10.1093/mnras/stx2120 FERMILAB-PUB-17-162-A-AE (accepted)

Optimized Clustering Estimators for BAO Measurements Accounting for Significant Redshift Uncertainty

Ashley J. Ross[⋆][1,2], Nilanjan Banik[3,4], Santiago Avila[2,5,6], Will J. Percival[2], Scott Dodelson[4], Juan Garcia-Bellido[5,6], Martin Crocce[7], Jack Elvin-Poole[8], Tommaso Giannantonio[9,10], Marc Manera[10,11], Ignacio Sevilla-Noarbe[12]

[1]*Center for Cosmology and AstroParticle Physics, The Ohio State University, Columbus, OH 43210, USA*
[2]*Institute of Cosmology & Gravitation, Dennis Sciama Building, University of Portsmouth, Portsmouth, PO1 3FX, UK*
[3] *Department of Physics, University of Florida, Gainesville, Florida 32611, USA*
[4] *Fermi National Accelerator Laboratory, Batavia, Illinois 60510, USA*
[5]*Departamento de Física Teórica, Módulo C-15, Facultad de Ciencias, Universidad Autónoma de Madrid, 28049 Cantoblanco, Madrid, Spain*
[6]*Instituto de Física Teórica, UAM-CSIC, Universidad Autonoma de Madrid, 28049 Cantoblanco, Madrid, Spain*
[7]*Institut de Ciències de l'Espai, IEEC-CSIC, Campus UAB, Carrer de Can Magrans, s/n, 08193 Bellaterra, Barcelona, Spain*
[8]*Jodrell Bank Center for Astrophysics, School of Physics and Astronomy, University of Manchester, Oxford Road, Manchester, M13 9PL, UK*
[9]*Universitäts-Sternwarte München, Fakultät für Physik, Ludwig-Maximilians-Universität München, Scheinerstrasse 1, 81679 München, Germany*
[10]*Kavli Institute for Cosmology, University of Cambridge, Institute of Astronomy, Madingley Road, Cambridge, CB3 0HA, UK*
[11]*Centre for Theoretical Cosmology, Department of Applied Mathematics and Theoretical Physics, Wilberforce Road, Cambridge CB3 0WA, UK*
[12]*Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas (CIEMAT), Madrid, Spain*

ABSTRACT

We determine an optimized clustering statistic to be used for galaxy samples with significant redshift uncertainty, such as those that rely on photometric redshifts. To do so, we study the baryon acoustic oscillation (BAO) information content as a function of the orientation of galaxy clustering modes with respect to their angle to the line-of-sight (LOS). The clustering along the LOS, as observed in a redshift-space with significant redshift uncertainty, has contributions from clustering modes with a range of orientations with respect to the true LOS. For redshift uncertainty $\sigma_z \geqslant 0.02(1 + z)$ we find that while the BAO information is confined to transverse clustering modes in the true space, it is spread nearly evenly in the observed space. Thus, measuring clustering in terms of the projected separation (regardless of the LOS) is an efficient and nearly lossless compression of the signal for $\sigma_z \geqslant 0.02(1 + z)$. For reduced redshift uncertainty, a more careful consideration is required. We then use more than 1700 realizations (combining two separate sets) of galaxy simulations mimicking the Dark Energy Survey Year 1 sample to validate our analytic results and optimized analysis procedure. We find that using the correlation function binned in projected separation, we can achieve uncertainties that are within 10 per cent of those predicted by Fisher matrix forecasts. We predict that DES Y1 should achieve a 5 per cent distance measurement using our optimized methods. We expect the results presented here to be important for any future BAO measurements made using photometric redshift data.

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

1 INTRODUCTION

Measurement of the location of the baryon acoustic oscillation (BAO) feature in the clustering of galaxies has proven to be a robust and precise method to measure the expansion history of the Universe and the properties of dark energy (see, e.g.,Weinberg et al. 2013; Aubourg et al. 2015; Ross et al. 2017; Vargas-Magaña et al. 2016). The most successful results have been obtained by

galaxy redshift surveys such as the SDSS-III Baryon Oscillation Spectroscopic Survey (Eisenstein et al. 2011; Dawson et al. 2013; Alam et al. 2016). Such surveys obtain precise redshift information (∼ 0.1 per cent) and therefore are able to resolve the BAO along and transverse to the line of sight, providing simultaneous measurement of the expansion rate, $H(z)$, and the angular diameter distance, $D_A(z)$.

Multi-band imaging surveys rely on photometric redshifts for radial information. The Dark Energy Survey (DES) and Large Synoptic Survey Telescope are two such surveys and will produce enor-

⋆ Email: ross.1333@osu.edu; Ashley.Jacob.Ross@gmail.com

mous imaging catalogs and will rely on photometric redshifts for their radial information and thus their BAO measurements. The precision that can be achieved for bright, red samples is typically 3 per cent (or better; Rozo et al. 2016); such precision is good enough to localize the galaxies for tomographic studies but removes most of the radial BAO information. The angular diameter distance information is depressed, but still measurable (c.f. Seo & Eisenstein 2003; Blake & Bridle 2005; Zhan et al. 2009 for forecasts and Padmanabhan et al. 2007; Estrada et al. 2009; Hütsi 2010; Sánchez et al. 2011; Seo et al. 2012; Carnero et al. 2012 for measurements).

The above references have focused on using angular clustering measurements in narrow redshift slices. As the available BAO information mainly affords a measurement of D_A, such a focus obtains most of the information. However, it clearly does not use the full information available, as the radial binning blends the data beyond that induced by the redshift error. Furthermore, cross-correlations between redshift slices are often ignored for simplicity. Determining the angular and physical separations of all galaxy pairs and weighting the information should allow more precise BAO measurements. Estrada et al. (2009) present an analysis to do so in the context of photometric galaxy clusters and Chaves-Montero et al. (2016) analyze the results one expects for surveys that obtain redshifts from narrow-band filters. Etherington & Thomas (2015); Etherington et al. (2017) have conducted similar studies on the effect of DES redshift uncertainties on measures of galaxy environment. Here, we primarily focus on the results for Gaussian[1] redshift uncertainties close to 0.03(1+z).

In this study, we develop a framework for an optimal BAO analysis of a large-scale structure sample in the presence of significant redshift uncertainty. The paper is outlined as follows: In Section 2 we study the distribution of BAO information within a given volume and in the presence of redshift uncertainty. This section begins with a discussion of how redshift errors impact the observed signal, then investigates the signal to noise using Fisher matrix formalism, and finally derives optimal weighting as a function of the redshift uncertainty. In Section 3 we describe how we simulate 1700 realizations of the DES Year 1 BAO galaxy sample (mocks) and measure their clustering statistics. In Section 4 we describe how we measure the BAO scale in these mocks and the results of these measurements. We conclude in Section 5 with a discussion summarizing the results and presenting a recommended approach.

The fiducial cosmology we use for this work is flat ΛCDM with $\Omega_{\mathrm{matter}} = 0.25$. Such a low matter density is ruled out by current observational constraints (see, e.g., Planck Collaboration et al. 2015). However, the cosmology we use is matched to that of the MICE (?Fosalba et al. 2015a,b; Carretero et al. 2015) N-Body simulation, which was used to calibrate the mock galaxy samples we employ to test and validate our methodology. Throughout, we will use a fiducial redshift of 0.8, as this is close to the effective redshift we expect for DES Y1 BAO studies.

2   BAO INFORMATION

In this section, we outline the factors that affect the distribution of BAO information. We begin by describing how the BAO signal changes as a function of the orientation of clustering modes with respect to their angle to the line-of-sight (LOS), and demonstrate the

---

[1] See, e.g., Asorey et al. (2016) for a study on effects of non-Gaussian redshift uncertainties.

effect in configuration space in terms of the observed orientation. We then investigate how the signal to noise, given by the Fisher matrix prediction, varies as a function of the true and observed LOS orientation. Here, we consider the 'true' LOS orientation to be that one would observe in the absence of redshift uncertainties (e.g., in an ideal galaxy redshift survey) and the observed LOS orientation being that observed due to redshift uncertainties (e.g., in a photometric redshift galaxy sample). In this formulation, the true (and the observed) LOS orientation is affected by redshift-space distortions. We test different assumptions about the number density, redshift uncertainty, and clustering amplitude within the sample. Finally, we discuss how to weight galaxies in any given volume, given that the number density, redshift uncertainty, and clustering amplitude are all likely to evolve significantly with redshift.

Defining a sphere around a given LOS, equal volume elements are defined by the cosine of the angle to the LOS, μ. Thus, in real, non-evolving, space, information is divided evenly in bins of μ. These facts make the separation, s, (or wave number k) and μ a convenient pair of variables to use when considering the information content of large-scale structure data. The component of any property along the LOS is denoted using $_{||}$ and that transverse is denoted using $_{\perp}$. In what follows, we will use μ to denote the true μ in redshift-space (with no redshift uncertainty) and μ_{obs} the observed μ, which differs from the true μ due to errors in the redshift determination.

2.1   Configuration-space signal with respect to the line-of-sight

Our BAO modeling approach is derived from that of Xu et al. (2012); Anderson et al. (2014); Ross et al. (2017). We will model the BAO signal that one can observe in configuration space rather than in Fourier space, though the linear matter power spectrum, $P_{\mathrm{lin}}(k)$, is our required input. We first obtain $P_{\mathrm{lin}}(k)$ from CAMB[2] (Lewis et al. 2000) and fit for the smooth 'no-wiggle' $P_{\mathrm{nw}}(k)$ via the Eisenstein & Hu (1998) fitting formulae. We account for redshift-space distortions (RSD) and non-linear effects via

$$P(k,\mu) = (1 + \mu^2\beta)^2 \left((P_{\mathrm{lin}} - P_{\mathrm{nw}})e^{-k^2\Sigma_{\mathrm{nl}}^2} + P_{\mathrm{nw}} \right), \quad (1)$$

where $\mu = \cos(\theta_{\mathrm{LOS}}) = k_{||}/k$,

$$\Sigma_{\mathrm{nl}}^2 = (1 - \mu^2)\Sigma_{\perp}^2/2 + \mu^2\Sigma_{||}^2/2, \quad (2)$$

and $\beta \equiv f/b$. This factor is set based on the galaxy bias, b, and effective redshift of the sample we are modeling, with f defined as the logarithmic derivative of the growth factor with respect to the scale factor. The factor $(1 + \beta\mu^2)^2$ is the 'Kaiser boost' (Kaiser 1987), which accounts for linear-theory RSD. The BAO damping parameters are fixed at $\Sigma_{||} = 10h^{-1}\mathrm{Mpc}$ and $\Sigma_{\perp} = 6h^{-1}\mathrm{Mpc}$, as were used in Ross et al. (2017) and are close to the expected values at $z = 0.6$ given by Seo & Eisenstein (2007); Seo et al. (2015). We study clustering at redshift ~0.8, but our conclusions are not sensitive to the precise Σ values and these values produce a good match between theory and simulation, as shown in Section 3.2.

Given $P(k,\mu)$, we determine the multipole moments

$$P_\ell(k) = \frac{2\ell + 1}{2} \int_{-1}^1 d\mu P(k,\mu) L_\ell(\mu), \quad (3)$$

---

[2] camb.info
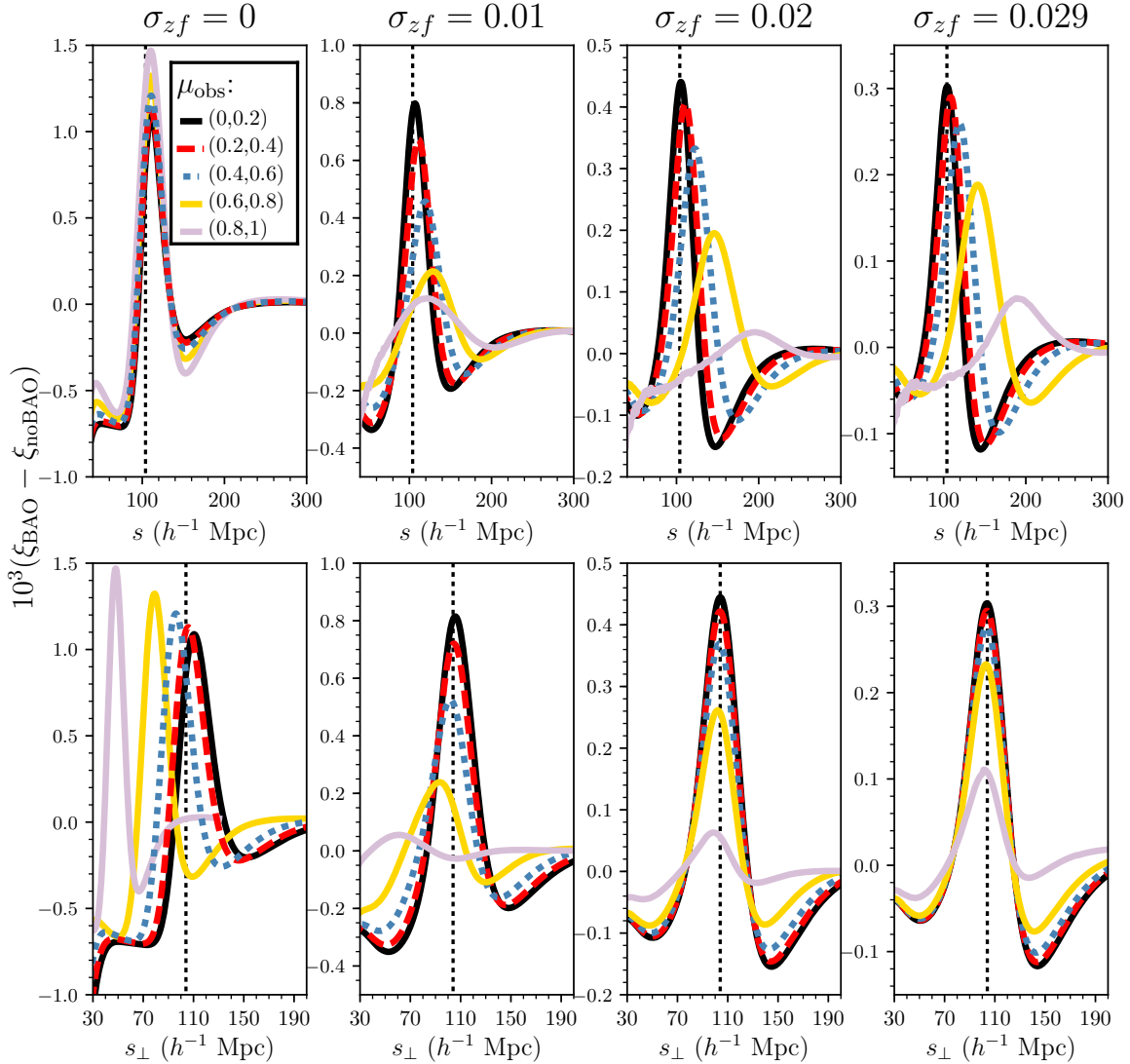
Figure 1. The top panels display the predicted BAO signal, separated by orientation of galaxy pairs with respect to the line of sight. This is orientation is parameterized by $\mu_{\rm obs}$, the cosine of the observed angle a pair of galaxies makes with respect to the line of sight (which differs from the true μ due to redshift errors). We assume a sample of galaxies at $z = 0.8$, and a bias of 1.8, with $\sigma_{zf} = \sigma_z/(1+z)$ increasing from 0 to 0.029 from left to right. In order to isolate the BAO signal, a model constructed from a smooth power spectrum has been subtracted. The bottom panels display the same information, but against the transverse separation, $s_\perp = s\sqrt{1 - \mu_{\rm obs}^2}$.

where $L_\ell(\mu)$ are Legendre polynomials. These are transformed to moments of the correlation function, ξ_ℓ, via

$$\xi_\ell(s) = \frac{i^\ell}{2\pi^2} \int dk k^2 P_\ell(k) j_\ell(ks). \qquad (4)$$

We then use

$$\xi(s,\mu) = \sum_\ell \xi_\ell(s) L_\ell(\mu). \qquad (5)$$

In order to create 'no BAO' models, we set $P_{\rm lin} = P_{\rm nw}$ in Eq. 1.

We can thus isolate the BAO signal by using $\xi - \xi_{\rm noBAO}$ (as in, e.g., figure 3 of Alam et al. 2016).

The preceding paragraph described the standard approach to BAO modeling, generally applied to the analysis of spectroscopic galaxy samples. We now extend these results under the assumption of significant redshift uncertainty in order to derive photometric redshift space clustering $\xi_{\rm phot}(s, \mu_{\rm obs})$. Given $\xi(s, \mu)$, we then convolve the results with the redshift uncertainty to convert them to $\xi_{\rm phot}(s, \mu_{\rm obs})$. In practice, this implies properly averaging over the probability distributions for $s_{\rm true}$ and $\mu_{\rm true}$ given the s and $\mu_{\rm obs}$

observed in photometric redshift space. We use slightly simplified approach and consider all pairs to be centered on some effective redshift, $z_{\rm eff}$ (for which we use 0.8). The uncertainty on the separation in redshift of a pair of galaxies is $\sqrt{2}\sigma_z$; this provides a Gaussian probability distribution, $G(z)$, (centered on $z_{\rm eff}$) for a redshift representing the change in redshift separation between the observed and true separations. This change in redshift separation is converted to a change in radial separation given the difference between comoving distance to $z_{\rm eff}$ and to z sampled from $G(z)$. At every observed $s, \mu_{\rm obs}$, we thus take the weighted mean of[3] $\xi(s_{\rm true}, \mu_{\rm true})$, evaluated using Eq. 5, assuming the observed $s, \mu_{\rm obs}$ are centered on $z_{\rm eff}$. In equation form

$$\xi_{\rm phot}(s, \mu_{\rm obs}) = \int {\rm d}z G(z)\xi(s_{\rm true}, \mu_{\rm true}), \qquad (6)$$

where $G(z)$ is the aforementioned Gaussian likelihood of width $\sqrt{2}\sigma_z$ centered on $z_{\rm eff}$. Using $\chi(z)$ as the comoving distance to redshift z, one can derive,

$$s_{\rm true}(z, s, \mu_{\rm obs}) = \left([\mu_{\rm obs}s + \chi(z_{\rm eff}) - \chi(z)]^2 + (1 - \mu_{\rm true}^2)s^2\right)^{1/2},$$
$$(7)$$

$$\mu_{\rm true} = (\mu_{\rm obs}s + \chi[z_{\rm eff}] - \chi[z]) / s_{\rm true} \qquad (8)$$

and thus evaluate Eq. 6. Evolution in the redshift uncertainty can be taken into account by evaluating Eq. 6 for each redshift uncertainty and taking the appropriately weighted mean of the results. Non-Gaussian redshift uncertainties can potentially be modeled through modifications in $G(z)$, but we leave this for future study.

We use the above to take averages over any given observed $\mu_{\rm obs}$ window to create any particular template:

$$\xi_{\rm phot,W}(s) = \int_0^1 d\mu_{\rm obs}W(\mu_{\rm obs})\xi_{\rm phot}(s, \mu_{\rm obs}). \qquad (9)$$

The $W(\mu_{\rm obs})$ should be normalized to integrate to 1, e.g., we will consider 'wedges' (Kazin et al. 2012, 2013) where $W(\mu_{\rm obs}) = 1/(\mu_{\rm obs,\ max} - \mu_{\rm obs,\ min})$ within a given $\mu_{\rm obs}$ range and 0 outside of it.

In Fig. 1, we display $\xi_{\rm phot,W}(s) - \xi_{\rm phot,W,noBAO}(s)$ for models with redshift uncertainty increasing in panels from left to right. The left-most panels display the results assuming no redshift uncertainty, as would be appropriate for a spectroscopic survey. Each panel shows the results for five $\mu_{\rm obs}$ bins of thickness $\Delta\mu_{\rm obs} = 0.2$. The top panels show the model as a function of redshift-space separation, s. One can observe that in all cases with redshift uncertinaty the BAO signal is diluted as $\mu_{\rm obs}$ increases. This dilution relative to the $\mu_{\rm obs} < 0.2$ case becomes greater as the redshift uncertainty decreases (e.g., compare the heights of the yellow curves representing the $0.6 < \mu_{\rm obs} < 0.8$ signal). This is due to the fact that in all cases the majority of the information is at low true μ (as explained in the following subsection), but when the redshift uncertainty is greater, more of the information is transferred to high observed $\mu_{\rm obs}$. For the same reason, the BAO peak also appears at greater observed s at greater $\mu_{\rm obs}$. This shift is more pronounced when the redshift uncertainty is greater. In the case with no redshift uncertainty, the signal is greatest at high μ due to the 'Kaiser boost' from RSD.

The bottom panels of Fig. 1 show the results in terms of the transverse component of the redshift-space separation, $s_\perp$. One can see that for $\mu_{\rm obs} < 0.8$ and $\sigma_z/(1+z) \geqslant 0.02$, the BAO signal is at

a nearly constant $s_\perp$. For $\mu_{\rm obs} > 0.8$, the signal strength is greatly diminished. This suggests that for $\sigma_z/(1+z) \geqslant 0.02$, measuring $\xi_{\rm phot}(s_\perp)$ and weighting the signal appropriately with $\mu_{\rm obs}$ should obtain a nearly optimal clustering estimator to use for BAO distance measurements.

For $\sigma_z/(1+z) = 0.01$, the peaks do not line up well when displayed as a function of the transverse separation. This is indicative of a significant amount of signal coming from radial clustering information. Clearly, at this level of redshift uncertainty, the optimal separation to use and the interpretation of results in terms of D_A and H measurements needs to be carefully considered.

2.2   Signal to noise with respect to the line-of-sight

We now use Fisher matrix formalism in order to investigate the distribution of BAO signal to noise. In this section, we employ the Seo & Eisenstein (2007) Fisher matrix techniques and work in Fourier space (though we will later return back to configuration space). Given the linear matter power spectrum $P_{\rm lin}(k)$ (implicitly at redshift, z, with no RSD, as in Eq. 1), the relative amount of BAO information, F, as a function of μ can be modeled as depending on galaxy sample's bias, b, number density, n, and redshift uncertainty, σ_z. Suppressing factors (e.g., survey volume) that do not affect the distribution of information, the relevant equation from Seo & Eisenstein (2007) is

$$F(\mu) \propto \int dk \frac{k^2 \exp[-2(k\Sigma_s)^{1.4}]\exp[-k^2(1-\mu^2)\Sigma_\perp^2 - k^2\mu^2\Sigma_{||}^2]}{(b^2 P_{\rm lin}(k)/P_{0.2} + 1/[nP_{0.2}R(k,\mu)])^2},$$
$$(10)$$

where Σ_s is $1/k_{\rm Silk}$ given in Eisenstein & Hu (1998); which accounts for 'Silk damping' (Silk 1968) in the BAO signature produced just after recombination. We use $\Sigma_s = 8.38h^{-1}$Mpc, the value Seo & Eisenstein (2007) use for the WMAP3 cosmology, which is a close match to our fiducial cosmology.[4] $P_{0.2}$ is the value of the power spectrum at $k = 0.2\ h$Mpc^{-1}, which includes the galaxy bias, so $b^2 P_{\rm lin}(k)/P_{0.2} = b^2 P_{\rm lin}(k)/b^2 P_{\rm lin}(0.2) = P_{\rm lin}(k)/P_{\rm lin}(0.2)$ independent of the galaxy bias.

The modulation of the power spectrum amplitude due to RSD and redshift uncertainty enters in the $R(k,\mu)$ term:

$$R(k,\mu) \equiv (1 + \beta\mu^2)^2 \exp(-k^2\mu^2\Sigma_z^2). \qquad (11)$$

The factor $(1 + \beta\mu^2)^2$ is the same 'Kaiser boost' (Kaiser 1987) as in Eq. 1. The redshift uncertainty, σ_z, is accounted for with the second factor, with $\Sigma_z = \frac{c}{H(z)}\frac{\sigma_z}{(1+z)}$, where c is the speed of light, and $H(z)$ is the Hubble parameter at redshift z. The factor in the numerator of Eq. 10 $\exp[-k^2(1-\mu^2)\Sigma_\perp^2 - k^2\mu^2\Sigma_{||}^2]$ adds an additional LOS dependence. This factor accounts for damping of the BAO feature due to non-linear structure growth. Again following Seo & Eisenstein (2007), but rearranging the factors, we use $\Sigma_\perp = 8.3\sigma_8 D(z)h^{-1}$Mpc and $\Sigma_{||} = (1 + f)\Sigma_\perp$, where $D(z)$ is the linear growth factor, normalized to be 1 at $z = 0$.

There are thus three separate factors that are a function of μ. Two are important even if there is no redshift uncertainty. The amplitude of the power spectrum is increased at high μ by RSD, thereby boosting the signal and thus the BAO information, while the BAO amplitude is damped to a greater extent at high μ, thereby decreasing the signal and the amount of BAO information. These two factors are evident in Eqs. 1 and 2. The degree to which one effect dominates over the other depends on the number density of the

---

[3] In order to be explicit, we use $\mu_{\rm true}$ here.

[4] Throughout, when using Seo & Eisenstein 2007 Fisher matrix forecasts, we will use the WMAP3 parameterization.
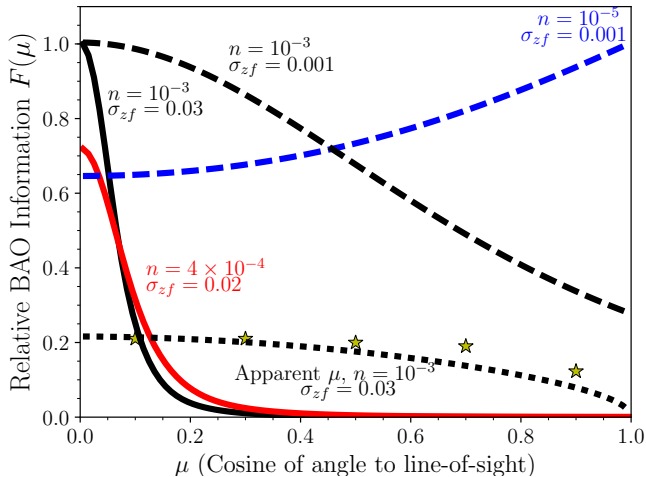
Figure 2. The relative amount of BAO information as a function of μ, the cosine of the angle to the line of sight, for a sample of galaxies at $z = 0.8$, with properties as labeled. We use $\sigma_{zf} = \sigma_z/(1+z)$ and the number density n is in units $h^3\mathrm{Mpc}^{-3}$. All cases use a galaxy bias of $b = 1.8$, except for the solid red curve, which assumes $b = 2$. The dashed curves represent redshift precision achievable from spectroscopic galaxy surveys, while the solid curves show results achievable from multi-band imaging surveys, such as the Dark Energy Survey. All results are in terms of the true μ, except for the dotted curve, which displays the information in terms of observed μ. Results are normalized as follows: The solid black and red curves and the black dashed curve are all divided by the same factor; thus, they are directly comparable. The blue dashed curve is normalized such that it has the same maximum value as the black dashed curve; it thus illustrates how the relationship between BAO information and μ changes with number density. The black dotted curve displays the same information as the black solid line, but with respect to observed μ; the area under the apparent μ curve is normalized to be twice that of the solid black curve (for legibility). The yellow stars represent what we find for mock galaxy samples, normalized so that the results match the dotted curve at $\mu = 0.1$.

sample in question, which can be understood given the expectation for the variance of the power spectrum (Feldman et al. 1994)

$$\sigma_P^2 \propto (P + 1/n)^2. \qquad (12)$$

In the low number density limit, boosting the overall amplitude of the power spectrum is most important, while in the high number density limit, the fractional uncertainty on the measured power, σ_P/P, is constant and thus it is only the modulation of the BAO amplitude that matters. The competing effects are illustrated in Fig. 2. This figure displays the relative BAO information as a function of μ for multiple scenarios. The black and blue dashed curves show identical cases, except that the black curve is for a number density of $10^{-3}h^3\mathrm{Mpc}^{-3}$ and the blue curve is for a number density of $10^{-5}h^3\mathrm{Mpc}^{-3}$ (the dashed curves have been normalized so that the maximum value is the same for each, otherwise the $n = 10^{-5}h^3\mathrm{Mpc}^{-3}$ would have considerably smaller amplitude at all scales).

The third factor that modulates the BAO information as a function of μ is the redshift uncertainty. In Fig. 2, we show the relative BAO information as a function of μ for two cases. The solid black curve displays the case where the number density is $10^{-3}h^3\mathrm{Mpc}^{-3}$, the galaxy bias is 1.8, and the redshift uncertainty is $\sigma_z/(1+z) = 0.03$. These characteristics are similar to a sample

of galaxies selected from DES Y1 data to be optimal for BAO measurements. The solid red curve displays the case where the number density is $4 \times 10^{-4}h^3\mathrm{Mpc}^{-3}$, the galaxy bias is 2, and the redshift uncertainty is $\sigma_z/(1+z) = 0.02$. These characteristics are similar to the DES Y1 high luminosity 'redmagic' (Rozo et al. 2016) sample. We have normalized the black and red curves by the same factor (so that the black curve has a maximum value of 1) and thus the BAO information for these cases can be directly compared. The total BAO information (the area under each curve) matches to within 1 per cent.

For each case representing DES Y1 data, the vast majority of the information is at low μ; for $\sigma_z/(1+z) = 0.03$, half of the information is reached at $\mu = 0.061$ while for $\sigma_z/(1+z) = 0.02$, half of the information is reached at $\mu = 0.08$. The relative amount of information in D_A compared to H can be determined by integrating Eq. 10 with respect to μ and weighting by a factor of μ^2 for H and $(1 - \mu^2)$ for D_A (Seo & Eisenstein 2007). There is more than a factor of 100 times more information in D_A than H in the $\sigma_z/(1+z) = 0.03$ example, to be compared to a factor of 8/3 for the spherically symmetric case with no redshift uncertainty (see, e.g., Seo & Eisenstein 2007; Ross et al. 2015). Put differently, using equations 9 and 10 of Ross et al. (2015) and the $\sigma_z/(1+z) = 0.03$ curve (solid black) in Fig. 2 as their $F(\mu)$, the power-law indices are 0.99 for D_A and 0.01 for H.

Importantly, all curves on Fig. 2 *except* for the dotted curve display the information in terms of the *true* μ distribution. This is implicit in the Fisher matrix formalism; power in the line of sight clustering modes is removed and this removes the information. However, this is the information with respect to the true orientation of the galaxies on the sky (i.e., assuming zero redshift uncertainty). The orientation we observe is in fact (strongly) affected by the redshift uncertainty, and we presented the effect of this on the measured signal in the previous subsection. In order to simulate how the signal to noise is distributed, we switch back to configuration space and determine the mapping between true and observed μ for galaxy pairs with separation $r = 100h^{-1}\mathrm{Mpc}$; we do this in the same manner as in Eq. 6. For each observed μ, we have a probability distribution in true μ, given the probability distribution for true redshift separation, itself given by the pair of observed redshifts and their uncertainties. We show this PDF for three observed μ cases in Fig. 3, all for a sample with $\sigma_z = 0.054$, centered at $z = 0.8$ (so at $z = 0.8$, $\sigma_z = 0.03[1 + z]$). One can see that for all observed μ, the bulk of the pairs will be at high true μ, but also a significant number will be from true $\mu \sim 0$, where the majority of the BAO information exists.

The dotted black curve in Fig. 2 thus displays the $F(\mu)$ relative BAO information in the solid black curve, but in terms of *observed* μ, accounting for the redshift uncertainty, using the mapping illustrated in Fig. 3. The redshift uncertainty distributes the BAO information over a wide range in observed μ. For example, for $\mu_{\mathrm{obs}} = 0$ the curve peaks at true $\mu = 0.84$, due to the fact that at $z = 0.8$ a redshift uncertainty of $0.03(1 + z)$ implies a mean radial dispersion of $157h^{-1}\mathrm{Mpc}$ (for our fiducial cosmology) and $157/\sqrt{157^2 + 100^2} = 0.84$. The information is displayed so that the area under dotted curve is twice that of the solid curve (purely for legibility; they represent the same total amount of information). However, as the *true* information is represented by the solid curve, the information in the observed μ bins must be highly covariant (i.e., the BAO information in $\mu_{\mathrm{obs}} < 0.2$ will be highly covariant with that in $0.2 < \mu_{\mathrm{obs}} < 0.4$ since it must be coming from similar ranges in true μ). This is a significant difference from the spectroscopic redshift-space scenario, where the information in different
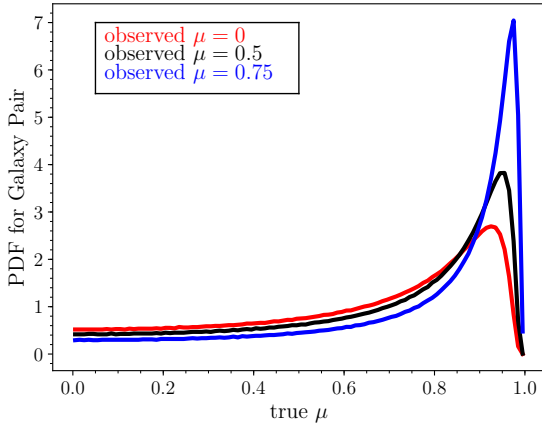
Figure 3. The configuration space mapping between observed and true μ for galaxy pairs of (true) separation $100h^{-1}$Mpc and redshift uncertainty $\sigma_z = 0.03(1 + z)$, centered at $z = 0.8$. Each curve peaks at large μ due to the fact that the observed radial separation is dispersed by $157h^{-1}$Mpc, on average.

μ bins is close to independent (Ross et al. 2015). The correlation arises due to the fact that at a number density of $10^{-3}h^3$Mpc^{-3}, the $\mu \sim 0$ clustering modes are significantly over-sampled, so that even when randomly distributing the information in observed μ, one expects to find a significant correlation in the observed μ bins.

In Ross et al. (2015), methodology was presented for obtaining optimal[5] windows $W(\mu_{\rm obs})$ to be applied to clustering measurements, and modeled as in Eq. 9, in order to best extract the information on D_A and H. Given the lack of information on H for a realistic photometric redshift survey, discovering a single optimal weighting function in $\mu_{\rm obs}$ should provide an optimized BAO distance measure heavily weighted to D_A; the potential gain in defining multiple windows to obtain separated D_A/H constraints is likely negligible. The dotted curve in Fig. 2 represents an optimal $W(\mu_{\rm obs})$ window (ignoring the covariance at different $\mu_{\rm obs}$). However, there is an implicit assumption in Ross et al. (2015) that the BAO feature is nearly constant in μ, in which it is trivial to combine information from any range in μ (e.g., given some optimal weighting function). In the preceding section, we investigated the signal as a function of $\mu_{\rm obs}$, showing the BAO signal to appear at the same $s_\perp$ independent of $\mu_{\rm obs}$. Thus, one expects the optimized clustering estimator for BAO from a photometric redshift survey to be $\int d\mu_{\rm obs} W_{\rm opt}(\mu_{\rm obs})\xi_{\rm phot}(s_\perp, \mu)$, where $W_{\rm opt}(\mu_{\rm obs})$ can be obtained using the methods described in this section.

2.3 Accounting for Variations in Redshift Uncertainty

In the previous sub-section, we outlined how redshift uncertainty affects the distribution of BAO information with respect to the LOS orientation. In this sub-section, we study how one can take variations in redshift uncertainty within a given survey volume into account. Throughout, we assume Gaussian redshift uncertainties; studying the impact of non-Gaussian redshift uncertainties is a possible extension of these results.

[5] Ross et al. (2015) used the variable F for this window, but we use W here so as not to confuse with the information F.

2.3.1   As a function of redshift

Generically, one expects that the number density, clustering amplitude, and redshift uncertainty will evolve with redshift for some selection of galaxies. In order to maximize signal to noise, one therefore wishes to properly weight the contribution from each volume. Given the variance on the power spectrum $\sigma_P^2 \propto (P + 1/n)^2$, the square of the signal to noise ratio[6] , F_P, for each volume is $F_P \propto (nP)^2/(1 + nP)^2$. Assuming the signal, P, is constant, this leads to the k and z dependent inverse-variance 'FKP' weight per galaxy based on Feldman et al. (1994):

$$w_{\rm FKP}(k, z) = \frac{1}{1 + n(z)P_g(k)}. \qquad (13)$$

$P_g(k)$ should be the measured amplitude of the power spectrum of the sample in question. Note that when changing from the inverse variance of the power to a per-galaxy weight, we divide by the number density n in the numerator and we take the square root of F_P due to the fact that the galaxy field is in effect squared when one obtains P.

For a sample with power that evolves as a function of redshift, we write $P_g = P_{\rm lin}b^2D^2$, where $P_{\rm lin}$ is the present day ($z = 0$) linear theory power spectrum. We now have $F_P \propto (nP_{\rm lin}b^2D^2)^2/(1 + nP_{\rm lin}b^2D^2)^2$. This leads to (Percival et al. 2004)

$$w_{\rm FKP}(k, z) = \frac{b(z)D(z)}{1 + n(z)P_{\rm lin}(k)b^2(z)D^2(z)}, \qquad (14)$$

where one factor of bD has been canceled due to the fact that each galaxy contributes its own bD to the signal.

In practice, it is most useful to assign a single weight per galaxy, which requires evaluating $P_{\rm lin}(k)$ at some effective k, $k_{\rm eff}$. To find $k_{\rm eff}$, we take the weighted mean using the Fisher information at each k value as the weight. The redshift uncertainty moves $k_{\rm eff}$ to lower values. For a number density $1.5 \times 10^{-3}h^3$Mpc^{-3} and a bias of 1.8 at $z = 0.65$, $k_{\rm eff}$ moves from $0.16h$Mpc^{-1} to $0.12h$Mpc^{-1} when comparing a sample with no redshift uncertainty to one with $\sigma_z = 0.03(1 + z)$.

In order to further account for redshift uncertainty, we recognize that in Eq. 10, at a given k a change in the redshift uncertainty can be equivalently modeled as a change in the shot-noise term that is strongly dependent on the true μ (but more weakly dependent on the observed μ, based on the results in the previous subsection). Thus, one can define an effective number density, $n_{\rm eff}$:

$$n_{\rm eff}(z) = n(z) \int d\mu R(k_{\rm eff}, \mu). \qquad (15)$$

This implies $F_P \propto (n_{\rm eff}P)^2/(1 + n_{\rm eff}P)^2$. One can define a FKP weight per galaxy accounting for redshift uncertainty by simply substituting $n_{\rm eff}(z)$ for $n(z)$ and using $k_{\rm eff}$ to evaluate $P_{\rm lin}$

$$w_{\rm FKP}(z) = \frac{b(z)D(z)}{1 + n_{\rm eff}(z)P_{\rm lin}(k_{\rm eff})b^2(z)D^2(z)}. \qquad (16)$$

We have verified this weights galaxies approximately correctly by comparing to the forecasted uncertainty in redshift shells, from which a weight per galaxy can also be determined.

As the redshift weights are simply inverse-variance weights, evaluating their expected impact is straight-forward. For instance,

[6] We use F_P to denote the square of the signal to noise ratio in a similar fashion as previous sections, but this now refers to power spectrum amplitude, rather than BAO, information.

if half of the volume is occupied by galaxies have $w_{\rm FKP} = 1$ and the other half $w_{\rm FKP} = 0.5$, the variance *per galaxy* is twice as large for the sample with $w_{\rm FKP} = 0.5$. In this case, one can determine with a simple numerical experiment — sampling Gaussians of the appropriate width — that, assuming each volume has equal variance (so one half of the volume would have twice the total number of galaxies), the improvement in precision achieved from using the inverse-variance weights per galaxy is 5 per cent. If instead half the volume has a per galaxy weight of $w_{\rm FKP} = 0.1$, the improvement is 30 per cent.[7]

2.3.2   At a given redshift

We have developed most of the tools and formalism required in order to probe the question of how to weight galaxies in a sample based on their photometric redshift uncertainty. The previous subsection derives weights assuming the redshift uncertainty can be treated as a constant at a given redshift. However, at a given redshift the photometric redshift uncertainty is likely to vary between galaxies based on their particular color or flux, we might therefore wish to weight galaxies differently based on these properties. This is analogous to how one might weight for varying galaxy bias within a population, as given by equation 28 of Percival et al. (2004).

Unfortunately, the dependence of Eq. 10 on the redshift uncertainty is not as simple as the dependence on the galaxy bias. It is highly dependent on μ; we must consider[8]

$$F_P \propto \frac{\left(nP\exp(-k^2\mu^2\Sigma_z^2)\right)^2}{\left(1 + nP\exp(-k^2\mu^2\Sigma_z^2)\right)^2}. \tag{17}$$

Previously, we integrated over μ in order to obtain $n_{\rm eff}$. It is clear that now the relative weight as a function of redshift uncertainty would change considerably as a function of μ. Thus, any optimal weighting is a function of Σ_z (which depends primarily on σ_z) and μ and it is not obvious how to apply such a weight on a per galaxy basis.[9]

Thus, to move forward we consider the possibility of two populations with different number density and redshift uncertainty. How should the galaxies with the worse redshift uncertainty be weighted to give the optimal result? We are free to arbitrarily apply weights, Z_i, to each population when combining them. Doing so, we calculate the weighted number density, $n(z)_{\sigma_z}$, and a weighted-mean redshift uncertainty[10], $\langle\sigma_z\rangle$, that we can enter into a Fisher matrix forecast (i.e., directly in to Eq. 10). For a general number of populations, the total number density will simply be $n_{\rm tot} = \sum n_i$; since we are considering a single redshift range, we drop the redshift dependence in what follows. If we are to weight each population, this must modulate the effective number density, n_{σ_z}, so that it is lower.

If we are to normalize the weight such that the mean weight is 1, i.e, $\sum_{i=0}^{i=N_{\rm tot}} n_i(\sigma_z)Z_i(\sigma_z) = n_{\rm tot}$, we can work out the effective number density, $n(z)_{\sigma_z}$, by considering Gaussian statistics. If one measures some parameter, X, $N_{\rm tot}$ times using the exact same experiment, then $\langle X \rangle = 1/N_{\rm tot}\sum_{i=0}^{i=N_{\rm tot}} X_i$ and the

uncertainty can be expressed by considering each element in the sum: $\sigma_{\rm tot}^2 = 1/N_{\rm tot}^2 \sum_{i=0}^{i=N_{\rm tot}} \sigma_{\rm d}^2$, given each measurement has the same variance $\sigma_{\rm d}^2$. Inspection reveals this leads to the standard $\sigma_{\rm tot}^{-2} \propto N_{\rm tot}$. We wish to determine $N_{\rm eff}$ for the case where the result of each measurement is weighted differently. With $N_{\rm tot}$ measurements, our estimate is $\langle X_{\rm weighted} \rangle = 1/N_{\rm tot}\sum_{i=0}^{i=N_{\rm tot}} Z_i X_i$ (since we have chosen the weights to be normalized such that $\langle Z \rangle = 1$).

The weighted uncertainty can be similarly written $\sigma_{\rm tot,weighted}^2 = 1/N_{\rm tot}^2 \sum_{i=0}^{i=N_{\rm tot}} (Z_i \sigma_d)^2$. Defining $N_{\rm eff}$ via $\frac{N_{\rm eff}}{N_{\rm tot}} \equiv \left(\frac{\sigma_{\rm tot}}{\sigma_{\rm tot,weighted}}\right)^2$ we obtain

$$\frac{N_{\rm eff}}{N_{\rm tot}} = \frac{\sum_{i=0}^{i=N_{\rm tot}} \sigma_d^2}{\sum_{i=0}^{i=N_{\rm tot}} (Z_i\sigma_d)^2}, \tag{18}$$

which we can divide σ_d out of to yield

$$\frac{N_{\rm eff}}{N_{\rm tot}} = \frac{N_{\rm tot}}{\sum_{i=0}^{i=N_{\rm tot}} Z_i^2}. \tag{19}$$

We take this into account to obtain our effective number density, $n(z)_{\sigma_z}$, which replaces $N_{\rm eff}$. Instead of $N_{\rm tot}$ samples, we have $N_{\rm tot}$ different weights Z being applied to portions of the sample with different redshift uncertainty, each with a number density $n(\sigma_z)$. We find

$$n(z)_{\sigma_z} = n_{\rm tot}(z)^2 / \sum_{i=0}^{i=N_{\rm tot}} n_i(\sigma_z)Z_i^2(\sigma_z). \tag{20}$$

The redshift uncertainty is simply the mean weighted redshift uncertainty

$$\langle\sigma_z\rangle = \sum_{i=0}^{i=N_{\rm tot}} Z_i(\sigma_z)\sigma_{z,i}n_i(\sigma_z) / \sum_{i=0}^{i=N_{\rm tot}} Z_i(\sigma_z)n_i(\sigma_z). \tag{21}$$

We start with a simple example. We consider a total sample with a number density $2 \times 10^{-3} h^3{\rm Mpc}^{-3}$, evenly divided between galaxies with $\sigma_z/(1 + z) = 0.03$ and $\sigma_z/(1 + z) = 0.05$. We assume a galaxy bias of 1.8 and a redshift of 0.8. For this particular case, we search for the Z_2/Z_1 that minimizes the forecasted BAO uncertainty and find that a relative weighting of 0.44 provides the optimal result. This optimal result is, however, only a 3 per cent improvement over the case where only the galaxies with $\sigma_z = 0.03(1 + z)$ are used. If both galaxy populations were to be used without any weighting, the weighted combination provides a 2 per cent improvement. If instead, we consider the case where the number density of the galaxies with $\sigma_z = 0.03(1 + z)$ is $10^{-4} h^3{\rm Mpc}^{-3}$, we find the relative weighting factor is instead 0.72 and the weighting allows a 0.4 per cent improvement in the constraints. These results are presented in Table 1, along with a few other examples.

We find that the relative gains from the optimal weighting procedure vary based on σ_{z1}/σ_{z2} and the respective number densities, e.g., the relative gain is the same for $\sigma_{z1} = 0.02, \sigma_{z2} = 0.04$ and $\sigma_{z1} = 0.03, \sigma_{z2} = 0.06$, though the relative weighting Z_2/Z_1 changes. One such example is shown in the top two rows of Table 1. In Fig. 4, we show the percentage improvement achieved by applying optimal weights to each sample as a function of σ_{z1}/σ_{z2}, given different choices for the number density of each sample (defined in this manner, the curves should be symmetric around $\sigma_{z1}/\sigma_{z2} = 1$ and we ignore $\sigma_{z1}/\sigma_{z2} < 1$). We find at most a 2.5 per cent improvement. The black curve turns over, as at a certain redshift uncertainty the improvement becomes relative to the uncertainty obtained from the sample with lower redshift uncertainty, rather than

---

[7]   A notebook with these calculations is here: https://github.com/ashleyjross/LSSanalysis/tree/master/notebooks .
[8]   As in the previous subsection, we are ignoring the Kaiser boost factors.
[9]   A potential option is to consider a $\mu_{\rm eff}$ for each σ_z, but we do not investigate that.
[10]   Note, this is different than the unweighted mean quantity $\bar\sigma_z$ in Table 1.
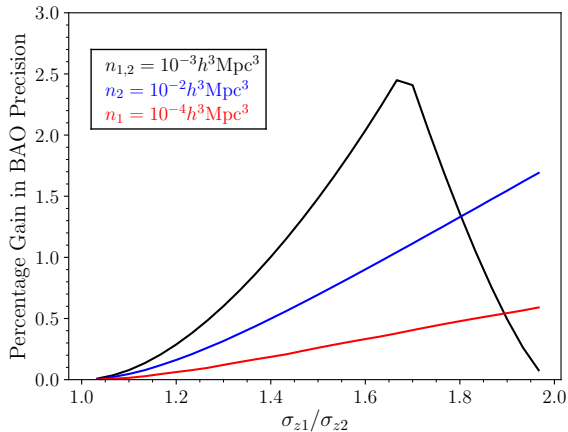
Figure 4. The percentage improvement obtained in the BAO precision as a function of the ratio between the photometric redshift uncertainties. All cases assume $z = 0.8$, $b = 1.8$, $n_1, n_2 = 10^{-3} h^3 \mathrm{Mpc}^{-3}$ unless labeled otherwise. The percentage gain is relative to the optimal choice between either weighting each sample evenly or omitting one sample.

Table 1. Given the properties of two samples occupying the same volume, we find the optimal weighting for each and the forecasted uncertainty for this combination divided by the forecasted uncertainty for each individual sample and their evenly-weighted combination. The σ_z are divided by $1 + z$ and the number densities, n, have units $10^{-3} h^3 \mathrm{Mpc}^{-3}$. The optimal relative weighting of the samples is denoted by Z_2/Z_1. The quantities $\sigma_{\mathrm{opt}}, \sigma_1, \sigma_2, \sigma_{1+2}$ refer to the Fisher matrix forecast for the BAO uncertainty. At most, we find that weighting for the redshift uncertainty improves the results by 2 per cent (the results in the right-most column). The number density used to obtain σ_{1+2} is $n_1 + n_2$ and redshift uncertainty is $\bar{\sigma}_z = 1/2 \sum_i \sigma_{z,i} n_i$; these would be the nominal quantities one uses in the case with no weighting.

$\sigma_{z,1}$	$\sigma_{z,2}$	$\bar{\sigma}_z$	n_1	n_2	Z_2/Z_1	$\frac{\sigma_{opt}}{\sigma_1}$	$\frac{\sigma_{opt}}{\sigma_2}$	$\frac{\sigma_{opt}}{\sigma_{1+2}}$
0.04	0.067	0.054	1	1	0.40	0.97	0.75	0.98
0.03	0.05	0.048	1	10	0.44	0.85	0.96	0.99
0.03	0.05	0.043	0.5	1	0.58	0.87	0.83	0.99
0.03	0.05	0.048	0.1	1	0.72	0.46	0.95	0.996
0.02	0.05	0.047	0.1	1	0.61	0.56	0.94	0.99
0.02	0.03	0.027	0.4	1	0.65	0.81	0.87	0.99

that of the equally-weighted combination. We find no turn-over for the cases where the number densities are not matched, though we have not tested beyond $\sigma_{z1}/\sigma_{z2} > 2$.

The discussion above provides a framework that one could expand upon to obtain an optimal set of weights based on the redshift uncertainty. For the cases considered in Table 1 and Fig. 4, we find that applying optimal weights provides at most a 2.5 per cent improvement over what would be considered the optimal sample selection. This suggests that for the number densities and redshift uncertainties achievable by multi-band imaging surveys such as DES, weighting by individual redshift uncertainties is unlikely to make a substantial impact on the recovered results. These results therefore validate our approach to treat galaxy samples as having a single redshift uncertainty, equal to the mean uncertainty of a given population of galaxies. Our simple tests reveal at most a 2.5 per cent

improvement over this approach (assuming one removes galaxies with redshift uncertainties that degrade the ensemble constraints).

3 MEASURING CLUSTERING IN MOCK GALAXY SAMPLES

In order to test and validate the analytic results we presented in the previous Section, we conduct clustering analyses on mock galaxy catalogs (mocks). In this section, we first describe how we created mocks that simulate DES Y1 data and then describe the clustering we measure for those mocks.
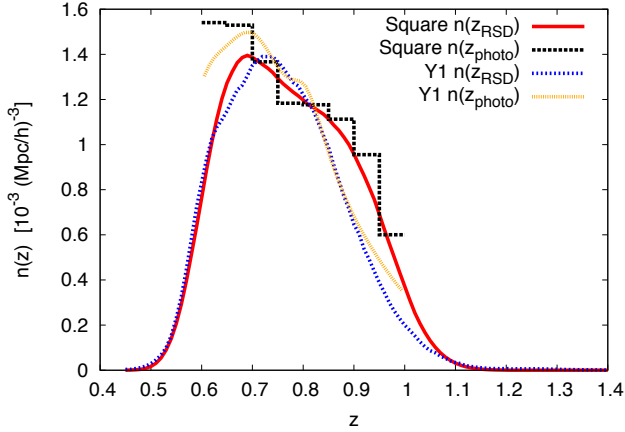
3.1 HALOGEN Mocks

We test and validate our analytical results using two sets of mocks, produced to have similar properties to the DES Y1 BAO sample: 'Square' mocks and 'Y1' mocks. We generate the mocks using HALOGEN (Avila et al. 2015), a method based on an exponential bias model applied to a 2nd-order Lagrangian perturbation theory (2LPT) density field. As a first step in their production, HALOGEN dark matter halo catalogs are produced such that the halo clustering and velocity distributions match those of the MICE (?Fosalba et al. 2015a,b; Carretero et al. 2015) N-Body simulation as a function of both mass and redshift.

We produce lightcone catalogs by super-imposing redshift shells (each of them obtained from a cubic snapshot of length 3072 $h^{-1}\mathrm{Mpc}$) of width $\Delta_z = 0.05$ in the interval $0.6 < z < 1$, and two additional slices from snapshots $z = 0.55$ and $z = 1.05$ to complete the boundaries of the lightcone (and for the Y1 mocks another two at $z = 0.3$ and 1.3). From each set of snapshots we produce eight independent halo lightcone catalogs to a depth of $z = 1.42$ and mass resolution of $M = 2.5 \cdot 10^{12} M_\odot/h$.

Our first set of mocks, denoted 'Square', simulate early estimates of the properties of the DES Y1 BAO sample using 504 halo catalogs. This data sample occupies 1420 deg^2 with a number density that decreased from $3 \times 10^{-3} h^3 \mathrm{Mpc}^{-3}$ at $z = 0.6$ to $8 \times 10^{-4} h^3 \mathrm{Mpc}^{-3}$ at $z = 1$ and redshift uncertainty $\sigma_z = 0.029(1+z)$ (these redshift uncertainties were later found to be optimistic). At the resolution of our halo catalogs, a maximum number density of $2 \times 10^{-3} h^3 \mathrm{Mpc}^{-3}$ is achievable at $z = 0.6$. Thus we increased the area in the footprint of our Square mock samples to $A_{\mathrm{mock}} = 1779$ deg^2. We used the Seo & Eisenstein (2007) Fisher matrix predictions to determine that this is the area required to be able to match the estimate of BAO signal to noise for the mocks and (early estimates of) the data as a function of redshift. Given this expanded area, the signal to noise as a function of redshift was matched by adjusting the number density in intervals of $\Delta_z = 0.05$, obtaining the density displayed in Fig. 5. Halos were then selected by mass, yielding the bias evolution represented in that Fig. 6. For simplicity, we use a *square* mask, selecting galaxies to be at the center of halos in with both ϕ and $\cos(\theta)$ in the interval $[0, 0.7361]$ (with θ, ϕ representing the latitudinal and longitudinal coordinates of an arbitrarily sized sphere). Given its geometry, we refer to this mock sample as the 'Square' sample. The process described above gave us mock galaxy catalogs with angular coordinates and cosmological redshifts. We added the effects of RSD by correcting the redshifts based on the peculiar velocity of each halo. We then added the redshift uncertainty by sampling a Gaussian centered at this redshift of width $\sigma_z = 0.029(1 + z)$.

Our second set of mocks resembles more closely the true nature of the final DES Y1 BAO sample (Crocce et al., in prep.), we

Figure 5. Number density of objects as a function of redshift before and after the application of the photo-z modeling, for both the 504 'Square' and 1200 'Y1' halogen mocks we use.
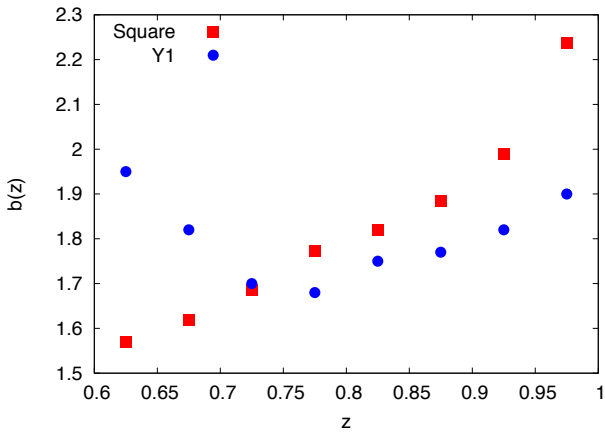
Figure 6. The bias as a function of redshift, for the 504 'Square' and 1200 'Y1' halogen mocks we use. The Y1 mocks are a closer match to the DES Y1 data, but the Square mocks use a simpler methodology and are thus useful for aiding theoretical understanding.

refer to it as the 'Y1' sample. It matches the footprint (with an area of 1420 deg^2), number density (evolving from $1.4 \times 10^{-3} h^3 \mathrm{Mpc}^{-3}$ at $z = 0.6$ to $4 \times 10^{-4} h^3 \mathrm{Mpc}^{-3}$ at $z = 1$), angular clustering, and redshift uncertainty (evolving from $0.028(1 + z)$ to $0.05(1 + z)$). The number density and bias of these mocks are also shown in Fig. 5 and 6. Other novelties in these mocks include the non-Gaussian modelling of the redshift uncertainty and a redshift-evolving Halo Occupation Distribution galaxy model, both fit to the DES Y1 data. In particular, to match our estimates of the DES Y1 dN/dz as a function of photometric redshift, a non-Gaussian PDF was used to assign the photometric redshift of a galaxy given its true redshift. Full details of the mocks and the data are found, respectively, in Avila et al. (in prep.) and Crocce et al. (in prep.). We analyze 1200 mocks for this Y1 sample.

3.2 Clustering Measurements

We measure the correlation functions of our mock galaxy samples, ξ, by converting angles and redshifts to physical distances; this

gives us a three dimensional map in 'photometric redshift space'. In this space, the apparent transverse and radial separations, $s_\perp$ and $s_{||}$, are calculated for pairs of galaxies and/between a synthetic random catalog matching the angular and redshift selections of the mocks. Given (normalized) pair counts, we calculate $\xi_{\mathrm{phot}}(s_\perp, s_{||})$ via Landy & Szalay (1993)

$$\xi_{\mathrm{phot}}(s_\perp, s_{||}) = \frac{DD(s_\perp, s_{||}) - 2DR(s_\perp, s_{||}) + RR(s_\perp, s_{||})}{RR(s_\perp, s_{||})}, \tag{22}$$

where D represents the galaxy sample and R represents the uniform random sample that simulates the selection function of the galaxies. $DD(s_\perp, s_{||})$ thus represent the number of pairs of galaxies with separation $s_\perp$ and $s_{||}$, within some bin tolerance. We use bins of $1h^{-1}\mathrm{Mpc}$. These are narrow enough to allow us to test many treatments with respect to $\mu_{\mathrm{obs}} = s_{||}/\sqrt{s_\perp^2 + s_{||}^2}$.

Our use of the clustering statistic $\xi(s_\perp)$ within some range of μ_{obs} is similar to the commonly used $w_p(r_p)$ statistic. We are using the coordinates $s_\perp, s_{||}$ to have the same meaning as r_p, π in, e.g., Zehavi et al. (2011). Using our convention, their equation 3 is

$$w_p(s_\perp) = 2 \int_0^\infty \mathrm{d}s_{||} \xi(s_\perp, s_{||}). \tag{23}$$

In practice, studies have employed some maximum $s_{||}$ value. Our approach is different in that we consider some window, $W(\mu_{\mathrm{obs}})$

$$\xi_{\mathrm{phot,F}}(s_\perp) = \int_0^1 \mathrm{d}\mu_{\mathrm{obs}} W(\mu_{\mathrm{obs}}) \xi_{\mathrm{phot}}(s_\perp, s_{||}), \tag{24}$$

normalized such that $\int_0^1 \mathrm{d}\mu_{\mathrm{obs}} W(\mu_{\mathrm{obs}}) = 1$. The main practical difference between our approach and that of $w_p(s_\perp)$ is that our results will be binned in terms of μ_{obs} and any cut on μ_{obs} results in a varying maximum $s_{||}$ as a function of $s_\perp$.

In Eq. 16, we defined a weight, w_{FKP}, to be applied as a function of redshift, given a calculation of $n_{\mathrm{eff}}(z)$ and k_{eff}. This weight is meant to properly account for the changing signal to noise in each redshift shell. We apply these weights to the galaxies and random points and thus each paircount is weighted by the multiplication of the two weights representing each pair.

Our $w_{FKP}(z)$ weights were defined as follows. Based on the sample characteristics presented in Section 3.1, we find that k_{eff} is within 10 per cent of $0.12h\mathrm{Mpc}^{-1}$ for each sample of mocks, independent of redshift. We thus use a value $P_0 = 4500h^{-3}\mathrm{Mpc}^3$, which is close to the linear redshift zero matter power spectrum at $k = 0.12h\mathrm{Mpc}^{-1}$ for our fiducial (MICE) cosmology. For the Square mocks, n_{eff} evolves from $1.8 \times 10^{-4}h^3\mathrm{Mpc}^{-3}$ to $6.9 \times 10^{-5}h^3\mathrm{Mpc}^{-3}$. Using Eq. 16, we find the $w_{FKP}(z)$ evolves from 0.82 to 1 for this sample (normalizing so that the maximum weight value is 1). For the Y1 mocks, we follow a similar process and find the weights evolve from 0.71 to 1. We therefore expect the use of these weights to have a minor effect on our results, as the arguments at the end of Section 2.3.1 suggest only a 5 per cent gain in precision when the weights differ by a factor of 2.

Fig. 7 displays the mean of the clustering in the Square mock samples, divided into 0.2 thick bins in μ_{obs} and binned in terms of the transverse separation $s_\perp$. We have subtracted 5×10^{-5} from each set of points and then multiplied by $s_\perp^2$. The error bars represent the standard deviation across the 504 mock realizations, divided by $\sqrt{504}$; i.e., this is the ensemble uncertainty. The curves represent the model described in previous sections, with a bias of 1.76 (consistent with the $b(z)$ given in Fig. 6 and our redshift weighting). The location of the BAO feature remains nearly constant as a function of μ_{obs}; this is because at high μ_{obs}, the vast
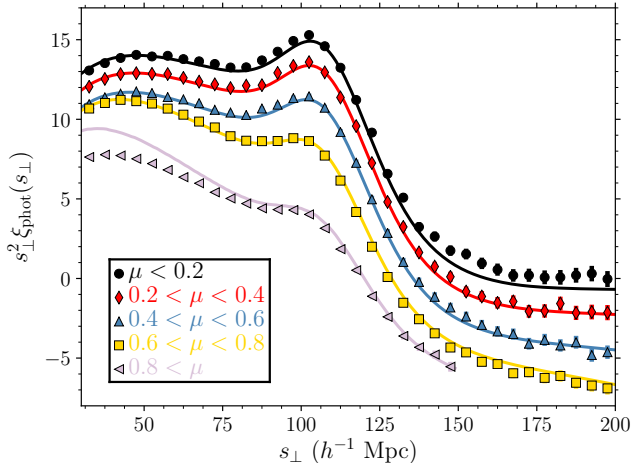
Figure 7. The predicted $\xi_{\text{phot}}(s_\perp)$ signal (solid curves; obtained via Eqs. 6-9) in bins of observed μ (labeled μ in the legend), the cosine of the angle to the line of sight, compared to the mean of 504 'Square' mock realizations. The mock points all have 5×10^{-5} subtracted from them, before multiplying by $s_\perp^2$. The error-bars are the standard deviation of the mocks samples divided by $\sqrt{504}$. The model curves assume a sample centered on $z = 0.8$ with $\sigma_z/(1+z) = 0.029$, and a bias of 1.76, as is approximately the case for the mock samples. The red curve and points are shifted vertically by 1 and the black curve and points by 2, for legibility.

majority of the information is from pairs with *true* μ values that are close to zero. While the BAO feature stays nearly constant (see Fig. 1), the shape changes considerably with μ_{obs}, such that the overall amplitude is decreased with increasing μ_{obs}. For $\mu_{\text{obs}} < 0.8$ and $s_\perp < 140h^{-1}$Mpc, the model curves are a good match to the mock measurements, given the 5×10^{-5} offset.[11] Our pair-counts were calculated for $r_{||} < 200h^{-1}$Mpc and thus there are no results for $\mu_{\text{obs}} > 0.8, s_\perp > 150h^{-1}$Mpc.

The match between the Y1 mocks and our theoretical calculation is not displayed, but is similarly good. The agreement is achieved despite the fact that the Y1 mocks do not assume Gaussian redshift uncertainties, while our modeling does. Further details can be found in Avila et al. (in prep.). For the Y1 mocks, we will compare results to those obtained from the angular correlation function, $w(\theta)$. We calculate $w(\theta)$ as in Eq. 22, except that we bin paircounts by their angular separation, θ. When doing so, we divide the measurements into four redshift bins between $0.6 < z_{\text{phot}} < 1.0$, each with thickness $\Delta_z = 0.1$.

4 BAO INFORMATION IN MOCK SAMPLES

In this section, we compare measurements of the BAO in mock samples for different μ_{obs} windows and test the extent to which the $\int d\mu_{\text{obs}} W_{\text{opt}}(\mu_{\text{obs}})\xi_{\text{phot}}(s_\perp, \mu_{\text{obs}})$ described in Section 2.2 is indeed the optimal estimator for BAO information.

In order to extract the BAO information, we fit to a model (c.f., Xu et al. 2012; Anderson et al. 2014; Ross et al. 2017)

$$\xi_{F,\text{mod}}(s_\perp) = B_F \xi_{F,\text{phot}}(s_\perp \alpha) + A_F(s_\perp) \qquad (25)$$

---

[11] We are unable to determine the reason for the mismatch between the model and mock results for $\mu_{\text{obs}} > 0.8, s_\perp < 80h^{-1}$Mpc, but given the lack of BAO signal for $\mu_{\text{obs}} > 0.8$ this does not impact strongly impact our results.

where $\xi_{F,\text{phot}}$ represents the use of a general μ_{obs} window (i.e., $\xi_{F,\text{phot}} = \int d\mu_{\text{obs}} W_F(\mu_{\text{obs}})\xi_{\text{phot}}(s_\perp, \mu_{\text{obs}})$) and $A_F(s_\perp) = a_{F,1}/s_\perp^2 + a_{F,2}/s_\perp + a_{F,3}$. In each case, the parameter B_F essentially sets the size of the BAO feature in the template. We apply a Gaussian prior of width $\log(B_F) = 0.4$ around the best-fit B_F in the range $50 < s < 80h^{-1}$Mpc with $A_F = 0$. Likelihoods for α are obtained by finding the minimum χ^2 on a grid of α between 0.8 and 1.2, with spacing 0.001, when marginalizing over the other model parameters. This is a close match to the methods used in BOSS (see, e.g., Xu et al. 2012; Anderson et al. 2014; Ross et al. 2017 and references therein). The χ^2 are determined in the standard manner, with the covariance determined from the sample of mocks being tested, i.e.,

$$C(s_1, s_2) = \frac{1}{N_{\text{mock}} - 1} \sum_i \left([\xi(s_1) - \langle\xi(s_1)\rangle][\xi(s_2) - \langle\xi(s_2)\rangle] \right) \qquad (26)$$

We start by using the 504 Square mocks in order to explore how the signal to noise varies as a function of μ_{obs}. While we have fewer Square mocks compared to the Y1 mocks, they have better expected precision and a simple constant $\sigma_z = 0.029(1+z)$, making them better suited for sub-dividing and comparing to theoretical expectation. We use $\xi_{\text{phot}}(s_\perp)$ measurements averaged in bins of 0.2 thickness in μ_{obs} to measure the BAO location and its uncertainty. The results are presented in the top rows of Table 2. The uncertainty we recover from the mean of the mock samples (denoted 'mean' in the table) increases slightly with increasing μ_{obs}, from 0.039 to 0.041 for $\mu_{\text{obs}} < 0.8$ and then sharply to 0.051 for $\mu_{\text{obs}} > 0.8$. These results for the uncertainty from the mean of the mocks are displayed with yellow stars in Fig. 2. One can see that the distribution of information with apparent μ_{obs} is close to our Fisher matrix analysis, except that we find a more gentle decrease in the information recovered as μ_{obs} increased than compared to the prediction.

The $\mu_{\text{obs}} > 0.8$ clustering contains significantly less BAO information than the rest of the μ_{obs} range. Given the lack of constraining power and the increased difficulty in its modeling, we ignore the $\mu_{\text{obs}} > 0.8$ data from further comparisons. Given the covariance between data at different μ_{obs} (described later in this section), we expect this has minimal impact on the BAO results we recover from the mocks.

If we instead look at the mean uncertainty for mocks where we find a 1σ bound in region $0.8 < \alpha < 1.2$, denoted 'samples' in the table, we find similar results. The mean uncertainty increases from 0.039 to 0.041 with increasing μ_{obs}. We also find the standard deviation of the recovered maximum likelihood α values increases and is in the range 0.042 to 0.044. Correspondingly, the fraction of mocks with 1σ bounds also decreases slightly, from 0.98 to 0.96. Including only mocks with 2σ bounds within the region $0.8 < \alpha < 1.2$, the mean uncertainties are, naturally, smaller by ~ 10 per cent. For this case, the standard deviations are closer to the mean uncertainty; this is consistent with the likelihoods for the better-constrained realizations being more Gaussian. Our results from splitting the clustering by μ_{obs} consistently show a slight decrease in BAO information content with increasing μ_{obs} in the range $0 < \mu_{\text{obs}} < 0.8$.

The results are improved when we combine all data with $\mu_{\text{obs}} < 0.8$. These results for the Square mocks are found in the middle rows of Table 2. Given the similarity of the BAO signal shown in the bottom-right panel of Fig. 1, we expect such a compression to preserve the strength of the BAO signal. The uncertainty on the mean of the mock samples decreases to 0.038 and

we find the mean of the uncertainty determined for each individual sample is the same (when considering the 1σ threshold). The standard deviation is 0.042 and the fraction of mocks with a 1σ bound matches that for $\mu_{\rm obs} < 0.2$ (0.98). The dotted line in Fig. 2 displays how we expect BAO information to be distributed in apparent $\mu_{\rm obs}$ and approximately matches what we find for the mocks divided in $\mu_{\rm obs}$ bins. This suggests that we should use that curve as a relative weight to apply as a function of $\mu_{\rm obs}$. We do so and find negligible improvement. For the mean of the mocks, we find a 0.2 per cent improvement in the recovered uncertainty (which is too small to be observable in Table 2). For the mock realizations, the biggest change is a 1 per cent improvement in the standard deviation obtained when limiting to those cases with 1σ bounds within $0.8 < \alpha < 1.2$.

The precision of the results for $\mu_{\rm obs} < 0.8$ are just slightly better than those obtained for more thin slices in $\mu_{\rm obs}$. This is due to fact that the results for different $\mu_{\rm obs}$ slices are highly correlated (as implied by the fact that the information is sharply peaked at low *true* μ). The correlation coefficients vary from 0.77 to 0.7 (these can be compared to correlation coefficients of ~ 0.15 for mocks simulating a spectroscopic redshift sample in Ross et al. 2015). One can construct a covariance matrix for the four $\mu_{\rm obs}$ bins and take the sum of its inverse to obtain a standard deviation from the combination of the $\mu_{\rm obs}$ bins. We use mock realizations with 2σ bounds and find 0.034, which can be compared to 0.036 for our $\mu_{\rm obs} < 0.8$ results. This suggests that $\xi(s_\perp)$ with $\mu_{\rm obs} < 0.8$ is close to an optimal compression of the BAO information distributed in the four $\mu_{\rm obs}$ bins, but does suffer ~ 10 per cent information loss ($[0.036/0.034]^2 = 1.12$).

For the Y1 mocks, we focus on $\mu_{\rm obs} < 0.8$ and weight evenly. The redshift uncertainties are greater for the Y1 mocks, suggesting the information should be spread more evenly in observed $\mu_{\rm obs}$ than for the Square mocks. We use 1200 Y1 realizations to define the covariance matrix. The uncertainty obtained for the mean of the mocks is just greater than five per cent and this is broadly consistent with the results obtained from the individual mock realizations. More than 90(70) per cent of the realizations have a $1(2)\sigma$ bound within $0.8 < \alpha < 1.2$. This suggests there is a good likelihood of obtaining a robust 5 per cent angular diameter distance measurement using BAO in the DES Y1 data.

For Y1 mocks, we can compare to $w(\theta)$ results (we do not have enough Square mocks for a robust covariance matrix for $w(\theta)$). We find a 13 per cent improvement in the uncertainty obtained on the mean when using $\xi_{\rm phot}(s_\perp)$. Here, for $w(\theta)$, we are applying the same BAO model and simply swapping $\xi_{\rm phot}$ & $s_\perp$ for w & θ. A more detailed comparison of methodologies will be presented with the DES collaboration analysis of the Y1 BAO signal. This study will include further tests to optimize each method. In principle, the precision of $w(\theta)$ results should converge to the $\xi_{\rm phot}(s_\perp)$ results as the redshift bin size is narrowed and all cross-correlations between redshift bins are included. The clear advantage of $\xi_{\rm phot}(s_\perp)$ is the smaller size of the data vector, which reduces the noise bias in the inverse covariance matrix (Hartlap et al. 2007; Dodelson & Schneider 2013; Percival et al. 2014).

The mean α values for the Square mocks are slightly biased compared to the expectation of $\alpha = 1$, though the Y1 mocks are less biased. Some of this bias likely comes from non-linear structure, which is expected produce shifts less than 0.005 (c.f., Seo et al. 2008; Padmanabhan et al. 2012). The biases we find for the Square mocks increase as $\mu_{\rm obs}$ gets smaller, which is the opposite as we would naively expect, given the modeling is more simple for $\mu = 0$. Accounting for the shift expected from non-linear structure

Table 2. Statistics for BAO fits on mocks. $\langle\alpha\rangle$ is either the BAO dilation-scale measured from the correlation function averaged over all of the mocks (denoted 'mean'), or the mean of the set of dilation-scales recovered from mocks with $> 1\sigma$ ($> 2\sigma$) BAO detections (denoted 'samples'). $\langle\sigma\rangle$ is the same for the uncertainty obtained from $\Delta\chi^2 = 1$ region. S is the standard deviation of the α recovered from the mock realizations with $> 1\sigma$ ($> 2\sigma$) BAO detections and $f(N_{\rm det})$ is the fraction of realizations with $> 1\sigma$ ($> 2\sigma$) detections. The label w_μ denotes that the results have been weighted in $\mu_{\rm obs}$ following the dotted line in Fig. 2.

case	$\langle\alpha\rangle$	$\langle\sigma\rangle$	S	$f(N_{\rm det})$
Square mocks, $\mu_{\rm obs}$ bins:				
mean, $\mu_{\rm obs} < 0.2$	1.007	0.039	-	-
samples, $\mu_{\rm obs} < 0.2$, 1σ	1.007	0.039	0.042	0.976
samples, $\mu_{\rm obs} < 0.2$, 2σ	1.006	0.037	0.037	0.885
mean, $0.2 < \mu_{\rm obs} < 0.4$, 1σ	1.006	0.040	-	-
samples, $0.2 < \mu_{\rm obs} < 0.4$, 1σ	1.005	0.039	0.042	0.976
samples, $0.2 < \mu_{\rm obs} < 0.4$, 2σ	1.004	0.038	0.038	0.901
mean, $0.4 < \mu_{\rm obs} < 0.6$	1.005	0.041	-	-
samples, $0.4 < \mu_{\rm obs} < 0.6$, 1σ	1.005	0.041	0.044	0.966
samples, $0.4 < \mu_{\rm obs} < 0.6$, 2σ	1.003	0.038	0.038	0.855
mean, $0.6 < \mu_{\rm obs} < 0.8$	1.005	0.041	-	-
samples, $0.6 < \mu_{\rm obs} < 0.8$, 1σ	1.004	0.041	0.044	0.958
samples, $0.6 < \mu_{\rm obs} < 0.8$, 2σ	1.003	0.037	0.039	0.813
mean, $\mu_{\rm obs} > 0.8$	0.993	0.051	-	-
Square mocks, $\mu_{\rm obs} < 0.8$:				
mean, w_μ	1.008	0.038	-	-
samples, w_μ, 1σ	1.008	0.038	0.041	0.976
samples, w_μ, 2σ	1.007	0.035	0.036	0.889
mean	1.008	0.038	-	-
samples, 1σ	1.008	0.038	0.042	0.978
samples, 2σ	1.007	0.035	0.036	0.883
Y1 mocks, $\mu_{\rm obs} < 0.8$:				
mean	1.005	0.053	-	-
samples, 1σ	1.003	0.047	0.053	0.910
samples, 2σ	1.000	0.042	0.044	0.706
mean $w(\theta)$, $\Delta z = 0.1$	1.009	0.061	-	-

growth, these biases are $\sim 0.1\sigma$ and thus not especially concerning for DES Y1. However, future data sets will provide considerably more precise results and will thus require improvements in the modeling if the biases are similar to what we find for the Square mocks. For the Y1 mocks, we observe that the fact that we have modeled the sample assuming Gaussian redshift uncertainties does not significantly bias our results, though it is possible this causes the estimated uncertainties to be smaller than the recovered standard deviations. This will be studied further in a future publication describing the DES Y1 BAO analysis.

We obtain a BAO uncertainty of 0.038 from the Square mocks and 0.053 for the Y1 mocks with $\mu_{\rm obs} < 0.8$. These can be compared to Fisher matrix predictions of 0.031 and 0.038. Our fiducial cosmology closely matches that used to define the Seo & Eisenstein (2007) Fisher matrix forecasts. As explained in Seo & Eisenstein (2007), this Fisher matrix prediction neglects to include additional BAO smearing due to projection effects, whose effect is described in Seo & Eisenstein (2003). Essentially, projection effects cause the BAO feature at different redshift to be combined and thus diluted. Emulating Seo & Eisenstein (2003), we estimate the projection effect should degrade our uncertainties by an additional ~ 10 per cent and we obtain forecasts of 0.033 for the Square mocks and 0.042 for the Y1 mocks. The values we recover from the mean

of the mocks are affected somewhat by the likelihoods being non-Gaussian. We can account for this by multiplying the covariance matrix by an arbitrary factor, f, small enough so that the recovered $\sigma/\sqrt{f}$ is constant as f is decreased (since for Gaussian likelihoods $\sigma(f) = \sqrt{f}\sigma(f = 1)$), i.e., we find the region where the likelihood is Gaussian and scale this appropriately. Doing so, we find a 'Gaussian' uncertainty of 0.047 for the mean of the Y1 mocks and no change for the Square mocks. This suggests we are recovering 80 per cent of the available BAO information in both cases[12]. Potential reasons we have not recovered the full 100 per cent include: our compression of the $\mu_{\rm obs} < 0.8$ BAO signal into one data vector, our cut at $\mu_{\rm obs} < 0.8$, our redshift weights being sub-optimal, and the fact that we are treating the redshift errors as Gaussian in the forecasts.

The results presented in this section validate our modeling of the BAO information with respect to the line of sight. As expected for redshift uncertainty greater than $0.02(1 + z)$, the information and the signal itself is found to be nearly constant in $\mu_{\rm obs}$, allowing one to combine the information at $\mu_{\rm obs} < 0.8$ into a single $\xi(s_\perp)$ measurement and extract nearly all of the available BAO information.

5   DISCUSSION

We have presented a concise analysis of the factors that affect the achievable precision of BAO distance scale measurements made using photometric redshift samples. We use this information in order to determine an optimal clustering estimator for BAO measurements to be applied to DES Y1 data.

Our analytic work comprised of three pieces:

• We first investigated the form of the BAO signal, in configuration space, by subtracting a smooth 'no BAO' model from the nominal model, accounting for redshift uncertainty. We found that the BAO feature is nearly constant as a function of observed $\mu_{\rm obs}$ when plotted against the transverse separation, $s_\perp$. See Fig. 1.

• We then utilized Fisher matrix predictions to understand the distribution of BAO with respect to the true and observed LOS. Using the cosine of the angle to the LOS, μ, we found that the information is confined primarily to nearly transverse clustering modes ($\mu = 0$). For a redshift uncertainty of $\sigma_z = 0.03(1 + z)$, there is more than 100 times as much $D_A(z)$ information as $H(z)$ information. However, when we consider the distribution of information in terms of the observed $\mu_{\rm obs}$, we find it is nearly evenly distributed up to $\mu_{\rm obs} \sim 0.8$. The redshift uncertainty causes true $\mu = 0$ information to be distributed over all $\mu_{\rm obs}$. This explains the results we obtained for the observed signal. See Fig. 2.

• Finally, we derived weights to apply to galaxies based on the evolution in redshift uncertainty, number density, and bias as a function of redshift. This weight is given by Eq. 16. Further, we learned that at a given redshift, weighting galaxies based on their particular redshift uncertainty provides minimal gains (see Table 1).

The above findings lead to the following recommendations for obtaining optimal BAO measurements from a photometric redshift sample:

• Given a parent sample of galaxy magnitudes, redshifts, and estimates of the redshift uncertainty, one should optimize the sample based on optimizing the Fisher forecast for the BAO uncertainty when applying color and magnitude cuts to change the number density and mean redshift of the sample. This matches the approach used for the DES Y1 BAO sample selection (Crocce et al. in prep.). Our results in Section 2.3.2 suggest that once selecting a sample based on these criteria, the treatment of variations in the redshift uncertainty within the sample should have negligible impact.

• Assuming the redshift uncertainty is $\geqslant 0.02(1 + z)$, measurements of the clustering with respect to the transverse separation should provide nearly optimal BAO constraints. Thus, one should measure $\xi_{\rm phot}(s_\perp, \mu_{\rm obs})$ and use Eq. 16 to weight each galaxy as function of redshift.

• One should then compress this information into the clustering estimator $\xi_{\rm opt}(s_\perp) = \int d\mu_{\rm obs} W_{\rm opt}(\mu_{\rm obs})\xi_{\rm phot}(s_\perp, \mu_{\rm obs})$, with $W_{\rm opt}(\mu_{\rm obs})$ determined from the Fisher matrix considerations described in Section 2.2. Additionally, one should consider the accuracy to which $\xi_{\rm phot}(s_\perp)$ can be modeled at high $\mu_{\rm obs}$ in order to inform any choice for a maximum $\mu_{\rm obs}$. One can then measure the angular diameter distance using $\xi_{\rm opt}(s_\perp)$.

We measured the clustering of mock samples intended to simulate DES Y1 data, using the approach advocated above. These measurements validated our approach and we found we were able to extract 80 per cent of the BAO information using measurements of $\xi_{\rm phot}(s_\perp)$ and $\mu_{\rm obs} < 0.8$. Our analytic and simulated results agreed that little BAO information is accessible for $\mu_{\rm obs} > 0.8$.

One important result that we obtained is that the expected precision on BAO measurements determined from $\mu_{\rm obs}$ bins of width $\Delta\mu_{\rm obs} = 0.2$ recover nearly the same uncertainty as those for $\mu_{\rm obs} < 0.8$. This is because the information between different $\mu_{\rm obs}$ is highly covariant. This suggests that measuring the BAO scale in different $\mu_{\rm obs}$ bins is likely to be an important test of the robustness of the results for the entire $\mu_{\rm obs}$ range, as disagreement might imply errors related to, e.g., the modeling assumptions or the knowledge of the photometric redshift uncertainty.

Our results show how one can construct a compressed and nearly optimal data vector to use for BAO measurements for data with significant redshift uncertainty. We expect this to be applied to the DES Y1 data in the near future, for which we predict a 5 per cent angular diameter distance measurement to be achieved. Given we extract 80 per cent of the BAO information, we expect uncertainties can be reduced by an additional 10 per cent. As this is modest, we expect that the most significant further gains will be achieved through more precise redshift information, either via improved photometric redshifts or methods that are able to reconstruct the radial density field. Accepting the redshift uncertainty as given from a catalog, we believe this work presents a close to optimized analysis that should inform any future BAO studies conducted with purely photometric data.

---

[12] Determining relative information achieved as compared to that possible as $(\sigma/\sigma_{\rm Fisher})^{-2}$.

REFERENCES

Alam, S., Albareti, F. D., Allende Prieto, C., et al. 2015, arXiv:1501.00963
Alam, S., Ata, M., Bailey, S., et al. 2016, arXiv:1607.03155
Anderson, L., Aubourg, E., Bailey, S., et al. 2012, MNRAS, 427, 3435
Anderson, L., Aubourg, E., Bailey, S., et al. 2014, MNRAS, 441, 24
Asorey, J., Carrasco Kind, M., Sevilla-Noarbe, I., Brunner, R. J., & Thaler, J. 2016, MNRAS, 459, 1293
Aubourg, É., Bailey, S., Bautista, J. E., et al. 2015, Phys. Rev. D., 92, 123516
Avila S., G.Murray S., Knebe A., Power C., Robotham A. S. and Garcia-Bellido, J., 2015, MNRAS, 450, 1856
Blake, C., & Bridle, S. 2005, MNRAS, 363, 1329
Carnero, A., Sánchez, E., Crocce, M., Cabré, A., & Gaztañaga, E. 2012, MNRAS, 419, 1689
Carretero J., Castander F. J., Gaztañaga E., Crocce M., Fosalba P., 2015, MNRAS, 447, 646
Crocce, M., Castander, F. J., Gaztañaga, E., Fosalba, P., & Carretero, J. 2015, MNRAS, 453, 1513
Dawson K., et al., 2013, AJ, 145, 10
de Simoni, F., Sobreira, F., Carnero, A., et al. 2013, MNRAS, 435, 3017
Dodelson, S., & Schneider, M. D. 2013, Phys. Rev. D, 88, 063537
Eisenstein, D. J., & Hu, W. 1998, ApJ, 496, 605
Eisenstein D.J., et al., 2011, AJ, 142
Estrada, J., Sefusatti, E., & Frieman, J. A. 2009, ApJ, 692, 265
Etherington, J., & Thomas, D. 2015, MNRAS, 451, 660
Etherington, J., Thomas, D., Maraston, C., et al. 2017, MNRAS, 466, 228
Feldman, H. A., Kaiser, N., & Peacock, J. A. 1994, ApJ 426, 23
Fosalba P., Crocce M., Gaztañaga E., Castander F. J., 2015a, MNRAS, 448, 2987
Fosalba P., Gaztañaga E., Castander F. J., Crocce M., 2015b, MNRAS, 447, 1319
Hartlap, J., Simon, P., & Schneider, P. 2007, A&A, 464, 399
Hütsi, G. 2010, MNRAS, 401, 2477
Kaiser, N. 1987, MNRAS, 227, 1
Kazin, E. A., Sánchez, A. G., & Blanton, M. R. 2012, MNRAS, 419, 3223
Kazin, E. A., Sánchez, A. G., Cuesta, A. J., et al. 2013, MNRAS, 435, 64
Landy S. D., Szalay A. S., 1993, ApJ, 412, 64
Lewis, A., Challinor, A., & Lasenby, A. 2000, ApJ, 538, 473
Manera, M., Scoccimarro, R., Percival, W. J., et al. 2013, MNRAS, 428, 1036
Chaves-Montero, J., Angulo, R. E., & Hernández-Monteagudo, C. 2016, arXiv:1610.09688
Padmanabhan, N., Schlegel, D. J., Seljak, U., et al. 2007, MNRAS, 378, 852
Padmanabhan, N., & White, M. 2009, Phys. Rev. D, 80, 063508
Padmanabhan, N., Xu, X., Eisenstein, D. J., et al. 2012, MNRAS, 427, 2132
Peebles, P. J. E. 1980, Research supported by the National Science Foundation. Princeton, N.J., Princeton University Press, 1980. 435 p.
Percival, W. J., Verde, L., & Peacock, J. A. 2004, MNRAS, 347, 645
Percival, W. J., Ross, A. J., Sánchez, A. G., et al. 2014, MNRAS, 439, 2531
Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2015, arXiv:1502.01589
Ross A. J., et al., 2015, MNRAS, 451, 1331
Ross, A. J., Beutler, F., Chuang, C.-H., et al. 2017, MNRAS, 464, 1168
Rozo, E., Rykoff, E. S., Abate, A., et al. 2016, MNRAS, 461, 1431
Rykoff E.S., et al., 2014, ApJ, 785, 104
Sánchez, E., Carnero, A., García-Bellido, J., et al. 2011, MNRAS, 411, 277
Seo, H.-J., & Eisenstein, D. J. 2003, ApJ, 598, 720
Seo, H.-J., & Eisenstein, D. J. 2007, ApJ, 665, 14
Seo, H.-J., Siegel, E. R., Eisenstein, D. J., & White, M. 2008, ApJ, 686, 13-24
Seo, H.-J., Ho, S., White, M., et al. 2012, ApJ, 761, 13
Seo H.-J., Beutler F., Ross A. J., & Saito S., 2016, 460, 2453
Silk, J. 1968, ApJ, 151, 459
Tegmark, M. 1997, Physical Review Letters, 79, 3806
Vargas-Magaña, M., Ho, S., Cuesta, A. J., et al. 2016, arXiv:1610.03506

Weinberg, D. H., Mortonson, M. J., Eisenstein, D. J., et al. 2013, Physics Reports, 530, 87

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

Zehavi, I., Zheng, Z., Weinberg, D. H., et al. 2011, ApJ, 736, 59

Zhan, H., Knox, L., & Tyson, J. A. 2009, ApJ, 690, 923