

Study of the connected four-point correlation function of galaxies from the DESI Data Release 1 luminous red galaxy sample

J. Hou^{1,2,3,4,*} R. N. Cahn,⁵ J. Aguilar,⁵ S. Ahlen,⁷ D. Bianchi,^{8,9} D. Brooks,¹⁰ T. Claybaugh,⁵ P. Doel,¹⁰ S. Ferraro,^{5,12} J. E. Forero-Romero,^{13,14} E. Gaztañaga,^{15,16,17} L. Le Guillou,²⁹ G. Gutierrez,¹⁸ K. Honscheid,^{19,20,21} D. Huterer,^{22,23} M. Ishak,²⁴ R. Joyce,²⁵ S. Juneau,²⁵ R. Kehoe,²⁶ D. Kirkby,²⁷ T. Kisner,⁵ A. Kremin,⁵ C. Lamman,²⁸ M. Landriau,⁵ A. de la Macorra,¹¹ M. Manera,^{30,31} A. de Mattia,⁶ R. Miquel,^{31,32} E. Mueller,³³ S. Nadathur,¹⁶ G. Niz,^{34,35} W. J. Percival,^{36,37,38} F. Prada,³⁹ I. Pérez-Ràfols,⁴⁰ A. J. Ross,^{19,21,41} G. Rossi,⁴² E. Sanchez,⁴³ D. Schlegel,⁵ M. Schubnell,^{22,23} H. Seo,⁴⁴ J. Silber,⁵ Z. Slepian,^{4,5} D. Sprayberry,²⁵ G. Tarlé,²³ B. A. Weaver,²⁵ and H. Zou⁴⁵

¹*Max-Planck-Institut für Extraterrestrische Physik,*

Postfach 1312, Giessenbachstrasse, 85748 Garching, Germany

²*Institute of Astronomy, University of Cambridge, Madingley Rd, Cambridge CB3 0HA, United Kingdom*

³*Kavli Institute for Cosmology Cambridge, Madingley Road, Cambridge CB3 0HA, United Kingdom*

⁴*Department of Astronomy, University of Florida, Gainesville, Florida 32611, USA*

⁵*Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, California 94720, USA*

⁶*IRFU, CEA, Université Paris-Saclay, F-91191 Gif-sur-Yvette, France*

⁷*Department of Physics, Boston University,*

590 Commonwealth Avenue, Boston, Massachusetts 02215 USA

⁸*Dipartimento di Fisica “Aldo Pontremoli,” Università degli Studi di Milano, Via Celoria 16, I-20133 Milano, Italy*

⁹*INAF-Osservatorio Astronomico di Brera, Via Brera 28, 20122 Milano, Italy*

¹⁰*Department of Physics and Astronomy, University College London, Gower Street, London WC1E 6BT, United Kingdom*

¹¹*Instituto de Física, Universidad Nacional Autónoma de México,*

Circuito de la Investigación Científica, Ciudad Universitaria, Ciudad de México Código Postal 04510, Mexico

¹²*University of California, Berkeley, 110 Sproul Hall No. 5800 Berkeley, California 94720, USA*

¹³*Departamento de Física, Universidad de los Andes,*

Carrera 1 No. 18A-10, Edificio Ip, CP 111711, Bogotá, Colombia

¹⁴*Observatorio Astronómico, Universidad de los Andes,*

Carrera 1 No. 18A-10, Edificio H, CP 111711 Bogotá, Colombia

¹⁵*Institut d’Estudis Espacials de Catalunya (IEEC),*

c/ Esteve Terradas 1, Edifici RDIT, Campus PMT-UPC, 08860 Castelldefels, Spain

¹⁶*Institute of Cosmology and Gravitation, University of Portsmouth,*

Dennis Sciama Building, Portsmouth, PO1 3FX, United Kingdom

¹⁷*Institute of Space Sciences, ICE-CSIC, Campus UAB,*

Carrer de Can Magrans s/n, 08913 Bellaterra, Barcelona, Spain

¹⁸*Fermi National Accelerator Laboratory, PO Box 500, Batavia, Illinois 60510, USA*

¹⁹*Center for Cosmology and AstroParticle Physics, The Ohio State University,*

191 West Woodruff Avenue, Columbus, Ohio 43210, USA

²⁰*Department of Physics, The Ohio State University,*

191 West Woodruff Avenue, Columbus, Ohio 43210, USA

²¹*The Ohio State University, Columbus, Ohio 43210, USA*

²²*Department of Physics, University of Michigan, 450 Church Street, Ann Arbor, Michigan 48109, USA*

²³*University of Michigan, 500 S. State Street, Ann Arbor, Michigan 48109, USA*

²⁴*Department of Physics, The University of Texas at Dallas,*

800 West Campbell Road, Richardson, Texas 75080, USA

²⁵*NSF NOIRLab, 950 North Cherry Avenue, Tucson, Arizona 85719, USA*

²⁶*Department of Physics, Southern Methodist University, 3215 Daniel Avenue, Dallas, Texas 75275, USA*

²⁷*Department of Physics and Astronomy, University of California, Irvine 92697, USA*

²⁸*Center for Astrophysics | Harvard and Smithsonian,*

60 Garden Street, Cambridge, Massachusetts 02138, USA

²⁹*Sorbonne Université, CNRS/IN2P3, Laboratoire de Physique Nucléaire et de Hautes Energies (LPNHE), FR-75005 Paris, France*

³⁰*Departament de Física, Serra Hünter, Universitat Autònoma de Barcelona,*

08193 Bellaterra (Barcelona), Spain

³¹*Institut de Física d’Altes Energies (IFAE), The Barcelona Institute of Science and Technology, Edifici Cn, Campus UAB, 08193, Bellaterra (Barcelona), Spain*

³²*Institució Catalana de Recerca i Estudis Avançats, Passeig de Lluís Companys, 23, 08010 Barcelona, Spain*

³³*Department of Physics and Astronomy, University of Sussex, Brighton BN1 9QH, United Kingdom*

³⁴*Departamento de Física, DCI-Campus León, Universidad de Guanajuato, Loma del Bosque 103, León, Guanajuato C. P. 37150, Mexico*

³⁵*Instituto Avanzado de Cosmología A. C.,*

San Marcos 11—Atenas 202, Magdalena Contreras, Ciudad de México C. P. 10720, Mexico

³⁶*Department of Physics and Astronomy, University of Waterloo, 200 University Ave W, Waterloo, Ontario N2L 3G1, Canada*

³⁷*Perimeter Institute for Theoretical Physics, 31 Caroline Street North, Waterloo, Ontario N2L 2Y5, Canada*

³⁸*Waterloo Centre for Astrophysics, University of Waterloo, 200 University Ave W, Waterloo, Ontario N2L 3G1, Canada*

³⁹*Instituto de Astrofísica de Andalucía (CSIC), Glorieta de la Astronomía, s/n, E-18008 Granada, Spain*

⁴⁰*Departament de Física, EEBE, Universitat Politècnica de Catalunya, c/Eduard Maristany 10, 08930 Barcelona, Spain*

⁴¹*Department of Astronomy, The Ohio State University, 4055 McPherson Laboratory, 140 W 18th Avenue, Columbus, Ohio 43210, USA*

⁴²*Department of Physics and Astronomy, Sejong University, 209 Neungdong-ro, Gwangjin-gu, Seoul 05006, Republic of Korea*

⁴³*CIEMAT, Avenida Complutense 40, E-28040 Madrid, Spain*

⁴⁴*Department of Physics and Astronomy, Ohio University, 139 University Terrace, Athens, Ohio 45701, USA*

⁴⁵*National Astronomical Observatories, Chinese Academy of Sciences, A20 Datun Road, Chaoyang District, Beijing 100101, People's Republic of China*



(Received 16 August 2025; accepted 6 November 2025; published 15 December 2025)

We present a measurement of the non-Gaussian four-point correlation function (4PCF) from the DESI DR1 luminous red galaxy (LRG) sample. For the gravitationally induced parity-even 4PCF, we detect a signal with a significance of 14.7σ using our fiducial setup. We assess the robustness of this detection through a series of validation tests, including auto and cross-correlation analyses, sky partitioning across multiple patch combinations, and variations in radial scale cuts. Due to the low completeness of the sample, we find that differences in fiber assignment implementation schemes can significantly impact estimation of the covariance and introduce biases in the data vector. After correcting for these effects, all tests yield consistent results. This is one of the first measurements of the connected 4PCF on the DESI LRG sample; the good agreement between the simulation and the data implies that the amplitude of the density fluctuation inferred from the connected 4PCF is consistent with the Planck Λ CDM cosmology. The methodology and diagnostic framework established in this work provide a foundation for interpreting parity-odd 4PCF.

DOI: [10.1103/1wy-758f](https://doi.org/10.1103/1wy-758f)

I. INTRODUCTION

Non-Gaussianity on cosmological scales can arise from late-time nonlinear gravitational evolution or from deviations in the initial conditions from the standard single-field, slow-roll inflation paradigm [1–3]. While two-point statistics can fully capture all information in a Gaussian

random field, higher-order statistics offer an efficient framework for capturing this non-Gaussian information.

Beyond the standard two-point statistics, a variety of tools have been developed, such as N -point correlation functions in position space, and their Fourier-space counterparts, the polyspectra. The lowest-order extension beyond two-point statistics is the three-point statistics. In recent years, significant developments have been made for both the three-point correlation function (3PCF) and the bispectrum [4–10]. In addition to N -point functions, alternative statistics have been explored to extract non-Gaussian features, such as galaxy skew spectra [11–14], density split statistics [15], the wavelet scattering transform [16–18], and marked two-point statistics [19,20].

*Contact author: jiamin.hou@ast.cam.ac.uk

Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI. Open access publication funded by the Max Planck Society.

While a range of statistics beyond the standard two-point function has been developed, most either focus on three-point statistics or extract additional non-Gaussian information through nonlinear transformations. As a result, they do not explicitly encode correlations among more than three fields. A key advantage of galaxy spectroscopic surveys is their sensitivity to the full three-dimensional galaxy spatial positions, making them naturally suited to probe symmetries such as parity, defined as spatial inversion. The sensitivity to the parity transformation can be explored using galaxy four-point functions [21–23] or various compressed forms of four-point statistics [24,25].

The galaxy four-point function can be decomposed into parity-even and parity-odd components. The even component itself contains both a disconnected part, composed of products of two-point functions; and a connected part arising from nonlinear gravitational evolution. The latter also encodes rich cosmological information in tightening cosmological parameters and understand primordial non-Gaussianity [26].

Previous work [27] quantified the detection significance for the gravitationally induced connected even 4PCF using the galaxy sample from Baryon Oscillation Spectroscopic Survey (BOSS [28]). In this paper, we study the galaxy sample from the Dark Energy Spectroscopic Instrument (DESI) [29–36], which currently provides the largest spectroscopic galaxy sample to date. In particular, we focus on the luminous red galaxy (LRG) sample, enabling a direct comparison of information gain relative to BOSS. Moreover, this work is crucial for understanding systematics in the 4PCF and for interpreting the significance of parity-odd measurements.

The paper is structured as follows. In Sec. II we provide an overview of the DESI DR1, the LRG sample, the simulations, and the evaluation of systematics, particularly those arising from sample incompleteness in the first-year data of the five-year survey program. In Sec. IV, we provide a review of the 4PCF estimator and the measurement on the real data. In Sec. V we discuss the methodology used in this paper in quantifying the detection significance. In Sec. VI we discuss the detection significance for the DR1 result. In Sec. VII we discuss various points regarding the scale dependence, usage of different covariance, impact of incompleteness. Finally, we conclude in Sec. VIII.

Our Fourier transform convention is $\tilde{f}(\mathbf{k}) = \int d^3x e^{-i\mathbf{k}\cdot\mathbf{x}} f(\mathbf{x})$ and $f(\mathbf{x}) = \int_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{x}} \tilde{f}(\mathbf{k})$, where $\int_{\mathbf{k}} \equiv \int d^3\mathbf{k}/(2\pi)^3$.

II. AN OVERVIEW OF DESI DR1

DESI is a ground-based, fiber-fed spectroscopic survey located at Kitt Peak National Observatory in Arizona [37,38]. Over its five-year duration, DESI will map 14,200 square degrees of the sky, aiming to collect spectra for approximately 40 million galaxies and quasars across

the redshift range $0.1 < z < 4.2$ [39]. This includes the Bright Galaxy Survey (BGS) targeting galaxies at $0.1 < z < 0.4$ [40]. During dark time, DESI observes luminous red galaxies (LRGs) in the range $0.4 < z < 1.1$ [41], emission-line galaxies (ELGs) at $0.8 < z < 1.6$ [42], and quasars (QSOs) spanning $0.8 < z < 2.1$ [43]. DESI uses robotic positioners to place optical fibers within the 7-square-degree field of view of its focal plane [44]. Each set of targets is assigned to a set of 5,000 fibers. The year-one (Y1) data were obtained during the first 13 months of main survey operation. Over 4.7 million objects covering a sky area of 7,500 square degrees are included in the clustering analyses [45]. For this work, we focus on a sample of 2,138,627 LRGs, which we briefly describe in the following section.

A. DESI DR1 LRG sample

All four classes of DESI targets, including the LRGs, were selected based on photometry from Data Release 9 (DR9) of the Legacy Survey (LS) imaging [46]. The Legacy Survey combines photometric data from the Beijing-Arizona Sky Survey (BASS [47]) and the Mayall z-band Legacy Survey (MzLS) for target selection in the Northern Galactic Cap (NGC). For the Southern Galactic Cap (SGC), target selection is based on imaging from the Dark Energy Camera (DECam [48]) and the Dark Energy Survey (DES [49]). Infrared photometry in the W1 and W2 bands from the WISE satellite [50,51] is used across the entire sky.

The DESI LRG sample is selected using g,r,z , and W1 flux measurements [41]. A series of photometric cuts is applied to the sample to reduce stellar contamination, select the most luminous galaxies across redshifts, maintain an approximately constant comoving number density in the range $0.4 < z < 0.8$, and ensure a high spectroscopic-redshift success rate. The selection criteria are independently optimized in the BASS/MzLS and DECam regions to yield a sample of passively evolving galaxies with an approximately constant number density of $5 \times 10^{-4} [\text{Mpc}^{-1}h]^3$ in the redshift range $0.4 < z < 0.8$. Beyond this range, the number density drops below $1 \times 10^{-4} [\text{Mpc}^{-1}h]^3$ by $z = 1.1$, primarily due to a threshold on the z -band fiber magnitude [41].

To obtain galaxy spectra, each target is assigned a fiber that directs its emitted light into a DESI spectrograph. A FIBERASSIGN code [52] is used to allocate fibers to designated targets based on the DESI targeting algorithm. This algorithm assigns fibers according to the merged-target-ledger (MTL) file, which records sample information such as sky coordinates, target IDs, and priorities. Among these, the priority label is particularly relevant; in this process, the LRG sample has intermediate priority, lower than that of the QSO sample but higher than that of the ELG sample.

Fiber assignment in DESI has two major impacts. First, each fiber can access only a limited patrol area; although

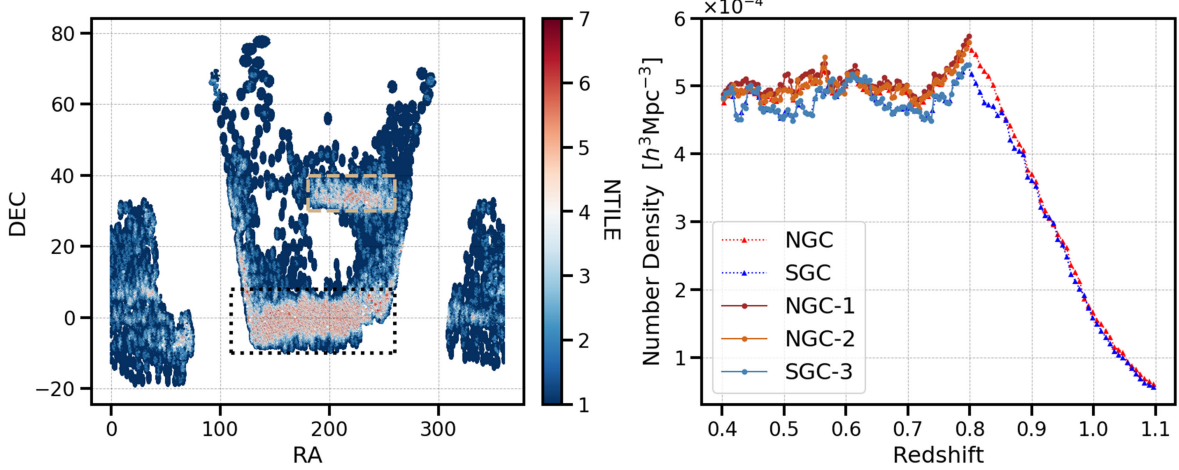


FIG. 1. *Left*: footprint of DESI-DR1 LRG for both the NGC and SGC. The color map shows the number of targets sharing the same unique identifier, which combines the tile ID and fiber location. This quantity is closely correlated with the sample completeness; higher values of NTILE indicate higher completeness. For this paper, we specifically studied two regions, the NGC₁ (dotted black box) and NGC₂ (dashed gray box). *Right*: number density as a function of redshift $n(z)$ for the NGC (red) and SGC (blue), including different patches with various cuts in both angular and radial directions. After applying completeness weights, the number densities as a function of redshift are consistent across various cuts. This demonstrates the effectiveness of the completeness weights at the density level in the radial direction.

adjacent patrol regions overlap and about 15% of the focal plane is reachable by two fibers, in dense regions the number of targets often exceeds the number of available fibers. Second, due to the physical size of the positioners, fibers cannot be placed too close together, a constraint commonly referred to as “fiber collisions,” which limits the ability to observe all targets, especially for closely spaced galaxy pairs in a single pass. Each pointing of the telescope brings approximately 25,000 targets into the focal plane, of which only 5,000 can be accommodated. As the survey progresses, each domain is visited multiple times. Early in the survey there is significant incompleteness. In the full five-year survey, these limitations will be largely mitigated by repeated observations.

In addition to fiber assignment incompleteness, other systematics such as density variation due to imaging

systematics, redshift failure also need to taken into account. These effects are currently accounted for by applying weights to either data or randoms at catalog level.

B. Choice of the LRG sample

A key consideration for the DR1 LRG sample is its completeness, defined as the ratio of observed galaxies to the targeted galaxies for a given region. Figure 1 shows the DESI DR1 LRG footprint, where the colormap indicates the number of targets sharing the same identifier (TILELOCID, which combines the tile ID and fiber location). The quantity NTILE denotes the number of tiles on which a target was observed, serving as an indicator of sample completeness; higher values correspond to more complete regions [53].

TABLE I. Statistics for the DR1 LRG sample, including the number of galaxies N_{gal} ; both the number density \bar{n}_g and fiducial volume V_{fid} are obtained from fitting the analytic covariance matrix. The effective volume V_{eff} is computed from the trace of the product of the inverse of the full analytic covariance and the covariance for the μ th patch (cf. Sec. V B). The difference between V_{fid} and V_{eff} arises from variations in number density.

Region	Selection	N_{gal}	$\bar{n}_g [h^{-1} \text{Mpc}]^{-3}$	$V_{\text{fid}} [h^{-1} \text{Gpc}]^3$	$V_{\text{eff}} [h^{-1} \text{Gpc}]^3$
Full-NGC	$0.4 < z < 1.1$	1,476,135	2.3×10^{-4}	2.54	2.54
NGC-1	$110 < \text{RA} < 260,$ $-10 < \text{DEC} < 8,$ $0.4 < z < 0.8$	434,054	3.5×10^{-4}	0.54	0.87
NGC-2	$180 < \text{RA} < 260,$ $30 < \text{DEC} < 40,$ $0.4 < z < 0.8$	119,034	2.9×10^{-4}	0.13	0.17
Full-SGC	$0.4 < z < 1.1$	662,492	1.5×10^{-4}	1.43	0.79
SGC-3	$0.4 < z < 0.8$	398,089	1.7×10^{-4}	0.72	0.48

We use the default clustering catalog with a redshift range $0.4 < z < 1.1$. Additionally, given the dependence on completeness and redshift evolution, we further split the LRG sample into three subregions. For the NGC, we select areas with higher completeness, defining two subregions: NGC-1 with angular cuts in both right ascension (RA) and declination (DEC); $110 < \text{RA} < 260$, $-10 < \text{DEC} < 8$, and a radial cut $0.4 < z < 0.8$; and NGC-2 with $180 < \text{RA} < 260$, $30 < \text{DEC} < 40$, and the same redshift range. For the SGC, we define a single subregion, SGC-3, using only the radial cut $0.4 < z < 0.8$, since there is no large variation in completeness (unlike NGC) and also to retain statistical constraining power. The geometric selection, the number of galaxies, and the corresponding volumes are listed in Table I.

III. SIMULATIONS

Two types of simulations are used for DESI DR1 for estimating the covariance matrix and studying systematics.

To estimate the covariance matrix, we use a suite of fast, approximate, but extensive EZmock [54,55] provided by the DESI collaboration. For the EZmock, the dark matter density field is based on the Zel'dovich approximation. The resulting density field is populated with galaxies using an effective-bias model to account for nonlinear evolution and galaxy bias. There are 1000 realizations for the box of size $L_{\text{box}} = 6h^{-1}$ Gpc, which are postprocessed to match the DESI DR1 geometry.

In addition, DESI also has a suite of high-resolution N -body simulations; AbacusSummit [56,57]. These have $L_{\text{box}} = 2h^{-1}$ Gpc box and use a fiducial cosmology $\Omega_c h^2 = 0.1200$, $\Omega_b h^2 = 0.02237$, $\sigma_8 = 0.811355$, $n_s = 0.9649$, $h = 0.6736$, $w_0 = -1$, $w_a = 0$. They use a halo occupation distribution model for the galaxy-halo connection, calibrated on the DESI Early Data Release (EDR [58]) described in [59] for the LRG sample. Since the mock box size is insufficient to encompass the full DESI DR1 LRG volume, the mocks are tiled to fill the survey footprint. As a result, they overestimate the cosmic variance and are therefore unsuitable for covariance estimation. Nevertheless, they remain valuable for studying systematics.

Both EZmock and Abacus simulations are postprocessed with three variations for the fiber assignment implementation. The ‘‘complete’’ mocks represent surveys capturing all galaxies that could have been targeted [60]. The ‘‘alternative MTL’’ (altMTL) mocks represent the most realistic fiber assignment; they were generated using the DESI FIBERASSIGN code [52] applied to tiles in the same ordering and cadence, with a feedback loop to the target list, mirroring the procedure used for the observed data. The ‘‘fast fiber assignment’’ (FFA) mocks emulate the fiber assignment process by repeatedly sampling from the average targeting probability of galaxies. This probability

is learned from the data as a function of the number of overlapping tiles and the local angular clustering.

Although the altMTL fiber-assignment approach is more realistic, it is substantially more expensive to run and imposes stricter simulation requirements; it requires the full dark-time sample, whereas the FFA algorithm does not and is much faster. Consequently, we have only 25 altMTL realizations applied to Abacus, an insufficient number to estimate the inverse of the covariance reliably. We therefore rely on FFA-based EZmock realizations for covariance estimation.

We will discuss in more detail the fiber assignment implementations and their impact on statistical fluctuations of the 4PCF (see Sec. VII D), which is crucial for interpreting the detection significance.

IV. FOUR-POINT CORRELATION FUNCTION AND ITS ESTIMATOR

Assuming homogeneity, the four-point correlation can be expressed as $\zeta(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3)$ by taking the origin at one galaxy. It can be decomposed into a parity-even part ζ_+ and a parity-odd part ζ_- ,

$$\zeta(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3) = \zeta_+(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3) + i\zeta_-(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3), \quad (1)$$

where ζ_+ and ζ_- are real. The parity-odd part in general vanishes for the standard model of cosmology (further discussion see [21–23]). The parity-even part includes both a disconnected contribution, given by products of two-point functions ξ and a connected contribution $\zeta_+^{(c)}$. If initially the distribution is given by a Gaussian random field, there is only a disconnected piece. Nonlinear gravitation evolution generates the connected piece:

$$\zeta_+(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3) = \zeta_+^{(c)}(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3) + [\xi(\mathbf{r}_1)\xi(\mathbf{r}_2 - \mathbf{r}_3) + 2 \text{cyc.}], \quad (2)$$

where we cyclically permute the contribution from the 2PCF.

A. Isotropic basis functions

A brute force measurement of the 4PCF scales as $\mathcal{O}(N_g^4)$, with N_g being the number of galaxies and is thus computationally challenging. To accelerate the process, we expand the 4PCF in a basis of isotropic functions $\mathcal{P}_{\ell_1 \ell_2 \ell_3}$ [61], which capture its angular behavior about one galaxy (the ‘‘primary’’), multiplied by functions that capture its dependence on tetrahedron side lengths r_1, r_2, r_3 from that primary. The radial coefficient functions are projections of the full distribution on the isotropic functions:

$$\begin{aligned} & \zeta_{\ell_1 \ell_2 \ell_3}(r_1, r_2, r_3) \\ &= \int d\hat{r}_1 d\hat{r}_2 d\hat{r}_3 \zeta(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3) \mathcal{P}_{\ell_1 \ell_2 \ell_3}^*(\hat{r}_1, \hat{r}_2, \hat{r}_3), \quad (3) \end{aligned}$$

where ℓ_i , for $i = 1, 2, 3$ are the ‘‘angular momenta’’ associated with the three direction vectors \mathbf{r}_i , and star denotes a complex conjugate. We define the angular differential $d\hat{r}$ by $\int d\hat{r} = 4\pi$.

To avoid an over-complete basis, we expand in $\mathcal{P}_{\ell_1 \ell_2 \ell_3}^*(\hat{r}_1, \hat{r}_2, \hat{r}_3)$ where $r_1 < r_2 < r_3$. The isotropic functions $\mathcal{P}_{\ell_1 \ell_2 \ell_3}(\hat{r}_1, \hat{r}_2, \hat{r}_3)$ of three arguments is given by

$$\begin{aligned} \mathcal{P}_{\ell_1 \ell_2 \ell_3}(\hat{r}_1, \hat{r}_2, \hat{r}_3) &= \sum_{m_1 m_2 m_3} (-1)^{\ell_1 + \ell_2 + \ell_3} \begin{pmatrix} \ell_1 & \ell_2 & \ell_3 \\ m_1 & m_2 & m_3 \end{pmatrix} \\ &\times Y_{\ell_1 m_1}(\hat{r}_1) Y_{\ell_2 m_2}(\hat{r}_2) Y_{\ell_3 m_3}(\hat{r}_3), \quad (4) \end{aligned}$$

where the $Y_{\ell_i m_i}(\hat{r}_i)$ are spherical harmonics and the m_i are the z -components of their angular momenta ℓ_i . The product of the three spherical harmonics is weighted by the Wigner 3- j symbol. The isotropic basis is fully separable in the \hat{r}_i , thus reducing the formal scaling of the 4PCF estimator to pairwise operations [62].

Applying a parity transformation $\mathbb{P}: (x, y, z) \rightarrow (-x, -y, -z)$ to a spherical harmonic yields $\mathbb{P}[Y_{\ell m}(\hat{r})] = (-1)^\ell Y_{\ell m}(\hat{r})$. Extending this relation to the isotropic basis function in Eq. (4), we find

$$\mathbb{P}[\mathcal{P}_{\ell_1 \ell_2 \ell_3}(\hat{r}_1, \hat{r}_2, \hat{r}_3)] = (-1)^{\ell_1 + \ell_2 + \ell_3} \mathcal{P}_{\ell_1 \ell_2 \ell_3}(\hat{r}_1, \hat{r}_2, \hat{r}_3). \quad (5)$$

With this basis, the parity behavior described in Eq. (1) can be isolated by examining the sum $\sum_i \ell_i$; the basis function is parity-even if the sum is even, and parity-odd if the sum is odd.

In practice, the connected 4PCF multipole estimator $\hat{\zeta}_{\ell_1 \ell_2 \ell_3}^{(c)}$ is obtained by subtracting from the full 4PCF multipole estimator $\hat{\zeta}_{\ell_1 \ell_2 \ell_3}$ the disconnected 4PCF estimator $\hat{\zeta}_{\ell_1 \ell_2 \ell_3}^{(dc)}$ as follows:

$$\hat{\zeta}_{\ell_1 \ell_2 \ell_3}^{(c)}(r_1, r_2, r_3) = \hat{\zeta}_{\ell_1 \ell_2 \ell_3}(r_1, r_2, r_3) - \hat{\zeta}_{\ell_1 \ell_2 \ell_3}^{(dc)}(r_1, r_2, r_3), \quad (6)$$

where the disconnected multipoles arise from the product of two two-point correlation functions (2PCFs, cf, Eq. (2); one depending on a single argument, and the other on a composite argument. After locally projecting both 2PCFs onto the spherical harmonic basis, the resulting estimators are summed, yielding global 2PCF multipoles with single angular momentum indices $\xi_{\ell m}$, and double indices $\xi_{\ell m, \ell' m'}$. The details of measuring the connected 4PCF and definition of the 2PCF multipoles $\xi_{\ell m}$ and $\xi_{\ell m, \ell' m'}$ are given in [27].

Hereafter, we will use ζ to denote the connected parity-even 4PCF multipoles and drop all the superscripts and subscripts for brevity.

B. Survey geometry correction

The survey geometry can affect galaxy clustering measurements. To account for this, a random catalog with identical survey geometry is used as a correction. For a discrete galaxy sample, the galaxy number density can be expressed as

$$n_{g,r}(\mathbf{x}) = \sum_{i=1}^{N_g} w_i^{g,r} \delta_D^{[3]}(\mathbf{x} - \mathbf{x}_i), \quad (7)$$

with $\{g, r\}$ denoting the galaxy or random catalog, w_i is the weight for each galaxy, $\delta_D^{[3]}(\mathbf{x} - \mathbf{x}_i)$ is the three-dimensional Dirac delta function. The number density can be converted to number counts $D(\mathbf{x}) = n_g(\mathbf{x})\Delta V$ and $R(\mathbf{x}) = n_r(\mathbf{x})\Delta V$, where ΔV is a volume element, which cancels out when taking the ratio. By forming the difference between the data and the random catalog, we can construct $N(\mathbf{x}) = D(\mathbf{x}) - R(\mathbf{x})$. The extended Landy-Szalay estimator for the 4PCF thus reads [63,64],

$$\begin{aligned} \hat{\zeta}(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3) &= \frac{\int d^3 \mathbf{x} N(\mathbf{x}) N(\mathbf{x} + \mathbf{r}_1) N(\mathbf{x} + \mathbf{r}_2) N(\mathbf{x} + \mathbf{r}_3)}{\int d^3 \mathbf{x} R(\mathbf{x}) R(\mathbf{x} + \mathbf{r}_1) R(\mathbf{x} + \mathbf{r}_2) R(\mathbf{x} + \mathbf{r}_3)} \\ &= \frac{\mathcal{N}(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3)}{\mathcal{R}(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3)}. \quad (8) \end{aligned}$$

Since we are interested in the coefficients of the 4PCF in the isotropic basis, we expand both sides of the equation in this basis. In particular, both the numerator and denominator on the right-hand side are also expanded in the isotropic basis. By doing so, Eq. (8) becomes

$$\begin{aligned} \hat{\zeta}_{\ell_1 \ell_2 \ell_3}(r_1, r_2, r_3) &= \sum_{\ell'_1 \ell'_2 \ell'_3} [\mathbf{M}^{-1}]_{\ell_1 \ell_2 \ell_3, \ell'_1 \ell'_2 \ell'_3}(r_1, r_2, r_3) \\ &\times \frac{\mathcal{N}_{\ell'_1 \ell'_2 \ell'_3}(r_1, r_2, r_3)}{\mathcal{R}_{000}(r_1, r_2, r_3)}. \quad (9) \end{aligned}$$

The mode decoupling matrix \mathbf{M} reads [63]

$$\begin{aligned} & \mathbf{M}_{\ell_1 \ell_2 \ell_3, \ell'_1 \ell'_2 \ell'_3}(r_1, r_2, r_3) \\ &= (4\pi)^{-3/2} (-1)^{\ell'_1 + \ell'_2 + \ell'_3} \sum_{L_1 L_2 L_3} \frac{\mathcal{R}_{L_1 L_2 L_3}(r_1, r_2, r_3)}{\mathcal{R}_{000}(r_1, r_2, r_3)} \\ &\times \prod_{i=1}^3 D_{\ell_i L_i \ell'_i} C_{000}^{\ell'_i L_i \ell_i} \begin{Bmatrix} \ell_1 & L_1 & \ell'_1 \\ \ell_2 & L_2 & \ell'_2 \\ \ell_3 & L_3 & \ell'_3 \end{Bmatrix}. \quad (10) \end{aligned}$$

Here we use the Wigner 9- j symbol and define the coefficients

$$\mathcal{D}_{\ell_1 \ell_2 \ell_3} = \sqrt{(2\ell_1 + 1)(2\ell_2 + 1)(2\ell_3 + 1)} \quad (11)$$

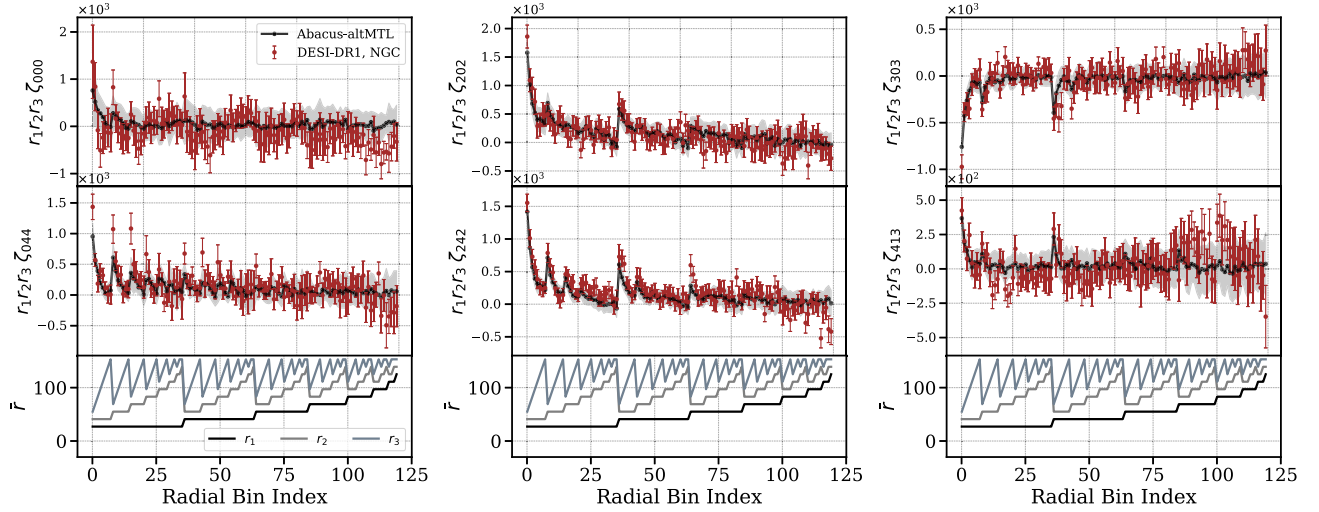


FIG. 2. Measurement of the connected 4PCFs using the DESI-DR1 LRG sample in NGC. The top two panels show a subset of the angular channels (ℓ_1, ℓ_2, ℓ_3) for the coefficients of the connected 4PCF, weighted by $r_1 r_2 r_3$. Points with error bars (red) are the measurement from the data; the black curves with shaded gray regions indicate the 1σ standard deviation from the Abacus simulations with altMTL fiber assignment scheme. The bottom panel shows the arrangement of the three radial bins.

and

$$C_{000}^{\ell_1 \ell_2 \ell_3} \equiv \begin{pmatrix} \ell_1 & \ell_2 & \ell_3 \\ 0 & 0 & 0 \end{pmatrix}. \quad (12)$$

C. Measurement of the connected four-point correlation functions

We apply the 4PCF estimator described in this section to the DESI DR1 LRG sample [65]. For each galaxy, we apply weights following [53]:

$$w_{\text{tot}} = w_{\text{comp}} w_{\text{sys}} w_{\text{zfail}} / \langle w_{\text{comp}} \rangle_{N_{\text{tile}}}, \quad (13)$$

where w_{comp} is the completeness weight, defined as the number of targets that compete for a given fiber. w_{sys} accounts for target density fluctuations due to imaging conditions, including stellar density, HI column density, z -band galaxy depth, r -band Point Spread Function (PSF) size, and W1 PSF depth, and is derived using a linear regression method [53]. w_{zfail} accounts for the redshift success rate and reduces spurious fluctuations in tracer density that are correlated with observational conditions in

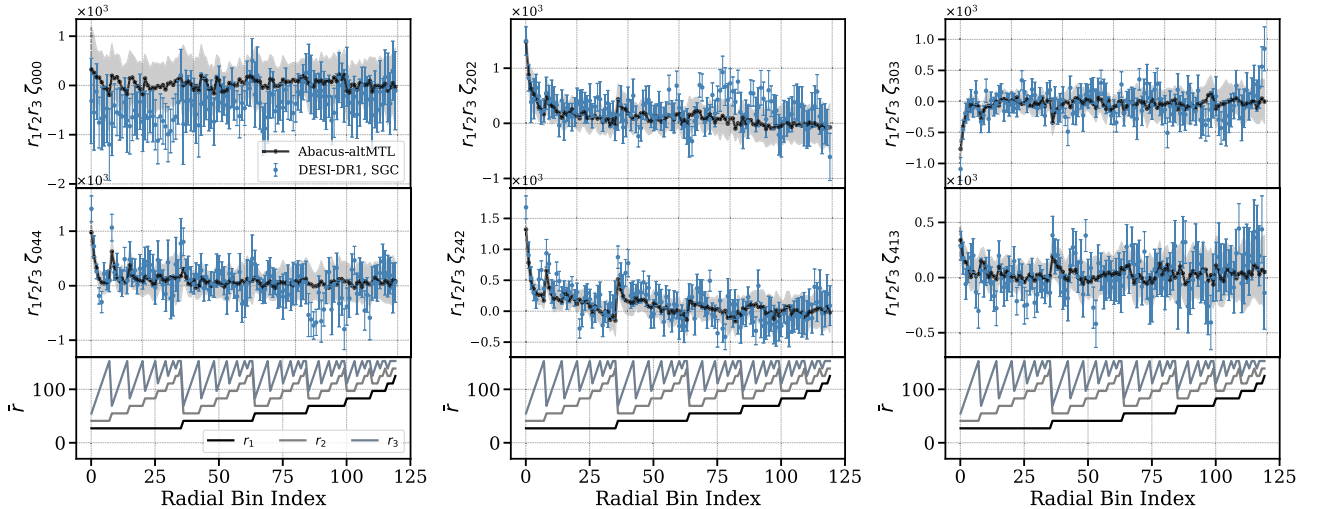


FIG. 3. Similar to Fig. 2, this figure shows the measurement of the connected 4PCFs using the DESI-DR1 LRG sample in SGC. The top two panels show a subset of the angular channels (ℓ_1, ℓ_2, ℓ_3) for the coefficients of the connected 4PCF, weighted by $r_1 r_2 r_3$. Points with error bars (blue) are the measurement from the data; the black curves with shaded gray regions indicate the 1σ standard deviation from the Abacus simulations with altMTL fiber assignment scheme. The bottom panel shows the arrangement of the three radial bins.

DESI spectra. $\langle w_{\text{comp}} \rangle_{N_{\text{tile}}}$ is the averaged completeness for a given number of overlapping tiles N_{tile} .

Additionally, we apply Feldman–Kaiser–Peacock (FKP) weights [66], which balance the trade-off between cosmic variance and the Poissonian shot noise:

$$w_{\text{FKP}} = (1 + n_0 \langle C_{\text{assign}} \rangle_{N_{\text{tile}}} P_0)^{-1}, \quad (14)$$

where we take $n_0 = 4 \times 10^{-4} (\text{Mpc}^{-1} h)^3$ and $P_0 = 10,000 (h^{-1} \text{Mpc})^3$, and $\langle C_{\text{assign}} \rangle_{N_{\text{tile}}}$ is the expectation value of assignment completeness for a given number of overlapping tiles.

Figures 2 and 3 show the measurement of the connected 4PCF using DESI-DR1 LRG sample in the NGC and SGC, respectively. We use the radial scales between $20h^{-1} \text{Mpc} < r < 160h^{-1} \text{Mpc}$, divided into ten bins. The top two panels show, for selected choices of (ℓ_1, ℓ_2, ℓ_3) the coefficients of the connected 4PCF, weighted by the product of $r_1 r_2 r_3$. Points with error bars are the measurement from the data; the black curves with shaded gray regions indicate the 1σ standard deviation from the *Abacus* simulations with altMTL fiber assignment scheme. The bottom panel shows the arrangement of the three radial bins. The radial bin index can be expressed as $i = i_1(n - i_2 + 1) + (i_2 - 1)$, with i_1 and i_2 being the first and second bin index, and n being the total number of radial bins.

We select these angular channels because they exhibit prominent features, notably a characteristic “sawtooth” pattern. We note that these features are closely related to the specific ordering of the three sides r_1, r_2, r_3 ; the signal-to-noise ratio is highest at small scales for the approximately equilateral configurations, *i.e.*, when the three bins are closest in value.

V. ANALYSIS METHODOLOGY

To quantify the gravity-induced connected 4PCF, we perform the analysis using the null hypothesis. With the assumption that the gravity-induced connected 4PCF is zero we assess the extent to which this hypothesis can be rejected. To do so, we compress the 4PCF data vector into a single scalar quantity, χ^2 , which involves the data vector, a model, and a covariance matrix.

The key challenge is accurate estimation of the covariance matrix. A common practice is to use simulations to estimate the covariance matrix. The simulations used for DESI DR1 are, however, not tuned to match higher-order clustering, and the approximate method may underestimate nonlinear effects from gravitational dynamics, leading to a data-simulation mismatch. Additionally, the sample covariance requires the number of simulations N_{sim} to be much larger than the length of the data vector. However, given that the length of the data vector is $\mathcal{O}(10^3)$, alternative covariance estimation methods are necessary.

In Sec. VC, we present the formalism for computing the analytic covariance matrix. The covariance matrix of the 4PCF includes both Gaussian and higher-order non-Gaussian contributions. Following [67], we keep only the leading Gaussian term, as that work showed that the Gaussian approximation to the covariance provides a reasonably good description even when nonlinearity is present at the level of the summary statistics. However, because the analytic approach does not account for survey systematics and geometry, we adopt a hybrid strategy; we use the analytic covariance to construct a basis in which to decompose the sample covariance estimated from simulations. In this way, the analytic covariance does not directly enter the significance quantification, while still enabling an efficient reduction in the dimensionality of the data vector.

However, there remains a potential concern that the covariance estimated from simulations may not fully capture the true fluctuations present in the data. To address the data-simulation mismatch in the covariance, we perform two tests; (1) autocorrelation (Sec. VA) and (2) cross-correlation (Sec. VB). The autocorrelation captures both the signal and statistical fluctuations, while the cross-correlation enables separation of these components. If the excess in χ^2 from the autocorrelation arises from an underestimated covariance in the simulation, the cross-correlation will yield a reduced significance. Comparing the two provides a cross-check on the robustness of the detection significance.

A. Autocorrelation

We define χ^2 , which quantifies the *overall* deviation from the null hypothesis for the parity-even signal and corresponds to the square of the detection significance (in units of σ) [68]:

$$\chi^2 = \hat{\zeta}^i \hat{C}_{ij}^{-1} \hat{\zeta}^j. \quad (15)$$

\hat{C} is an estimator of the covariance matrix, which can be derived analytically (see Sec. VC) or from simulations. $\hat{\zeta}$ is the measured 4PCF from the data, which is composed of a gravity-induced connected 4PCF signal ζ_g . The components of the data vector—the values of $\zeta_{(\ell_1, \ell_2, \ell_3)}(r_1, r_2, r_3)$ —are labeled ζ^i . Following [22,27], we set ζ_g to be zero for the null hypothesis test.

We assume that the observed connected 4PCF is composed of the following terms:

$$\hat{\zeta} = \zeta_g + \hat{\zeta}_s + \hat{e}, \quad (16)$$

where ζ_g represents the true gravitationally-induced signal, while $\hat{\zeta}_s$ accounts for systematic effects, including survey systematics and inaccuracies in modeling nonlinear gravitational effects. The third term \hat{e} corresponds to statistical fluctuations arising from cosmic variance.

Broadly, systematic effects can impact the signal in two ways. On the one hand, they may introduce a nonvanishing bias, on the other hand, they can have nonvanishing correlations; $\hat{\zeta}_s = \zeta_s + \hat{e}_s$, with ζ_s being a systematic bias that survives in the ensemble average, and the stochastic \hat{e}_s term introduces a correction in addition to the cosmic variance:

$$\langle \hat{\zeta}_s \rangle = \zeta_s \neq 0, \quad \langle \hat{e}_s^i \hat{e}_s^j \rangle \equiv \Delta C_{ss}^{ij} \neq 0 \quad (17)$$

The statistical fluctuation becomes zero under the ensemble average, and has nonvanishing correlation:

$$\langle \hat{e} \rangle = 0, \quad \langle \hat{e}^i \hat{e}^j \rangle \equiv C^{ij}, \quad (18)$$

here, C represents the intrinsic statistical fluctuation for the data C_{data} or the simulations C_{mock} . Moreover, the cross-correlation between the stochastic component of the systematics and the statistical fluctuation of the density field could also be present:

$$\langle \hat{e}_s^i \hat{e}^j \rangle \equiv \Delta C_s^{ij}. \quad (19)$$

In the presence of systematics, the inverse of the covariance receives an additional correction,

$$\begin{aligned} \hat{C}^{-1} &\rightarrow (C + \Delta C_s + \Delta C_{ss})^{-1} \\ &\approx C^{-1} - C^{-1}(\Delta C_s + \Delta C_{ss})C^{-1}. \end{aligned} \quad (20)$$

Given Eqs. (16)–(18), we find the expectation value of the product of the 4PCF observables is

$$\langle \hat{\zeta}^i \hat{\zeta}^j \rangle = \sum_{a,b \in \{g,s\}} \zeta_a^i \zeta_b^j + (C^{ij} + \Delta C_s^{ij} + \Delta C_{ss}^{ij}). \quad (21)$$

As a result, the expectation of the χ^2 defined in Eq. (15) under the null assumption, that there is no connected contribution to 4PCF, becomes

$$\begin{aligned} \langle \chi^2 \rangle &= \langle \hat{\zeta}^i (\hat{C}^{-1})_{ij} \hat{\zeta}^j \rangle \\ &= (\hat{C}^{-1})_{ij} (\zeta_g^i \zeta_g^j + 2\zeta_s^i \zeta_g^j + \zeta_s^i \zeta_s^j) + \text{Tr}[\hat{C}^{-1}C], \end{aligned} \quad (22)$$

where $(\hat{C}^{-1})_{ij} \zeta_s^i \zeta_g^j = (\hat{C}^{-1})_{ij} \zeta_g^i \zeta_s^j$, given that the covariance matrix is symmetric $C_{ij} = C_{ji}$.

The difference between the χ^2 from the data and the mocks is given by

$$\begin{aligned} \langle \chi_{\text{data}}^2 \rangle - \langle \chi_{\text{mock}}^2 \rangle &= \langle (\hat{\zeta}_{\text{data}} - \zeta_g)^i (\hat{C}^{-1})_{ij} (\hat{\zeta}_{\text{data}} - \zeta_g)^j \rangle \\ &\quad - \langle (\hat{\zeta}_{\text{mock}} - \zeta_g)^i (\hat{C}^{-1})_{ij} (\hat{\zeta}_{\text{mock}} - \zeta_g)^j \rangle \\ &\equiv \hat{C}_{ij}^{-1} (\Delta_{gg}^{ij} + 2\Delta_{gs}^{ij} + \Delta_{ss}^{ij}) \\ &\quad + \text{Tr}[\hat{C}^{-1}(C_{\text{data}} - C_{\text{mock}})], \end{aligned} \quad (23)$$

with the residual difference in cross-correlation products between data and mock catalogs Δ_{ab}^{ij} defined as

$$\Delta_{ab}^{ij} \equiv \zeta_a^{\text{data},i} \zeta_b^{\text{data},j} - \zeta_a^{\text{mock},i} \zeta_b^{\text{mock},j}, \quad (24)$$

where $a \in \{g, s\}$ and $b \in \{g, s\}$ represent gravity-induced “g” or systematic-induced “s.”

The expectation values of $\langle \chi_{\text{data}}^2 \rangle$ and $\langle \chi_{\text{mock}}^2 \rangle$ are equal only if the mocks perfectly reproduce nonlinear gravitational evolution, all systematics are fully corrected (i.e., $\Delta_{ab}^{ij} = 0$), and the covariance matrix is estimated without bias. In that case, the χ^2 statistic effectively quantifies the excess signal arising from the gravity-induced connected 4PCF. In practice, however, these conditions are not always satisfied, and residual systematics may be misinterpreted as gravity-induced nonlinearity, impacting both the signal and the covariance estimation.

For the DESI-Y1 LRG sample, one major concern is the fiber assignment implementation scheme. Overall, an incomplete sample has larger statistical fluctuations; in particular, we find that different fiber assignment implementations can lead to a difference in covariance of up to 20% to 30% (see Sec. VIII D).

For these reasons, in the autocorrelation tests we compare the results using four sets of simulations: (1) EZmock with FFA; (2) Abacus with altMTL (reference); (3) Abacus with FFA; (4) “complete” Abacus. These sets of tests allow us to isolate the effects due to inaccurate estimation of statistical fluctuations. Moreover, they allow us to understand the contribution due to incomplete fiber assignment implementation effects, and potential misinterpretation of gravity-induced non-Gaussianity.

B. Cross-correlation test

As an additional test of the detection significance, we follow [22,69], and divide the full survey into different smaller angular patches, with data vectors labeled $\zeta^{\mu,i}, \zeta^{\nu,j}, \dots$. The idea is that if there is an underlying signal, it should show correlation across different patches of the sky.

In particular, [69] proposed to further divide χ^2 in Eq. (15) into two terms; (1) χ_{\times}^2 , which captures spatially-correlated physics such as gravity-induced connected 4PCF (the signal in our context) and (2) χ_{Δ}^2 , which is sensitive to the mismatch in statistical fluctuations between the data and the mocks.

The signal-sensitive statistic χ_{\times}^2 is defined as

$$\chi_{\times}^2 = \frac{1}{\mathcal{N}_{\times}} \sum_{\mu < \nu} \hat{\zeta}_s^{\mu,i} (\hat{C}^{-1})_{ij} \hat{\zeta}_s^{\nu,j}, \quad (25)$$

with the normalization factor

$$\mathcal{N}_{\times} \equiv N_p(N_p - 1)/2, \quad (26)$$

with N_p being the number of patches, and the normalization given by the combinatorial factor.

Under the assumption that different patches are uncorrelated, the underlying density field and the systematics are also uncorrelated. The expectation value of the cross-correlation between the 4PCF of two different patches reads

$$\langle \hat{\zeta}_s^{\mu,i} \hat{\zeta}_s^{\nu,j} \rangle = \sum_{a,b \in \{g,s\}} \zeta_a^{\mu,i} \zeta_b^{\nu,j} + \hat{C}^{\mu,ij} \delta_{\mu\nu}^K. \quad (27)$$

Here we explicitly include the systematics term ζ_s , which can potentially generate a connected parity-even 4PCF at the signal level.

The expectation value of the signal-sensitive statistic $\langle \chi_{\times}^2 \rangle$ is

$$\begin{aligned} \langle \chi_{\times}^2 \rangle &= \frac{1}{\mathcal{N}_{\times}} \sum_{\mu < \nu} \langle \hat{\zeta}_s^{\mu,i} (\hat{C}^{-1})_{ij} \hat{\zeta}_s^{\nu,j} \rangle \\ &= (\hat{C}^{-1})_{ij} \left\{ \zeta_g^i \zeta_g^j + \frac{1}{\mathcal{N}_{\times}} \sum_{\mu < \nu} (\zeta_g^{\mu,i} \zeta_s^{\nu,j} + \zeta_s^{\mu,i} \zeta_g^{\nu,j} + \zeta_s^{\mu,i} \zeta_s^{\nu,j}) \right\}. \end{aligned} \quad (28)$$

Here we impose that the signal from gravitational evolution is identical across different patches, while allowing the systematics in the μ th and ν th patches to differ in general. Compared to the parity-odd case in [69], the expectation of the cross statistics for simulations does not center around zero, since the simulations do include gravitationally-induced non-Gaussian effects. Moreover, the additional systematics-induced term can potentially contaminate the cross-statistic. Notably, the cross-correlation between the signal and the systematics may lead to a negative correlation.

Comparing the difference between χ_{\times}^2 for data and mocks, we arrive at the following:

$$\begin{aligned} &\langle \chi_{\times,\text{data}}^2 \rangle - \langle \chi_{\times,\text{mock}}^2 \rangle \\ &= (\hat{C}^{-1})_{ij} \left\{ (\zeta_{g,\text{data}}^i \zeta_{g,\text{data}}^j - \zeta_{g,\text{mock}}^i \zeta_{g,\text{mock}}^j) \right. \\ &\quad + \frac{1}{\mathcal{N}_{\times}} \sum_{\mu < \nu} [(\zeta_{g,\text{data}}^{\mu,i} \zeta_{s,\text{data}}^{\nu,j} - \zeta_{g,\text{mock}}^{\mu,i} \zeta_{s,\text{mock}}^{\nu,j}) \\ &\quad + (\zeta_{s,\text{data}}^{\mu,i} \zeta_{g,\text{data}}^{\nu,j} - \zeta_{s,\text{mock}}^{\mu,i} \zeta_{g,\text{mock}}^{\nu,j}) \\ &\quad \left. + (\zeta_{s,\text{data}}^{\mu,i} \zeta_{s,\text{data}}^{\nu,j} - \zeta_{s,\text{mock}}^{\mu,i} \zeta_{s,\text{mock}}^{\nu,j}) \right\} \\ &\equiv \hat{C}_{ij}^{-1} (\Delta_{gg}^{ij} + \tilde{\Delta}_{gs}^{ij} + \tilde{\Delta}_{sg}^{ij} + \tilde{\Delta}_{ss}^{ij}), \end{aligned} \quad (29)$$

where we define the patchwise residual difference in cross-correlation products between data and mock catalogs $\tilde{\Delta}_{ab}^{ij}$ is defined as

$$\tilde{\Delta}_{ab}^{ij} \equiv \frac{1}{\mathcal{N}_{\times}} \sum_{\mu < \nu} \zeta_a^{\mu,i} \zeta_b^{\nu,j} - \zeta_a^{\text{mock},\mu} \zeta_b^{\text{mock},\nu}, \quad (30)$$

for $a \in \{g, s\}$ and $b \in \{g, s\}$. Note that when $a = b = g$, the first pure gravity-induced term reduces to the definition in Eq. (24).

If the systematics in the μ th and the ν th patches are identical, Eq. (29) trivially reduces to the first term of Eq. (23). However, when systematics vary between patches, auto and cross-statistics will yield different results due to their distinct computational approaches. Autocorrelation statistics compute the 4PCF with systematics applied globally across the entire survey volume, while cross-statistics compute the 4PCF with systematics applied separately to each individual patch before combining results. These two approaches can yield significantly different outcomes when systematic effects vary sufficiently between patches, particularly when the systematic variations are comparable to or larger than the underlying cosmological signal.

Similarly to Eq. (23), if the simulations accurately reproduce gravitational nonlinearity and systematics, Δ_{ab}^{ij} should vanish, and the difference between the two cross statistics should approach zero. In Eq. (29) we also see that gravity-induced nonlinearity is not the only cause for the cross-statistic to deviate from zero; if the systematics were not faithfully implemented, they can also potentially lead to a data-mock mismatch. For this purpose, we applied different tests in Sec. VI to quantify the impact of systematics at the signal level.

Note that the parity-even case studied here differs from the test applied to the parity-odd case presented in [69], where a discrepancy between the mock and data indicates a parity-violating signal. In the test for the gravitationally induced 4PCF, we are interested in both the agreement between the data and mocks, and in any deviation of the data minus mock distribution from the null distribution, where the latter indicates the detection significance.

As a further consistency check, we now turn to the statistical mismatch-sensitive statistic χ_{Δ}^2 defined as

$$\chi_{\Delta}^2 = \frac{1}{\mathcal{N}_{\Delta}} \sum_{\mu < \nu} (\hat{\zeta}_s^{\mu} - \hat{\zeta}_s^{\nu})^i (\hat{C}^{-1})_{ij} (\hat{\zeta}_s^{\mu} - \hat{\zeta}_s^{\nu})^j, \quad (31)$$

with \mathcal{N} being the normalization factor

$$\mathcal{N}_{\Delta} \equiv (N_p - 1) \sum_{\mu} \frac{V_{\text{eff}}}{V_{\text{eff}}^{\mu}}. \quad (32)$$

Following [69], we define the effective volume to be

$$V_{\text{eff}} \equiv \frac{V_{\text{fid}} N_{d.o.f}}{\text{Tr}[C_{\text{th}}^{-1} \hat{C}_{\text{mock}}]}, \quad (33)$$

where V_{fid} is the fiducial volume of the survey, obtained by fitting the analytic covariance to the EZmock. $N_{d.o.f}$ is the

number of degrees of freedom for the data vector, \mathbb{C}_{th} is the analytic covariance (see Sec. V C). If the analytic covariance is a good approximation of the mock covariance $\hat{\mathbb{C}}_{\text{mock}}$, $V_{\text{eff}} \approx V_{\text{fid}}$. However, in our case, due to the difference in the number density in different patches, there could be a significant difference between the two volumes.

The effective volume for each patch is similarly defined as $V_{\text{eff}}^{\mu} \equiv V_{\text{fid}} N_{d.o.f} / \text{Tr}[\mathbb{C}_{\text{th}}^{-1} \hat{\mathbb{C}}^{\mu}]$. Again, if the analytic covariance is a good approximation of the mock covariance, the definition of the effective volume for μ th patch reduces to a rescaling of the fiducial volume.

The expectation value of the mismatch-sensitive statistic is given as

$$\begin{aligned} \langle \chi_{\Delta}^2 \rangle &= \frac{1}{\mathcal{N}_{\Delta}} \sum_{\mu < \nu} \langle (\hat{\zeta}^{\mu} - \hat{\zeta}^{\nu})^i (\hat{\mathbb{C}}^{-1})_{ij} (\hat{\zeta}^{\mu} - \hat{\zeta}^{\nu})^j \rangle \\ &= \frac{1}{\mathcal{N}_{\Delta}} \left\{ (N_p - 1) \sum_{\mu} \text{Tr}[\hat{\mathbb{C}}^{-1} \hat{\mathbb{C}}^{\mu}] \right. \\ &\quad \left. + \sum_{\mu < \nu} \Delta \zeta_s^{\mu\nu, i} (\hat{\mathbb{C}}^{-1})_{ij} \Delta \zeta_s^{\mu\nu, j} \right\}, \end{aligned} \quad (34)$$

to obtain the second line, we used the definition of the trace $\text{Tr}[\hat{\mathbb{C}}^{-1} \hat{\mathbb{C}}^{\mu}] = \sum_i [\hat{\mathbb{C}}^{-1} \hat{\mathbb{C}}^{\mu}]_{ii} = \sum_{ij} (\hat{\mathbb{C}}^{-1})_{ij} \hat{\mathbb{C}}_{ij}^{\mu}$. We note that the deterministic gravity terms in Eq. (27) cancel between the different patches, while we still keep the potential difference in the systematics-induced bias across different patches $\Delta \zeta_s^{\mu\nu, i} \equiv \zeta_s^{\mu, i} - \zeta_s^{\nu, i}$.

From Eq. (34), we can see that the normalization coefficient in Eq. (32) is defined to recover the number of degrees of freedom of the data vector. We note that the definition assumes that the statistical fluctuations of different patches only differ by an overall rescaling, and the systematics-induced correction is negligible. However, the variation in the number densities and the difference in the fiducial and the effective volume as shown in Table I indicate that the statistical fluctuation across different patches may be more than an overall scaling factor. Moreover, there is no guarantee the systematics-induced bias should vanish. Consequently, the center of the χ_{Δ}^2 can deviate from the number of degrees of freedom.

The difference between χ_{Δ}^2 for data and mocks is given as follows:

$$\langle \chi_{\Delta, \text{data}}^2 \rangle - \langle \chi_{\Delta, \text{mock}}^2 \rangle = \frac{N_p - 1}{\mathcal{N}_{\Delta}} \sum_{\mu} \text{Tr}[\hat{\mathbb{C}}^{-1} (\mathbb{C}_{\text{data}}^{\mu} - \mathbb{C}_{\text{mock}}^{\mu})], \quad (35)$$

which quantifies the mismatch between the statistical fluctuation of the data and mocks.

C. Analytic covariance matrix

In this section, we provide a brief review of the analytic covariance matrix derived following the method outlined in [67]. The covariance for the NPCF coefficients in the isotropic basis reads

$$\begin{aligned} \mathbb{C}_{\Lambda, \Lambda'}(R, R') &= (4\pi)^{3N/2} \int \frac{s^2 ds}{V} \sum_G \sum_{\Lambda'', \mathcal{L}_G} (-1)^{\sum_i (-\ell_i - \ell'_i + \ell''_i)/2} \mathcal{B}_{\mathcal{L}_G, \Lambda}^{G^{-1}} \\ &\quad \times \mathcal{G}^{\mathcal{L}_G \Lambda' \Lambda''} \mathcal{D}_{\Lambda''} \mathcal{C}_0^{\Lambda''} \prod_{i=0}^{N-1} f_{\ell_{G_i} \ell'_i \ell''_i}(r_{G_i}, r'_i, s) |_{r_{G_0} = r'_0 = 0}. \end{aligned} \quad (36)$$

We let $\Lambda \equiv \{\ell_0, \dots, \ell_{N-1}\}$ and $\Lambda' \equiv \{\ell'_0, \dots, \ell'_{N-1}\}$ denoting the angular momenta associated with the position vectors $\mathbf{R} \equiv \{\mathbf{r}_0, \dots, \mathbf{r}_{N-1}\}$ and $\mathbf{R}' \equiv \{\mathbf{r}'_0, \dots, \mathbf{r}'_{N-1}\}$. \mathbf{s} is the vector denote the separation between the two copies of the N -tuplets of the density fields. Its corresponding primary angular momenta are $\Lambda'' = \{\ell''_0, \dots, \ell''_{N-1}\}$. The $N!$ permutations are given by the notation G . The permutation G can cause the order of the r_{G_i} to differ from the canonical ordering by length. This accounted for by defining $\mathcal{B}_{\mathcal{L}_G, \Lambda}^{G^{-1}}$ through

$$\mathcal{P}_{\mathcal{L}_G}(\hat{r}_{G_1}, \hat{r}_{G_2}, \dots) = \sum_{\Lambda} \mathcal{B}_{\mathcal{L}_G, \Lambda}^{G^{-1}} \mathcal{P}_{\Lambda}(\hat{r}_1, \hat{r}_2, \dots), \quad (37)$$

where the (r_1, r_2, \dots) are ordered by length.

For $N = 4$ with only three nonzero elements, the reordering operator reduces to a product of Kronecker deltas, $\delta_{\Lambda_i \mathcal{L}_{G_i}}^{\mathbf{K}}$, and is given by

$$\mathcal{B}_{\mathcal{L}_G, \Lambda}^{G^{-1}} = (-1)^{\sum_{i=0}^3 \ell_i (1 - \mathcal{E}_G)/2} \prod_{i=0}^{N-1} \delta_{\Lambda_i \mathcal{L}_{G_i}}^{\mathbf{K}}, \quad (38)$$

where $\mathcal{E}_G = 1$ for even permutations and $\mathcal{E}_G = -1$ for odd permutations. The generalized Gaunt integral reads [61]

$$\mathcal{G}^{\mathcal{L}_G \Lambda' \Lambda''} = (4\pi)^{-2} \left[\prod_{i=0}^3 \mathcal{D}_{\ell_{G_i} \ell'_i \ell''_i}^{\mathbf{P}} \mathcal{C}_{000}^{\ell_{G_i} \ell'_i \ell''_i} \right] \mathcal{Q}^{\mathcal{L}_G \Lambda' \Lambda''}, \quad (39)$$

with

$$\mathcal{Q}^{\mathcal{L}_G \Lambda' \Lambda''} = \prod_{i=0}^3 \sum_{m_{G_i} m'_i m''_i} c_{m_{G_i} m'_i m''_i}^{\ell_{G_i} \ell'_i \ell''_i} c_{M_G}^{\mathcal{L}_G} c_{M'}^{\Lambda'} c_{M''}^{\Lambda''}, \quad (40)$$

where

$$c_{m_1 m_2 m_3}^{\ell_1 \ell_2 \ell_3} \equiv \begin{pmatrix} \ell_1 & \ell_2 & \ell_3 \\ m_1 & m_2 & m_3 \end{pmatrix}. \quad (41)$$

In the case of $N = 4$, we have

$$\begin{aligned} \mathcal{C}_M^\Lambda &\equiv \mathcal{C}_{m_0 m_1 (m_*) m_2 m_3}^{\ell_0 \ell_1 (\ell_*) \ell_2 \ell_3} \\ &= (-1)^{\sum_{i=0}^3 \ell_i + \ell_*} \mathcal{D}_{\ell_*} \sum_{m_*} (-1)^{m_*} \mathcal{C}_{m_0 m_1 m_*}^{\ell_0 \ell_1 \ell_*} \mathcal{C}_{-m_* m_2 m_3}^{\ell_* \ell_2 \ell_3}, \end{aligned} \quad (42)$$

and similarly,

$$\begin{aligned} \mathcal{C}_0^{\Lambda''} &\equiv \mathcal{C}_{0000}^{\ell_0'' \ell_1'' (\ell_*'') \ell_2'' \ell_3''} \\ &= (-1)^{\sum_{i=0}^3 \ell_i'' + \ell_*''} \mathcal{D}_{\ell_*''} \mathcal{C}_{000}^{\ell_0'' \ell_1'' \ell_*''} \mathcal{C}_{000}^{\ell_*'' \ell_2'' \ell_3''} \end{aligned} \quad (43)$$

and

$$\begin{aligned} \mathcal{D}_{\Lambda''} &= \mathcal{D}_{\ell_0'' \ell_1'' \ell_2'' \ell_3''} \\ &\equiv \sqrt{(2\ell_0'' + 1)(2\ell_1'' + 1)(2\ell_2'' + 1)(2\ell_3'' + 1)}. \end{aligned} \quad (44)$$

Here we choose to fix $\ell_0' = 0$ and permute only ℓ' 's. Depending on the position of zeros for the angular momentum ℓ_{G0} , they will lead to seemingly different structure in the 6- j and 9- j symbols. In Appendix C, we discuss these three cases and show the unification of the coefficients for these three different cases.

The f -function is defined as

$$f_{\ell_i \ell_i' \ell_i''}(r_i, r_i', s) \equiv \int \frac{k^2 dk}{2\pi^2} j_{\ell_i}(kr_i) j_{\ell_i'}(kr_i') j_{\ell_i''}(ks) P(k), \quad (45)$$

where $P(k)$ is the power spectrum. In the case of discretized samples, the power spectrum also includes the shot noise; $P(k) \rightarrow P(k) + 1/\bar{n}$, with \bar{n} being the galaxy number density.

Finally, we fit the analytic covariance for its galaxy number density \bar{n}_g and the sample volume V_{fid} to the EZmock-FFA mocks following [67]. The fitted values are summarized in Table I. This fitted covariance can partially mitigate effects such as survey geometry and gravitational nonlinearity that are not fully modeled in the analytic approach. However, it does not capture all contributions, including potential systematic effects, which can lead to deviation from the Gaussian distribution assumption (see Sec. VII C).

D. Dimensionality reduction with eigendecomposition

Given the high incompleteness of the DESI DR1 sample, the mock-based covariance offers a more reliable estimate of the covariance matrix. However, due to the high dimensionality of the data vector and the limited number of the available mocks, data vector compression is needed before we proceed with the estimating the detection significance. For this purpose, we use the analytic covariance to compress the data vector. The analytic covariance matrix can be decomposed as

$$\mathbb{C}_{\text{th}} = U \Lambda U^T, \quad (46)$$

with Λ being the diagonal matrix of eigenvalues with the diagonal entries λ_i , and U the orthonormal matrix of eigenvectors, the 4PCF coefficients can be projected onto the eigenbasis of the analytic covariance as

$$\tilde{\zeta} = U^T \zeta. \quad (47)$$

Next, we select the components that have the highest signal-to-noise ratio

$$\text{SNR}_i = \frac{|\tilde{\zeta}_i|}{\lambda_i^{1/2}}, \quad (48)$$

where $\tilde{\zeta}_i$ is the mean of the 4PCF estimated from the simulations.

This compression method enables us to construct a hybrid covariance matrix by decomposing the 4PCFs measured from simulations (specifically, the 1,000 EZmock with FFA implementation used in this work) into the eigenbasis of the analytic covariance matrix. We select the top N_{eig} eigenmodes with the highest signal-to-noise ratio and use the corresponding rotated data vectors $\tilde{\zeta}$ to build a covariance matrix in the compressed basis:

$$\hat{\mathbb{C}}_{\text{hybrid},ij} = \frac{1}{N_{\text{mock}} - 1} \sum_{n=1}^{N_{\text{mock}}} (\tilde{\zeta}_i^{(n)} - \langle \tilde{\zeta} \rangle_i) (\tilde{\zeta}_j^{(n)} - \langle \tilde{\zeta} \rangle_j). \quad (49)$$

Due to the data compression, even though the covariance matrix is estimated from mocks and subject to the condition $N_{\text{sim}} \gg N_{d.o.f}$, the reduced size of the data vector allows for a smaller number of simulations to be sufficient.

In Sec. VI, we will compare the results obtained using both the analytic covariance and the hybrid approach, discussing the differences in the detected significance and the implications of assuming a Gaussian likelihood.

VI. RESULTS

In this section, we first present the detection significance based on the full autocorrelation statistic χ^2 , as defined in Eq. (15). However, as discussed in Sec. VA, the full autocorrelation may be affected by potential data-mock mismatches. To address this, we perform an additional test using the cross statistic χ^2_{\times} , described in Sec. VB.

Figure 4 presents the detection significance for the full DESI DR1 LRG sample, combining both the NGC and SGC regions. Each plot corresponds to a different number of eigenvalues, $N_{\text{eig}} = \{50, 100, 200\}$. We used the range $20h^{-1} \text{ Mpc} < r < 160h^{-1} \text{ Mpc}$ with ten equally spaced radial bins. This range and spacing is used in all the following analyses. The vertical lines represent the statistics from the DESI data. The predictions assuming the absence of any connected 4PFC are shown as black curves. For

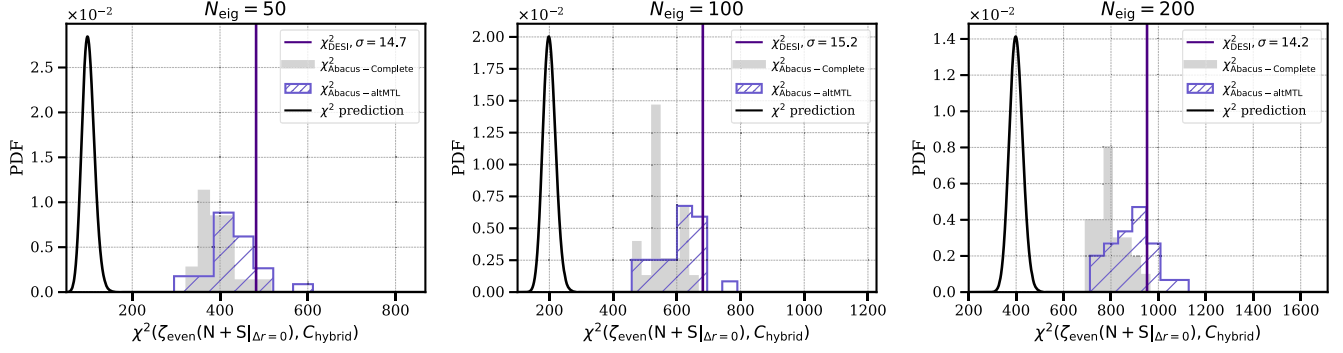


FIG. 4. Distribution of the autocorrelation χ^2 for the parity-even connected 4PCF, using the full DESI DR1 LRG sample, combining the NGC and SGC. The panel shows the comparison between simulations and data, as well as the detection significance of the gravitationally-induced connected 4PCF. Each plot corresponds to the number of eigenvalues used; $N_{\text{eig}} = \{50, 100, 200\}$. We used the range $20h^{-1} \text{ Mpc} < r < 160h^{-1} \text{ Mpc}$. The data vector and the corresponding EZmock-FFA-based covariance are both compressed according to Sec. V D. The vertical lines represent the statistics for the DESI data; the filled gray histograms correspond to the Abacus-Complete; the hashed histogram corresponds to the Abacus-altMTL; the black curves show the theoretical predictions of a χ^2 distribution, which represents the null-hypothesis distribution of the test statistic. We evaluated the significance using both the probability-to-exceed and the Wilson-Hilferty transformation, finding Gaussian-equivalent significances of 14.7σ , 15.2σ , and 14.2σ . In addition, we rescale all the distributions by the ratio of the fiber implementation given by the FFA and the altMTL fiber implementation, which reduces the χ^2 by 15%–20%.

further consistency checks, we show the distributions from the Abacus simulations; one using the altMTL implementation (hatched histogram), and the other from the Complete mocks without systematic effects (filled histogram), which allow us to isolate the impact of pure nonlinear gravitational evolution. We find that the two distributions from simulations are consistent with each other and with the DESI data, indicating that systematic effects do not artificially enhance the detection significance or lead to a misinterpretation of the gravity-induced connected parity-even 4PCF.

Since the observed χ_{obs}^2 is much higher than the number of degrees of freedom for the data vector, $\chi_{\text{obs}}^2 \gg \chi_{\text{d.o.f}}^2$, a probability-to-exceed (PTE)-based significance estimation will run into numerical instability. We therefore assess significance using the Wilson-Hilferty transformation [70] as a statistical approximation that maps a chi-squared distributed variable into an approximately normal distributed variable. Suppose $X \sim \chi_{\text{obs}}^2$, the Wilson-Hilferty transformation reads

$$\sigma_{\text{WH}} \equiv \frac{(X/N_{\text{d.o.f}})^{1/3} - (1 - 2/(9N_{\text{d.o.f}}))}{\sqrt{2/(9N_{\text{d.o.f}})}}. \quad (50)$$

For DESI DR1 LRG sample, we obtain Gaussian-equivalent values of 14.7σ , 15.2σ , and 14.2σ as shown in Fig. 4. We note that the Wilson-Hilferty transformation could yield a more conservative estimation of the significance. Alternatively, one can obtain a more accurate estimate of the tail probability using a saddle point approximation. The two approaches yield similar answers.

Since we applied the data compression scheme as in Sec. V D, the covariance matrix is essentially based on the EZmock-FFA. As we will see in Sec. VII D, the FFA fiber assignment implementation leads to an underestimation of the covariance compared to the altMTL implementation. We therefore rescale all distributions by the ratio between the FFA and altMTL samples, leading to a 15%–20% reduction in χ^2 . Additionally, we also apply the Hartlap factor $M \equiv (N_{\text{m}} - N_{\text{b}} - 2)/(N_{\text{m}} - 1)$ to debias the inverse of the sample covariance [71].

In addition to the autocorrelation, we perform a cross-correlation test as discussed in Sec. V B. Figure 5 shows the detection significance of the cross-correlation statistic χ_{\times}^2 for the DESI DR1 LRG sample between NGC and SGC. Each plot shows results for a different number of eigenvalues, $N_{\text{eig}} = \{50, 100, 200\}$. Both the data vector and the associated EZmock-FFA-based covariance matrix are compressed as described in Sec. V D. Vertical lines indicate the χ_{\times}^2 values obtained from the DESI data. The filled gray histograms represent the distribution from the Abacus-Complete mocks, while the histogram marked with circles corresponds to the Abacus-altMTL mocks. The theoretically predicted distribution of the signal-sensitive cross statistic is shown as the black curve. Under the null hypothesis, i.e., the signal term in Eq. (27) vanishes $\langle \hat{\xi}^{\mu} \hat{\xi}^{\nu} \rangle_{\xi^{(c)=0}} \rightarrow \delta_{\mu\nu}^{\text{K}} C^{\mu}$, leading to $\langle \chi_{\times, \text{null}}^2 \rangle = 0$. As a result, the black curve of Fig. 5 is centered around zero.

Next, to further quantify the significance of the signal-sensitive cross statistic, we evaluate its variance, which is defined by

$$\text{Var}(\chi_{\times, \text{null}}^2) = \langle (\chi_{\times, \text{null}}^2)^2 \rangle - \langle \chi_{\times, \text{null}}^2 \rangle^2, \quad (51)$$

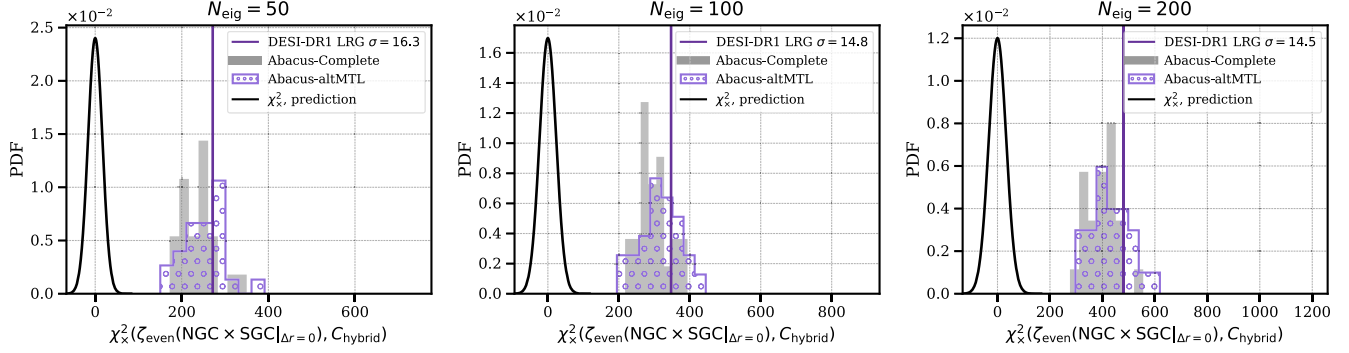


FIG. 5. Distribution of the cross-correlation χ^2_x for the parity-even connected 4PCF, using the full DESI DR1 LRG sample, combining the NGC and SGC. The black curves show the theoretical predictions for the cross statistic under the null-hypothesis of the test statistic, obtained by assuming a Gaussian distribution centered at zero with a variance given by Eq. (54). The rest of the labeling is identical to that in the preceding figure.

where the first term is

$$\langle (\chi^2_{x,\text{null}})^2 \rangle = \frac{1}{N_x^2} \sum_{\substack{\mu < \nu \\ \sigma < \rho}} \langle \hat{\zeta}_i^\mu (\hat{C}^{-1})_{ij} \hat{\zeta}_j^\nu \hat{\zeta}_k^\rho (\hat{C}^{-1})_{kl} \hat{\zeta}_l^\sigma \rangle. \quad (52)$$

This is to be evaluated in the Gaussian approximation where the four-point expectation is a sum of products of two-point expectations. The second term of Eq. (51) cancels the contribution where $\mu = \nu$ and $\sigma = \rho$. The remaining four-point expectation is

$$\langle \hat{\zeta}_i^\mu \hat{\zeta}_j^\nu \hat{\zeta}_k^\rho \hat{\zeta}_l^\sigma \rangle_{\substack{\mu < \nu \\ \sigma < \rho}} = \delta_{\mu\sigma}^K \hat{C}_{ii}^\mu \delta_{\nu\rho}^K \hat{C}_{jk}^\nu. \quad (53)$$

Inserting this, we have for the variance of the signal-sensitive cross statistic under the null hypothesis (also see [69])

$$\begin{aligned} \text{Var}(\chi^2_{x,\text{null}}) &= \frac{1}{N_x^2} \sum_{\mu < \nu} \text{Tr}[\mathbf{C}^{-1} \hat{C}^\mu \mathbf{C}^{-1} \hat{C}^\nu] \\ &\approx \frac{N_{d.o.f}}{N_x^2} \sum_{\mu < \nu} \frac{V_{\text{fid}}^2}{V_{\text{eff}}^\mu V_{\text{eff}}^\nu}, \end{aligned} \quad (54)$$

where V_{eff}^μ is the effective volume for each patch, defined in Eq. (33). Here we choose to normalize all the patches with respect to the full NGC, and thus V_{fid} corresponds to the volume of the full NGC. Here we approximate the distribution of χ^2_x as a Gaussian distribution, with Eq. (54), the width of the black curve is $\sigma(\chi^2_x)|_{N_{\text{eig}}=50} = 16.6$, $\sigma(\chi^2_x)|_{N_{\text{eig}}=100} = 23.5$, and $\sigma(\chi^2_x)|_{N_{\text{eig}}=200} = 33.2$. To test the estimation of the variance for the cross statistic in the absence of a real signal, we also analyze the parity-odd modes in Appendix A. The good agreement between the theoretical prediction and the mock distribution in Fig. 14 demonstrates the validity of the estimation.

As in Fig. 4, a fiber correction factor is also applied to account for the use of the FFA configuration in the EZmock

implementation. Comparing the significance found in Fig. 4 and Fig. 5, we found good agreement between both the autocorrelation and the cross-correlation.

As a further consistency check, we test null cross-correlation $\chi^2_{x,\text{null}}$ for the DESI DR1 LRG sample between NGC and SGC as shown in Fig. 7. This statistic is targeted at distinguishing the data-mock mismatch. As before, each plot here corresponds to the number of eigenvalues $N_{\text{eig}} = \{50, 100, 200\}$. The data vector and the corresponding EZmock-FFA-based covariance are both compressed according to Sec. VD. The vertical lines represent the statistics for the DESI data, the filled gray histograms correspond to the Abacus-Complete mocks, the histogram marked with circles corresponds to the Abacus-altMTL mocks. Here, we also apply the fiber correction factor. We find good agreement between the Abacus-altMTL mocks and data across the three numbers of eigenvalues, indicating that the

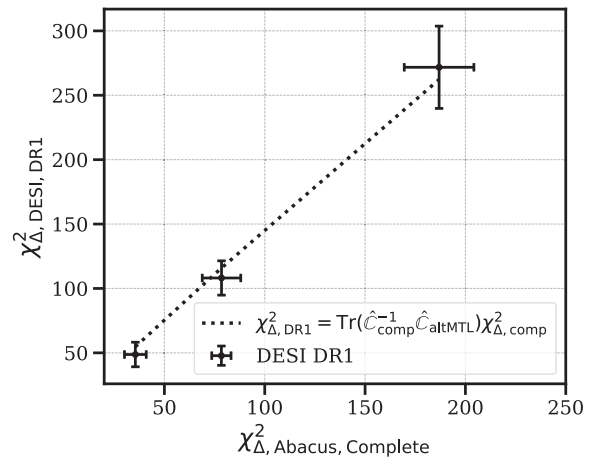


FIG. 6. χ^2_Δ for data and Abacus-Complete mocks. The relation between the two data sets can well be described using the ratio between the statistical fluctuation due to the fiber assignment effects. This demonstrates that the second term in Eq. (34) is negligible.

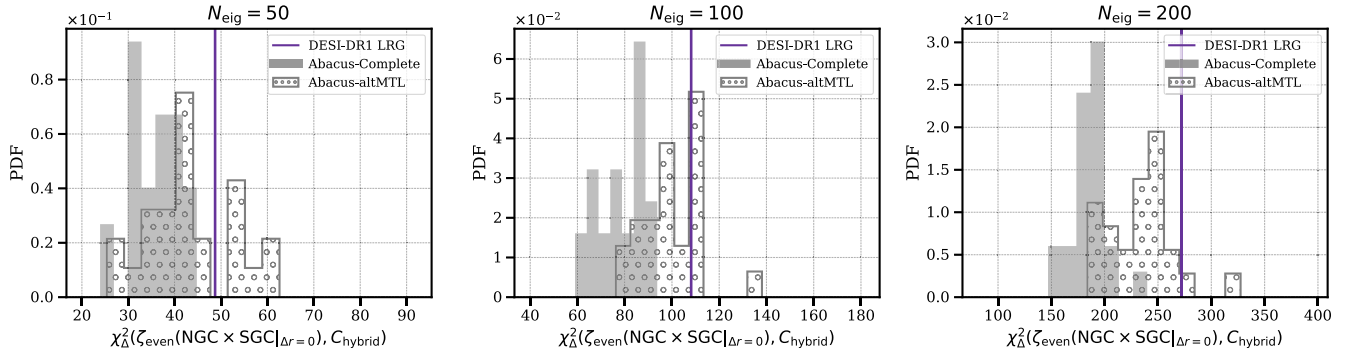


FIG. 7. Distribution of the χ^2_{Δ} statistic for the parity-even connected 4PCF, using the full DESI DR1 LRG sample, combining the NGC and SGC. Each plot corresponds to the number of eigenvalues $N_{\text{eig}} = \{50, 100, 200\}$. The data vector and the corresponding EZmock-FFA-based covariance are both compressed according to Sec. V D. The vertical lines represent the statistics for the DESI data; the filled gray histograms correspond to the Abacus-Complete mocks. The histogram marked by circles corresponds to the Abacus-altMTL mocks.

statistical fluctuations estimated from the mocks are consistent with the intrinsic fluctuations in the data. Moreover, the distribution also centers around the number of degrees of freedom, as predicted in Eq. (34).

At the same time, we find deviation between the data and the Abacus-Complete mocks and the deviation becomes larger as we increase the number of eigenvalues. The difference could potentially arise from two effects; larger statistical fluctuation due to incompleteness (see Sec. VII D) and systematics-induced bias [see second term of Eq. (34)]. To distinguish these two effects, we further compare the observed χ^2_{Δ} for the DESI DR1 LRG sample

and the Abacus-Complete mocks as shown in Fig. 6, with the error bars derived from the Abacus mocks. We find that the slope of the dashed line is well described by the ratio between the fiber assignment using altMTL-assigned Abacus mocks and the one without fiber assignment, demonstrating that the systematic bias in Eq. (34) does not contribute significantly and has a subdominant impact on the signal.

We highlight that despite the good agreement between the data and Abacus-altMTL, the center value of the distribution is higher than the number of degrees of freedom [cf. Eq. (34)], which may be a result of the large variation in incompleteness across different patches.

Figure 8 displays the detection significance for the DESI DR1 LRG sample, shown separately for the NGC and SGC, for $N_{\text{eig}} = 50$. The vertical lines indicate the statistics from the DESI data; the line-hatched histograms correspond to the Abacus-Complete mocks, the circle-hatched histograms to the Abacus-altMTL mocks, and the black curves represent the theoretical predictions. The distributions for both mocks are very close to each other, indicating that systematics do not artificially enhance the detection significance. While here we only display the result for $N_{\text{eig}} = 50$, we also find very similar results for $N_{\text{eig}} = 100$ and $N_{\text{eig}} = 200$.

Using Eq. (50), we find the significances for the NGC and SGC to be 12.9σ and 7.4σ . Approximating the two distributions to be independent Gaussian detections, the combined significance is approximately $\sqrt{\sigma_{\text{NGC}}^2 + \sigma_{\text{SGC}}^2} \approx 14.9\sigma$, which is in good agreement with the significance obtained from analyzing the combined sample directly.

VII. DISCUSSION

In this section, we discuss alternative configurations for the detection significance, including the scale dependence, impact of individual patches as listed in Table I, assumption

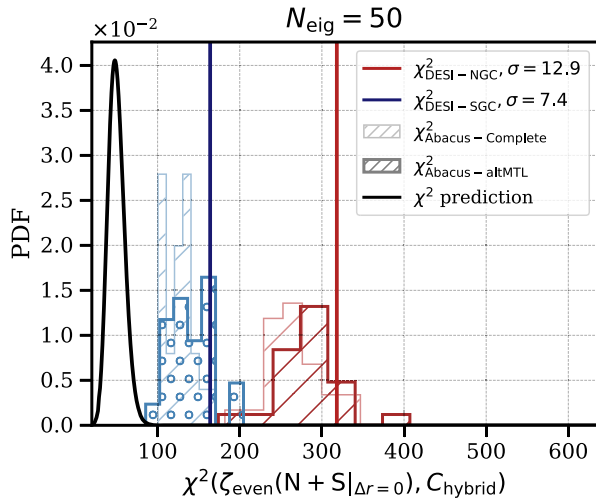


FIG. 8. Distribution of the autocorrelation χ^2 for the parity-even connected 4PCF. Here we show the DESI DR1 LRG sample for the NGC (red) and SGC (blue) separately with $N_{\text{eig}} = 50$. The vertical lines represent the statistics for the DESI data. The thinner-lined histograms correspond to the Abacus-Complete mock. The thicker-lined histograms are for the Abacus-altMTL. The black curve shows the theoretical prediction in the absence of a connected 4PCF.

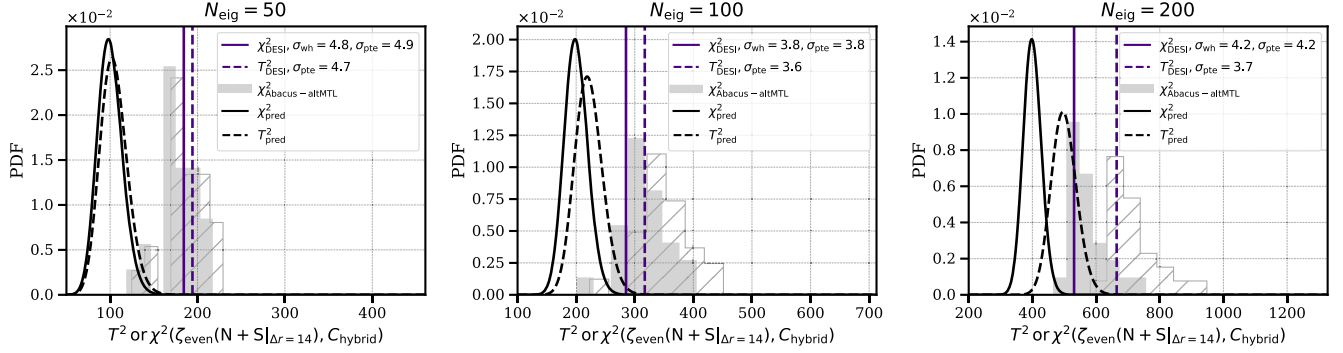


FIG. 9. Distribution of the autocorrelation χ^2 and T^2 for the parity-even connected 4PCF, using the full DESI DR1 LRG sample, combining the NGC and SGC. Each plot corresponds to the number of eigenvalues $N_{\text{eig}} = \{50, 100, 200\}$. We applied additional radial cuts $\Delta r = 14h^{-1}$ Mpc, to ensure that used galaxies do not fall in adjacent radial bins. We quantified the detection significance using two statistical distributions: the T^2 (dashed) and the χ^2 (solid). The vertical lines represent the statistics for the DESI data, the histograms correspond to the EZmock, and the black curves show the theoretical predictions under the null hypothesis. For the χ^2 distribution, we evaluated the significance using both the probability-to-exceed and the Wilson-Hilferty transformation, finding Gaussian-equivalent significances of 4.8σ , 3.8σ , and 4.2σ . The probability-to-exceed method yielded nearly identical values. For T^2 , we used only the probability-to-exceed approach and obtained detection significances of 4.7σ , 3.6σ , and 3.7σ .

of the Gaussianity of the likelihood, and the impact of sample incompleteness.

A. Scale dependence

Figure 9 shows the detection significance for the full DESI DR1 LRG sample combining NGC and SGC. Each plot corresponds to the number of eigenvalues $N_{\text{eig}} = \{50, 100, 200\}$. We applied additional radial cuts $\Delta r = 14h^{-1}$ Mpc to enforce a minimum separation between the radial bins, to ensure that used galaxies do not fall into adjacent radial bins.

Here we additionally quantified the detection significance using two statistical distributions: the T^2 (dashed) and the χ^2 (solid). The T^2 distribution is used because the covariance matrix is estimated from simulations, which modifies the likelihood to follow a T^2 form [72]. This distribution has been adopted in previous analyses [22,27], with an explicit expression given, e.g., in Eq. (20) of [22].

The vertical lines represent the statistics for the DESI data, the histograms correspond to the EZmock, and the black curves show the theoretical predictions for the null hypothesis. Both the data and the simulations show strong evidence for the presence of a connected four-point correlation. Due to the reduced detection, we can assess the significance using both the Wilson-Hilferty transformation and the probability-to-exceed by computing the area as the fraction of the distribution that lies beyond the detection significance observed in the data.

For the χ^2 distribution, we evaluated the significance using both methods, finding Gaussian-equivalent significances of 4.8σ , 3.8σ , and 4.2σ using the Wilson-Hilferty transformation. The probability-to-exceed method yielded nearly identical values. For T^2 , we used only the probability-to-exceed approach and obtained detection significances of

4.7σ , 3.6σ , and 3.7σ . Compared to the full range of scales in Fig. 4, we find a reduction in the detection, as expected, since the gravity-induced signal should be manifest mostly on small scales.

B. NGC-1, NGC-2, SGC-3

In this section, we examine the significance individually for NGC-1, NGC-2, and SGC-3. As before, we focus on the case with $N_{\text{eig}} = 100$, since the results do not show notable variation with the number of eigenmodes. Figure 10 shows the significance for the cross-correlation of the three combinations. From left to right, we have NGC-1 \times NGC-2, NGC-1 \times SGC-3, and NGC-2 \times NGC-3, with significance of 3.6σ , 8.5σ , and 1.6σ , respectively. In addition, we also find good agreement between the data measurement and the mock distribution, with or without fiber assignment. Due to the additional geometry cuts for each patch, the volume for each patch is smaller than the full volume, and we account for it when computing the variance of χ^2_{Δ} [cf. Eq. (54)]. Approximating the three distributions to be independent Gaussian detections, the combined significance is $\sigma_{\text{tot}} \approx \sqrt{\sigma_{\text{ngc-1}}^2 + \sigma_{\text{ngc-2}}^2 + \sigma_{\text{sgc-3}}^2} \approx 9.4$, which is less compared to the fiducial result due to the angular and radial cuts.

C. Testing Gaussianity of the likelihood

When analyzing high-dimensional data vectors, one needs to be careful about the assumption of Gaussianity in the likelihood function. Many statistical tests applied to the data rely on this assumption, under which the information can be effectively compressed into a single χ^2 statistic and interpreted in terms of a detection significance, expressed in units of standard deviations relative to

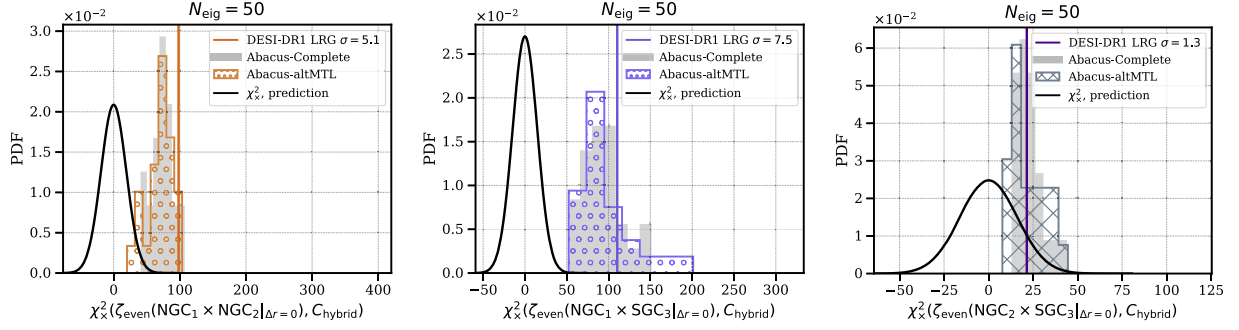


FIG. 10. Distribution of the cross-correlation χ^2_x for the parity-even connected 4PCF of the three patches: NGC-1, NGC-2, SGC-3. Here we only show $N_{\text{eig}} = 50$, given that the significance varies only marginally for different numbers of eigenvalues. The filled histograms correspond to the Abacus-Complete mocks. The hatched histograms and histograms marked with circles are for the Abacus-altMTL mocks. The vertical line corresponds to the measurement from the data and the black curves show the theoretical predictions for the χ^2_x distribution with null hypothesis. For all three plots, we also apply the fiber correction factor.

a Gaussian distribution. However, when the likelihood deviates significantly from Gaussianity, this interpretation may no longer be valid, and one must be cautious in drawing conclusions based on the χ^2 (or T^2)-derived significance.

Following [27], we assess the validity of the Gaussian likelihood assumption by examining the distribution of Cholesky-normalized data vectors, where the data vector is transformed into an uncorrelated mode space [73]. Figure 11 presents the distribution of the normalized

components of the data vector in mode space, sampled across all eigenmodes and simulations, for different choices of degrees of freedom. The black dashed line denotes the standard normal distribution $\mathcal{N}(0, \sigma_{\text{std}})$. In the left plot, this mode space normalization is performed using the inverse Cholesky decomposition L^{-1} of the hybrid covariance matrix $\hat{C}_{\text{hybrid}} = LL^T$, where the hybrid covariance is based on the EZmock, following Sec. VD. In the left panel, the mocks are split into two sets: one subset is used to estimate the covariance matrix, and the other to generate the test

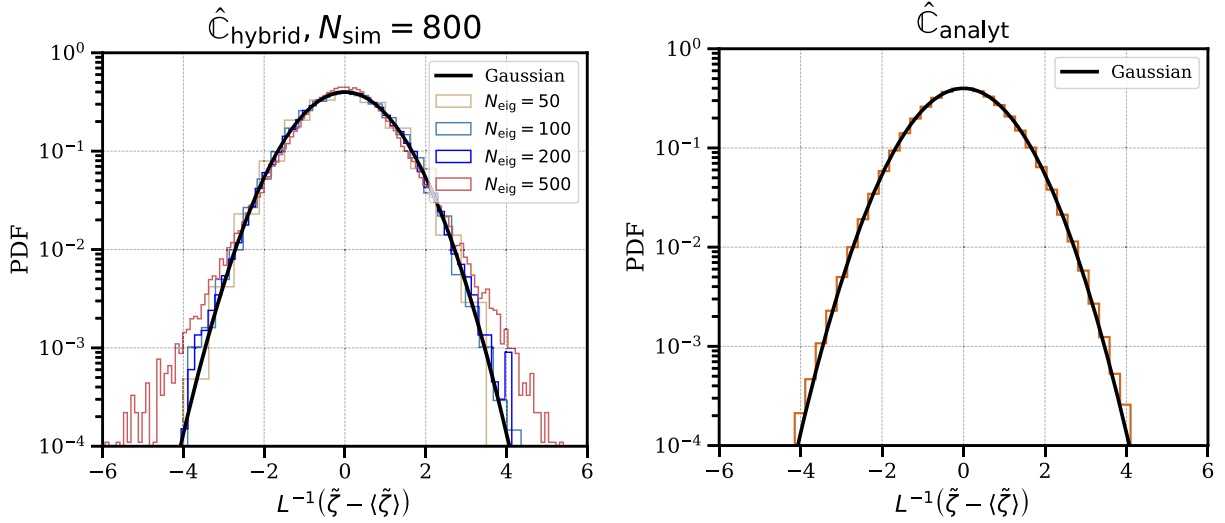


FIG. 11. The histograms show the distribution of the normalized components of the data vector in mode space. The black dashed line denotes the standard normal distribution $\mathcal{N}(0, 1)$. In the left, the mode space normalization is performed using the inverse Cholesky decomposition L^{-1} of the hybrid covariance matrix $\hat{C}_{\text{hybrid}} = LL^T$, where \hat{C}_{hybrid} is obtained from the EZmock, following Sec. VD. Here, we split the mock into two sets: one subset is used to estimate the covariance matrix, and the other to generate the test data vectors. We use 800 mocks for covariance estimation and 200 as test data. We vary different choices of N_{eig} (indicated by color). When the condition $N_{\text{cov}} \gg N_{\text{eig}}$ is not well-satisfied, the normalized data vector shows deviations from Gaussianity, indicating that the Gaussian likelihood assumption breaks down in this regime. In the right panel, we apply a Cholesky decomposition to the analytic covariance matrix and use the resulting Cholesky factor to normalize the unprojected data vector from the EZmock. Despite various simplified assumptions in the analytic covariance, we find good agreement with the Gaussian likelihood.

data vectors. We use 800 mocks for covariance estimation and 200 for testing. We find that when the condition $N_{\text{cov}} \gg N_{\text{eig}}$ is not well-satisfied, the normalized data vector shows deviations from Gaussianity, indicating that the Gaussian likelihood assumption breaks down in this regime.

In the right panel, we test the Gaussian likelihood assumption with the analytic covariance matrix. As above, here we apply the Cholesky decomposition to the analytic covariance matrix $\hat{C}_{\text{analyt}} = LL^T$, and use the Cholesky factor L to normalize the data vectors in the mode space. Although the analytic approach relies on simplified treatments of survey geometry and sample incompleteness, it nevertheless shows good consistency with a Gaussian likelihood. We further discuss the detection significance in Appendix B.

D. Impact of incompleteness due to statistical fluctuations

Due to the limited number of fibers and the large sample, the DR1 sample exhibits low completeness and is therefore sensitive to the details of fiber assignment implementation. As discussed in Sec. II, two fiber assignment implementation schemes were applied to the DESI mocks; the FFA and the altMTL. In the following, we briefly describe these two implementations and examine their impact on the covariance matrices.

We start by briefly describing the altMTL implementation. Each target sample has its own ‘‘Merged Target Ledger’’ [39], which tracks the observation history of the corresponding targets, including targets’ metadata, observation status, redshift information, priority, and etc., [52]. In addition to the priority, a subpriority field is generated to

assign small additional priority which FIBERASSIGN uses to reproducibly resolve collisions. This is a random number in the range 0 to 1. With the altMTL approach, the default is the inverse completeness weighting $w_{\text{comp}} = (f_{\text{TLID}}f_{\text{tile}})^{-1}$, with f_{TLID} the number of targets competing for a single fiber and f_{tile} is a tile-based completeness that quantifies the fraction of targets that are considered to have been ‘‘observed’’ within a group of overlapping tiles [cf. Eq. (13)].

The FFA is an emulator that learns the mapping between an input and a FIBERASSIGN-processed galaxy catalog, specifically predicting the probability of a target being assigned a fiber based on its properties and tiling information. This mapping function, \mathcal{F} , depends on two key variables: the close-friend count, n_{cfc} , and the number of tiles, n_{tile} . To compute n_{cfc} , a friend-of-friend (FOF) algorithm is applied to identify galaxies that are closely separated in angular space. The learning algorithm is optimized by comparing to the reference algorithm, in this case the output of the FIBERASSIGN code. With the FFA approach, while the default weight is also the inverse completeness weighting, it does not decompose f_{TLID} and f_{tile} , but rather learns the average effect of a mixed priory targets such as QSO and LRG as a function of number of overlapping tile n_{tile} .

Figure 12 shows the impact of different fiber assignment schemes on the statistical fluctuations. Left panel shows the diagonal elements of the covariance matrix from the Abacus altMTL (black), Abacus FFA (orange), and Abacus complete (gray). Right panel shows the ratio between the diagonal elements of Abacus simulations with different fiber assignment implementations and the complete sample. We find the average ratio between the Abacus altMTL to the complete sample (black) to be $\langle C_{\text{altMTL}}/C_{\text{comp}} \rangle \approx 1.12$,

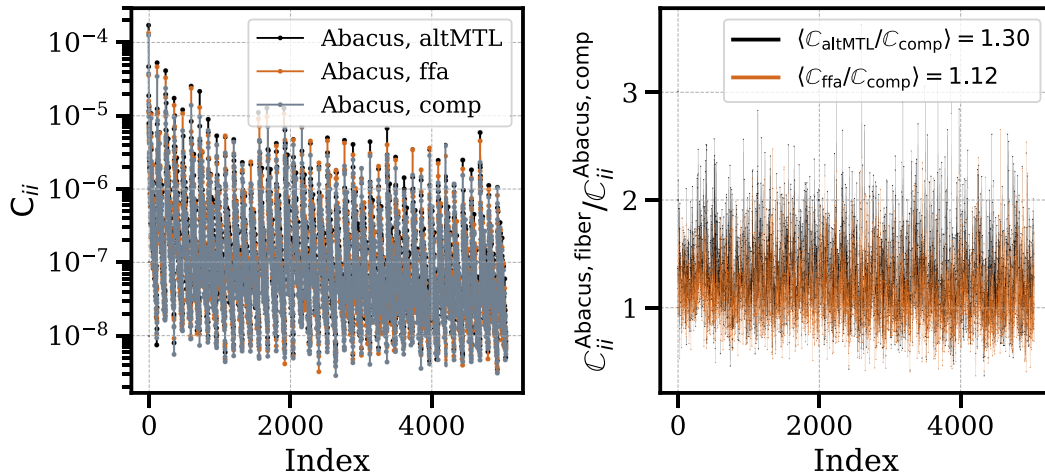


FIG. 12. The plot shows the statistical fluctuations in the Abacus simulations under different fiber assignment implementations. *Left*: Diagonal elements of the covariance matrix, with different colors indicating altMTL (black), FFA (orange), and a set of mocks without fiber assignment (complete, gray). *Right*: Ratio between the fiber-implemented mocks and the complete mocks. We find that fiber assignment induces an overall rescaling of the covariance matrix, increasing its amplitude by 12% and 30% for altMTL and FFA, respectively. The covariance is measured from the mocks for the NGC; we find a comparable ratio in the SGC.

whereas the average ratio between the *Abacus* FFA to the complete sample (orange) to be $\langle C_{\text{FFA}}/C_{\text{comp}} \rangle \approx 1.30$.

A reduced covariance in the FFA mocks was also observed in [53,74]. Several factors may contribute to the differences in statistical fluctuations as outlined in [75]; altMTL is a more realistic implementation, whereas FFA is based on an approximation. FFA exhibits a less pronounced collision window (i.e., the angular scales affected by the fiber collision effect) and therefore yields fewer zero-probability pairs compared to the altMTL mocks. For a fixed number density, this implies that the FFA-based mocks exhibit smaller density fluctuations, consistent with the observed reduction in covariance. We leave a detailed quantitative investigation to future work. However, since we use the FFA-based EZmock for covariance estimation, we account for the differences between the two fiber assignment schemes by applying an additional correction factor to our significance, as discussed in Sec. VI.

VIII. SUMMARY

In this paper we presented the measurement of the nonlinear gravity-induced parity-even connected 4PCF on the DESI DR1 LRG sample and quantified its detection significance. We found a 14.7σ significance for our fiducial configuration using an autocorrelation test with the number of eigenvalues $N_{\text{eig}} = 50$. Increasing the number of eigenmodes by a factor of two or four changes the significance by $\pm 3\%$, indicating robustness to the choice of eigenmode cutoff in our data compression method.

In addition, we tested various analysis setups, including scale dependence by imposing additional radial bin cuts, partitioning the sky into different patches and testing various combinations, and performing both auto and cross-correlation tests. Excluding smaller scales reduces the significance to 4.8σ , as expected, since small scales are more sensitive to nonlinear gravitational evolution. In contrast, combining sky patches in different ways or using the cross-correlation test yields results consistent with the fiducial configuration. In particular, the two cross statistics allow us to isolate the impact of general systematics by comparing the data to the *Abacus*-complete mocks. We find that systematics have subdominant impact on the signal.

To compute the significance, we compared the χ^2 and T^2 distributions and evaluated the probability using both the probability-to-exceed method and the Wilson-Hilferty transformation, finding only a marginal impact on the resulting detection significance.

In this paper, we also briefly discuss the Gaussian likelihood assumption and assess its validity by examining the distribution of Cholesky-normalized data vectors in mode space. We find that the Gaussian assumption holds well for simulation-based hybrid covariance estimates, provided that the number of simulations significantly exceeds the number of degrees of freedom by at least a factor of two $N_{\text{sim}} > 2N_{d.o.f.}$. For the analytic covariance,

despite the simplifying assumptions involved, we still find good agreement with the Gaussian likelihood. However, the autocorrelation test is more affected by the data-mock mismatch when using the analytic covariance, as discussed in Appendix B. We therefore present the hybrid method as our main result:

- (i) *Comparison to BOSS*: We reevaluate the detection significance for the BOSS CMASS sample using the Wilson-Hilferty transformation [Eq. (50)] and obtain 10.5σ , 10.4σ and 10.5σ using 10, 50, and 100 eigenvalues. We note that when the observed χ^2_{obs} lies in the extreme tail of the distribution, the probability-to-exceed method can become numerically unstable due to machine precision limitations. The BOSS CMASS sample contains 777,202 galaxies with a completeness exceeding 95%. Scaling the significance with the square root of the galaxy number yields an expected significance of approximately 17σ for DESI. After accounting for the impact of fiber assignment reduces this estimate, bringing it into agreement with our findings in this paper.
- (ii) *Fiber assignment and sample incompleteness*: We find that the fiber assignment implementation has a significant impact on the covariance (see Sec. VII D). This is likely due to the highly incomplete DESI DR1 sample. To account for this, we introduce a correction factor in our significance estimation. Nevertheless, when using the hybrid covariance, we find consistent results between the fiber-assigned mocks and the corresponding complete mocks without fiber assignment. The impact of fiber assignment is expected to be reduced in future data releases with increased sample completeness;
- (iii) *Imaging systematics*: For DESI LRG, major imaging systematics were identified in the DECam region, including $E(B - V)$ maps, r -band depth, and stellar density [53]. Since SGC uses entirely DECam photometry while NGC uses a mix of DECam and BASS/MzLS photometry, we tested cross-correlations between NGC and SGC regions to assess whether these DECam-specific systematics bias clustering measurements. The consistency between NGC-SGC cross-correlations and their respective autocorrelations demonstrates that these imaging systematics do not substantially contaminate the clustering signal used for cosmological inference. In addition, the angular integral constraint [76] was also considered but showed marginal impact on LRG measurements when using linear regression weights [53]. Since this effect primarily manifests on large scales, we do not explicitly test it here, though it should be verified for other tracers when quantifying the connected 4PCF;
- (iv) *Parity-odd 4PCF measurements*: Unlike the parity-even 4PCF, the parity-odd signal has a much

lower signal-to-noise ratio and is largely noise dominated, making its detection particularly sensitive to the covariance estimate. However, the DESI DR1 LRG sample suffers from high incompleteness, and the available mocks have limitations; the EZmock underestimates the covariance, while the Abacus mocks do not have sufficient volume. As a result, despite performing both autocorrelation and cross-correlation tests, we cannot draw a definitive conclusion regarding the parity-odd result in the DR1 LRG sample. We plan to revisit this analysis with future data releases. Consistency tests based on the parity-odd 4PCF analysis are presented in Appendix A.

ACKNOWLEDGMENTS

We thank Eiichiro Komatsu, Drew Jamieson, Ariel Sánchez and Sánchez’s group for useful discussions. We also thank Zachary Brown, Dragon Huterer, Alex Krolewski, Lado Samushia, and Benjamin Weaver for providing helpful comments on the manuscript. J. H. thanks Jue Fox for his office support. J. H. has received funding from the European Union’s Horizon 2020 research and innovation program under the Marie Skłodowska-Curie Grant Agreement No. 101025187. We acknowledge UFIT Research Computing for providing computational resources and support that have contributed to the research results reported in this publication. The work of R. N. C. was supported in part by the Director, Office of Science, Office of High Energy Physics, of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231. This material is based upon work supported by the U.S. Department of Energy (DOE), Office of Science, Office of High-Energy Physics, under Contract No. DE-AC02-05CH11231, and by the National Energy Research Scientific Computing Center, a DOE Office of Science User Facility under the same contract. Additional support for DESI was provided by the U.S. National Science Foundation (NSF), Division

of Astronomical Sciences under Contract No. AST-0950945 to the NSF’s National Optical-Infrared Astronomy Research Laboratory; the Science and Technology Facilities Council of the United Kingdom; the Gordon and Betty Moore Foundation; the Heising-Simons Foundation; the French Alternative Energies and Atomic Energy Commission (CEA); the National Council of Humanities, Science and Technology of Mexico (CONAHCYT); the Ministry of Science, Innovation and Universities of Spain (MICIU/AEI/10.13039/501100011033), and by the DESI Member Institutions: [77]. The authors are honored to be permitted to conduct scientific research on I’oligam Du’ag (Kitt Peak), a mountain with particular significance to the Tohono O’odham Nation.

Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the U.S. National Science Foundation, the U.S. Department of Energy, or any of the listed funding agencies.

DATA AVAILABILITY

The data that support the findings of this article are openly available [78].

APPENDIX A: PARITY-ODD 4PCF

In this section, we consider the parity-odd 4PCF, focusing on the consistency between the statistical fluctuations observed in simulations and those predicted by theoretical expectations. We present the consistency check for both the autocorrelation and the cross-correlation tests.

Figure 13 shows the autocorrelation for the parity-odd 4PCF combining the NGC and SGC. We do not show the measurements from the DESI data in the parity-odd case. We find that the correction due to the fiber assignment reduces the data χ^2 by 10%. For the autocorrelation test,

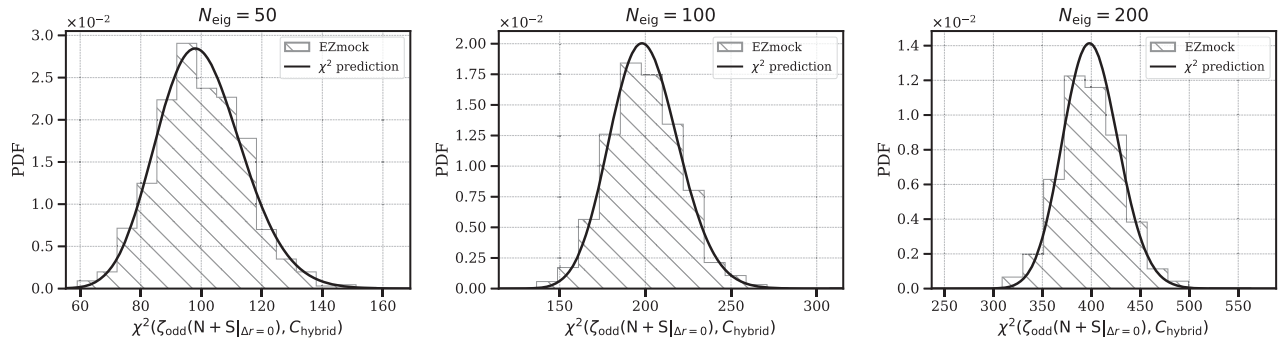


FIG. 13. Distribution of χ^2 for the parity-odd 4PCF combining the NGC and SGC using EZmock only. We do not display measurements from DESI Y1 LRG. The fiber correction reduces the data χ^2 by 10% in the odd sector. The black curve is the theoretical prediction in the absence of parity violation. For different choices of eigenmodes, we find the EZmock distributions agree well with the theoretical χ^2 distribution.

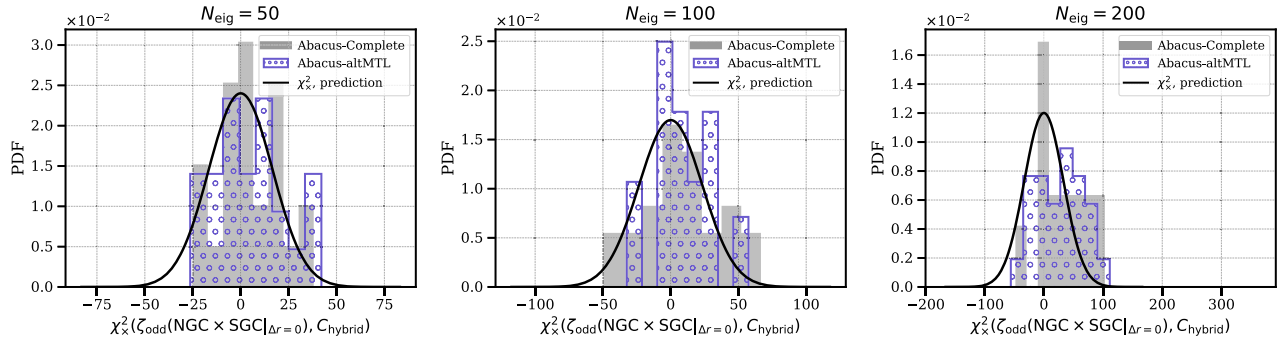


FIG. 14. Cross correlation χ^2_x of the parity-odd 4PCF between NGC and SGC using EZmock. We do not display measurements from DESI Y1 LRG. The EZmock-FFA-based covariance is compressed according to Sec. V D. The filled gray histograms correspond to the Abacus-Complete mocks; the histogram marked with circles corresponds to the Abacus-altMTL mocks. The black curves show the predictions for the cross statistic under the null hypothesis, assuming a zero-mean Gaussian distribution with variance given by Eq. (54).

we use only the EZmock-FFA, which has sufficient volume to reproduce the statistical fluctuation due to cosmic variation. We do not include the Abacus results here, as the box side length of the Abacus mocks is only one-third that of the EZmock simulations. This leads to increased statistical fluctuations and artificially high χ^2 values. The black curve is the theoretical prediction for a χ^2 distribution. The EZmock χ^2 distribution is well-described by the theoretical prediction.

Figure 14 shows the cross correlation χ^2_x of the parity-odd 4PCF between NGC and SGC using EZmock only. Since statistical fluctuations do not contribute to the cross-correlation, the smaller volume of the Abacus mocks does not have a direct consequence on the inferred χ^2_x , and we include all available mocks for this test.

The EZmock-FFA-based covariance is compressed according to Sec. V D. The filled gray histograms correspond to the Abacus-Complete mocks, and the histogram marked with circles corresponds to the Abacus-altMTL mocks. The black curves show the predictions for the cross statistic under the null hypothesis, assuming a Gaussian distribution with a zero mean with variance given

in Eq. (54). A fiber correction factor is applied to account for the fact that the covariance is calibrated with respect to EZmock-FFA.

We note that the good agreement between the theoretical prediction (black curve) and the mock distribution demonstrates that Eq. (54) provides a reliable estimate of the cross-statistic χ^2_x in the absence of a real signal.

APPENDIX B: DETECTION SIGNIFICANCE WITH ANALYTIC COVARIANCE

In this section, we show the results using the analytic covariance. In Fig. 15, we show the detection significance using the autostatistic (left), cross-statistic χ^2_x (middle), and cross-statistic χ^2_Δ (right) of the parity-even 4PCF measured from the full DESI DR1 LRG sample in both the NGC and SGC, using the analytic covariance matrix. In all three panels, the vertical lines are the measurements from the DESI DR1 LRG sample, while the hatched histograms correspond to the Abacus-altMTL and the gray histograms correspond to the Abacus-complete mocks.

As discussed in Sec. V A, the autocorrelation can suffer from data-mock mismatch, leading to an overestimation

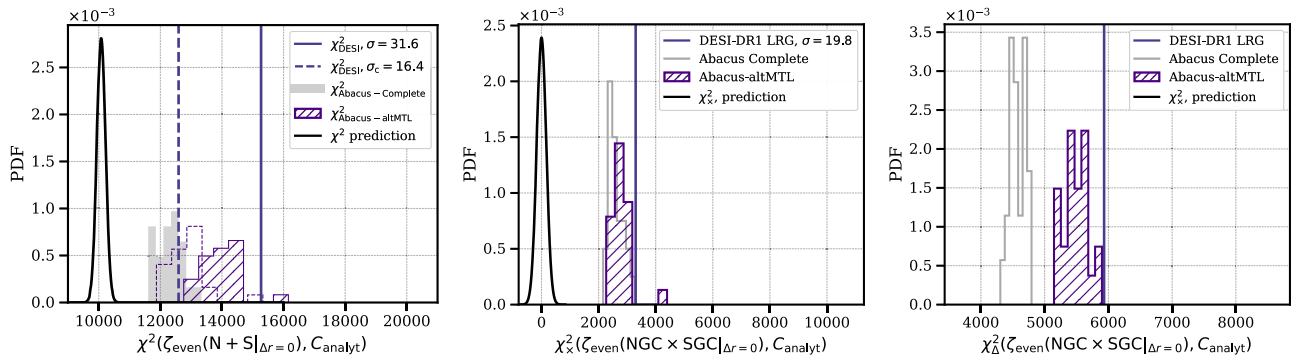


FIG. 15. Autostatistic (left), cross-statistic χ^2_x (middle), and cross-statistic χ^2_Δ (right) of the parity-even 4PCF measured from the full DESI DR1 LRG sample in both the NGC and SGC, using the analytic covariance matrix. The vertical lines represent the statistics for the DESI data. The histograms correspond to the Abacus mocks and the black curves show the theoretical predictions in the absence of a connected 4PCF.

of the detection significance. In the left panel, the uncorrected value shown by the solid vertical line corresponds to a 31.6σ detection significance, which is 60% higher compared to the cross-statistic in the middle panel. At the same time, there is a discrepancy between the data and the mock distribution, as well as a disagreement between the two Abacus mocks. To account for this, and following the procedure in Sec. VB, we apply χ_{Δ}^2 to correct for the mismatch. Specifically, we first compute the difference between the two Abacus mocks, $\delta\chi_{\Delta,1}^2 \equiv \chi_{\Delta,\text{altmtl}}^2 - \chi_{\Delta,\text{comp}}^2$, which quantifies the systematic-induced fluctuations. We then compute the difference between the data and the Abacus-altMTL, $\delta\chi_{\Delta,2}^2 \equiv \chi_{\Delta,\text{iron}}^2 - (\chi_{\Delta,\text{comp}}^2)$, which characterizes the residual fluctuation.

In the left panel, we apply $\delta\chi_{\Delta,1}^2$ to the Abacus-altMTL distribution (dashed histogram) and $\delta\chi_{\Delta,2}^2$ to the data (dashed vertical line). After these corrections, the mock distributions show good agreement with each other, and the resulting detection significance is 16.4σ , consistent with the value obtained using the compressed method.

APPENDIX C: GENERALIZED GAUNT INTEGRAL

In this section, we take a different approach to derive the analytic covariance, without explicitly introducing the ‘‘H’’ permutation as done in [67]. From Eq. (36), we first recognize that the prefactor of interest is

$$\begin{aligned} \mathcal{F}^{\mathcal{L}_G \Lambda' \Lambda''} &= \mathcal{E}_G \mathcal{G}^{\mathcal{L}_G \Lambda' \Lambda''} \mathcal{D}_{\Lambda''} \mathcal{C}_0^{\Lambda''} \\ &= (4\pi)^{-2} \mathcal{E}_G \prod_{i=0}^3 \mathcal{D}_{\ell_{G_i} \ell'_i \ell''_i} \mathcal{C}_{000}^{\ell_{G_i} \ell'_i \ell''_i} \\ &\quad \times \mathcal{Q}^{\mathcal{L}_G \Lambda' \Lambda''} \mathcal{D}_{\Lambda''} \mathcal{C}_0^{\Lambda''}, \end{aligned} \quad (\text{C1})$$

where \mathcal{E}_G denotes a permutation of $\{\ell_1, \ell_2, \ell_3\}$, with $\mathcal{E}_G = 1$ for an even permutation of the three indices, and with $\mathcal{E}_G = -1$ for an odd permutation. The generalized Gaunt integral $\mathcal{G}^{\mathcal{L}_G \Lambda' \Lambda''}$ is defined in Eq. (39). The coefficients $\mathcal{D}_{000}^{\ell_{G_i} \ell'_i \ell''_i}$ and $\mathcal{C}_{000}^{\ell_{G_i} \ell'_i \ell''_i}$ are defined in Eqs. (11) and (12), respectively, while $\mathcal{D}_{\Lambda''}$ and $\mathcal{C}_0^{\Lambda''}$ are defined in Eqs. (44) and (43). As mentioned in Sec. VC, we fix $\ell'_1 = 0$, such that $\ell'_* \rightarrow \ell'_1$ and $m'_* \rightarrow m'_1$. Given the definition of $\mathcal{Q}^{\mathcal{L}_G \Lambda' \Lambda''}$ in Eq. (40), we have

$$\begin{aligned} \mathcal{Q}^{\mathcal{L}_G \Lambda' \Lambda''} &= \sum_{MM'M''} (-1)^{\ell_{G0} + \ell'_1 - m_{G0} - m'_1} \mathcal{D}_{\ell_{G0} \ell'_1}^{-1} \delta_{\ell_{G0} \ell''_0}^{\text{K}} \delta_{m_{G0} - m''_0}^{\text{K}} \\ &\quad \times \sum_{\ell_* \ell'_*, m_* m''_*} (-1)^{\sum_{i=0}^3 \ell_i + \ell'_i + \ell''_i + \ell_* + \ell'_* + \ell''_*} \mathcal{D}_{\ell_* \ell'_* \ell''_*} \\ &\quad \times (-1)^{m_* + m'_1 + m''_*} \left[\prod_{i=1}^3 \mathcal{C}_{m_{G_i} m'_i m''_i}^{\ell_{G_i} \ell'_i \ell''_i} \right] \mathcal{C}_{m'_1 m'_2 m'_3}^{\ell'_1 \ell'_2 \ell'_3} \\ &\quad \times \mathcal{C}_{m_{G0} m_{G1} m_*}^{\ell_{G0} \ell_{G1} \ell_*} \mathcal{C}_{-m_* m_{G2} m_{G3}}^{\ell_* \ell_{G2} \ell_{G3}} \mathcal{C}_{m''_0 m''_1 m''_*}^{\ell''_0 \ell''_1 \ell''_*} \mathcal{C}_{-m''_* m''_2 m''_3}^{\ell''_* \ell''_2 \ell''_3}. \end{aligned} \quad (\text{C2})$$

Depending on the position of the zero among ℓ_{G_i} , for $i = \{0, 1, 2, 3\}$, different solutions can in principle arise. We consider the following cases: (1) $\ell_{G0} = 0$, where an entire column of the 9- j symbol vanishes; (2) $\ell_{G1} = 0$, where two elements of the first 9- j symbol are zero; and (3) $\ell_{G2} = 0$, where one element in each 9- j symbol is zero. The fourth case, where $\ell_{G3} = 0$, is trivially degenerate with case (3) and is therefore not treated separately.

For case (1), since we set $\ell_{G0} = 0$, $\ell''_0 \rightarrow 0$. Therefore, we have

$$\begin{aligned} \mathcal{Q}_I^{\mathcal{L}_G \Lambda' \Lambda''} &= (-1)^{\sum_{i=1}^3 \ell_{G_i} + \ell'_i + \ell''_i} \sum_{MM'M''} \left[\prod_{i=1}^3 \mathcal{C}_{m_{G_i} m'_i m''_i}^{\ell_{G_i} \ell'_i \ell''_i} \right] \\ &\quad \times \mathcal{C}_{m_{G1} m_{G2} m_{G3}}^{\ell_{G1} \ell_{G2} \ell_{G3}} \mathcal{C}_{m'_1 m'_2 m'_3}^{\ell'_1 \ell'_2 \ell'_3} \mathcal{C}_{m''_1 m''_2 m''_3}^{\ell''_1 \ell''_2 \ell''_3} \\ &= \begin{Bmatrix} \ell_{G1} & \ell_{G2} & \ell_{G3} \\ \ell'_1 & \ell'_2 & \ell'_3 \\ \ell''_1 & \ell''_2 & \ell''_3 \end{Bmatrix}, \end{aligned} \quad (\text{C3})$$

where we relabel $\ell_* \rightarrow \ell_{G1}$, $\ell'_* \rightarrow \ell'_1$, and $\ell''_* \rightarrow \ell''_1$. The phase $(-1)^{\ell_{G1} + \ell'_1 + \ell''_1}$ vanishes due to the presence of $\mathcal{C}_{000}^{\ell_{G1} \ell'_1 \ell''_1}$. Given the relabeling, we find

$$\begin{aligned} \mathcal{D}_{\Lambda''} \mathcal{C}_0^{\Lambda''} &= \mathcal{D}_{\ell''_1 \ell''_2 \ell''_3} \sqrt{2\ell''_1 + 1} (-1)^{\ell''_1} \mathcal{C}_{000}^{0 \ell''_1 \ell''_1} \mathcal{C}_{000}^{\ell''_1 \ell''_2 \ell''_3} \\ &= \mathcal{D}_{\ell''_1 \ell''_2 \ell''_3} \mathcal{C}_{000}^{\ell''_1 \ell''_2 \ell''_3}. \end{aligned} \quad (\text{C4})$$

Combining Eqs. (C3) and (C4), the prefactor for case (1) reads

$$\begin{aligned} \mathcal{F}_I^{\mathcal{L}_G \Lambda' \Lambda''} &= (4\pi)^{-2} \mathcal{E}_G \prod_{i=1}^3 \mathcal{D}_{\ell_{G_i} \ell'_i \ell''_i} \mathcal{C}_{000}^{\ell_{G_i} \ell'_i \ell''_i} \\ &\quad \times \begin{Bmatrix} \ell_{G1} & \ell_{G2} & \ell_{G3} \\ \ell'_1 & \ell'_2 & \ell'_3 \\ \ell''_1 & \ell''_2 & \ell''_3 \end{Bmatrix} \mathcal{D}_{\ell''_1 \ell''_2 \ell''_3} \mathcal{C}_{000}^{\ell''_1 \ell''_2 \ell''_3}. \end{aligned} \quad (\text{C5})$$

For case (2), we set $\ell_{G1} = 0$, therefore $\ell_{G0} \rightarrow \ell_{G1}$, $\ell''_0 \rightarrow \ell_{G1}$, $\ell'_* \rightarrow \ell'_1$, $\ell''_1 \rightarrow \ell'_1$, and $\ell_* \rightarrow \ell_{G1}$. We find

$$\begin{aligned} \mathcal{Q}_{II}^{\mathcal{L}_G \Lambda' \Lambda''} &= \sum_{MM'M''} (\mathcal{D}_{\ell_{G1} \ell'_1})^{-2} \mathcal{D}_{\ell_{G1} \ell'_1 \ell''_1} \mathcal{C}_{m_{G1} m_{G2} m_{G3}}^{\ell_{G1} \ell_{G2} \ell_{G3}} \mathcal{C}_{m'_1 m'_2 m'_3}^{\ell'_1 \ell'_2 \ell'_3} \\ &\quad \times \mathcal{C}_{m_{G2} m'_2 m''_2}^{\ell_{G2} \ell'_2 \ell''_2} \mathcal{C}_{m_{G3} m'_3 m''_3}^{\ell_{G3} \ell'_3 \ell''_3} \sum_{\ell''_1 m''_1} \mathcal{C}_{m_{G1} m'_1 m''_1}^{\ell_{G1} \ell'_1 \ell''_1} \mathcal{C}_{-m''_1 m'_2 m''_3}^{\ell''_1 \ell'_2 \ell''_3} \\ &= (\mathcal{D}_{\ell_{G1} \ell'_1})^{-2} \mathcal{D}_{\ell_{G1} \ell'_1 \ell''_1} \begin{Bmatrix} \ell_{G1} & \ell_{G2} & \ell_{G3} \\ \ell'_1 & \ell'_2 & \ell'_3 \\ \ell''_1 & \ell''_2 & \ell''_3 \end{Bmatrix}, \end{aligned} \quad (\text{C6})$$

where we have relabeled $\{\ell''_0, \ell_*\} \rightarrow \ell_{G1}$ and $\{\ell''_1, \ell'_*\} \rightarrow \ell'_1$ in the first line and $\ell''_* \rightarrow \ell''_1$ for the last line, which leads to a prefactor of $(\mathcal{D}_{\ell_{G1} \ell'_1})^2$. For case (2), we find

$$\mathcal{D}_{\Lambda''} \mathcal{C}_0^{\Lambda''} = \mathcal{D}_{\ell_1^{\prime\prime} \ell_2^{\prime\prime} \ell_3^{\prime\prime}} \mathcal{D}_{\ell_{G1} \ell_1^{\prime\prime}} (-1)^{\ell_1^{\prime\prime}} \mathcal{C}_{000}^{\ell_{G1} \ell_1^{\prime\prime}} \mathcal{C}_{000}^{\ell_1^{\prime\prime} \ell_2^{\prime\prime} \ell_3^{\prime\prime}}. \quad (\text{C7})$$

Combining Eqs. (C6) and (C7), the prefactor for case (2) becomes

$$\begin{aligned} \mathcal{F}_{\text{II}}^{\mathcal{L}_G \Lambda' \Lambda''} &= (4\pi)^{-2} \mathcal{E}_G \mathcal{D}_{\ell_{G1} \ell_1^{\prime\prime} \ell_1^{\prime\prime}} \mathcal{D}_{\ell_{G1} 0 \ell_{G1}} \mathcal{C}_{000}^{\ell_{G1} 0 \ell_{G1}} \mathcal{D}_{0 \ell_1^{\prime\prime} \ell_1^{\prime\prime}} \mathcal{C}_{000}^{0 \ell_1^{\prime\prime} \ell_1^{\prime\prime}} \\ &\times \prod_{i=2}^3 \mathcal{D}_{\ell_{Gi} \ell_i^{\prime\prime} \ell_i^{\prime\prime}} \mathcal{C}_{000}^{\ell_{Gi} \ell_i^{\prime\prime} \ell_i^{\prime\prime}} (\mathcal{D}_{\ell_{G1} \ell_1^{\prime\prime}})^{-2} \left\{ \begin{array}{ccc} \ell_{G1} & \ell_{G2} & \ell_{G3} \\ \ell_1^{\prime\prime} & \ell_2^{\prime\prime} & \ell_3^{\prime\prime} \\ \ell_1^{\prime\prime} & \ell_2^{\prime\prime} & \ell_3^{\prime\prime} \end{array} \right\} \\ &\times \mathcal{D}_{\ell_1^{\prime\prime} \ell_2^{\prime\prime} \ell_3^{\prime\prime}} \mathcal{D}_{\ell_{G1} \ell_1^{\prime\prime}} (-1)^{\ell_1^{\prime\prime}} \mathcal{C}_{000}^{\ell_{G1} \ell_1^{\prime\prime}} \mathcal{C}_{000}^{\ell_1^{\prime\prime} \ell_2^{\prime\prime} \ell_3^{\prime\prime}}, \quad (\text{C8}) \end{aligned}$$

since $\mathcal{C}_{000}^{\ell_{G1} 0 \ell_{G1}} = (-1)^{\ell_{G1}} \mathcal{D}_{\ell_{G1}}^{-1}$ and $\mathcal{C}_{000}^{0 \ell_1^{\prime\prime} \ell_1^{\prime\prime}} = (-1)^{\ell_1^{\prime\prime}} \mathcal{D}_{\ell_1^{\prime\prime}}^{-1}$, the coefficients cancels with $\mathcal{D}_{\ell_{G1} \ell_1^{\prime\prime}}$ which arises in Eq. (C7). The resulting phase $(-1)^{\ell_{G1} + \ell_1^{\prime\prime} + \ell_1^{\prime\prime}}$ also cancels due to the presence of $\mathcal{C}_{000}^{\ell_{G1} \ell_1^{\prime\prime} \ell_1^{\prime\prime}}$. As a result, we have

$$\begin{aligned} \mathcal{F}_{\text{II}}^{\mathcal{L}_G \Lambda' \Lambda''} &= (4\pi)^{-2} \mathcal{E}_G \prod_{i=1}^3 \mathcal{D}_{\ell_{Gi} \ell_i^{\prime\prime} \ell_i^{\prime\prime}} \mathcal{C}_{000}^{\ell_{Gi} \ell_i^{\prime\prime} \ell_i^{\prime\prime}} \\ &\times \left\{ \begin{array}{ccc} \ell_{G1} & \ell_{G2} & \ell_{G3} \\ \ell_1^{\prime\prime} & \ell_2^{\prime\prime} & \ell_3^{\prime\prime} \\ \ell_1^{\prime\prime} & \ell_2^{\prime\prime} & \ell_3^{\prime\prime} \end{array} \right\} \mathcal{D}_{\ell_1^{\prime\prime} \ell_2^{\prime\prime} \ell_3^{\prime\prime}} \mathcal{C}_{000}^{\ell_1^{\prime\prime} \ell_2^{\prime\prime} \ell_3^{\prime\prime}}, \quad (\text{C9}) \end{aligned}$$

where case (1) and (2) share the same coefficients $\mathcal{F}^{\mathcal{L}_G \Lambda' \Lambda''}$.

For case (3), we set $\ell_{G2} = 0$, such that $\ell_{G0} = \ell_{G2}$, $\ell_0^{\prime\prime} = \ell_{G2}$, $\ell_*^{\prime\prime} = \ell_1^{\prime\prime}$, $\ell_* = \ell_{G3}$, $\ell_2^{\prime\prime} = \ell_2^{\prime\prime}$. We therefore find

$$\begin{aligned} \mathcal{Q}_{\text{III}}^{\mathcal{L}_G \Lambda' \Lambda''} &= \mathcal{D}_{\ell_{G2} \ell_1^{\prime\prime} \ell_2^{\prime\prime} \ell_{G3}}^{-1} \mathcal{D}_{\ell_{G3} \ell_1^{\prime\prime} \ell_2^{\prime\prime}} (-1)^{\ell_{G2} + \ell_1^{\prime\prime} + \ell_2^{\prime\prime} + \ell_{G3}} \\ &\times \sum_{\ell_*^{\prime\prime} \ell_* m_* m_*^{\prime\prime}} (-1)^{\sum_{i=0}^3 \ell_i + \ell_i^{\prime\prime} + \ell_i^{\prime\prime} + \ell_{G3} + \ell_1^{\prime\prime} + \ell_2^{\prime\prime}} \\ &\times \sum_{m'_s} (-1)^{m_{G2} + m'_1 + m'_2 + m_{G3}} (-1)^{m_{G3} + m'_1 + m'_2} \\ &\times \mathcal{C}_{m_{G1} m'_1 m_1^{\prime\prime}}^{\ell_{G1} \ell_1^{\prime\prime} \ell_1^{\prime\prime}} \mathcal{C}_{m_{G2} m'_2 m_2^{\prime\prime}}^{\ell_{G2} \ell_2^{\prime\prime} \ell_2^{\prime\prime}} \mathcal{C}_{m_1^{\prime\prime} m'_2 m_3^{\prime\prime}}^{\ell_1^{\prime\prime} \ell_2^{\prime\prime} \ell_3^{\prime\prime}} \mathcal{C}_{m_{G2} m'_1 m_*^{\prime\prime}}^{\ell_{G2} \ell_1^{\prime\prime} \ell_*^{\prime\prime}} \\ &\times \mathcal{C}_{-m_*^{\prime\prime} m'_2 m_3^{\prime\prime}}^{\ell_*^{\prime\prime} \ell_2^{\prime\prime} \ell_3^{\prime\prime}} \mathcal{C}_{m_{G3} m'_3 m_3^{\prime\prime}}^{\ell_{G3} \ell_3^{\prime\prime} \ell_3^{\prime\prime}}, \quad (\text{C10}) \end{aligned}$$

where the 3- j connection structure is represented by the type-II Yutsis diagram in Fig. 16. Here all the phases involving summation of angular momenta ℓ 's and m 's all cancel with each other.

To restore the 9- j symbol, we need to rearrange the product of the 3- j symbols; $\mathcal{C}_{m_{G2} m'_1 m_*^{\prime\prime}}^{\ell_{G2} \ell_1^{\prime\prime} \ell_*^{\prime\prime}} \mathcal{C}_{-m_*^{\prime\prime} m'_2 m_3^{\prime\prime}}^{\ell_*^{\prime\prime} \ell_2^{\prime\prime} \ell_3^{\prime\prime}}$. In order to do so, we first recognize the property of the integral for the product of four spherical harmonics:

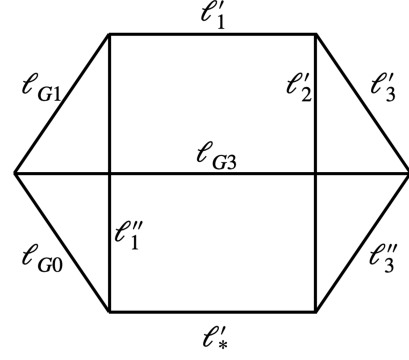


FIG. 16. Yutsis diagram for the type-II 9- j symbol.

$$\begin{aligned} &\int d\hat{n} Y_{\ell_{G2} m_{G2}}(\hat{n}) Y_{\ell_1^{\prime\prime} m_1^{\prime\prime}}(\hat{n}) Y_{\ell_2^{\prime\prime} m_2^{\prime\prime}}(\hat{n}) Y_{\ell_3^{\prime\prime} m_3^{\prime\prime}}(\hat{n}) \\ &= \sum_{\ell_*^{\prime\prime} m_*^{\prime\prime}} (-1)^{m_*^{\prime\prime}} \mathcal{G}_{m_{G2} m_2^{\prime\prime} m_*^{\prime\prime}}^{\ell_{G2} \ell_2^{\prime\prime} \ell_*^{\prime\prime}} \mathcal{G}_{-m_*^{\prime\prime} m_1^{\prime\prime} m_3^{\prime\prime}}^{\ell_*^{\prime\prime} \ell_1^{\prime\prime} \ell_3^{\prime\prime}} \\ &= \sum_{\ell_*^{\prime\prime} m_*^{\prime\prime}} (-1)^{m_*^{\prime\prime}} \mathcal{G}_{m_{G2} m_1^{\prime\prime} m_*^{\prime\prime}}^{\ell_{G2} \ell_1^{\prime\prime} \ell_*^{\prime\prime}} \mathcal{G}_{-m_*^{\prime\prime} m_2^{\prime\prime} m_3^{\prime\prime}}^{\ell_*^{\prime\prime} \ell_2^{\prime\prime} \ell_3^{\prime\prime}}, \quad (\text{C11}) \end{aligned}$$

where the Gaunt integral $\mathcal{G}_{m_1 m_2 m_3}^{\ell_1 \ell_2 \ell_3}$ involves product of 3 j symbols both with all m 's being zero $m_1 = m_2 = m_3 = 0$ as well as the ones being nonzero. For the 3 j symbols with all m 's being zero, they are defined as

$$\begin{pmatrix} \ell_1 & \ell_2 & \ell_3 \\ 0 & 0 & 0 \end{pmatrix} = \begin{cases} 0, & J \text{ odd,} \\ \frac{(-1)^{L/2} (L/2)! ((L-2\ell_i)!)^{1/2}}{((L+1)!)^{1/2} (L/2-\ell_i)!}, & J \text{ even,} \end{cases} \quad (\text{C12})$$

where $L \equiv \ell_1 + \ell_2 + \ell_3$. As a result, we can re-order the arrangement of the product

$$\mathcal{C}_{000}^{\ell_{G2} \ell_2^{\prime\prime} \ell_*^{\prime\prime}} \mathcal{C}_{000}^{\ell_*^{\prime\prime} \ell_2^{\prime\prime} \ell_3^{\prime\prime}} = \mathcal{C}_{000}^{\ell_{G2} \ell_2^{\prime\prime} \ell_*^{\prime\prime}} \mathcal{C}_{000}^{\ell_*^{\prime\prime} \ell_2^{\prime\prime} \ell_3^{\prime\prime}}. \quad (\text{C13})$$

Combining with Eq. (C11), we find that

$$\mathcal{C}_{m_{G2} m_2^{\prime\prime} m_*^{\prime\prime}}^{\ell_{G2} \ell_2^{\prime\prime} \ell_*^{\prime\prime}} \mathcal{C}_{-m_*^{\prime\prime} m_1^{\prime\prime} m_3^{\prime\prime}}^{\ell_*^{\prime\prime} \ell_1^{\prime\prime} \ell_3^{\prime\prime}} = \mathcal{C}_{m_{G2} m_1^{\prime\prime} m_*^{\prime\prime}}^{\ell_{G2} \ell_1^{\prime\prime} \ell_*^{\prime\prime}} \mathcal{C}_{-m_*^{\prime\prime} m_2^{\prime\prime} m_3^{\prime\prime}}^{\ell_*^{\prime\prime} \ell_2^{\prime\prime} \ell_3^{\prime\prime}}. \quad (\text{C14})$$

By relabeling $\ell_*^{\prime\prime} \rightarrow \ell_2^{\prime\prime}$, case-III can be written as

$$\mathcal{Q}_{\text{III}}^{\mathcal{L}_G \Lambda' \Lambda''} = \mathcal{D}_{\ell_{G2} \ell_2^{\prime\prime}}^{-1} \mathcal{D}_{\ell_2^{\prime\prime}} \left\{ \begin{array}{ccc} \ell_{G1} & \ell_{G2} & \ell_{G3} \\ \ell_1^{\prime\prime} & \ell_2^{\prime\prime} & \ell_3^{\prime\prime} \\ \ell_1^{\prime\prime} & \ell_2^{\prime\prime} & \ell_3^{\prime\prime} \end{array} \right\}. \quad (\text{C15})$$

For case (3), we find

$$\mathcal{D}_{\Lambda''} \mathcal{C}_0^{\Lambda''} = \mathcal{D}_{\ell_1^{\prime\prime} \ell_2^{\prime\prime} \ell_3^{\prime\prime}} \mathcal{D}_{\ell_{G2} \ell_2^{\prime\prime}} (-1)^{\ell_2^{\prime\prime}} \mathcal{C}_{000}^{\ell_{G2} \ell_2^{\prime\prime} \ell_1^{\prime\prime}} \mathcal{C}_{000}^{\ell_1^{\prime\prime} \ell_2^{\prime\prime} \ell_3^{\prime\prime}}. \quad (\text{C16})$$

Combining Eqs. (C15) and (C16), we find

$$\mathcal{F}_{\text{III}}^{\mathcal{L}_G \Lambda' \Lambda''} = (4\pi)^{-2} \mathcal{E}_G \prod_{i=1}^3 \mathcal{D}_{\ell_{G_i} \ell'_i \ell''_i} \mathcal{C}_{000}^{\ell_{G_i} \ell'_i \ell''_i} \left\{ \begin{array}{ccc} \ell_{G1} & \ell_{G2} & \ell_{G3} \\ \ell'_1 & \ell'_2 & \ell'_3 \\ \ell''_1 & \ell''_2 & \ell''_3 \end{array} \right\} \mathcal{D}_{\ell''_1 \ell''_2 \ell''_3} \mathcal{C}_{000}^{\ell''_1 \ell''_2 \ell''_3}, \quad (\text{C17})$$

Comparing Eqs. (C5), (C9), and (C17), we find that the coefficients can be unified, regardless of where the zero appears in the 3n-j symbols.

-
- [1] J. Maldacena, *J. High Energy Phys.* **05** (2003) 013.
[2] N. Bartolo, E. Komatsu, S. Matarrese, and A. Riotto, *Phys. Rep.* **402**, 103 (2004).
[3] X. Chen, *Adv. Astron.* **2010**, 638979 (2010).
[4] I. Szapudi, *Astrophys. J. Lett.* **605**, L89 (2004).
[5] Z. Slepian and D. J. Eisenstein, *Mon. Not. R. Astron. Soc.* **454**, 4142 (2015).
[6] R. Scoccimarro, *Phys. Rev. D* **92**, 083532 (2015).
[7] N. S. Sugiyama, S. Saito, F. Beutler, and H.-J. Seo, *Mon. Not. R. Astron. Soc.* **484**, 364 (2018).
[8] G. D'Amico, Y. Donath, M. Lewandowski, L. Senatore, and P. Zhang, *J. Cosmol. Astropart. Phys.* **05** (2024) 059.
[9] O. H. E. Philcox and M. M. Ivanov, *Phys. Rev. D* **105**, 043517 (2022).
[10] A. Eggemeier, N. Lee, R. Scoccimarro, B. Camacho-Quevedo, A. Pezzotta, M. Crocce, and A. G. Sánchez, *Phys. Rev. D* **112**, 063532 (2025).
[11] M. Schmittfull and A. M. Dizgah, *J. Cosmol. Astropart. Phys.* **03** (2021) 020.
[12] A. M. Dizgah, H. Lee, M. Schmittfull, and C. Dvorkin, *J. Cosmol. Astropart. Phys.* **04** (2020) 011.
[13] J. Hou, A. Moradinezhad Dizgah, C. Hahn, and E. Massara, *J. Cosmol. Astropart. Phys.* **03** (2023) 045.
[14] J. Hou, A. M. Dizgah, C. Hahn, M. Eickenberg, S. Ho, P. Lemos, E. Massara, C. Modi, L. Parker, and B. R.-S. Blancard, *Phys. Rev. D* **109**, 103528 (2024).
[15] E. Paillas, Y.-C. Cai, N. Padilla, and A. G. Sánchez, *Mon. Not. R. Astron. Soc.* **505**, 5731 (2021).
[16] M. Eickenberg *et al.*, arXiv:2204.07646.
[17] B. Régaldo-Saint Blancard, C. Hahn, S. Ho, J. Hou, P. Lemos, E. Massara, C. Modi, A. Moradinezhad Dizgah, L. Parker, Y. Yao, and M. Eickenberg, *Phys. Rev. D* **109**, 083535 (2024).
[18] G. Valogiannis and C. Dvorkin, *Phys. Rev. D* **105**, 103534 (2022).
[19] M. White, *J. Cosmol. Astropart. Phys.* **11** (2016) 057.
[20] R. K. Sheth, A. J. Connolly, and R. Skibba, arXiv:astro-ph/0511773.
[21] R. N. Cahn, Z. Slepian, and J. Hou, *Phys. Rev. Lett.* **130**, 201002 (2023).
[22] J. Hou, Z. Slepian, and R. N. Cahn, *Mon. Not. R. Astron. Soc.* **522**, 5701 (2023).
[23] O. H. E. Philcox, *Phys. Rev. D* **106**, 063501 (2022).
[24] D. Jamieson, A. Caravano, J. Hou, Z. Slepian, and E. Komatsu, *Mon. Not. R. Astron. Soc.* **533**, 2582 (2024).
[25] D. Jeong and M. Kamionkowski, *Phys. Rev. Lett.* **108**, 251301 (2012).
[26] D. Gualdi, H. Gil-Marín, and L. Verde, *J. Cosmol. Astropart. Phys.* **07** (2021) 008.
[27] O. H. E. Philcox, J. Hou, and Z. Slepian, arXiv:2108.01670.
[28] K. S. Dawson *et al.*, *Astron. J.* **145**, 10 (2013).
[29] M. Levi, C. Bebek, T. Beers, R. Blum, R. Cahn, D. Eisenstein, B. Flaugher, K. Honscheid, R. Kron, O. Lahav, P. McDonald, N. Roe, and D. Schlegel (DESI Collaboration), arXiv:1308.0847.
[30] A. Aghamousa *et al.* (DESI Collaboration), arXiv:1611.00036.
[31] A. Aghamousa *et al.* (DESI Collaboration), arXiv:1611.00037.
[32] B. Abareshi *et al.* (DESI Collaboration), *Astron. J.* **164**, 207 (2022).
[33] M. Abdul-Karim *et al.* (DESI Collaboration), arXiv:2503.14745.
[34] A. G. Adame *et al.* (DESI Collaboration), *J. Cosmol. Astropart. Phys.* **07** (2025) 028.
[35] M. Abdul-Karim *et al.* (DESI Collaboration), *Phys. Rev. D* **112**, 083515 (2025).
[36] E. F. Schlafly *et al.* (DESI Collaboration), *Astron. J.* **166**, 259 (2023).
[37] C. Poppett *et al.*, *Astron. J.* **168**, 245 (2024).
[38] J. Guy *et al.*, *Astron. J.* **165**, 144 (2023).
[39] A. D. Myers *et al.*, *Astron. J.* **165**, 50 (2023).
[40] C. Hahn *et al.*, *Astron. J.* **165**, 253 (2023).
[41] R. Zhou *et al.*, *Astron. J.* **165**, 58 (2023).
[42] A. Raichoor *et al.*, *Astron. J.* **165**, 126 (2023).
[43] E. Chaussidon *et al.*, *Astrophys. J.* **944**, 107 (2023).
[44] T. N. Miller *et al.* (DESI Collaboration), *Astron. J.* **168**, 95 (2024).
[45] A. G. Adame *et al.* (DESI Collaboration), *J. Cosmol. Astropart. Phys.* **09** (2025) 008.
[46] A. Dey *et al.*, *Astron. J.* **157**, 168 (2019).
[47] H. Zou, X. Zhou, X. Fan, T. Zhang, Z. Zhou, J. Nie, X. Peng, I. McGreer, L. Jiang, A. Dey, D. Fan, B. He, Z. Jiang, D. Lang, M. Lesser, J. Ma, S. Mao, D. Schlegel, and J. Wang, *Publ. Astron. Soc. Pac.* **129**, 064101 (2017).
[48] B. Flaugher *et al.* (DES Collaboration), *Astron. J.* **150**, 150 (2015).
[49] The Dark Energy Survey Collaboration, arXiv:astro-ph/0510346.
[50] E. L. Wright *et al.*, *Astron. J.* **140**, 1868 (2010).
[51] D. Lang, *Astron. J.* **147**, 108 (2014).
[52] J. Lasker *et al.* (DESI Collaboration), *J. Cosmol. Astropart. Phys.* **01** (2025) 127.
[53] A. G. Adame *et al.* (DESI Collaboration), *J. Cosmol. Astropart. Phys.* **07** (2025) 017.

- [54] C.-H. Chuang, F.-S. Kitaura, F. Prada, C. Zhao, and G. Yepes, *Mon. Not. R. Astron. Soc.* **446**, 2621 (2015).
- [55] C. Zhao *et al.*, *Mon. Not. R. Astron. Soc.* **503**, 1149 (2021).
- [56] L. H. Garrison, D. J. Eisenstein, and P. A. Pinto, *Mon. Not. R. Astron. Soc.* **485**, 3370 (2019).
- [57] N. A. Maksimova, L. H. Garrison, D. J. Eisenstein, B. Hadzhiyska, S. Bose, and T. P. Satterthwaite, *Mon. Not. R. Astron. Soc.* **508**, 4017 (2021).
- [58] A. G. Adame *et al.* (DESI Collaboration), *Astron. J.* **168**, 58 (2024).
- [59] S. Yuan *et al.*, *Mon. Not. R. Astron. Soc.* **530**, 947 (2024).
- [60] For the Abacus-Complete mocks, we additionally down-sample it to match the number density of the altMTL mocks, so that their statistical fluctuations are comparable as those in the real data.
- [61] R. N. Cahn and Z. Slepian, *J. Phys. A* **56**, 325204 (2023).
- [62] In the large- N_g limit, with typical number of galaxies of modern galaxy surveys reaches $\mathcal{O}(10^6)$, the algorithm scales linearly with the number of objects. This is because the computational cost is dominated by assembling all galaxy quartets for each primary galaxy, whereas the computation of the harmonic expansion coefficients is sub-dominant (see also [63]).
- [63] O. H. E. Philcox, Z. Slepian, J. Hou, C. Warner, R. N. Cahn, and D. J. Eisenstein, *Mon. Not. R. Astron. Soc.* **509**, 2457 (2021).
- [64] Z. Slepian and D. J. Eisenstein, *Mon. Not. R. Astron. Soc.* **454**, 4142 (2015).
- [65] We use the GPU version of the public code <https://github.com/oliverphilcox/encore> with extensions including ℓ_{\max} for the 4PCF measurement.
- [66] H. A. Feldman, N. Kaiser, and J. A. Peacock, *Astrophys. J.* **426**, 23 (1994).
- [67] J. Hou, R. N. Cahn, O. H. E. Philcox, and Z. Slepian, *Phys. Rev. D* **106**, 043515 (2022).
- [68] Here we apply the Einstein summation notation.
- [69] A. Krolewski, S. May, K. Smith, and H. Hopkins, *J. Cosmol. Astropart. Phys.* 08 (2024) 044.
- [70] E. B. Wilson and M. M. Hilferty, *Proc. Natl. Acad. Sci. U.S.A.* **17**, 684 (1931).
- [71] J. Hartlap, P. Simon, and P. Schneider, *Astron. Astrophys.* **464**, 399 (2007).
- [72] E. Sellentin and A. F. Heavens, *Mon. Not. R. Astron. Soc.* **456**, L132 (2016).
- [73] W. Press, *Numerical Recipes 3rd Edition: The Art of Scientific Computing*, Numerical Recipes: The Art of Scientific Computing (Cambridge University Press, Cambridge, England, 2007).
- [74] D. Forero-Sánchez *et al.*, *J. Cosmol. Astropart. Phys.* 04 (2025) 055.
- [75] D. Bianchi *et al.*, *J. Cosmol. Astropart. Phys.* 04 (2025) 074.
- [76] E. Chaussidon, C. Yèche, A. de Mattia, C. Payerne, P. McDonald *et al.*, *J. Cosmol. Astropart. Phys.* 06 (2025) 029.
- [77] <https://www.desi.lbl.gov/collaborating-institutions>.
- [78] [10.5281/zenodo.16729072](https://zenodo.org/record/16729072).