6% is the rms expected with 110 degrees of freedom and 299 mocks, as discussed in the main text. We have blocked out the upper half because the matrix is symmetric by construction and this can create the misleading appearance by eye of more pattern than there is in reality. The mapping of the 4-D compressed covariance tensor to 2-D is the same as in Figure 2.

lower left corner and the radial bins vary faster than the multipoles.

While there are 2048 MULTIDARK-PATCHY mocks extant, 299 mocks were sufficient for the present analysis. These mocks are only used to solve for two parameters (survey volume and shot noise), so the system is well-constrained algebraically. Note that even all 2048 mocks would be insufficient if one sought to determine the full covariance matrix empirically, as it has $110^2/2 = 6,050$ independent elements, many more than one per mock.

6.3 Details of Analytic Covariance Matrix Numerical Implementation

Here we briefly detail our numerical implementation of the compressed version of equations (25)-(27). As discussed in SE15b, the covariance can be efficiently compressed following the prescription of §4 by compressing the f tensors equations (26) and (27) first when their arguments include either r_2 or r_2' or both. However for some terms in the covariance the tensors only have r_1 and r_1' as arguments, meaning we still need the full uncompressed tensors. We form these tensors on a linearly spaced grid in k with 3000 points from $10^{-4}\ h/\mathrm{Mpc}$ to $2\ h/\mathrm{Mpc}$. These tensors also involve a dummy variable r which is integrated over at the end of the covariance calculation (see equation (25)); for this we use a linearly spaced grid with 4100 points from $10^{-5}\ \mathrm{Mpc}/h$ to $1000\ \mathrm{Mpc}/h$. A number of different endpoints for each

grid and number of points for each grid were tested and the results found to be well-converged for the choices above.

Binning adds an additional complication to the covariance calculation; as discussed in SE15b, one can analytically bin-average the spherical Bessel functions within the f tensors (equations (26) and (27)) to incorporate this. In practice we found that simply numerically averaging them over the required bins was more accurate. To obtain the f tensors, we use PYTHON's internal implementation of the spherical Bessel functions to produce lookup tables; these need only be computed once even if one wishes to test a variety of cosmological parameter values in the 3PCF predictions.

Finally, for the covariance matrix, we wish to match as closely as possible the observed redshift-space galaxy power spectrum. We do so by rescaling $kP(k)/(2\pi^2)$ to be $90\ h^{-2}\ \mathrm{Mpc}^2$ at its maximum to match the observed CMASS power spectrum (Anderson et al. 2014). This normalization incorporates redshift-space distortions and linear bias as they enter the power spectrum; empirically we find the required rescaling factor is 4.43.

7   RESULTS

7.1   Mock and Data Results

For both the CMASS data and the 299 MULTIDARK-PATCHY mocks, we used the algorithm of SE15b to measure the 3PCF in 18 radial bins of 10 Mpc/h width each, covering triangle sides from 0 to 180 Mpc/h. Our compression scheme means that there are no allowed r_2 bins for $r_1 < 70$ Mpc/h (§4), so after compression we have 11 bins for r_1 going from 70 Mpc/h to 180 Mpc/h. The run time was 2 hours per mock on a 36-core node, about 20k core-hours for the full computation. Computing the 3PCF of data and mocks was by far the dominant time cost of the calculation, with all other calculations taking negligible time. These include computing the covariance matrix and fitting for the best volume and shot noise, as well as fitting redshift-space linear and non-linear bias and the integral constraint parameter c to the data and mock results.

For all fits, we use 10 multipoles, from $l = 0$ to 9, and 11 radial bins at each multipole. These fits for the mean of the 299 MULTIDARK-PATCHY mocks are displayed in Figures 4 and 5. The error bars plotted are from the diagonal of the covariance matrix divided by $\sqrt{299}$ as is appropriate for a measurement of the mean of 299 mocks. We have multiplied the compressed 3PCF multipoles by $(r_1/10)^4$ to remove the large-scale fall-off of the 3PCF, which from the hierarchical ansatz we expect to scale like the square of the 2PCF, $\zeta \sim \xi^2 \propto r^{-4}$ (Groth & Peebles 1977). Notice the good match between the data and the PT. In particular, note the features at the BAO scale of $r_1 = 100$ Mpc/h: a significant peak and trough around the BAO scale in $\ell = 1$, and bumps in $\ell = 0$ and 4. As discussed in §5.1, we expect the BAO to be an order-unity fractional change in the dipole $\ell = 1$, and this is indeed borne out.

This fit shows that lowest-order PT in our compressed multipole basis appears to describe the 3PCF of these mocks fairly well. In particular, it supports our earlier claim that the primary effect of RSD in this compressed basis is a roughly multipole and side-length-independent rescaling of the 3PCF.

Nonetheless, the fit's $\chi^2/\mathrm{d.o.f.} = 405.53/107 = 3.79 > 1$, meaning that the PT is not a perfect match to these 299 mocks. We suspect this imperfect fit stems both from higher-order corrections to our leading-order PT model as well as possible scale or multipole-dependent effects of RSD. Nonetheless, this χ^2 for the mocks suggests that the model should be more than sufficient to fit the data well. Given that $\chi^2 \propto \mathbf{C}^{-1} \propto V$, we expect the model only begins to be insufficient for surveys with $V \sim 80\mathrm{x}$ that of SDSS DR12. Indeed, the PT model is an adequate fit to the data as we will soon show.

Table 1 gives our best-fit values of b_1, b_2, and c for the mean of the 299 mocks. Again note these biases are redshift-space quantities. The error bars quoted in Table 1 are from integrating over the probability distribution given by $P \propto \exp[-\chi^2/2]$ to compute $\sigma_{b_1} = \sqrt{\langle b_1^2 \rangle - \langle b_1 \rangle^2}$ from $\langle b_1 \rangle$ and $\langle b_1^2 \rangle$ and the same for b_2 and c. They are the error bars we might expect if we did the measurement on a volume 299 times that of the SDSS DR12 CMASS sample. To verify our analysis procedure, we also examined the probability contours in $b_1 - b_2$ space having marginalized over c, checking that the contours were consistent with our quoted error bars on b_1 and b_2. We do not show this plot here though we will display it for the CMASS data.

Finally, Rodríguez-Torres et al. (2015; see Figure 14) measure the linear bias for a set of 100 MULTIDARK-PATCHY mocks using the 2PCF on scales up to 60 Mpc/h. That work's model involves only linear bias and uses the Kaiser (1987) formula to remove the effects of RSD. They find some scale-dependence to the bias; the largest scale they measure is that most directly comparable to the present work. At this scale, the real-space $b_1 \simeq 1.8$, comparable to our inferred real-space bias of $1.90 - 2.00$ if the rescaling factors of §5.2 are applied to the redshift-space mock bias $b_1 = 2.390$. The difference between our inferred real-space b_1 and Rodríguez-Torres et al.'s measured value should not be over-interpreted for several reasons. First, the RSD factors used to convert our redshift-space bias are approximate; second, the bias model of Rodríguez-Torres et al. (2015) includes only linear bias (ours includes non-linear bias as well), and third, they compute the bias by comparing the measured 2PCF to that from the dark matter halos, while we compare the measured 3PCF to the PT model of §5.1.

We fit the CMASS data in the same way as the mocks, with the results displayed Figures 6 and 7. The error bars plotted are again from the diagonal of the covariance matrix. Again notice the match between data and PT. Though the match looks visually worse for the CMASS data than for the mocks, in fact the $\chi^2/\mathrm{d.o.f.} = 107.64/107 = 1.01$, meaning the PT model is adequate for the dataset, whereas for the mocks it was not.

In the case of the mocks, one might have worried that, since they are based on 2LPT, by construction PT would fit them well. However, the fair match between data and model shows that on these large scales and in our compressed basis, leading-order PT describes the 3PCF with reasonable fidelity.

The best-fit parameters for the data are in Table 1. We measure the redshift-space linear bias b_1 with 2.60% precision at fixed σ_8. Further, in concert with the power spectrum or 2PCF, our work can be used to place a constraint on σ_8.
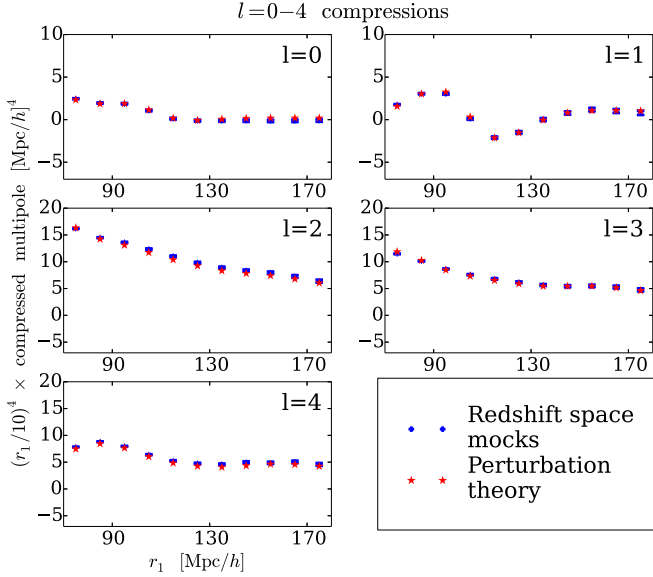
Figure 4. Fit of PT predictions, computed using the physical power spectrum, to the compressed 3PCF's multipoles $\ell = 0 - 4$ for the mean of 299 MULTIDARK-PATCHY mocks. The BAO scale is $r_1 = 100$ Mpc/h; note the pronounced peak and trough around this scale in $\ell = 1$ and the modest bumps in $\ell = 0$ and 4. The tiny error bars on the mocks' mean are shown; they are the covariance's diagonal scaled down as appropriate for the mean of 299 mocks, as further discussed in the main text.

Figure 5. Same as Figure 4 but now for multipoles $\ell = 5 - 9$. Notice the clear bump at the BAO scale of $r_1 = 100$ Mpc/h in both $\ell = 5$ and 6. Note the similarity of these higher ℓ panels to each other, and indeed to $\ell = 4$ in Figure 4. This similarity indicates there is less information in the higher multipoles, for reasons further discussed in §3 and S5.1. Essentially, prior to cyclic summing, the leading-order 3PCF has only $\ell = 0, 1$, and 2 terms, and so in our model higher multipoles stem from the purely geometric effect of cyclic summing. RSD can in principle introduce higher-multipole angular structure even at the pre-cyclic level, but it seems this effect is not substantial.

If the power spectrum, $P \propto b_1^2 \sigma_8^2$, and the redshift-space to real-space rescaling are known to high accuracy, then dividing $b_1^3 \sigma_8^4$ by $P^{3/2}$ isolates σ_8 with 7.8% precision (this approach was originally proposed in Fry 1994).

We obtain essentially no constraint on b_2. On these large scales, the 3PCF in our compressed basis seems to be insensitive to the redshift-space non-linear bias aside from indicating its existence. Figure 8 shows the probability of a given b_1 and b_2 having marginalized over the integral constraint (encoded in c; see §5.3). The elliptical appearance of the iso-probability contours means that b_1 and b_2 are roughly Gaussian-distributed. The ellipses drawn in red and light blue show 68% and 95% containment regions and do not assume Gaussianity; we simply integrate over the region until we reach these containments. The greater length of the ellipses in the b_2 direction illustrates that we do not obtain much constraint on b_2.

7.2 Searching for the BAO

To determine the significance of a BAO signal in our compressed multipole measurements of the 3PCF, we fit PT predictions for the 3PCF computed using the no-wiggle model to both mocks and data, and compute the $\Delta\chi^2$ relative to our fits of §7.1, which used the physical power spectrum equation (19). We emphasize that the BAO significance always stems from comparing the no-wiggle model to the physical power spectrum model. In the present work we do not fit for the BAO scale itself to extract distance information, but this will be a direction of future work.

For the mean of the MULTIDARK-PATCHY mocks, we find a clear preference for the BAO model. The comparison

is between Figures 4 (with BAO) and 9 (without BAO). The $\Delta\chi^2$ is 3234.34, meaning if we had a survey volume 299 times as large as CMASS we would expect a 56.9σ BAO detection even in our compressed 3PCF. The reason for the large χ^2 penalty of the no-BAO model over the model with BAO can be seen by visual comparison of Figures 4 and 9 around the BAO scale of $r_1 = 100$ Mpc/h, most prominently in $\ell = 1$ but also in $\ell = 0$ and 4.

For the CMASS data, we again find a preference for the BAO, with $\Delta\chi^2 = 7.58$, meaning a 2.8σ preference for the BAO. One can see comparing Figures 6 and 10 that both physical and no-wiggle models fit the data reasonably well, but that around the BAO scale of $r_1 = 100$ Mpc/h the no-wiggle model fits less well. Scaling the 56.9σ detection from the 299 mocks' mean down by $\sqrt{299}$ to mirror the volume of CMASS, we expect on average a 3.29σ detection of the BAO, indicating our result from the CMASS data is plausible. We expect the BAO feature's significance to increase when we use the full multipole coefficients of the 3PCF rather than the compressions presented here, a direction of future work.

Finally, b_2 changes from physical to no-wiggle models for both the mocks and the data. This change likely occurs because the no-wiggle power spectrum differs slightly from the physical power spectrum on small scales and b_2 changes to absorb this difference.
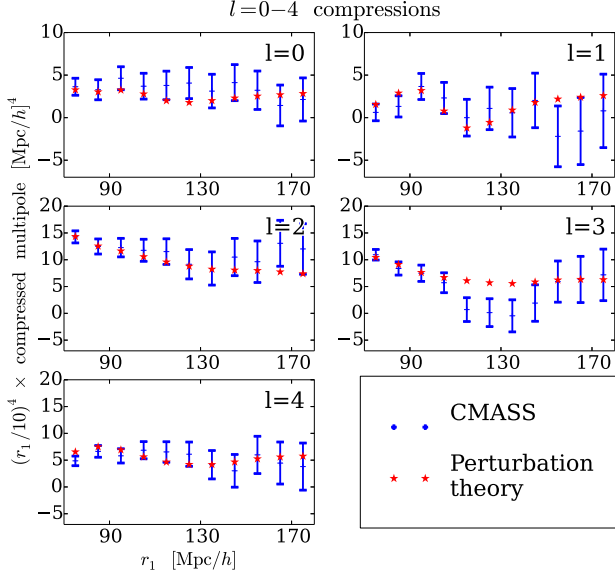
Figure 6. A fit of PT predictions, computed using the physical power spectrum, to the compressed 3PCF's multipoles $\ell = 0 - 4$ for the CMASS sample. Notice the peaks in $l = 0, 1$, and 4 around the BAO scale of $r_1 = 100$ Mpc/h. In particular compare the $\ell = 1$ panel here with that of Figure 4 to aid in identifying the peak and trough the BAO induce in the 3PCF's dipole moment. The points in the peak are anti-correlated with those in the trough, as shown in Figure 2 (second tile on the diagonal). These points are therefore more constraining than the error bars shown would suggest. The error bars plotted are the diagonal of the covariance matrix, and the $\chi^2/$d.o.f. $= 107.64/107$.
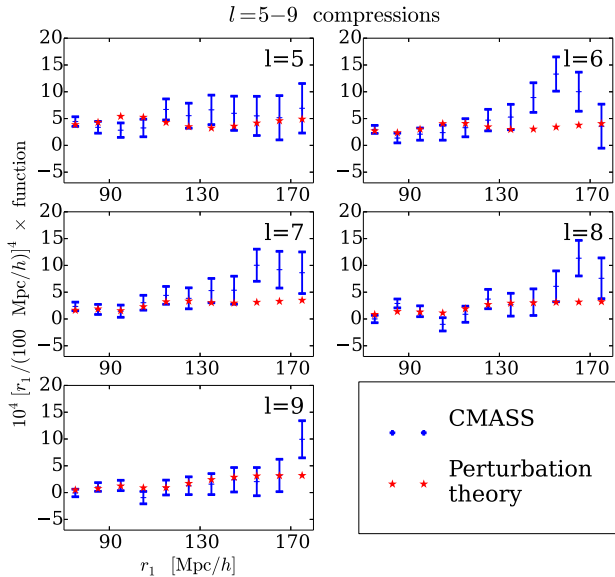
Figure 7. Same as Figure 6 but now for multipoles $\ell = 5 - 9$. These higher multipoles appear noisier than their lower-ℓ counterparts, as indicated by the larger number of points more than 1σ distant from the model. While the error bars are similar in magnitude to those in Figure 6, the signal is reduced relative to the largest in Figure 6 (i.e. $\ell = 2$ and 3).
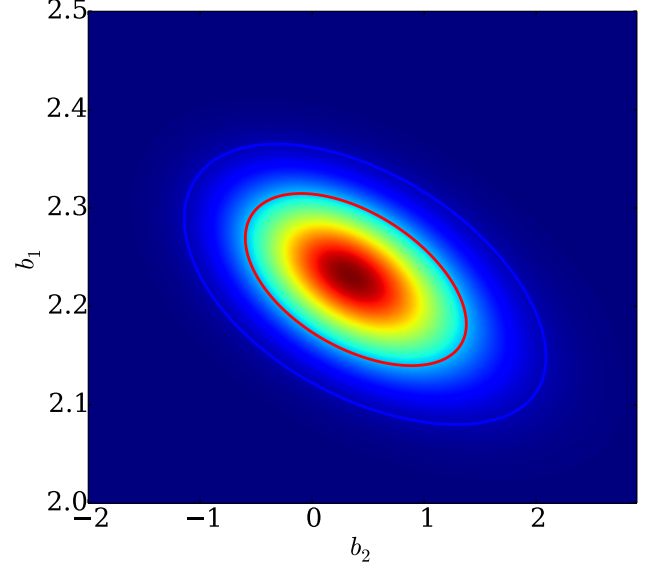
Figure 8. The probability contours for the redshift-space biases b_1 and b_2 having marginalized over the integral constraint. The red ellipse contains 68% of the probability and the light blue 95%. One can see that our measurement obtains a good constraint on b_1 but has very little constraining power on b_2, a conclusion borne out quantitatively by the large error bar on b_2 relative to that on b_1 quoted in Table 1.

	b_1	b_2	c	$\Delta\chi^2$ no BAO vs. BAO
MULTIDARK-PATCHY mocks	2.390 ±0.003	0.32 ±0.04	0.0000 ±0.0006	3234.34
CMASS SDSS DR12 sample	2.23 ±0.06	0.3 ±0.7	−0.023 ±0.007	7.58

Table 1. Table of best-fit parameters for MULTIDARK-PATCHY mocks and CMASS data. The biases are redshift-space quantities, and c encodes the integral constraint (§5.3). The last column describes the χ^2 penalty a no-BAO model (§7.2) pays over a model with BAO.

8   CONCLUSIONS

We have measured the large-scale 3PCF of the SDSS DR12 CMASS sample of 777,202 LRGs. The novel multipole algorithm of SE15b permits us for the first time to take advantage of all triangle configurations. We have used a compression scheme first developed in SE15a to reduce the dimension of the full 3PCF multipole coefficients and to avoid the triangles for which PT is likely invalid. We have shown that in this basis the analytic covariance matrix of SE15b, which assumes a Gaussian Random Field density and a boundary-free survey, is a good match to the empirical covariance matrix derived from 299 MULTIDARK-PATCHY mock catalogs. Using our analytic covariance matrix with volume and shot noise derived from the mocks, we have fit for the redshift-space linear and non-linear bias as well as a constant to marginalize over possible failure to satisfy the integral constraint. We measure the redshift-space linear bias with 2.6% precision. **We also find a 2.8σ preference for the BAO in the data** by comparing a physical model with BAO to the Eisenstein & Hu (1998) "no-wiggle" model.
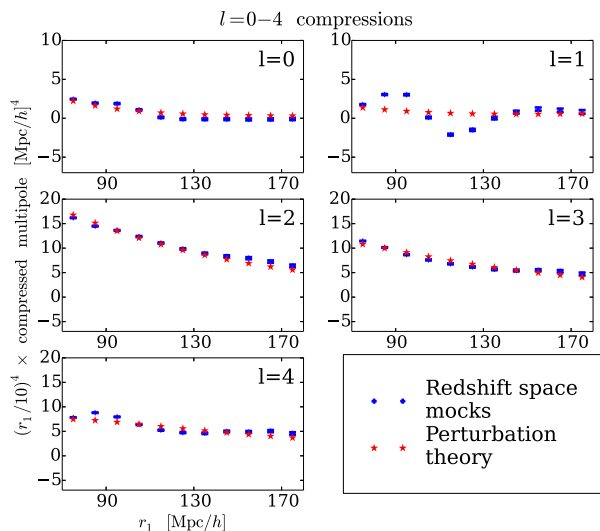
The present work uses the largest number of galaxies to

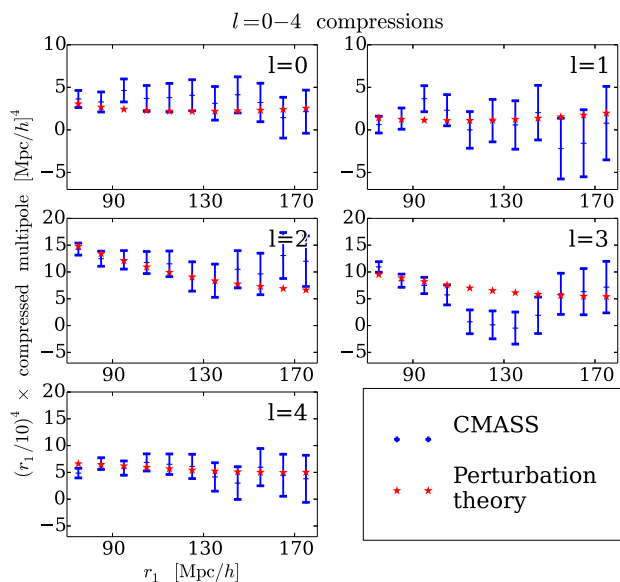**Figure 9.** A fit of PT predictions, computed using the no-wiggle power spectrum, to the compressed 3PCF's multipoles $\ell = 0 - 4$ for the mean of 299 MultiDark-Patchy mocks. In general the PT with no-wiggle is not a bad fit except around the BAO scale of $r_1 = 100$ Mpc/h; this is especially so in $\ell = 0, 1$, and 4. This indicates a clear preference for a BAO feature in the 3PCF. The best-fit parameter values are $b_1 = 2.362 \pm 0.003$, $b_2 = 0.16 \pm 0.03$, $c = -0.0020 \pm 0.0006$ with $\chi^2/\mathrm{d.o.f} = 3639.88/107$. The error bars plotted are the diagonal of the covariance matrix scaled down appropriately for the mean of 299 mocks.

Figure 10. A fit of PT predictions, computed using the no-wiggle power spectrum, to the compressed 3PCF's multipoles $\ell = 0 - 4$ for the CMASS sample. Notice the failure of the no-wiggle model especially in $l = 0, 1$, and 4 at the BAO scale of $r_1 = 100$ Mpc/h. The parameter values are $b_1 = 2.21 \pm 0.06$, $b_2 = 0.1 \pm 0.6$ and $c = -0.022 \pm 0.007$ and $\chi^2/\mathrm{d.o.f.} = 115.22/107$. The error bars plotted are the diagonal of the covariance matrix.

date for the 3PCF by a factor of roughly 4, and is unique in exploiting all triangle configurations. We also measure the 3PCF on significantly larger scales (roughly a factor of 2) than previous works. Our error bar on the redshift-space linear bias is competitive with that found on the real-space linear bias from recent bispectrum studies such as Gil-Marín et al. (2015), which measures the combination $b_1^{1.40}\sigma_8(z_{\mathrm{eff}})$ with 3.5% precision (their Table 2). If σ_8 is fixed, as we do in the present work, their measurement translates to 2.6% precision on b_1.

Gil-Marín et al. (2015) use $690,827$ LRGs in the CMASS sample of DR11 for their bispectrum, finding $b_1^{1.40}\sigma_8(z_{\mathrm{eff}}) = 1.672 \pm 0.060$. This work uses a detailed, complex model for removing RSD to obtain the real-space bias. Using our value of σ_8 to convert $b_1^{1.40}\sigma_8$ to b_1 translates this measurement to a real-space linear bias $b_1 = 2.03$. Gil-Marín et al. (2015) also measure b_1 and b_2 from the power spectrum, finding $b_1 = 2.086$ and $b_2 = 0.902$; they do not quote errorbars on the parameters derived from the power spectrum.

Our measured redshift-space b_1 is likely enhanced by $\sim 20 - 25\%$ relative to its real-space value based on the rescaling factors computed in §5.2. Reducing our value of 2.23 by these factors leads to an inferred real-space $b_1 \sim 1.77 - 1.85$. However, we caution that the rescaling factors used here are estimates and so there is some uncertainty in how our redshift-space bias maps to a real-space bias. The sample of Gil-Marín et al. (2015) is a proper subset of our own (though $\sim 90\%$), so one would not necessarily expect the bias values to be equal. Furthermore, the bias model of Gil-Marín et al. (2015) is Lagrangian and includes a tidal tensor bias, while our model is Eulerian and does not include this bias. Consequently any quantitative comparison should be taken with considerable caution at present.

From the projected 2PCF, Guo et al. (2013) infer a real-space linear bias of $b_1 = 2.16 \pm 0.01$ using subsamples from the CMASS sample of DR9 with a maximum of $260,000$ galaxies. Guo et al. (2013) uses projected separations of $3 - 25$ Mpc/h and a detailed model of non-linear structure formation. That work's cosmology is WMAP7, with $\sigma_8(z = 0) = 0.8$, whereas in the present work $\sigma_8(z = 0) = 0.8288$. Rescaling the Guo et al. (2013) measurement to our σ_8 yields a real-space linear bias of $b_1 = 2.1$. Note that the projected measurement is insensitive to RSD.

From the full 2PCF, Ross et al. (2014) uses $540,505$ CMASS galaxies in DR10 to measure the real-space linear bias, finding $b_1 = 1.96 \pm 0.05$. $\sigma_8(z = 0) = 0.8$ for this work, and rescaling as discussed above reduces this value to $b_1 = 1.89$, comparable to our expected real-space range of $b_1 \sim 1.77 - 1.85$. We caution that the galaxy samples used in Guo et al. (2013) and in Ross et al. (2014) are again proper subsets of our own (roughly 30% and 70%, respectively) so one would not necessarily expect the bias values to be equal.

The 2PCF and 3PCF depend on different combinations of b_1 and σ_8 (Fry 1994), so measuring the two in concert can permit isolating a value for each. In the present work, we have not pursued a combined fit to the 2PCF and 3PCF; future work might explore this approach. Such work would require the covariance between the 2PCF and the 3PCF, a topic explored for the Fourier-space analogs in Sefusatti et al. (2006), who find that squeezed triangles dominate this covariance. As our compression scheme avoids such trian-

gles, it may be beneficial in the context of joint 2PCF-3PCF fitting.

Regarding squeezed triangles, the position-dependent 2-point correlation function, developed in Chiang et al. (2015), likely offers interesting complementary information not captured by our compression. This technique evaluates the 2PCF within sub-volumes conditioned on the mean overdensity of those subvolumes, and is therefore sensitive to an integral of the 3PCF dominated by squeezed triangles. Chiang et al. (2015) have already applied the technique to the SDSS DR10 CMASS sample to measure the non-linear bias b_2. In future it would be worthwhile to explore joint parameter constraints using the position-dependent 2PCF and our compressed 3PCF.

We will also explore more complicated bias models, such as Lagrangian schemes (shown to be more accurate in Chan et al. 2012) that include tidal tensor bias (McDonald & Roy 2009; Baldauf et al. 2012). Adding a tidal tensor bias might modestly raise our error bars on the measured biases. However it is unlikely to alter the BAO significance because the tidal tensor enters at $\ell = 0$ and 2 pre-cyclically whereas the BAO's dominant contribution is at $\ell = 1$ (see §5). Additional modeling of RSD to permit precise extraction of the real-space bias values will also be important moving forward. It may also be worthwhile to further explore the impact of observational systematics on the 3PCF, for instance by re-analyzing the sample with systematic weights such as the stellar density weight turned off.

Finally, in future work we will extend our BAO search from the compressed statistic to the full multipole moments of the 3PCF. We developed the compressed approach to ease handling of the covariance matrix and fitting as well as to avoid regimes where PT is invalid, but it may be overly conservative. In practice we have found that the covariance matrix and fitting are computationally by far not the limiting steps in our analysis, and the good agreemeent between the compressed PT and our measurements argues that there is some room to incorporate more strongly squeezed triangles before PT breaks down.

Regarding the BAO, Gaztañaga et al. (2009)'s finding of a $2 - 3\sigma$ preference using a single triangle configuration measured from $40,000$ LRGs suggests that our measurement of the 3PCF might contain an even more significant BAO feature than we have found. Gaztañaga et al. (2009) did find a high value of the baryon fraction $f_b = 0.28$, significantly larger than the Planck-compatible value $f_b = 0.16$ used in the present work; this is likely related to the strength of their BAO feature. Since our work uses many triangle configurations but also compresses, the significance of their BAO evidence cannot be simply scaled up as the square root of survey volume or galaxy number. Nonetheless, this is an additional motivation for future work examining analysis variations in the context of the BAO.

REFERENCES

Abazajian K. N. et al., 2009, ApJS, 182, 543.

Aihara H. et al., 2011, ApJS, 193, 29.

Alam S. et al., 2015, preprint (arXiv:1501.00963).

Anderson et al. 2012, MNRAS 427, 4, 3435-3467.

Anderson et al. 2014, MNRAS 441, 1, 24-62.

Baldauf T., Seljak U., Desjacques V. & McDonald P., 2012, PRD 86, 8.

Bardeen J.M., Steinhardt P.J. & Turner M.S., 1983, PRD, 28, 4, 679-693.

Bernardeau F., Colombi S., Gaztañaga E., Scoccimarro R., 2002, Phys. Rep., 367, 1.

Blake C. & Glazebrook K., 2003, ApJ 594, 2, 665-673.

Blanton M. et al., 2003, AJ, 125, 2276.

Bolton A. et al., 2012, AJ, 144, 144.

Burden A., Percival W.J., Manera M., Cuesta A.J., Vargas-Magaña M. & Ho S., 2014, MNRAS, 445, 3, 3152-3168.

Chan K.C., Scoccimarro R. & Sheth R.K., 2012, PRD 85, 8, 083509.

Chiang C.-T., Wagner C., Sánchez A.G., Schmidt F. & Komatsu E., 2015, JCAP 09, 28.

Chuang C.-H. et al., 2015, MNRAS 452, 1, 686-700.

Cuesta A.J. et al., 2015, preprint (arXiv:1509.06371).

Dawson K.S. et al., 2013, AJ, 145, 10.

Dekel A. & Rees M.J., 1987, Nature 326:455-462.

Desjacques V. & Seljak U., 2010, Class. and Quant. Grav. 27, 12.

Doi M. et al., 2010, AJ, 139, 1628.

Eisenstein D.J. & Hu W., 1998, ApJ 496, 605.

Eisenstein D.J., Hu W. & Tegmark M., 1998, ApJ 504:L57-L60.

Eisenstein D.J., Seo H.-J., Sirko E. & Spergel D., 2007, ApJ 664, 2, 675-679.

Eisenstein D.J., Seo H.-J. & White M., 2007, ApJ 664, 2, 660-674.

Eisenstein D.J. et al., 2011, AJ, 142, 72.

Feldman H.A., Frieman J.A., Fry J.N. & Scoccimarro R., 2001, PRL 86, 1434.

Fry J.N. & Seldner M., 1982, ApJ, 259, 474.

Fry J.N., 1984, ApJ 279, 499-510.

Fry J.N. & Gaztañaga E., 1993, ApJ 413, 2, 447-452.

Fry J.N., 1994, PRL 73, 215.

Fukugita M. et al., 1996, AJ, 111, 1748.

Gaztañaga E. & Scoccimarro R., 2005, MNRAS, 361, 824.

Gaztañaga E., Cabré A., Castander F., Crocce M. & Fosalba P., 2009, MNRAS 399, 2, 801-811.

Giannantonio T. et al., 2014, PRD 89, 2, 023511.

Gil-Marín H., Noreña J., Verde L., Percival W.J., Wagner C., Manera M. & Schneider D.P., 2015, MNRAS 451, 1, 539-580.

Gunn J.E. et al., 1998, AJ, 116, 3040.

Gunn J.E. et al., 2006, AJ, 131, 2332.

Guo H. et al., 2013, ApJ 767, 2, 122.

Guo H. et al., 2014, ApJ 780, 2, 139.

Guo H. et al., 2015, MNRAS 449, 1, L95-L99.

Groth E. J. & Peebles P. J. E., 1977, 217, 385.

Hamilton A.J.S, 1998, in "The Evolving Universe: Selected Topics on Large-Scale Structure and on the Properties of Galaxies," Dordrecht: Kluwer.

Hu W. & Haiman Z., 2003, PRD 68, 6, 063004.

Jackson J.C., 1972, MNRAS 156, 1.

Jing Y.P. & Börner G., 2004, ApJ 607, 140.

Kaiser N. 1984, ApJL 284, L9-12.

Kaiser N., 1987, MNRAS 227, 1.

Kitaura F.-S. & Heß S., 2013, MNRAS 435, 1, L78-L82.

Kitaura F.-S., Yepes G. & Prada F., 2014, MNRAS 439, L21.

Kitaura F.-S., Gil-Marín H., Scóccola C.G., Chuang C.-H., Müller V., Yepes G. & Prada F., 2015a, MNRAS 450, 2, 1836-1845.

Kitaura F.-S. et al., 2015b, preprint (arXiv:1509.06400).

Kulkarni G., Nichol R., Sheth R., Seo H.-J., Eisenstein D.J. & Gray A., 2007, MNRAS 378, 3, 1196-1206.

Lewis A., 2000, ApJ, 538, 473.

Linder E.V., 2003, PRD 68, 8, 083504.

Lupton R., Gunn J.E., Ivezić Z., Knapp G. & Kent S., 2001, "Astronomical Data Analysis Software and Systems X", v. 238, 269.

Marín F.A., 2011, ApJ 737, 2, 97.

Marín F.A. et al., 2013, MNRAS 432, 4 2654.

McBride C., Connolly A.J., Gardner J.P., Scranton R., Newman J., Scoccimarro R., Zehavi I., Schneider D.P., 2011a, ApJ, 726, 13.

McBride K., Connolly A.J., Gardner J.P., Scranton R., Scoccimarro R., Berlind A., Marin F., Schneider D.P., 2011b, ApJ, 739, 85

McDonald P. & Roy A., 2009, JCAP 0908, 020.

Nichol R.C. et al., 2006, MNRAS 368, 4, 1507-1514.

Osumi K., Ho S., Eisenstein D.J. & Vargas-Magaña M., 2015, preprint (arXiv:1505.00782).

Padmanabhan N. et al., 2008, ApJ, 674, 1217.

Padmanabhan N., White M. & Cohn J.D., 2009, PRD 79, 6, 063523.

Pan J. & Szapudi I., 2005, MNRAS 362, 4, 1363-1370.

Peebles P.J.E. & Groth E.J., 1975, ApJ, 196, 1.

Peebles P.J.E., 1980, The Large-Scale Structure of the Universe: Princeton University Press, Princeton.

Percival W.J. et al., 2014, MNRAS 439, 3, 2531-2541.

Pier J.R. et al., 2003, AJ, 125, 1559.

Planck Collaboration, Paper XIII, 2015, preprint (arXiv:1502.01589).

Planck Collaboration, Paper XVII, 2015, preprint (arXiv:1502.01592).

Reid B. et al., 2016, MNRAS 455, 2, 1553-1573.

Rodríguez-Torres S. et al., 2015, preprint (arXiv:1509.06404).

Ross A.J. et al., 2012, MNRAS 424, 1, 564-590.

Ross A.J. et al., 2013, MNRAS 428, 2, 1116-1127.

Ross A.J. et al., 2014, MNRAS 437, 2, 1109-1126.

Ross A.J. et al., 2015, in prep.

Schmittfull M.M., Baldauf T. & Seljak U, 2015, PRD 91, 4, 043530.

Schmittfull M.M., Feng Y., Beutler F., Sherwin B. & Yat Chu, M., 2015, preprint (arXiv:1508.06972).

Schmittfull M.M., Regan D.M. & Shellard E.P.S., 2013, PRD, 88, 6, 063512.

Scoccimarro R., 2000, ApJ 544, 593.

Scoccimarro R., Couchman H.M.P. & Frieman J.A., 1999, ApJ 517:531-540.

Scoccimarro R., Feldman H., Fry J.N. & Frieman J.A., 2001, ApJ 546, 2, 652-664.

Scoccimarro R., 2015, preprint (arXiv:1506.02729).

Sefusatti E., Crocce M., Pueblas S. & Scoccimarro R., 2006, PRD 74, 2, 023522.

Sefusatti E. & Komatsu E., 2007, PRD 76, 8, 083004.

Seo H.J. & Eisenstein D.J., 2003. ApJ 598, 2, 720-740.

Slepian Z. & Eisenstein D.J., 2015a, MNRAS 448, 1, 9-26.

Slepian Z. & Eisenstein D.J., 2015b, MNRAS 454, 4, 4142-4158.

Slepian Z. & Eisenstein D.J., 2015c, MNRASL 455, 1, L31-L35.

Slepian Z. & Eisenstein D.J., 2015d, preprint (arXiv:1509.08199).

Smee S. et al., 2013, AJ, 146, 32.

Smith J.A. et al., 2002, AJ, 123, 2121.

Starobinsky A.A., 1982, Phys. Lett. B, 117, 3-4, 175-178.

Szapudi I., Szalay A., 1998, ApJ, 494, L41.

Szapudi I., 2004, ApJ, 605, L89.

Verde L. et al., 2002, MNRAS, 335, 432.

Wang Y., Yang X., Mo H.J., van den Bosch F.C. & Chu Y., 2004, MNRAS, 353, 287.

Xia J.-Q. , Baccigalupi C., Matarrese S., Verde L. & Viel M., 2011, JCAP 8, 33.

Xia J.-Q., Bonaldi A., Baccigalupi C., De Zotti G., Matarrese S., Verde L. & Viel M., 2010, JCAP 8, 13.

Xu X., Padmanabhan N., Eisenstein D.J., Mehta K.T. & Cuesta A.J., 2012, MNRAS 427, 3, 2146-2167.

York D.G. et al., 2000, AJ, 120, 1579.