

Clustering with general photo- z uncertainties: Application to Baryon Acoustic Oscillations

Kwan Chuen Chan^{1,2*}, Ismael Ferrero^{3,4,5}, Santiago Avila^{6,7}, Ashley J. Ross⁸,
Martin Crocce^{4,5}, Enrique Gaztañaga^{4,5}

¹ School of Physics and Astronomy, Sun Yat-sen University, 2 Daxue Road, Tangjia, Zhuhai, 519082, China

² CSST Science Center for the Guangdong-Hongkong-Macau Greater Bay Area, SYSU, China

³ Institute of Theoretical Astrophysics, University of Oslo. P.O. Box 1029 Blindern, NO-0315 Oslo, Norway

⁴ Institut d'Estudis Espacials de Catalunya (IEEC), 08034 Barcelona, Spain

⁵ Institute of Space Sciences (ICE, CSIC), Campus UAB, Carrer de Can Magrans, s/n, 08193 Barcelona, Spain

⁶ Instituto de Física Teórica UAM/CSIC, Universidad Autónoma de Madrid, 28049 Madrid, Spain

⁷ Departamento de Física Teórica, Universidad Autónoma de Madrid, 28049 Madrid, Spain

⁸ Center for Cosmology and Astro-Particle Physics, The Ohio State University, Columbus, OH 43210, USA

Accepted XXX. Received YYY; in original form ZZZ

ABSTRACT

Photometric data can be analyzed using the three-dimensional correlation function ξ_p to extract cosmological information via e.g., measurement of the Baryon Acoustic Oscillations (BAO). Previous studies modeled ξ_p assuming a Gaussian photo- z approximation. In this work we improve the modeling by incorporating realistic photo- z distribution. We show that the position of the BAO scale in ξ_p is determined by the photo- z distribution and the Jacobian of the transformation. The latter diverges at the transverse scale of the separation s_\perp , and it explains why ξ_p traces the underlying correlation function at s_\perp , rather than s , when the photo- z uncertainty $\sigma_z/(1+z) \gtrsim 0.02$. We also obtain the Gaussian covariance for ξ_p . Due to photo- z mixing, the covariance of ξ_p shows strong off-diagonal elements. The high correlation of the data causes some issues to the data fitting. Nonetheless, we find that either it can be solved by suppressing the largest eigenvalues of the covariance or it is not directly related to the BAO. We test our BAO fitting pipeline using a set of mock catalogs. The data set is dedicated for Dark Energy Survey Year 3 (DES Y3) BAO analyses and includes realistic photo- z distributions. The theory template is in good agreement with mock measurement. Based on the DES Y3 mocks, ξ_p statistic is forecast to constrain the BAO shift parameter α to be 1.001 ± 0.023 , which is well consistent with the corresponding constraint derived from the angular correlation function measurements. Thus ξ_p offers a competitive alternative for the photometric data analyses.

Key words: cosmology: observations - (cosmology:) large-scale structure of Universe

1 INTRODUCTION

Imaging surveys infer the redshift of galaxy samples, photo- z , using broadband filters. Ongoing and future large-scale structure surveys that collect enormous amount of photo- z data include Kilo-Degree Survey (KiDS)¹, Dark Energy Survey (DES)², Hyper Suprime-Cam (HSC)³, Rubin Ob-

servatory's Legacy Survey of Space and Time (LSST)⁴, Euclid⁵, and the Chinese Survey Space Telescope (CSST)⁶. These surveys typically use several bandpasses, e.g. DES has *grizY* and the accuracy of the resultant photo- z is about $\sigma \sim 0.03(1+z)$ [see e.g. Crocce et al. (2019); Carnero Rosell et al. (2021)], but it can be improved if more stringent criteria are applied (Roza et al. 2016)]. Although photometric

* E-mail: chankc@mail.sysu.edu.cn (KCC)

¹ <http://kids.strw.leidenuniv.nl>

² <https://www.darkenergysurvey.org>

³ <https://www.naoj.org/Projects/HSC>

⁴ <https://www.lsst.org>

⁵ <https://www.euclid-ec.org>

⁶ <http://www.nao.cas.cn/csst>

surveys yield less precise redshift than spectroscopic ones ($\sigma \sim 10^{-3}$), they are more efficient to collect a large volume of survey data with deep magnitude.

Baryonic Acoustic Oscillations (BAO) (Peebles & Yu 1970; Sunyaev & Zeldovich 1970) are the primordial acoustic features imprinted in the distribution of the large-scale structure. In the early universe, photons and baryons form a tightly coupled plasma, and acoustic oscillations are excited in it. After the recombination, hydrogen atoms form, the acoustic waves get stalled because there is no medium for propagation. The acoustic patterns are preserved in the distribution of galaxies and they correspond to the sound horizon at the drag epoch, which is about 150 Mpc. Formation of the BAO is governed by well-understood linear physics, see Bond & Efstathiou (1984, 1987); Hu & Sugiyama (1996); Hu et al. (1997); Dodelson (2003) for the details of the cosmic microwave background physics. Given its robustness, BAO has been regarded as a standard ruler in cosmology, see e.g. Weinberg et al. (2013); Aubourg et al. (2015). BAO has been measured in numerous spectroscopic data analyses. It was first clearly detected by Eisenstein et al. (2005); Cole et al. (2005) and followed up in many subsequent studies, e.g. Gaztanaga et al. (2009); Percival et al. (2010); Beutler et al. (2011); Blake et al. (2012); Anderson et al. (2012); Kazin et al. (2014); Ross et al. (2015); Alam et al. (2017); Alam et al. (2020).

The sample size in terms of volume and depth from photometric surveys may compensate its imprecision in redshift information and give competitive measurements (Seo & Eisenstein 2003; Blake & Bridle 2005; Amendola et al. 2005; Benitez et al. 2009; Zhan et al. 2009; Chaves-Montero et al. 2018). Indeed, there are several detections of the BAO feature in the galaxy distribution using photometric data from SDSS (Padmanabhan et al. 2007; Estrada et al. 2009; Hütsi 2010; Seo et al. 2012; Carnero et al. 2012; de Simoni et al. 2013), DES [Y1 (Abbott et al. 2019) and Y3 (Abbott et al. 2021), hereafter DES Y3], and DECaLS (Sridhar et al. 2020).

Because of the photo- z uncertainty, it is natural to perform the analyses of the photometric data using angular statistics such as the angular correlation function (Coupon et al. 2012; Crocce et al. 2016; Abbott et al. 2019; Elvin-Poole et al. 2018; Vakili et al. 2020) or the angular power spectrum (Padmanabhan et al. 2007; Seo et al. 2012; Camacho et al. 2019; Nicola et al. 2020). In particular, these two statistics were adopted as the fiducial statistical tools to measure BAO in DES Y3 analyses. However, given enormous amount photo- z data available, it is worthwhile exploring new methods to enhance information extraction, especially from the BAO. Ross et al. (2017) proposed to measure the BAO using the three-dimensional (3D) correlation function characterized by the transverse scale of the separation. This statistic has been applied to photo- z data to get interesting BAO measurements (Abbott et al. 2019; Sridhar et al. 2020). The photo- z uncertainty of the sample must be folded in the theory template. To simplify the modeling, the photo- z distribution was assumed to be Gaussian in Ross et al. (2017). The actual photo- z distribution has non-negligible deviation from Gaussian distribution, e.g. Crocce et al. (2019); Zhou et al. (2021); Carnero Rosell et al. (2021). This assumption may risk introducing bias, and so this statistic was not adopted in the fiducial DES Y1 and Y3 analyses.

In this work, we improve the modeling of the 3D clustering by incorporating general photo- z distribution and present a Gaussian covariance for it.

This paper is organized as follows. In Sec. 2, we first review the computation of the angular correlation and then show that the general angular correlation can be turned into the 3D correlation function ξ_p . We consider the Gaussian photo- z limit and use it to understand the effect of photo- z uncertainties on ξ_p in Sec. 3. In Sec. 4, we derive the Gaussian covariance for ξ_p using the Gaussian covariance for the angular correlation function. In Sec. 5.1, we present the numerical results on the template and covariance, and compare them with mock results whenever possible. We discuss the issue in data fitting caused by highly correlated data in Sec. 5.2.1, and then present the BAO fit results in Sec. 5.2.2. We conclude in Sec. 6. In Appendix A, we test our pipeline for the case of spectroscopic data. The default cosmology adopted in this work is the MICE cosmology (Fosalba et al. 2015; Crocce et al. 2015), which is a flat Λ CDM with $\Omega_m = 0.25$, $\Omega_\Lambda = 0.75$, $h = 0.7$, and $\sigma_8 = 0.8$.

2 ANGULAR AND THREE-DIMENSIONAL TWO-POINT CORRELATION

In this section, we first review the computation of the angular correlation function, w , in the presence of general photo- z uncertainty. We then use the general angular correlation function to derive the 3D two-point correlation ξ_p .

2.1 Angular correlation function w_p

Because in imaging surveys, angles are well determined but there are typically significant redshift uncertainties, it is natural to use angular quantities. It is useful to first clarify the relation between volume number density and tomographic bin angular number density.

Suppose that there are dN objects in a volume element parametrized by the redshift interval dz and the solid angle $d\Omega$, the volume number density $n_{(3)}$ is given by

$$n_{(3)}(\mathbf{r}) = \frac{dN}{dzd\Omega}. \quad (1)$$

The tomographic angular number density $n_{(2)}$ is defined as

$$n_{(2)}(z, \hat{\mathbf{r}}) = \frac{dN}{d\Omega} = n_{(3)}(\mathbf{r})dz. \quad (2)$$

We can go on to define the volume density contrast $\delta_{(3)}$ and angular density contrast $\delta_{(2)}$

$$\delta_{(3)} = \frac{n_{(3)} - \bar{n}_{(3)}}{\bar{n}_{(3)}}, \quad (3)$$

$$\delta_{(2)} = \frac{n_{(2)} - \bar{n}_{(2)}}{\bar{n}_{(2)}}, \quad (4)$$

where $\bar{n}_{(3)}$ and $\bar{n}_{(2)}$ are their respective mean values. From Eq. (2), it is clear that $\delta_{(3)} = \delta_{(2)}$.

We start with the effect of the photo- z on the galaxy number density:

$$n_p(z_p, \hat{\mathbf{r}}) = \int dz f(z|z_p) n_{(3)}(z, \hat{\mathbf{r}}), \quad (5)$$

where $f(z|z_p)dz$ is the probability that the true redshift of the galaxy falls within $[z, z + dz]$ given that its photo- z is

z_p , and n_p denotes the number density in photo- z space. For photometric data, $f(z|z_p)$ is of central importance and it can be estimated by the photo- z codes using template fitting or machine learning methods [see e.g. Benitez (2000); Ilbert et al. (2006); Sadeh et al. (2016); De Vicente et al. (2016)].

At the background level, we have

$$\bar{n}_p(z_p) = \int dz f(z|z_p) \bar{n}_{(3)}(z), \quad (6)$$

where \bar{n}_p and $\bar{n}_{(3)}$ are the angular mean number density in the photo- z and true redshift space, respectively. The perturbation is given by

$$\begin{aligned} n_p(z_p, \hat{\mathbf{r}}) &= \bar{n}_p(z_p) [1 + \delta_p(z_p, \hat{\mathbf{r}})] \\ &= \int dz f(z|z_p) \bar{n}_{(3)}(z) [1 + \delta_{(3)}(z, \hat{\mathbf{r}})], \end{aligned} \quad (7)$$

and we get

$$\delta_p(z_p, \hat{\mathbf{r}}) = \int dz f(z|z_p) \frac{\bar{n}_{(3)}(z)}{\bar{n}_p(z_p)} \delta_{(3)}(z, \hat{\mathbf{r}}). \quad (8)$$

Therefore, the appropriate true redshift distribution for the density contrast is

$$\phi(z|z_p) = f(z|z_p) \frac{\bar{n}_{(3)}(z)}{\bar{n}_p(z_p)}. \quad (9)$$

It is easy to see that $\phi(z|z_p)$ indeed furnishes a probability density. Here we assume that there are sufficiently many particles in the coarse-grained scale of interest so that the photo- z effect can be modeled by the distribution ϕ . In this regime, the effect of the photo- z is to average $\delta_{(3)}$ over z .

Using Eq. (9), the photo- z angular correlation function w_p can be written as

$$\begin{aligned} w(\theta, z_p, z'_p) &\equiv \langle \delta_p(z_p, \hat{\mathbf{r}}) \delta_p(z'_p, \hat{\mathbf{r}}') \rangle \\ &= \int dz \phi(z|z_p) \int dz' \phi(z'|z'_p) \langle \delta_{(3)}(z, \hat{\mathbf{r}}) \delta_{(3)}(z', \hat{\mathbf{r}}') \rangle, \end{aligned} \quad (10)$$

with $\theta \equiv \cos^{-1}(\hat{\mathbf{r}} \cdot \hat{\mathbf{r}}')$. By expressing $\delta_{(3)}$ in terms of its Fourier modes, we have

$$\begin{aligned} w(\theta, z_p, z'_p) &= \int dz \phi(z|z_p) \int dz' \phi(z'|z'_p) \int \frac{d^3 k}{(2\pi)^3} \\ &\quad \times e^{i\mathbf{k} \cdot [\mathbf{r}(z) - \mathbf{r}'(z')]} P(\mathbf{k}, z, z'). \end{aligned} \quad (11)$$

Utilizing the variable $\mathbf{s} = \mathbf{r} - \mathbf{r}'$ and applying the Rayleigh expansion of the plane wave

$$e^{i\mathbf{k} \cdot \mathbf{s}} = 4\pi \sum_{l=0}^{\infty} \sum_{m=-l}^l i^l j_l(ks) Y_{lm}(\hat{\mathbf{k}}) Y_{lm}^*(\hat{\mathbf{s}}), \quad (12)$$

we have

$$\begin{aligned} w(\theta, z_p, z'_p) &= \int dz \phi(z|z_p) \int dz' \phi(z'|z'_p) \int \frac{dk k^2}{(2\pi)^3} \\ &\quad \times \sum_{l,m} 4\pi i^l j_l(ks) Y_{lm}(\hat{\mathbf{s}}) \int d\hat{\mathbf{k}} Y_{lm}^*(\hat{\mathbf{k}}) P(\mathbf{k}, z, z'). \end{aligned} \quad (13)$$

Note that we have only one integral involving the spherical Bessel function because we expand the plane wave for \mathbf{s} .

Eq. (13) accounts for the photo- z uncertainty exactly and it is valid for the curved sky. We only need to plug in the appropriate power spectrum $P(\mathbf{k}, z, z')$. However, the redshift-space power spectrum in the distant observer limit

is often used, and so this effectively limits the results to the flat sky case. As long as the local line of sight (LOS) is used, the effect is mild, e.g. Yoo & Seljak (2015).

In general, because of the azimuthal symmetry about the LOS direction $\hat{\mathbf{e}}$, the power spectrum can be expanded about the LOS direction as

$$P(\mathbf{k}, z, z') = \sum_{\text{even } \ell} P_{\ell}(k, z, z') \mathcal{L}_{\ell}(\hat{\mathbf{k}} \cdot \hat{\mathbf{e}}), \quad (14)$$

where \mathcal{L}_{ℓ} is the Legendre polynomial and P_{ℓ} is the multipole coefficient. Parity invariance further restricts the sum to even ℓ only. Then the angular correlation can be written as

$$\begin{aligned} w(\theta, z_p, z'_p) &= \sum_{\ell} i^{\ell} \int dz \phi(z|z_p) \int dz' \phi(z'|z'_p) \mathcal{L}_{\ell}(\hat{\mathbf{s}} \cdot \hat{\mathbf{e}}) \\ &\quad \times \int \frac{dk k^2}{2\pi^2} j_{\ell}(ks) P_{\ell}(k, z, z'), \end{aligned} \quad (15)$$

where j_{ℓ} is the spherical Bessel function. To speed up the computation, the j_{ℓ} integrals are evaluated using FFTLog (Hamilton 2000).

For the linear redshift-space power spectrum with linear bias (Kaiser 1987)

$$\begin{aligned} P(\mathbf{k}, z, z') &= [b + f(\hat{\mathbf{k}} \cdot \hat{\mathbf{e}})^2] [b' + f'(\hat{\mathbf{k}} \cdot \hat{\mathbf{e}})^2] \\ &\quad \times D(z) D(z') P_m(k, 0, 0), \end{aligned} \quad (16)$$

where b is the linear bias, f is $d \ln D / d \ln a$ with D being the linear growth factor, and $P_m(k, 0, 0)$ is the matter power spectrum evaluated at $z = 0$, we have (Hamilton 1992; Cole et al. 1994; Crocce et al. 2011)

$$\begin{aligned} w(\theta, z_p, z'_p) &= \int dz \phi(z|z_p) D(z) \int dz' \phi(z'|z'_p) D(z') \\ &\quad \times \sum_{\ell=0,2,4} i^{\ell} A_{\ell} \mathcal{L}_{\ell}(\hat{\mathbf{s}} \cdot \hat{\mathbf{e}}) \int \frac{dk k^2}{2\pi^2} j_{\ell}(ks) P_m(k, 0, 0), \end{aligned} \quad (17)$$

with

$$A_{\ell}(z_p, z'_p) = \begin{cases} bb' + \frac{1}{3}(bf' + b'f) + \frac{1}{5}ff' & \text{for } \ell = 0, \\ \frac{2}{3}(bf' + b'f) + \frac{4}{7}ff' & \text{for } \ell = 2, \\ \frac{8}{35}ff' & \text{for } \ell = 4. \end{cases} \quad (18)$$

In DES Y3, an anisotropic BAO damping factor is applied to the BAO feature, i.e.

$$\begin{aligned} P(\mathbf{k}, z, z') &= [b + f(\hat{\mathbf{k}} \cdot \hat{\mathbf{e}})^2] [b' + f'(\hat{\mathbf{k}} \cdot \hat{\mathbf{e}})^2] D(z) D(z') \\ &\quad \times [(P_{\text{lin}} - P_{\text{nw}}) e^{-k^2 \Sigma^2(\mu)} + P_{\text{nw}}], \end{aligned} \quad (19)$$

where P_{lin} and P_{nw} denotes the linear power spectrum with and without BAO wiggles respectively. Because Σ is a function of μ , we can no longer express the first three multipoles analytically, instead we compute them numerically. As our primary goal here is to model the BAO feature, we shall use Eq. (19). But we will use Eqs. (17) and (18) to measure the bias parameters.

2.2 From w to ξ_p

Because of the statistical rotational invariance, the angular correlation function $w(\theta, z_{\text{eff}})$ for the sample at the effective redshift z_{eff} can be parametrized by the angle θ subtended by two points on the sky at the observer. In addition, the

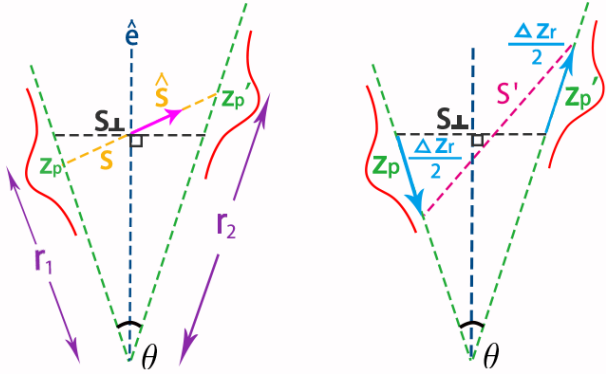


Figure 1. Left panel: The relation between the angular correlation function specified by two true redshift distributions conditional on photo- z z_p and z'_p , subtending an angle θ at the observer. The corresponding 3D correlation function is described by the separation vector \mathbf{s} relative to the line of sight direction $\hat{\mathbf{e}}$. Right panel: The integral over the two distributions can be approximated by a 1D integral over Δz_r or the separation s' .

redshifts (or photo- z) of these two points are also specified, e.g. in the form of tomographic bins, $w(\theta, z, z')$ furnishes a representation of the 3D correlation function. On the other hand, the 3D correlation function ξ in real space depends only on the distance between the two points, $s = |\mathbf{r}_1 - \mathbf{r}_2|$, thanks to the statistical translational and rotational invariance. The redshift space distortion along the LOS direction breaks the rotational invariance in redshift space and ξ is usually parametrized as a function of s and μ , the latter is the dot product between direction of the separation, $\hat{\mathbf{s}}$ and the LOS $\hat{\mathbf{e}}$, or in terms of s_\perp and s_\parallel , the component of \mathbf{s} perpendicular and parallel to the LOS. In practice, both w and ξ can be computed by assigning density to a grid (2D pixelized grid cells for w and 3D grid cells for ξ) or by comparing pair counts between the data catalog and the random one.

Because $\delta_{(2)} = \delta_{(3)}$, it is clear that the general cross angular correlation function and the 3D correlation are equal provided a cosmology is assumed to convert angle and redshift to distance and the redshift bin is fine enough so that the projection effect along the LOS direction is negligible. Depending on how it is parametrized, the correlation function goes by different names. We shall consider the 3D correlation function $\xi_p(s, \mu) = \langle \delta_p(\mathbf{r}_1) \delta_p(\mathbf{r}_2) \rangle$ and its projective version characterized by the transverse scale, $\xi_p(s_\perp, \mu) \equiv \langle \delta_p(\mathbf{r}_1) \delta_p(\mathbf{r}_2) \rangle_\parallel$, where the notation $\langle \dots \rangle_\parallel$ means that we average over $s_\parallel \equiv s\mu$ by stacking together the pairs with the same $s_\perp \equiv s\sqrt{1 - \mu^2}$.

We obtain $\xi_p(s, \mu)$ by mapping from w . We first evaluate the general cross correlation between redshift bin i and j separated by angle θ_k bin, $w_{ij}(\theta_k)$. From its redshift bins and angular separation, s and μ for this pair can be obtained. This allows us to allocate this angular correlation into the appropriate s and μ bins. The relation between w and ξ_p is schematically shown on the left panel of Fig. 1. Because ξ_p is measured by averaging over all possible grid cell pairs (or data pairs in the language of pair counting) in the survey data region, we loop over all the $w_{ij}(\theta_k)$ pairs that fit into

the data region and give rise to the same s and μ bins. Explicitly, we have

$$\xi_p(s, \mu) = \frac{\sum_{ijk} f_{ijk} w_{ij}(\theta_k; z_i, z_j)}{\sum_{ijk} f_{ijk}}, \quad (20)$$

where f_{ijk} denotes the weight for all the cross bin pairs $w_{ij}(\theta_k)$ falling into the s and μ bins. To approximately mimic the pair counting in the computation of ξ_p , we shall use the weight $f_{ijk} = N_i N_j$ with N_i (N_j) being the number of particles in bin i (j). More accurate weighting can be achieved by counting pairs as in the case of random catalog, but it would be much more computationally expensive.

Here we describe binning into s and μ bin as an example. For $\xi_p(s_\perp, \mu)$, the procedures are the same except that we compute s_\perp from the redshifts of the bin pairs and their angular separation, and then bin the correlation based on s_\perp and μ .

The measurement of the ξ_p statistics requires a fiducial cosmology input, while the tomographic bin correlation in term of angle and redshift can evade the cosmology dependence (Montanari & Durrer 2012). However, the advantage of the ξ_p statistic is that it can effectively summarize the information in multiple tomographic bins in a single correlation measurement, and it is easier to borrow the techniques developed for the spectroscopic data analysis.

We turn to comment on the definition of the LOS direction associated with a pair of galaxies on the sky. For the angular correlation function, a natural definition of the LOS direction is the angle bisector of the angle subtended by the pairs at the observer. In this ‘‘angle-centric’’ definition, the dot product between LOS direction $\hat{\mathbf{e}}_\theta$ and the unit separation vector $\hat{\mathbf{s}}$ can be expressed as

$$\hat{\mathbf{e}}_\theta \cdot \hat{\mathbf{s}} = \frac{r_2 - r_1}{2s} \sqrt{\frac{(r_1 + r_2 + s)(r_1 + r_2 - s)}{r_1 r_2}}, \quad (21)$$

where r_1 and r_2 are the distances to the galaxy pair respectively, and s is the length of the separation vector. On the other hand, it is natural to align the LOS with the midpoint of the separation for the 3D correlation function. For this ‘‘separation-centric’’ viewpoint, the dot product between LOS direction $\hat{\mathbf{e}}_s$ and $\hat{\mathbf{s}}$ can be written as

$$\hat{\mathbf{e}}_s \cdot \hat{\mathbf{s}} = \frac{(r_2 - r_1)(r_2 + r_1)}{s\sqrt{2(r_1^2 + r_2^2) - s^2}}. \quad (22)$$

We have checked that these two definitions are essentially identical for our purpose. For example, at $z_{\text{eff}} = 0.8$ and $s = 100 \text{ Mpc } h^{-1}$, the agreement between these two definitions are better than 10^{-3} . Thus we can use them interchangeably.

3 UNDERSTANDING THE IMPACT OF PHOTO-Z BY A GAUSSIAN CASE STUDY

Ross et al. (2017) used a simple Gaussian setup to study the impact of photo- z uncertainty on the BAO information. The apparatus can be constructed by the following arguments. Suppose that the position vectors of two points are \mathbf{r}_1 and \mathbf{r}_2 and due to photo- z error, they have additional displacement $\Delta\mathbf{r}_1$ and $\Delta\mathbf{r}_2$. The total separation vector is

$$\begin{aligned} \mathbf{s}_p &= (\mathbf{r}_2 - \mathbf{r}_1) + (\Delta\mathbf{r}_2 - \Delta\mathbf{r}_1) \\ &= \mathbf{s}_{\parallel} + \Delta\mathbf{r}_2 - \Delta\mathbf{r}_1 + \mathbf{s}_{\perp}, \end{aligned} \quad (23)$$

where $\mathbf{s}_{t\perp}$ and $\mathbf{s}_{t\parallel}$ denote the perpendicular and parallel components of $\mathbf{r}_2 - \mathbf{r}_1$ w.r.t. the LOS direction without photo- z error. The photo- z error only affects the direction parallel to the LOS in the plane-parallel limit. If we assume that $\Delta\mathbf{r}_1$ and $\Delta\mathbf{r}_2$ are independent Gaussian random variables with zero mean and constant variance σ^2 , then $\Delta\mathbf{r}_2 - \Delta\mathbf{r}_1$ is Gaussian distributed with zero mean and variance $2\sigma^2$. Thus we can model the effects of the Gaussian photo- z error on the 3D correlation function by a 1D Gaussian distribution.

It is instructive to check under what conditions our general expression would reduce to this result. We will see that this exercise gives valuable insight into our results.

We now assume that the conditional true redshift distributions in Eq. (10) are Gaussian, namely

$$\phi(z|z_p) = P_G(z - z_p; 0, \sigma^2) \equiv \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(z-z_p)^2}{2\sigma^2}}. \quad (24)$$

By adopting the variables

$$z = z_p + \Delta z, \quad z' = z'_p + \Delta z', \quad (25)$$

w_p can be written as

$$w_p(\theta, z_p, z'_p) = \frac{1}{2\pi\sigma^2} \int d\Delta z e^{-\frac{\Delta z^2}{2\sigma^2}} \int d\Delta z' e^{-\frac{\Delta z'^2}{2\sigma^2}} \times w(\theta, z_p + \Delta z, z'_p + \Delta z'). \quad (26)$$

After further changing to the ‘‘Center-of-Mass’’ variables

$$\Delta z_c = \frac{\Delta z + \Delta z'}{2}, \quad \Delta z_r = \Delta z - \Delta z', \quad (27)$$

we have

$$w_p(\theta, z_p, z'_p) = \frac{1}{2\pi\sigma^2} \int d\Delta z_c \int d\Delta z_r \exp\left(-\frac{\Delta z_c^2}{\sigma^2} - \frac{\Delta z_r^2}{4\sigma^2}\right) \times w\left(\theta, z_p + \Delta z_c + \frac{\Delta z_r}{2}, z'_p + \Delta z_c - \frac{\Delta z_r}{2}\right). \quad (28)$$

To proceed, we neglect the Δz_c dependence in w as it gives an overall shift of the pair in redshift⁷ and we arrive at

$$w_p(\theta, z_p, z'_p) \approx \int d\Delta z_r \frac{1}{\sqrt{4\pi\sigma^2}} \exp\left(-\frac{\Delta z_r^2}{4\sigma^2}\right) \times w\left(\theta, z_p + \frac{\Delta z_r}{2}, z'_p - \frac{\Delta z_r}{2}\right). \quad (29)$$

By mapping to s and μ , we can get the corresponding equation for ξ_p , i.e.

$$\xi_p(s(\theta, z_p, z'_p), \mu(\theta, z_p, z'_p)) \approx \int d\Delta z_r P_G(\Delta z_r; 0, 2\sigma^2) \times \xi\left(s'\left(\theta, z_p + \frac{\Delta z_r}{2}, z'_p - \frac{\Delta z_r}{2}\right), \mu'\left(\theta, z_p + \frac{\Delta z_r}{2}, z'_p - \frac{\Delta z_r}{2}\right)\right). \quad (30)$$

This equation invites the interpretation on the right panel of Fig. 1.

Alternatively we can use the separation distance s' (shown on the right panel of Fig. 1) as the integration variable

$$\xi_p(s(\theta, z_p, z'_p), \mu(\theta, z_p, z'_p)) \approx \int ds' \left(\frac{ds'}{d\Delta z_r}\right)^{-1} P_G(\Delta z_r(s'); 0, 2\sigma^2) \xi(s', \mu'). \quad (31)$$

⁷ The standard 3D correlation analysis neglects the redshift evolution within the redshift bin.

We can think of s' as a function of Δz_r . When Δz_r increases from the negative end, s' first decreases and it reaches the minimum value $s' = s_\perp$ before starting to increase. The scale $s' = s_\perp$ divides the integral into two parts, and they should be evaluated separately.

The inverse of the Jacobian is given by

$$\frac{ds'}{d\Delta z_r} = \frac{c}{2s'} \left[\frac{r(z_2)}{H(z_2)} - \frac{r(z_1)}{H(z_1)} - \cos\theta \left(\frac{r(z_1)}{H(z_2)} - \frac{r(z_2)}{H(z_1)} \right) \right], \quad (32)$$

where $r(z)$ and $H(z)$ respectively represent the comoving distance and Hubble rate evaluated at redshift z , c is the light speed, and z_1 (z_2) denotes $z_p - \frac{\Delta z_r}{2}$ ($z'_p + \frac{\Delta z_r}{2}$). It is important to note that this derivative vanishes when $z_1 = z_2$, which corresponds to $s' = s_\perp$.

Eq. (30) reveals that the photo- z correlation function is a projection of the underlying true correlation with a Gaussian weight of zero mean and variance $2\sigma^2$, in agreement with Eq. (6) of Ross et al. (2017). In Fig. 2, we show the wedge correlation function obtained with the Gaussian photo- z distribution. The effective redshift of the two conditional true redshift distributions centered at z_p and z'_p respectively, is taken to be $z_{\text{eff}} = 0.8$. The standard deviation of the distribution is assumed to be $\sigma = (1 + z_{\text{eff}})\sigma_z$ and a suite of σ_z values are considered: 0.001, 0.01, 0.02, and 0.04. In this plot, for simplicity the linear real-space correlation is used. We have compared the results computed with the approximation Eq. (31) and by directly integrating over the 2D integral. We find that the 1D approximation is in good agreement with the exact result, but deviations become apparent for high σ_z .

Interestingly, while for small σ_z such as 0.001, the BAO bump peaks at the true sound horizon size, s_{sh} , it shifts to larger values as σ_z increases. Also for large σ_z , it is clear that the BAO bump moves to larger s as μ increases. In the lower panels of Fig. 2, we show ξ_p as a function of the transverse scale $s_\perp = s\sqrt{1 - \mu^2}$. As observed in Ross et al. (2017), when ξ_p is plotted against s_\perp , the BAO position appears to line up at $s_\perp = s_{\text{sh}}$ for $\sigma_z \gtrsim 0.02$. Thus this plot suggests that the BAO peak in the upper panel (the usual way to plot ξ) is directly related to the scale s_\perp for $\sigma_z \gtrsim 0.02$.

To shed light on these results, we show in Fig. 3 the three factors of the integrand of Eq. (31): the absolute value of the Jacobian $|d\Delta z_r/ds'|$, the Gaussian weight P_G and the underlying correlation function ξ . We shall shortly explain these terms in details. In this illustration, we take $s(z_p, z'_p) = 100 \text{ Mpc } h^{-1}$.

For small σ_z such as 0.001, which is close to the spectroscopic redshift case, the Gaussian weight is narrowly distributed about the scale $s(z_p, z'_p)$. The distribution P_G as a function of s' is in fact mildly tilted for small μ and it becomes more symmetric when μ increases. As we mentioned, Eq. (32) vanishes at $s' = s_\perp$, and hence the Jacobian diverges. Except for small μ such as $\mu \lesssim 0.1$, the divergence occurs at the scale where P_G is essentially zero and hence it is irrelevant. Therefore, in this case the BAO scale in ξ_p is controlled by P_G , and it corresponds to the true sound horizon well.

As σ_z increases, the P_G distribution becomes more tilted. Among the cases shown, for $\sigma_z \geq 0.01$, s' is clearly not a monotonic function of Δz_r , as evident from the multi-

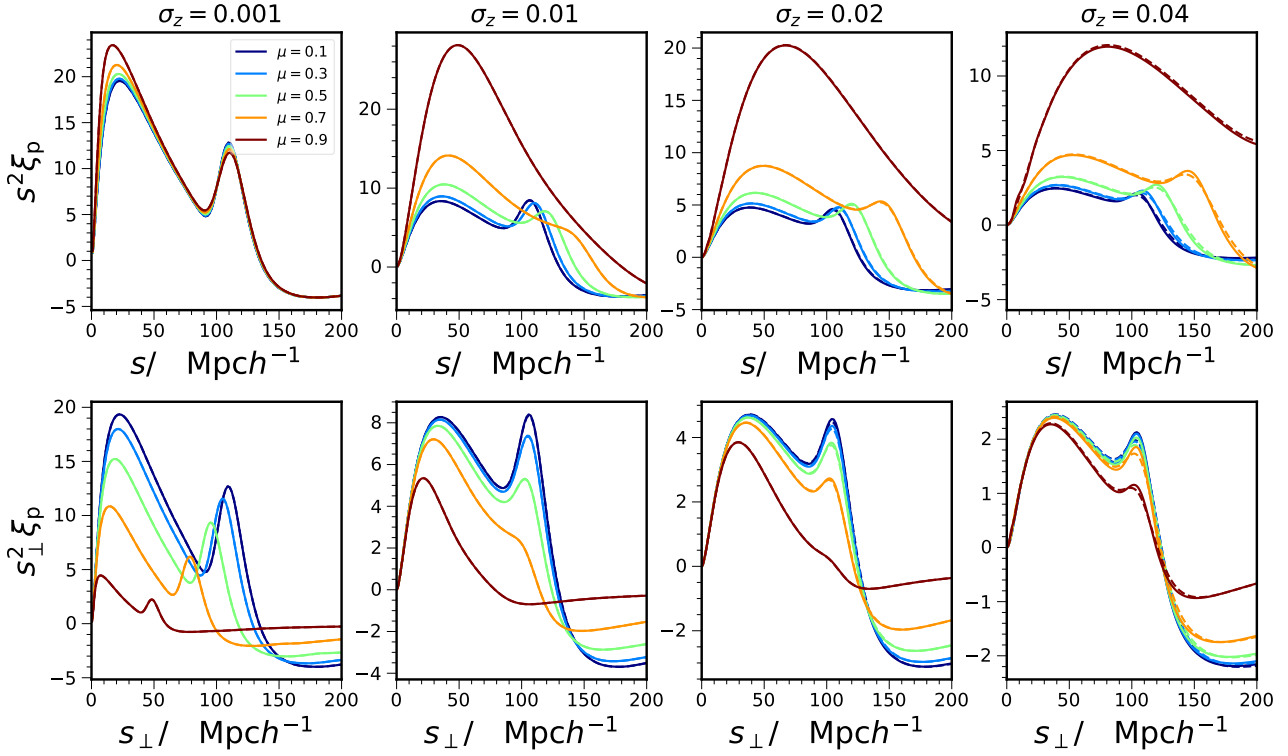


Figure 2. The wedge correlation function computed with Gaussian photo- z distributions. The upper and low panels show the results for using the separation s and the transverse scale s_{\perp} as the independent variables respectively. The 1D approximation Eq. (31) (solid) and the exact 2D integral results (dashed) agree with each other well, but small deviation occurs for large value of σ_z . Four values of σ_z are displayed (left to right panels). Different colors indicate wedge correlation of different values of μ .

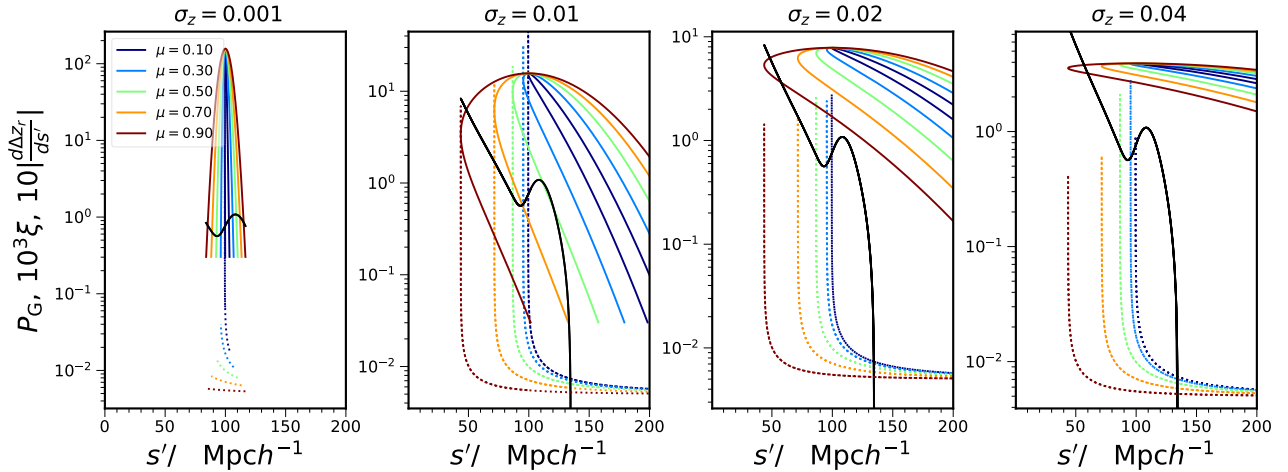


Figure 3. The factors in the photo- z projection integral [Eq. (31)] as a function of the separation s' . As shown on the right panel of Fig. 1, as Δz_r increases, s' first decreases until it reaches the minimum value $s' = s_{\perp}$ and then it increases. Both the Gaussian weight P_G (solid, color) and absolute value of the Jacobian $|d\Delta z_r/ds'|$ (multiplied by a factor of 10, dotted, color) correspond to the paths traversed as Δz_r increases. The point where all the P_G curves with different μ meet corresponds to $s' = s(z_p, z'_p)$ (chosen to be $100 \text{ Mpc } h^{-1}$ here). The Jacobian diverges at $s' = s_{\perp}$ and this indicates as the vertical line in the plot. The verticle part for the Jacobian touches the P_G curve as the latter turns at $s' = s_{\perp}$. The underlying correlation function (multiplied by a factor 10^3 , solid black) is also overplotted. See the text for further explanations.

true redshift distribution of the neighboring bins may overlap with each other for photometric galaxies, there is no cross photo- z (or cross-bin in tomographic analysis) shot noise contribution because shot noise arises from self-correlation and each galaxy is assigned one and only one photo- z .

Our method approximates the 3D correlation and the shot noise by the tomographic version. For the signal, when the tomographic bin size is small compared to photo- z uncertainty, the results are convergent. For the shot noise, we approximate the Dirac delta in redshift in Eq. (40) by a Kronecker delta and the volume density by the tomographic one.

5 MOCK TESTS

In this section, we first test the theory template and the Gaussian covariance for ξ_p against mock catalogs in Sec. 5.1 and then test the BAO fit pipeline and present the fit results in details in Sec. 5.2. To do so, we use the ICE-COLA mocks (Ferrero et al. 2021), which is a dedicated mock catalog for the DES Y3 BAO analyses. We briefly describe it here, and refer readers to Ferrero et al. (2021) for more details.

The ICE-COLA mocks are derived from the COLA simulations which is based on the COLA method (Tassev et al. 2013) and implemented by the ICE-COLA code (Izard et al. 2016). The COLA method combines the second order Lagrangian perturbation theory with the particle-mesh simulation technique to ensure that the large-scale modes remains accurate when coarse simulation time steps are used. The simulation consists of 2048^3 particles in a cube of side length $1536 \text{ Mpc } h^{-1}$, matching to the mass resolution of MICE Grand challenge N -body simulations (Fosalba et al. 2015; Crocce et al. 2015). To generate the whole sky light-cone simulation at $z \sim 1.4$, the simulation box is replicated three times in each Cartesian direction (altogether 64 copies). The replication has important consequences for the covariance measurements, and we will further comment on it later on. Mock galaxies are assigned to the ICE-COLA halos using a hybrid Halo Occupation Distribution and Halo Abundance Matching recipe, following a similar approach in Avila et al. (2018). The mocks are calibrated with the redshift distribution and the angular clustering amplitude of the true galaxy samples. From the full-sky light-cone mocks, four maps with the DES Y3 footprint are extracted. Thus there are 1952 mock catalogs in total, although we will only use a subset of them.

A key challenge is to assign realistic photo- z 's to the mock galaxies. To do so the photo- z distribution is estimated using 8362 galaxies with known spectroscopic redshift from VIPERS survey (Guzzo et al. 2014). Using these galaxies, which have both the spectroscopic redshift z_s and the photo- z from DES, the two-dimensional distribution $P(z_p, z_s)$ can be estimated. By sampling $P(z_p, z_s)$ conditioned on the true redshift z_s , the candidate mock galaxies are put to the appropriate photometric bins. A further requirement is that the mock galaxies must match the abundance of the observed photometric galaxies. Suppose there are $N(z_{pi})$ galaxies in the photometric bin z_{pi} , the most luminous $N(z_{pi})P(z_{pi}, z_{sj})$ galaxies from the bin (z_{pi}, z_{sj}) is taken as the mock sample. By construction, the mock

galaxies will match the observed galaxy abundance and the conditional true redshift distribution.

As the cosmology adopted by the mock catalog is the MICE cosmology, the theory templates, i.e. the theoretical model described in Sec. 2.2, and the Gaussian covariance are computed in this cosmology, although we will also consider the template computed in Planck cosmology (Planck Collaboration et al. 2020).

5.1 Template and covariance

In this subsection, we use the ICE-COLA mocks to test the theories developed in Sec. 2 and Sec. 4. The theory calculations require the conditional true redshift distribution and the bias parameters measured from mocks. Thus we shall first present the measurement of these quantities before confronting the theories with measurements. In the Appendix A, we check the template calculation using the spectroscopic correlation function.

The spread of the conditional true redshift distribution generally increases with redshift. As ϕ depends on the number density in true redshift and the spectroscopic data is not available outside $[0.6, 1.1]$, we cannot reliably calibrate ϕ for all bins. Here we approximate ϕ by f . The preliminary estimate suggests that the effect is small, but when the spectroscopic density is rapidly changing such as for the high redshift bins, it should be taken into account. We shall present more detailed analysis in future work. In the standard tomographic analysis, the samples are divided into a number of redshift bins and their auto (and cross) correlation functions are considered. For example, in the DES Y3 BAO analyses, the samples in the photo- z range $[0.6, 1.1]$ are divided into five tomographic bins with bin width $\Delta z = 0.1$. For 3D correlation function, because ϕ is continuously sampled by the galaxy pair counts, we have to use fine photo- z bins in order to sample the ϕ accurately and smoothly.

To investigate the number of photo- z bin width required, we consider a number of ϕ 's measured from the ICE-COLA mocks in the photo- z range $[0.6, 1.1]$. We divide this photo- z range into redshift bins of equal spacing. In Fig. 4, we show a sample of ϕ obtained with 5 and 50 photo- z bins. We also overplot the corresponding Gaussian approximation with the same mean and variance as the true ϕ . The Gaussian approximation is more accurate for the five-bin case as it is averaged over a larger redshift range. At high redshift, ϕ is more irregular and the Gaussian approximation is less accurate.

The photo- z spread is often characterized by

$$\sigma_z = \frac{\sigma_p}{1+z}, \quad (41)$$

where σ_p is the standard deviation of ϕ . We measure σ_z from ϕ for different number of photo- z bins and the results are displayed in Fig. 5. While σ_z is overestimated in the 5-bin case, the 20-bin results are consistent with the ones from the 50-bin case except that the 50-bin results are more noisy. This suggests that when the bin width $\Delta z = 0.025 \lesssim \sigma_z$, the photo- z bin is fine enough to probe the intrinsic photo- z uncertainty. We emphasize that σ_z is used for sanity checks only, we shall use the full distribution in the test of the analysis.

Another ingredient required is the bias parameters. We

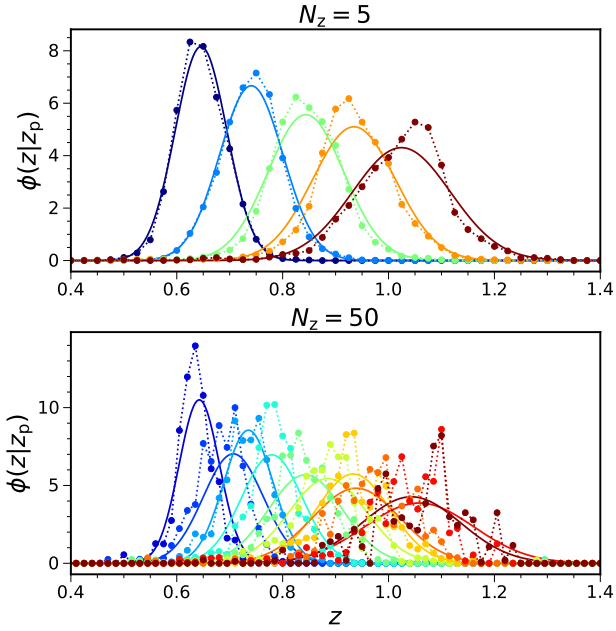


Figure 4. The conditional true redshift distributions measured from COLA mocks (filled circles) with 5 (upper panel) and 50 (lower panel) photo- z bins. The Gaussian approximation is also overlotted (solid curves). For clarity, in the case of 50 bins, only those whose bin number is divisible by 5 are shown.

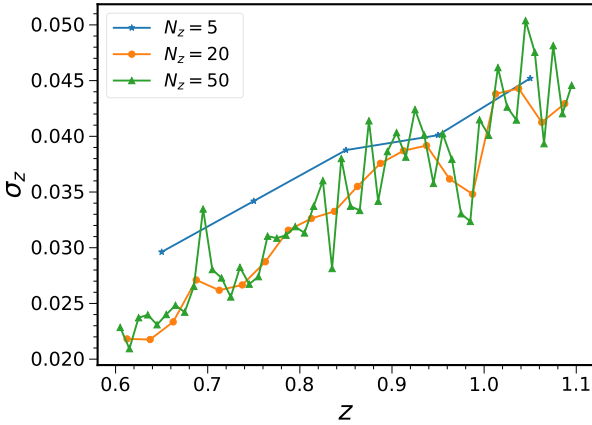


Figure 5. σ_z [Eq. (41)] of the conditional true redshift distribution for 5 (blue), 20 (orange), and 50 (green) photo- z bins.

measure the linear bias parameters by fitting the auto-angular correlation [Eq. (17) and (18)] to the mock measurements as in Chan et al. (2018); Abbott et al. (2019). The fit is performed by fitting to the mean of mocks in the angular range of $[0.5, 1.5]$ degree. We show the bias parameters obtained with different number of tomographic bins in Fig. 6. Similar to the case of σ_z , the 20-bin and 50-bin cases show consistent result with the 50-bin ones showing more

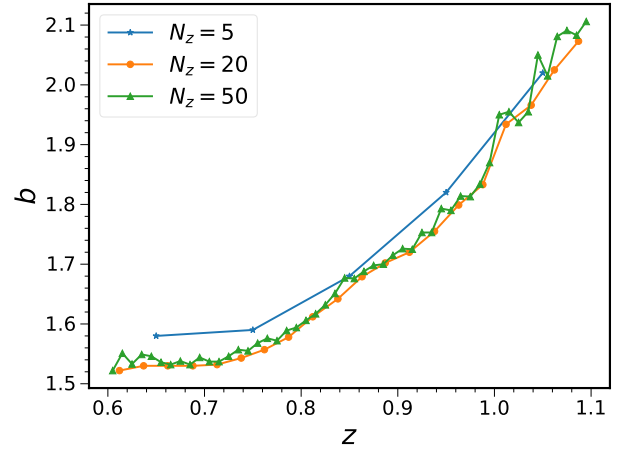


Figure 6. The bias parameters measured from 5 (blue), 20 (orange), and 50 (green) photo- z bins.

random fluctuations, while the five-bin results are slightly higher.

We compare the theory template against the mock measurements in Fig. 7. The 3D correlation ξ_p is measured using the Landy-Szalay estimator (Landy & Szalay 1993) via the public code CUTE (Alonso 2012). Limited by the computing power, we only measure the correlation function from 384 ICE-COLA mocks. Judging from the size of the error bars, which are estimated by the standard error of the mean, the fidelity of the results are well sufficient for our purpose here. We estimate the correlation by considering the pairs with maximum radial separation $120 \text{ Mpc } h^{-1}$ and transverse separation $175 \text{ Mpc } h^{-1}$. To ensure good sampling of the projected correlation within the maximum radial separation, we only consider ξ_p up to $\mu = 0.8$. We further divide the μ range $[0, 0.8]$ into three equal spacing “wedges”. Both ξ_p versus s and s_\perp are shown.

For the theory template, we first compute $w_{ij}(\theta)$ in 50 photo- z bins with angular bin width of 0.02 degree. Because the bin width in the radial direction is wide ($\Delta r \sim 20 \text{ Mpc } h^{-1}$), we linearly interpolate the resultant angular correlation to a larger redshift grid of 500 bins ($\Delta z = 0.005$) before binning all the pairs $w_{ij}(\theta_k)$ into the appropriate $\xi(s, \mu)$ bin. It is important, especially for the projected correlation function, to ensure that the maximum ranges in the radial and transverse directions used in template construction match those in the measurements. We find that the template obtained is of slightly lower amplitude than the mock measurement (by about 3%), the amplitude deficit is probably due to sparse sampling in the radial direction. The template shown in Fig. 7 has been multiplied by an overall amplitude factor. Besides this amplitude adjustment, the theory templates are in good agreement with mock results.

In the conventional angular tomographic analysis, the BAO features manifest in different angular scales, it is less direct to see the photo- z effect on the BAO. To take advantage of the fact that the BAO information is condensed to a single bump in ξ_p , we plot in Fig. 8 the photometric correlation ξ_p against s_\perp and the spectroscopic ξ against s . We have shown the results for both MICE and Planck cos-

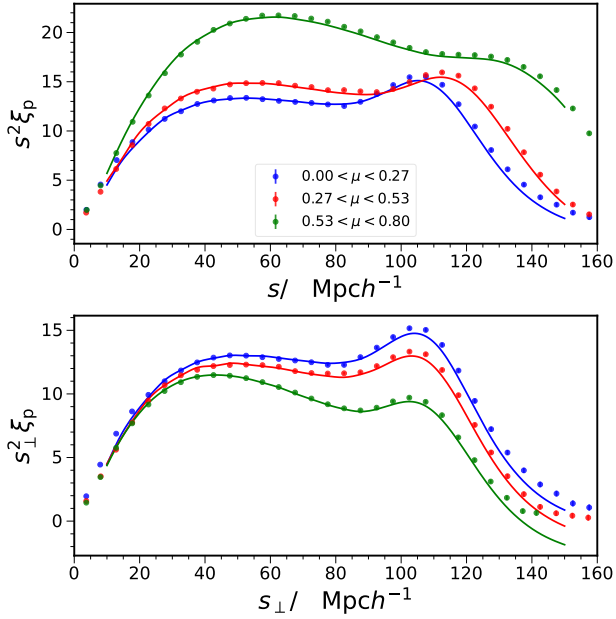


Figure 7. The wedge correlation function ξ_p as a function of the separation s (upper panel) and the transverse separation s_\perp (lower panel). The μ range in $[0, 0.8]$ is divided into 3 equal spacing bin, $[0, 0.27]$ (blue), $[0.27, 0.53]$ (red), $[0.53, 0.8]$ (green). The measurements (filled circles) are the mock mean and the error bars estimated by the standard error are too small to be seen. The theory templates are the solid curves.

mology. The spectroscopic correlation ξ is the dark matter real space correlation with isotropic BAO damping, obtained by an inverse Fourier transform of an analog of Eq. (19) at $z_{\text{eff}} = 0.83$. We also show the sound horizon r_d in these two cases for comparison. Relative to the spectroscopic correlation, the BAO feature in ξ_p is less strong. Note that the BAO peak does not precisely correspond to r_d , and it is on different sides relative to the peak for the spectroscopic and the photometric cases. A small bias could arise because of the subtle shape differences between these two cosmologies. For example, the broadband terms introduced to take care of the shape difference, if their contribution is rapidly changing, it may slightly shift the BAO position. This is more likely to occur in the photometric case given the features are less sharp than in the spectroscopic one.

We now turn to the covariance. In Fig. 9, we show the correlation coefficient matrix for ξ_p

$$\rho_{ij} \equiv \frac{C_{ij}}{\sqrt{C_{ii}C_{jj}}}, \quad (42)$$

where C_{ij} is an element of the covariance matrix. Both the results for s and s_\perp are shown. In this plot we have used the Gaussian covariance. We have considered a single μ bin in the range of $[0, 0.8]$. These plots demonstrate that the covariance of ξ_p has strong off-diagonal elements, especially for s . Real space results are known to be more correlated than the Fourier space ones. We argue that the situation for ξ_p is exacerbated by the photo- z mixing.

It is instructive to contrast the results with the Gaussian covariance for the angular correlation function, which

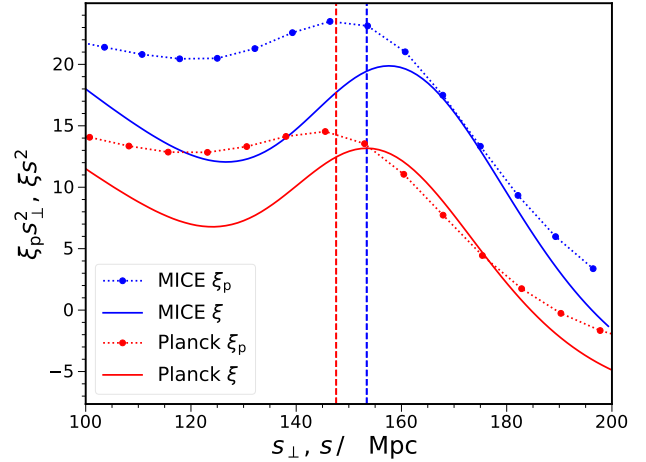


Figure 8. The photometric correlation function ξ_p versus s_\perp (dotted line with filled circles) and the dark matter real space spectroscopic correlation ξ against s (solid) in MICE (blue) and Planck (red) cosmology, respectively. The sound horizons at the drag epoch in these cosmologies are also overplotted as dashed vertical lines (in the same color code). To enhance the BAO feature, they have been multiplied by the appropriate distance squared.

is plotted in Fig. 10. The setup of the calculations are the same as those in the fiducial DES Y3 analyses. There are five photo- z bins, and in each bin, θ ranges from 0.5 to 5 degree with angular bin width of 0.2 degree. The photo- z mixes modes in the radial direction. Because the width of the redshift bin ($\Delta z = 0.1$) is wide compared to the photo- z uncertainty $\sigma [\sim 0.03(1+z)]$, the photo- z mixing does not affect the covariance in the same bin much. The cross correlation between different redshift bins is mainly caused by photo- z mixing, and the only significant correlation is between neighboring redshift bins. Again because the redshift bin is wide compared to σ , the mixing between the neighboring bins is mild, and so even the neighboring bin covariance is weak. On the other hand, as photo- z mixing scale can be as high as $\sim 100 \text{ Mpc } h^{-1}$, for ξ_p there can be substantial mixing between the underlying correlation from different scales. Thus the covariance for ξ_p can be significantly higher than that for w .

As we mentioned, construction of the ICE-COLA mock catalogs involves replicating the simulation box to form a full-sky light-cone mock. The repeated structures cause the covariance to be anomalously higher than the correct one, and this prevents us from using the mock covariance as the benchmark to test our Gaussian theory covariance. See Ferrero et al. (2021) for more discussions. Thus in DES Y3, the analytic Gaussian covariance is the default choice. Although we cannot directly test the ξ_p covariance against mock results, the code used to compute the Gaussian covariance for w , which is the backbone of the ξ_p results, has been tested in Chan et al. (2018); Avila et al. (2018); Abbott et al. (2021). Moreover, we have verified that the amplitude ratio difference between the Gaussian theory covariance and the mock covariance are at a similar level for w and ξ_p , albeit quantitative comparison is not possible.

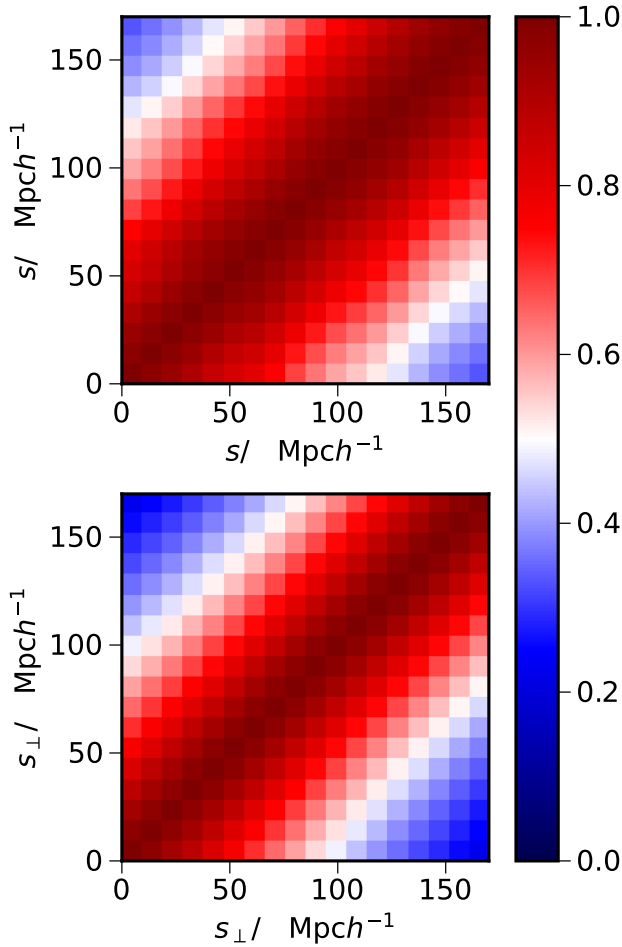


Figure 9. The correlation coefficient for ξ_p . The upper panel shows the results for s and the lower one for s_\perp . The results are for one μ bin in the range of $[0, 0.8]$.

5.2 BAO fit

In this subsection, we apply our theory template and covariance to extract the BAO scale from the mock catalogs. With the variable s_\perp , the BAO scale is almost equal to the true one, we adopt ξ_p as a function of s_\perp in the BAO fit. To measure the BAO scale, as in standard BAO analyses [e.g. Xu et al. (2012); Anderson et al. (2014); Ross et al. (2017)], we apply the following general model

$$M(s_\perp) = BT(\alpha s_\perp) + \sum_i \frac{A_i}{s_\perp^i}, \quad (43)$$

where T denotes the theory template and B gives the freedom to adjust the amplitude. The parameter α in the argument of T is the key parameter we are after. It allows us to shift the BAO scale computed in the fiducial cosmology to match that in the data. The polynomial in $1/s_\perp$ is designed to absorb residual broadband effects due to imperfectness in modeling, changes in correlation function shape as a result

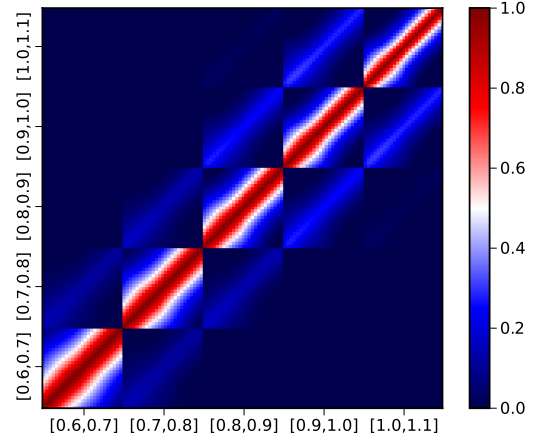


Figure 10. The correlation coefficient for w . The data in the redshift range $[0.6,1,1]$ is split into five tomographic bins with bin size 0.1.

of the differences between the fiducial cosmology and the actual one, and residual broadband systematic effects.

The default configuration for the BAO fit is as follows: the broadband terms consisting of $\sum_i A_i/s_\perp^i$ with $i = 0, 1,$ and 2 ; the fit range $[40, 140] \text{ Mpc } h^{-1}$; the bin width $\Delta s = 3 \text{ Mpc } h^{-1}$, and the Gaussian theory covariance. We use 384 mocks for the BAO fit test. Different configurations for BAO fit are checked.

We look for the best fit α by the maximum likelihood estimator following the procedures in Chan et al. (2018). We first analytically fit the linear parameter A_i , and then minimize the resultant χ^2 w.r.t. B subject to the constraint $B > 0$. At the end of the operations, we are left with the residual χ^2 as a function of α . The best fit α corresponds to the point where the residual χ^2 attains its minimum and the symmetric $1-\sigma$ error bar is derived from the $\Delta\chi^2 = 1$ criterion. Besides we also cross check our results with the Markov Chain Monte Carlo (MCMC) method, in which all the parameters are fitted simultaneously.

5.2.1 Highly correlated data fit

The photo- z mixing causes ξ_p to be strongly correlated, as evident in Fig. 9. Here we first describe some fitting problems caused by the high correlation, namely the apparently bad fit results and the alarmingly large χ^2 per degree of freedom (χ^2/dof). The former issue is caused by the largest eigenvalues of the covariance and the latter is due to the small ones. We discuss a few approaches to modify the covariance, and find that the first issue can be resolved by suppressing the largest eigenvalues. For the second issue, while it is not practical to solve it by adjusting the eigenvalues, we demonstrate that the BAO feature is well-fit by our model and the problem is not directly related to the BAO.

To begin with, we present some peculiar fit results. One of the odd examples appears in the BAO fit to the mean of the mocks, which are shown in Fig. 11. When the MICE cosmology template is used, the resultant fit is very good

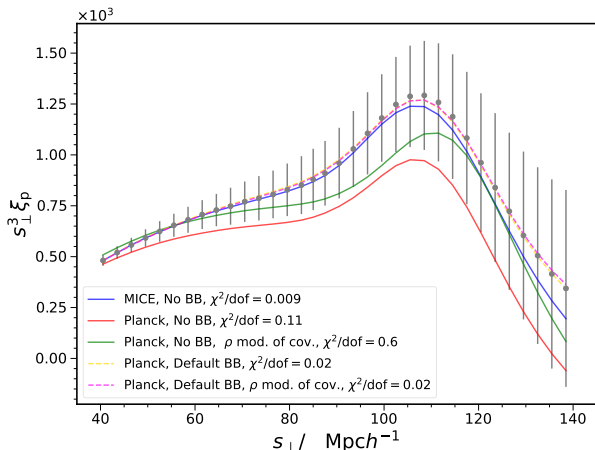


Figure 11. The BAO fit results to the mean of the mock measurements (grey data points with error bars). The fit obtained with the MICE template without broadband terms (solid, red) is compared with the Planck template without broadband terms result (solid, green). The latter case is below all the data points. With the inclusion of the broadband terms, the Planck template (dashed, yellow) also offers good fit. The results for the modified covariance are also shown [No broadband terms (solid, green) and with broadband terms (dashed, violet)].

and it is not surprising that the χ^2/dof is very small (0.009). However, when the template computed in Planck cosmology (Planck Collaboration et al. 2020) is used, we find that the resultant χ^2/dof is still small (0.11), but the fit is manifestly poor by eye because the best fit model is below *all* the data points and most of the data points are more than $1\text{-}\sigma$ away from the best fit.

Although this issue only appears when the model Eq. (43) does not include the broadband terms at all and disappears otherwise, it does indicate that something abnormal happens to the fit⁸. In the other extreme, when we fit to the individual mocks, we find that the fit results are often fine by eye as the best fit model fall within all the $1\text{-}\sigma$ errors, but χ^2/dof is alarmingly large, some of them reach as high as 2. Note that the second issue we describe here appears for generic fit configurations, in particular for both Gaussian and mock covariance fit. We point out that the second problem is already implicit in Abbott et al. (2019) and Sridhar et al. (2020), as their ξ_p fit results look decent by eye but appreciably large χ^2/dof were reported. We attribute these problems to the high correlation of the data induced by the photo- z mixing.

Although the covariance of ξ_p has strong off-diagonal elements, its inverse, the precision matrix Ψ could be approximated by a band matrix with a few non-vanishing elements near the diagonal. The upper panel of Fig. 12 shows the elements of Ψ , and it is approximately a pentadiagonal band matrix, i.e. a band matrix has non-zero elements only for the main diagonal, and the first two upper and two lower diagonals. We can use this to intuitively understand why the

⁸ The issue is not caused by the fitting method. Usage of the MCMC method also gives the same results.

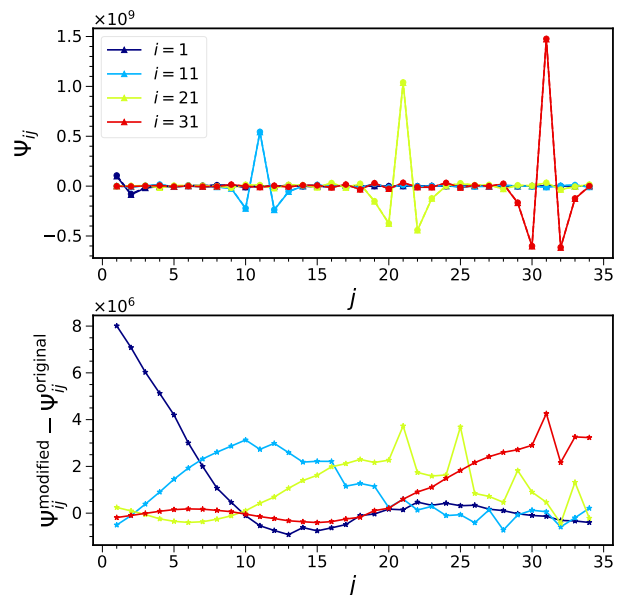


Figure 12. Upper panel: The elements of the precision matrix Ψ_{ij} plotted against j for fixed i . The elements for the original Ψ (triangles) and the one obtained by modifying the covariance matrix (circles, see the text for details) are compared. Lower panel: The differences between the modified Ψ_{ij} and the original one are displayed. The effect of the modified covariance is mainly to shift Ψ to the positive side. As the large eigenvalue modes correspond to eigenvectors of long “wavelength”, it also weakly affects the zones far from the diagonal.

best fit may not be driven to the data points as close as possible. By assuming Ψ to be pentadiagonal, the χ^2 can be written as

$$\chi^2 = \sum_{j=1}^N (\Delta_{j-2} \Psi_{j-2,j} + \Delta_{j-1} \Psi_{j-1,j} + \Delta_j \Psi_{j,j} + \Delta_{j+1} \Psi_{j+1,j} + \Delta_{j+2} \Psi_{j+2,j}) \Delta_j, \quad (44)$$

where Δ denotes the difference between the model and the data. This expression can accommodate the results near the “boundaries” as well if $\Psi_{i,j} \equiv 0$ whenever $i, j < 1$ or $> N$. Now suppose $\Delta_j > 0$, because $\Psi_{j,j} > 0$, to minimize χ^2 , Δ_j is attracted to zero. For diagonal covariance, this is the sole contribution. However, because Ψ_{j-2} , Ψ_{j-1} , Ψ_{j+1} , and Ψ_{j+2} are negative, it is favorable for Δ_{j-2} , Δ_{j-1} , Δ_{j+1} , and Δ_{j+2} to be repelled to as negative as possible. Of course these expressions are coupled, more careful analysis is required, but it is likely that the competition between the diagonal terms and the off-diagonal ones lead to a best fit that does not precisely fall on the data.

To look into the covariance, we plot its eigenvalues in Fig. 13. We find that the spectrum spans three orders of magnitude, and most of them are of relatively small values. As a comparison we also show the diagonal elements of C , its range is much more limited. For the fit to the mean of mocks using the Planck template, we find that the fit is dominated by the largest eigenvalue mode. Because of its (excessively) large eigenvalue, its contribution to the χ^2 is very small, so that even the fit is visually bad, χ^2/dof is still small. On the

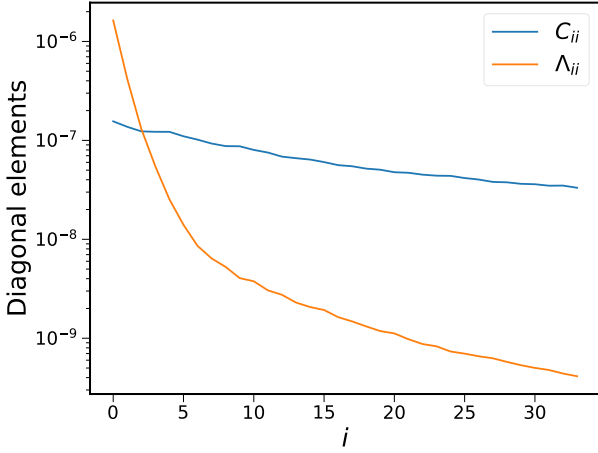


Figure 13. The eigenvalue spectrum of the covariance matrix (orange) and the diagonal elements of covariance matrix (blue). While C_{ii} is in its natural order, the eigenvalues have been arranged in descending order.

other hand, for the fit to individual mocks, the large χ^2/dof values are caused by significant contributions from modes with small eigenvalues.

The problem of fitting highly correlated data is also encountered in the lattice field theory because lattice simulation data are highly correlated. Yoon et al. (2013) surveyed a few methods proposed by the lattice field theory community to tackle this problem. Michael (1994) advocated ignoring the off-diagonal elements in the covariance, but this is clearly a too violent modification. A more mild method is to remove the eigenmodes deemed problematic. Both removing the largest eigenvalue modes (Thacker & Lepage 1991) and the smallest ones (Drummond & Horgan 1993; Kilcup 1994) have been proposed. For the largest eigenvalue modes removal, it was argued that these modes have small contribution to χ^2 . The smallest eigenvalue modes removal is often favored in the spirit of the singular value decomposition (Press et al. 2007). A physical argument in support of it is that those modes are generally highly oscillatory and so they are not conducive to getting a smooth fit. In a similar vein, Bernard et al. (2002) proposed to modify instead the correlation coefficient Eq. (42). They replace the eigenvalues of ρ smaller than certain threshold λ_{th} by λ_{th} . A new covariance matrix can be constructed using the original eigenvectors and the modified eigenvalues. Modifying the eigenmodes is more attractive as it keeps all the directions in the covariance space constrained.

We test the results obtained with different covariance recipes. While for most of the fit, different approaches give similar results, in a small fraction (order of 10%) the differences are quite pronounced. In Fig. 14, we show the fit results from one of the mocks, in which the differences are clearly visible. The one obtained with the original covariance (blue) is manifestly not optimal. The best fit is 1.006 ± 0.017 in this case. In the other extreme, when the diagonal approximation (red) is used, it is not surprising that the best fit is driven close to data points, but the best fit we get is

1.002 ± 0.032 , with the error bar substantially larger than the original one. By suppressing the largest eigenvalues of the covariance (green) or ρ (yellow), we can also get an apparently better fit. We replace the eigenvalues larger than certain threshold λ_{th} by λ_{th} . We take λ_{th} to be 3×10^{-8} and 1 for the covariance and correlation coefficient modification respectively. The thresholds are chosen so that as few eigenvalues are affected as possible and the affected ones account for about 10% of them. The best fits we get are 1.003 ± 0.016 and 1.005 ± 0.017 respectively.

In Fig. 12 we also show the elements of Ψ obtained by suppressing the largest eigenvalues of ρ . Overall, the size of the modification is small, of the order 10^{-3} . The main effect of the covariance modification is to shift Ψ to the positive side so that the attraction to zero is enhanced and the repulsion to negative infinity is weakened. Note that the large eigenvalue modes correspond to eigenvectors that are oscillatory in nature with long “wavelength”, while the small eigenvalue ones tend to be well localized. Thus the modification also weakly affects the the zone far from the diagonal. Furthermore, we find that the ρ modification recipe often results in better agreement with the original one compared to the covariance modification, hence we will adopt ρ modification in the following analysis. In Fig. 11, we also show the results obtained with ρ modification. Without the broadband terms, the best fit now passes through the data points, although the BAO position appears to be biased. With the default broadband terms (A_i with $i = 0, 1, 2,$ and 3), the bias is removed.

On the other hand, the large χ^2/dof problem is caused by the small eigenvalues. The nominal χ^2/dof for the example shown in Fig. 14 are similar for different recipes, around 2.7, except for the diagonal approximation, which gives $\chi^2/\text{dof} \sim 0.06$. Since only the largest eigenvalues are modified, the χ^2 is only increased slightly compared to the original one. As it is clear from the spectrum in Fig. 13, most of the eigenvalues are small and of similar values, we need to remove or modify a large fraction of eigenmodes before the effect becomes noticeable. Removing 40% of the eigenmodes with the smallest eigenvalues, the naive χ^2/dof can be reduced to roughly 1. This represents a very substantial modification of the covariance matrix. Moreover, after eigenmodes removal or modification, there is no proper theoretical basis for χ^2/dof anymore (Bernard et al. 2002). Thus we shall not pursue to reduce χ^2/dof by adjusting the covariance matrix.

Using the example in Fig. 14 again, we check directly if the BAO scale is well fitted by the model in Fig. 15. To decorrelate the data, we rotate the data vector by the matrix U , which is the orthogonal matrix that diagonalizes the covariance, i.e. $C = U^T \Lambda U$. To highlight the BAO feature in the rotated coordinates, the no-BAO template has been subtracted from both the data vector and the best fit model. Note that as the data vector is mixed by U , they no longer strictly correspond to the original distance r . Here we only use it as a convenient label for the x -axis. Nonetheless, from the difference with the no-BAO template, the location of the BAO feature can be inferred. In this basis, the BAO feature shows up as a series of wiggles in $r \gtrsim 110 \text{ Mpc } h^{-1}$ instead of a single peak in the usual configuration representation. The error bars on the data are given by the eigenvalues in Λ and they are not correlated in this basis. We have also shown

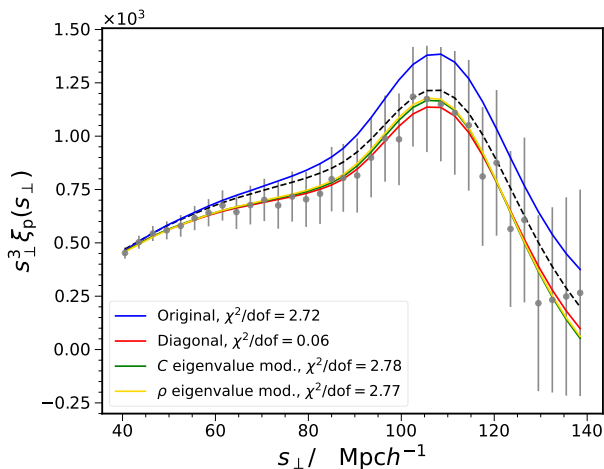


Figure 14. The best fit to the ξ_p measurement from one of the mocks (circles, the error bars are $1\text{-}\sigma$ error from the diagonal elements of the covariance). This mock is chosen because the differences among different covariance recipes are pronounced. The results obtained with the original covariance matrix (blue), diagonal approximation (red), suppressing the largest eigenvalues in the covariance (green), and suppressing the largest eigenvalues in the correlation coefficient (yellow) are compared. The raw template (black, dashed) is also plotted.

error bars from the modified covariance according to the ρ modification recipe. We see that the modification reduces the error bars around the BAO scale. To highlight the source of large χ^2/dof , we show the ratio with respect to the error bar in the lower panel of Fig. 15. Both the results obtained with the original error bars and the modified ones are shown, which coincide with each other except for $r > 130 \text{ Mpc } h^{-1}$. Significant deviations of the best fit model from the data points occur for $r \lesssim 100 \text{ Mpc } h^{-1}$, while in the BAO zone, the model and the data agree with each other well for both types of error bars. Thus we conclude that the large χ^2/dof is caused by the part of the fit that is not directly related to the BAO.

5.2.2 Fit results

In this section, we present the BAO fit results in details. We will consider various metrics to quantify the accuracy of the fit and different variation of the fit configurations. Although these are important, they can be tedious to read. The key default result is given in the first line in Table 1. In particular, the ξ_p constraint on α is $\langle \alpha \rangle \pm \sigma_{\text{std}} = 1.001 \pm 0.023$. We have checked that the results are robust against variation of the fit configurations. The default setup uses the Gaussian covariance modified with the ρ modification prescription. We shall also contrast them with the corresponding results for the angular correlation. They are also consistent with the w results even though ξ_p statistic has considerably larger $\langle \chi^2 \rangle / \text{dof}$.

The default case results for ξ_p fit are shown in the first row in Table 1. As a reference, we also show the corresponding results for the angular correlation function, which are obtained from a joint five-tomographic fit. See Table II in

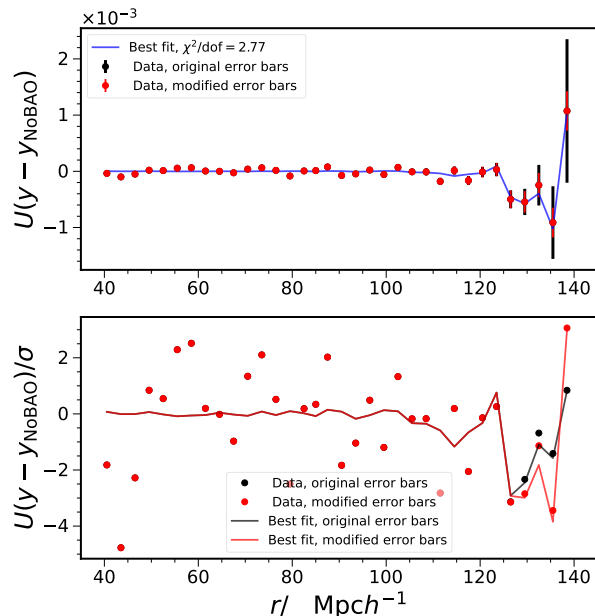


Figure 15. Upper panel: The data (filled circles) and the best-fit model (solid blue) shown in the rotated basis. The orthogonal matrix U decorrelates the correlation in the error. To highlight the BAO feature, the no-BAO template has been subtracted. In this basis, the BAO feature manifests as a series of wiggles for $r \gtrsim 110 \text{ Mpc } h^{-1}$. The uncorrelated error bars from the original covariance (black) and the modified covariance obtained from the ρ modification scheme (red) are displayed. The modified error bars are reduced compared to original ones only in the large r range. The data used in this plot is based on the same mock data used in Fig. 14. Lower panel: The results are further divided by the error bars to show the significance of the deviations between the model and the data points. The results obtained with the original error bars (black) and the modified error bars (red) overlap with each other and deviations show up only for $r > 130 \text{ Mpc } h^{-1}$. In the BAO zone, the model and the data points agree with each other well. The large χ^2/dof is driven by the data points lying outside the BAO feature.

Abbott et al. (2021) for more details on the w fit. The mean of the best fit $\langle \alpha \rangle$ is 1.001, with small bias which can be attributed to nonlinear evolution (Crocco & Scoccimarro 2008; Padmanabhan & White 2009). Two measures of the spread of the distribution of the best fit α are shown: the standard deviation σ_{std} and the half of the width between 16 and 84 percentile of the distribution, σ_{68} , which is less sensitive to the tails of the distribution. We find that σ_{std} is larger than σ_{68} by a small amount (0.001), suggesting the influence by the tails. As mentioned, the $1\text{-}\sigma$ error bar is derived from the $\Delta\chi^2 = 1$ condition, and for the error bars to be meaningful, $\langle \sigma_\alpha \rangle$ should agree with the measures of the spread of the distribution. In the case of Gaussian likelihood, they are equal to each other. We find that the mean of the error bar is slightly smaller than the spread of the distribution. These numbers are marginally better than those for w . The bias in the best fit $\langle \alpha \rangle$ is smaller (1.004 for w) and the measures of spread and error are smaller by ~ 0.001 . These (mild) improvements could arise from the fact that the ξ_p statistic

Table 1. BAO fit results on the COLA mocks with different configurations in the analysis. The default setup is as follows: the broadband terms consisting of $\sum_i A_i/s_{\perp}^i$ with $i = 0, 1, \text{ and } 2$, the fit range $[40, 140] \text{ Mpc } h^{-1}$, the bin width $\Delta s = 3 \text{ Mpc } h^{-1}$, and Gaussian theory covariance with eigenvalues of ρ suppressed. See the texts for details.

case	$\langle \alpha \rangle$	σ_{std}	σ_{68}	$\langle \sigma_{\alpha} \rangle$	frac. encl. $\langle \alpha \rangle$	$\langle d_{\text{norm}} \rangle$	$\sigma_{d_{\text{norm}}}$	mean of mocks	$\langle \chi^2 \rangle / \text{dof}$
Default	1.001	0.023	0.022	0.020	61%	-0.005	1.18	1.002 ± 0.019	58.7/29 (2.03)
w fit	1.004	0.024	0.023	0.021	62%	-0.024	1.11	1.004 ± 0.021	93.1/89 (1.05)
$A_i, i = 0$	1.001	0.023	0.022	0.019	59%	0.032	1.19	1.002 ± 0.019	62.9/31 (2.03)
$A_i, i = 0, 1$	1.001	0.023	0.021	0.021	64%	0.003	1.12	1.003 ± 0.020	58.6/30 (1.95)
$A_i, i = 0, 1, 2, -1$	1.002	0.023	0.021	0.020	60%	-0.005	1.17	1.002 ± 0.019	56.8/28 (2.03)
Fit range $[70, 130] \text{ Mpc } h^{-1}$	1.000	0.027	0.024	0.022	59%	0.084	1.20	1.003 ± 0.021	36.4/16 (2.28)
$\Delta s = 10 \text{ Mpc } h^{-1}$	1.001	0.025	0.024	0.020	57%	-0.007	1.25	1.001 ± 0.020	10.2/5 (2.05)
$\Delta s = 5 \text{ Mpc } h^{-1}$	1.003	0.025	0.023	0.020	59%	-0.022	1.17	1.002 ± 0.020	32.1/15 (2.14)
$\Delta s = 2 \text{ Mpc } h^{-1}$	1.002	0.024	0.022	0.020	60%	0.008	1.12	1.002 ± 0.019	81.0/45 (1.80)
$\Delta s = 1 \text{ Mpc } h^{-1}$	1.001	0.025	0.023	0.019	57%	0.019	1.37	1.002 ± 0.018	143.1/95 (1.51)
Planck template	0.958	0.022	0.020	0.019	65%	-0.003	1.14	0.957 ± 0.018	53.7/29 (1.85)
Original Gaussian cov.	1.002	0.023	0.022	0.021	64%	-0.005	1.13	1.003 ± 0.020	57.7/29 (2.00)
COLA cov	1.001	0.023	0.022	0.021	65%	-0.009	1.16	1.002 ± 0.021	47.9/29 (1.65)
MCMC fit	1.001	0.024	0.023	0.020	59%	-0.004	1.18	1.003 ± 0.020	
$N_{\mu} = 3$	0.999	0.028	0.025	0.017	47%	-0.024	1.60	0.999 ± 0.017	139.0/89 (1.56)
$N_z = 2$	1.000	0.023	0.022	0.018	59%	-0.014	1.24	1.001 ± 0.018	98.1/59 (1.66)

makes use of the cross correlation in the data as well, not only the tomographic bin auto-correlation for w as in the study of DES Y3.

Another indicator for the accuracy of 1- σ error bar is the fraction of times that it encloses $\langle \alpha \rangle$, which is 68% in the case of Gaussian likelihood. In line with the error bar being slightly smaller than the spread, the fraction enclosing $\langle \alpha \rangle$ is slightly lower than the Gaussian expectation. The normalized deviation, $d_{\text{norm}} \equiv (\alpha - \langle \alpha \rangle) / \sigma_{\alpha}$ is designed to study the correlation between the deviation of the best fit α from the ensemble mean and the error bar derived. The mean $\langle d_{\text{norm}} \rangle$ is close to zero, but the standard deviation $\sigma_{d_{\text{norm}}}$ exceeds unity by 18%. This trend is consistent with the previous observation that the spread of the distribution is slightly larger than the error estimate. In contrast, for w , the fraction of time enclosing $\langle \alpha \rangle$ is 62%, and $\langle d_{\text{norm}} \rangle = 0.02$ and $\sigma_{d_{\text{norm}}} = 1.12$. Both suggest that ξ_p results are less Gaussian than the w ones.

As a reference, we also show the fit to the mean measurement of the mocks. To show the results for a “typical” fit, we have used the covariance corresponding to a single survey volume. The results are similar to $\langle \alpha \rangle \pm \langle \sigma_{\alpha} \rangle$. However, we note that for w these different measurement are almost always equal to each other, suggesting that the averaging and fitting operations are more commutative for w .

Finally, the $\langle \chi^2 \rangle / \text{dof}$ (~ 2) is also shown for reference, but we should bear in mind the complications mentioned in the previous subsection. On the other hand, for w , $\langle \chi^2 \rangle / \text{dof}$ is 1.05. We attribute this to the fact that the covariance for w is much less correlated.

A number of tests on the variation of the fitting configurations have been performed. We have tested the number of the broadband terms $\sum_i A_i/s_{\perp}^i$ with $i = 0, i = 0, 1, \text{ and } i = 0, 1, 2, -1$. Besides minor fluctuations, the fit results are

not sensitive to the number of broadband terms used. This is because the broadband terms are not degenerate with α , but they are degenerate among themselves. When a narrower fitting range $[70, 130] \text{ Mpc } h^{-1}$ is used, the constraining power is reduced and so the spread of the distribution is enlarged, especially for σ_{std} . As the error bar σ_{α} increases only mildly, the difference between the error bar and the spread of the distribution increases.

The bin width size $\Delta s = 10, 5, 2, \text{ and } 1 \text{ Mpc } h^{-1}$ are compared with the fiducial setting $\Delta s = 3 \text{ Mpc } h^{-1}$. Among the cases shown, σ_{std} for $\Delta s = 3 \text{ Mpc } h^{-1}$ is the lowest and it is also more consistent with σ_{68} and $\langle \sigma_{\alpha} \rangle$. Note that χ^2 / dof are smaller for $\Delta s = 2$ and $1 \text{ Mpc } h^{-1}$, but this is not vital here as we already mentioned that the BAO region is well-fit by our model, the reduction in χ^2 / dof is primarily due to the region lying outside the BAO.

To test if the expected shift in α can be recovered when the cosmology of the template is different from that of the mocks, we generate a template computed with the Planck cosmology (Planck Collaboration et al. 2020). We can estimate the expected α by

$$\alpha = \frac{r_d}{D_M} \Big|_{\text{MICE}} = 0.959, \quad (45)$$

$$\frac{r_d}{D_M} \Big|_{\text{Planck}}$$

where r_d is the sound horizon at the drag epoch and D_M is the comoving angular diameter distance evaluated at the effective redshift $= 0.835$ following DES Y3. For MICE cosmology, $r_d = 153.4 \text{ Mpc}$ and $D_M(z_{\text{eff}}) = 2959.7 \text{ Mpc}$, while for Planck cosmology $r_d = 147.6 \text{ Mpc}$ and $D_M(z_{\text{eff}}) = 2967.0 \text{ Mpc}$.

Unlike other tests, when an alternative fiducial cosmology is assumed, the measurement of ξ_p should be performed again as the distance has to be computed in the new cosmology.

Table 2. Comparison of the α values obtained with the Planck and the MICE templates in fitting to the MICE cosmology data. The simple estimates using Eq. (45) are shown in the first row, and the mock measurement in the second. As a reference, the mock measurements for w from Abbott et al. (2021) are also shown in the last row.

	Planck	MICE
Simple estimate	0.959	1
ξ_p mock results	0.958	1.001
w mock results	0.966	1.004

ogy. Furthermore, we mentioned that the Hubble parameter h is contained in the unit $\text{Mpc } h^{-1}$ and the distance does not explicitly depend on it.

In Table 2, we compare the values of the α parameter obtained by the simple estimate using Eq. (45) and the mock measurement results. In contrast, we also show similar results for w from Abbott et al. (2021). Interestingly, we find that the ξ_p results are closer to the simple estimate than the ones from w . The deviation from the simple estimate could stem from the nonlinear correction and other systematic effects, and they affect the results from both templates in a similar manner.

We have shown the results obtained with the original Gaussian covariance. The ρ modification prescription only affects a small fraction of the mocks and the results are statistically similar to the original Gaussian case. Moreover, although the COLA covariance suffers from the overlap issue, we also show its results for reference. Despite the overlap issue, the results, especially the measures of the spread and the error bar, are similar to the default ones. Because we are fitting the mocks with the same covariance derived from it, the $\chi^2/\text{dof} = 1.65$ is smaller, but still it is much larger than 1. Overall, these suggest the consistency of our results across the covariance used.

The sequential maximum likelihood fitting is applied mostly for its efficiency and convenience. We have checked our results using MCMC, implemented by `emcee` (Foreman-Mackey et al. 2013), in which all the parameters are fit simultaneously. Similar to the conclusion for w in Chan et al. (2018), we find that the MCMC fit gives consistently similar results for ξ_p .

Furthermore, we have considered splitting the data in the μ range [0,0.8] into three equal- μ bins and dividing the redshift range [0.6,1.1] into two sub-redshift bins: [0.6, 0.85] and [0.85, 1.1]. We have also included cross- μ bin or cross sub-redshift bin covariances. In both cases we find that the discrepancy between σ_{std} and $\langle\sigma_\alpha\rangle$ increases. The issue is more serious for $N_\mu = 3$ case, and these invite the interpretation that the discrepancy increases as the number of bins increases. We also see similar trend in the smallest Δs bin case. These could be attributed to the high covariance of the data as the correlated data are more susceptible to random fluctuations. By dividing the data into more bins, there are larger random fluctuations. On one hand, there are larger fluctuations in the best fit resulting in an increase in the spread of the distribution. On the other hand, the error bar derived from the local curvature of the likelihood function

could be underestimated as the likelihood is less smooth and the curvature around the minimum is increased.

In Table 1, we already compared the ξ_p results in the photo- z range [0.6,1.1] against the w results obtained using the joint fit with five tomographic bins, each of bin width $\Delta z_p = 0.1$. It is instructive to compare the performance of ξ_p and w for individual tomographic bins, and the results are shown in Table 3. In this test the intrinsic photo- z distribution is comparable to the bin width. Unlike the joint-bin fit, there is no cross-bin information. The difference between these statistics is that the data is further projected onto a sphere for w . Because of the limitation in computing power, only 100 mocks are used in this comparison. Overall, the results for w are more stable across redshift bins. Except for the first bin, the measures of the spread of the distribution are larger for w , and this can be interpreted as information loss due to projection. However, we note that for ξ_p , $\langle\sigma_\alpha\rangle$ tends to be smaller than the measures of the spread by a larger amount. Besides, $\sigma_{d_{\text{norm}}}$ is also more significantly larger than 1. These are caused by the increase in noise as the data size reduces and are consistent with the trends for $N_{\text{mu}} = 3$ and $N_z = 2$ in Table 1. The χ^2/dof for ξ_p is still substantially larger than 1 although it is smaller than the joint fit case, while for w , it remains close to 1. In summary, ξ_p makes use of more information and hence potentially more constraining, but it is more susceptible to noise in the data.

Before closing this section, we would like to comment further on the dual roles of χ^2 : the characterization of the goodness of fit and the model differentiation. The χ^2/dof is often used to indicate the goodness of the fit⁹. We discussed previously that the χ^2/dof is substantially larger than 1 although the model offers a good fit to the BAO. A separate question is if we can use the χ^2 value for model differentiation, especially for claiming the significance of the BAO detection. Although we may not be able to use the χ^2 distribution results, in principle we can use the mocks to answer the question that if there are no BAO signals in the data, how much the chance is that the BAO template yields a χ^2 value smaller than certain threshold.

6 CONCLUSIONS

There are numerous large-scale structure photometric surveys ongoing or upcoming. One of the important probes to extract cosmological information from these surveys is via the measurement of the BAO. A common way to measure BAO in photometric data is to divide the sample into tomographic bins and measure their auto (and cross) correlation.

It has been suggested to analyze the photometric data as the spectroscopic ones through the 3D correlation function ξ_p . Previous modeling was limited to a Gaussian photo- z approximation. In this work, we generalize the modeling to accommodate general photo- z distribution. This eliminates the possibility of biasing the results by assuming Gaussian

⁹ Strictly speaking, the degrees of freedom is only defined for linear models as stressed in Andrae et al. (2010), while the parameter α (and B) appears nonlinearly. Here, we followed the usual practice to assume that the degrees of freedom can be approximated defined for our model.

Table 3. Comparison of the ξ_p and w BAO fit results, for which only the data in a single tomographic bin is used.

case	$\langle\alpha\rangle$	σ_{std}	σ_{68}	$\langle\sigma_\alpha\rangle$	frac. encl. $\langle\alpha\rangle$	$\langle d_{\text{norm}}\rangle$	$\sigma_{d_{\text{norm}}}$	mean of mocks	$\langle\chi^2\rangle/\text{dof}$
1, ξ_p	0.996	0.050	0.049	0.036	45%	0.051	1.48	1.002 ± 0.039	44.3/29 (1.53)
1, w	0.996	0.049	0.045	0.045	66%	-0.026	1.18	0.995 ± 0.049	17.4/17 (1.02)
2, ξ_p	0.997	0.044	0.040	0.033	52%	-0.025	1.33	0.999 ± 0.037	44.2/29 (1.53)
2, w	1.000	0.049	0.047	0.045	67%	-0.053	1.19	0.999 ± 0.049	16.3/17 (0.96)
3, ξ_p	1.001	0.048	0.039	0.035	56%	-0.029	1.38	1.003 ± 0.035	44.8/29 (1.54)
3, w	0.998	0.050	0.039	0.041	67%	-0.042	1.20	0.997 ± 0.043	17.6/17 (1.03)
4, ξ_p	0.992	0.043	0.035	0.032	60%	-0.006	1.45	0.993 ± 0.034	39.1/29 (1.35)
4, w	1.005	0.050	0.050	0.043	59%	-0.067	1.12	1.003 ± 0.043	17.2/17 (1.01)
5, ξ_p	1.009	0.039	0.033	0.038	69%	-0.056	1.03	1.007 ± 0.039	43.7/29 (1.51)
5, w	1.005	0.045	0.042	0.042	66%	-0.027	1.06	1.006 ± 0.047	18.9/17 (1.11)

photo- z approximation. To do so, we bin the general cross angular correlation between different redshift bins $w_{ij}(\theta_k)$ using the variables s and μ (or s_\perp and μ) like in usual 3D correlation function analysis. Our approach highlights the connection between the angular correlation function and 3D correlation function.

To help understanding the impact of the photo- z uncertainties on the correlation function, especially the BAO, we specialize the calculations to the Gaussian photo- z case. When $\sigma \gtrsim 0.02$, the BAO peak in the correlation shows up at the true BAO scale only if the transverse scale of the separation, s_\perp , is used as the independent variable. We show that this results from the interplay between the photo- z distribution and the Jacobian of the transformation. The latter formally diverges at s_\perp , and for $\sigma \gtrsim 0.02$, it dominates the integral, causing ξ_p to trace the underlying correlation function at the scale s_\perp . Large-scale photometric surveys equipped with broadband filters such as DES typically have $\sigma \gtrsim 0.02$. In this case, it is favorable to adopt the parametrization $\xi_p(s_\perp, \mu)$, which effectively projects the 3D correlation function to the transverse direction. Unfortunately, this also means that ξ_p only probes the transverse BAO and cannot be used to directly constrain the Hubble parameter.

Our approach also enables us to derive the Gaussian covariance for ξ_p from the Gaussian covariance for the angular correlation function w . Due to photo- z mixing, the covariance of the ξ_p statistics has strong off-diagonal elements. This high correlation causes problems to the data fitting. One of the problems is that some of the best fit results are manifestly poor, but we find that they can be resolved by suppressing the large eigenvalue modes in the covariance. Another issue is that the mean χ^2/dof for the fit is large, even though the fit is apparently good. While this second issue cannot be easily resolved by adjusting the covariance, we find that the issue is primarily caused by the small eigenvalue modes and they are not directly related to the BAO scale. We conclude that the BAO scale is well-fit by our model (Fig. 15).

We have used a set of dedicated DES Y3 mocks to test our pipeline. This set of mocks include realistic photo- z distribution among other things. We find that the theory template is in good agreement with the mock measurement. After verifying our methodology, we apply our pipeline to the

BAO fit in the mock catalog. Overall we find that the results are consistent with the ones derived from w even with small improvements because ξ_p makes use of the cross redshift bin correlation as well. However, the angular correlation function measurements are less correlated with $\chi^2/\text{dof} \sim 1$. As ξ_p offers an alternative way to extract the BAO information, we shall apply it to the DES Y3 samples. This can serve as important cross check as different statistics have different sensitivities to the potential systematics. Some preliminary results on ξ_p have been presented in Abbott et al. (2021) on the BAO samples, but a detailed quantitative analysis is still lacking. This is particularly interesting, as the D_M measurement in Abbott et al. (2021) shows stimulating deviation from the Planck results. Measuring D_M using alternative statistics on various samples helps establish the results.

ACKNOWLEDGMENTS

We thank Fengji Hou for useful discussions. K.C.C. acknowledges the support from the National Science Foundation of China under the grant 11873102, the science research grants from the China Manned Space Project with NO.CMS-CSST-2021-B01, and the Science and Technology Program of Guangzhou, China (No. 202002030360). S.A. is supported by the MICUES project, funded by the EU H2020 Marie Skłodowska-Curie Actions grant agreement no. 713366 (InterTalentum UAM) and “EU-HORIZON-2020-776247 Enabling Weak Lensing Cosmology (EWC)”.

Funding for the DES Projects has been provided by the U.S. Department of Energy, the U.S. National Science Foundation, the Ministry of Science and Education of Spain, the Science and Technology Facilities Council of the United Kingdom, the Higher Education Funding Council for England, the National Center for Supercomputing Applications at the University of Illinois at Urbana-Champaign, the Kavli Institute of Cosmological Physics at the University of Chicago, the Center for Cosmology and Astro-Particle Physics at the Ohio State University, the Mitchell Institute for Fundamental Physics and Astronomy at Texas A&M University, Financiadora de Estudos e Projetos, Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro, Conselho Nacional de Desenvolvimento Científico e Tecnológico and the Ministério da Ciência, Tecnologia

e Inovação, the Deutsche Forschungsgemeinschaft and the Collaborating Institutions in the Dark Energy Survey.

The Collaborating Institutions are Argonne National Laboratory, the University of California at Santa Cruz, the University of Cambridge, Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas-Madrid, the University of Chicago, University College London, the DES-Brazil Consortium, the University of Edinburgh, the Eidgenössische Technische Hochschule (ETH) Zürich, Fermi National Accelerator Laboratory, the University of Illinois at Urbana-Champaign, the Institut de Ciències de l'Espai (IEEC/CSIC), the Institut de Física d'Altes Energies, Lawrence Berkeley National Laboratory, the Ludwig-Maximilians Universität München and the associated Excellence Cluster Universe, the University of Michigan, NSF's NOIRLab, the University of Nottingham, The Ohio State University, the University of Pennsylvania, the University of Portsmouth, SLAC National Accelerator Laboratory, Stanford University, the University of Sussex, Texas A&M University, and the OzDES Membership Consortium.

Based in part on observations at Cerro Tololo Inter-American Observatory at NSF's NOIRLab (NOIRLab Prop. ID 2012B-0001; PI: J. Frieman), which is managed by the Association of Universities for Research in Astronomy (AURA) under a cooperative agreement with the National Science Foundation.

The DES data management system is supported by the National Science Foundation under Grant Numbers AST-1138766 and AST-1536171. The DES participants from Spanish institutions are partially supported by MICINN under grants ESP2017-89838, PGC2018-094773, PGC2018-102021, SEV-2016-0588, SEV-2016-0597, and MDM-2015-0509, some of which include ERDF funds from the European Union. IFAE is partially funded by the CERCA program of the Generalitat de Catalunya. Research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Program (FP7/2007-2013) including ERC grant agreements 240672, 291329, and 306478. We acknowledge support from the Brazilian Instituto Nacional de Ciência e Tecnologia (INCT) do e-Universo (CNPq grant 465376/2014-2).

This manuscript has been authored by Fermi Research Alliance, LLC under Contract No. DE-AC02-07CH11359 with the U.S. Department of Energy, Office of Science, Office of High Energy Physics.

DATA AVAILABILITY

The data underlying this article will be shared on reasonable request to the corresponding author.

REFERENCES

- Abbott T., et al., 2019, *Mon. Not. Roy. Astron. Soc.*, 483, 4866
 Abbott T., et al., 2021 ([arXiv:2107.04646](https://arxiv.org/abs/2107.04646))
 Alam S., et al., 2017, *MNRAS*, 470, 2617
 Alam S., et al., 2020 ([arXiv:2007.08991](https://arxiv.org/abs/2007.08991))
 Alonso D., 2012 ([arXiv:1210.1833](https://arxiv.org/abs/1210.1833))
 Amendola L., Quercellini C., Giallongo E., 2005, *MNRAS*, 357, 429
 Anderson L., et al., 2012, *MNRAS*, 427, 3435
 Anderson L., et al., 2014, *MNRAS*, 441, 24
 Andrae R., Schulze-Hartung T., Melchior P., 2010 ([arXiv:1012.3754](https://arxiv.org/abs/1012.3754))
 Aubourg É., et al., 2015, *Phys. Rev.*, D92, 123516
 Avila S., et al., 2018, *Mon. Not. Roy. Astron. Soc.*, 479, 94
 Benitez N., 2000, *Astrophys. J.*, 536, 571
 Benitez N., et al., 2009, *Astrophys. J.*, 691, 241
 Bernard C., et al., 2002, *Phys. Rev. D*, 66, 094501
 Beutler F., et al., 2011, *MNRAS*, 416, 3017
 Blake C., Bridle S., 2005, *MNRAS*, 363, 1329
 Blake C., et al., 2012, *MNRAS*, 425, 405
 Bond J. R., Efstathiou G., 1984, *ApJL*, 285, L45
 Bond J. R., Efstathiou G., 1987, *MNRAS*, 226, 655
 Camacho H., et al., 2019, *Mon. Not. Roy. Astron. Soc.*, 487, 3870
 Carnero Rosell A., et al., 2021 ([arXiv:2107.05477](https://arxiv.org/abs/2107.05477))
 Carnero A., Sánchez E., Crocce M., Cabré A., Gaztañaga E., 2012, *MNRAS*, 419, 1689
 Chan K. C., et al., 2018, *Mon. Not. Roy. Astron. Soc.*, 480, 3031
 Chaves-Montero J., Angulo R. E., Hernández-Monteagudo C., 2018, *Mon. Not. Roy. Astron. Soc.*, 477, 3892
 Cole S., Fisher K. B., Weinberg D. H., 1994, *MNRAS*, 267, 785
 Cole S., et al., 2005, *MNRAS*, 362, 505
 Coupon J., et al., 2012, *Astron. Astrophys.*, 542, A5
 Crocce M., Scoccimarro R., 2008, *Phys. Rev. D*, 77, 023533
 Crocce M., Cabré A., Gaztañaga E., 2011, *MNRAS*, 414, 329
 Crocce M., Castander F. J., Gaztañaga E., Fosalba P., Carretero J., 2015, *MNRAS*, 453, 1513
 Crocce M., et al., 2016, *Mon. Not. Roy. Astron. Soc.*, 455, 4301
 Crocce M., et al., 2019, *MNRAS*, 482, 2807
 De Vicente J., Sánchez E., Sevilla-Noarbe I., 2016, *Mon. Not. Roy. Astron. Soc.*, 459, 3078
 Dodelson S., 2003, *Modern cosmology*. Academic Press, Amsterdam
 Drummond I. T., Horgan R. R., 1993, *Phys. Lett. B*, 302, 271
 Eisenstein D. J., et al., 2005, *ApJ*, 633, 560
 Elvin-Poole J., et al., 2018, *Phys. Rev. D*, 98, 042006
 Estrada J., Sefusatti E., Frieman J. A., 2009, *ApJ*, 692, 265
 Ferrero I., et al., 2021 ([arXiv:2107.04602](https://arxiv.org/abs/2107.04602))
 Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, *PASP*, 125, 306
 Fosalba P., Crocce M., Gaztañaga E., Castander F. J., 2015, *MNRAS*, 448, 2987
 Gaztañaga E., Cabre A., Hui L., 2009, *Mon. Not. Roy. Astron. Soc.*, 399, 1663
 Guzzo L., et al., 2014, *A&A*, 566, A108
 Hamilton A. J. S., 1992, *ApJL*, 385, L5
 Hamilton A. J. S., 2000, *MNRAS*, 312, 257
 Hu W., Sugiyama N., 1996, *ApJ*, 471, 542
 Hu W., Sugiyama N., Silk J., 1997, *Nature*, 386, 37
 Hütsi G., 2010, *MNRAS*, 401, 2477
 Ilbert O., et al., 2006, *Astron. Astrophys.*, 457, 841
 Izard A., Crocce M., Fosalba P., 2016, *MNRAS*, 459, 2327
 Kaiser N., 1987, *MNRAS*, 227, 1
 Kazin E. A., et al., 2014, *MNRAS*, 441, 3524
 Kilcup G., 1994, *Nucl. Phys. B Proc. Suppl.*, 34, 350
 Landy S. D., Szalay A. S., 1993, *ApJ*, 412, 64
 Lewis A., Challinor A., 2007, *Phys. Rev. D*, 76, 083005
 Limber D. N., 1953, *ApJ*, 117, 134
 LoVerde M., Afshordi N., 2008, *Phys. Rev. D*, 78, 123506
 Michael C., 1994, *Phys. Rev. D*, 49, 2616
 Montanari F., Durrer R., 2012, *Phys. Rev. D*, 86, 063503
 Nicola A., et al., 2020, *JCAP*, 03, 044
 Padmanabhan N., White M., 2009, *Phys. Rev. D*, 80, 063508
 Padmanabhan N., et al., 2007, *MNRAS*, 378, 852
 Peebles P. J. E., 1980, *The Large-Scale Structure of the Universe*. Princeton University Press, New Jersey
 Peebles P. J. E., Yu J. T., 1970, *ApJ*, 162, 815
 Percival W. J., et al., 2010, *MNRAS*, 401, 2148

- Planck Collaboration et al., 2020, *A&A*, **641**, A6
- Press W. H., Teukolsky S. A., Vetterling W. T., Flannery B. P., 2007, *Numerical Recipes 3rd Edition: The Art of Scientific Computing*, 3 edn. Cambridge University Press, New York, NY, USA
- Ross A. J., Samushia L., Howlett C., Percival W. J., Burden A., Manera M., 2015, *MNRAS*, **449**, 835
- Ross A. J., et al., 2017, *Mon. Not. Roy. Astron. Soc.*, **472**, 4456
- Rozo E., et al., 2016, *MNRAS*, **461**, 1431
- Sadeh I., Abdalla F. B., Lahav O., 2016, *Publ. Astron. Soc. Pac.*, **128**, 104502
- Salazar-Albornoz S., et al., 2017, *MNRAS*, **468**, 2938
- Scott D., Srednicki M., White M., 1994, *ApJ*, **421**, L5
- Seo H.-J., Eisenstein D. J., 2003, *ApJ*, **598**, 720
- Seo H.-J., et al., 2012, *ApJ*, **761**, 13
- Sridhar S., et al., 2020, *ApJ*, **904**, 69
- Sunyaev R. A., Zeldovich Y. B., 1970, *Astrophysics and Space Science*, **7**, 3
- Tassev S., Zaldarriaga M., Eisenstein D. J., 2013, *JCAP*, **2013**, 036
- Thacker B. A., Lepage G. P., 1991, *Phys. Rev. D*, **43**, 196
- Vakili M., et al., 2020 ([arXiv:2008.13154](https://arxiv.org/abs/2008.13154))
- Weinberg D. H., Mortonson M. J., Eisenstein D. J., Hirata C., Riess A. G., Rozo E., 2013, *Phys. Rep.*, **530**, 87
- Xu X., Padmanabhan N., Eisenstein D. J., Mehta K. T., Cuesta A. J., 2012, *MNRAS*, **427**, 2146
- Yoo J., Seljak U., 2015, *MNRAS*, **447**, 1789
- Yoon B., Jang Y.-C., Jung C., Lee W., 2013, *J. Korean Phys. Soc.*, **63**, 145
- Zhan H., Knox L., Tyson J. A., 2009, *ApJ*, **690**, 923
- Zhou R., et al., 2021, *MNRAS*, **501**, 3309
- de Simoni F., et al., 2013, *MNRAS*, **435**, 3017

APPENDIX A: CHECKS ON THE SPECTROSCOPIC SAMPLE

Although our goal is to consistently include photo- z in our modeling, it is helpful to check our methodology on the redshift-space correlation function.

To test our pipeline, we use the real-space and redshift-space correlation function measured from the ICE-COLA mocks. In order to verify our template method, we also compute the standard linear redshift-space 3D correlation function template directly, which can be obtained by replacing ϕ in Eq. (17) by a Dirac delta distribution. The real-space template can be retrieved by further setting the f -terms in Eq. (18) to zero. In Fig. A1, we show the real-space correlation function measurement and the corresponding best-fit template with the bias as the fitting parameter. The real-space template is in good agreement with the measurements for $r \gtrsim 20 \text{ Mpc } h^{-1}$. Note that the BAO feature is stronger in the template as the linear power spectrum is used. The best-fit effective bias is 1.63, which roughly corresponds to the bias value at the effective redshift 0.83 shown in Fig. 6 (or Fig. A2 below). Moreover, we compare the redshift-space correlation measurement with the corresponding template from the standard direct computation. The wedge correlation function is averaged over μ in the same way as ξ_p in the main text. We find that the agreement is mildly worse than the real-space result, and the BAO feature damping appears to be weaker than in the real-space case. In the computation, we have used the best-fit effective bias parameter from the real-space result so as to compare with the template obtained by mapping w to ξ below. We could have used the

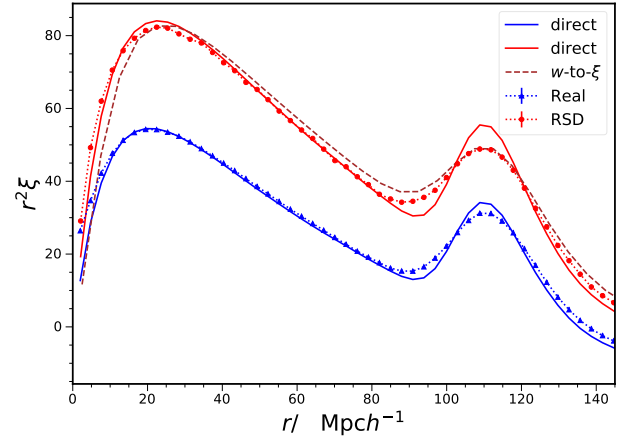


Figure A1. The real-space (blue, triangles) and redshift-space (red, circles) correlation measured from the ICE-COLA mocks. The standard linear templates obtained by direct computation (solid) are compared with the corresponding measurements. We also show the redshift-space template obtained by mapping w to ξ (brown, dashed).

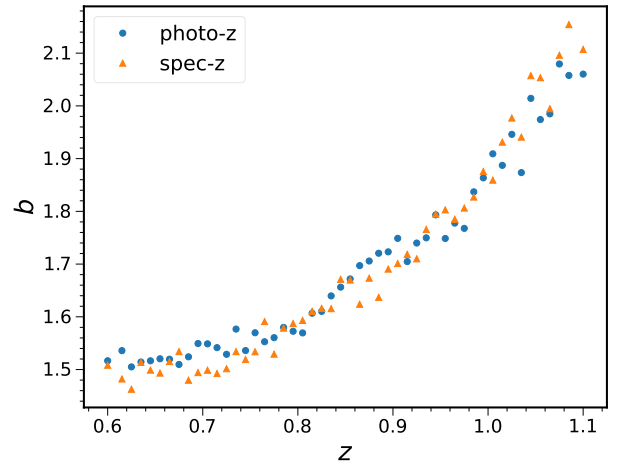


Figure A2. The best-fit bias obtained from the photometric sample (blue, circles) and spectroscopic sample (orange, triangles).

bias parameter obtained by fitting to the redshift-space correlation and the agreement on small scales would be slightly improved.

We also compute the spectroscopic template following the method of mapping w to ξ discussed in the main text. Similar to the analysis for the photo- z sample, we first measure the bias parameters of the redshift-space sample using the angular correlation function. In Fig. A2, we compare the bias parameters measured from the photo- z angular correlation function and the spectroscopic angular correlation. In both cases, we divide the data in the redshift range $[0.6, 1.1]$ into 50 equal redshift bins. For the spectroscopic sample, the window function is a top-hat of width $\Delta z = 0.01$. We

Table A1. The BAO fit results for the spectroscopic sample. For each of the case shown, the first entry denotes the type of template, mock mean or theory, and the second one represents the type of covariance, mock or Gaussian covariance.

Case	$\langle\alpha\rangle$	σ_{std}	σ_{68}	$\langle\sigma_\alpha\rangle$	frac. encl. $\langle\alpha\rangle$	$\langle d_{\text{norm}}\rangle$	$\sigma_{d_{\text{norm}}}$	mean of mocks	$\langle\chi^2\rangle/\text{dof}$
Mock, Mock	1.000	0.010	0.010	0.007	70%	-0.029	1.60	1.000 ± 0.009	24.0/29 (0.83)
Mock, Gaussian	1.000	0.010	0.011	0.005	52%	-0.022	1.43	1.001 ± 0.011	30.9/29 (1.07)
Theory, Mock	1.001	0.013	0.014	0.011	74%	-0.005	1.25	1.001 ± 0.012	36.6/29 (1.26)
Theory, Gaussian	1.002	0.014	0.014	0.007	44%	-0.023	1.38	1.002 ± 0.007	43.1/29 (1.49)

find that the recovered tomographic bin bias parameters are consistent with each other, and this demonstrates that fine photo- z bin measurements can indeed recover the underlying intrinsic bias.

We go on to use the bias parameters to compute the 3D correlation. For the case of spectroscopic redshift, we compute the cross angular correlation function $w_{ij}(\theta)$ with the sampling width $\Delta z = 0.00333$ and $\Delta\theta = 0.02^\circ$. We find that finer sampling in z does not improve the results. Note that in this case we have used the linear power spectrum with BAO damping. The template is in decent agreement with the measurement, but in the intermediate range 30 to 80 $\text{Mpc } h^{-1}$ it is slightly worse than the direct computation result. The difference could be caused by the extra binning and averaging in the w to ξ mapping and these may result in smoothing of the correlation function. However, this is less of an issue in the photo- z case as the intrinsic photo- z mixing window is much stronger.

In Table A1, we show the BAO fit results for the spectroscopic samples. We have shown four scenarios with different template and covariance combinations. The template is either the theory one or the mock mean. The theory template is obtained by mapping w to ξ . As there is small difference between the mock measurement and the theory, we also consider directly using the mean of the mock measurement as the template. In passing, in the main text this possibility is not presented because we found that using the photo- z ξ_p mock mean as the template leads to a non-smooth likelihood. Two covariances are considered, the mock covariance and the Gaussian one. The Gaussian covariance is computed in a similar way as the photo- z case with 50 bins in redshift.

When both the template and the covariance are from mocks, it could be the best case scenario. But we still find that the error estimate is smaller than the spread of the distribution and d_{norm} deviates from Gaussianity significantly. This suggests that the likelihood is non-Gaussian. Moreover, the $\langle\chi^2\rangle/\text{dof}$ is small relative to 1. This could be caused by issues of the mocks such as the overlapping issue discussed in the main text. When the mock covariance is replaced by the Gaussian one, $\langle\sigma_\alpha\rangle$ is further reduced and the fraction enclosing $\langle\alpha\rangle$ is lowered. Interestingly, $\langle\chi^2\rangle/\text{dof}$ is closest to unity in this case. When the theory template is used with the mock covariance, the spread of the best-fit α and $\langle\chi^2\rangle/\text{dof}$ increase. These can be attributed to the difference between the template and the mock mean. Finally, when both the theory template and the Gaussian covariance are used, the $\langle\chi^2\rangle/\text{dof}$ is the highest. We note that it is still smaller than the photo- z case by 0.5.

This paper has been typeset from a $\text{T}_{\text{E}}\text{X}/\text{L}^{\text{A}}\text{T}_{\text{E}}\text{X}$ file prepared by the author.