

# Measurement of Parity-Odd Modes in the Large-Scale 4-Point Correlation Function of SDSS BOSS DR12 CMASS and LOWZ Galaxies

Jiamin Hou,<sup>1\*</sup> Zachary Slepian,<sup>1,2†</sup> & Robert N. Cahn<sup>2</sup>

<sup>1</sup>*Department of Astronomy, University of Florida, Gainesville, FL 32611, USA*

<sup>2</sup>*Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA*

9 June 2022

## ABSTRACT

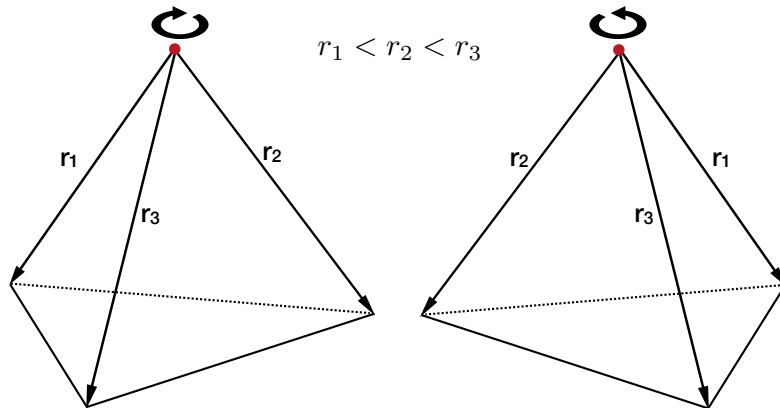
A tetrahedron is the simplest shape that cannot be rotated into its mirror image in 3D. The 4-Point Correlation Function (4PCF), which quantifies excess clustering of quartets of galaxies over random, is the lowest-order statistic sensitive to parity violation. Each galaxy defines one vertex of the tetrahedron. Parity-odd modes of the 4PCF probe an imbalance between tetrahedra and their mirror images. We measure these modes from the largest currently available spectroscopic samples, the 280,067 Luminous Red Galaxies (LRGs) of BOSS DR12 LOWZ ( $\bar{z} = 0.32$ ) and the 803,112 LRGS of BOSS DR12 CMASS ( $\bar{z} = 0.57$ ). In LOWZ we find  $3.1\sigma$  evidence for a non-zero parity-odd 4PCF, and in CMASS we detect a parity-odd 4PCF at  $7.1\sigma$ . Gravitational evolution alone does not produce this effect; parity-breaking in LSS, if cosmological in origin, must stem from the epoch of inflation. We have explored many sources of systematic error and found none that can produce a spurious parity-odd *signal* sufficient to explain our result. Underestimation of the *noise* could also lead to a spurious detection. Our reported significances presume that the mock catalogs used to calculate the covariance sufficiently capture the covariance of the true data. We have performed numerous tests to explore this issue. The odd-parity 4PCF opens a new avenue for probing new forces during the epoch of inflation with 3D LSS; such exploration is timely given large upcoming spectroscopic samples such as DESI and Euclid.

## 1 INTRODUCTION

The laws of nature respect certain symmetries; the physical processes governed by them are invariant under the corresponding transformations. Parity transformation (P), which reverses the sign of each coordinate axis, had been thought to be such a symmetry. Indeed, the electromagnetic and strong interactions are invariant under P. However, this symmetry is broken in the weak interaction (Lee & Yang 1956; Wu et al. 1957). Sakharov (1967) showed that the matter-antimatter asymmetry of the Universe requires that the combination CP of P and charge-conjugation (C) symmetry be broken. The currently-known CP violation is inadequate to explain the observed matter-antimatter asymmetry. Whatever additional CP violation is responsible may involve pure P violation as well.

Most of the cosmological studies of parity invariance to date have focused on Cosmic Microwave Background (CMB) polarization data (Lue et al. 1999; Kamionkowski & Souradeep 2011; Shiraishi et al. 2011; Minami & Komatsu 2020) or on gravitational waves (Jackiw & Pi 2003; Yunes et al. 2010; Wang et al. 2013; Zhu et al. 2013; Nishizawa & Kobayashi 2018; Orlando et al. 2021). A recent study on parity violation reported  $2.4\sigma$  evidence using cosmic birefringence (where the two polarization states of a wave propagate differently) (Minami & Komatsu 2020). This analysis was refined in Diego-Palazuelos et al. (2022), where a significance was ultimately not reported because of uncertainty in polarized foregrounds. Although helical primordial gravitational waves could leave an imprint on the cross spectrum between CMB polarization  $E$  and  $B$  modes or on the cross spectrum between temperature and  $B$  modes, these observables are suppressed by the two-dimensional nature of the CMB (Masui et al. 2017). Probes using three dimensions, such as the orientation of galaxy spins or odd modes (with respect to the line of sight) of the redshift-space clustering (Jeong & Kamionkowski 2012; Dai et al. 2013; Schmidt et al. 2014) have been proposed as alternatives.

Cosmological parity violation can be induced by a Chern-Simons-like interaction with a graviton (Alexander & Yunes 2009; Soda et al. 2011; Dyda et al. 2012) or a  $U(1)$  gauge field (Sorbo 2011; Barnaby et al. 2011; Özsoy 2021). In the latter case, the very same term can also give rise to non-vanishing parity-odd curvature perturbation polyspectra (Abolhasani et al. 2016; Bartolo et al. 2015; Shiraishi 2016). These perturbations would seed large-scale structure persisting to the late-time Universe.



**Figure 1.** Parity transformation applied to a tetrahedron formed by a quartet of galaxies. Each vertex represents a galaxy. Choosing one galaxy (red dot) as our primary, the quartet is defined by the three vectors to the remaining vertices,  $\mathbf{r}_1$ ,  $\mathbf{r}_2$ , and  $\mathbf{r}_3$ . For a quartet, the subscripts are fixed by requiring that  $r_1 \leq r_2 \leq r_3$ . When viewing the tetrahedron from the primary (red) looking down along each vector  $\mathbf{r}_i$ , the direction in which one reads going from smallest to largest side ( $r_1$  to  $r_3$ ) defines a handedness, either clockwise or counterclockwise. Here, the tetrahedron on the left, as viewed from the primary, is clockwise. Parity transformation in 3D is a reflection about a plane and then a  $180^\circ$  rotation about the vector perpendicular to that plane, and converts the clockwise tetrahedron at left to the counterclockwise one at right. When one averages over rotations, as in this work, only the mirroring matters.

Recently, [Cahn et al. \(2021\)](#) proposed using 3D Large-Scale Structure (LSS) as a probe of parity violation via the galaxy 4-Point Correlation Function (4PCF).

Four galaxies can be taken as the vertices of a tetrahedron, the lowest-order 3D shape that cannot be rotated into its mirror image, rendering the 4PCF sensitive to parity violation. An illustration of a galaxy quartet and how we define parity on it, is shown in Fig. 1. Following [Cahn et al. \(2021\)](#), we expand the 4PCF in the isotropic (*i.e.* rotation-averaged) basis functions of [Cahn & Slepian \(2020\)](#). In the standard inflationary paradigm ([Albrecht & Steinhardt 1982](#); [Linde 1982, 1983](#)), we would not expect a parity-odd 4PCF. The initial density fluctuations are a Gaussian random field ([Bernardeau et al. 2002](#)), which then evolves under gravity and forms galaxies at late times. Gravity, and even the baryonic physics of galaxy formation, is parity-conserving. Hence, detection of a parity-odd 4PCF of cosmological origin would be evidence that parity violation was present before the known forces dominated the evolution of the matter distribution.

We present here a measurement of the parity-odd modes of the 4PCF measured using the Baryon Oscillation Spectroscopic Survey (BOSS; [BOSS collaboration et al. 2017](#)) of Sloan Digital Sky Survey (SDSS)-III ([Dawson et al. 2013](#); [Eisenstein et al. 2011](#)). [Philcox et al. 2021b](#) presented the parity-even 4PCF measurement on the same dataset and found an  $8.1\sigma$  detection of a non-Gaussian 4PCF (expected in the standard picture of gravitationally evolved structure formation, *e.g.* [Bernardeau et al. 2002](#)). A progenitor of the algorithm and approach here used was employed to measure the 3-Point Correlation Function (3PCF) of BOSS DR12 CMASS ([Slepian et al. 2017a,b,c](#)), and extended to use Fourier transforms in [Slepian & Eisenstein \(2015c\)](#); [Portillo et al. \(2018\)](#) and to the anisotropic 3PCF in [Slepian & Eisenstein \(2018\)](#); [Friesen et al. \(2017\)](#); [Garcia & Slepian \(2020\)](#). The 4PCF has been measured before in just a few works (in 2D; [Fry & Peebles 1978](#); averaged over internal angles ([Sabiu et al. 2019](#)), as well as in Fourier space and with a degree of compression (integrated trispectrum; [Gualdi & Verde 2022](#)), but never separated into parity-odd modes; the history of NPCFs is reviewed in [Peebles \(2001\)](#).

Given that as yet no detection of parity-odd physics in large-scale structure has been made and that a number of proposed theoretical models can produce it, in this work we pursue a model-independent analysis. Hence we lack an *a priori* expectation for the shape of the signal. While such using such a model could strengthen any detection significance by correlating data in different modes, it would inevitably tie our detection significance to a particular model, which we wish to avoid. In contrast to typical analyses *e.g.* of the 2PCF for Baryon Acoustic Oscillations (BAO) or the 3PCF for BAO or galaxy biasing, in this work we cannot identify systematic errors simply by observing departures from an expected template for the cosmological signal. We must thus pay especial attention to systematics. We make extensive use of both mock catalogs and analytics to assess whether any systematics can produce spurious parity-odd 4PCF modes.

The present work is organized as follows. §2 outlines the multiple analyses of varying complexity used to increase confidence in our measurements. §3 reviews the basis used to decompose the 4PCF. We also present a toy model to illustrate the relation between parity and tetrahedra. §4 describes the data, simulations, and covariance matrix. §5 then presents two different paths to obtaining a detection significance, and their outcomes. We also present analysis of cross-correlations between spatially separated regions. §6 outlines our systematics tests on mocks as well as analytic work. §7 concludes. A number of Appendices present details of the work.

## 2 GUIDING PRINCIPLES FOR THE ANALYSIS

To isolate the potentially parity-violating component of the 4PCF, we expand the correlation function in two distinct sets of isotropic basis functions, one that is parity-even and one that is parity-odd. These are constructed from products of three spherical harmonics with angular-momentum indices  $\ell_1, \ell_2, \ell_3$ . Isotropy requires that the  $\ell_i$  satisfy the triangular inequality. If the sum of the  $\ell_i$  is even, the product is parity-even and if the sum is odd, the product is parity-odd. In this analysis, only the parity-odd elements are used. Each basis element is a function of three radial distances, the length of the sides from a chose vertex among the four defining a tetrahedron. In practice, the radial distances are binned.

Ideally, one would like to capture as much information as possible, using many narrow radial bins, and working to some high value of the  $\ell_i$ . This would increase the difficulty of evaluating all the independent amplitudes, but in fact the technique of Slepian & Eisenstein (2015b); Philcox et al. (2021a) makes this manageable, as we discuss in §3. The real challenge is in determining the covariance matrix. This is especially critical when looking for parity violation, where establishing a statistically significant non-zero signal is the crux and hence understanding the inevitable statistical fluctuations is essential.

A fairly modest choice of separating the radial variable into ten bins and including  $\ell_i$  only up to  $\ell_{\max} = 4$  results in  $23 \times 120 = 2,760$  independent amplitudes; we also consider eighteen radial bins, which produces  $816 \times 120 = 18,768$  4PCF amplitudes. Determining the covariance matrix is thus a formidable challenge. In order to invert a sampling covariance matrix, as required for calculating  $\chi^2$ , we need at least many mocks as the dimensionality of the data vector, which is in excess of the 2,000 available to us for BOSS.

We have chosen three ways to obtain the covariance matrix. First, the NPCF covariance matrix can be calculated analytically under the assumption of a Gaussian Random Field (GRF) as shown in Hou et al. (2021a). The GRF does not on average have any parity-violation at the *signal* level, but it can still have non-zero fluctuations in parity-odd modes. Therefore the GRF is still the leading-order contribution to the *covariance* of the parity-odd modes. This is simply the statement that a signal may be zero but its root-mean-square may not. We fit the analytic template covariance matrix by varying the number density and volume with respect to the covariance matrix derived from the mocks.

With this analytic template in hand, we may (i) directly compute the  $\chi^2$  of the data using the adjusted analytic covariance matrix. An alternative (ii) is to compress the data vector to reduce its dimensionality. The eigenvectors of the analytical covariance matrix with the smallest eigenvalues represent the linear combinations of basis functions that have the smallest statistical uncertainties. We then expand the measured 4PCF using just the  $N_{\text{eig}}$  best expansion functions. We may also determine the  $N_{\text{eig}} \times N_{\text{eig}}$  covariance matrix directly from the mocks. Since  $N_{\text{eig}}$  is much less than the number of mocks this covariance matrix is invertible. Finally, we may (iii) use the empirical covariance matrix from the mocks directly, with no involvement of the analytic template at all, by considering many fewer channels than in (i) and (ii) (we lower  $\ell_{\max}$  to be  $\ell_{\max} = 2$ ). Substantial reduction of the number of channels is required to enable inverting the empirical covariance matrix employed in this approach, and so we lose statistical power. We thus treat (iii) as a test rather than as giving us a main result of our analysis. The reliability of all three approaches above can be assessed using the mocks themselves by verifying that their  $\chi^2$  (or  $T^2$ ) values match the expected distribution.

Finally, our choice of fiducial cosmology in this work is consistent with the results of BOSS collaboration et al. 2017. We use a geometrically flat  $\Lambda$ CDM cosmology with matter density (in units of the critical density)  $\Omega_m = 0.31$ , baryon density (in units of the critical density)  $\Omega_b h^2 = 0.022$ , Hubble constant  $h \equiv H_0/(100 \text{ km/s/Mpc}) = 0.676$ , root-mean-square fluctuations on  $8h^{-1}\text{Mpc}$  spheres of  $\sigma_8 = 0.8$ , scalar spectral tilt  $n_s = 0.96$  and sum of the neutrino masses  $\sum m_\nu = 0.06 \text{ eV}$ .

## 3 METHOD FOR 4PCF MEASUREMENT

The 4PCF estimator, indicated by a hat, is

$$\begin{aligned} \hat{\zeta}(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3) &\equiv \langle \delta(\mathbf{s})\delta(\mathbf{s} + \mathbf{r}_1)\delta(\mathbf{s} + \mathbf{r}_2)\delta(\mathbf{s} + \mathbf{r}_3) \rangle \\ &= \int \frac{d\mathbf{s}}{V} \delta(\mathbf{s})\delta(\mathbf{s} + \mathbf{r}_1)\delta(\mathbf{s} + \mathbf{r}_2)\delta(\mathbf{s} + \mathbf{r}_3), \end{aligned} \quad (1)$$

where angle brackets denotes an ensemble average of the density fluctuations field  $\delta(\mathbf{s}) \equiv \rho(\mathbf{s})/\bar{\rho} - 1$ , with  $\rho(\mathbf{s})$  the density field and  $\bar{\rho}$  the average density.<sup>1</sup> Invoking ergodicity, the ensemble average may be replaced by an integral of spatial position  $\mathbf{s}$  over the volume  $V$ . This integration results in a function dependent only on the relative separation vectors  $\mathbf{r}_1$ ,  $\mathbf{r}_2$ , and  $\mathbf{r}_3$ . We also average over joint rotations of these vectors. In practice, the density fluctuation field is computed from discrete galaxy data, appropriately weighted.

Since the 3D distribution of galaxies is assumed to be isotropic on cosmological scales (ignoring redshift-space distortions), the isotropic basis (Cahn & Slepian 2020) is an efficient means of systematically extracting the cosmological information. The

<sup>1</sup> The 4PCF after rotation-averaging has six degrees of freedom, so we will only require certain combinations of the arguments on the lefthand side of the estimator.

isotropic basis functions required to measure an NPCF are given by products of  $(N - 1)$  spherical harmonics  $Y_{\ell m}(\hat{\mathbf{r}})$  combined according to angular momentum addition. In particular, for the 4PCF ( $N = 4$ ) we require the three-argument basis functions, which are

$$\mathcal{P}_{\ell_1 \ell_2 \ell_3}(\hat{\mathbf{r}}_1, \hat{\mathbf{r}}_2, \hat{\mathbf{r}}_3) = \sum_{m_1 m_2 m_3} \mathcal{C}_{m_1 m_2 m_3}^{\ell_1 \ell_2 \ell_3} Y_{\ell_1 m_1}(\hat{\mathbf{r}}_1) Y_{\ell_2 m_2}(\hat{\mathbf{r}}_2) Y_{\ell_3 m_3}(\hat{\mathbf{r}}_3). \quad (2)$$

The factorizability of these functions is important to the speed-up of the 4PCF algorithm (Philcox et al. 2021a); in practice it enables us to compute the 4PCF as a sum over the spherical harmonic coefficients  $a_{\ell_i m_i}$  of the density field about a given primary galaxy at  $\mathbf{s}$ .

Each unit vector  $\hat{\mathbf{r}}_i$  is associated with one angular momentum  $\ell_i$  whose  $z$ -component is  $m_i$ . The weight is

$$\mathcal{C}_{m_1 m_2 m_3}^{\ell_1 \ell_2 \ell_3} \equiv (-1)^{\ell_1 + \ell_2 + \ell_3} \begin{pmatrix} \ell_1 & \ell_2 & \ell_3 \\ m_1 & m_2 & m_3 \end{pmatrix}. \quad (3)$$

The 3- $j$  symbol enforces the triangular inequality  $|\ell_1 - \ell_2| \leq \ell_3 \leq \ell_1 + \ell_2$ , because the total angular momentum must be zero for an isotropic function.<sup>2</sup> Under the parity operator, denoted  $\hat{\mathbb{P}}$ , a spherical harmonic  $Y_{\ell m}$  transforms as  $(-1)^\ell$ , so the three-argument isotropic functions transform as

$$\begin{aligned} \hat{\mathbb{P}}[\mathcal{P}_{\ell_1 \ell_2 \ell_3}(\hat{\mathbf{r}}_1, \hat{\mathbf{r}}_2, \hat{\mathbf{r}}_3)] &\equiv \mathcal{P}_{\ell_1 \ell_2 \ell_3}(-\hat{\mathbf{r}}_1, -\hat{\mathbf{r}}_2, -\hat{\mathbf{r}}_3) \\ &= (-1)^{\ell_1 + \ell_2 + \ell_3} \mathcal{P}_{\ell_1 \ell_2 \ell_3}(\hat{\mathbf{r}}_1, \hat{\mathbf{r}}_2, \hat{\mathbf{r}}_3) = \mathcal{P}_{\ell_1 \ell_2 \ell_3}^*(\hat{\mathbf{r}}_1, \hat{\mathbf{r}}_2, \hat{\mathbf{r}}_3). \end{aligned} \quad (4)$$

Thus the basis functions are real if  $\ell_1 + \ell_2 + \ell_3$  is even and imaginary if the sum is odd.

The isotropic functions also satisfy an orthonormality relation, which follows from that of the spherical harmonics. We have

$$\int d\hat{\mathbf{r}}_1 d\hat{\mathbf{r}}_2 d\hat{\mathbf{r}}_3 \mathcal{P}_{\ell_1 \ell_2 \ell_3}(\hat{\mathbf{r}}_1, \hat{\mathbf{r}}_2, \hat{\mathbf{r}}_3) \mathcal{P}_{\ell'_1 \ell'_2 \ell'_3}^*(\hat{\mathbf{r}}_1, \hat{\mathbf{r}}_2, \hat{\mathbf{r}}_3) = \delta_{\ell_1 \ell'_1}^K \delta_{\ell_2 \ell'_2}^K \delta_{\ell_3 \ell'_3}^K. \quad (5)$$

The Kronecker delta  $\delta_{\ell_i \ell'_i}^K$  is unity when its subscripts are equal and zero otherwise.

The 4PCF estimator defined in Eq. (1) can be expanded into the basis of isotropic functions (see Eq. 2), where the expansion coefficients depend only on the  $r_i$  and are given by orthogonality as

$$\hat{\zeta}_{\ell_1 \ell_2 \ell_3}(r_1, r_2, r_3) = \int \frac{d\mathbf{s}}{V} \delta(\mathbf{s}) \int d\hat{\mathbf{r}}_1 d\hat{\mathbf{r}}_2 d\hat{\mathbf{r}}_3 \delta(\mathbf{s} + \mathbf{r}_1) \delta(\mathbf{s} + \mathbf{r}_2) \delta(\mathbf{s} + \mathbf{r}_3) \mathcal{P}_{\ell_1 \ell_2 \ell_3}^*(\hat{\mathbf{r}}_1, \hat{\mathbf{r}}_2, \hat{\mathbf{r}}_3). \quad (6)$$

To avoid an over-complete basis, the radial arguments  $r_i$  are ordered as  $r_1 \leq r_2 \leq r_3$ , as further discussed in Cahn & Slepian (2020).

To construct a density fluctuation field from the discrete galaxy counts, and also to account for the survey geometry, we use a generalized Landy-Szalay estimator (Landy & Szalay 1993; Szapudi & Szalay 1998, see also Kerscher et al. 2000) as first outlined for the angular momentum basis in Slepian & Eisenstein (2015b) and further developed in Philcox et al. (2021a,b). It is

$$\hat{\zeta}(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3) = \frac{\mathcal{N}(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3)}{\mathcal{R}(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3)}, \quad (7)$$

where  $\mathcal{N} \equiv (D - R)^4$  and  $\mathcal{R} \equiv R^4$ , and these powers are shorthand for expanding by the binomial theorem and letting each  $D$  and  $R$  be evaluated at a different spatial position.  $D$  means a particle drawn from the ‘‘data’’ and  $R$  means a particle drawn from the ‘‘random’’ catalog (a spatially uniform catalog cut by the survey geometry). As outlined in Slepian & Eisenstein (2015b), we may estimate the numerator and denominator separately (*i.e.* compute each separately averaging over the whole survey). Doing so gives optimally weighted estimates of each in the shot-noise limit, as discussed in Slepian & Eisenstein 2015b §4, equations 24-26 and surrounding text. Multiplying Eq. (7) through by  $\mathcal{R}$ , expanding each side of the resulting relation in the isotropic basis, reducing a product of two isotropic basis functions to a sum over single ones, and finally taking an inverse to solve the linear system so obtained (Slepian & Eisenstein 2015b; Philcox et al. 2021a), we find the edge-corrected 4PCF estimator as

$$\hat{\zeta}_{\ell_1 \ell_2 \ell_3}(r_1, r_2, r_3) = \sum_{\ell'_1 \ell'_2 \ell'_3} [\mathbf{M}^{-1}]_{\ell_1 \ell_2 \ell_3, \ell'_1 \ell'_2 \ell'_3} \mathcal{N}_{\ell'_1 \ell'_2 \ell'_3}(r_1, r_2, r_3) \frac{\mathcal{R}_{000}(r_1, r_2, r_3)}{\mathcal{R}_{000}(r_1, r_2, r_3)}. \quad (8)$$

We note that there is no mixing in the radial variables; survey geometry does not change lengths. Our notation indicates a given element of  $\mathbf{M}^{-1}$ . This latter is the inverse of the coupling matrix  $\mathbf{M}$  describing how survey geometry breaks the orthogonality of our basis functions, much as Fourier modes are orthogonal only on an infinite domain.  $\mathcal{N}$  denotes a measurement from the ‘‘data-minus-random’’ catalog, and  $\mathcal{R}_{000}$  is the  $\ell_1 = \ell_2 = \ell_3 = 0$  expansion coefficient of  $\mathcal{R} \equiv R^4$ , *i.e.* the randoms’ 4PCF.  $\mathcal{N}$  and  $\mathcal{R}$  are evaluated by replacing  $\delta$  in Eq. (6) with their definitions, given below Eq. (7).

<sup>2</sup> We recall that a 3- $j$  symbol with zeros in the bottom row demands that  $\ell_1 + \ell_2 + \ell_3$  be even, but with non-zero  $m_i$  there is no such requirement, allowing odd sums of the  $\ell_i$  and hence parity-odd basis functions.

The coupling matrix has elements

$$\mathbf{M}_{\ell_1 \ell_2 \ell_3, \ell'_1 \ell'_2 \ell'_3}(r_1, r_2, r_3) = (4\pi)^{-3/2} (-1)^{\ell'_1 + \ell'_2 + \ell'_3} \sum_{L_1 L_2 L_3} \frac{\mathcal{R}_{L_1 L_2 L_3}(r_1, r_2, r_3)}{\mathcal{R}_{000}(r_1, r_2, r_3)} \prod_{i=1}^3 D_{\ell_i L_i \ell'_i}^{\text{P}} C_{000}^{\ell_i L_i \ell'_i} \left\{ \begin{array}{ccc} \ell_1 & L_1 & \ell'_1 \\ \ell_2 & L_2 & \ell'_2 \\ \ell_3 & L_3 & \ell'_3 \end{array} \right\}, \quad (9)$$

with the coefficient

$$\mathcal{D}_{\ell_1 \ell_2 \ell_3}^{\text{P}} = \sqrt{(2\ell_1 + 1)(2\ell_2 + 1)(2\ell_3 + 1)} \quad (10)$$

which depends on the product of the *primary* (hence the superscript ‘‘P’’) angular momenta.<sup>3</sup> The matrix in curly brackets in Eq. (10) is a Wigner 9-*j* symbol. The factor  $C_{000}^{\ell_i L_i \ell'_i}$  (defined in Eq. 3 preceding it guarantees that  $\ell_i, \ell'_i$ , and  $L_i$  can be combined to make a zero total angular momentum state. It also requires that  $\ell_i + \ell'_i + L_i$  is even.

Regarding the edge correction, we note that formally  $\mathbf{M}$  is infinite, but we have found in practice truncating it at one angular momentum beyond that used for the physical analysis is suitable (Slepian & Eisenstein 2015b, Philcox et al. 2021a). In this work we use  $\ell_{\text{max}} = 4$  for our analysis but work to  $\ell = 5$  on all  $\ell_i$  for the edge correction.<sup>4</sup> Further details regarding the suitability of, when performing the edge correction, truncating at an  $\ell$  one above that used for the analysis, are in Slepian & Eisenstein (2015b). Ultimately this suitability stems from the rough tri-diagonality of the edge-correction matrix (see their §4.2. and our Fig. 28).

### 3.1 Illustration With Toy Tetrahedra

To understand the parity-odd measurement more intuitively, we study cubic boxes of side length  $L_{\text{box}} = 1000 h^{-1} \text{Mpc}$  with tetrahedra tuned to produce particular parity signals. To fill the boxes as fully as possible, yet at the same time have tetrahedra with minimum side length of order  $10 h^{-1} \text{Mpc}$ , which is similar to the situation in our BOSS dataset, we choose the three sides extending from the primary to be roughly  $r_1 \sim 10 h^{-1} \text{Mpc}$ ,  $r_2 \sim 20 h^{-1} \text{Mpc}$ , and  $r_3 \sim 30 h^{-1} \text{Mpc}$ . We require that the minimum separation between primaries be twice the longest side of the tetrahedron (*i.e.*  $60 h^{-1} \text{Mpc}$ ) in order to minimize any overlap between tetrahedra. Finally, we have  $N_{\text{tets}} \sim 1,500$  tetrahedra within each cubic box. Fig. 2 shows an example box and also an example tetrahedron and its partner under parity transformation.

In 3D, parity transformation is equivalent to a mirror reflection across a 2D plane (which flips the sign of the coordinate axis perpendicular to that plane) plus a  $180^\circ$  rotation around this latter axis. For simplicity, in Fig. 2 we depict the parity transformation simply as a mirror reflection, since our basis, being isotropic, always averages over 3D rotations and is thus insensitive to a  $180^\circ$  rotation. Put another way, only the mirroring fundamentally alters the shape of a tetrahedron; the  $180^\circ$  rotation only changes its orientation in absolute space.<sup>5</sup> In particular, one can imagine the mirroring as taking one side and ‘‘pulling it through’’ the tetrahedron from being on one side of the plane formed by the other two sides, to being on the other side of this plane.

In practice, we must allow all four vertices of each tetrahedron a chance to be the primary (Philcox et al. 2021a). However, for this toy tetrahedron illustration we restricted to bins in side length (radial bins) such that  $9 h^{-1} \text{Mpc} < r_i < 30 h^{-1} \text{Mpc}$  for all  $r_i$  so that nearly always only one of the four vertices will satisfy the radial bins required for each tetrahedron. This means that the contribution to the signal will stem only from the isotropic function evaluated on unit vectors extending from a *single* primary; this renders it easier to understand the measured signal in this toy model.

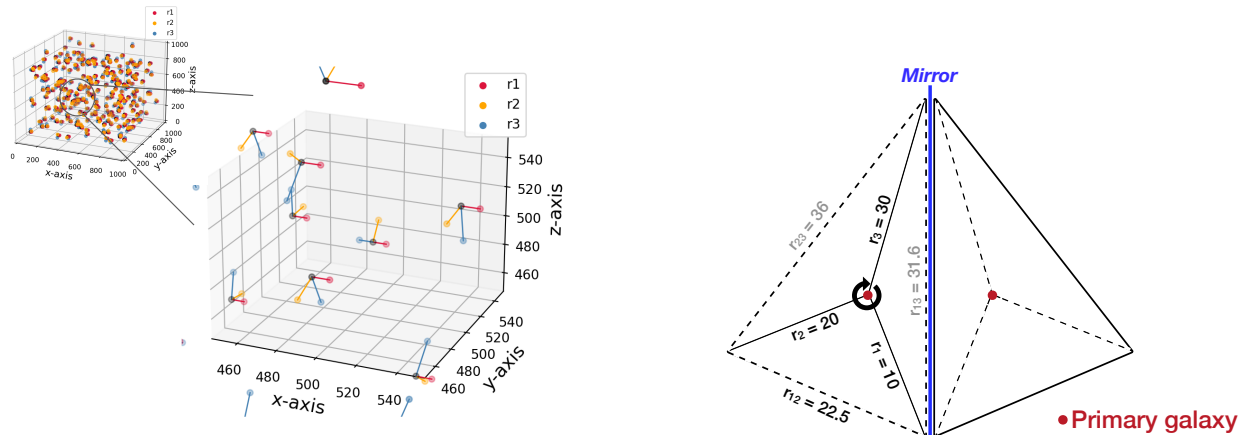
We set  $\hat{\mathbf{r}}_1$  to be along  $\hat{\mathbf{x}}$ ,  $\hat{\mathbf{r}}_2$  to be along  $\hat{\mathbf{y}}$ , and  $\hat{\mathbf{r}}_3$  to be along  $\hat{\mathbf{z}}$ . The tetrahedron is therefore clockwise at the only primary allowed in this toy model; our convention is presented in Fig. 1. A parity transformation simply means that we interchange the sides, so that  $\hat{\mathbf{r}}_2$  aligns with the  $x$ -axis and  $\hat{\mathbf{r}}_1$  aligns with the  $y$ -axis, while  $\hat{\mathbf{r}}_3$  is unchanged. Again, characterization as ‘‘clockwise’’ or ‘‘counterclockwise’’ depends on at which vertex one sits, as discussed in Fig. 1, but here is unambiguous. By restricting the radial bins we use, we force only one galaxy to be the primary; the other choices of primary will not lead to sides extending from them that fall within our chosen bins. Thus, we can ensure that the tetrahedra for this test are perfectly counterclockwise as viewed from the single primary allowed.

We produce toy boxes in three configurations. The first is filled with only counterclockwise tetrahedra (enforced by the radial bin restriction); the second has only clockwise tetrahedra, and the third has an equal mix. To render each toy box somewhat more realistic, we randomly rotate each vertex by an angle  $\theta \in [-180^\circ, 180^\circ]$  around each of the three Cartesian coordinates. We also add random numbers  $\Delta r_1 \in [0, 1]$ ,  $\Delta r_2 \in [-2, 2]$ , and  $\Delta r_3 \in [-1, 0]$  (in  $h^{-1} \text{Mpc}$ ) to, respectively,  $r_1$ ,  $r_2$ , and  $r_3$ . We

<sup>3</sup> Were one to measure an NPCF for  $N \geq 5$ , one would require isotropic basis functions of four arguments or more, and these basis functions require specification of intermediate angular momenta fixing how the primary momenta are coupled (further detailed in Cahn & Slepian 2020). The distinction between primary and intermediate angular momenta is not needed in the present work, but for consistency we retain the superscript ‘‘P.’’

<sup>4</sup> This truncation does not induce spurious parity-odd modes; if it did, we would see them when we edge-correct our mock catalogs.

<sup>5</sup> In general parity transformation and mirror reflection are distinct operations.



**Figure 2.** *Left:* A box with side length  $L_{\text{box}} = 1000h^{-1}\text{Mpc}$  filled with tetrahedra (upper left, small panel). Each of them has a *unique* primary (black) from which the three sides with respective lengths  $r_1 \sim 10h^{-1}\text{Mpc}$  (red),  $r_2 \sim 20h^{-1}\text{Mpc}$  (yellow), and  $r_3 \sim 30h^{-1}\text{Mpc}$  (blue) extend. The larger panel on the left is a zoom-in on the full box to display the tetrahedra more clearly. *Right:* A sketch of a “clockwise” tetrahedron and its “counterclockwise” mirror image. The primary is in red. Our convention on clockwise and counterclockwise is detailed in Fig. 1. On the left, the red point is closer to us than all the others. Thus the tetrahedron on the left is clockwise, as looking down from the primary, we go clockwise as we move from the smallest side to the largest. On the right, the primary (in red) is behind the other galaxies, so looking down from it towards them will reverse the handedness. Thus the rightmost tetrahedron is counterclockwise as viewed from the primary.

choose these ranges such that we always have  $r_1 < r_2 < r_3$ . Hence the parity of a given tetrahedron will not be flipped by these additions.

Fig. 3 shows the 4PCF of these illustrative boxes. We may approximate each tetrahedron as a sphere about the primary of radius roughly  $20h^{-1}\text{Mpc}$ , with volume  $V_{\text{tet}}$ , to estimate the expected 4PCF amplitude. We define  $n$  as the local number density due to a given tetrahedron,  $n = 1/V_{\text{tet}}$ , and  $\bar{n}$  as the average number density in the box,  $\bar{n} = N_{\text{tets}}/L_{\text{box}}^3$ . We find  $\langle \delta^4 \rangle = [n/\bar{n} - 1]^4 \approx 2 \times 10^5$ . The lowest-lying parity-odd isotropic basis function is  $\mathcal{P}_{111} = -3i/[\sqrt{2}(4\pi)^{3/2}] \hat{\mathbf{r}}_1 \cdot (\hat{\mathbf{r}}_2 \times \hat{\mathbf{r}}_3)$  (Cahn & Slepian 2020). As expected, the counterclockwise-only box has a positive projection onto this function, while the clockwise-only box has a negative projection. The mixed box is consistent with zero projection onto  $\mathcal{P}_{111}$  on average. We can analytically predict the ratios among the 4PCF coefficients for different channels  $\ell_1, \ell_2, \ell_3$ . We compare the mean ratios of the measured 4PCF coefficients for several combinations to these predictions and find good agreement.<sup>6</sup> This also serves as an additional test of our code (the code is further discussed in §4.1).

### 3.2 Internal Cancellation

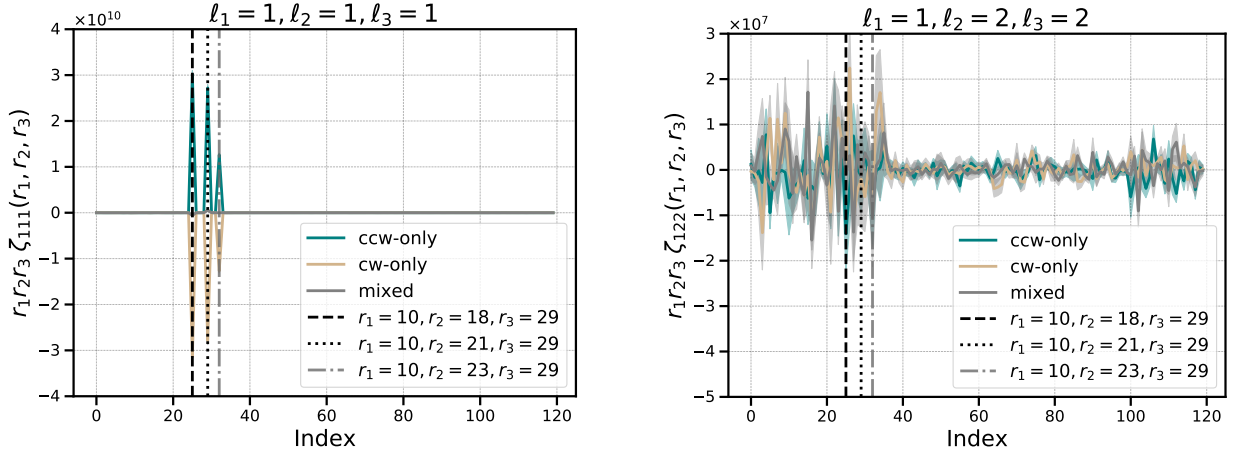
For a given tetrahedron, in practice (but not in our illustrative boxes above), each of the four vertices gets a chance to serve as the primary about which the isotropic basis function expansion is computed. Some of these vertices will be “clockwise”, some “counterclockwise”. Hence if co-added into the same channel and triple-bin there will be “internal cancellation” and consequent reduction of any parity-odd signal. However, if the radial bins are made fine enough then each vertex, in virtue of the presumably unique lengths of the sides extending from it, will be accumulated to a different triple-bin.<sup>7</sup> Thus finer binning can reduce the internal cancellation and increase the signal. This is much the same as in a configuration-space BAO search, where fine enough bins must be chosen that the BAO feature in the 2PCF is not averaged out by all being added into a single bin.

## 4 DATASET AND COVARIANCE

We use the final galaxy catalog of the Baryon Oscillation Spectroscopic Survey (BOSS), from the twelfth data release (DR12) of the Sloan Digital Sky Survey-III (SDSS-III). The catalog is split into the North Galactic Cap (NGC) and the South Galactic Cap (SGC). The catalog contains two samples, CMASS and LOWZ, which were selected via the SDSS multicolor photometry

<sup>6</sup> As an example, the analytically predicted ratio for angular momenta  $\{1, 1, 1\}$  and  $\{1, 3, 3\}$  is  $\zeta_{111}/\zeta_{133} = -1.07$ , and we measure from the data  $\bar{\zeta}_{111}/\bar{\zeta}_{133} = -1.07$ , with  $\bar{\zeta}$  denoting the bin-averaging.

<sup>7</sup> Save for isosceles or equilateral tetrahedra, which we exclude from our analysis in any case.



**Figure 3.** Here we display 4PCF measurements from the illustrative toy boxes. The left and right panels each show a different channel, as indicated in their titles. In both panels, the “clockwise-only” box results are in blue, the “counterclockwise-only” box results are in brown, and the “mixed” box results are in grey. *Left:* Projection onto  $\mathcal{P}_{111}$ . As discussed in the main text, we expect a negative projection for the clockwise box and a positive projection for the counterclockwise box. The mixed will have on average zero projection. These expectations are borne out and indeed analytic calculation yields agreement with the measured signals. *Right:* Projection onto  $\mathcal{P}_{122}(\hat{\mathbf{r}}_1, \hat{\mathbf{r}}_2, \hat{\mathbf{r}}_3) \propto \hat{\mathbf{r}}_1 \cdot (\hat{\mathbf{r}}_2 \times \hat{\mathbf{r}}_3)(\hat{\mathbf{r}}_2 \cdot \hat{\mathbf{r}}_3)$ . By construction the scalar triple product  $\hat{\mathbf{r}}_1 \cdot (\hat{\mathbf{r}}_2 \times \hat{\mathbf{r}}_3)$  is unity for all three of our illustrative boxes. The amplitudes on the righthand side are not strictly zero because there are still a small number of tetrahedra formed by connecting secondaries around one primary with secondaries around another. Additionally, the amplitude fluctuations divide into two envelopes at bin index 36. This bin index corresponds to raising the index for  $r_1$  (the smallest side and the slowest-varying one).

and cover a redshift range  $0.15 < z < 0.7$ . CMASS and LOWZ use similar target selection algorithms (Eisenstein et al. 2001; Cannon et al. 2006). The target selection algorithm provides samples that are mainly composed of Luminous Red Galaxies (LRGs). For CMASS the selection algorithm is further tuned to select massive objects uniformly in redshift (Reid et al. 2016), which results in an approximately mass-limited sample down to a stellar mass  $M \sim 10^{11.3} h^{-1} M_{\odot}$  (Maraston et al. 2013). The majority of CMASS is LRGs ( $\sim 74$  percent), while the rest is late-type spirals (Masters et al. 2011). LOWZ consists primarily of LRGs (Parejko et al. 2013). Despite the difference in target selection, the LRGs in two samples has similar stellar mass distribution (Maraston et al. 2013). We apply a redshift cut of  $0.43 < z < 0.7$  to CMASS, which results in a redshift tail from LOWZ at redshift  $0.43 < z < 0.5$ . This tail ( $\sim 15\%$  of the entire “CMASS” sample) slightly raises the purity of the CMASS sample by adding more LRGs. To ensure that the LOWZ sample as used here is independent of the CMASS one, we apply a redshift cut of  $0.2 < z < 0.4$  to the former. This cut produces a separation of about  $70 h^{-1} \text{Mpc}$  between the lower edge of CMASS and the upper edge of LOWZ. Fig. 4 shows the number density for each sample as a function of redshift. Finally, we note that the early LOWZ target selection was not uniform due to the use of different iterations of the galaxy-star separation algorithm (Reid et al. 2016). Therefore we do not include those early chunks in this analysis.

#### 4.1 4PCF from Data and Mocks

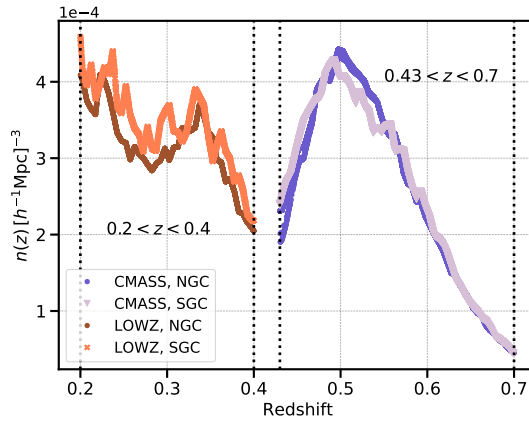
For our main analyses, we considered tetrahedra with side lengths  $r_1$ ,  $r_2$  and  $r_3$  from the primary ranging from  $20 h^{-1} \text{Mpc}$  to  $160 h^{-1} \text{Mpc}$  inclusive, split into ten linearly-spaced bins, and also from  $20 h^{-1} \text{Mpc}$  to  $164 h^{-1} \text{Mpc}$  inclusive, split into eighteen linearly-spaced bins. We also explored a coarser binning (six bins), presented as a test in Appendix E. As a result the sides of the tetrahedron that do not include the primary can range from zero to  $320 h^{-1} \text{Mpc}$ . We expand the parity-odd 4PCF in the 23 angular channels with  $\ell_i \leq 4$  given in Cartesian form in Appendix A. For the edge corrections, we include all functions with  $\ell_i \leq 5$ , as further discussed in §3. We also compute the even-parity modes (for which we do not reproduce the basis functions here), as they are needed within the edge-correction, despite that our actual analysis focuses solely on the parity-odd ones.

To each galaxy we apply a total weight  $w$  given by

$$w = w_{\text{FKP}} w_{\text{sys}} (w_{\text{noz}} + w_{\text{cp}} - 1), \quad (11)$$

with the systematic weight  $w_{\text{sys}}$  (a combined weight for stellar density and seeing), the redshift failure weight  $w_{\text{noz}}$ , and the fiber collision weight  $w_{\text{cp}}$  (subscript “cp” for “close pairs”). The FKP weight (Feldman et al. 1994) is  $w_{\text{FKP}} = [1 + n(z)P_0]^{-1}$ , with  $P_0 = 10^4 [h^{-1} \text{Mpc}]^3$ ;  $n(z)$  is the weighted number density at the given redshift. We use the public random catalog provided by BOSS<sup>8</sup>, which is fifty times the size of the data catalog. We split it into 32 chunks such that each chunk is about 1.5 to 2 times

<sup>8</sup> <https://data.sdss.org/sas/dr12/boos/lss/>



**Figure 4.** The number density  $n$  as a function of redshift  $z$  for the two BOSS samples used in this work. The LOWZ ( $0.2 < z < 0.4$ ) North Galactic Cap (NGC) is in brown and South Galactic Cap (SGC) is in orange. For CMASS ( $0.43 < z < 0.7$ ), the NGC is in purple and the SGC is in lavender. We intentionally do not allow redshift overlap between the two samples, and the redshift gap  $\Delta z = 0.03$  corresponds to a comoving radial separation of  $73 h^{-1} \text{Mpc}$ . This separation means that the samples are fairly independent; the 2PCF  $\xi$  between a point in LOWZ and CMASS would be of order 1% at this scale. The covariance between two “worst-case” tetrahedra, one in each sample (where each is very close to the respective edge), is of order  $\xi^4$ . Few tetrahedra are near enough either edge to be significantly correlated with those in the other slice. The plot shows that LOWZ has both a more uniform selection function and a somewhat higher average number density than CMASS. It also lacks the strongly decaying tail with increasing redshift that CMASS displays. These points are important when assessing the possible impact of systematics on each sample and when addressing any differences between the detection significances in the two samples.

the size of the data; the reason is discussed in Slepian & Eisenstein (2015b) and Philcox et al. (2021b) though the detailed splitting does not affect the 4PCF algorithm speed notably. Each chunk is first randomly shuffled and then normalized so that its weighted sum matches the sum of the completeness weights  $w_c$  of the data.

We use 2,000 public MultiDark PATCHY lightcone mocks (PATCHY hereafter; Kitaura et al. 2016; Rodríguez-Torres et al. 2017). These mocks use second-order Lagrangian Perturbation Theory (2LPT) plus a spherical collapse prescription calibrated on full N-body simulations, and the mocks are additionally calibrated to match the 2-point and some 3-point statistics of the observed BOSS CMASS and LOWZ samples. The PATCHY mocks include realistic survey geometry as well as many observational effects, further detailed in Rodríguez-Torres et al. (2017).

The 4PCF measurement for the BOSS data catalog and the 32 ( $D - R$ ) catalogs alone takes 0.33 GPU-hours for the measurement where we use ten radial bins, and 0.65 GPU-hours for the measurement where we use eighteen radial bins; both timings are using 69 NVIDIA A100 GPUs (further discussed below) and include edge-correction, which is done on the CPU linked to the GPU and takes negligible extra time.

All calculations were done with the GPU-accelerated NPCF code CADENZA (Slepian et al. 2021), built on the CPU-based NPCF code ENCORE (Philcox et al. 2021a). CADENZA uses the representation of the basis functions in spherical harmonics, for reasons outlined in §3. Tests of ENCORE are described in (Philcox et al. 2021a), and CADENZA was verified to obtain exactly the same outputs on a fiducial data set used for testing. As noted already, we used 69 NVIDIA A100 GPUs simultaneously on the HiPerGator cluster at the University of Florida. Overall, all of the computations performed for this work would have taken roughly 7.5 million CPU hours, or about 20 months of continuously running on a cluster of 500 CPU cores. In practice, since CADENZA is about  $140\times$  faster than ENCORE for the 4PCF, we required about 54,000 GPU-hours, or about 1 month if one ran continuously on 69 GPUs.

## 4.2 Covariance Matrix

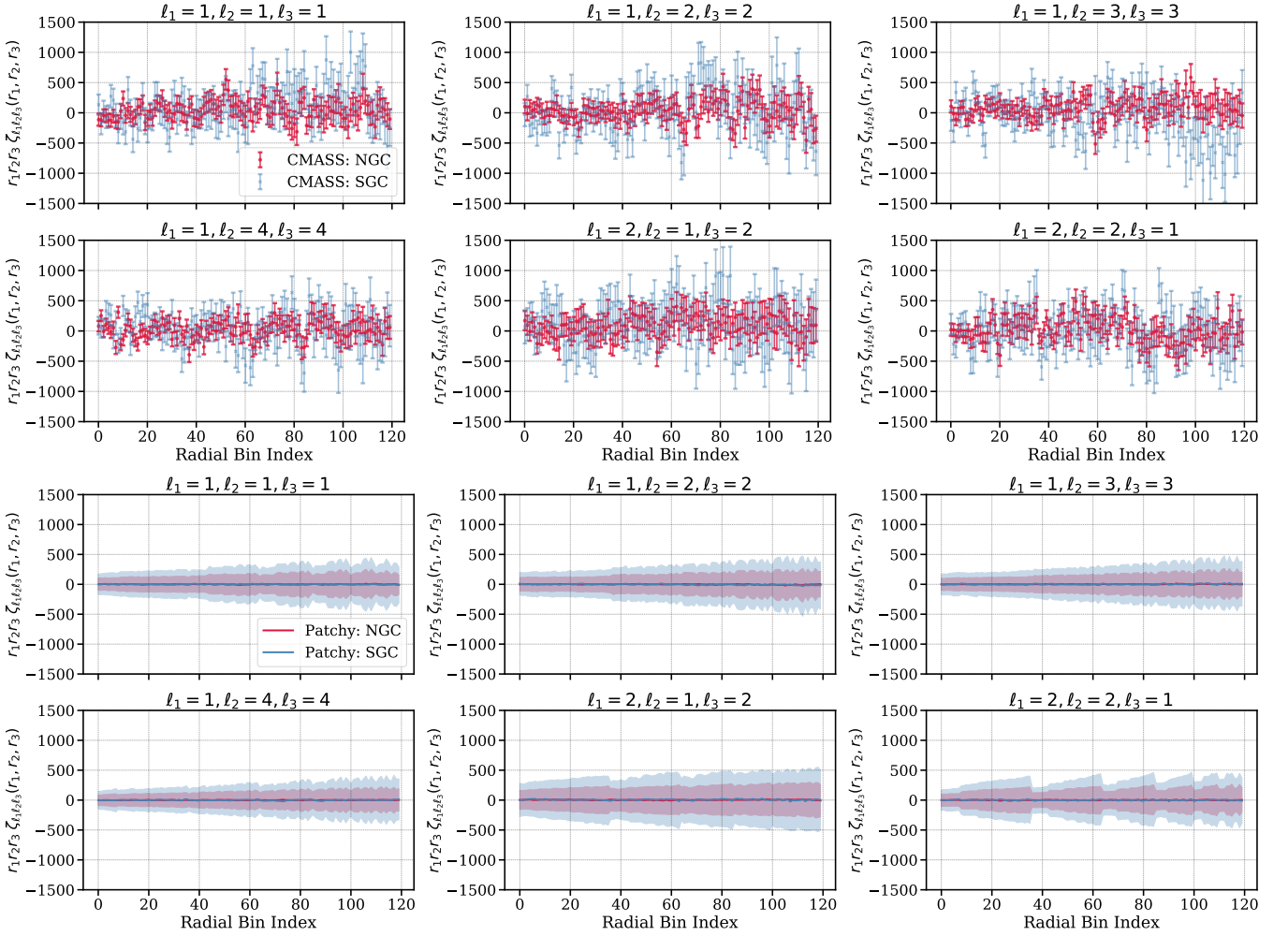
### 4.2.1 Analytic Covariance

The covariance matrix for the 4PCF may be computed analytically if we assume that the density is a Gaussian Random Field (Hou et al. 2021a). It is expressed in terms of  $f$ -functions

$$f_{\ell_1 \ell_2 \ell_3}(r_1, r_2, s) \equiv \int \frac{k^2 dk}{2\pi^2} P(k) j_{\ell_1}(kr_1) j_{\ell_2}(kr_2) j_{\ell_3}(ks), \quad (12)$$

with  $j_\ell(kr)$  the spherical Bessel function of order  $\ell$  and  $P(k)$  the galaxy power spectrum. The power spectrum is taken to be isotropic; the effect of RSD (after averaging over the line of sight) is simply to boost the amplitude, which is embedded in the input power spectrum and also partially absorbed in our later fitting for the best amplitude. Fast approaches do exist to evaluate these  $f$ -integrals (*e.g.* Slepian et al. 2019) but here we use direct computation.



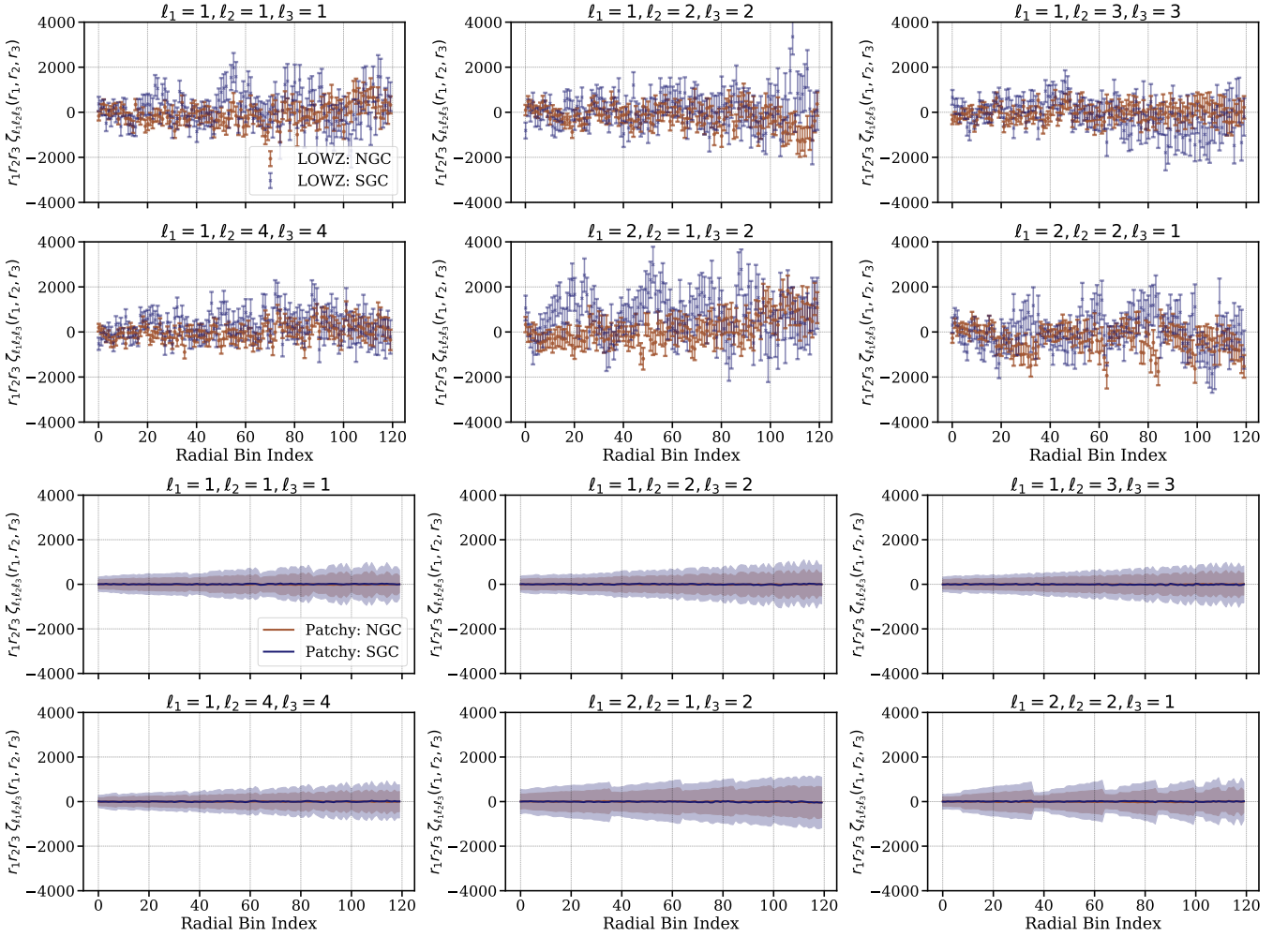


**Figure 5.** *Upper:* The parity-odd 4PCF for the BOSS CMASS data with **ten** radial bins, with NGC in red and SGC in blue, showing only the six lowest-lying channels. The error bars are the root-mean-square of the PATCHY mocks (over our set of 2,000). Here for legibility we focus on the ten-bin results; the eighteen-bin results are in Appendix C, as are the remaining channels for the ten-bin results. We have mapped the three radial bins to a single index, with  $r_3$  varying the fastest and  $r_1$  the slowest. This is done in all similar plots that follow. *Lower:* The mean of 2,000 PATCHY mocks (solid curves) for both NGC (red) and SGC (blue). The shaded region is the rms expected for a single mock; this set of panels is intended to display the region around zero that the 4PCF of a dataset with no true parity-odd signal might inhabit.

The analytic covariance matrix couples each vertex of an (unprimed) tetrahedron to a vertex in another (primed) tetrahedron. Since in our approach, the primary vertex is distinguished from the other three, in the covariance expression there are two pieces: in one the two primary vertices, unprimed and primed, are paired with each other; in the other they are not. The final result will be a sum given by products involving Wigner- $n_j$  symbols and the  $f$ -functions of Eq. (12). Below we simply reproduce the final result derived in Hou et al. (2021a), for brevity here we define  $\Lambda \equiv \{\ell_1, \ell_2, \ell_3\}$  and  $\Lambda' \equiv \{\ell'_1, \ell'_2, \ell'_3\}$ .

$$\text{Cov}_{\Lambda, \Lambda'}(r_1, r_2, r_3; r'_1, r'_2, r'_3) = \text{Cov}_{\Lambda, \Lambda'}^{\text{I}}(r_1, r_2, r_3; r'_1, r'_2, r'_3) + \text{Cov}_{\Lambda, \Lambda'}^{\text{II}}(r_1, r_2, r_3; r'_1, r'_2, r'_3). \quad (13)$$

$$\begin{aligned} \text{Cov}_{\Lambda, \Lambda'}^{\text{I}}(r_1, r_2, r_3; r'_1, r'_2, r'_3) &= (4\pi)^4 \sum_G (-1)^{\Sigma(\Lambda)(1-\varepsilon_G)/2} \sum_{L_1 L_2 L_3} \mathcal{D}_{L_1 L_2 L_3}^{\text{P}} C_{000}^{L_1 L_2 L_3} \begin{Bmatrix} \ell_{G1} & \ell_{G2} & \ell_{G3} \\ \ell'_1 & \ell'_2 & \ell'_3 \\ L_1 & L_2 & L_3 \end{Bmatrix} \\ &\times \int \frac{s^2 ds}{V} \prod_{i=1}^3 \left[ (-1)^{(-\ell_{Gi} - \ell'_i + L_i)/2} \mathcal{D}_{\ell_i \ell'_i L_i}^{\text{P}} C_{000}^{\ell_{Gi} \ell'_i L_i} \xi(s) f_{\ell_{Gi} \ell'_i L_i}(r_{Gi}, r'_i, s) \right], \end{aligned} \quad (14)$$



**Figure 6.** *First and second rows:* measurement of the parity-odd 4PCF using **ten** radial bins for the BOSS LOWZ data in the NGC (brown) and SGC (blue). We show the same six channels as in Fig. 5, and the errorbars are again the rms over the mocks. Again for legibility we focus on the ten-bin results; the eighteen-bin results are in Appendix G, as are the remaining channels for the ten-bin analysis. *Third and fourth rows:* The mean of 2,000 PATCHY mocks (solid curves) for both NGC (brown) and SGC (blue). The shaded region is the rms expected for a single mock.

$$\begin{aligned}
 \text{Cov}_{\Lambda, \Lambda'}^{\text{II}}(r_1, r_2, r_3; r'_1, r'_2, r'_3) &= (4\pi)^4 \sum_{G, H} (-1)^{\Sigma(\Lambda')(1-\varepsilon_H)/2} \sum_{L_1 L_2 L_3} \mathcal{D}_{L_1 L_2 L_3}^{\text{P}} \mathcal{C}_{000}^{L_1 L_2 L_3} \begin{Bmatrix} \ell_{G1} & \ell_{G2} & \ell_{G3} \\ \ell'_{H1} & \ell'_{H2} & \ell'_{H3} \\ L_1 & L_2 & L_3 \end{Bmatrix} \\
 &\times \int \frac{s^2 ds}{V} \prod_{i=1}^3 \left[ (-1)^{(-\ell_{Gi} - \ell'_{Hi} + L_i)/2} \mathcal{D}_{\ell_{Gi} \ell'_{Hi} L_i}^{\text{P}} \mathcal{C}_{000}^{\ell_{Gi} \ell'_{Hi} L_i} \right] \\
 &\times f_{\ell_{G1} 0 \ell_{G1}}(r_{G1}, 0, s) f_{0 \ell'_{H1} \ell'_{H1}}(0, r'_{H1}, s) f_{\ell_{G2} \ell'_{H2} L_2}(r_{G2}, r'_{H2}, s) f_{\ell_{G3} \ell'_{H3} L_3}(r_{G3}, r'_{H3}, s).
 \end{aligned} \tag{15}$$

In Case I, the sum over  $G$  includes all six permutations, while in Case II the sum over  $G$  includes only cyclic permutations. The sum over  $H$  includes all six permutations.  $\varepsilon_G$  is the Levi-Civita symbol, giving a negative sign if the permutation  $G$  is odd.  $\Sigma(\Lambda') \equiv \ell'_1 + \ell'_2 + \ell'_3$ . The coefficients  $\mathcal{C}_{m_1 m_2 m_3}^{\ell_1 \ell_2 \ell_3}$  and  $\mathcal{D}_{\ell_1 \ell_2 \ell_3}^{\text{P}}$  are given in Eq. (3) and Eq. (10), respectively. For the parity-even modes of the 4PCF, the phases in Eqs. (14) and (15) are identically unity because the sum of the three angular momenta is always even. However, these phases do matter for the parity-odd 4PCF.

We also note that, apart from the leading Gaussian contribution, the covariance should also contain non-Gaussian contributions as well as the coupling of the redshift-space power spectrum multiple moments to different angular channels of the covariance. Any combination of multipoles that after multiplying out is rotation-invariant (*i.e.* has total angular momentum zero) can contribute. An expression for this RSD-added analytic covariance is in Hou et al. (2021a) Appendix E., there it is also shown that these redshift-space induced coupling can also increase the off-diagonals of the covariance. Hou et al. (2021a) showed that the rms of the off-diagonals after comparison of the analytic covariance to the mock-based covariance is exactly

the same in real-space and redshift-space, both for lognormal mocks and for QUIJOTE simulations (*cf.* figure 6, 7 and figure 10, 11). Therefore, off-diagonal enhancement by redshift-space induced coupling is in principle not significant enough to greatly affect our analysis. Furthermore, some of the redshift-space-induced effects can be absorbed in the overall normalization, as much of their effect after isotropic averaging is to rescale the overall power spectrum. Nevertheless, we caution that missing these other terms could in principle reduce the accuracy of our analytic covariance relative to that from the mocks

Fitting the volume and number-density to match the mocks' covariance can be done without inverting this latter, important to avoid as that inverse will be biased unless the number of degrees of freedom much smaller than the number of mocks. For the fitting, we maximize a likelihood based on the Kullback-Leibler (KL) divergence (Kullback & Leibler 1951). One can show that with the likelihood Eq. (55) of Hou et al. (2021a), the optimal volume is at a given number density is uniquely determined, so the optimization is only a 1D problem, making it efficient. There is no guarantee that the best-fitting number density and volume for the covariance of the even-parity 4PCF (*e.g.* as found in Philcox et al. 2021b) and for the covariance of the parity-odd 4PCF should be the same. We summarize the values obtained for each sample and each radial binning in Table 1.

In Fig. 7 we compare for CMASS NGC the parity-odd 4PCF analytic covariance matrix to that from PATCHY for ten radial bins. We have mapped the channel and triple-bin to a 1D index, so the covariance may be plotted as a 2D matrix (one 1D index for the unprimed channels and triple-bins, another for the primed). In the left panel we show a direct comparison of the correlation matrix, which is the covariance matrix normalized by its diagonal. The similarity between the analytic correlation matrix (upper triangle) and the mock correlation matrix (lower triangle) shows that the analytic approach can well capture the covariance's features. In the right panel we compare the diagonal elements of the covariance matrices and also show the ratio between the mock and analytic covariance diagonals. Again this uses our mapping of the angular momenta and radial bins into a 1D index. Despite the similarity between the two diagonals in their behavior with increasing index, there is non-negligible variation. On average, the mean of the ratio between the mock and analytic covariance diagonal elements is roughly 1.05 for CMASS NGC. Fig. 8 shows the same comparisons for the LOWZ NGC. Again, we see that the analytic covariance can well describe that of the mocks. The mean of the ratio is 1.01 for the LOWZ NGC.

To further quantify the similarity between the analytic and mock covariance, we define a **half-inverse test** matrix as

$$\Phi \equiv \mathbf{C}_{\text{analytic}}^{-1/2} \mathbf{C}_{\text{patchy}} \mathbf{C}_{\text{analytic}}^{-1/2} - \mathbb{1}, \quad (16)$$

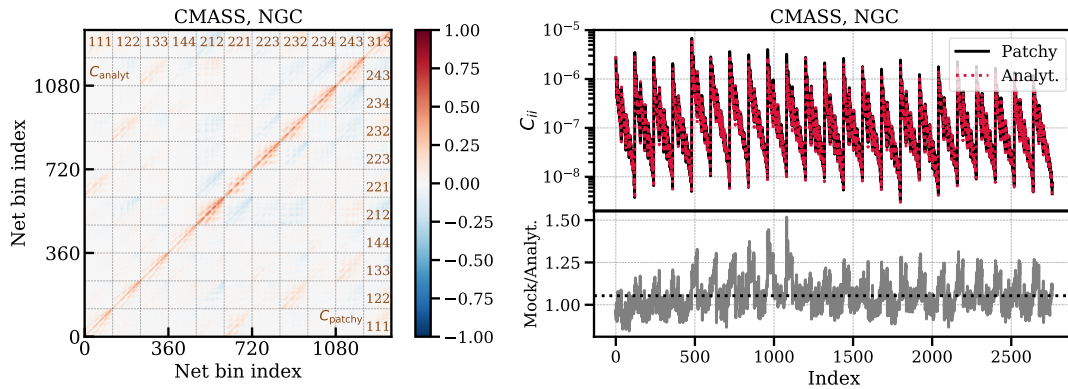
where  $\mathbf{C}_{\text{analytic}}^{-1/2}$  is the square root of the inverse of the analytic covariance matrix, and  $\mathbf{C}_{\text{patchy}}$  is the covariance from the mocks. We use half-inverses so that the test is symmetric. Were the two covariance matrices identical,  $\Phi$  would be the zero matrix (we note that the identity matrix is subtracted). One can show that the optimal volume at a given number density (as discussed above) will imply that  $\Phi$  is traceless.

Fig. 9 shows the lower half of the matrix  $\Phi$  for the half-inverse test of CMASS NGC (left) and LOWZ NGC (right), with the standard deviation shown on the upper triangle. The standard deviation for all elements is  $\sigma_{\text{all}} = 0.06$  for both CMASS NGC and LOWZ NGC, which is expected as it scales as  $1/\sqrt{N_{\text{mocks}}}$  (we used  $N_{\text{mocks}} = 2,000$  for this test). Each element of the covariance matrix follows a Wishart (1928) distribution, and the variance of the diagonal elements for this distribution is two times that of the off-diagonal elements. We do find that the standard deviation  $\sigma_{\text{diag}}$  is approximately two times that of the off-diagonal,  $\sigma_{\text{non-diag}}$ . However, we also find residuals in the diagonal elements in  $\Phi$  for both the left and right panels, and larger residuals for LOWZ NGC than for CMASS NGC. These indicate that fitting the volume and number density to best match the mock covariance is imperfect. As mentioned in §2, one of our analysis approaches (“compressed”) will only require a smooth estimate of the covariance, and turns out to be insensitive to its amplitude. Furthermore, even when employing the analytical covariance matrix directly, we may compare the  $\chi^2$  from data only to the distribution of  $\chi^2$  values for the mocks, both computed using the same covariance matrix. This comparison should minimize any bias in the detection significance computed using the analytic covariance.

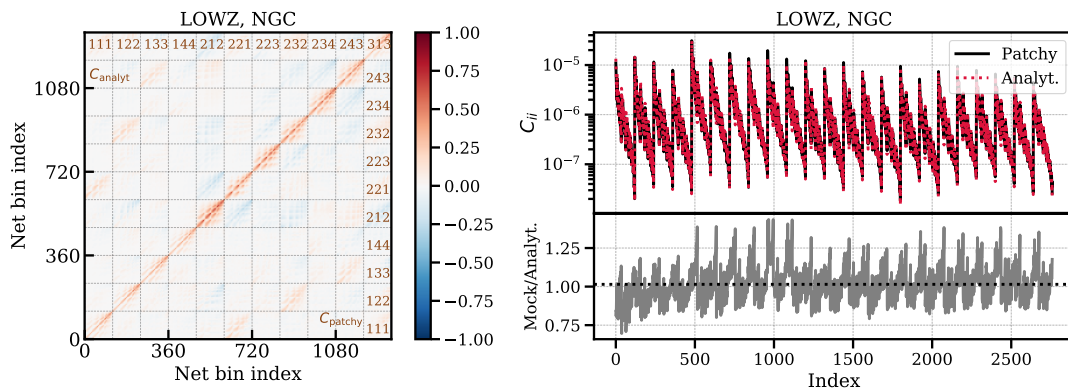
#### 4.2.2 Mock Covariance Matrix

In this work, we assume that the PATCHY mocks match well the covariance structure of the observed data. This is a reasonable assumption as the PATCHY mocks are calibrated to match the 2-point and some 3-point statistics measured from BOSS. Here we outline in detail aspects of PATCHY that could cause a mismatch between the mock covariance and the data's true covariance. In the end, we do not find any strong reason to believe that such a mismatch is present, but that is after performing a number of tests motivated by the below points.

(i) PATCHY uses approximate methods to simulate nonlinear structure formation, and underestimates small-scale power ( $s \lesssim 20 h^{-1}\text{Mpc}$ ) especially for galaxies with high stellar mass (Kitaura et al. 2016; see their Fig. 4). Lacking power at smaller scales could have two impacts on the covariance. First, it reduces the “squeezed” tetrahedra (where the side lengths from the primary are close to each other in size and hence galaxies can become very close to each other as well as the angles at the primary vertex become small). This could produce an underestimate of the covariance on small scales. Second, this lack of small-scale power also may reduce the derived covariance at all scales, as the covariance comes from averaging the product of



**Figure 7.** Comparisons of the analytic parity-odd 4PCF covariance with that from the PATCHY mocks, all for CMASS NGC and ten radial bins. For greater legibility, we show only 11 of the 23 channels used in our analysis, and we have mapped each unprimed combination of angular momenta and radial bins to a 1D index, and the same for each primed; the full set of variables and indices involved is given in Eq. (13), and the volume and number density here used are given in Table 1. *Left:* analytic correlation matrix (upper triangle) and correlation matrix from the PATCHY mocks (lower triangle). The sub-blocks represent the covariance between a pair of channels; each is  $120 \times 120$  as for ten radial bins, there are 120 radial bin combinations in each channel. These sub-blocks show how the covariance changes as the side lengths of each tetrahedron vary at fixed channel pair. The similarity between the upper and lower triangles shows that the analytic covariance captures the cross-covariance between different radial bins as well as between different channels. *Right:* diagonals of the PATCHY covariance matrix (solid black) and the analytic covariance (dashed red) for CMASS NGC. The average ratio is 1.05.

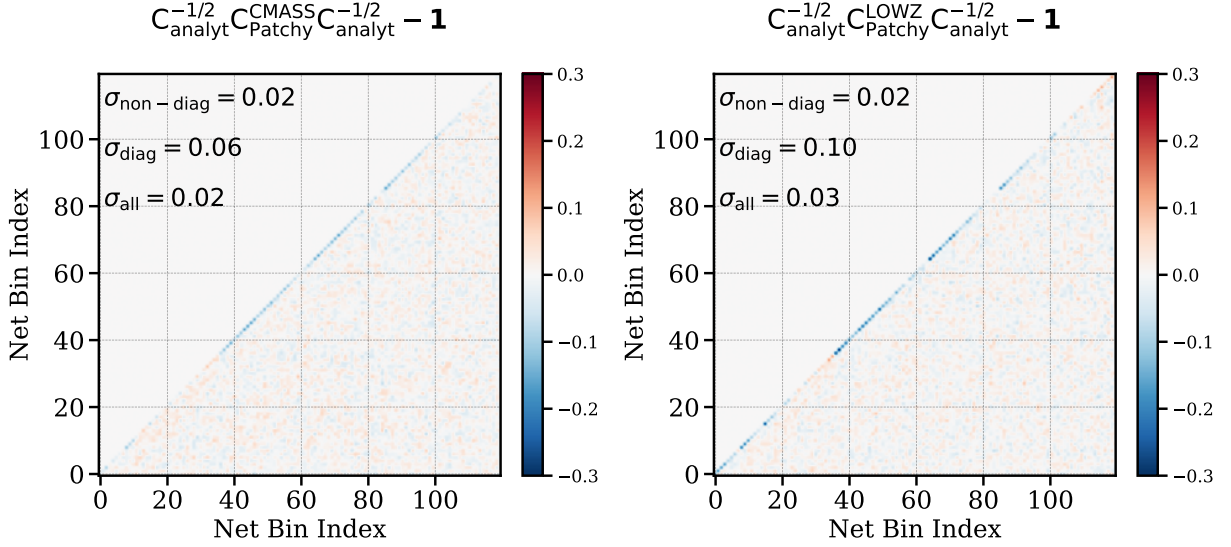


**Figure 8.** Here we show the same comparisons as made in Fig. 7 but for the LOWZ NGC. Again, in the left panel we see that the analytic and mock correlation matrices have very similar patterns. The average ratio between the mock and analytic covariance diagonals is 1.01 as indicated by the dotted black horizontal line in the lower right panel.

unprimed and primed tetrahedra over all possible separations between them, including small separations wherein points on the primed one may be close to those on the unprimed one. Indeed, this is the origin of  $s$  in the integral Eq. (12).

To assess the impact of the above, we do an analysis where we force each side to be different in length from the others by at least one full bin width; this at least addresses the first point in (i). We still find a significant detection of parity-odd 4PCF in all samples. The second point is more challenging to deal with, although we did not have a direct test against it; its impact can be inferred from the parity-even *connected* 4PCF measurement (which tracks only the contribution due to non-linear evolution). The level of agreement between PATCHY and BOSS CMASS suggests that PATCHY actually does well-capture the non-linearity on the scales relevant for our analysis.

(ii) PATCHY may not reproduce all systematics that may be present in data. We computed the 2PCF and found that the LOWZ SGC has an amplitude that is higher on average than the mean of the PATCHY mocks by  $1\sigma$ . However, it is quite consistent with the spread we see in the measured 2PCF for a set of 1,000 mocks. In addition, as discussed in Ross et al. (2016) and Kitaura et al. (2016) there are deviations between PATCHY and the BOSS 2PCF at scales greater than roughly  $100 h^{-1}\text{Mpc}$ . These 2PCF bins are highly correlated though and the actual deviation may be less severe than the visual impression. Again we found that the spread of the 2PCF of 1,000 mocks covered the BOSS data well. Nonetheless, to make our analysis robust, we also performed our parity-odd analysis with a cut on the maximum side lengths from the primary so that no side of the tetrahedron could exceed  $160 h^{-1}\text{Mpc}$ . We still found evidence for an parity-odd 4PCF consistent with what one would



**Figure 9.** Half-inverse test  $\Phi$  (Eq. 16) of the analytic covariance matrix at  $\{\Lambda, \Lambda'\} = \{111, 111\}$  for CMASS NGC (*left*) and LOWZ NGC (*right*). We show only half of the matrix in the lower triangle and indicate the standard deviation of the matrix elements in the upper triangle; the test is symmetric so there is no information lost by deleting the latter. Here we compare the analytic covariance matrix to that from 2,000 PATCHY mocks; the overall standard deviation is roughly  $1/\sqrt{N_{\text{mocks}}} \approx 0.02$ . It is expected that the standard deviation of the diagonal is double, as further discussed in the main text; we see slight deviation from this relation in both panels.

expect given the significant reduction in the constraining power in this analysis due to the much smaller number of triple-bins it permits.

We also briefly consider that the number density of the survey is not equal to the true number density of the Universe; this would be a failure of the total integral constraint. Such a failure would produce a correction to the observed power spectrum, such that a term  $P_{ic} \geq 0$  is subtracted (see equation (29) of Beutler et al. 2014). The PATCHY mocks would not have this correction, and hence have a larger power spectrum than the observed data. Thus, the covariance estimated from them would be *larger* than it should be. Since this error, if present, would cause an *over-estimation* of the covariance (and thus a spuriously decreased detection significance), and also since given BOSS’s large volume it is expected to be a small effect, we do not explore it further.

#### 4.2.3 Consistency of Parity-Even Modes

##### Impact of possible systematics

As will be discussed further in §6, distortions in the radial or the angular directions that are not captured in the mocks can mean that a mock-based covariance underestimates the true covariance. Here we study the sensitivity of the parity-even sector to such contamination. For a distortion of the radial selection function, we apply the contamination directly to the mocks and find their mean detection significance of a connected-even parity 4PCF (when computed with an undistorted covariance, as would be the case if the data had an unaccounted-for change) is shifted by  $\sim 0.6\sigma$  for ten radial bins and  $\sim 2.4\sigma$  for eighteen radial bins. We also explore an angular contamination as described in §6.1.5; this increases the even-parity detection significance difference between data and mean of the mocks to  $0.9\sigma$  for the ten bin and  $1.8\sigma$  for the eighteen bin.

##### Self-calibration of covariance

Our detection significance relies on correctly estimating the covariance. Since the PATCHY mocks have no intrinsic parity-violating mechanism, they cannot be used to check our pipeline regarding signal, but only to estimate the covariance. Yet it is possible that the mocks do not contain all the systematics present in the real data, and this lack could result in mis-estimation of the covariance.

However, any unknown data-only systematic might well impact the parity-even modes as well as the parity-odd modes. Hence, we can see if the mocks and data are consistent in the parity-even sector at the signal level, where both do have a signal due to non-linear evolution. If we see that the even-parity signal from the data is consistent with that from the mocks, this suggests that there are unlikely to be unaccounted-for systematics. It also would argue that we have reasonably estimated the even-sector covariance, as an underestimate of it would show up as a severe tension between the signal in the mocks and in

**Table 1.** Summary of the datasets and binning schemes used. We give the effective redshift  $z_{\text{eff}}$  and the effective volume  $V_{\text{eff}}$  and number density  $\bar{n}$  derived by fitting our analytic covariance to that from the mocks. The effective volume and number density are different for each number of radial bins because we re-fit the covariance for each radial binning scheme.

Bins	$z_{\text{eff}}$	CMASS		LOWZ	
		NGC	SGC	NGC	SGC
		0.57		0.32	
18	$V_{\text{eff}}^{\text{eff}}$ [ $h^{-1}\text{Gpc}$ ] <sup>3</sup>	2.50	0.79	0.31	0.12
	$\bar{n} \times 10^4$ [ $h^3\text{Mpc}^{-3}$ ]	1.4	1.4	2.0	2.0
10	$V_{\text{eff}}^{\text{eff}}$ [ $h^{-1}\text{Gpc}$ ] <sup>3</sup>	1.71	0.54	0.23	0.09
	$\bar{n} \times 10^4$ [ $h^3\text{Mpc}^{-3}$ ]	1.8	1.8	2.6	2.6
6	$V_{\text{eff}}^{\text{eff}}$ [ $h^{-1}\text{Gpc}$ ] <sup>3</sup>	1.36	0.44	-	-
	$\bar{n} \times 10^4$ [ $h^3\text{Mpc}^{-3}$ ]	2.2	2.2	-	-

**Table 2.** The detection significances using the analytic covariance for the split sky (into NGC and SGC) and joint sky (NGC + SGC). The latter treats a given channel and radial bin combination in N and S as two separate degrees of freedom, *i.e.* we assume vanishing covariance between N and S. All of the quoted detection significances are from comparing the  $\chi^2$  of the data to a Gaussian fit to the histogram of the mocks'  $\chi^2$ ; see lower rightmost panels of Figs. 10-13 (ten and eighteen bin results) and of Fig. E1 (six bin results, CMASS only).

Bins	$\sigma_{\text{eff}}$ , Detection Significance	CMASS		LOWZ	
		NGC	SGC	NGC	SGC
18	split	4.7	5.4	2.5	2.0
	joint	7.1		3.1	
10	split	3.3	2.6	1.8	1.9
	joint	4.0		2.7	
6	split	-0.7	1.2	-	-
	joint	0.4		-	

the data. Since our procedures for obtaining the even and odd-sector covariances are exactly the same, such a finding would in turn build confidence that our odd-sector covariance is correct.

Furthermore, as a highly conservative approach, we can ask what factor the covariance in the even-sector would need to be rescaled by to force the data to agree within *e.g.*  $3\sigma$  or  $1\sigma$  with the mocks. We can then apply this rescaling to our odd-parity covariance and ask how it impacts our detection significance in the odd sector. We emphasize that this is a highly conservative check, not a correction procedure; it is perfectly possible that the even-parity data signal is *e.g.*  $3\sigma$  different from that in the mocks just by chance, without implying that there is any error that should be corrected in either sector. Thus, while we summarize the results of this idea in Table 4, one should not take these rescalings too literally.

For CMASS with ten radial bins, we found excellent agreement between the data and mock distribution both using the analytic covariance and the compressed method (also see Philcox et al. (2021b), where a slightly different sample selection was applied). Therefore, we do not present any rescaling factor for the detection significance for the parity-odd modes.

For CMASS with eighteen radial bins, we found a discrepancy in the even-parity sector between the data and the mock distribution of  $4.9\sigma$  when using the analytic covariance. Rescaling the even covariance such that the data and mocks agree at the  $3\sigma$  level, we then propagate the same factor to the parity-odd measurement. We also repeat this procedure to enforce agreement at the  $1\sigma$  level in the even sector. We summarize these results in Table 4.

Regarding the discrepancy found in the eighteen-bin test, we ascribe it to the possible breakdown of the GRF assumption of the analytic covariance as one goes to smaller scales, especially for the even sector, as discussed in §4.2. In particular, the parity-even covariance requires including the leading higher-order covariance and connected estimator contribution. These contributions are naturally absent when estimating the parity-odd covariance. Further, when using the compressed method with  $N_{\text{eig}} = 800$ , the discrepancy is reduced to  $1.3\sigma$ .

**Table 3.** Correlation coefficients between NGC and SGC for CMASS and LOWZ samples; see §5.2 and Figs. 15-17. We do not see significant cross-correlation between NGC and SGC in either binning scheme or sample; the possible reason is further discussed in the main text, as is the reason we cannot perform a CMASS × LOWZ analysis.

		CMASS		LOWZ	
Bins	Correlation	NGC	SGC	NGC	SGC
18	$r_p$	0.002		-0.017	
	$p$ -value	0.752		0.017	
10	$r_p$	-0.014		0.000	
	$p$ -value	0.460		0.985	

**Table 4.** Here we summarize the rescaled detection significance in the parity-odd modes by forcing agreement in the parity-even sector at the levels indicated in the leftmost column. These results are for CMASS and eighteen radial bins; see §4.2.3. We explore this rescaling both for our analytic covariance analysis and our compressed analysis with 800 eigenvalues; in the last line, no rescaling is required to attain  $3\sigma$  consistency in the even sector. Generally, it is notable that an odd detection more or less remains in all cases even after rescaling.

Consistency in Even-Parity Sector	CMASS 18 bins	Rescaling Factor	Rescaled Odd Detection Significance
1 standard deviation	Analytic Covariance	0.88	$2.0\sigma$
	$N_{\text{eig}} = 800$	0.98	$4.0\sigma$
3 standard deviation	Analytic Covariance	0.94	$4.6\sigma$
	$N_{\text{eig}} = 800$	-	$4.4\sigma$

## 5 RESULTS ON CMASS AND LOWZ DATA

Here, we first compute the significance of our observed parity-odd 4PCF using the methods briefly outlined in §2, and then explore the cross-correlation of NGC and SGC in each sample.

### 5.1 Statistical Significance

We quantify the statistical significance of the parity-violating amplitudes using three approaches: a direct analysis harnessing the analytic covariance calibrated against PATCHY (§5.1.1), a compressed analysis using the analytic covariance just to select a reduced basis with few enough degrees of freedom that the mocks can then provide the covariance (§5.1.2), and finally, by lowering  $\ell_{\text{max}}$  enough so that we may use the direct approach but with covariance from the mocks (§5.1.5).

#### 5.1.1 Direct Analysis

With the analytic covariance, adjusted in density and volume to best fit the covariance given by the mocks, we evaluate

$$\chi^2 = \left( \hat{\zeta} - \zeta_{\text{model}} \right)^T \mathbf{C}^{-1} \left( \hat{\zeta} - \zeta_{\text{model}} \right). \quad (17)$$

where  $\hat{\zeta}$  and  $\zeta_{\text{model}}$  are respectively the various parity-violating amplitudes,  $\hat{\zeta}_{\ell_1 \ell_2 \ell_3}(r_1, r_2, r_3)$  and  $\zeta_{\text{model}}$  is a model for them. To investigate the null hypothesis that there is no parity-violation, we set all elements of  $\zeta_{\text{model}}$  to zero.  $\mathbf{C}^{-1}$  is the inverse covariance matrix. If the underlying data have Gaussian behavior with  $N_d$  degrees of freedom, the resulting distribution is the  $\chi^2$  distribution

$$f(\chi^2) = \frac{1}{2^{N_d/2} \Gamma(N_d/2)} (\chi^2)^{N_d/2} e^{-\chi^2/2}. \quad (18)$$

The adequacy of the covariance obtained by fitting the analytic covariance to the data sets can be assessed by examining the distribution of  $\chi^2$  values from the mocks when run through our analysis, and seeing if it matches the predicted  $\chi^2$  distribution.<sup>9</sup>

#### 5.1.2 Compressed Analysis

The alternative means of obtaining an invertible covariance matrix by reducing the dimensionality of the problem, the “compressed” analysis, is described in §2. This scheme was introduced by Scoccimarro et al. (1999) to quantify detection significance,

<sup>9</sup> We expect that radially-binning the 4PCF amplitudes results in a multivariate Gaussian distribution on each triple-bin. Thus a  $\chi^2$  distribution is appropriate here.

and the same method was applied to the even-parity connected 4PCF analysis of [Philcox et al. \(2021b\)](#). We diagonalize the analytic covariance, writing  $C = ODO^{-1}$  where  $D$  is diagonal and  $O$  is an orthogonal matrix (we note that  $O^{-1} = O^T$ ).

The columns of  $O$  then provide an eigenbasis for the analytic covariance. We would like to select the subset of these eigenvectors that will be most useful in detecting a signal. If one has a model for the expected signal, one can rank the eigenvectors by their signal-to-noise, and choose  $N_{\text{eig}}$  of them with the best S/N. In the absence of an anticipated signal, we choose the  $N_{\text{eig}}$  eigenvectors associated with the *smallest* eigenvalues. These eigenvectors have the lowest noise. The analysis is then restricted to the space spanned by these eigenvectors. With  $N_{\text{eig}} \ll N_{\text{mocks}}$ , the covariance matrix for this restricted analysis can be determined entirely from the mocks, while retaining the domain with the greatest information.

While this procedure provides an invertible covariance matrix, the inverse of the mock covariance matrix is unbiased only when  $N_{\text{eig}} \ll N_{\text{mocks}}$ , where here  $N_{\text{eig}}$  is the number of degrees of freedom, *i.e.* plays the role of  $N_d$  in §5.1.1. Consequently, the  $\chi^2$ -distribution we used in §5.1.1 is modified. To distinguish this distribution from the  $\chi^2$  distribution, we refer to the variable  $T^2$ , defined analogously to  $\chi^2$ , as

$$T^2 = \left( \hat{\zeta} - \zeta_{\text{model}} \right)^T C^{-1} \left( \hat{\zeta} - \zeta_{\text{model}} \right). \quad (19)$$

whose distribution is ([Sellentin & Heavens 2016](#))

$$f(T^2) = \frac{\Gamma(N_{\text{mocks}}/2)}{\Gamma(N_{\text{eig}}/2) \Gamma([N_{\text{mocks}} - N_{\text{eig}}]/2)} \frac{(N_{\text{mocks}} - 1)^{-N_{\text{eig}}/2} (T^2)^{N_{\text{eig}}/2 - 1}}{(T^2/(N_{\text{mocks}} - 1) + 1)^{N_{\text{mocks}}/2}}. \quad (20)$$

This reduces to the  $\chi^2$  distribution for  $N_{\text{mocks}} \rightarrow \infty$ .

Since the data from CMASS and LOWZ in the NGC and SGC are statistically independent given the large physical separations of both NGC and SGC and of CMASS and LOWZ, we can simply add their values of  $\chi^2$  or  $T^2$ , while increasing the number of degrees of freedom correspondingly.

### 5.1.3 Detection Significances from Direct & Compressed Methods

Fig. 10 displays the  $T^2$  or  $\chi^2$  distributions for CMASS (both NGC and SGC) for the two approaches described above (direct and compressed). The detection significance changes as we vary the number of eigenvalues, but this is expected: as  $N_{\text{eig}}$  rises, more information is added (unless the S/N is pathological), but at the same time the potential bias of the inverse covariance is increasing. This latter can be assessed by looking for when the distribution of the mocks'  $T^2$  values begins to deviate from the  $T^2$  distribution expected (Eq. 20). Fig. 11 shows the same information but for LOWZ, which generally shows comparable detection significances compared to CMASS both at each number of eigenvalues in the compressed analysis, and in the direct analysis.

### 5.1.4 Using Finer Binning

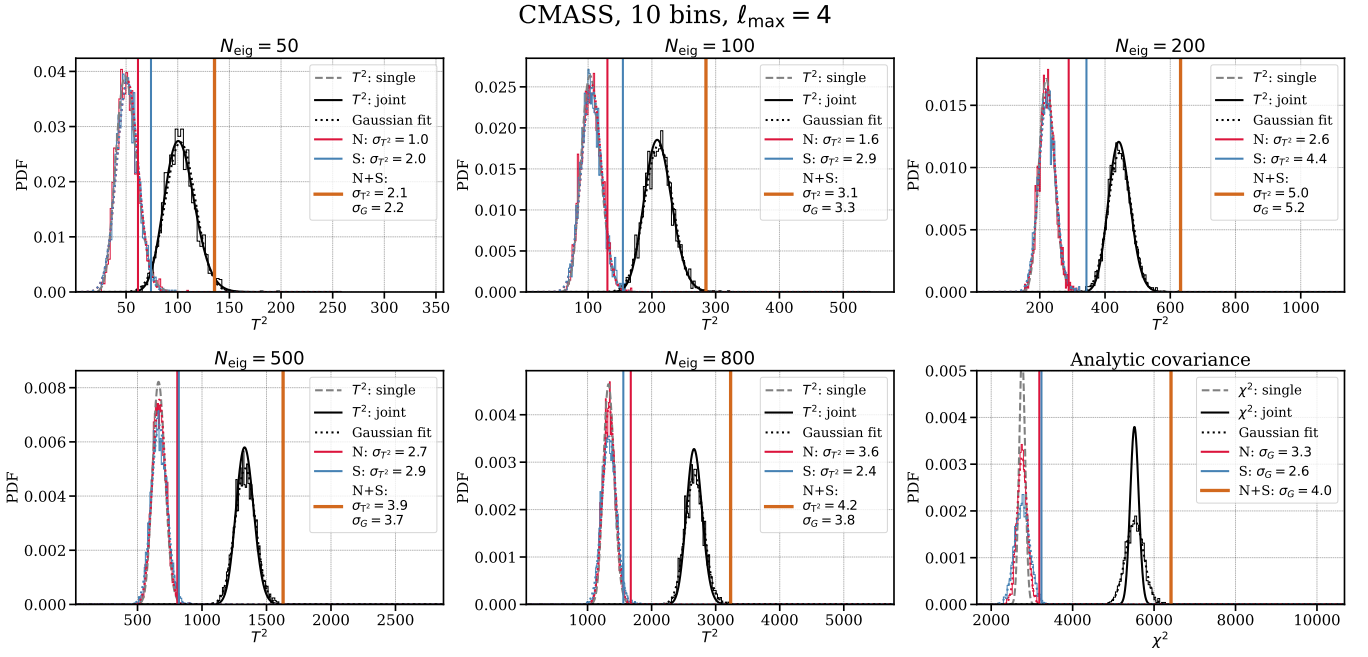
Since two cancelling contributions can occur from a single tetrahedron when they fall into the same radial bins (§3.2), increasing the number of bins could increase the power of the analysis, at the cost of further enlarging the covariance matrix. So motivated, we now use eighteen linearly-spaced radial bins from  $r_{\text{min}} = 20 h^{-1}\text{Mpc}$  to  $r_{\text{max}} = 164 h^{-1}\text{Mpc}$ , leading to a bin width of  $8 h^{-1}\text{Mpc}$ , roughly double the radial resolution of our previous analyses. We refit the volumes and number densities for the covariance matrices of CMASS NGC and SGC and LOWZ NGC and SGC as in §4.2, and these numbers are in Table 1. Other than these new inputs to the covariance matrix, all else is the same in our analysis, and we display the results in Fig. 12 (CMASS) and Fig. 13 (LOWZ). Generally, for each sample, the detection significance rises at each  $N_{\text{eig}}$  relative to the ten-bin analysis, and the same is true for the “direct” significances (where the analytic covariance were used). This increase is expected given our arguments about “internal cancellation”. We summarize the significances in Table 2.

### 5.1.5 Reduced- $\ell_{\text{max}}$ Analysis

A pure-mock covariance is ideal deriving the detection significance, yet the high dimensional data vector leads to a non-invertible sampling matrix due to the limited number of mocks. Reducing the number of channels considered makes it possible to use a covariance matrix taken directly from the mocks, with no need for the analytic one.

For  $\ell_{\text{max}} = 1$ , there is just one parity-odd channel, and for  $\ell_{\text{max}} = 2$  there are four. Fig. 14 shows the detection significances for CMASS and LOWZ with  $\ell_{\text{max}} = 1$  in the first and second rows and  $\ell_{\text{max}} = 2$  in the third and fourth rows. For  $\ell_{\text{max}} = 1$  the histograms of the mocks'  $T^2$  values agree well with the expected  $T^2$  distribution, both for the compressed and the pure mock-based covariance approaches. However the histogram of the mocks'  $\chi^2$  values from the direct approach with analytic covariance deviates from the expected  $\chi^2$  distribution, which indicates that the analytic covariance is imperfect. Nonetheless, all three methods show consistent detection significance. For  $\ell_{\text{max}} = 2$ , all three methods are still consistent. However, we notice the changes in detection significance of CMASS and LOWZ are different when going from  $\ell_{\text{max}} = 1$  to  $\ell_{\text{max}} = 2$ . We





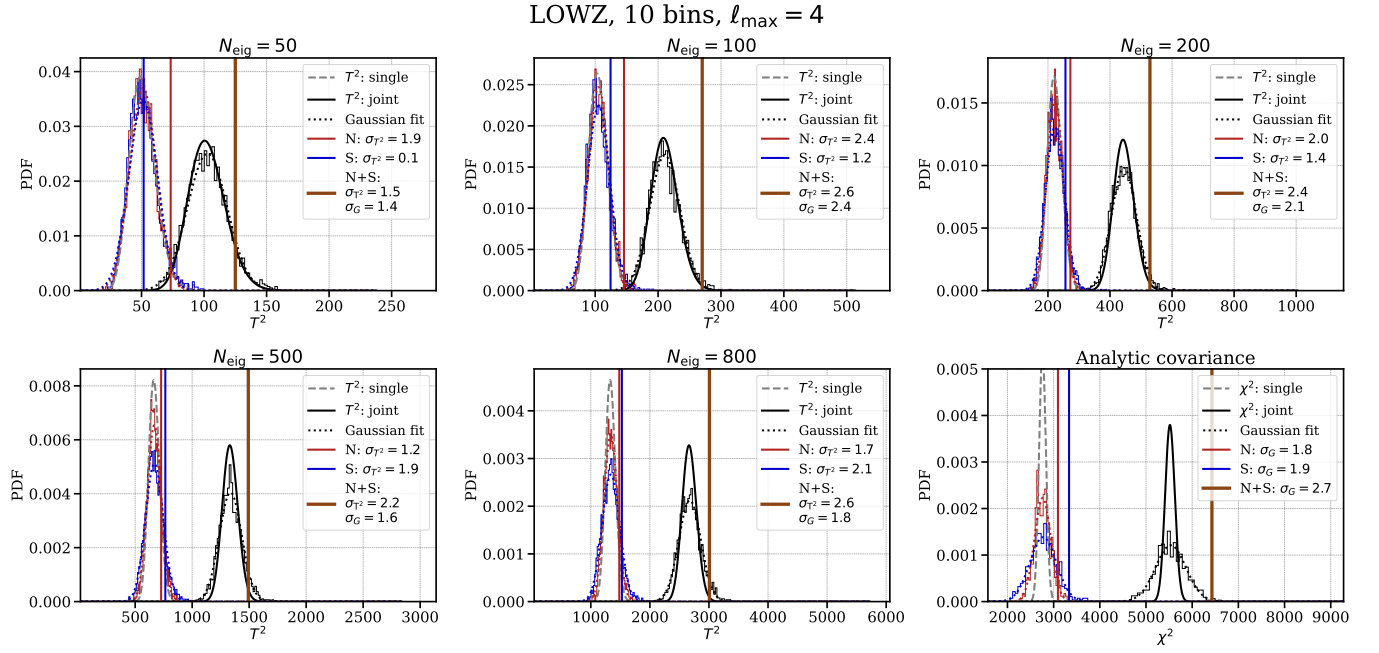
**Figure 10.**  $T^2$  and  $\chi^2$  distributions for CMASS data using **ten** radial bins. The first five plots are obtained using the data compression scheme of §5.1.2 with different numbers of eigenvalues  $N_{\text{eig}}$  as indicated in the plot titles. If there is no parity-breaking, we expect a  $T^2$  distribution (Eq. 20) with  $N_{\text{eig}}$  degrees of freedom (solid black). The vertical lines in the first five plots are the  $T^2$  values from CMASS for the NGC (red), SGC (blue), and joint NGC and SGC (black). For comparison, we show histograms of the  $T^2$  values from the PATCHY mocks for each case. We interpret the increasing deviations of the mocks’ histograms from the  $T^2$  distribution as  $N_{\text{eig}}$  rises as evidence that the covariance becomes less well-determined for larger  $N_{\text{eig}}$ . This is expected as for  $N_{\text{eig}} = 800$ , there are only about three mocks per degree of freedom (we use 2,000 mocks). In the lower rightmost plot, we use the direct approach of §5.1.1, which employs the analytic covariance. Here, the expected distribution if there is no parity-breaking is a  $\chi^2$  distribution (Eq. 18; solid black). We show a histogram of the mocks’  $\chi^2$  values for comparison (orange) as well as a Gaussian fit to it (dot-dashed). This histogram is wider than the predicted  $\chi^2$  distribution (which the mocks should match as they have no parity-breaking). We interpret this as an indication that the analytic covariance may not be unbiased. To be conservative, we compute a detection significance  $\sigma_G$  by comparing the data  $\chi^2$  values to the width of the Gaussian fit to the mocks. In fact we quote  $\sigma_G$  in all panels above, but it generally agrees with that computed from the  $T^2$  distribution,  $\sigma_{T^2}$ . We do not quote a detection significance computed with respect to the  $\chi^2$  distribution (as would naively be appropriate in the lower rightmost panel) because the distribution of mocks is clearly mismatched with a  $\chi^2$  distribution.

suspect this is due to the difference in the two samples, given their redshifts, number density, which we will leave for future exploration. After all, this reduced  $\ell$  analysis serves as a robustness check of the pure analytic covariance matrix approach.

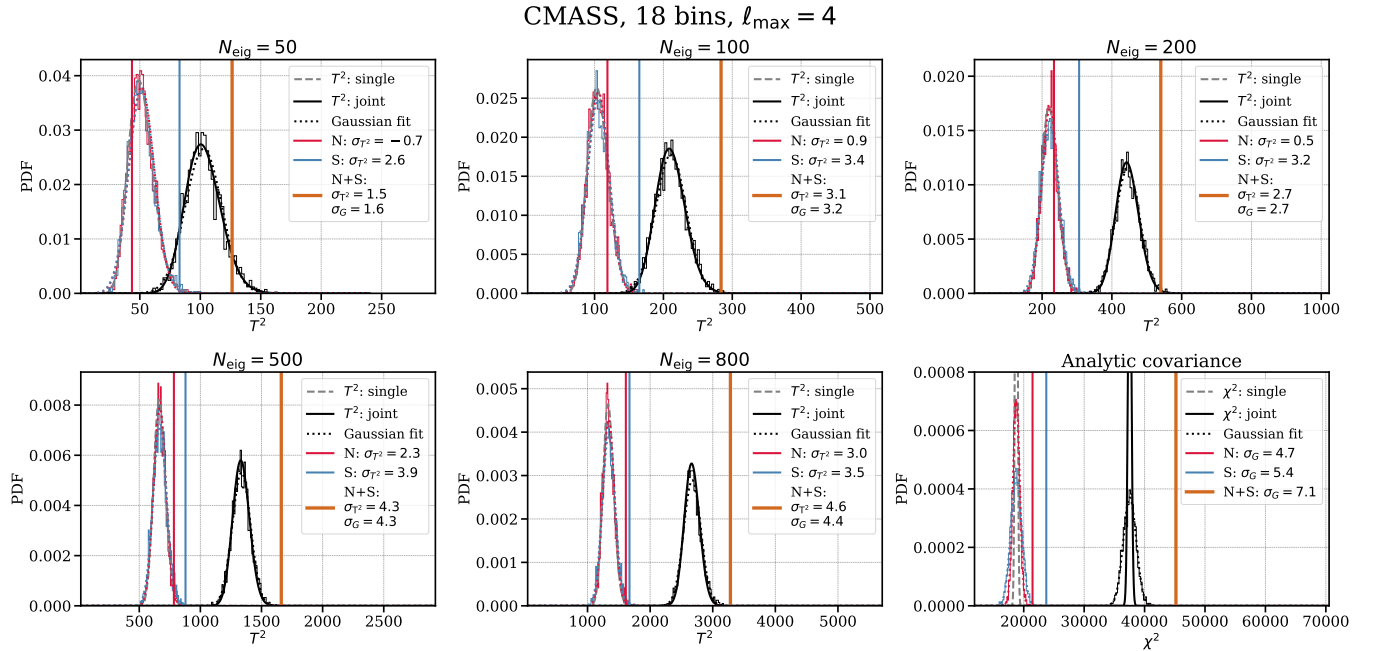
## 5.2 Cross-Correlation Between NGC and SGC in Each Sample

A cosmological parity-odd correlation is expected to appear isotropically when decomposed using our choice of basis function. This would manifest itself as a correlation between the signals in the SGC and NGC. Noise in the observed 4PCF, would however not be correlated between hemispheres or between samples, and would reduce any underlying cross-correlation. We compute the Pearson correlation coefficient  $r_p$  between the signals in the SGC and NGC, with some modifications which we now describe. The naive Pearson coefficient does not incorporate covariance between different points *within* each data stream being cross-correlated, nor does it account for varying errors from point to point. For instance, two neighboring triple-bins within a given channel in NGC are likely to be much more correlated with each other than two very different triple-bins. Thus, if it happens that one such triple-bin is highly correlated with its analog in SGC, then the neighboring combination in NGC is also likely correlated with the neighboring combination in SGC. This is not unphysical, but it should be taken as less strong evidence for correlation than if two highly independent triple-bins in NGC *each* showed strong correlation with their analogs in SGC. The Pearson coefficient would not distinguish. This motivates us to rotate our measured data streams from NGC and from SGC to a basis where each data vector (*within* the NGC or SGC) is independent; we thus work in the eigenbasis of the analytic covariance matrix. This procedure will be affected by any issues with the covariance matrix estimation, which is present for both analytic or mock covariance.

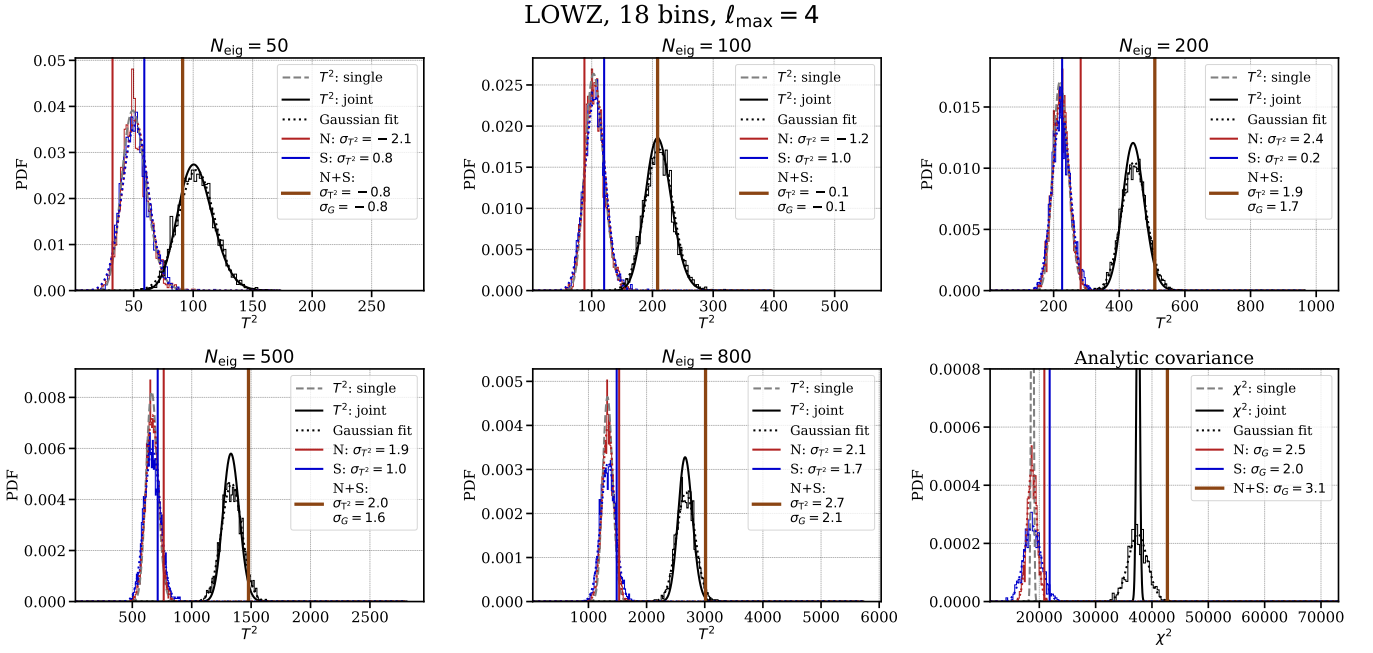
Another issue with naive Pearson correlation is that some channels and triple-bins have larger statistical errors than others. They may therefore more easily be outliers and drive a spurious Pearson correlation. To correct for this, once we are in the



**Figure 11.** Similar to Fig. 10 but for LOWZ; again using ten radial bins. As for CMASS, in the compressed approach, the detection significance first increases with increasing number of eigenvalues and then drops. In general the detection significance is comparable to CMASS for both compressed and direct analyses and for split and joint sky. Again in the direct method we see noticeable deviation of the mocks' histogram from a  $\chi^2$  distribution; this again suggests that the analytic covariance may not be unbiased. We also see noticeable deviation at  $N_{\text{eig}} = 800$  (and even at  $N_{\text{eig}} = 500$ ), to a much greater extent than for the same  $N_{\text{eig}}$  values in CMASS (*cf.* Fig. 10).



**Figure 12.** Similar to Fig. 10 but for CMASS with eighteen bins. Using finer binning we find even higher detection significance, both in our compressed and direct analyses (*cf.* Fig. 10). Unlike the ten-bin CMASS analysis, which peaked at  $N_{\text{eig}} = 200$ , here the detection significance monotonically increases with  $N_{\text{eig}}$ . This behavior is likely due to how the covariance matrix structure and signal-to-noise in each channel and bin combination balance with the penalty paid for adding more degrees of freedom if they do not contain much additional S/N. Given that this balance can change depending on the radial binning, we would not have expected that the behavior of the ten-bin and eighteen-bin compressed analyses with  $N_{\text{eig}}$  would be uniform.



**Figure 13.** Similar to Fig. 11 but for LOWZ with **eighteen** bins. As is the case for CMASS, the detection significance increases when using the finer binning (*cf.* Fig. 11). However, the increase is more moderate than the increase for CMASS. The detection significance rises monotonically with  $N_{\text{eig}}$ , in agreement with the trend for the ten-bin LOWZ analysis of Fig. 11.

“independent” basis as above, taking the square-root of the associated covariance matrix eigenvalue as its error-bar, we divide each value by its error-bar.

For the above procedure, if we had the true covariance, rotating into the true eigenbasis would preserve any correlations perfectly. However, as discussed in §4.2 the analytic covariance is likely imperfect. We believe it to be close enough to the true covariance to preserve correlations within a given sample (CMASS or LOWZ) after rotation, but that trying to rotate into the different eigenbases implied by the covariances of the two different samples and then cross-comparing would not be robust. In particular, the different number densities in CMASS and LOWZ (see Table 1) will non-trivially impact the eigenvectors for each, likely randomizing any underlying correlations. Given this issue (explored more fully in Appendix C), we do not report results of a CMASS cross LOWZ analysis in this work.

Having outlined how we render our measured 4PF coefficients suitable for a cross-correlation analysis, we now define the statistic we use, the Pearson correlation coefficient  $r_p$ . It is<sup>10</sup>

$$r_p = \frac{\sum_{i=1}^{N_d} (d_{1,i} - \langle \mathbf{d}_1 \rangle_{\text{all}}) (d_{2,i} - \langle \mathbf{d}_2 \rangle_{\text{all}})}{\sqrt{\sum_{i=1}^{N_d} (d_{1,i} - \langle \mathbf{d}_1 \rangle_{\text{all}})^2 \sum_{i=1}^{N_d} (d_{2,i} - \langle \mathbf{d}_2 \rangle_{\text{all}})^2}}, \quad (21)$$

where  $d_i$  is the  $i^{\text{th}}$  element of the 4PCF data vector  $\mathbf{d}_i$  with  $i = 1$  for the NGC and  $i = 2$  for the SGC. Both are in the decorrelated, normalized basis for their respective hemisphere.  $\langle \dots \rangle_{\text{all}}$  denotes the average over the elements of the vector, and  $N_d$  is the number of degrees of freedom.

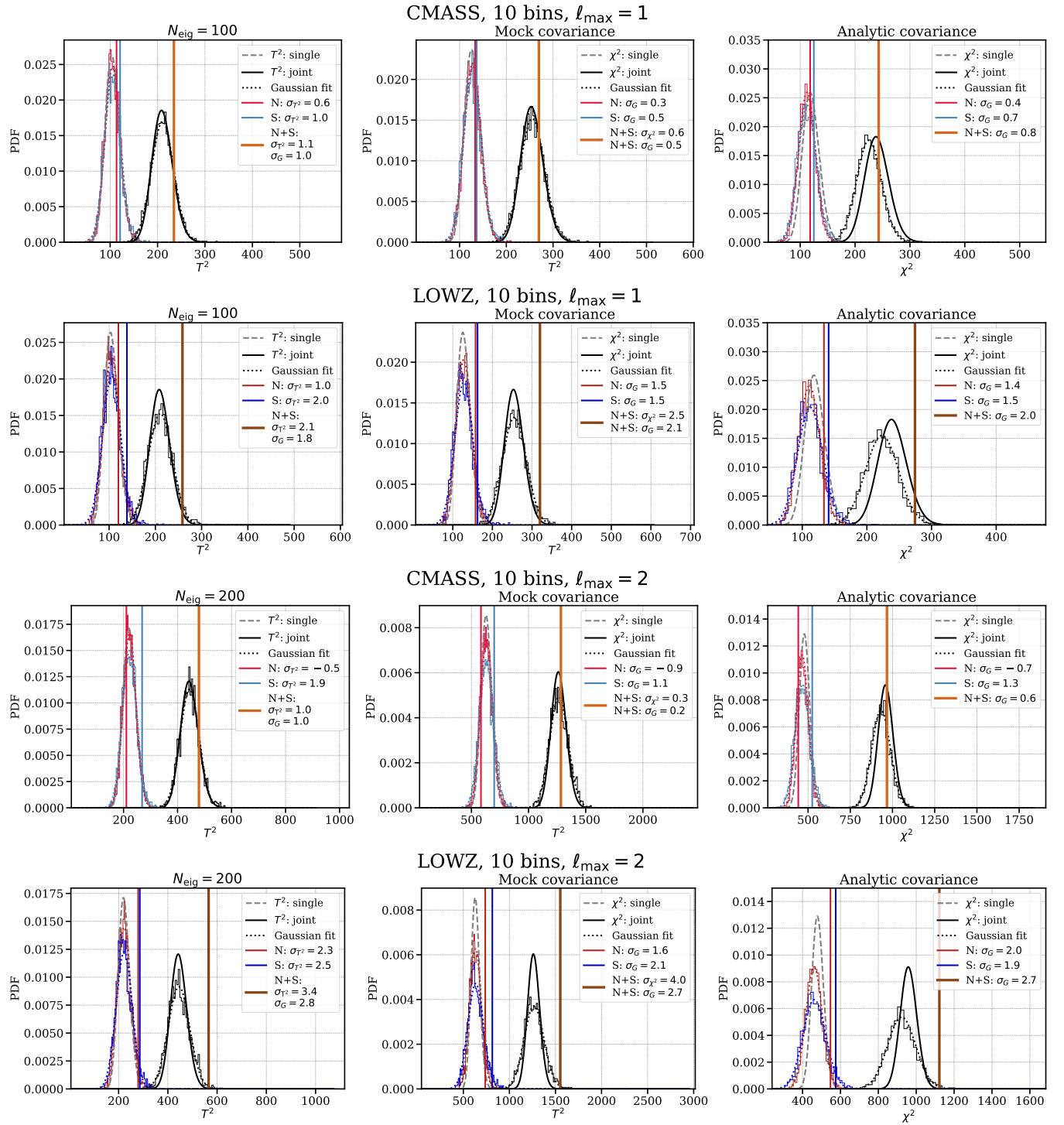
The PDF for the Pearson correlation coefficient if there is no correlation is (Kenney 1947; Hotelling 1953)

$$f(r_p) = \frac{(1 - r_p^2)^{(N_d-4)/2}}{B(1/2, (N_d - 2)/2)}, \quad (22)$$

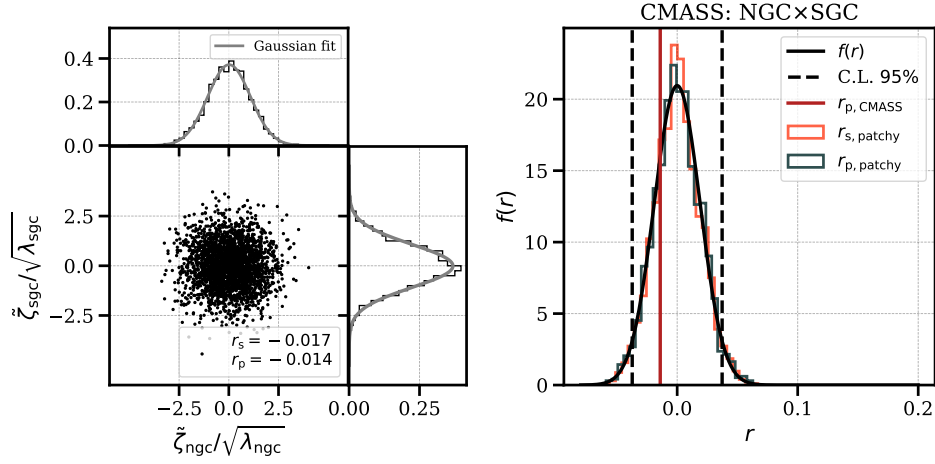
where B is the Beta function.

Fig. 15 shows the correlation analysis results for CMASS using ten bins. The left panel is a scatter plot of the (decorrelated and scaled) data for CMASS NGC and SGC. The histograms in the extended panels can be well-fit by a Gaussian, showing this assumption of the Pearson coefficient is satisfied. We find no statistically significant correlation at the 95% confidence level. We also report the Spearman’s rank coefficient,  $r_s$ , in the Figure (see footnote 10). The right panel shows the PDF for  $r_p$  (black curve; see Eq. 21) if the two datasets are both normally-distributed. As noted above, a correlation at some level

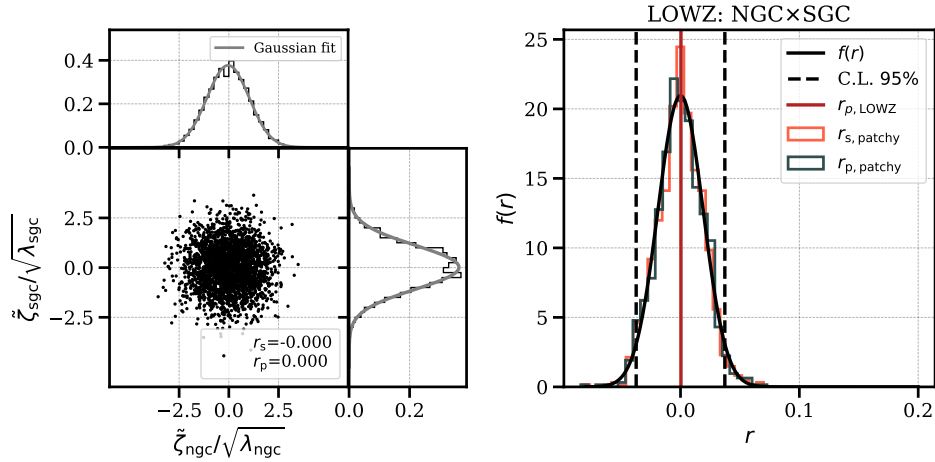
<sup>10</sup> The Pearson correlation coefficient is best for quantifying bivariate normally-distributed data. In the scaled, decorrelated basis, the data vectors indeed satisfy this assumption. We also quote the Spearman’s rank correlation coefficient, which identifies correlations between any two monotonic datasets by searching for correlations in the space of their ranks. It is more general than Pearson but also more difficult to manipulate analytically (as we do with the Pearson correlation coefficient in Appendix B). The results of both are consistent.



**Figure 14.** Reduced  $\ell_{\max}$  analysis for ten radial bins. We use three different methods of obtaining the covariance matrix but always with  $\ell_{\max}$  restricted to 1 or 2. The left column uses the compression scheme. The central column uses the covariance matrix obtained directly from the mocks, which is possible for this reduced  $\ell_{\max}$  analysis. The third column use the covariance obtained from adjusting the analytical covariance to fit, as well as possible, the purely mock-based covariance. The solid curves show the expected distributions for  $T^2$  or  $\chi^2$  while the histograms show the distributions obtained from the mocks. The solid lines indicate results from the data. There is overall good agreement in detection significance between the purely mock-based covariance (central column) and the analytic covariance (right column).



**Figure 15.** *Left:* Scatter plot for the 4PCF with **ten bins** measured from CMASS in the decorrelated basis (described in §5.2) and scaled by the covariance matrix eigenvalues, with both Pearson ( $r_p$ ) and Spearman ( $r_s$ ) correlation coefficients given in the legend. The extended panes in the left panel display the histogram overplotted with a Gaussian fit. *Right:* PDF  $f(r)$  (black curve; Eq. 22) for the Pearson correlation coefficient and the Pearson and Spearman correlation coefficients obtained from PATCHY (orange and blue histograms). The 95% confidence level (dashed vertical lines) corresponds to an  $r_p = 0.037$ . Our Pearson  $r_p$  of -0.014 corresponds to a  $p$ -value of 0.460, so there is no statistically significant correlation detected.



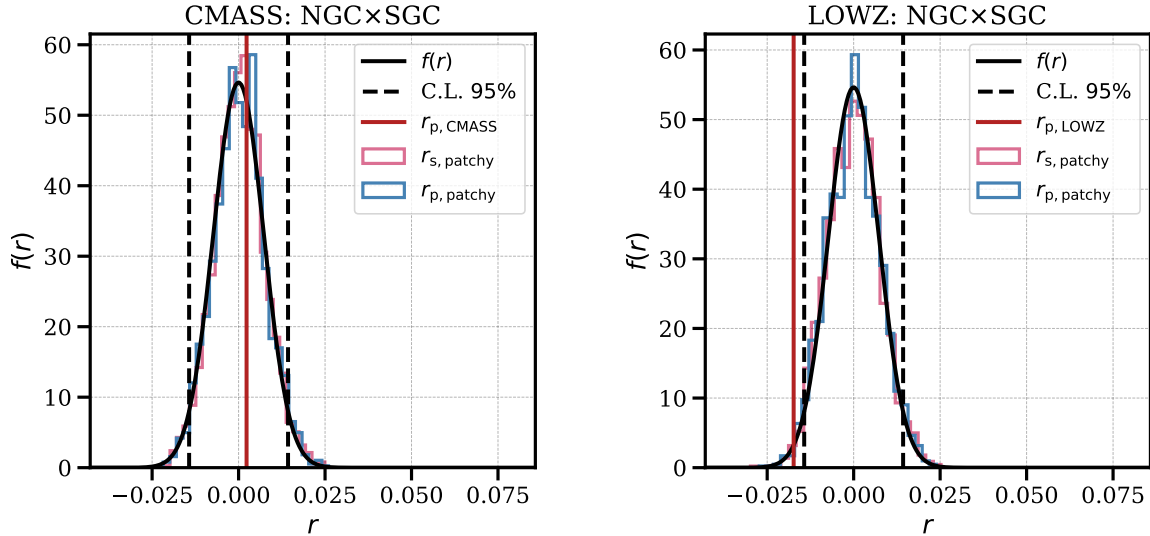
**Figure 16.** Same as Fig. 15 but for LOWZ, again with **ten bins**. Again we detect no statistically significant correlation.

might be expected if the signal were of cosmological origin. Meanwhile, there are some systematics where such a correlation would *not* be expected, and the same goes for some instrumental artifacts. On the other hand, artifacts at the telescope *could* correlate NGC and SGC.

Fig. 16 is similar to Fig. 15 but for the LOWZ sample. Fig. 17 shows the correlation analysis results using eighteen bins, where the data here also follow a Gaussian distribution (although we do not show them explicitly). In the right panel, for LOWZ, the correlation coefficient is a negative number at more than 95% confidence. However, we also find that the correlation coefficients are sensitive to the covariance matrix. Changing the number density used in the analytic covariance matrix by 10% can flip the LOWZ correlation coefficient to a positive number (but one consistent with zero). Overall, we thus conclude that there is no evidence for a positive correlation between the N and S for LOWZ or for CMASS when using eighteen bins. Our correlation study results for ten and eighteen bins and LOWZ and CMASS are summarized in Table 3.

### 5.2.1 Linear Toy Model in the Eigenbases to Explore $\chi^2$ and $r_p$

To gain intuition for the high detection significance and the low correlation between the NGC and SGC in the LOWZ, we construct a toy model. Henceforth we use the Pearson correlation coefficients and  $\chi^2$  obtained from ten bins results. The purpose is not to constrain real physics, but simply to gain intuition for whether the values of  $\chi^2$  and  $r_p$  found in our data are plausibly consistent with each other.



**Figure 17.** PDF of  $f(r_p)$  for **eighteen bins**; the left panel displays results using CMASS, the right using LOWZ. The correlation coefficient for CMASS is consistent with zero, but we see a negative correlation for LOWZ; this is further discussed in the main text.

We assume the signal is given by a toy model that is linear in the eigenbases of the analytic covariance matrices parametrized by a slope  $m$  and an intercept  $b$  as  $s = mx + b$ , where  $x$  is the *index* of each eigenvalue. We produce two sets of fake data. Each is drawn randomly from a Gaussian distribution, the mean of which is the linear model, and the width of which is the square-root of the eigenvalue of the analytic covariance matrix for either CMASS or LOWZ as appropriate.

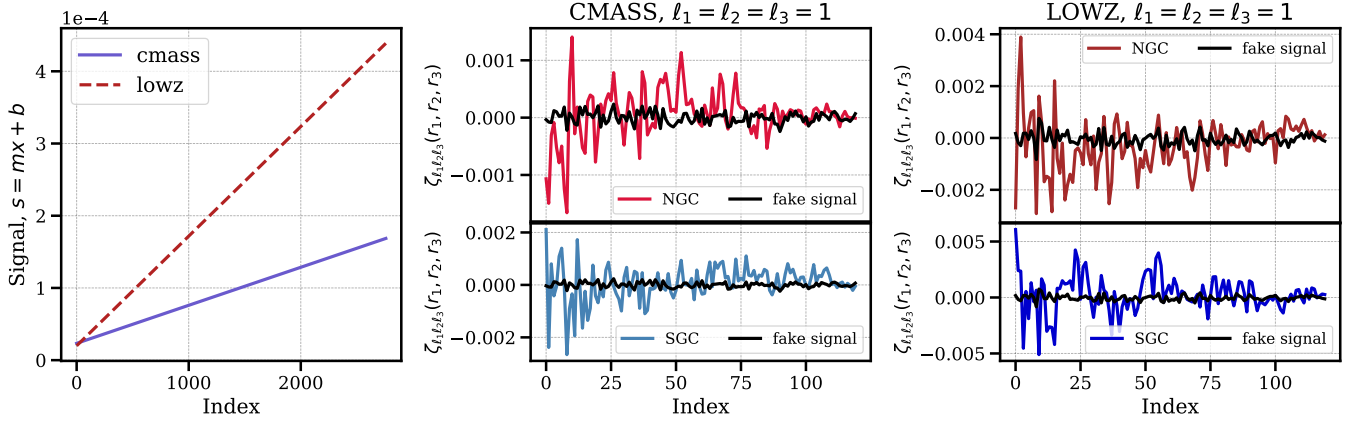
We separately maximize the log-likelihood for the data vectors  $\mathbf{d}_{\text{cmass}} = \{r_{p,\text{cmass}}, \chi_{\text{cmass}}^2\}$  and  $\mathbf{d}_{\text{lowz}} = \{r_{p,\text{lowz}}, \chi_{\text{lowz}}^2\}$ , respectively. If there is a true primordial signal, we might expect that the parity-odd 4PCF in LOWZ has overall amplitude  $2\times$  that in CMASS, folding in the difference in growth rates and linear biases. Therefore we search over different ranges in the slope and the intercept for each:  $m_{\text{cmass}} \in [1 \times 10^{-8}, 2 \times 10^{-7}]$ , the intercept  $b_{\text{cmass}} \in [1 \times 10^{-5}, 5 \times 10^{-4}]$  for CMASS and  $m_{\text{lowz}} \in [2 \times 10^{-8}, 4 \times 10^{-7}]$ , the intercept  $b_{\text{lowz}} \in [2 \times 10^{-5}, 1 \times 10^{-3}]$ . In Fig. 19 is the log-likelihood plotted for slope as a function of intercept. We find the best fit slope and intercept (the bluest region in the log-likelihood):  $m_{\text{cmass}} = 5.3 \times 10^{-8}$  and  $b_{\text{cmass}} = 1.1 \times 10^{-4}$  for CMASS;  $m_{\text{lowz}} = 1.5 \times 10^{-7}$  and  $b_{\text{lowz}} = 2.7 \times 10^{-4}$  for LOWZ. In Fig. 20 we show a comparison of the correlation coefficients and  $\chi^2$  values obtained from the real data to the distribution of those from the fake data. The fake data distribution in different colors are generated with (i) zero signal (ii) a signal with the best-fitting parameters and (iii) a signal with both a steeper slope and a higher intercept. This plot shows that it is possible to explain the low correlation and noticeable detection significance with a simple linear model.<sup>11</sup> This model's purpose is **not** to constrain any non-standard inflationary physics. Rather we wish to demonstrate that it is generically possible to have high  $\chi^2$  and low correlation coefficients, even with a simple functional form for the model.

## 6 POTENTIAL SYSTEMATICS

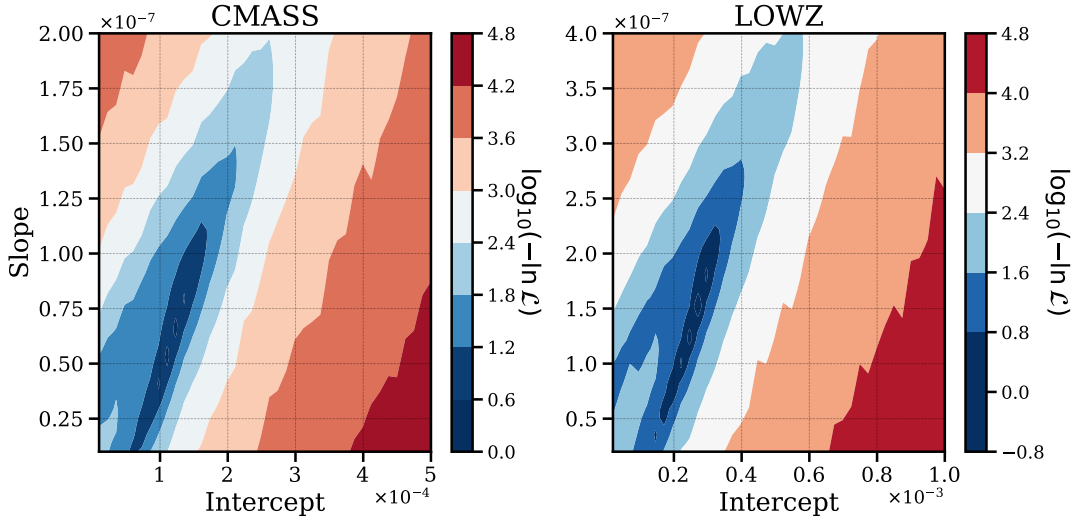
Parity-breaking in large-scale structure is not expected in the standard cosmological paradigm. Hence it is important to assess whether systematic errors could produce an apparent parity-breaking signal. We must consider systematics both at the signal level, and at the covariance level. These latter must be considered because they could cause an underestimate of the covariance leading to an overestimate of the detection significance.

We divide these systematics into three categories: survey-related effects, observer-related effects, and procedure-related effects. We discuss these in respectively §6.1, §6.2, and §6.3. We have already considered the possible impact of systematics on the covariance matrix in §4.2.2.

<sup>11</sup> The linear model can be regarded as a first-order Taylor expansion of some true physical model. We imagine some true model in the space of the physical 4PCF amplitudes; then rotate this model into the eigenbasis of the covariance and take a Taylor series for it there.



**Figure 18.** Exploration of a linear model for the underlying signal (vs. the eigenvalue index) for both CMASS and LOWZ (with ten radial bins). *Left:* Toy models for CMASS with  $m = 5.3 \times 10^{-8}$ ,  $b = 1.1 \times 10^{-4}$  (dashed red) and LOWZ (solid purple) with  $m = 1.5 \times 10^{-7}$ ,  $b = 2.7 \times 10^{-4}$  (these are our best-fit values; see Fig. 19 for discussion of the fitting and likelihood). *Middle:* Toy model (black curve; “fake signal”) transformed back to the observed basis plotted with the measured CMASS 4PCF for  $\ell_1 = \ell_2 = \ell_3$ ; upper row is NGC and lower is SGC. *Right:* Analog of the middle panel but for LOWZ.



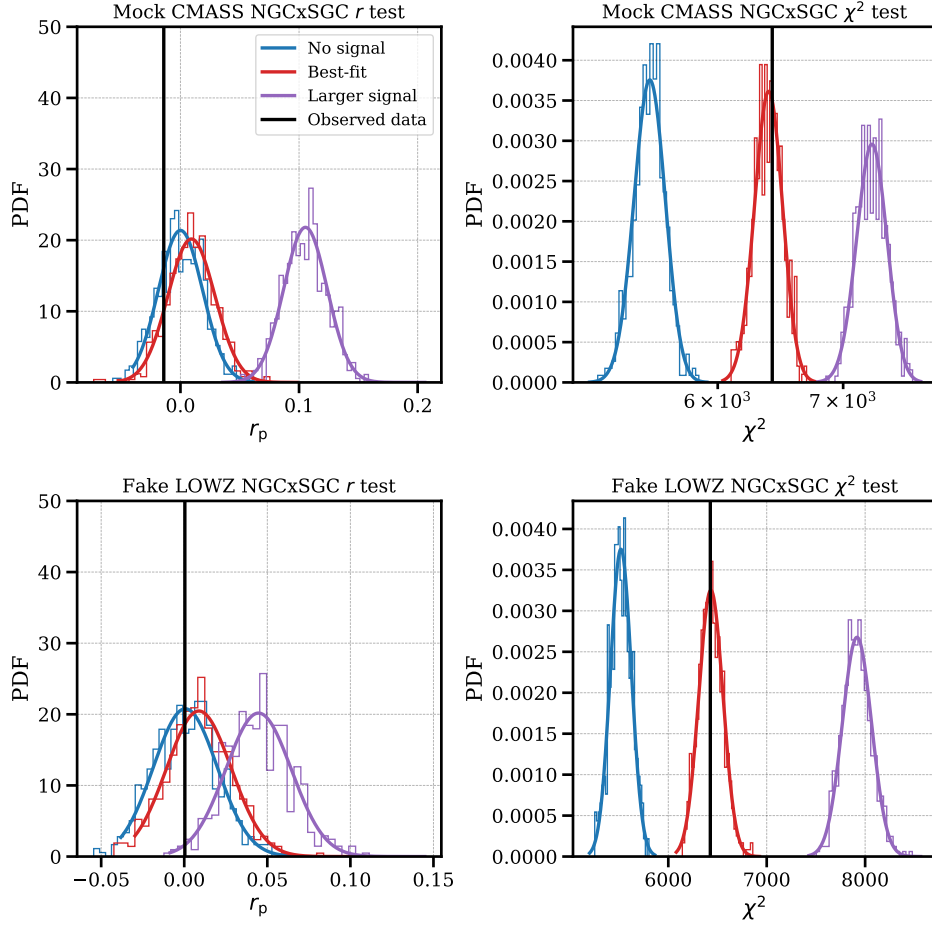
**Figure 19.** Log-likelihood for our two-parameter linear toy model  $s = mx + b$  (“fake signal”). We search for  $m$  and  $b$  so as to make the observed  $\chi^2$  and  $r_p$  consistent with each other in each sample. The fitting is independent for CMASS and LOWZ. In the eigenbasis of the 4PCF analytic covariance for either CMASS or LOWZ, we produce our linear model and then add Gaussian noise drawn using these eigenvalues to produce 200 realizations of our “fake signal”, each with an associated  $\chi^2$  and  $r_p$ . This then lets us obtain the covariance matrix of  $\chi^2$  and  $r_p$  in this toy model. We may then compute, for a given  $m$ , and  $b$ , their likelihood given the actual observed  $\chi^2$  and  $r_p$  from the data. The blue region is the highest likelihood. We display these results for both CMASS NGC cross SGC and LOWZ NGC cross SGC. It is notable that the best-fit parameters for LOWZ are roughly two times those for CMASS; this did not have to be the case, as the fitting is independent for each sample, but is in fact the scaling we would naively expect were the parity-odd signal truly of primordial origin. Our interpretation of this linear toy model is further discussed in §5.2.

## 6.1 Survey-Related Effects

Spectroscopic surveys provide galaxies’ 3D positions, but estimating the correlation functions requires the excess probability *over random* of finding an N-tuplet of galaxies in a given configuration. This in turn demands precise knowledge of the selection function.

### 6.1.1 A Toy Model for Errors in the Selection Function

Despite that great care has been taken in BOSS to model the impact of astrophysical foregrounds (*e.g.* stellar density), or observational conditions on the survey selection function, we still seek an intuition for the impact of an imperfectly-modeled selection function on the parity-odd modes. In particular, our goal is to see if an even-parity 4PCF is converted into an odd



**Figure 20.** The Pearson correlation coefficients  $r_p$  (*left column*) and  $\chi^2$  values (*right column*) observed in our data (black lines) shown over the distribution of the same from the “fake signal” generated by our “linear-in-eigenbasis” toy model. We show results from  $m = b = 0$  in blue, from the best-fit  $m$  and  $b$  in red (Fig. 18 displays these models and gives the best-fit values in the caption, and the likelihood used is shown in Fig. 19), and from a larger “fake” signal in purple to indicate the possible range of outcomes. The fake data at each index (see Fig. 18) is drawn from a Gaussian distribution with mean given by the linear model and width given by the square-root of the eigenvalue of the analytic covariance matrices. CMASS results are shown in the upper row and LOWZ in the lower. We see that there exists a “fake signal” that does explain the observed  $r_p$  and  $\chi^2$  from the real data. While our “fake signal” should not be taken too literally, this does indicate that, generically, our observed  $\chi^2$  and  $r_p$  are not necessarily in tension with each other.

parity one by an improperly corrected selection function. For this goal, it is sufficient to begin with a spatially unclustered density field (which would produce only an even-parity 4PCF).

The selection function encapsulates how our selection criteria (*e.g.* color and magnitude cuts) samples the underlying distribution of galaxies in the universe. The underlying density field can be inferred from the observed number density  $n(\mathbf{s})$  and the selection function  $f^{\text{obs}}(\mathbf{s})$ . Since we have no access to the underlying distribution, we cannot disentangle how our selection criteria shape the sample we end up observing from any underlying variation in the universe (usually specifically along the line of sight). In this discussion we allow the deviation of the selection function to be general and depend on 3D position  $\mathbf{s}$ .

Our estimate of the true number density of objects is

$$\hat{n}^{\text{true}}(\mathbf{s}) = \frac{n(\mathbf{s})}{\hat{f}^{\text{obs}}(\mathbf{s})}, \quad (23)$$

where  $n$  is the observed number density and  $\hat{f}^{\text{obs}}$  our estimate of the selection function. The actual true number density is

$$n^{\text{true}}(\mathbf{s}) = \frac{n(\mathbf{s})}{f^{\text{obs}}(\mathbf{s})}, \quad (24)$$

where  $f^{\text{obs}}$  is the true selection function.

Hence we may relate our estimated true number density to the actual true number density as

$$\hat{n}^{\text{true}}(\mathbf{s}) = \frac{\hat{f}^{\text{obs}}(\mathbf{s})}{f^{\text{obs}}(\mathbf{s})} n^{\text{true}}(\mathbf{s}) \equiv g(\mathbf{s}) n^{\text{true}}(\mathbf{s}). \quad (25)$$



Our density fluctuation field is then

$$\delta(\mathbf{s}) = \frac{\hat{n}^{\text{true}}(\mathbf{s}) - \bar{n}}{\bar{n}} = g(\mathbf{s}) - 1. \quad (26)$$

We have assumed that we correctly estimated  $\bar{n}$  and also that  $n^{\text{true}}(\mathbf{s}) \equiv \bar{n}$ , *i.e.* in this simple toy model, there is no underlying clustering. We require that the integral of  $g$  over the survey will be unity; we use this condition to normalize  $g$  shortly.

We now make a toy model for  $g$  to enable assessment of its possible impact. We take it that it is factorizable as

$$g(\mathbf{s}) = g_x(x)g_y(y)g_z(z) \quad (27)$$

and that along each direction it is a power law as

$$g_x = I^{-1/3} (x_0 + \Delta x)^{p_x}, \quad g_y = I^{-1/3} (y_0 + \Delta y)^{p_y}, \quad g_z = I^{-1/3} (z_0 + \Delta z)^{p_z}, \quad (28)$$

where we have written  $\Delta x$ ,  $\Delta y$ , and  $\Delta z$  as displacements around a central point  $(x_0, y_0, z_0)$  and have defined the normalization  $I$  as

$$I \equiv \frac{L_x^{p_x+1} L_y^{p_y+1} L_z^{p_z+1}}{(p_x+1)(p_y+1)(p_z+1)}. \quad (29)$$

This normalization ensures that the integral of  $g$  over the survey is unity and the  $L_i$  are the lengths of the box sides in each direction.

The primary galaxy for our computation of the spherical harmonic coefficients  $a_{\ell m}$  (see §3, as well as Philcox et al. 2021a for detailed discussion of how these coefficients enter our algorithm) is at  $\mathbf{s}_0$  and each secondary galaxy is displaced by  $\mathbf{r}_i$  from the primary. If we now assume that the true density field is uniform, its change due to our modulation will be (for the primary)

$$\delta(\mathbf{s}_0) = g(\mathbf{s}_0) - 1. \quad (30)$$

For the secondaries (in this case, the three field positions around the primary, which, with it, build up that primary's contribution to a given 4PCF channel) we have, by taking a Taylor series for  $g$  to leading order:

$$\delta(\mathbf{s}_i) = g(\mathbf{s}_0) + \nabla g(\mathbf{s})|_{\mathbf{s}_0} \cdot \mathbf{r}_i - 1, \quad (31)$$

with  $\mathbf{r}_i \equiv \mathbf{s}_i - \mathbf{s}_0$ . We now assess whether this systematic can contribute to the parity-odd 4PCF. When we multiply out the three secondaries, we can produce terms involving one, two, or three factors of  $\nabla g$ . All of the parity-odd basis functions are proportional to  $\hat{\mathbf{r}}_1 \cdot (\hat{\mathbf{r}}_2 \times \hat{\mathbf{r}}_3)$ , so only the term involving three factors of  $\nabla g$  may possibly contribute.

We now compute the projection of this term onto our basis. Denoting it  $\zeta_g$ , we have

$$\begin{aligned} \zeta_g &= \int d\hat{\mathbf{r}}_1 d\hat{\mathbf{r}}_2 d\hat{\mathbf{r}}_3 (\nabla g(\mathbf{s})|_{\mathbf{s}_0} \cdot \mathbf{r}_1)(\nabla g(\mathbf{s})|_{\mathbf{s}_0} \cdot \mathbf{r}_2)(\nabla g(\mathbf{s})|_{\mathbf{s}_0} \cdot \mathbf{r}_3) \hat{\mathbf{r}}_1 \cdot (\hat{\mathbf{r}}_2 \times \hat{\mathbf{r}}_3) \\ &= r_1 r_2 r_3 \nabla g(\mathbf{s})|_{\mathbf{s}_0} \cdot (\nabla g(\mathbf{s})|_{\mathbf{s}_0} \times \nabla g(\mathbf{s})|_{\mathbf{s}_0}) = 0, \end{aligned} \quad (32)$$

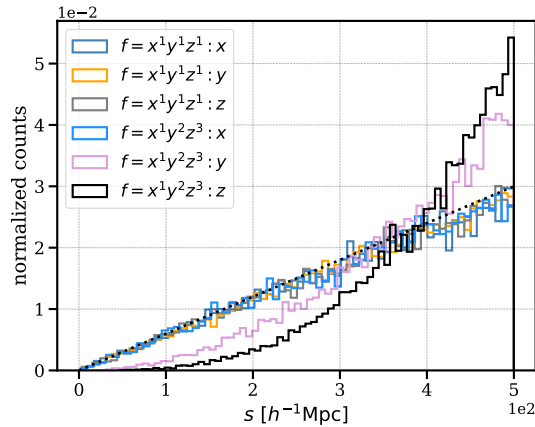
where we used a vector integral identity to obtain the second line.

Though the above shows that on average (in an infinite volume), a power-law modulation would produce no spurious parity-odd signal, such a modulation could increase the fluctuations observed. If such a systematic were present in the data but not the mocks, then the covariance used (from mocks or calibrated by mocks) would be an underestimate, and this could produce an apparent detection. Hence, we explore this power-law toy model numerically to assess if such an effect might occur. We construct 100 mocks whose points have a spatially random distribution, each in a cubic box with side length  $L_{\text{box}} = 500 h^{-1} \text{Mpc}$  and number density  $\bar{n} = 3 \times 10^{-4} [h^{-1} \text{Mpc}]^{-3}$ . The distribution of the mock number density along the  $x$ -,  $y$ -, and  $z$ -axes is shown in Fig. 21. We investigate two cases,  $p_x = 1, p_y = 1, p_z = 1$  and  $p_x = 1, p_y = 2, p_z = 3$ . We also create a set of 100 mocks with the same box-size and number density but with a uniform distribution along the three axes. For all cases, we use the same random catalog with a uniform distribution at all spatial positions to convert the density field into a density fluctuation field (the only relevant part of edge correction for a periodic box).

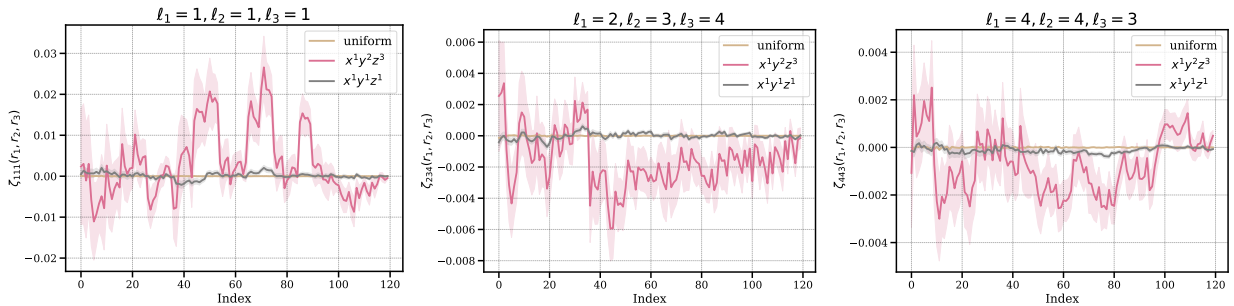
Fig. 22 shows the 4PCF coefficients measured from the no-clustering mocks. From left to right, we plot the channels  $\{\ell_1 = 1, \ell_2 = 1, \ell_2 = 1\}$ ,  $\{\ell_1 = 2, \ell_2 = 3, \ell_2 = 4\}$ , and  $\{\ell_1 = 4, \ell_2 = 4, \ell_2 = 3\}$ . In each panel, we compare a parity-odd 4PCF coefficient for an unmodulated sample and two differently-modulated samples. As expected, the 4PCFs measured from the uniform sample have very small amplitudes. The  $\chi^2$  for the uniform sample is comparable to those modulated samples. Overall, this means that the power-law modulation would not produce a spurious detection of parity-odd modes at the signal level.

### 6.1.2 Density Distortions, Selection Function, and Fiber Collisions

In theory, one can expect parity-odd signals to be induced solely by 3D modulations of the density field, because such a modulation could convert a parity-even mode into a parity-odd mode. A 2D screen uncorrelated with the underlying density field can be separately averaged over rotations, and the rotation will force any parity-odd piece of the screen to vanish. An example of an uncorrelated screen is that given by the veto mask, which was designed to remove angular regions contaminated by foreground bright stars, plate holes, *etc.* and is not correlated with the underlying 3D galaxy density.



**Figure 21.** Normalized number count distribution along the  $x$ -,  $y$ -, and  $z$ -axes of mocks with a spatially-random underlying distribution but modulated by a power-law in each direction (§6.1.1, Eq. 28), as indicated in the legend; the function  $f$  shows the modulation along each axis, and the variable after the colon which axis is plotted. We show two cases along each of the three axes; powers  $p_x = p_y = p_z = 1$  and then  $p_x = 1, p_y = 2, p_z = 3$ .



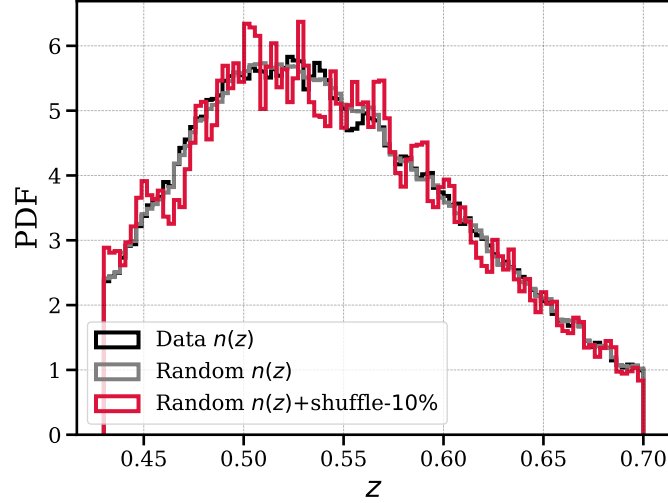
**Figure 22.** Several parity-odd 4PCF coefficients as measured from our spatially-random power-law-modulated mocks (§6.1.1). We show results from 100 mocks with number density  $\bar{n} = 3 \times 10^{-4} [h^{-1} \text{Mpc}]^{-3}$  for no modulation (tan),  $x^1 y^2 z^3$  (red), and  $xyz$  (grey). We find that the  $\chi^2$  for the power-law-modulated mocks is smaller than that for the unmodulated ones, implying that this modulation does not produce a significant parity-odd 4PCF at the signal level.

Another 2D effect is “fiber collisions” (FC). BOSS uses fibers to guide the light of the observed objects from the focal plane to the spectrograph. Each fiber has a limited angular diameter ( $62''$ ) and multiple galaxies that fall within this radius cannot be resolved within one exposure. In the data fiber collisions are corrected by upweighting the nearest angular neighbor (Ross et al. 2012; Reid et al. 2016). This approach can potentially introduce long wavelength mode-coupling because it identifies neighbors only on an angular basis. Although one can model the fiber-collision effect as a top-hat function (Hahn et al. 2017), the exact physical scale could still be redshift-dependent (see Hou et al. 2021b for quasars, which span a wide redshift range).

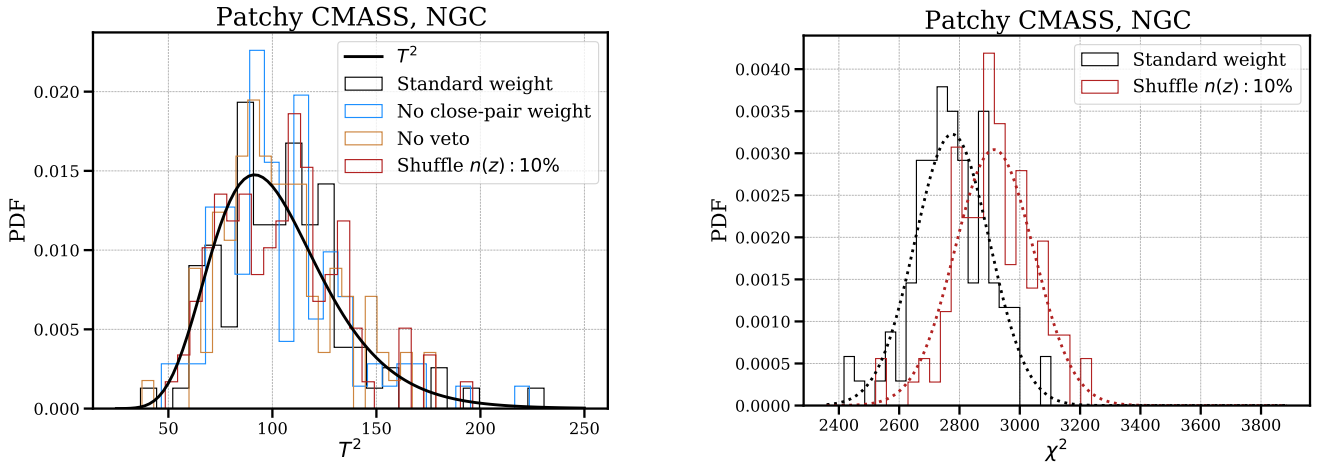
To demonstrate the impact of these angular effects on the parity-odd modes, we turn off the veto mask weights implemented in the PATCHY mocks as well as the fiber collision weights. Further, by switching on and off the fiber collision weight we can see the impact of the nearest-neighbor upweighting on the parity-odd detection significance. The left panel of Fig. 24 shows the distribution of the resulting  $T^2$  values compared to that expected. When the covariance is inferred from the contaminated mocks themselves, there is good agreement between the contaminated mocks’  $T^2$  distribution and that for the standard case. We therefore conclude that these effects do not induce parity-odd modes at the signal level. In particular, we also find that mocks without veto mask and fiber collision weights both result in a lower detection significance. This is in contrast to mocks contaminated with a distorted radial selection function, which we discuss next.

In addition to the angular effects, we also distort the radial selection function by introducing a 10% scatter in the  $n(z)$  of the randoms. At each redshift in Fig. 23, we draw a value from a Gaussian with mean 10% and width 5%, and then decide randomly and with equal probability whether to give it a plus or a minus sign. We then use this value to set a new probability of a random particle at that redshift being selected. This procedure results in the red curve in Fig. 23 for the shuffled randoms’  $n(z)$ . The FKP weights for the randoms are also recomputed after this.

The distribution of  $T^2$  values for the PATCHY mocks with shuffled  $n(z)$  when we use the compression scheme with the covariance inferred from the contaminated mocks agrees well with that obtained using the standard weights (black histogram)

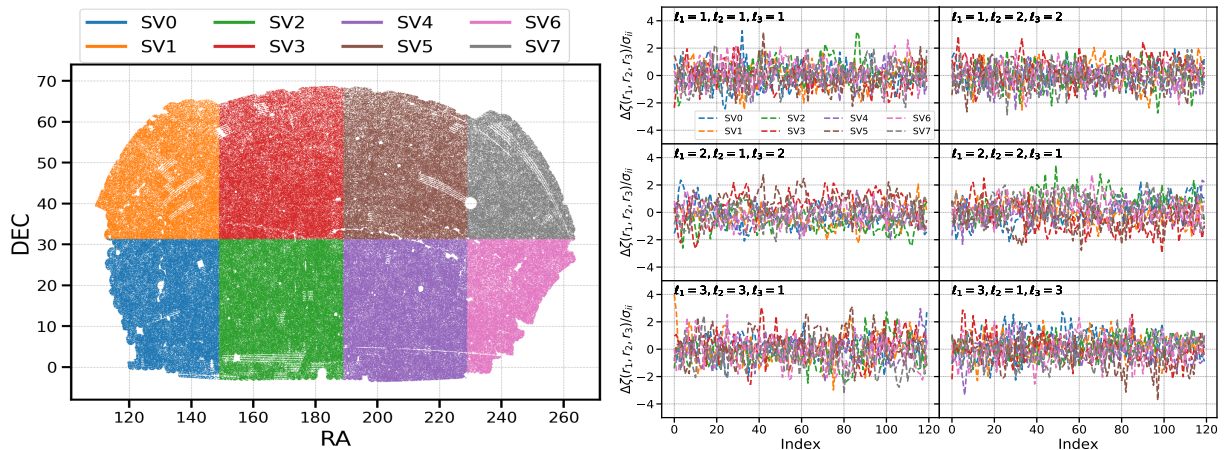


**Figure 23.** PDF of the radial selection function for PATCHY NGC. The selection function applied to the PATCHY mocks is designed to match that from the BOSS data for both the mock data catalog (black histogram) and the random catalog (grey histogram). To test whether an imperfect inference of the radial selection function can lead to a spurious parity-odd signal, we deliberately distorted the  $n(z)$  for the random catalog (red histogram), by shifting it by 10% on average, where the percentage shift is drawn from a Gaussian with that mean and width of 5%. We show the distribution of  $T^2$  values for PATCHY mocks analyzed using these “shuffled” randoms in Fig. 24; it is consistent with no detection of a parity-odd 4PCF.



**Figure 24.** Histograms of statistics for PATCHY mocks’ 4PCFs analyzed with different weighting schemes. *Left:*  $T^2$  values calculated with 50 eigenvalues using our data compression scheme (§5.1.2) with the covariance matrix inferred from the systematics-integrated mocks themselves. The black curve is the expected  $T^2$  distribution. The histogram for the PATCHY mocks with the standard weighting scheme is in black, with the close pair weight turned off to assess the impact of fiber collisions in blue, with the veto mask turned off to assess the impact of the angular mask in orange, and finally, in red, the results of shuffling the randoms’ redshift distribution by 10% as described in Fig. 23. None of these alterations to the 4PCF analysis produce a distribution of  $T^2$  values that are inconsistent with that expected under the null hypothesis of no-parity-odd 4PCF (black curve). *Right:*  $\chi^2$  values calculated using the analytic covariance calibrated to mocks analyzed with the standard weights. Here we can see that the  $\chi^2$  distribution has a higher signal mean for the shuffled mocks than for the standard mocks, suggesting that although such a systematic does not induce a parity-odd *signal*, it might lead to underestimation of the covariance.

and also with the predicted  $T^2$  distribution (black curve) as shown in the left panel of Fig. 24. However, the radially-distorted randoms lead to an increase in the covariance. As a result, the mock distribution shifts towards higher  $\chi^2$  when we use an analytic covariance matrix calibrated with respect to the standard mocks (right panel of Fig. 24). In other words, had this systematic been present in our BOSS data, the covariance matrix inferred from the standard PATCHY mocks would have been underestimated. In table 5 we estimate how much this effect could have been affecting our detection significance for **ten** and **eighteen** bin tests, respectively.



**Figure 25.** *Left:* Footprint of CMASS NGC showing how we split the data into eight patches, each roughly  $40^\circ$  in right ascension and  $35^\circ$  in declination. “SV” stands for sub-volume, as each patch extends into the redshift direction. *Right:* For six parity-odd channels, we display the difference between the 4PCF coefficients of the eight sub-volumes created by extending these angular patches in the redshift direction, and those from the full NGC sky. We divide these differences by the standard deviation of the PATCHY results. The colors for the eight sub-volumes match in the left and right panels. Overall, we see that no one angular region stands out as producing a great deal more parity-odd signal than another. We offer a more rigorous demonstration of this claim in Fig. 26, which shows the  $\chi^2$  for each sub-volume.

### 6.1.3 Magnification Bias

We consider magnification bias qualitatively. If the underlying density field produces non-zero amplitudes only in even-parity 4PCF channels, it is sufficient to ask whether magnification bias will preferentially affect tetrahedrons of one handedness. Much like the method of image charges, we may consider a notional pair: a tetrahedron and its mirror image. As long as we can prove that for any such pair, there is no effect, no parity-breaking will be introduced.

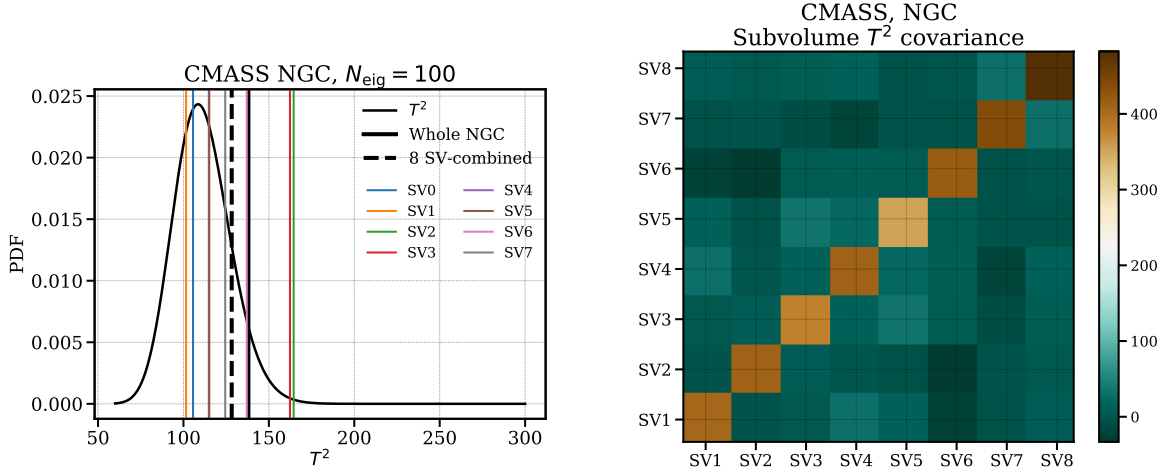
Consider for simplicity a tetrahedron for which the three galaxies defining a triangular base are at the same redshift, and that the fourth point is at a higher redshift. If the three “base” points are closely-enough concentrated, they may magnify the fourth galaxy behind them and on average make it more likely to be selected than otherwise. However, this will be the case also for the mirror image of this tetrahedron. Indeed, magnification will be invariant to sign-flip of the two plane-of-sky coordinates; at most, it can be parity-odd in  $z$ , but any such contribution from the  $z$  behavior alone will vanish after rotation-averaging.

### 6.1.4 Splitting the Sky into Angular Patches

To assess if our detection is coming from one part of the sky preferentially, which might indicate a systematic, we split the sky into eight angular patches, each of roughly  $40 \times 35 \text{ deg}^2$  (right ascension, RA, times declination, DEC), and run our analysis on each independently. Each angular patch implies a subvolume (hence SV) as we extend it out in the redshift direction. The right panel shows some examples of the 4PCF measurement and compares them to the standard deviation of each subvolume. We estimate the covariance of each subvolume by splitting the PATCHY mocks in the same way. Since the subvolumes receive modulations from Fourier modes larger than them, “beat-coupling” leads to additional corrections to the covariance matrix (Li et al. 2014; Putter et al. 2012). Thus a naive rescaling of the covariance by its volume relative to the total survey volume would likely give an underestimate and thence produce artificially high  $T^2$  values.

Fig. 26 shows the  $T^2$  computed from the entire BOSS CMASS sky (thick dashed vertical black line) and the eight subvolumes (thin colored vertical lines). The expected  $T^2$  distribution is the black curve. We combine the eight subvolumes by taking an inverse-variance weighted sum of the  $T^2$  values. Using the diagonals of the covariance matrix (see the right panel of Fig. 26) we find the combined  $T^2 = 128$ , which differs from the one from the entire BOSS CMASS sky  $T^2 = 138$ . We repeat this using the full covariance of the subvolumes and find  $T^2 = 129$ , quite similar to the result using just the diagonal. We see that subvolume 2 and 3 have the highest detection significance compared to the average; given the errorbar of the  $\chi^2$  shown in the right panel of Fig. 26, this corresponds to a  $1.8\sigma$  difference. We also note that combining the signal at the estimator level is not a simple linear operation due to the edge correction (see also Eq. 8). Moreover, since we use the same analytic covariance for different subvolumes, the amount of information picked out by the ranked eigenvalues could differ for each subvolume. We therefore expect that there is a difference between analyzing the data on the whole sky or splitting it and then combining.

We also repeat this same test on the SGC and on the LOWZ sample. Overall, we do not find any subvolume has a particularly high or low detection significance compared to the average.



**Figure 26.** *Left:* Here we show the  $T^2$  values from our angular patch test as described in Fig. 25. The  $T^2$  value for the entire CMASS NGC is the thick black line, the eight sub-volumes'  $T^2$  values are in thin, colored lines, and the combined  $T^2$  from the eight sub-volumes is in dashed black. The predicted  $T^2$  distribution (under the null hypothesis of no parity-odd signal) with 100 degrees of freedom is the black curve; we have used the compressed analysis described in §5.1.2. *Right:* Covariance for the  $T^2$  values of the eight sub-volumes, estimated using the PATCHY mocks. This covariance is needed to properly co-add the  $T^2$  values for the eight sub-volumes to report the combined  $T^2$  given in the lefthand panel (dashed line there). The slight non-vanishing off-diagonal indicates correlations between the eight sub-volumes, but this turns out not to affect substantially the combined  $T^2$  value. We may observe that the combined  $T^2$  value differs from that for the whole NGC by about 10%; we believe this is due to missing the produced by quartets that are split across more than one sub-volume, which go uncounted in our sub-volume analysis. We would expect that adding them in (as the whole NGC analysis does) would indeed raise the  $\chi^2$ .

### 6.1.5 Redshift Failure

The BOSS fibers are not completely randomly positioned in an observing tile. The resolving power of the fibers is impacted by their positions on the CCD cameras. Those fibers near the edge of the spectrograph slit are less likely to obtain a high-quality spectrum than are more central fibers (see §4.2.2 of Smeed et al. 2013). Consequently, redshift failure occurs in a way that depends on 2D position on the telescope plate. This dependence is not wholly symmetric under a parity transformation of the  $x$  and  $y$  axes on the plate.

Now, when a redshift failure occurs, the nearest angular neighbor of the failed galaxy is upweighted. The neighbor's redshift should be a fair draw from the redshift distribution of the survey. However, redshift failure is statistically slightly more likely for a galaxy at higher redshift, as it will on average be fainter, and hence more affected by the reduction in spectrum quality from being at the edge of a CCD. Thus, replacing it by upweighting a galaxy fairly drawn from the survey's redshift distribution can actually produce a *bias*. Effectively, we will systematically be pulling galaxies towards lower redshift as a function of 2D position on the plate. This 2D function, as already noted, is not fully symmetric under  $(x, y) \rightarrow (-x, -y)$  and hence, in combination with the coupling to  $z$  produced by the above, can in principle imprint a parity-odd signal on the density.

We now seek to explore this issue. CMASS has 1.8% of its objects undergo redshift failures, while LOWZ has only 0.3% do so (Ross et al. 2012). Thus we mainly focus on CMASS. Since the redshift failure galaxies are already removed from the catalog, we explore this effect by creating *more* redshift failures and seeing whether it appreciably alters our parity-odd signal.

We first infer the redshift efficiency according to the inverse of redshift failure weight, then we apply a fit and assign each fiber a redshift efficiency rate accordingly. We apply a magnitude cut with  $\text{mag}_{\text{cut}} = 19.20$  and from which a galaxy subsample is selected weighted by the fiber efficiency rate such that the subsample is 1.8% of the total galaxies. We then identify the faint and bad fiber galaxies' neighbors and select the ones that are closest ones in angular distance. Finally we upweight those galaxies' redshift failure weights by one and renormalize the random catalog's weights such that it matches the total weights of the new data catalog. We repeat the above steps for both NGC and SGC.

Here we do not transfer the total systematic weights, as they cannot be simply propagated to the nearest neighbours. Assuming the total systematic weights to be unity for those galaxies has an impact on the total weights of only 0.1%, so we do not expect an issue to arise due to this. We also retain the current estimation of the sector completeness, as this cannot be easily redone at the catalog level.

We found that doubling the redshift failures increased the total parity-odd detection on the joint sky by  $2.5\sigma$  when using eighteen radial bins and  $1.5\sigma$  when using ten radial bins, both with the analytic covariance; these results are summarized in Table 5. This increase could be due to the 3D modulation of the density field producing a parity-odd mode as discussed above, but it could also simply be due to increased *variance* in the data (*i.e.* that any parity-odd *signal* so produced would average to

zero in the infinite limit). This increased variance could stem from our inability to correct the sector completeness to account for the removal of more objects.

Given that the angular separation between galaxies and their nearest-neighbours is typically much smaller than the sky coverage of a sector, and most of the upweighted galaxies and the removed galaxies are still within the same sector and this, we might expect that the impact of uncorrected sector completeness in our test setup is small. However, we need to quantify its impact. In order to do so, we applied a uniform selection in redshift and selected 1.8% of those galaxies that are close to bad fibers. We proceed by upweighting their neighbours. There we observed  $2.2\sigma$  and  $0.5\sigma$  increase in detection significance for the joint sky using eighteen bins and ten bins, respectively. Hence the increase in the detection significance is actually dominated by the imperfect compensation between the random catalog and the real data produced by the uncorrected sector completeness in our test setup, which leads to a spurious variation in the density field. This is qualitatively similar to what we find with the power-law toy model in §6.1.1.) Furthermore, redshift failures also impact the parity-even modes; this impact (for our “doubling” test) is given in Table 5. The good agreement between this mode in the true BOSS data and in the distribution of the mocks (Philcox et al. 2021b) demonstrates that redshift failure at the level it is actually occurring in BOSS should not have a strong impact on the detection significance.

## 6.2 Observer-Related Effects

### 6.2.1 Redshift-Space Distortions (RSD)

Galaxies’ radial distances are inferred from their redshifts, however, their peculiar velocities caused by local gravitational environment add a component in addition to their cosmological redshifts. This distortion in galaxies’ positions are known as redshift-space distortions (RSD).

In order to understand the impact of the RSD on parity-odd modes, we first consider the “global” line of sight approximation and take the  $\hat{\mathbf{z}}$  as the line of sight. RSD modulate the density by the usual Kaiser (1987) factor, which is even-parity with respect to the line of sight. The primary galaxy itself is at the origin and so does not carry any angular momentum; multiplying it with three density fields, with each of them being parity-even, will be parity-even. Given that rotation averaging does not alter the parity, RSD considered in the “global” line of sight approximation cannot produce parity-odd modes. We can easily extend this logic to a “local” line of sight by taking the line of sight to be that to the primary. The modulation in the overdensity field due to the RSD will be proportional to  $(\hat{\mathbf{n}} \cdot \hat{\mathbf{k}})^2$ , but again results only in parity-even combinations. Given that rotation averaging does not alter the parity, certainly RSD considered in the “local” line of sight approximation cannot produce parity-odd modes.<sup>12</sup> We also note Jeong & Schmidt (2019) have considered parity-odd bispectrum induced by galaxies’ radial velocity by introducing explicit coupling to the line of sight degree of freedom, by doing so the total rotational invariant system includes the additional line of sight vector, which by itself has no constraint on parity as long as the combined system has total zero angular momentum. In our work, instead, the line of sight dependencies are already averaged over and we are not sensitive to the parity-odd mode due to the RSD as in Jeong & Schmidt (2019).<sup>13</sup>

To investigate the impact of RSD, we also use the PATCHY mocks in real- and redshift-space. For this test, the PATCHY mocks used are in a cubic box of side length  $L_{\text{box}} = 2,500h^{-1}\text{Mpc}$  and with number density  $\bar{n} = 6.6 [h^{-1}\text{Mpc}]^{-3}$  at redshift  $z = 0.57$ . The galaxies are shifted using the exact 3D velocities of each galaxy in the mock. A comparison of the 4PCF in real- and in redshift-space is shown in Fig. 27. We do not observe an overall enhancement in the amplitude for the 4PCF.

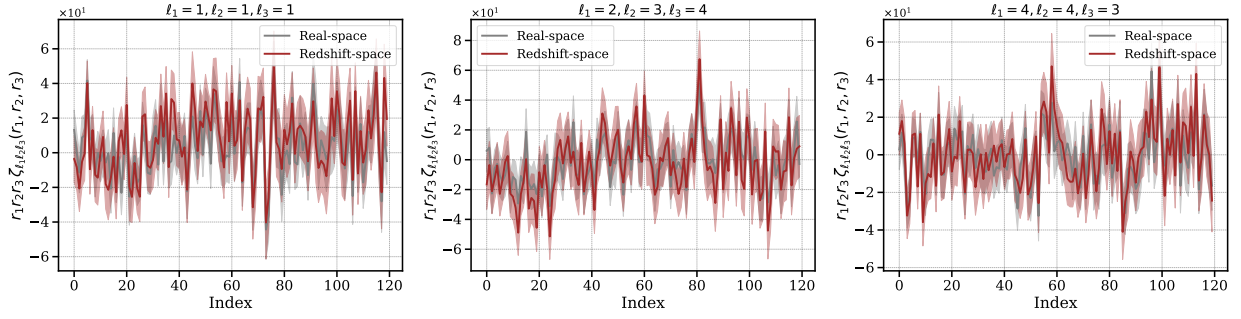
### 6.2.2 Tidal Alignment

The large-scale tidal field impacts galaxies’ orientations, which might lead to an anisotropic galaxy selection. For instance, in the case of spirals, an imaging survey might be more likely to detect a face-on rather than edge on disk. Even LRGs are stretched ellipsoids, with aspect ratios of order 2:1, and so orientation could impact selection, though the effect would likely be smaller. For LOWZ, the sample is quite pure (LRGs), but for CMASS, the sample is roughly 26% disk blue galaxies. We might thus expect any anisotropic selection would manifest more strongly in CMASS.

However, given that the observer’s orientation is random with respect to the tidal field (and thus the galaxy being selected’s orientation), we suspect that such an effect could not produce parity-odd modes. If it did, we might expect a stronger detection of them in CMASS for the reason noted above.

<sup>12</sup> Beutler et al. (2019) find parity-odd modes in the power spectrum multipoles (*i.e.* odd  $\ell$ ; with respect to the line of sight) but that is an artificial consequence of the way that the Yamamoto estimator, which selects one of the two galaxies to define the line of sight, is implemented in Fourier space. Even in configuration space, the Yamamoto estimator does not produce parity-odd modes with respect to the line of sight, as shown in Slepian & Eisenstein (2015a)

<sup>13</sup> We have calculated that each term in their galaxy bias model that sources an odd-parity bispectrum will not lead to an odd-parity 4PCF after averaging with respect to the line of sight.



**Figure 27.** parity-odd 4PCF channels  $\{\ell_1 = 1, \ell_2 = 1, \ell_3 = 1\}$ ,  $\{\ell_1 = 2, \ell_2 = 3, \ell_3 = 4\}$ , and  $\{\ell_1 = 4, \ell_2 = 4, \ell_3 = 3\}$ . We compare the measured 4PCF coefficients from 99 PATCHY mocks in real-space (grey) and in redshift-space (red) with uniform geometry. The error bar is inferred from the standard deviation of the 99 mocks.

However, for completeness, we here work out what a tidal model might do if the selection depended upon it. Following Hirata (2009), the observed galaxy density fluctuation can be expressed as

$$\delta_g^{\text{obs}} + 1 = (\delta_g^{\text{true}} + 1) (1 + \epsilon(\hat{\mathbf{n}}|\mathbf{x})), \quad (33)$$

where  $\epsilon$  is the selection function at position  $\mathbf{x}$ , and it depends on the viewing direction  $\hat{\mathbf{n}}$ . The average of the selection function  $\epsilon$  over the unit sphere is zero. By treating the large-scale tidal field as a perturbation, the lowest-order Fourier-space contribution (after Taylor-expanding) is (Hirata 2009 equation (30)):

$$\epsilon(\hat{\mathbf{n}}|\mathbf{k}) = A \left[ (\hat{\mathbf{n}} \cdot \hat{\mathbf{k}})^2 - \frac{1}{3} \right] \delta_m(\mathbf{k}) \quad (34)$$

where  $A$  is an amplitude and  $\delta_m$  is the matter overdensity. We see that the selection function  $\epsilon$  is parity-even, and so multiplying it onto a single density does not alter the parity of any spherical harmonic expansion coefficient associated with that density field. Hence, if there is on average signal in only parity-even channels in the underlying galaxy 4PCF, modulating each density point by  $\epsilon(\hat{\mathbf{n}}|\mathbf{k})$  above will not convert them into signal in parity-odd channels.

### 6.3 Procedure-Related Effects

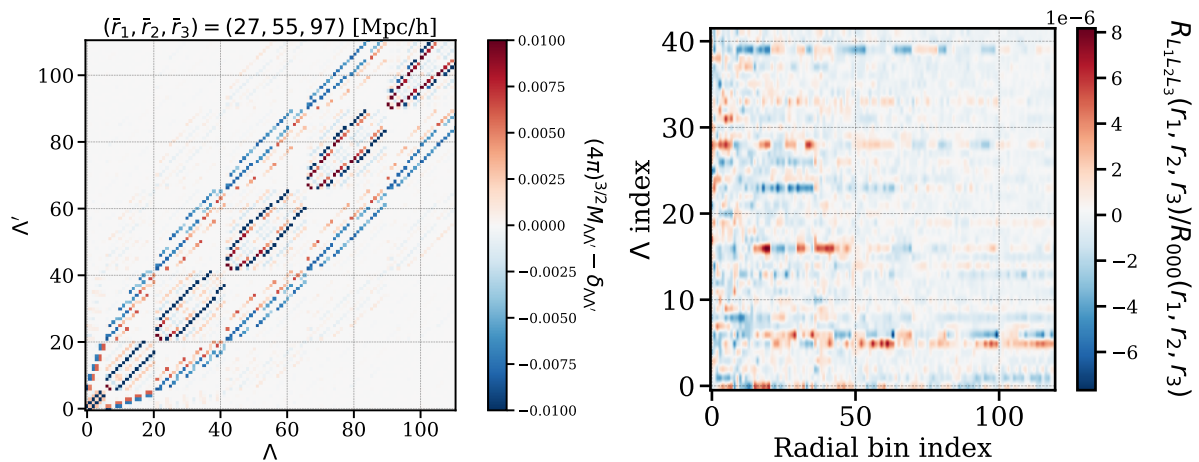
#### 6.3.1 Fiducial Cosmology

In order to convert the sky coordinates to Cartesian coordinate we need to assume a fiducial cosmology. Using a fiducial cosmology that differs from the true cosmology may lead to distortions in the clustering.

At the signal level, if any galaxy quadruplets have their parity flipped by such an error, it is equally likely to happen for a given tetrahedron and for its mirror image. Therefore, on average, even if a “parity-flip” occurs, it will be cancelled out after averaging over all galaxies in the survey. The radial bin separation  $\Delta r$  we choose is  $14 h^{-1} \text{Mpc}$  when we use ten radial bins and  $8 h^{-1} \text{Mpc}$  when we use eighteen radial bins. Beginning with the cosmology of Planck Collaboration et al. (2020) and moving  $\pm 3\sigma$ , the change in the angular diameter distance is less than 1%, which will not be resolved by our radial bin choice. At the covariance level, a different cosmology could lead to a different estimate of the best-fit number density and volume (for the analytic covariance) and would also change the power spectrum shape. Even though we use the same cosmology to convert the galaxies in the real data and the PATCHY mocks, the statistical fluctuations do not necessarily scale the same way as we assume a fiducial cosmology that differs from the true one. We can see a 6% change in the volume when distorting our fiducial cosmology by  $3\times$  the constraints on  $\Omega_m$  obtained in Planck Collaboration et al. (2020). Such a difference corresponds to an overestimation of the detection significance by  $\sim 2.5\sigma$  (see Table 5). However, if there is an obvious difference between our choice of fiducial cosmology and the underlying one, it will also be visible in the comparison of the parity-even 4PCF measurement. The fact that we saw good agreement between the observed parity-even 4PCF and the mock distribution suggests wrong fiducial cosmology is not relevant here.

#### 6.3.2 Survey Geometry and Edge Correction

As discussed in §3, the 4PCF measurement requires correction for the survey geometry as given in Eq. (8). In principle there can be mixing between parity-even and parity-odd channels due to the survey geometry. The left panel of Fig. 28 shows the coupling matrix  $\mathbf{M}_{\ell_1 \ell_2 \ell_3, \ell'_1 \ell'_2 \ell'_3}$  as in Eq. (9) from PATCHY mocks with CMASS NGC geometry at a radial bin combination  $(r_1, r_2, r_3) = (27, 55, 97) [h^{-1} \text{Mpc}]$ . This coupling matrix has off-diagonal elements, which means that the parity-even mode can couple to the parity-odd mode. To assess the correction of the parity-odd modes due to the survey geometry, we plot the



**Figure 28.** *Left:* Edge-correction matrix from PATCHY mocks with CMASS NGC geometry at a given radial bin  $(\bar{r}_1, \bar{r}_2, \bar{r}_3) = (27, 55, 97)$  [ $h^{-1}$  Mpc]. Each  $\Lambda = \{\ell_1, \ell_2, \ell_3\}$  denotes a specific angular channel for both even and odd parity. The non-vanishing off-diagonal features indicate that the survey geometry couples even and odd modes. *Right:* The odd coefficients of the randoms normalized by  $\mathcal{R}_{000}$ . We have mapped the odd angular momentum triplets given by  $\Lambda$  to a 1D index (vertical axis), and also so mapped on the horizontal axis. Here we see the dimensionless parity-even-to-parity-odd mixing is of order  $10^{-6}$ , showing that the dominant change to the odd-parity modes from edge-correction is due to  $\mathcal{R}_{000}$ , *i.e.* overall renormalization from dividing by the mean number density.

ratio of the random coefficients to the lowest-order one,  $\mathcal{R}_{\ell_1 \ell_2 \ell_3} / \mathcal{R}_{000}$  (right panel of Fig. 28). These ratios are of order  $10^{-6}$ ; the dominant edge correction is thus simply  $\mathcal{R}_{000}$ , which does not mix parities. In addition, the fact that the PATCHY mocks (which also get edge-corrected) are consistent with the predicted distribution indicates that survey geometry correction is not inducing spurious parity-odd modes.

We also may develop a toy model for the edge-correction factors based on Slepian & Eisenstein (2015b) §4.3, where the impact of a planar edge on a uniform density sphere was considered. We consider a sphere of radius  $R$  some distance away from a planar edge of the survey, and orient the  $z$ -axis perpendicular to that plane. We may define a critical angle cosine  $\mu_c$ ; for  $\mu < \mu_c$ , a point will fall outside the survey. We may compute the  $a_{\ell m}$  as a function of  $\mu_c$ , then form the 4PCF estimate, and finally average over all possible  $\mu_c$ , which corresponds to averaging over how far away the sphere’s center is from the edge. The choice of  $z$ -axis means that only  $m = 0$   $a_{\ell m}$  are non-zero. The setup is described more extensively in Slepian & Eisenstein (2015b) §4.3.

Proceeding in this way, we may find a general form for  $a_{\ell 0}$ , which is given in Slepian & Eisenstein (2015b). Here, we will only compute the first few leading edge correction factors, so we find it more useful to simply quote the first few  $a_{\ell 0}$  explicitly. We have  $a_{00} = \sqrt{\pi} [\mu_c + 1]$ ,  $a_{10} = \sqrt{3\pi} [\mu_c^2 - 1]/2$ , and  $a_{20} = \sqrt{5\pi} [\mu_c^3 - \mu_c]/2$ . Cubing  $a_{00}$ , we find that  $\mathcal{R}_{000} = \pi^{3/2} [15/4]$ ; the 3- $j$  symbol involved in  $\mathcal{R}_{000}$  is unity. Similarly, we obtain  $\mathcal{R}_{011} = -\pi^{3/2} \sqrt{3} [7/40]$ ,  $\mathcal{R}_{112} = -\pi^{3/2} \sqrt{3}/32$ , and  $\mathcal{R}_{022} = -\pi^{3/2} \sqrt{5} (33/1120)$ . These lead to, respectively,  $f_{011} = 8\%$ ,  $f_{112} = -1\%$ , and  $f_{022} = 2\%$ , where these are the ratios of the  $\mathcal{R}$  with the indicated angular momenta to  $\mathcal{R}_{000}$ . Furthermore, we note that only of order 20% of spheres in the survey impinge on a boundary at all for the BOSS footprint; hence these factors would be expected to scale down by roughly a factor of five. These estimates show that the edge correction is not a substantial alteration to the measured 4PCF, and indeed that the dominant effect is that of  $\mathcal{R}_{000}$ , with the higher-order edge-correction factors contributing of order a few percent at most (after division by the factor of five).

## 7 CONCLUSION

In this work we have used a novel, recently-suggested method (Cahn et al. 2021) to search for parity-violation in 3D large-scale structure. This method exploits the fact that the lowest-order shape that cannot be rotated into its mirror image in 3D is a tetrahedron, and thus counting tetrahedra and looking for excesses of one type over another can reveal parity-violation. We have detected strong evidence for it in both the CMASS and LOWZ samples of BOSS, and in both the NGC and SGC of each, summarized in Table 2. We have explored numerous systematics to assess whether any could produce either a spurious parity-odd signal, or a substantial change to the covariance matrix used such that we would underestimate the statistical errorbars so as to make a spurious detection. We have not found any systematics that seem to do either. We have also performed a number of tests of our covariance matrix and pursued a number of analysis variations both as regards covariance matrix and radial binning, scales, and angular momenta used.



**Table 5.** Here we summarize the shifts in detection significance, in units of  $\sigma$ , due to various systematics for both the parity-odd (top two rows) and parity-even (bottom two rows) modes, including redshift failure (§6.1.5), the impact of the fiducial cosmology (§6.3.1), and a distorted selection function (§6.1.2). We used the N+S of CMASS and the analytic covariance for these calculations.

Parity	Number of Bins	Redshift Failure	Fiducial Cosmology	Distortion in Selection Function
Odd	10	1.5	$\sim 1.5$	$\sim 1.0$
	18	2.5	$\sim 2.5$	$\sim 1.7$
Even	10	0.9	$\sim 0.7$	$\sim 0.6$
	18	1.8	$\sim 1.9$	$\sim 2.4$

Overall, the compact takeaway from this work is that CMASS with eighteen radial bins has a  $7\sigma$  detection of parity-odd modes when using the analytic covariance with all degrees of freedom, and a roughly  $4\sigma$  detection when using of order 500 eigenvalues plus a covariance estimated directly from the mocks. Given the large number of degrees of freedom in our analytic-covariance analysis, underestimating the noise by 20% in every degree of freedom would be enough to produce this detection spuriously. In the “compressed” analysis with 500 eigenvalues, we would have had to underestimate the covariance by 30% in every degree of freedom to spuriously produce our  $4\sigma$  result. Such a substantial underestimate seems unlikely but given that the result, if cosmological in origin, would be significant, caution remains warranted.<sup>14</sup>

We now briefly revisit the key technical aspects of our analysis and of systematics search.

### Radial Binning

In this work we tried three different radial binning schemes and found different detection significances. We hypothesize that the increasing detection significance is due to reduction of internal cancellation. The finest binning we use in this work has  $\Delta s = 8 h^{-1}\text{Mpc}$ , which is roughly half of the average inter-particle separation for BOSS. In this work, our main goal is to explore a possible parity-odd signal; we leave the exploration of the optimal radial binning to future work.

### CMASS vs. LOWZ Detections

CMASS and LOWZ are both mainly composed of LRGs, but the two samples behave differently with regard to i) sensitivity of the detection significance to varying radial binning and ii) sensitivity of the detection significance to varying  $\ell_{\text{max}}$ . These different behaviors may be due to differences in the number density, mean redshift, galaxy population, and redshift failure rates of the two samples.

### Correlation Across the Sky

We do not observe a statistically significant positive cross-correlation between the NGC and SGC in each dataset, and we find that this test is sensitive to the radial binning. We explored a simple phenomenological model to understand if high detection significance but low cross-correlation is generically possible, and it seems so. These two measures are algebraically independent so this finding was not outside the realm of expectation.

### Systematics

While we have shown that nearly all the systematics considered do not produce parity-odd modes at the *signal* level (*i.e.* they would contribute vanishingly in the infinite-volume limit), some can enhance the variance of the observed data relative to the covariance from the mocks if they are present in the data but not the mocks. If substantial enough, this in principle could produce a spurious detection. In practice, we have not found any systematic that, even if grossly larger than what we believe to be present in the data, could increase the variance enough to explain our results. We have also looked for consistency in the even-parity sector as a way to guard against any systematics’ unbeknownst to us increasing the variance.

### Implications

Given the unexpected nature of the detected signal, it is premature to speculate on the implications. Following the advice of Newton: *hypotheses non fingo*.<sup>15</sup> Nonetheless, it is worth placing our results in the context of possible theoretical models that could explain them, should the signal here be of genuinely cosmological origin. Besides Chern-Simons-like interactions (Deser et al. 1982b,a; Witten 1989) with couplings of different forms and various extensions (Freidel et al. 2005; Alexander et al. 2006; Liu et al. 2020; Li & Zhao 2022), string-sourced perturbations (Pogosian & Wyman 2008), primordial vortices (Vilenkin 1978;

<sup>14</sup> We chose 500 eigenvalues for discussion here because that strikes a balance between detection significance yet has several mocks per element in the covariance. This has 1,000 degrees of freedom because we treat eigenvalues from North and South as separate degrees of freedom.

<sup>15</sup> “I contrive no hypotheses,” from Newton’s “General Scholium” added to the *Principia* of 1713.

Vilenkin & Leahy 1982; Brizard et al. 2000), broken symmetry during phase transition (’t Hooft 1974; Quashnock et al. 1989; Baym et al. 1996), anomalies in gravity (Alvarez-Gaume & Witten 1984; Contaldi et al. 2008; Kuntz 2019) can all produce parity-violation. Among these mechanisms, the simplest relevant for us is where there is a non-vanishing parity-odd trispectrum (Fourier-space analog of the 4PCF) in the scalar sector (Shiraishi et al. 2014; Shiraishi 2016), as curvature perturbations directly seed the galaxy density perturbations. Extending such models to include the full wave-vector dependence as well as “late-time” effects such as galaxy biasing and redshift-space distortions will be an important step on the road to comparing models to data and determining if the signal here observed is consistent with any such models. Especially given that the detection significance we report here is a cumulative result from hundreds to thousands of degrees of freedom, a model comparison will provide much more constraining power as it will enable use of shape information.

## Outlook

There are two main directions we suggest on the data side to further pursue the parity-odd sector. First, it is worth validating the measurement on different surveys with different designs. Upcoming surveys such as Dark Energy Spectroscopic Instrument (DESI; DESI Collaboration et al. 2016), Euclid (Laureijs et al. 2011), and Rubin (LSST Science Collaboration et al. 2009), which will provide much-larger samples and with different systematics, will offer expanded insight into the signal’s true origin. In particular, if our signal is of truly cosmological origin, it would likely be detectable in DESI at much higher statistical significance due to the larger volume. Thus, while in the present paper, making an error in our covariance matrix at the 20-30% level would produce a spurious detection, in DESI, we expect one would have to make a much larger error in the covariance. Since this is unlikely, finding evidence for this signal in DESI would be a strong indicator that either the signal is real, or that it is due to some as yet-undiscovered parity-breaking systematic shared between BOSS and DESI. Yet, given that DESI has a different imaging survey, a different instrument and method for fiberizing objects and taking spectra, and is on a different telescope than BOSS, this latter possibility also seems unlikely. Second, it is worth pursuing further approaches to search out systematics. For example, assuming a linear relationship between the systematics and the observed galaxy density field, one can study cross-correlations between them to reveal residual systematics (Ho et al. 2012). Alternatively, likelihood-based forward modelling can also be applied to search for unknown systematics (Lavaux et al. 2019). The first approach would require non-trivial extension of our current NPCF algorithm in order to compute cross correlations between fields, and the second requires injecting inferred systematics maps into mock catalogs. We leave this for future work.

Whether the detection of the current work ultimately turns out to be cosmological in origin or not, we hope that the careful and extensive treatment of systematics here offers a well-delineated avenue forward for future studies of parity-breaking early-Universe physics using 3D large-scale structure.

## ACKNOWLEDGEMENTS

We thank K. Dawson, H. Ding, D.P. Finkbeiner, A. Ginsburg, R. Guzmán, D. Jeong, A. Krolewski, E. Lada, W. Percival, O. Philcox, S. Portillo, D. Richardson, A.J. Ross, U. Seljak, D.J. Schlegel, M. Slepian, D. Slepian, D. Spergel, C. Steinhardt, C. Telesco, and W. Xue for useful conversations. We especially thank D.J. Eisenstein, A. de Mattia and A. G. Sánchez for extensive comments on the paper. We also thank A. Chuang, F.S. Kitaura, and C. Zhao for providing access to the PATCHY mocks. We thank E. Deumens of UF Research Computing for assistance using HiPerGator. Finally, we thank the members of the Slepian group for thoughtful discussions throughout the course of this work: N. Brown, J. Chellino, M. Hansen, F. Kamalinejad, K. Meigs, W. Ortolà Leonard, J. Sunseri, and C. Zhang.

**APPENDIX A: PARITY-ODD BASIS FUNCTIONS IN CARTESIAN REPRESENTATION**

Below we list parity-odd basis Functions in Cartesian representation for angular momentum up to  $\ell_{\max} = 4$ .

$$\begin{aligned}
 \mathcal{P}_{111}(\hat{\mathbf{r}}_1, \hat{\mathbf{r}}_2, \hat{\mathbf{r}}_3) &= -i \frac{3}{\sqrt{2}} (4\pi)^{-3/2} \hat{\mathbf{r}}_1 \cdot (\hat{\mathbf{r}}_2 \times \hat{\mathbf{r}}_3), \\
 \mathcal{P}_{122}(\hat{\mathbf{r}}_1, \hat{\mathbf{r}}_2, \hat{\mathbf{r}}_3) &= i \sqrt{\frac{45}{2}} (4\pi)^{-3/2} \hat{\mathbf{r}}_1 \cdot (\hat{\mathbf{r}}_2 \times \hat{\mathbf{r}}_3) (\hat{\mathbf{r}}_2 \cdot \hat{\mathbf{r}}_3), \\
 \mathcal{P}_{133}(\hat{\mathbf{r}}_1, \hat{\mathbf{r}}_2, \hat{\mathbf{r}}_3) &= -i \frac{15}{4} \sqrt{7} (4\pi)^{-3/2} \hat{\mathbf{r}}_1 \cdot (\hat{\mathbf{r}}_2 \times \hat{\mathbf{r}}_3) \left[ (\hat{\mathbf{r}}_2 \cdot \hat{\mathbf{r}}_3)^2 - \frac{1}{5} \right], \\
 \mathcal{P}_{144}(\hat{\mathbf{r}}_1, \hat{\mathbf{r}}_2, \hat{\mathbf{r}}_3) &= +i \frac{21\sqrt{15}}{4} (4\pi)^{-3/2} \hat{\mathbf{r}}_1 \cdot (\hat{\mathbf{r}}_2 \times \hat{\mathbf{r}}_3) \left[ (\hat{\mathbf{r}}_2 \cdot \hat{\mathbf{r}}_3)^3 - \frac{3}{7} (\hat{\mathbf{r}}_2 \cdot \hat{\mathbf{r}}_3) \right], \\
 \mathcal{P}_{223}(\hat{\mathbf{r}}_1, \hat{\mathbf{r}}_2, \hat{\mathbf{r}}_3) &= -i 15 \sqrt{\frac{5}{8}} (4\pi)^{-3/2} \hat{\mathbf{r}}_1 \cdot (\hat{\mathbf{r}}_2 \times \hat{\mathbf{r}}_3) \left[ (\hat{\mathbf{r}}_1 \cdot \hat{\mathbf{r}}_3) (\hat{\mathbf{r}}_2 \cdot \hat{\mathbf{r}}_3) - \frac{1}{5} \hat{\mathbf{r}}_1 \cdot \hat{\mathbf{r}}_2 \right], \\
 \mathcal{P}_{234}(\hat{\mathbf{r}}_1, \hat{\mathbf{r}}_2, \hat{\mathbf{r}}_3) &= +i \frac{4}{7 \cdot 15} (4\pi)^{-3/2} \hat{\mathbf{r}}_1 \cdot (\hat{\mathbf{r}}_2 \times \hat{\mathbf{r}}_3) \left[ (\hat{\mathbf{r}}_1 \cdot \hat{\mathbf{r}}_3) (\hat{\mathbf{r}}_2 \cdot \hat{\mathbf{r}}_3)^2 - \frac{1}{7} \hat{\mathbf{r}}_1 \cdot \hat{\mathbf{r}}_3 - \frac{2}{7} (\hat{\mathbf{r}}_1 \cdot \hat{\mathbf{r}}_2) (\hat{\mathbf{r}}_2 \cdot \hat{\mathbf{r}}_3) \right], \\
 \mathcal{P}_{333}(\hat{\mathbf{r}}_1, \hat{\mathbf{r}}_2, \hat{\mathbf{r}}_3) &= i \frac{5 \cdot 105}{6\sqrt{6}} (4\pi)^{-3/2} \hat{\mathbf{r}}_1 \cdot (\hat{\mathbf{r}}_2 \times \hat{\mathbf{r}}_3) \left[ (\hat{\mathbf{r}}_1 \cdot \hat{\mathbf{r}}_2) (\hat{\mathbf{r}}_1 \cdot \hat{\mathbf{r}}_3) (\hat{\mathbf{r}}_2 \cdot \hat{\mathbf{r}}_3) \right. \\
 &\quad \left. - \frac{1}{5} [(\hat{\mathbf{r}}_1 \cdot \hat{\mathbf{r}}_2)^2 + (\hat{\mathbf{r}}_1 \cdot \hat{\mathbf{r}}_3)^2 + (\hat{\mathbf{r}}_2 \cdot \hat{\mathbf{r}}_3)^2] + \frac{2}{25} \right], \\
 \mathcal{P}_{344}(\hat{\mathbf{r}}_1, \hat{\mathbf{r}}_2, \hat{\mathbf{r}}_3) &= -i \frac{15 \cdot 49}{4} \sqrt{\frac{5}{11}} (4\pi)^{-3/2} \hat{\mathbf{r}}_1 \cdot (\hat{\mathbf{r}}_2 \times \hat{\mathbf{r}}_3) \left[ (\hat{\mathbf{r}}_1 \cdot \hat{\mathbf{r}}_2) (\hat{\mathbf{r}}_1 \cdot \hat{\mathbf{r}}_3) (\hat{\mathbf{r}}_2 \cdot \hat{\mathbf{r}}_3)^2 \right. \\
 &\quad \left. - \frac{1}{5} (\hat{\mathbf{r}}_2 \cdot \hat{\mathbf{r}}_3)^3 + \frac{31}{245} (\hat{\mathbf{r}}_2 \cdot \hat{\mathbf{r}}_3) - \frac{3}{49} (\hat{\mathbf{r}}_1 \cdot \hat{\mathbf{r}}_2) (\hat{\mathbf{r}}_1 \cdot \hat{\mathbf{r}}_3) - \frac{2}{7} (\hat{\mathbf{r}}_1 \cdot \hat{\mathbf{r}}_3)^2 (\hat{\mathbf{r}}_2 \cdot \hat{\mathbf{r}}_3) - \frac{2}{7} (\hat{\mathbf{r}}_1 \cdot \hat{\mathbf{r}}_2)^2 (\hat{\mathbf{r}}_2 \cdot \hat{\mathbf{r}}_3) \right]. \quad (\text{A1})
 \end{aligned}$$

The eight functions above are the fundamental set and the remaining fifteen are generated by interchanging the arguments.

**APPENDIX B: DISCUSSION OF PEARSON CORRELATION COEFFICIENT BETWEEN TWO DATASETS**

Consider two datasets  $\mathbf{d}_1 = \mathbf{s}_1 + \mathbf{n}_1 - \langle \mathbf{s}_1 \rangle_{\text{all}} \equiv \Delta \mathbf{s}_1 + \mathbf{n}_1$  and  $\mathbf{d}_2 = \mathbf{s}_2 + \mathbf{n}_2 - \langle \mathbf{s}_2 \rangle_{\text{all}} \equiv \Delta \mathbf{s}_2 + \mathbf{n}_2$ , with  $\mathbf{s}$  the signal-to-noise ratio and  $\mathbf{n}$  the noise. These are data vectors—in our case, over all channels and radial bins.  $\langle \mathbf{s}_j \rangle_{\text{all}}$  denotes an average over all channels and radial bins.

Given two datasets, each with a certain intrinsic statistical noise drawn from Gaussian distribution, what is the maximum allowed correlation between them? In the decorrelated basis, we can treat each data point independently. We recall from Eq. (21) the definition of the Pearson correlation coefficient and insert in it the data vectors as defined above, finding

$$r_p = \frac{\sum_{i=1}^{N_{\text{bin}}} (d_1^{(i)} - \langle \mathbf{d}_1 \rangle_{\text{all}}) (d_2^{(i)} - \langle \mathbf{d}_2 \rangle_{\text{all}})}{\sqrt{\sum_{i=1}^{N_{\text{bin}}} (d_1^{(i)} - \langle \mathbf{d}_1 \rangle_{\text{all}})^2 \sum_{i=1}^{N_{\text{bin}}} (d_2^{(i)} - \langle \mathbf{d}_2 \rangle_{\text{all}})^2}}. \quad (\text{B1})$$

Assuming a linear relation between the signals from the two samples, *i.e.*  $\mathbf{s}_2 = b \mathbf{s}_1$ , that the average observed signal is given by  $\langle \mathbf{d}_j \rangle_{\text{all}} = N_{\text{bin}}^{-1} \sum_{i=1}^{N_{\text{bin}}} s_j^{(i)}$  for  $j = 1, 2$ , and that the noise vanishes after averaging, we have for the numerator

$$\sum_{i=1}^{N_{\text{bin}}} (d_1^{(i)} - \langle \mathbf{d}_1 \rangle_{\text{all}}) (d_2^{(i)} - \langle \mathbf{d}_2 \rangle_{\text{all}}) = \langle \mathbf{s}_1 \mathbf{s}_2 \rangle_{\text{all}} + \langle \mathbf{s}_1 \mathbf{n}_2 \rangle_{\text{all}} + \langle \mathbf{s}_2 \mathbf{n}_1 \rangle_{\text{all}} + \langle \mathbf{n}_1 \mathbf{n}_2 \rangle_{\text{all}} - \langle \mathbf{s}_1 \rangle_{\text{all}} \langle \mathbf{s}_2 \rangle_{\text{all}}. \quad (\text{B2})$$

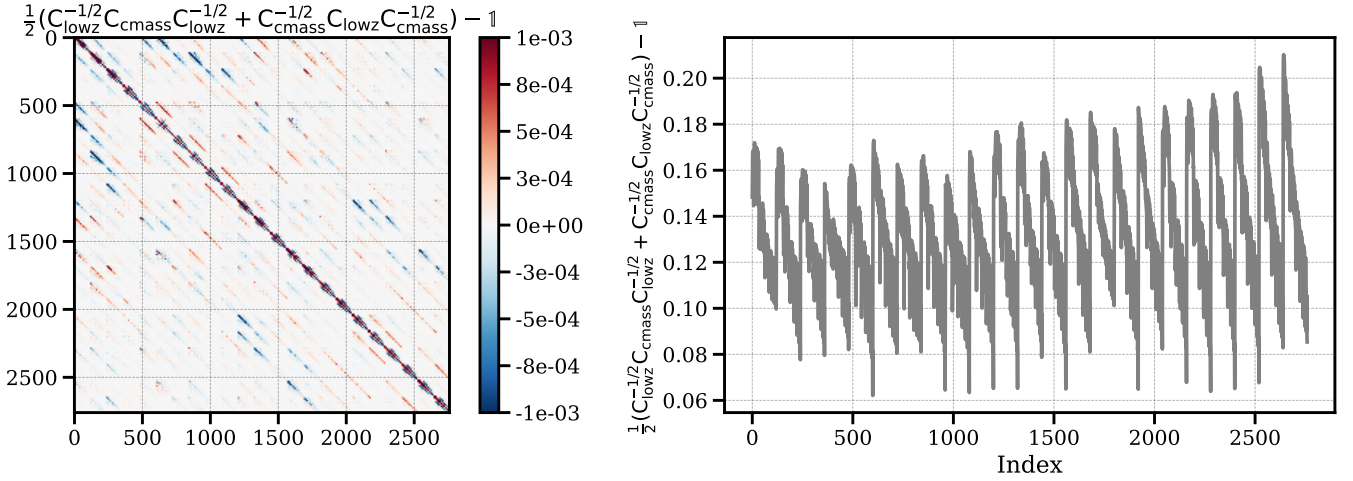
For the denominator, we use that

$$\begin{aligned}
 \sum_{i=1}^{N_{\text{bin}}} (d_j^{(i)} - \langle \mathbf{d}_j \rangle_{\text{all}})^2 &= \langle \mathbf{s}_j^2 \rangle_{\text{all}} - \langle \mathbf{s}_j \rangle_{\text{all}}^2 + \langle \mathbf{n}_j^2 \rangle_{\text{all}} + 2 \langle \mathbf{s}_j \mathbf{n}_j \rangle_{\text{all}} \\
 &\equiv \sigma_{\mathbf{s}_j}^2 + \sigma_{\mathbf{n}_j}^2 + 2 \langle \mathbf{s}_j \mathbf{n}_j \rangle_{\text{all}}. \quad (\text{B3})
 \end{aligned}$$

The correlation coefficient is then

$$r_p = \frac{\langle \mathbf{s}_1 \mathbf{s}_2 \rangle_{\text{all}} - \langle \mathbf{s}_1 \rangle_{\text{all}} \langle \mathbf{s}_2 \rangle_{\text{all}} + \langle \mathbf{s}_1 \mathbf{n}_2 \rangle_{\text{all}} + \langle \mathbf{n}_1 \mathbf{s}_2 \rangle_{\text{all}} + \langle \mathbf{n}_1 \mathbf{n}_2 \rangle_{\text{all}}}{\sqrt{(\sigma_{\mathbf{s}_1}^2 + 2 \langle \mathbf{s}_1 \mathbf{n}_1 \rangle_{\text{all}} + \sigma_{\mathbf{n}_1}^2)(\sigma_{\mathbf{s}_2}^2 + 2 \langle \mathbf{s}_2 \mathbf{n}_2 \rangle_{\text{all}} + \sigma_{\mathbf{n}_2}^2)}}. \quad (\text{B4})$$

First we discuss the low signal-to-noise limit for  $\mathbf{s}_1 \ll \mathbf{n}_1$  and  $\mathbf{s}_2 \ll \mathbf{n}_2$ . We assume the correlation between the signal and noise is zero, and that the signal covariance is much smaller than that of the noise, *i.e.*  $\langle \mathbf{s}_1 \mathbf{s}_2 \rangle_{\text{all}} \ll \langle \mathbf{n}_1 \mathbf{n}_2 \rangle_{\text{all}}$ . We find in this



**Figure C1.** Comparison of the CMASS and LOWZ analytic covariance matrices, but with their volumes set to be equal. We do this because an overall scaling will not adversely impact a cross-correlation analysis, but non-trivial differences in the matrices do and it is these we want to display. We have mapped each set of channels and radial bins into 1D indices to make a 2D plot. The *lefthand* panel shows a symmetrized test; notably, the diagonal is non-vanishing after subtracting the identity matrix. The non-vanishing off-diagonal elements present in the lefthand panel indicate that the eigenbases of the two matrices also differ, as we might expect given that the shot noise impacts the covariance non-trivially and the two matrices are computed with different shot-noises (see Table 1). The *right-hand* panel shows just the diagonal of our test matrix; the residual is likely due to the slightly different number densities.

limit that

$$r_p = \frac{\langle \mathbf{n}_1 \mathbf{n}_2 \rangle_{\text{all}}}{\sigma_{\mathbf{n}_1} \sigma_{\mathbf{n}_2}} = 0. \quad (\text{B5})$$

where to obtain the last equality we assumed that the noise is independent in each bin (as is assumed for a Pearson analysis, and as is the case once we have rotated to the decorrelated basis as discussed in the main text, §5.2).

Second, in the high signal-to-noise limit that  $\mathbf{n}_1 \ll \mathbf{s}_1$  and  $\mathbf{n}_2 \ll \mathbf{s}_2$ , and now with  $\langle \mathbf{n}_1 \mathbf{n}_2 \rangle_{\text{all}} \ll \langle \mathbf{s}_1 \mathbf{s}_2 \rangle_{\text{all}}$ , we find

$$r_p = \frac{\langle \mathbf{s}_1 \mathbf{s}_2 \rangle_{\text{all}} - \langle \mathbf{s}_1 \rangle_{\text{all}} \langle \mathbf{s}_2 \rangle_{\text{all}}}{\sigma_{\mathbf{s}_1} \sigma_{\mathbf{s}_2}}. \quad (\text{B6})$$

We note that  $\sigma$  here is the rms of the relevant vector *over radial bins and channels*, not any kind of statistical error. It is simply a measure of how much that vector varies as the angular momenta and the radial bins are changed. This equation thus shows that, even if one had a signal of very high amplitude (and hence high S/N and high detection significance), if one had little variation over channels and radial bins, one would obtain a low  $r_p$ . Put simply, the detection significance and  $r_p$  are independent.

## APPENDIX C: COMPARISON OF CMASS AND LOWZ ANALYTIC COVARIANCES

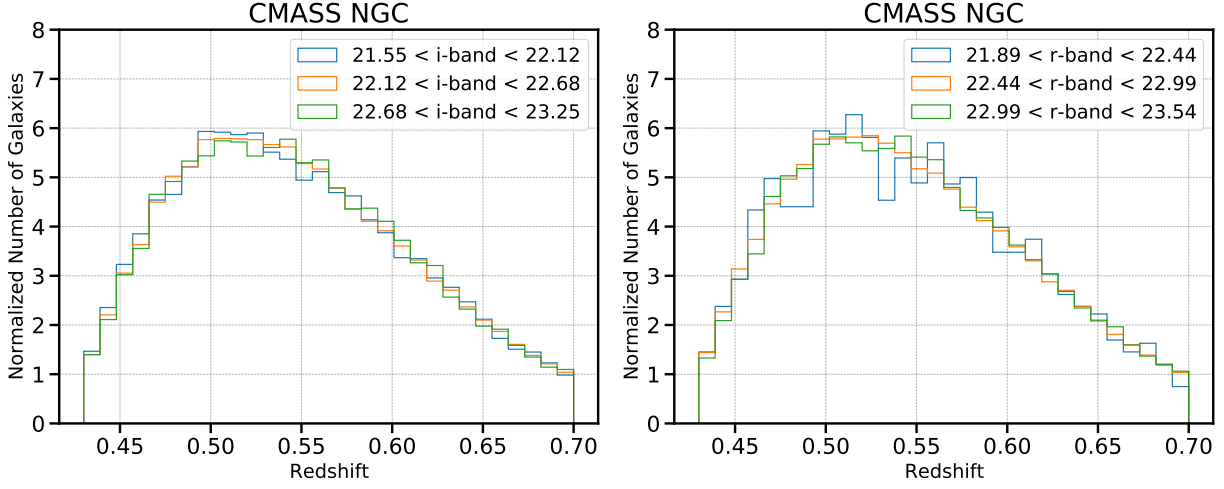
Here we compare the analytic covariances for the two samples (Fig. C1). This test shows that they differ non-trivially, thus precluding a cross-correlation analysis between the two samples if we wished to diagonalize as described in §5.2.

## APPENDIX D: REDSHIFT DEPENDENCE ON IMAGING DEPTH

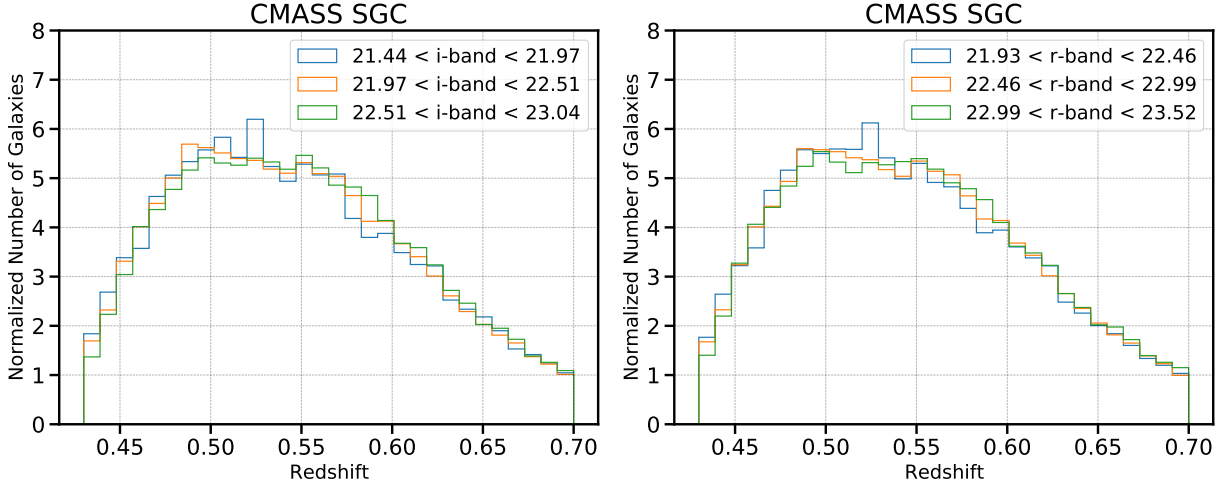
In this section we check how imaging depth affect number density as a function of redshift depends. In particular, we focus on the CMASS sample. Fig. D1 and D2 illustrate the redshift dependence on *r*-band and *i*-band imaging depth for northern and southern hemisphere, respectively. From both plots for *r*-band and *i*-band we saw that  $n(z)$  has very similar distribution for the three imaging depth bins, which implies that imaging depth does not have a strong impact on  $n(z)$ .

## APPENDIX E: COARSER RADIAL BINNING

Here we show the results of using a coarser radial binning by a factor of roughly two relative to our ten-bin case. We generally see lower detection significance, consistent with the theory that “internal cancellation” occurs to a greater extent in this coarser binning. Our results are displayed in Fig. E1.



**Figure D1.** Redshift dependence on imaging depth for CMASS NGC. *Left:* The normalized galaxy number counts as a function of redshift for the *i*-band, where we split the sample into three bins in imaging depth. *Right:* Same as the left but for the *r*-band. Here we can see that the three bins in *i*- and *r*-band have very similar  $n(z)$ , which implies that the imaging depth is unlikely to impact our analysis.



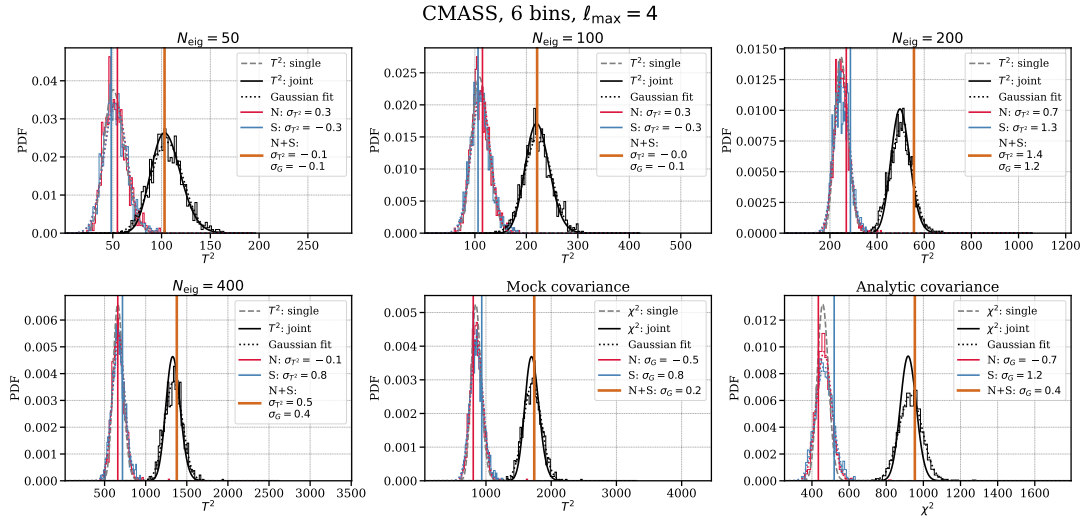
**Figure D2.** Same as Fig. D1 but for the CMASS SGC. Again, the *i*- and *r*-band have very similar  $n(z)$  distributions for the three bins, which implies that image depth does not strongly impact our analysis for SGC.

## APPENDIX F: MAXIMUM SCALE CUT AND MINIMUM BIN SEPARATION

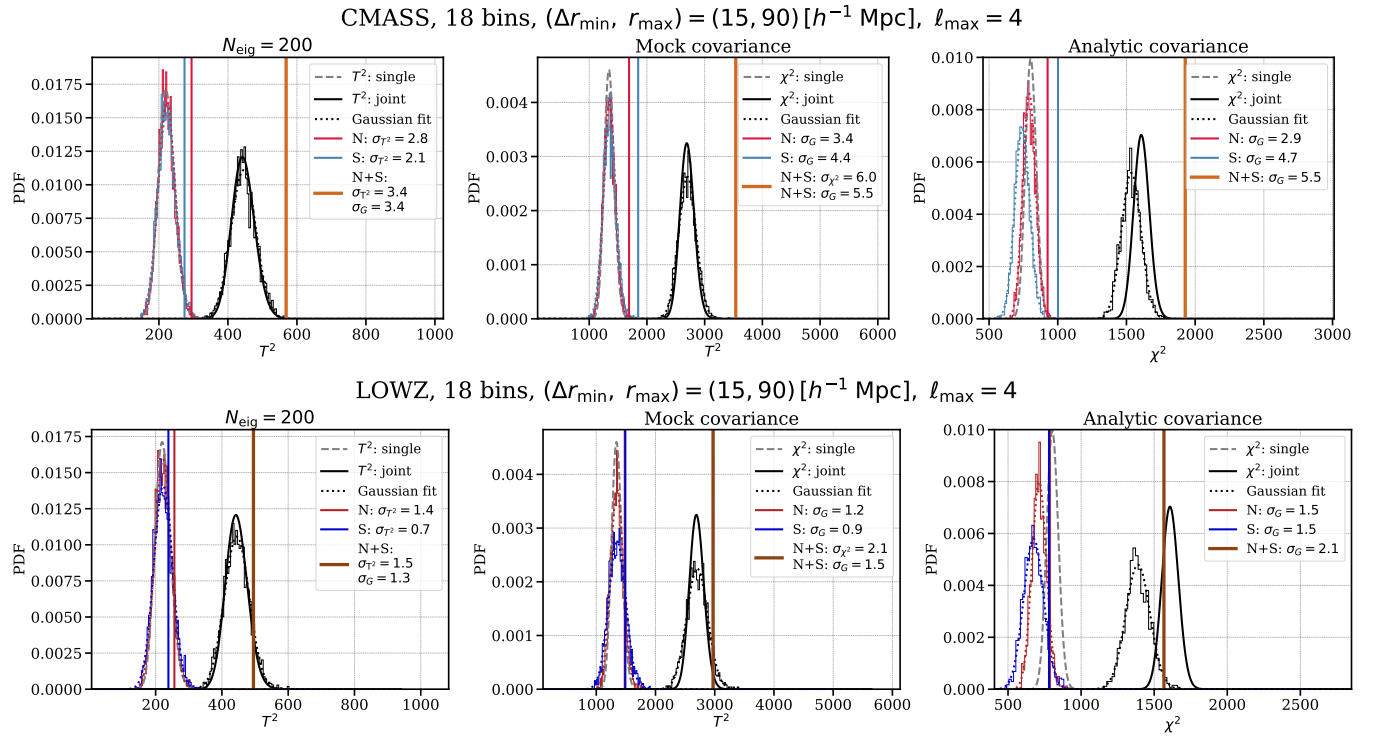
To explore whether use of small scales makes a difference in our analysis (as these are the scales on which the mocks may be most likely to imperfectly mirror the data, due to approximate treatment of non-linear structure formation), we test the detection significances by applying further restrictions to the radial bins. First, we force the minimum radial bin separation to be  $\Delta r \geq 15 h^{-1}\text{Mpc}$  such that the results are less sensitive to small scales. Second, we vary the maximum radial bin  $r_{\text{max}} = 90 h^{-1}\text{Mpc}$  and  $r_{\text{max}} = 130 h^{-1}\text{Mpc}$  such that we are less affected by any mismatch between mocks and data around the BAO position and towards larger scales (as was seen in some of the 2PCF measurements).

The results are shown in Figs. F1 and F2. In all these cases the detection significance reduces compared to the fiducial case (without minimum bin separation or maximum radial bin) given that the number degrees of freedom is reduced. However, we still observe non-negligible detection significances for all cases. It is worth pointing out that there are in total only  $35 \times 23 = 805$  degrees of freedom when using  $r_{\text{max}} = 90 h^{-1}\text{Mpc}$ , which allows us to use the mock covariance directly. Despite the mismatch between the analytic covariance matrix and the expected  $\chi^2$  distribution, there is no substantial difference in the detection significances.

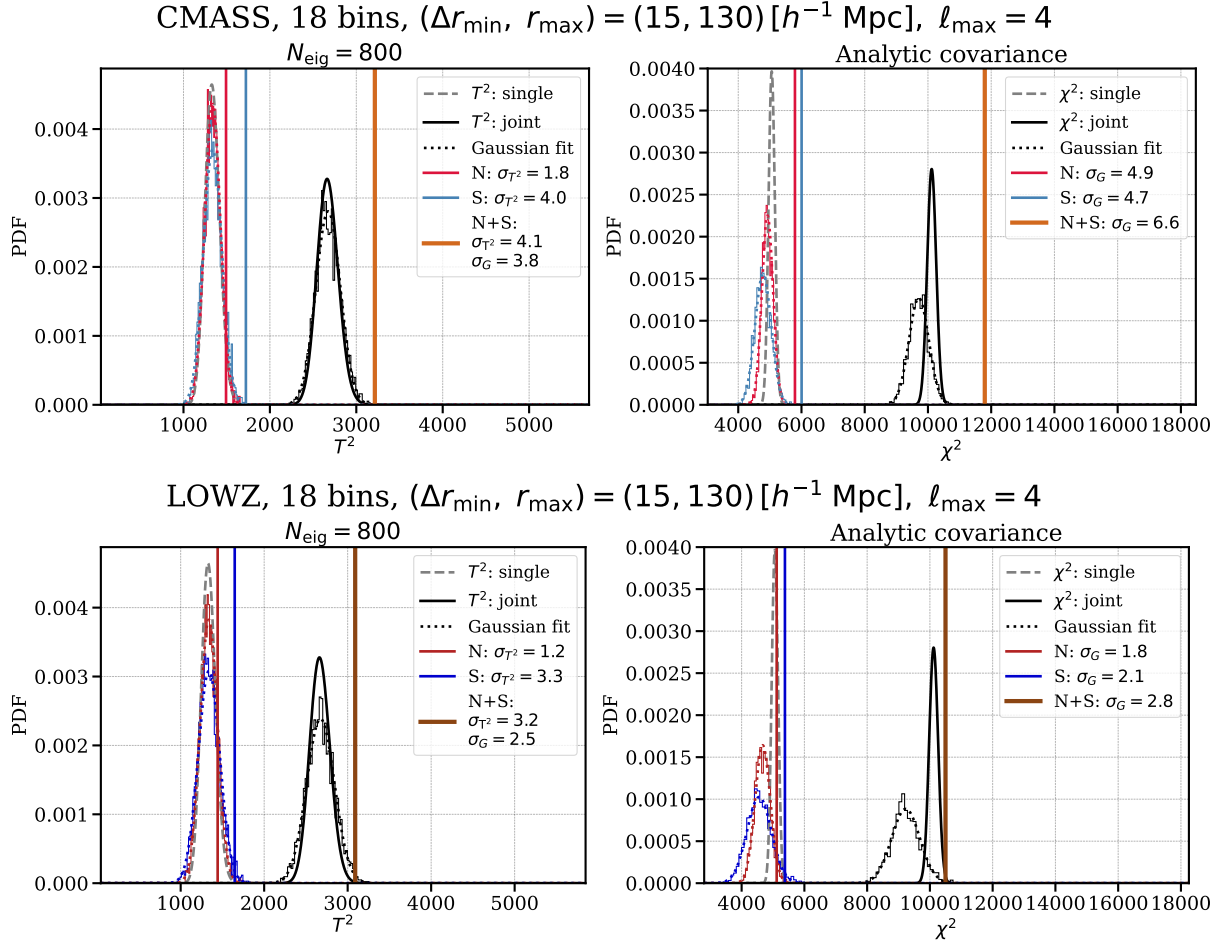
## APPENDIX G: ALL ANGULAR CHANNELS



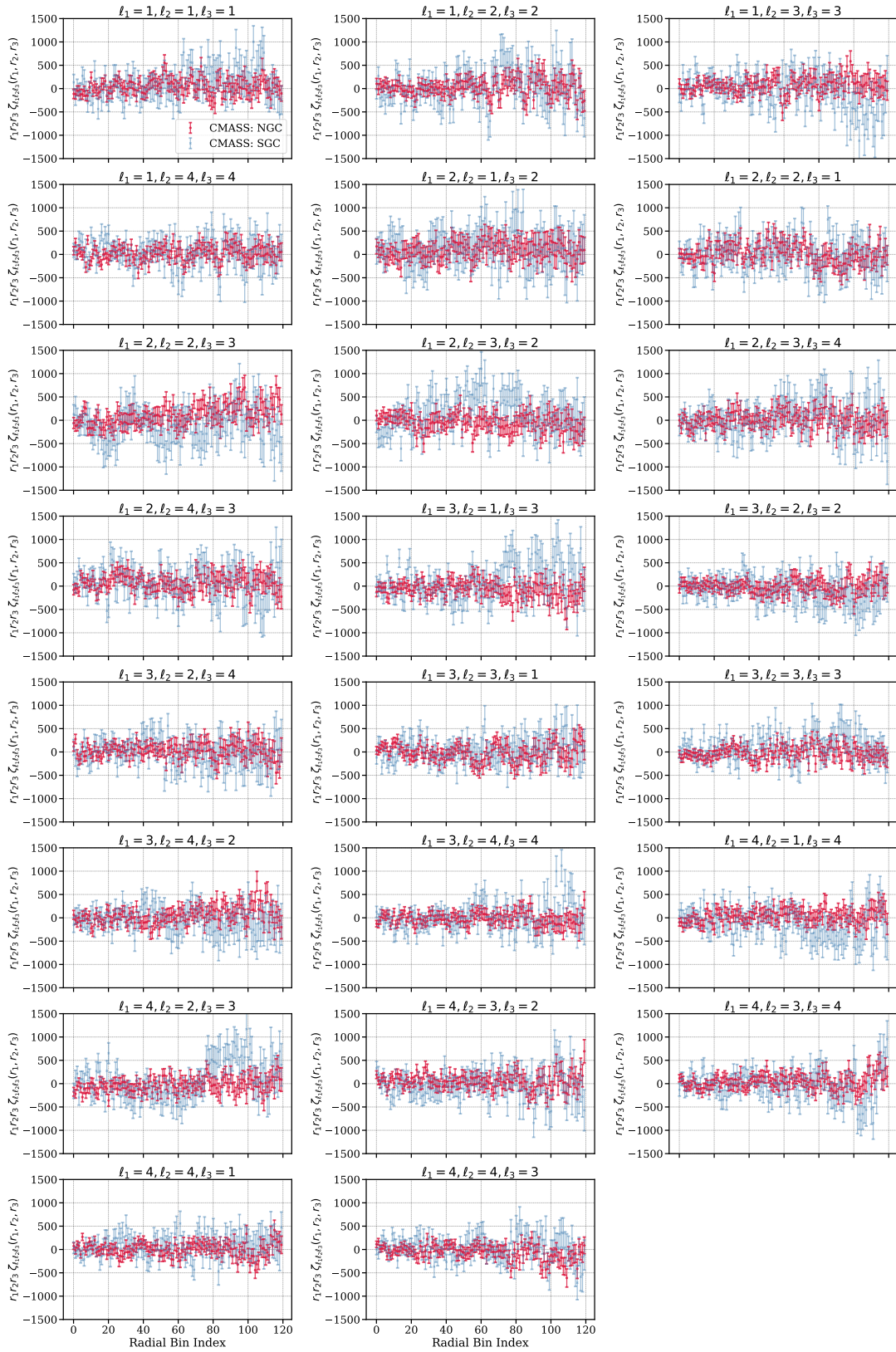
**Figure E1.** Here we show a compressed analysis (§5.1.2), and direct approaches with both the mock covariance (lower middle panel) and the analytic covariance (lower right panel) for a 4PCF with just six radial bins. This substantially reduces the number of degrees of freedom, permitting use of the mock covariance. However, the detection significance is also degraded; we attribute this to much larger “internal cancellation” (§3.2) than in our ten- and eighteen-bin analyses.



**Figure F1.** Upper: CMASS sample with minimum radial bin separation  $\Delta r = 15 h^{-1} \text{Mpc}$  and maximum radial bin  $r_{\max} = 90 h^{-1} \text{Mpc}$ . The leftmost column uses the data compression method with  $N_{\text{eig}} = 200$ . The middle column directly uses the mock covariance. The rightmost column uses the analytic covariance. Lower: Same as the upper row but for LOWZ. In all these cases the detection significance is reduced relative to that in our fiducial analysis due to having fewer degrees of freedom.

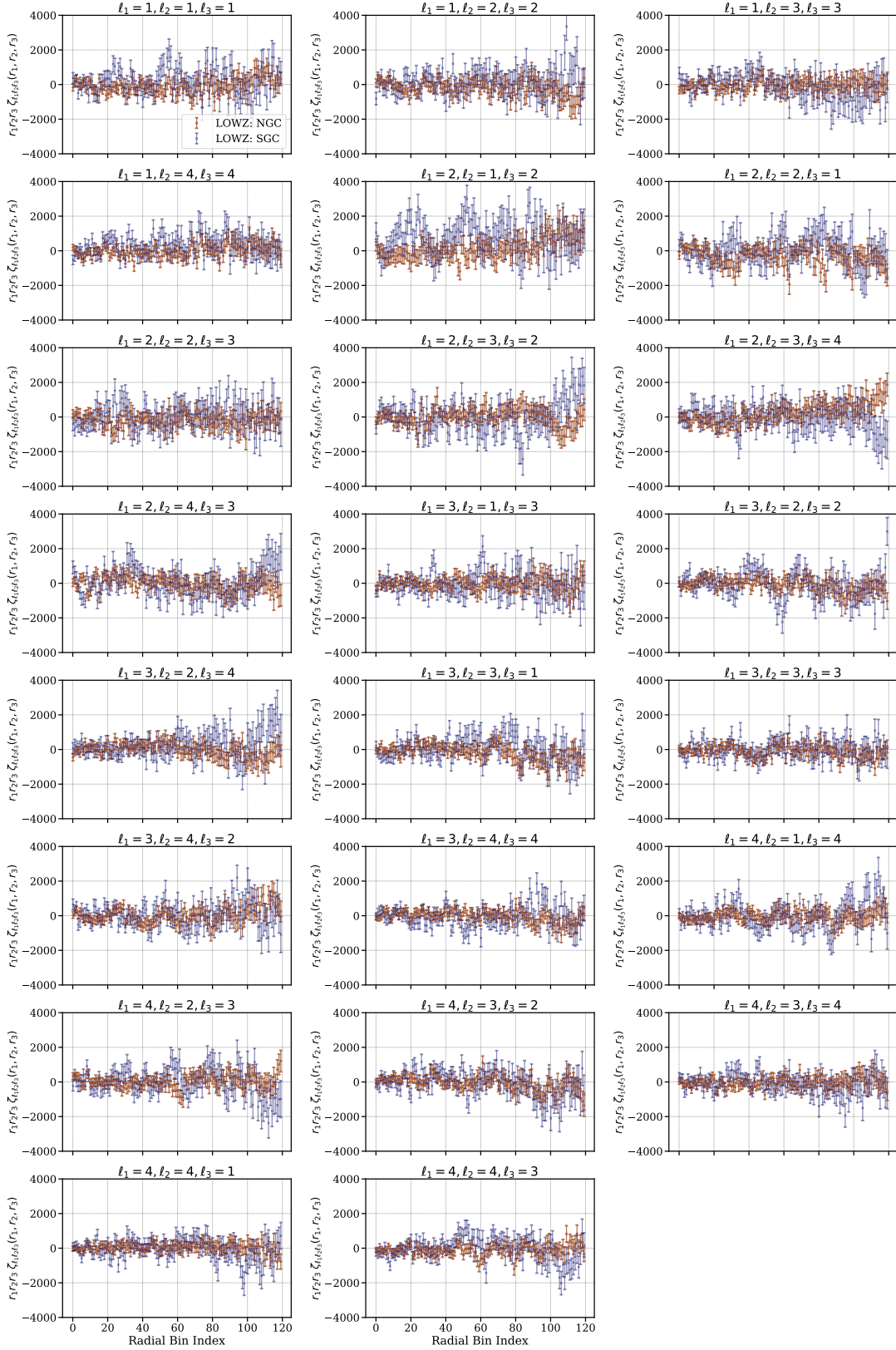


**Figure F2.** *Upper:* CMASS sample with minimum radial bin separation  $\Delta r = 15 h^{-1} \text{ Mpc}$  and maximum radial bin  $r_{\max} = 130 h^{-1} \text{ Mpc}$ . The left column uses the data compression method with  $N_{\text{eig}} = 200$ . The right column uses the analytic covariance. *Lower:* Same as the upper row but for LOWZ. As in Fig. F1 the detection significance is reduced compared to that in our fiducial analysis.

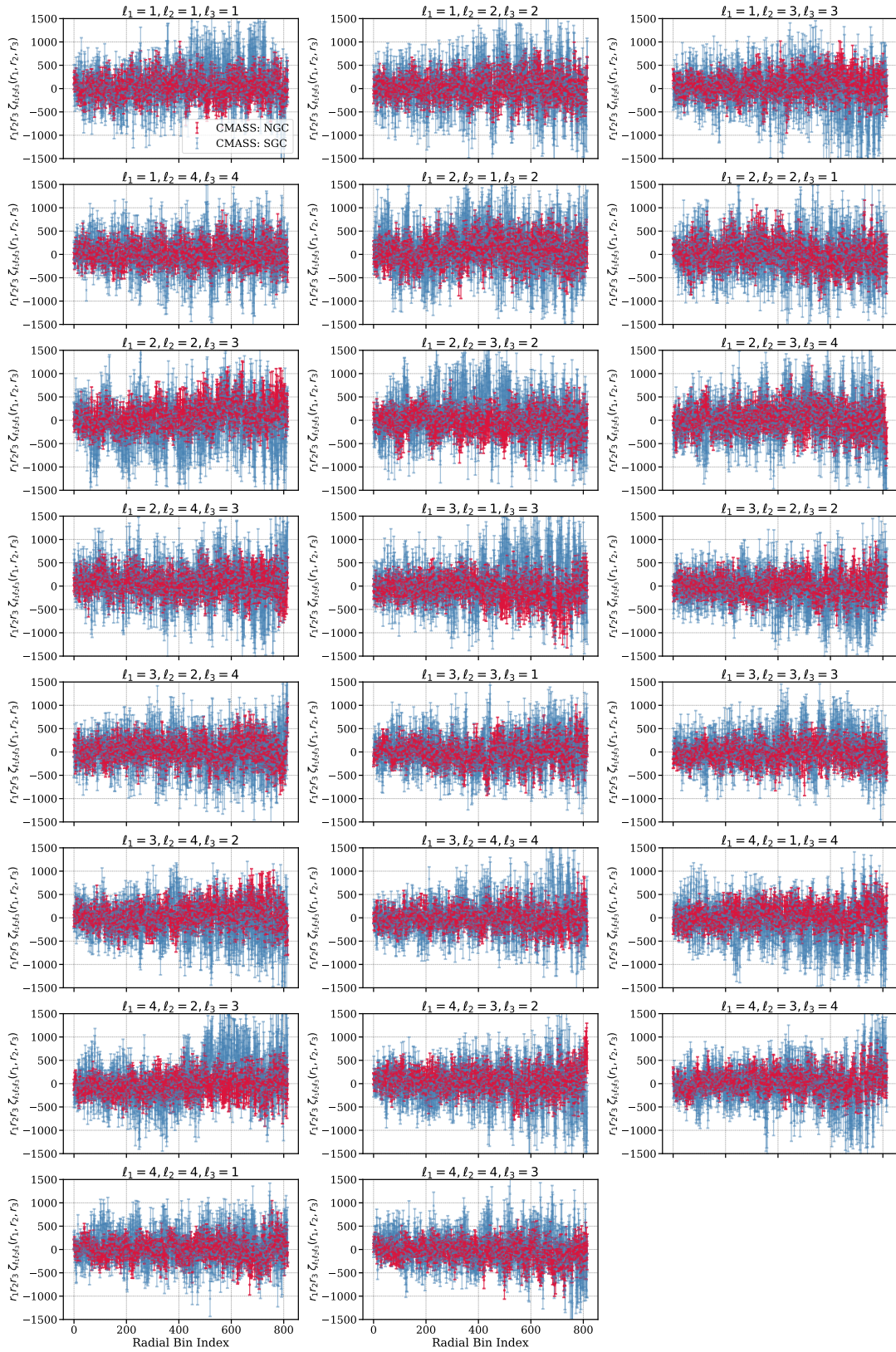


**Figure G1.** The parity-odd 4PCF for the BOSS CMASS data, with NGC in red and SGC in blue. The plot includes all the angular channels for ten radial bins. The error bars are the rms of the PATCHY mocks.

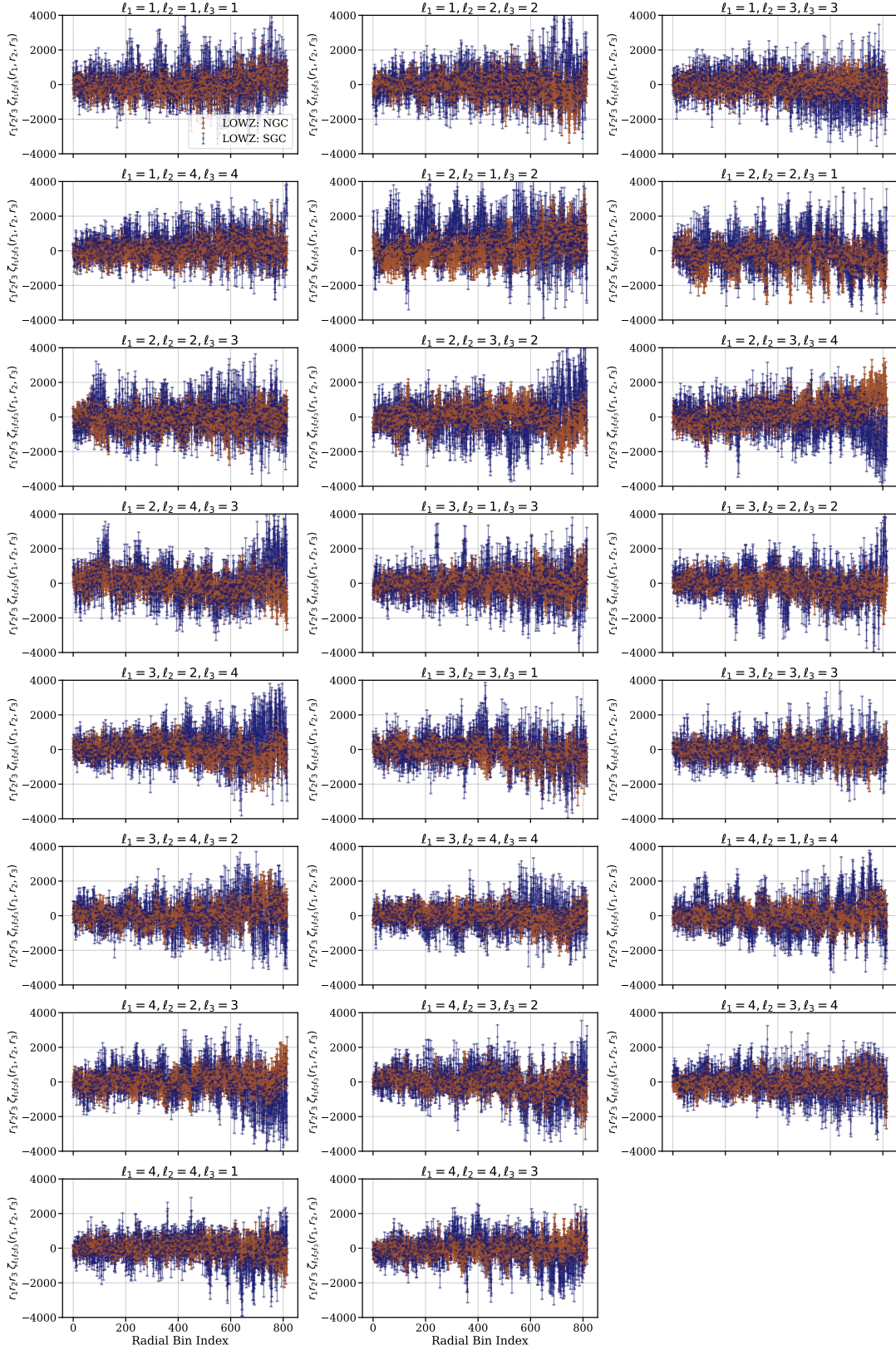




**Figure G2.** The parity-odd 4PCF for the BOSS LOWZ data including all the angular channels for **ten** radial bins. NGC is in brown and SGC in blue; the error bars are the rms of the PATCHY mocks.



**Figure G3.** The parity-odd 4PCF for the BOSS CMASS data including all the angular channels for **eighteen** radial bins. NGC is in red and SGC in blue; the error bars are the rms of the PATCHY mocks.



**Figure G4.** The parity-odd 4PCF for the BOSS LOWZ data including all the angular channels for **eighteen** radial bins. NGC is in brown and SGC in blue; the error bars are the rms of the PATCHY mocks.

## REFERENCES

- Abolhasani A. A., Akhshik M., Emami R., Firouzjahi H., 2016, *J. Cosmology Astropart. Phys.*, 2016, 020–020
- Albrecht A., Steinhardt P. J., 1982, *Phys. Rev. Lett.*, 48, 1220
- Alexander S., Yunes N., 2009, *Phys. Rep.*, 480, 1
- Alexander S. H. S., Peskin M. E., Sheikh-Jabbari M. M., 2006, *Physical Review Letters*, 96
- Alvarez-Gaume L., Witten E., 1984, *Nucl. Phys. B*, 234, 269
- BOSS collaboration et al., 2017, *MNRAS*, 470, 2617
- Barnaby N., Namba R., Peloso M., 2011, *J. Cosmology Astropart. Phys.*, 2011, 009
- Bartolo N., Matarrese S., Peloso M., Shiraishi M., 2015, *J. Cosmology Astropart. Phys.*, 2015, 039–039
- Baym G., Bödeker D., McLerran L., 1996, *Phys. Rev. D*, 53, 662
- Bernardeau F., Colombi S., Gaztañaga E., Scoccimarro R., 2002, *Physics Reports*, 367, 1–248
- Beutler F., et al., 2014, *MNRAS*, 443, 1065
- Beutler F., Castorina E., Zhang P., 2019, *Journal of Cosmology and Astroparticle Physics*, 2019, 040–040
- Brizard A. J., Murayama H., Wurtele J. S., 2000, *Phys. Rev. E*, 61, 4410
- Cahn R. N., Slepian Z., 2020, arXiv e-prints, p. arXiv:2010.14418
- Cahn R. N., Slepian Z., Hou J., 2021, arXiv e-prints, p. arXiv:2110.12004
- Cannon R., et al., 2006, *MNRAS*, 372, 425
- Contaldi C. R., Magueijo J., Smolin L., 2008, *Physical Review Letters*, 101
- DESI Collaboration et al., 2016, arXiv e-prints, p. arXiv:1611.00036
- Dai L., Jeong D., Kamionkowski M., 2013, *Physical Review D*, 88
- Dawson K. S., et al., 2013, *AJ*, 145, 10
- Deser S., Jackiw R., Templeton S., 1982a, *Phys. Rev. Lett.*, 48, 975
- Deser S., Jackiw R., Templeton S., 1982b, *Annals of Physics*, 140, 372
- Diego-Palazuelos P., et al., 2022, Cosmic Birefringence from Planck Data Release 4 (arXiv:2201.07682)
- Dyda S., Flanagan É. É., Kamionkowski M., 2012, *Phys. Rev. D*, 86, 124031
- Eisenstein D. J., et al., 2001, *AJ*, 122, 2267
- Eisenstein D. J., et al., 2011, *AJ*, 142, 72
- Feldman H. A., Kaiser N., Peacock J. A., 1994, *The Astrophysical Journal*, 426, 23
- Freidel L., Minic D., Takeuchi T., 2005, *Phys. Rev. D*, 72, 104002
- Friesen B., et al., 2017, *Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis*
- Fry J. N., Peebles P. J. E., 1978, *ApJ*, 221, 19
- García K., Slepian Z., 2020, Improving the Line of Sight for the Anisotropic 3-Point Correlation Function of Galaxies: Centroid and Unit-Vector-Average Methods Scaling as  $\mathcal{O}(N^2)$  (arXiv:2011.03503)
- Galdi D., Verde L., 2022, Integrated trispectrum detection from BOSS DR12 NGC CMASS (arXiv:2201.06932)
- Hahn C., Scoccimarro R., Blanton M. R., Tinker J. L., Rodríguez-Torres S. A., 2017, *MNRAS*, 467, 1940
- Hirata C. M., 2009, *MNRAS*, 399, 1074
- Ho S., et al., 2012, *ApJ*, 761, 14
- Hotelling H., 1953, *Journal of the royal statistical society series b-methodological*, 15, 193
- Hou J., Cahn R. N., Philcox O. H. E., Slepian Z., 2021a, arXiv e-prints, p. arXiv:2108.01714
- Hou J., et al., 2021b, *MNRAS*, 500, 1201
- Jackiw R., Pi S. Y., 2003, *Phys. Rev. D*, 68, 104012
- Jeong D., Kamionkowski M., 2012, *Physical Review Letters*, 108
- Jeong D., Schmidt F., 2019, arXiv e-prints, p. arXiv:1906.05198
- Kaiser N., 1987, *MNRAS*, 227, 1
- Kamionkowski M., Souradeep T., 2011, *Phys. Rev. D*, 83, 027301
- Kenney J. F., 1947, *Mathematics of statistics*.
- Kerscher M., Szapudi I., Szalay A. S., 2000, *ApJ*, 535, L13
- Kitaura F.-S., et al., 2016, *MNRAS*, 456, 4156–4173
- Kullback S., Leibler R. A., 1951, *The Annals of Mathematical Statistics*, 22, 79
- Kuntz I., 2019, *Foundations of Physics*, 49, 191
- LSST Science Collaboration et al., 2009, LSST Science Book, Version 2.0 (arXiv:0912.0201)
- Landy S. D., Szalay A. S., 1993, *ApJ*, 412, 64
- Laureijs R., et al., 2011, arXiv e-prints, p. arXiv:1110.3193
- Lavaux G., Jasche J., Leclercq F., 2019, arXiv e-prints, p. arXiv:1909.06396
- Lee T. D., Yang C. N., 1956, *Phys. Rev.*, 104, 254
- Li M., Zhao D., 2022, *Physics Letters B*, 827, 136968
- Li Y., Hu W., Takada M., 2014, *Physical Review D*, 89
- Linde A. D., 1982, *Physics Letters B*, 108, 389
- Linde A. D., 1983, *Physics Letters B*, 129, 177
- Liu T., Tong X., Wang Y., Xianyu Z.-Z., 2020, *Journal of High Energy Physics*, 2020
- Lue A., Wang L., Kamionkowski M., 1999, *Physical Review Letters*, 83, 1506–1509
- Maraston C., et al., 2013, *MNRAS*, 435, 2764–2792
- Masters K. L., et al., 2011, *MNRAS*, 418, 1055
- Masui K. W., Pen U.-L., Turok N., 2017, *Physical Review Letters*, 118
- Minami Y., Komatsu E., 2020, *Phys. Rev. Lett.*, 125, 221301
- Nishizawa A., Kobayashi T., 2018, *Phys. Rev. D*, 98, 124018
- Orlando G., Pironi M., Ricciardone A., 2021, *J. Cosmology Astropart. Phys.*, 2021, 069

- Özsoy O., 2021, *Phys. Rev. D*, **104**, 123523
- Parejko J. K., et al., 2013, *MNRAS*, **429**, 98
- Peebles P. J. E., 2001, The Galaxy and Mass N-Point Correlation Functions: a Blast from the Past ([arXiv:astro-ph/0103040](https://arxiv.org/abs/astro-ph/0103040))
- Philcox O. H. E., Slepian Z., Hou J., Warner C., Cahn R. N., Eisenstein D. J., 2021a, *MNRAS*,
- Philcox O. H. E., Hou J., Slepian Z., 2021b, arXiv e-prints, p. [arXiv:2108.01670](https://arxiv.org/abs/2108.01670)
- Planck Collaboration et al., 2020, *A&A*, **641**, A6
- Pogosian L., Wyman M., 2008, *Physical Review D*, **77**
- Portillo S. K. N., Slepian Z., Burkhart B., Kahraman S., Finkbeiner D. P., 2018, *The Astrophysical Journal*, **862**, 119
- Putter R. d., Wagner C., Mena O., Verde L., Percival W. J., 2012, *Journal of Cosmology and Astroparticle Physics*, 2012, 019–019
- Quashnock J. M., Loeb A., Spergel D. N., 1989, *ApJ*, **344**, L49
- Reid B., et al., 2016, *MNRAS*, **455**, 1553
- Rodríguez-Torres S. A., et al., 2017, *MNRAS*, **468**, 728–740
- Ross A. J., et al., 2012, *Monthly Notices of the Royal Astronomical Society*, **424**, 564–590
- Ross A. J., et al., 2016, *Monthly Notices of the Royal Astronomical Society*, **464**, 1168–1191
- Sabiu C. G., Hoyle B., Kim J., Li X.-D., 2019, *The Astrophysical Journal Supplement Series*, **242**, 29
- Sakharov A. D., 1967, *JETP Lett.*, **5**, 24
- Schmidt F., Pajer E., Zaldarriaga M., 2014, *Physical Review D*, **89**
- Scoccimarro R., Couchman H. M. P., Frieman J. A., 1999, *The Astrophysical Journal*, **517**, 531–540
- Sellentin E., Heavens A. F., 2016, *MNRAS*, **456**, L132
- Shiraishi M., 2016, *Physical Review D*, **94**
- Shiraishi M., Nitta D., Yokoyama S., Ichiki K., Takahashi K., 2011, *Progress of Theoretical Physics*, **125**, 795
- Shiraishi M., Komatsu E., Peloso M., 2014, *Journal of Cosmology and Astroparticle Physics*, 2014, 027–027
- Slepian Z., Eisenstein D. J., 2015a, A new look at lines of sight: using Fourier methods for the wide-angle anisotropic 2-point correlation function ([arXiv:1510.04809](https://arxiv.org/abs/1510.04809))
- Slepian Z., Eisenstein D. J., 2015b, *Monthly Notices of the Royal Astronomical Society*, **454**, 4142–4158
- Slepian Z., Eisenstein D. J., 2015c, *Monthly Notices of the Royal Astronomical Society: Letters*, **455**, L31–L35
- Slepian Z., Eisenstein D. J., 2018, *Monthly Notices of the Royal Astronomical Society*, **478**, 1468–1483
- Slepian Z., et al., 2017a, *Monthly Notices of the Royal Astronomical Society*, **468**, 1070–1083
- Slepian Z., et al., 2017b, *Monthly Notices of the Royal Astronomical Society*, **469**, 1738–1751
- Slepian Z., et al., 2017c, *Monthly Notices of the Royal Astronomical Society*, **474**, 2109–2115
- Slepian Z., Li Y., Schmittfull M., Vlah Z., 2019, arXiv e-prints, p. [arXiv:1912.00065](https://arxiv.org/abs/1912.00065)
- Slepian Z., Warner C., Hou J., Cahn R. N., 2021, CADENZA: Harmonics on the GPU for N-Point Correlation Functions
- Smee S. A., et al., 2013, *The Astronomical Journal*, **146**, 32
- Soda J., Kodama H., Nozawa M., 2011, *Journal of High Energy Physics*, 2011, 67
- Sorbo L., 2011, *J. Cosmology Astropart. Phys.*, 2011, 003
- Szapudi I., Szalay A. S., 1998, *ApJ*, **494**, L41
- Vilenkin A., 1978, *Phys. Rev. Lett.*, **41**, 1575
- Vilenkin A., Leahy D. A., 1982, *ApJ*, **254**, 77
- Wang A., Wu Q., Zhao W., Zhu T., 2013, *Phys. Rev. D*, **87**, 103512
- Wishart J., 1928, *Biometrika*, **20A**, 32
- Witten E., 1989, *Communications in Mathematical Physics*, **121**, 351
- Wu C. S., Ambler E., Hayward R. W., Hoppes D. D., Hudson R. P., 1957, *Phys. Rev.*, **105**, 1413
- Yunes N., O’Shaughnessy R., Owen B. J., Alexander S., 2010, *Phys. Rev. D*, **82**, 064017
- Zhu T., Zhao W., Huang Y., Wang A., Wu Q., 2013, *Phys. Rev. D*, **88**, 063508
- ’t Hooft G., 1974, *Nuclear Physics B*, **79**, 276