

# Analysing baryon acoustic oscillations in sparse spectroscopic samples via cross-correlation with dense photometry

A. Patej<sup>1</sup>\*† and D. J. Eisenstein<sup>2</sup>

<sup>1</sup>Steward Observatory, University of Arizona, 933 N. Cherry Ave., Tucson, AZ 85719, USA

<sup>2</sup>Harvard-Smithsonian Center for Astrophysics, 60 Garden St, Cambridge, MA 02138, USA

Accepted 2018 April 4. Received 2018 March 21; in original form 2017 September 12

## ABSTRACT

We develop a formalism for measuring the cosmological distance scale from baryon acoustic oscillations (BAO) using the cross-correlation of a sparse redshift survey with a denser photometric sample. This reduces the shot noise that would otherwise affect the autocorrelation of the sparse spectroscopic map. As a proof of principle, we make the first on-sky application of this method to a sparse sample defined as the  $z > 0.6$  tail of the Sloan Digital Sky Survey’s (SDSS) BOSS/CMASS sample of galaxies and a dense photometric sample from SDSS DR9. We find a  $2.8\sigma$  preference for the BAO peak in the cross-correlation at an effective  $z = 0.64$ , from which we measure the angular diameter distance  $D_M(z = 0.64) = (2418 \pm 73 \text{ Mpc})(r_s/r_{s, \text{fid}})$ . Accordingly, we expect that using this method to combine sparse spectroscopy with the deep, high-quality imaging that is just now becoming available will enable higher precision BAO measurements than possible with the spectroscopy alone.

**Key words:** distance scale – large-scale structure of Universe.

## 1 INTRODUCTION

The signature of baryon acoustic oscillations (BAOs) is a peak in the clustering of galaxies at large scales that is a relic of propagating sound waves in the early Universe and whose location depends on the distance the waves were able to cover prior to recombination (for a recent review, see Weinberg et al. 2013). Major redshift surveys like the Sloan Digital Sky Survey’s (SDSS) Baryon Oscillation Spectroscopic Survey (BOSS; Dawson et al. 2013) and the 2dF Galaxy Redshift Survey (2dFGRS; Colless et al. 2001) have analysed the clustering of galaxies in order to detect this feature and use it to measure the cosmological quantities  $D_M(z)$ , the angular diameter distance, and  $H(z)$ , the Hubble parameter (Cole et al. 2005; Eisenstein et al. 2005). BOSS has already attained percent level precision distance measurements using the BAO method at  $z = 0.57$  (Anderson et al. 2014).

However, most detections of the BAO feature in samples of galaxies thus far have occurred at fairly low redshifts, where these galaxy redshift surveys are able to attain sufficiently high densities of galaxies with spectroscopic redshift measurements. New measurements of quasar clustering have recently enabled BAO detections to higher redshifts (Ata et al. 2017) using eBOSS (Dawson et al. 2016), while future surveys such as the Dark Energy Spectroscopic Instrument (DESI) (Levi et al. 2013) will push the availability of spectra at higher redshifts by targeting quasars in addition to lower redshift galaxies, for which new approaches will be required to fully leverage these sparser data (see also Passaglia, Manzotti & Dodelson 2017).

To this end, in this paper, we develop the formalism of a method for measuring the BAO signal with a sparse sample of galaxies with spectroscopic redshifts that consists of cross-correlating the sparse data with a dense photometric sample to enhance the signal. Our work is related to concepts previously introduced by Nishizawa, Oguri & Takada (2013), but we more fully develop the configuration-space statistic, deriving explicit forms for the correlation function and the covariance matrix. We then make the first application of this method to data from SDSS to verify whether we can detect the BAO peak. In what follows, we use a flat  $\Lambda$ CDM cosmology with  $\Omega_m = 0.314$  and  $h = 0.67$ , similar to Alam et al. (2017).

\* E-mail: [apatej@email.arizona.edu](mailto:apatej@email.arizona.edu)

† Einstein Fellow

## 2 MOTIVATION

We begin with a pedagogical motivation of when employing the cross-correlation is useful, with a more detailed calculation to follow in Section 3. As we are interested in large physical scales, for this section and the remainder of the paper, we assume that the galaxy density field is described by a Gaussian random field and a linear bias model. For a spectroscopic survey of galaxy density  $n_s$ , the variance on each Fourier mode is  $\sigma_s^2 = (P_s + 1/n_s)^2$ , so that the signal to noise (S/N) is given by (Weinberg et al. 2013):

$$\frac{S}{N} = \frac{n_s P_s}{1 + n_s P_s}, \quad (1)$$

where  $P_s$  denotes the power spectrum  $P_s(k)$ . From this, we immediately see that the S/N per mode approaches 1 when  $n_s P_s \gg 1$  but decreases when  $n_s P_s < 1$ , so that the relevant quantity to track is  $nP$ .

Let us now assume we have a photometric data set with some density  $n_p$  and power spectrum  $P_p$ . Both the spectroscopic and photometric samples are biased tracers of the underlying matter distribution such that

$$\delta_s(\mathbf{x}) = b_s \delta(\mathbf{x}), \quad \delta_p(\mathbf{x}) = b_p \delta(\mathbf{x}). \quad (2)$$

Accordingly, we may write the power spectra as  $P_s(k) = b_s^2 P(k)$  and  $P_p(k) = b_p^2 (n_p/N_p) P(k)$ , with the factor of  $n_p/N_p = 1/L_p$ , where  $L_p$  is the line-of-sight spatial depth of the photometry, approximately accounting for the fact that we are converting the three-dimensional  $P(k)$  to the two-dimensional  $P_p(k)$ .

When we cross-correlate a photometric sample with a spectroscopic sample, we obtain the power spectrum of the cross-correlation as  $P_x = (n_p/N_p) b_s b_p P$  with variance  $\sigma_x^2 = P_x^2 + \sigma_{s,2} \sigma_{p,2}$ , where the subscript ‘2’ indicates projection, so that the signal-to-noise becomes

$$\left(\frac{S}{N}\right)^2 = \frac{P_x^2}{P_x^2 + (b_s^2 [n_s/N_s] P + 1/N_s)(b_p^2 [n_p/N_p] P + 1/N_p)}, \quad (3)$$

$$\left(\frac{S}{N}\right)^2 = \left(1 + \frac{n_s}{N_s} \frac{N_p}{n_p} \left[1 + \frac{1}{n_s b_s^2 P}\right] \left[1 + \frac{1}{n_p b_p^2 P}\right]\right)^{-1}, \quad (4)$$

$$\left(\frac{S}{N}\right)^2 = \left(1 + \frac{L_p}{L_s} \left[1 + \frac{1}{n_s b_s^2 P}\right] \left[1 + \frac{1}{n_p b_p^2 P}\right]\right)^{-1}. \quad (5)$$

Consequently, we see that with the cross-correlation, when  $n_s P_s$  gets small, we can, to a certain extent, compensate with the  $n_p P_p$  term. More specifically, while equation (1) goes as  $n_s$  for small  $n_s$ , equation (5) goes as  $n_s^{1/2}$ .

To see when this method would improve upon the results of using the autocorrelation, let us estimate the number of modes that we could measure. For a spectroscopic sample with constant density, the effective number of modes  $\mathcal{N}_s$  in a shell of width  $\Delta k$  is given by

$$\mathcal{N}_s = \frac{AL_s}{2(2\pi)^3} 4\pi k^2 \Delta k \times \left(\frac{n_s P_s}{1 + n_s P_s}\right)^2. \quad (6)$$

In the cross-correlation of the spectroscopy with a photometric sample, we have  $k_z = 0$ , which means that

$$\mathcal{N}_p = \frac{A}{(2\pi)^2} 2\pi k \Delta k \times \left(1 + \frac{L_p}{L_s} \left[1 + \frac{1}{n_s b_s^2 P}\right] \left[1 + \frac{1}{n_p b_p^2 P}\right]\right)^{-1}. \quad (7)$$

We are looking for the case wherein  $\mathcal{R} = \mathcal{N}_s/\mathcal{N}_p < 1$ , which means that

$$\mathcal{R} = \frac{kL_s}{2\pi} \left(\frac{1}{1 + 1/(n_s b_s^2 P)}\right)^2 \times \left(1 + \frac{L_p}{L_s} \left[1 + \frac{1}{n_s b_s^2 P}\right] \left[1 + \frac{1}{n_p b_p^2 P}\right]\right), \quad (8)$$

and if we assume that  $n_s b_s^2 P \ll 1$  (the limit in which the autocorrelation method begins to lose information), this simplifies to

$$\mathcal{R} = \frac{kL_s}{2\pi} (n_s b_s^2 P)^2 \times \left(\frac{L_p}{L_s} [n_s b_s^2 P]^{-1} \left[1 + \frac{1}{n_p b_p^2 P}\right]\right), \quad (9)$$

$$\mathcal{R} = \frac{kL_p}{2\pi} \left(\frac{1 + 1/(n_p b_p^2 P)}{1/(n_s b_s^2 P)}\right). \quad (10)$$

Let us denote  $\tau_s = n_s b_s^2 P$  and  $\tau_p = n_p b_p^2 P$ , so that

$$\mathcal{R} = \frac{kL_p}{2\pi} \frac{1 + \tau_p^{-1}}{\tau_s^{-1}}. \quad (11)$$

To satisfy  $\mathcal{R} < 1$ , we require

$$\tau_s^{-1} > \frac{kL_p}{2\pi} (1 + \tau_p^{-1}). \quad (12)$$

We can usually get photometric samples of high density, corresponding to the limit  $\tau_p^{-1} \rightarrow 0$ . Accordingly, for a reasonable photometric depth of  $L_p = 500 h^{-1} \text{ Mpc}$  and  $k = 0.15 h \text{ Mpc}^{-1}$ , we require  $\tau_s \lesssim 0.08$ . We additionally note that this method can be beneficial even when the condition that  $\tau_s \lesssim 0.08$  is *not* satisfied; shot noise always degrades the BAO precision, particularly when  $\tau_s < 1$ . In that case, even though the autocorrelation will give a more precise distance, the cross-correlation is still adding some new information since the individual modes are not being well measured by the autocorrelation.

### 3 MATHEMATICAL FORMALISM

#### 3.1 Framework

We now more formally derive our method. Let us assume that we have two samples of galaxies: a spectroscopic sample with fractional overdensity field  $\delta_s(\mathbf{r})$ , and the other a photometric sample. For this latter sample, we can calculate a projected overdensity field as

$$\Delta_p(\boldsymbol{\theta}) = \frac{\int dr' r'^2 n_p(r') \delta_p(\mathbf{r}')}{\int dr' r'^2 n_p(r')}. \quad (13)$$

Defining the normalization factor as

$$\mathcal{N} \equiv \int dr' r'^2 n_p(r'), \quad (14)$$

this may be written simply as

$$\Delta_p(\boldsymbol{\theta}) = \frac{1}{\mathcal{N}} \int dr' r'^2 n_p(r') \delta_p(\mathbf{r}'). \quad (15)$$

The correlation of  $\delta_s$  and  $\Delta_p$ , noting that  $\mathbf{r} = (r, \boldsymbol{\theta})$ , is then

$$W(\mathbf{R}) = \langle \delta_s(\mathbf{r}) \Delta_p(\boldsymbol{\theta} + \mathbf{R}/r) \rangle, \quad (16)$$

$$W(\mathbf{R}) = \frac{\int d^3r \delta_s(\mathbf{r}) n_s(r) \mathcal{W}(r, \eta) \int d^2\theta' \Delta_p(\boldsymbol{\theta}') \mathcal{N} \mathcal{S}(\boldsymbol{\theta}, \boldsymbol{\theta}', \mathbf{R}, r)}{\int d^3r n_s(r) \mathcal{W}(r, \eta) \int d^2\theta' \mathcal{N} \mathcal{S}(\boldsymbol{\theta}, \boldsymbol{\theta}', \mathbf{R}, r)}. \quad (17)$$

Here,  $\mathcal{W}$  is a weighting function that may depend on  $r$  and other variables independent of  $\mathbf{r}$ , collectively referred to as  $\eta$ , and  $\mathcal{S}$  is a separation function that depends on the angular coordinates of the galaxies. If we focus on a single separation, then the separation function is a Dirac delta function:

$$\mathcal{S}(\boldsymbol{\theta}, \boldsymbol{\theta}', \mathbf{R}, r) = \delta_D^{(2)}(\mathbf{r}\boldsymbol{\theta}' - (\mathbf{r}\boldsymbol{\theta} + \mathbf{R})). \quad (18)$$

Using equations (15) and (18) and writing out explicitly  $\mathbf{r} = (r, \boldsymbol{\theta})$ , equation (17) becomes

$$W(\mathbf{R}) = \frac{\int dr r^2 n_s(r) \mathcal{W}(r, \eta) \int d^2\theta \delta_s(r, \boldsymbol{\theta}) \int d^2\theta' \delta_D^{(2)}(\mathbf{r}\boldsymbol{\theta}' - (\mathbf{r}\boldsymbol{\theta} + \mathbf{R})) \int dr' r'^2 n_p(r') \delta_p(r', \boldsymbol{\theta}')}{\int dr r^2 n_s(r) \mathcal{W}(r, \eta) \int d^2\theta \int d^2\theta' \delta_D^{(2)}(\mathbf{r}\boldsymbol{\theta}' - (\mathbf{r}\boldsymbol{\theta} + \mathbf{R})) \int dr' r'^2 n_p(r')}. \quad (19)$$

The delta function collapses the  $\theta'$  integral in both the numerator and denominator, giving  $1/r^2$ , and in the denominator the  $\theta$  integral can be immediately evaluated to a constant, yielding

$$W(\mathbf{R}) = \frac{\int dr n_s(r) \mathcal{W}(r, \eta) \int d^2\theta \delta_s(r, \boldsymbol{\theta}) \int dr' r'^2 n_p(r') \delta_p(r', \boldsymbol{\theta} + \mathbf{R}/r)}{4\pi \int dr n_s(r) \mathcal{W}(r, \eta) \int dr' r'^2 n_p(r')}. \quad (20)$$

Rearranging the order of integration gives

$$W(\mathbf{R}) = \frac{\int dr n_s(r) \mathcal{W}(r, \eta) \int dr' r'^2 n_p(r') \int d^2\theta \delta_s(r, \boldsymbol{\theta}) \delta_p(r', \boldsymbol{\theta} + \mathbf{R}/r)}{4\pi \int dr n_s(r) \mathcal{W}(r, \eta) \int dr' r'^2 n_p(r')}. \quad (21)$$

With an eye towards an application to catalogue data, we note that the correlation function is constructed from pair counts of galaxies, normalized by pair counts of random galaxies, so that

$$W(\mathbf{R}) = \frac{N_s N_p}{R_s R_p}. \quad (22)$$

This allows us to identify the numerator and denominator of equation (21) with  $N_s N_p$  and  $R_s R_p$ , where both terms are a function of  $R$ , the separation between the galaxies in the pair.

#### 3.2 Expectation value

Taking the expectation value of equation (21) yields

$$\langle W(\mathbf{R}) \rangle = \frac{\int dr n_s(r) \mathcal{W}(r, \eta) \int dr' r'^2 n_p(r') \int d^2\theta \langle \delta_s(r, \boldsymbol{\theta}) \delta_p(r', \boldsymbol{\theta} + \mathbf{R}/r) \rangle}{4\pi \int dr n_s(r) \mathcal{W}(r, \eta) \int dr' r'^2 n_p(r')}. \quad (23)$$

Focusing on the numerator,

$$\langle N_s N_p(\mathbf{R}) \rangle = \int dr n_s(r) \mathcal{W}(r, \eta) \int dr' r'^2 n_p(r') \int d^2\theta \langle \delta_s(r, \boldsymbol{\theta}) \delta_p(r', \boldsymbol{\theta} + \mathbf{R}/r) \rangle, \quad (24)$$

we can use the Fourier transform,

$$\delta(\mathbf{r}) = \frac{1}{(2\pi)^{3/2}} \int d^3k \delta_k e^{i\mathbf{k}\cdot\mathbf{r}}, \quad (25)$$

as well as the formalism of Dekel & Lahav (1999), which relates the  $\delta_g = \delta_s, \delta_p$  fields to the underlying matter density field  $\delta \equiv \delta_m$  as

$$\delta_g = b_g \delta + \epsilon_g, \quad (26)$$

to write the expectation value of two  $\delta$ s as

$$\langle \delta(r, \boldsymbol{\theta}) \delta(r', \boldsymbol{\theta} + \mathbf{R}/r) \rangle = \frac{b_s b_p}{(2\pi)^3} \int \int d^3k d^3k' \langle \delta_k \delta_{k'} \rangle e^{i\mathbf{k}\cdot\mathbf{r}} e^{i\mathbf{k}'\cdot\mathbf{r}'} e^{i\mathbf{k}'\cdot(\boldsymbol{\theta} + \mathbf{R}/r)} e^{i\mathbf{k}'\cdot\mathbf{r}'}, \quad (27)$$

where we are using the flat sky approximation and have written  $k = \sqrt{k_\perp^2 + k_\parallel^2}$ , with  $k_\perp$  indicating the transverse components of  $k$  and  $k_\parallel$  indicating the line-of-sight component. Then, using

$$\langle \delta_k \delta_{k'} \rangle = \delta_D^{(3)}(\mathbf{k} + \mathbf{k}') P(k), \quad (28)$$

and integrating over  $k$  yields

$$\langle \delta(r, \boldsymbol{\theta}) \delta(r', \boldsymbol{\theta} + \mathbf{R}/r) \rangle = \frac{b_s b_p}{(2\pi)^3} \int d^3k' P(k') e^{-i\mathbf{r}\cdot\mathbf{k}'_\perp \cdot \boldsymbol{\theta}} e^{-i\mathbf{r}k'_\parallel} e^{i\mathbf{r}'\cdot\mathbf{k}'_\perp \cdot (\boldsymbol{\theta} + \mathbf{R}/r)} e^{i\mathbf{r}'k'_\parallel}, \quad (29)$$

$$= \frac{b_s b_p}{(2\pi)^3} \int d^3k' P(k') e^{-i\mathbf{r}\cdot\mathbf{k}'_\perp \cdot \boldsymbol{\theta}} e^{i\mathbf{r}'\cdot\mathbf{k}'_\perp \cdot (\boldsymbol{\theta} + \mathbf{R}/r)} e^{-i(r-r')k'_\parallel}. \quad (30)$$

Thus,

$$\langle N_s N_p \rangle = b_s b_p \int dr n_s(r) \mathcal{W}(r, \eta) \int dr' r'^2 n_p(r') \int \frac{d^2\theta}{(2\pi)^3} \int d^3k' P(k') e^{-i\mathbf{r}\cdot\mathbf{k}'_\perp \cdot \boldsymbol{\theta}} e^{i\mathbf{r}'\cdot\mathbf{k}'_\perp \cdot (\boldsymbol{\theta} + \mathbf{R}/r)} e^{-i(r-r')k'_\parallel}. \quad (31)$$

If we assume that the line-of-sight weighting is much larger than the BAO scale, then we may approximate  $k = \sqrt{k_\perp^2 + k_\parallel^2} \approx k_\perp$  and split up the last integral:

$$\langle N_s N_p \rangle = b_s b_p \int dr n_s(r) \mathcal{W}(r, \eta) \int dr' r'^2 n_p(r') \int \frac{d^2\theta}{(2\pi)^3} \int d^2k'_\perp P(k'_\perp) e^{-i\mathbf{r}\cdot\mathbf{k}'_\perp \cdot \boldsymbol{\theta}} e^{i\mathbf{r}'\cdot\mathbf{k}'_\perp \cdot (\boldsymbol{\theta} + \mathbf{R}/r)} \int dk'_\parallel e^{-i(r-r')k'_\parallel}, \quad (32)$$

$$= b_s b_p \int dr n_s(r) \mathcal{W}(r, \eta) \int dr' r'^2 n_p(r') \int \frac{d^2\theta}{(2\pi)^3} \int d^2k'_\perp P(k'_\perp) e^{-i\mathbf{r}\cdot\mathbf{k}'_\perp \cdot \boldsymbol{\theta}} e^{i\mathbf{r}'\cdot\mathbf{k}'_\perp \cdot (\boldsymbol{\theta} + \mathbf{R}/r)} 2\pi \delta_D(r - r'), \quad (33)$$

$$= b_s b_p \int dr r^2 n_s(r) \mathcal{W}(r, \eta) n_p(r) \int \frac{d^2\theta}{(2\pi)^2} \int d^2k'_\perp P(k'_\perp) e^{-i\mathbf{r}\cdot\mathbf{k}'_\perp \cdot \boldsymbol{\theta}} e^{i\mathbf{r}'\cdot\mathbf{k}'_\perp \cdot (\boldsymbol{\theta} + \mathbf{R}/r)}. \quad (34)$$

Now, the  $\boldsymbol{\theta}$  terms cancel, making the integral over that variable trivial:

$$\langle N_s N_p \rangle = b_s b_p \int dr r^2 n_s(r) \mathcal{W}(r, \eta) n_p(r) \frac{4\pi}{(2\pi)^2} \int d^2k'_\perp P(k'_\perp) e^{i\mathbf{k}'_\perp \cdot \mathbf{R}/r}, \quad (35)$$

$$= \frac{b_s b_p}{\pi} \int dr r^2 n_s(r) \mathcal{W}(r, \eta) n_p(r) \int d^2k'_\perp P(k'_\perp) e^{i\mathbf{k}'_\perp \cdot \mathbf{R}}. \quad (36)$$

We can then identify:

$$w_p(\mathbf{R}) \equiv \frac{1}{2\pi} \int d^2k'_\perp P(k'_\perp) e^{i\mathbf{k}'_\perp \cdot \mathbf{R}}, \quad (37)$$

where  $w_p$  is the projected correlation function (Davis & Peebles 1983), which has been used extensively in galaxy clustering analyses (e.g. Coil et al. 2008; Zehavi et al. 2011; Nuza et al. 2013) as it is less affected by redshift-space distortions. The function  $w_p$  can be calculated as the line-of-sight integral of the three-dimensional correlation function,

$$w_p(r_\perp) = \int_{-\infty}^{\infty} dr_\parallel \xi(r_\perp, r_\parallel). \quad (38)$$

Returning to our derivation, we find

$$\langle N_s N_p \rangle = 2b_s b_p \int dr r^2 n_s(r) \mathcal{W}(r, \eta) n_p(r) w_p(\mathbf{R}). \quad (39)$$

Finally, this yields

$$\langle W(\mathbf{R}) \rangle = \frac{b_s b_p}{2\pi} w_p(\mathbf{R}) \frac{\int dr r^2 n_s(r) \mathcal{W}(r, \eta) n_p(r)}{\int dr n_s(r) \mathcal{W}(r, \eta) \int dr' r'^2 n_p(r')}. \quad (40)$$

To gain some intuition for the behaviour of  $\langle w \rangle$ , we assume that the distributions of the spectroscopic and photometric galaxies are given by boxcar functions of lengths  $L_s$  and  $L_p$ , respectively, which represent the depths of the survey, and that  $L_s < L_p$ . Let us also assume the weighting is independent of  $r$ , and equal to 1. In that case, the integral normalization gives

$$\frac{\int dr r^2 n_s(r) \mathcal{W}(r, \eta) n_p(r)}{\int dr n_s(r) \mathcal{W}(r, \eta) \int dr' r'^2 n_p(r')} = \frac{\int_0^{L_s} dr r^2}{\int_0^{L_s} dr \int_0^{L_p} dr' r'^2} = \frac{L_s^3}{L_p^3 L_s} = \frac{L_s^2}{L_p^3} \sim \frac{1}{L}. \quad (41)$$

Accordingly, we see that this factor roughly scales as the inverse of the depth of the survey.

### 3.3 Binned correlation function and covariance matrix

We develop the calculation further in the Appendices, but quote the results that will be relevant in the following sections here. While so far we have focused on a single separation  $\mathbf{R}$ , in most cosmological theories, the clustering is independent of the angle on the sky and therefore it is usual for one to integrate over circular annuli, whose bounds we denote by  $R_1$  and  $R_2$ . Written in terms of the Bessel function  $J_1$ , our binned correlation function is

$$w(R_1, R_2) = \frac{b_s b_p}{2\pi^2 (R_2^2 - R_1^2)} \frac{\int dr r^2 n_s(r) \mathcal{W}(r, \eta) n_p(r)}{\int dr n_s(r) \mathcal{W}(r, \eta) \int dr' r'^2 n_p(r')} \int d^2 k'_\perp P(k'_\perp) \left[ \frac{R_2 J_1(k'_\perp R_2) - R_1 J_1(k'_\perp R_1)}{k'_\perp} \right], \quad (42)$$

while our covariance matrix is

$$\begin{aligned} C(R_1, R_2, S_1, S_2) &= \frac{b_s^2 b_p^2 \left[ \int dr_1 n_s(r_1) \mathcal{W}(r_1, \eta) \int dr'_1 r'^2 n_p(r'_1) \int dr_2 n_s(r_2) \mathcal{W}(r_2, \eta) \int dr'_2 r'^2 n_p(r'_2) \right]^{-1}}{\pi^2 \mathcal{A} (R_2^2 - R_1^2) (S_2^2 - S_1^2)} \\ &\times \left( \int dr_1 n_s(r_1)^2 \mathcal{W}(r_1, \eta)^2 \int dr'_1 r'^2 n_p(r'_1)^2 \int \frac{d^2 u}{u^2} P\left(\frac{u}{r_1}\right) P\left(\frac{u}{r'_1}\right) \right. \\ &\times \left[ R_2 J_1\left(\frac{u R_2}{r_1}\right) - R_1 J_1\left(\frac{u R_1}{r_1}\right) \right] \left[ S_2 J_1\left(\frac{u S_2}{r_1}\right) - S_1 J_1\left(\frac{u S_1}{r_1}\right) \right] \\ &+ \int dr_1 r_1 n_s(r_1) \mathcal{W}(r_1, \eta) n_p(r_1) \int dr'_1 r'_1 n_s(r'_1) \mathcal{W}(r'_1, \eta) n_p(r'_1) \int \frac{d^2 u}{u^2} P\left(\frac{u}{r_1}\right) P\left(\frac{u}{r'_1}\right) \\ &\times \left[ R_2 J_1\left(\frac{u R_2}{r_1}\right) - R_1 J_1\left(\frac{u R_1}{r_1}\right) \right] \left[ S_2 J_1\left(\frac{u S_2}{r'_1}\right) - S_1 J_1\left(\frac{u S_1}{r'_1}\right) \right] \Big). \end{aligned} \quad (43)$$

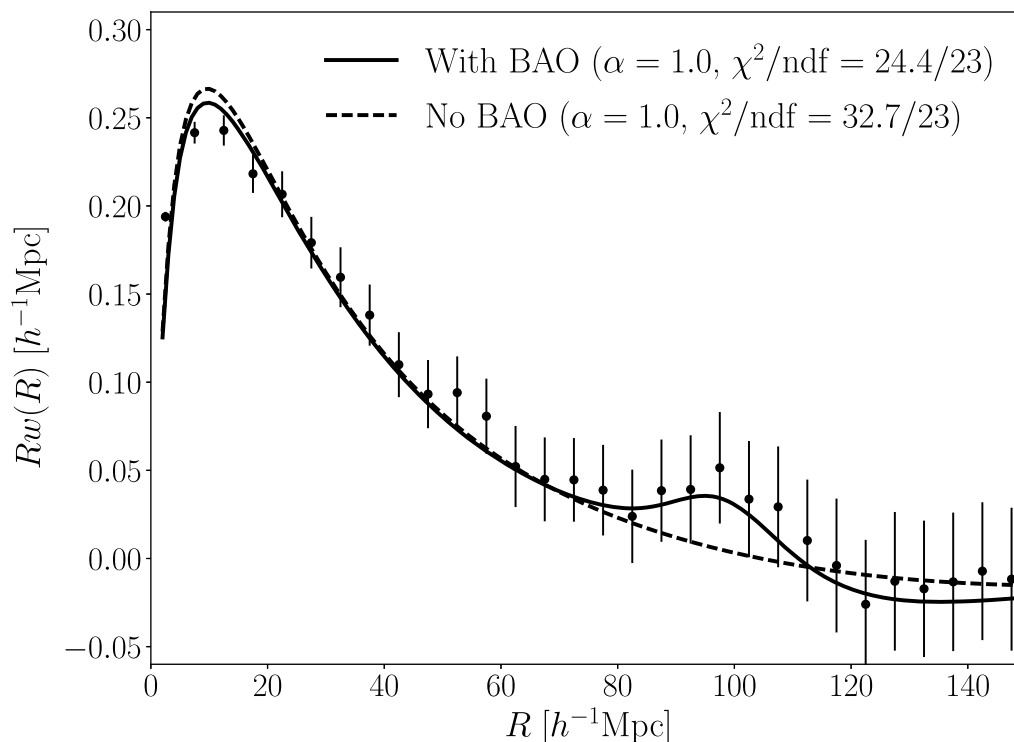
This expression can be numerically integrated up to the factors of  $b_s$  and  $b_p$ , which may not be known for the data sets under consideration. A form for  $n_p(r)$  needs to be assumed as well, and can be estimated from smaller areas with existing spectroscopic catalogues or high-quality photometric redshifts. Care must be taken when integrating the Bessel functions, as in the first term, both sets depend on  $r_1$ , whereas the second term also includes  $r'_1$ ; the integration over  $u$  thus needs to account for the range of arguments that will be encountered in order to appropriately sample the oscillations.

## 4 DATA

In order to test this method, we use two samples from the Sloan Digital Sky Survey. First, for our spectroscopic data set, we consider the colour cut-defined CMASS sample of galaxies from Data Release 12 of SDSS-III (Eisenstein et al. 2011; Dawson et al. 2013; Alam et al. 2015; Reid et al. 2016) for our spectroscopic sample. These data were obtained with a 2.5 m telescope at Apache Point Observatory (York et al. 2000; Gunn et al. 2006) whose instrumentation is well documented (Fukugita et al. 1996; Smith et al. 2002; Doi et al. 2010; Smee et al. 2013) and for which a set of data processing pipelines has been developed (Lupton et al. 2001; Pier et al. 2003; Padmanabhan et al. 2008; Bolton et al. 2012; Weaver et al. 2015). To obtain a fairly sparse sample of spectra for the validation of the method in the next section, we restrict the redshift range, focusing only on those galaxies for which  $z > 0.6$ . This yields 205 367 galaxies, and we select a set of random points 10 times larger than this for the construction of the correlation function. We note that while this  $z > 0.6$  data is fairly sparse, it is not a sample for which  $\tau < 0.08$ . The photometric sample used in this paper comes from SDSS DR9 (Ahn et al. 2012) imaging data and comprises 6594 677 galaxies [see Law-Smith & Eisenstein (2017) for a detailed description of the colour cuts used to define this sample]; we use a catalogue of randoms five times larger than this to construct the correlation function.

In practice, the *projected* cross-correlation function between the spectroscopic and photometric samples can be constructed via the Landy–Szalay estimator (Landy & Szalay 1993),

$$w(R) = \frac{D_s D_p(R) - D_s R_p(R) - D_p R_s(R) + R_s R_p(R)}{R_s R_p(R)}, \quad (44)$$



**Figure 1.** The projected correlation function  $w$  plotted against two theoretical curves: the solid black line is the predicted projected correlation function in the presence of BAO, while the dashed violet curve is the expectation without BAO. The data show a bump near  $100 h^{-1}$  Mpc consistent with the BAO signal.

where  $D$  indicates the data and  $R$  denotes the random points;  $p$  and  $s$  refer to the photometry and spectroscopy, respectively. While the Landy–Szalay estimator is normally defined as providing an estimate of  $\xi(r)$ , the full three-dimensional correlation function, our method involves a photometric sample for which only the on-sky coordinates are known while the radial position is unknown. Accordingly, when we compute the correlation function, we use only spectroscopic galaxies as our central points and then calculate pair counts in two-dimensional annular bins around the central galaxy, assuming that the photometric galaxies in the bin are at the same redshift as the spectroscopic galaxy to convert angular separations into physical distances, yielding  $w(R)$ .

We compute errors on our measurements in two ways. First, we divide the spectroscopic survey area into 150 roughly commensurate regions using Voronoi tessellation and then employ the jackknife method. Secondly, we use the analytical derivation presented in the preceding sections and test the explicit form for the covariance matrix (equation 43) in fitting the functional form of  $w(R)$  given by equation (42).

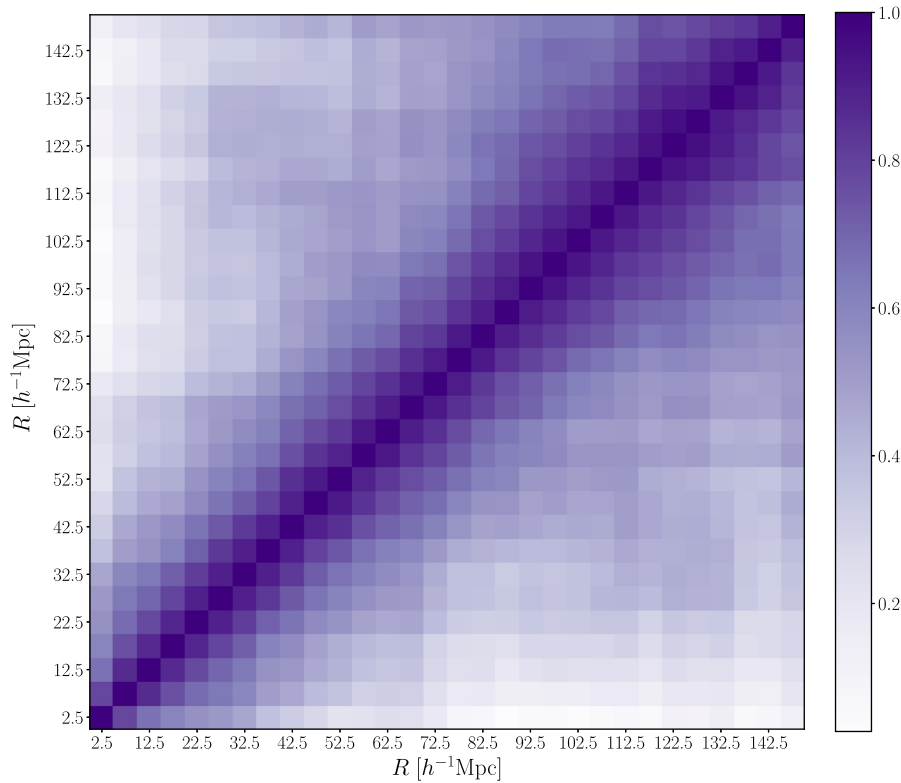
## 5 RESULTS

### 5.1 Fitting for the BAO peak

We compute the projected correlation function  $w(R)$  using the Landy–Szalay estimator with a binning of  $5 h^{-1}$  Mpc; the result is plotted in Fig. 1. We fit two computed projected correlation functions, one derived using a CAMB power spectrum (Lewis, Challinor & Lasenby 2000) that includes the BAO, and the other a smooth, BAO-less power spectrum derived using the prescriptions of Eisenstein & Hu (1998), to the data.

In both theoretical prescriptions of the correlation function, we include a nuisance parameter in the form of an additive constant  $w_0$ , which corresponds to excess power at very large scales and which is left a free parameter. Accordingly, we fit to a function of the form  $w(R) = aw_c(R) + w_0$ , where  $w_c(R)$  is the calculated correlation function and  $a$  and  $w_0$  are parameters to be found, to data at  $R \geq 27.5 h^{-1}$  Mpc. Prior to fitting, we normalize  $w_c$  to the value of the data at  $R = 27.5 h^{-1}$  Mpc. We use the jackknife covariance matrix shown in Fig. 2 in performing the fits. The results of our fits suggest that the parameter  $w_0$  is around  $1 \times 10^{-4}$ , indicating only a modest level of systematics.

We further consider the impact of a BAO peak shift, quantitatively provided by a dilation factor (denoted  $\alpha$ ) on our results. Scaling  $P(k) \rightarrow P(k/\alpha)$  for both the BAO and BAO-less cases, we recompute the correlation functions and fit them to the data, allowing for  $\alpha$  to vary in the range of  $0.8 \lesssim \alpha \lesssim 1.25$ ; the results for five values of  $\alpha$  are shown in Fig. 3. The data prefer a value of  $\alpha = 1.00$ , as indicated by the  $\chi^2$  curves for the fits that are shown in Fig. 4. A shift of  $\alpha = 1.00$  gives a  $\chi^2/\text{ndf} = 24.4/23$ , while the best-fitting value for the non-BAO curve has  $\chi^2/\text{ndf} = 32.3/23$  (at  $\alpha = 0.89$ ), yielding a  $2.8\sigma$  preference for the BAO. We do note, however, that as we are using a jackknife method, we have not included corrections for the impact of noise of the inversion of the covariance matrix such as is described in Hartlap, Simon & Schneider (2007) and Percival et al. (2014) in the context of mock catalogues.



**Figure 2.** The reduced covariance matrix for the cross-correlation, estimated from Voronoi jackknife samples and used to fit theoretical curves to the computed Landy–Szalay  $w$ .

Interpreting the  $\Delta\chi^2 = 4$  range as a  $2\sigma$  error, we find a rough estimate of about 0.03 for the  $1\sigma$  precision of  $\alpha$ . However, we caution that the low significance of the acoustic peak detection likely implies that the posterior of  $\alpha$  has non-Gaussian tails that would loosen the constraining power.

Our results align well with the prediction for the correlation function with BAO at intermediate-to-large scales. We note that while we do not have to consider redshift-space distortions with  $w$ , since we do not account for where the photometric galaxies are located along the line of sight, there are non-linear corrections that could have some impact at small scales ( $R \lesssim 10 h^{-1}$  Mpc), but these are not expected to yield major changes at large scales. Accordingly, our measurements yield a feature consistent with the BAO signal at  $R \sim 100 h^{-1}$  Mpc in the projected correlation function.

In order to measure the comoving angular diameter distance  $D_M(z) = (1+z)D_A(z)$ , where  $D_A$  is the physical angular diameter distance, we require an estimate of the effective redshift for our sample, which should be the mean redshift of the spectroscopic and photometric pairs. While we know the spectroscopic redshift distribution, we do not have detailed knowledge of the photometric redshift distribution. If the photometric galaxies had the same distribution as the spectroscopic sample, then the mean pair redshift would be the volume-weighted average of  $n_s^2(z)$ . However, the photometric sample is 1 mag fainter than the spectroscopic sample and has a slightly redder colour selection, so we expect the distribution of these galaxies to extend to higher redshift. Alternatively, if the photometric galaxies had a uniform  $n_p(z)$ , then the mean pair redshift would be the volume-weighted average of  $n_s(z)$ . We find these measurements give similar results and we adopt the average of the mean redshifts of  $n_s(z)$  and  $n_s^2(z)$ ,  $z = 0.64$ , as the effective redshift for our sample.

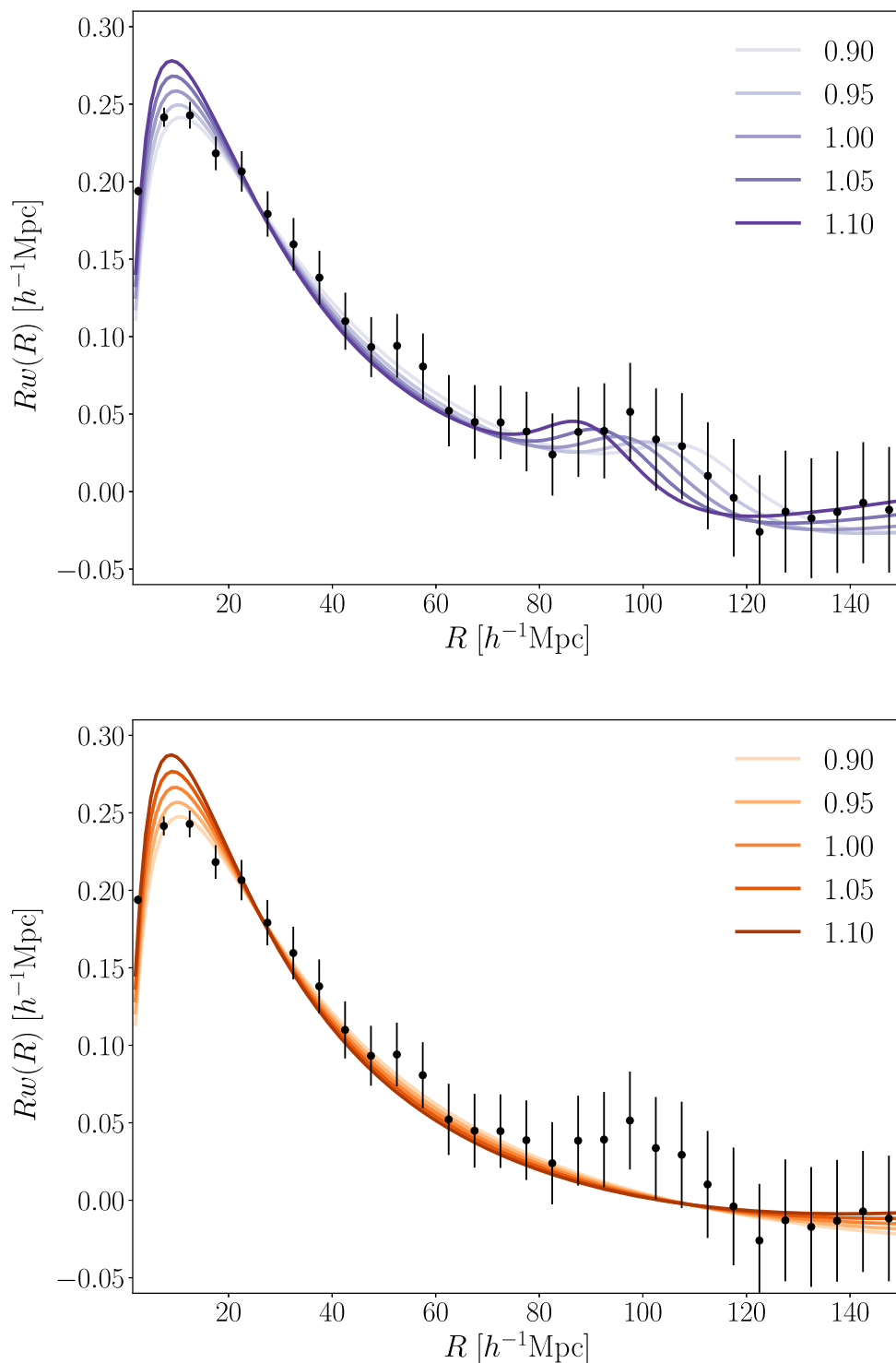
From the dilation parameter  $\alpha$ , we then obtain  $D_M(z)$  as (see e.g. Alam et al. 2017):

$$\alpha = \frac{D_M(z)r_{s,\text{fid}}}{D_M^{\text{fid}}(z)r_s}, \quad (45)$$

where  $r_s$  denotes the sound horizon and the subscript ‘fid’ refers to our fiducial cosmology. With the results above, we find  $D_M(z = 0.64) = (2418 \pm 73 \text{ Mpc})(r_s/r_{s,\text{fid}})$ . As a comparison, Alam et al. (2017) find  $D_M(z = 0.57) = (2179 \pm 35 \text{ Mpc})(r_s/r_{s,\text{fid}})$ . We note that our result is at least partially covariant with their high-redshift sample.

## 5.2 Comparison to the theoretical covariance matrix

In obtaining the above results, we have relied on the covariance matrix that we calculated from jackknife resampling. We now return to the analytical expression for the covariance matrix derived in Appendix C, albeit in an idealized form with some unknown quantities. In

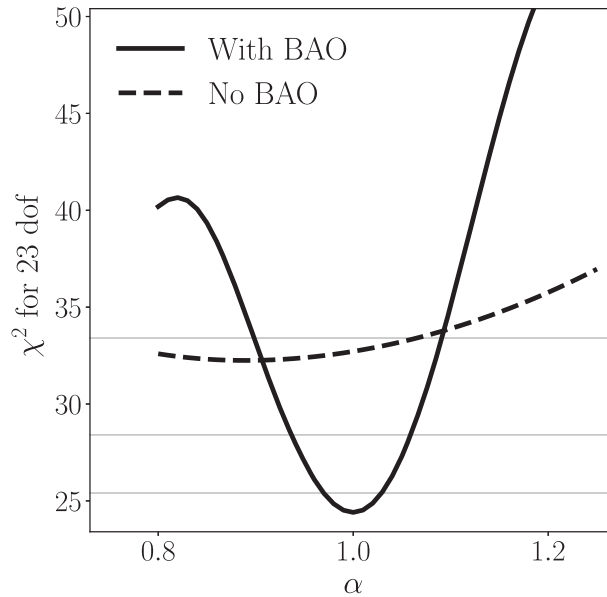


**Figure 3.** *Top:* the projected correlation functions with BAO for various values of the peak shift parameter,  $\alpha$ , fitted to the data. The data prefer a value of  $\alpha = 1.00$ . *Bottom:* the projected correlation functions for models without the BAO.

particular, in our derivation, we neglected to include shot noise and our matrix is only known up to the factor of  $b_s b_p$ , the product of the spectroscopic and photometric bias parameters, which is not known for our sample.

We accounted for these unknown factors in our expression as follows: first, to determine the rough value of  $b_s b_p$ , we took the ratio of the diagonals of the covariance matrix computed from equation (43) and those of the jackknife matrix. This comparison yielded roughly a factor of 2 in the ratio between large and small scales, indicating that the two matrices are not perfectly matched. The sense of the difference is that the jackknife matrix has larger diagonals at larger scales, which might be the result of additional systematic errors at large angles. We





**Figure 4.** The  $\chi^2$  values of the fits illustrated in Fig. 3; the grey lines denote the  $1\sigma$ ,  $2\sigma$ , and  $3\sigma$  bounds. The data clearly prefer a model that includes the BAO peak with  $\alpha = 1.00$ .

therefore chose to normalize the theory matrix at smaller scales ( $\lesssim 60 h^{-1}$  Mpc) where we expect the result to be less affected by large-scale systematics.

Secondly, we established the constant value of the shot noise  $P_{\text{shot}}$  to use by matching the eigenvalues of both matrices under the transformation  $P(k) \rightarrow P(k) + P_{\text{shot}}$  in equation (43). By choosing a modest value of  $P_{\text{shot}}$ , we could bring the eigenvalues into agreement without yielding noticeable deviations in the diagonal ratio. Following this procedure, we found  $P_{\text{shot}} = 1500 (h^{-1} \text{ Mpc})^3$ , a value that is commensurate with the level of shot noise in the BOSS survey. With this assumption, we found good agreement between the theoretical and jackknife matrices. Upon fitting the data, we found a preference for the BAO peak with  $\alpha = 0.99$  ( $\chi^2 = 15.7$  for 23 degrees of freedom) with  $\Delta\chi^2 = 2.9$ . The  $\chi^2$  suggests we may be slightly overnormalizing the theoretical matrix.

Furthermore, as noted above, we do see indications of the presence of large-angle observational systematic errors; accordingly, we have opted for relying on the results obtained by using the jackknife matrix. While it is beyond the scope of this work, future applications will require careful analyses of the survey systematics, which will need to be accounted for in the analytical calculation of the covariance matrix.

## 6 SUMMARY AND CONCLUSIONS

In this work, we have laid out a detailed formalism for a method of detecting the BAO in sparse spectroscopic data sets by cross-correlating the spectroscopy with photometry, for which greater densities of galaxies can be attained at far smaller cost. We analytically computed the projected correlation function and presented an explicit form for the covariance matrix, which can be calculated numerically and used to fit models to the data. We then applied the method to high-redshift data from SDSS.

We find that this method can reveal the BAO signal in sparse data; in our test data set, which relied on a spectroscopic sample consisting of the  $z > 0.6$  tail of the CMASS galaxy distribution cross-correlated with SDSS DR9 photometry, we found a  $2.8\sigma$  preference for the BAO and obtained a distance measurement of  $D_M(z = 0.64) = (2418 \pm 73 \text{ Mpc})(r_s/r_{s, \text{fid}})$ . This implies that cross-correlations of spectroscopy and photometry will provide a viable means with which to extend BAO measurements to higher redshifts even before the completion of major future surveys.

Of particular interest may be the application of this method to detect the BAO in the clustering of high-redshift quasars observed by present and upcoming surveys like the extended Baryon Oscillation Spectroscopic Survey (eBOSS) and the Dark Energy Spectroscopic Instrument (DESI). The eBOSS quasar sample, for example, is expected to have an average  $nP = 0.040$  (Zhao et al. 2016) over the redshift range  $0.6 < z < 2.2$ , which would make this method an effective way to detect the BAO in that sample. These data can be combined with photometry from the DECam Legacy Survey (DECaLS), the Mosaic  $z$ -band Legacy Survey (MzLS), and the Beijing-Arizona Sky Survey (BASS), which are currently obtaining deep, high-quality imaging of the SDSS survey area in order to provide target selection for DESI. Additionally, this method could be employed to measure the transverse BAO scale at  $z \sim 2$  using O III emission line galaxies that will be discovered by *Euclid* (e.g. Mehta et al. 2015).

## ACKNOWLEDGEMENTS

We would like to thank Aaron Bray for providing us with the photometric data files used in this paper and for generating the corresponding random catalogue. We additionally thank Lehman Garrison for generating the CAMB power spectrum used in our calculations.

AP acknowledges support from the NASA Einstein Fellowship program under Grant No. PF6-170157. DJE was supported as a Simons Foundation Investigator. This work was also supported by U.S. Department of Energy Grant No. DE-SC0013718, and made use of tools from the PYTHON packages NUMPY (van der Walt, Colbert & Varoquaux 2011), SCIPY (Oliphant 2007), and MATPLOTLIB (Hunter 2007), and also relied on ROOT (Brun & Rademakers 1997, see also <http://root.cern.ch/>).

This paper relies on data from SDSS-III. Funding for SDSS-III has been provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Science Foundation, and the U.S. Department of Energy Office of Science. The SDSS-III web site is <http://www.sdss3.org/>.

SDSS-III is managed by the Astrophysical Research Consortium for the Participating Institutions of the SDSS-III Collaboration including the University of Arizona, the Brazilian Participation Group, Brookhaven National Laboratory, Carnegie Mellon University, University of Florida, the French Participation Group, the German Participation Group, Harvard University, the Instituto de Astrofísica de Canarias, the Michigan State/Notre Dame/JINA Participation Group, Johns Hopkins University, Lawrence Berkeley National Laboratory, Max Planck Institute for Astrophysics, Max Planck Institute for Extraterrestrial Physics, New Mexico State University, New York University, Ohio State University, Pennsylvania State University, University of Portsmouth, Princeton University, the Spanish Participation Group, University of Tokyo, University of Utah, Vanderbilt University, University of Virginia, University of Washington, and Yale University.

## REFERENCES

- Ahn C. P. et al., 2012, *ApJS*, 203, 21  
 Alam S. et al., 2015, *ApJS*, 219, 12  
 Alam S. et al., 2017, *MNRAS*, 470, 2617  
 Anderson L. et al., 2014, *MNRAS*, 441, 24  
 Ata M. et al., 2017, *MNRAS*, 473, 4773  
 Bolton A. S. et al., 2012, *AJ*, 144, 144  
 Brun R., Rademakers F., 1997, *Nucl. Inst. Methods Phys. Res. A*, 389, 81  
 Coil A. L. et al., 2008, *ApJ*, 672, 153  
 Cole S. et al., 2005, *MNRAS*, 362, 505  
 Colless M. et al., 2001, *MNRAS*, 328, 1039  
 Davis M., Peebles P. J. E., 1983, *ApJ*, 267, 465  
 Dawson K. S. et al., 2013, *AJ*, 145, 10  
 Dawson K. S. et al., 2016, *AJ*, 151, 44  
 Dekel A., Lahav O., 1999, *ApJ*, 520, 24  
 Doi M. et al., 2010, *AJ*, 139, 1628  
 Eisenstein D. J., Hu W., 1998, *ApJ*, 496, 605  
 Eisenstein D. J. et al., 2005, *ApJ*, 633, 560  
 Eisenstein D. J. et al., 2011, *AJ*, 142, 72  
 Fukugita M., Ichikawa T., Gunn J. E., Doi M., Shimasaku K., Schneider D. P., 1996, *AJ*, 111, 1748  
 Gunn J. E. et al., 2006, *AJ*, 131, 2332  
 Hartlap J., Simon P., Schneider P., 2007, *A&A*, 464, 399  
 Hunter J. D., 2007, *Comput. Sci. Eng.*, 9, 90  
 Landy S. D., Szalay A. S., 1993, *ApJ*, 412, 64  
 Law-Smith J., Eisenstein D. J., 2017, *ApJ*, 836, 87  
 Levi M. et al., 2013, preprint (arXiv: [1308.0847](https://arxiv.org/abs/1308.0847))  
 Lewis A., Challinor A., Lasenby A., 2000, *ApJ*, 538, 473  
 Lupton R., Gunn J. E., Ivezić Z., Knapp G., Kent S., 2001, in Harnden F. R., Jr., Francis A. P., Harry E. P. *Astronomical Data Analysis Software and Systems X*, Vol. 238, Astron. Soc. Pac., San Francisco, p. 269,  
 Mehta V. et al., 2015, *ApJ*, 811, 141  
 Nishizawa A. J., Oguri M., Takada M., 2013, *MNRAS*, 433, 730  
 Nuza S. E. et al., 2013, *MNRAS*, 432, 743  
 Oliphant T. E., 2007, *Comput. Sci. Eng.*, 9, 10  
 Padmanabhan N. et al., 2008, *ApJ*, 674, 1217  
 Passaglia S., Manzotti A., Dodelson S., 2017, *Phys. Rev. D*, 95, 123508  
 Percival W. et al., 2014, *MNRAS*, 439, 2531  
 Pier J. R., Munn J. A., Hindsley R. B., Hennessy G. S., Kent S. M., Lupton R. H., Ivezić Ž, 2003, *AJ*, 125, 1559  
 Reid B. et al., 2016, *MNRAS*, 455, 1553  
 Smee S. A. et al., 2013, *AJ*, 146, 32  
 Smith J. A. et al., 2002, *AJ*, 123, 2121  
 van der Walt S., Colbert S. C., Varoquaux G., 2011, *Comput. Sci. Eng.*, 13, 22  
 Weaver B. A., Blanton M. R., Brinkmann J., Brownstein J. R., Stauffer F., 2015, *PASP*, 127, 397  
 Weinberg D. H., Mortonson M. J., Eisenstein D. J., Hirata C., Riess, A. G., Rozo E., 2013, *Phys. Rep.*, 530, 87  
 York D. G. et al., 2000, *AJ*, 120, 1579  
 Zehavi I. et al., 2011, *ApJ*, 736, 59

**APPENDIX A: COVARIANCE**

In this Appendix, we develop the covariance matrix for the unbinned statistic, which is given by

$$c(\mathbf{R}, \mathbf{S}) \equiv \langle [W(\mathbf{R}) - \langle W(\mathbf{R}) \rangle][W(\mathbf{S}) - \langle W(\mathbf{S}) \rangle] \rangle = \langle W(\mathbf{R})W(\mathbf{S}) \rangle - \langle W(\mathbf{R}) \rangle \langle W(\mathbf{S}) \rangle. \quad (\text{A1})$$

Recalling the fractional form of  $W$  from equation (40), the numerator of the first term in equation (A1), which we will refer to as  $N_C$  for simplicity (and denote the denominator by  $D_C$ ), may be written out as

$$\langle N_C \rangle = D_C \langle W(\mathbf{R})W(\mathbf{S}) \rangle \quad (\text{A2})$$

$$\begin{aligned} &= \int dr_1 n_s(r_1) \mathcal{W}(r_1, \eta) \int dr'_1 r_1'^2 n_p(r'_1) \int dr_2 n_s(r_2) \mathcal{W}(r_2, \eta) \int dr'_2 r_2'^2 n_p(r'_2) \int d^2\theta_1 \\ &\quad \times \int d^2\theta_2 \langle \delta_s(r_1, \boldsymbol{\theta}_1) \delta_p(r'_1, \boldsymbol{\theta}_1 + \mathbf{R}/r_1) \delta_s(r_2, \boldsymbol{\theta}_2) \delta_p(r'_2, \boldsymbol{\theta}_2 + \mathbf{S}/r_2) \rangle. \end{aligned} \quad (\text{A3})$$

Again using the Dekel & Lahav (1999) formalism so that  $\delta_g = b_g \delta + \epsilon_g$ , this immediately becomes

$$\begin{aligned} \langle N_C \rangle &= b_s^2 b_p^2 \int dr_1 n_s(r_1) \mathcal{W}(r_1, \eta) \int dr'_1 r_1'^2 n_p(r'_1) \int dr_2 n_s(r_2) \mathcal{W}(r_2, \eta) \int dr'_2 r_2'^2 n_p(r'_2) \int d^2\theta_1 \\ &\quad \times \int d^2\theta_2 \langle \delta(r_1, \boldsymbol{\theta}_1) \delta(r'_1, \boldsymbol{\theta}_1 + \mathbf{R}/r_1) \delta(r_2, \boldsymbol{\theta}_2) \delta(r'_2, \boldsymbol{\theta}_2 + \mathbf{S}/r_2) \rangle. \end{aligned} \quad (\text{A4})$$

The expectation value may be expressed as a Fourier transform:

$$\begin{aligned} \langle \delta(r_1, \boldsymbol{\theta}_1) \delta(r'_1, \boldsymbol{\theta}_1 + \mathbf{R}/r_1) \delta(r_2, \boldsymbol{\theta}_2) \delta(r'_2, \boldsymbol{\theta}_2 + \mathbf{S}/r_2) \rangle &= \frac{1}{(2\pi)^6} \int \int \int \int d^3k d^3k' d^3q d^3q' \langle \delta_{\mathbf{k}} \delta_{\mathbf{k}'} \delta_{\mathbf{q}} \delta_{\mathbf{q}'} \rangle e^{i r_1 \mathbf{k}_{\perp} \cdot \boldsymbol{\theta}_1} e^{i r'_1 \mathbf{k}'_{\perp} \cdot (\boldsymbol{\theta}_1 + \mathbf{R}/r_1)} \\ &\quad \times e^{i r_2 \mathbf{q}_{\perp} \cdot \boldsymbol{\theta}_2} e^{i r'_2 \mathbf{q}'_{\perp} \cdot (\boldsymbol{\theta}_2 + \mathbf{S}/r_2)} e^{i r_1 k_{\parallel}} e^{i r'_1 k'_{\parallel}} e^{i r_2 q_{\parallel}} e^{i r'_2 q'_{\parallel}}. \end{aligned} \quad (\text{A5})$$

Assuming a Gaussian random field and using Wick's theorem, we can write the expectation value as

$$\langle \delta_{\mathbf{k}} \delta_{\mathbf{k}'} \delta_{\mathbf{q}} \delta_{\mathbf{q}'} \rangle = \langle \delta_{\mathbf{k}} \delta_{\mathbf{k}'} \rangle \langle \delta_{\mathbf{q}} \delta_{\mathbf{q}'} \rangle + \langle \delta_{\mathbf{k}} \delta_{\mathbf{q}} \rangle \langle \delta_{\mathbf{k}'} \delta_{\mathbf{q}'} \rangle + \langle \delta_{\mathbf{k}} \delta_{\mathbf{q}'} \rangle \langle \delta_{\mathbf{k}'} \delta_{\mathbf{q}} \rangle. \quad (\text{A6})$$

The first term in this expansion can immediately be identified with  $\langle w(\mathbf{R}) \rangle \langle w(\mathbf{S}) \rangle$ , which cancels the second term in equation (A1). Then, we express the remaining factors in terms of the power spectrum:

$$\langle \delta_{\mathbf{k}} \delta_{\mathbf{q}} \rangle \langle \delta_{\mathbf{k}'} \delta_{\mathbf{q}'} \rangle = \delta_D^{(3)}(\mathbf{k} + \mathbf{q}) P(q) \delta_D^{(3)}(\mathbf{k}' + \mathbf{q}') P(q'), \quad (\text{A7})$$

$$\langle \delta_{\mathbf{k}} \delta_{\mathbf{q}'} \rangle \langle \delta_{\mathbf{k}'} \delta_{\mathbf{q}} \rangle = \delta_D^{(3)}(\mathbf{k} + \mathbf{q}') P(q') \delta_D^{(3)}(\mathbf{k}' + \mathbf{q}) P(q). \quad (\text{A8})$$

Let us denote the Fourier transform of the first of these terms by  $A$  and that of the second by  $B$ .

Beginning with term  $A$ , we find

$$A = \frac{1}{(2\pi)^6} \int \int \int \int d^3k d^3k' d^3q d^3q' \langle \delta_{\mathbf{k}} \delta_{\mathbf{q}} \rangle \langle \delta_{\mathbf{k}'} \delta_{\mathbf{q}'} \rangle e^{i r_1 \mathbf{k}_{\perp} \cdot \boldsymbol{\theta}_1} e^{i r'_1 \mathbf{k}'_{\perp} \cdot (\boldsymbol{\theta}_1 + \mathbf{R}/r_1)} e^{i r_2 \mathbf{q}_{\perp} \cdot \boldsymbol{\theta}_2} e^{i r'_2 \mathbf{q}'_{\perp} \cdot (\boldsymbol{\theta}_2 + \mathbf{S}/r_2)} e^{i r_1 k_{\parallel}} e^{i r'_1 k'_{\parallel}} e^{i r_2 q_{\parallel}} e^{i r'_2 q'_{\parallel}}, \quad (\text{A9})$$

$$\begin{aligned} &= \frac{1}{(2\pi)^6} \int \int \int \int d^3k d^3k' d^3q d^3q' \delta_D^{(3)}(\mathbf{k} + \mathbf{q}) P(q) \delta_D^{(3)}(\mathbf{k}' + \mathbf{q}') P(q') e^{i r_1 \mathbf{k}_{\perp} \cdot \boldsymbol{\theta}_1} e^{i r'_1 \mathbf{k}'_{\perp} \cdot (\boldsymbol{\theta}_1 + \mathbf{R}/r_1)} e^{i r_2 \mathbf{q}_{\perp} \cdot \boldsymbol{\theta}_2} e^{i r'_2 \mathbf{q}'_{\perp} \cdot (\boldsymbol{\theta}_2 + \mathbf{S}/r_2)} \\ &\quad \times e^{i r_1 k_{\parallel}} e^{i r'_1 k'_{\parallel}} e^{i r_2 q_{\parallel}} e^{i r'_2 q'_{\parallel}}, \end{aligned} \quad (\text{A10})$$

$$= \frac{1}{(2\pi)^6} \int \int d^3k d^3k' P(k) P(k') e^{i r_1 \mathbf{k}_{\perp} \cdot \boldsymbol{\theta}_1} e^{i r'_1 \mathbf{k}'_{\perp} \cdot (\boldsymbol{\theta}_1 + \mathbf{R}/r_1)} e^{-i r_2 \mathbf{k}_{\perp} \cdot \boldsymbol{\theta}_2} e^{-i r'_2 \mathbf{k}'_{\perp} \cdot (\boldsymbol{\theta}_2 + \mathbf{S}/r_2)} e^{i k_{\parallel} (r_1 - r_2)} e^{i k'_{\parallel} (r'_1 - r'_2)}. \quad (\text{A11})$$

We can again separate the integrals:

$$A = \frac{1}{(2\pi)^6} \int \int d^2k_{\perp} d^2k'_{\perp} P(k_{\perp}) P(k'_{\perp}) e^{i r_1 \mathbf{k}_{\perp} \cdot \boldsymbol{\theta}_1} e^{i r'_1 \mathbf{k}'_{\perp} \cdot (\boldsymbol{\theta}_1 + \mathbf{R}/r_1)} e^{-i r_2 \mathbf{k}_{\perp} \cdot \boldsymbol{\theta}_2} e^{-i r'_2 \mathbf{k}'_{\perp} \cdot (\boldsymbol{\theta}_2 + \mathbf{S}/r_2)} \int dk_{\parallel} e^{i k_{\parallel} (r_1 - r_2)} \int dk'_{\parallel} e^{i k'_{\parallel} (r'_1 - r'_2)}, \quad (\text{A12})$$

$$= \frac{1}{(2\pi)^4} \int \int d^2k_{\perp} d^2k'_{\perp} P(k_{\perp}) P(k'_{\perp}) e^{i r_1 \mathbf{k}_{\perp} \cdot \boldsymbol{\theta}_1} e^{i r'_1 \mathbf{k}'_{\perp} \cdot (\boldsymbol{\theta}_1 + \mathbf{R}/r_1)} e^{-i r_2 \mathbf{k}_{\perp} \cdot \boldsymbol{\theta}_2} e^{-i r'_2 \mathbf{k}'_{\perp} \cdot (\boldsymbol{\theta}_2 + \mathbf{S}/r_2)} \delta_D(r_1 - r_2) \delta_D(r'_1 - r'_2). \quad (\text{A13})$$

We may now insert this expression into equation (A4) to take advantage of the Dirac delta functions:

$$\begin{aligned} \langle A \rangle &= b_s^2 b_p^2 \int dr_1 n_s(r_1) \mathcal{W}(r_1, \eta) \int dr'_1 r_1'^2 n_p(r'_1) \int dr_2 n_s(r_2) \mathcal{W}(r_2, \eta) \int dr'_2 r_2'^2 n_p(r'_2) \frac{1}{(2\pi)^4} \int d^2\theta_1 \int d^2\theta_2 \\ &\quad \times \int \int d^2k_{\perp} d^2k'_{\perp} P(k_{\perp}) P(k'_{\perp}) e^{i r_1 \mathbf{k}_{\perp} \cdot \boldsymbol{\theta}_1} e^{i r'_1 \mathbf{k}'_{\perp} \cdot (\boldsymbol{\theta}_1 + \mathbf{R}/r_1)} e^{-i r_2 \mathbf{k}_{\perp} \cdot \boldsymbol{\theta}_2} e^{-i r'_2 \mathbf{k}'_{\perp} \cdot (\boldsymbol{\theta}_2 + \mathbf{S}/r_2)} \delta_D(r_1 - r_2) \delta_D(r'_1 - r'_2), \end{aligned} \quad (\text{A14})$$

$$\begin{aligned}
 &= \frac{b_s^2 b_p^2}{(2\pi)^4} \int dr_1 n_s(r_1)^2 \mathcal{W}(r_1, \eta)^2 \int dr'_1 r_1'^4 n_p(r'_1)^2 \int \int d^2 k_\perp d^2 k'_\perp P(k_\perp) P(k'_\perp) e^{i(r'_1/r_1) \mathbf{k}'_\perp \cdot \mathbf{R}} e^{-i(r'_1/r_1) \mathbf{k}'_\perp \cdot \mathbf{S}} \\
 &\quad \times \int d^2 \theta_2 \int d^2 \theta_1 e^{i r_1 \mathbf{k}_\perp \cdot (\theta_1 - \theta_2)} e^{i r'_1 \mathbf{k}'_\perp \cdot (\theta_1 - \theta_2)}. \tag{A15}
 \end{aligned}$$

To perform the angular integrals, we define the substitution  $\boldsymbol{\gamma} = \boldsymbol{\theta}_1 - \boldsymbol{\theta}_2$ , which yields

$$\begin{aligned}
 \langle A \rangle &= \frac{b_s^2 b_p^2}{(2\pi)^4} \int dr_1 n_s(r_1)^2 \mathcal{W}(r_1, \eta)^2 \int dr'_1 r_1'^4 n_p(r'_1)^2 \int \int d^2 k_\perp d^2 k'_\perp P(k_\perp) P(k'_\perp) e^{i(r'_1/r_1) \mathbf{k}'_\perp \cdot \mathbf{R}} e^{-i(r'_1/r_1) \mathbf{k}'_\perp \cdot \mathbf{S}} \\
 &\quad \times \int d^2 \theta_2 \int d^2 \gamma e^{i(r_1 \mathbf{k}_\perp + r'_1 \mathbf{k}'_\perp) \cdot \boldsymbol{\gamma}}. \tag{A16}
 \end{aligned}$$

The  $\boldsymbol{\gamma}$  term will integrate to a Dirac delta function, whereas the lone  $\int d^2 \theta_2$  integral will yield the survey area, denoted  $\mathcal{A}$ . Accordingly,

$$\langle A \rangle = \frac{b_s^2 b_p^2}{(2\pi)^4} \int dr_1 n_s(r_1)^2 \mathcal{W}(r_1, \eta)^2 \int dr'_1 r_1'^4 n_p(r'_1)^2 \int \int d^2 k_\perp d^2 k'_\perp P(k_\perp) P(k'_\perp) e^{i(r'_1/r_1) \mathbf{k}'_\perp \cdot (\mathbf{R} - \mathbf{S})} \mathcal{A} (2\pi)^2 \delta_D^{(2)}(r_1 \mathbf{k}_\perp + r'_1 \mathbf{k}'_\perp), \tag{A17}$$

$$= \frac{b_s^2 b_p^2}{(2\pi)^2} \mathcal{A} \int dr_1 n_s(r_1)^2 \mathcal{W}(r_1, \eta)^2 \int dr'_1 r_1'^2 n_p(r'_1)^2 \int d^2 k_\perp P(k_\perp) P(r_1 \mathbf{k}_\perp / r'_1) e^{-i \mathbf{k}_\perp \cdot (\mathbf{R} - \mathbf{S})}. \tag{A18}$$

Now, let  $\mathbf{u} = r_1 \mathbf{k}_\perp$ , so that:

$$\langle A \rangle = \frac{b_s^2 b_p^2}{(2\pi)^2} \mathcal{A} \int dr_1 n_s(r_1)^2 \mathcal{W}(r_1, \eta)^2 \int dr'_1 r_1'^2 n_p(r'_1)^2 \int \frac{d^2 u}{r_1^2} P\left(\frac{u}{r_1}\right) P\left(\frac{u}{r'_1}\right) e^{i u \cdot (\mathbf{S} - \mathbf{R}) / r_1}, \tag{A19}$$

$$= \frac{b_s^2 b_p^2}{(2\pi)^2} \mathcal{A} \int \frac{dr_1}{r_1^2} n_s(r_1)^2 \mathcal{W}(r_1, \eta)^2 \int dr'_1 r_1'^2 n_p(r'_1)^2 \int d^2 u P\left(\frac{u}{r_1}\right) P\left(\frac{u}{r'_1}\right) e^{i u \cdot (\mathbf{S} - \mathbf{R}) / r_1}. \tag{A20}$$

We proceed along similar lines for term  $B$ .

$$B = \frac{1}{(2\pi)^6} \int \int \int d^3 k d^3 k' d^3 q d^3 q' \langle \delta_k \delta_{q'} \rangle \langle \delta_{k'} \delta_q \rangle e^{i r_1 \mathbf{k}_\perp \cdot \boldsymbol{\theta}_1} e^{i r'_1 \mathbf{k}'_\perp \cdot (\boldsymbol{\theta}_1 + \mathbf{R} / r_1)} e^{i r_2 \mathbf{q}_\perp \cdot \boldsymbol{\theta}_2} e^{i r'_2 \mathbf{q}'_\perp \cdot (\boldsymbol{\theta}_2 + \mathbf{S} / r_2)} e^{i r_1 \mathbf{k}_\parallel} e^{i r'_1 \mathbf{k}'_\parallel} e^{i r_2 \mathbf{q}_\parallel} e^{i r'_2 \mathbf{q}'_\parallel}, \tag{A21}$$

$$\begin{aligned}
 &= \frac{1}{(2\pi)^6} \int \int \int d^3 k d^3 k' d^3 q d^3 q' \delta_D^{(3)}(\mathbf{k} + \mathbf{q}') P(q') \delta_D^{(3)}(\mathbf{k}' + \mathbf{q}) P(q) e^{i r_1 \mathbf{k}_\perp \cdot \boldsymbol{\theta}_1} e^{i r'_1 \mathbf{k}'_\perp \cdot (\boldsymbol{\theta}_1 + \mathbf{R} / r_1)} e^{i r_2 \mathbf{q}_\perp \cdot \boldsymbol{\theta}_2} e^{i r'_2 \mathbf{q}'_\perp \cdot (\boldsymbol{\theta}_2 + \mathbf{S} / r_2)} \\
 &\quad \times e^{i r_1 \mathbf{k}_\parallel} e^{i r'_1 \mathbf{k}'_\parallel} e^{i r_2 \mathbf{q}_\parallel} e^{i r'_2 \mathbf{q}'_\parallel}, \tag{A22}
 \end{aligned}$$

$$= \frac{1}{(2\pi)^6} \int \int d^3 k d^3 k' P(k) P(k') e^{i r_1 \mathbf{k}_\perp \cdot \boldsymbol{\theta}_1} e^{i r'_1 \mathbf{k}'_\perp \cdot (\boldsymbol{\theta}_1 + \mathbf{R} / r_1)} e^{-i r_2 \mathbf{k}_\perp \cdot \boldsymbol{\theta}_2} e^{-i r'_2 \mathbf{k}'_\perp \cdot (\boldsymbol{\theta}_2 + \mathbf{S} / r_2)} e^{i(r_1 - r'_2) \mathbf{k}_\parallel} e^{i(r'_1 - r_2) \mathbf{k}'_\parallel}. \tag{A23}$$

We split the integrals, assuming small angles:

$$B = \frac{1}{(2\pi)^6} \int \int d^2 k_\perp d^2 k'_\perp P(k_\perp) P(k'_\perp) e^{i r_1 \mathbf{k}_\perp \cdot \boldsymbol{\theta}_1} e^{i r'_1 \mathbf{k}'_\perp \cdot (\boldsymbol{\theta}_1 + \mathbf{R} / r_1)} e^{-i r_2 \mathbf{k}_\perp \cdot \boldsymbol{\theta}_2} e^{-i r'_2 \mathbf{k}'_\perp \cdot (\boldsymbol{\theta}_2 + \mathbf{S} / r_2)} \int d k_\parallel e^{i(r_1 - r'_2) \mathbf{k}_\parallel} \int d k'_\parallel e^{i(r'_1 - r_2) \mathbf{k}'_\parallel}, \tag{A24}$$

$$= \frac{1}{(2\pi)^4} \int \int d^2 k_\perp d^2 k'_\perp P(k_\perp) P(k'_\perp) e^{i r_1 \mathbf{k}_\perp \cdot \boldsymbol{\theta}_1} e^{i r'_1 \mathbf{k}'_\perp \cdot (\boldsymbol{\theta}_1 + \mathbf{R} / r_1)} e^{-i r_2 \mathbf{k}_\perp \cdot \boldsymbol{\theta}_2} e^{-i r'_2 \mathbf{k}'_\perp \cdot (\boldsymbol{\theta}_2 + \mathbf{S} / r_2)} \delta_D(r_1 - r'_2) \delta_D(r'_1 - r_2). \tag{A25}$$

Plugging this into equation (A4) yields

$$\begin{aligned}
 \langle B \rangle &= b_s^2 b_p^2 \int dr_1 n_s(r_1) \mathcal{W}(r_1, \eta) \int dr'_1 r_1'^2 n_p(r'_1) \int dr_2 n_s(r_2) \mathcal{W}(r_2, \eta) \int dr'_2 r_2'^2 n_p(r'_2) \frac{1}{(2\pi)^4} \int d^2 \theta_1 \int d^2 \theta_2 \\
 &\quad \times \int \int d^2 k_\perp d^2 k'_\perp P(k_\perp) P(k'_\perp) e^{i r_1 \mathbf{k}_\perp \cdot \boldsymbol{\theta}_1} e^{i r'_1 \mathbf{k}'_\perp \cdot (\boldsymbol{\theta}_1 + \mathbf{R} / r_1)} e^{-i r_2 \mathbf{k}_\perp \cdot \boldsymbol{\theta}_2} e^{-i r'_2 \mathbf{k}'_\perp \cdot (\boldsymbol{\theta}_2 + \mathbf{S} / r_2)} \delta_D(r_1 - r'_2) \delta_D(r'_1 - r_2), \tag{A26}
 \end{aligned}$$

$$\begin{aligned}
 &= \frac{b_s^2 b_p^2}{(2\pi)^4} \int dr_1 r_1^2 n_s(r_1) \mathcal{W}(r_1, \eta) n_p(r_1) \int dr'_1 r_1'^2 n_s(r'_1) \mathcal{W}(r'_1, \eta) n_p(r'_1) \int \int d^2 k_\perp d^2 k'_\perp P(k_\perp) P(k'_\perp) e^{i(r'_1/r_1) \mathbf{k}'_\perp \cdot \mathbf{R}} e^{-i(r_1/r'_1) \mathbf{k}_\perp \cdot \mathbf{S}} \\
 &\quad \times \int d^2 \theta_1 e^{i(r_1 \mathbf{k}_\perp + r'_1 \mathbf{k}'_\perp) \cdot \boldsymbol{\theta}_1} \int d^2 \theta_2 e^{-i(r_1 \mathbf{k}_\perp + r'_1 \mathbf{k}'_\perp) \cdot \boldsymbol{\theta}_2}, \tag{A27}
 \end{aligned}$$

$$\begin{aligned}
 &= \frac{b_s^2 b_p^2}{(2\pi)^4} \int dr_1 r_1^2 n_s(r_1) \mathcal{W}(r_1, \eta) n_p(r_1) \int dr'_1 r_1'^2 n_s(r'_1) \mathcal{W}(r'_1, \eta) n_p(r'_1) \int \int d^2 k_\perp d^2 k'_\perp P(k_\perp) P(k'_\perp) e^{i(r'_1/r_1) \mathbf{k}'_\perp \cdot \mathbf{R}} e^{-i(r_1/r'_1) \mathbf{k}_\perp \cdot \mathbf{S}} \\
 &\quad \times \mathcal{A} (2\pi)^2 \delta_D^{(2)}(r_1 \mathbf{k}_\perp + r'_1 \mathbf{k}'_\perp), \tag{A28}
 \end{aligned}$$

$$= \frac{b_s^2 b_p^2}{(2\pi)^2} \mathcal{A} \int dr_1 r_1^2 n_s(r_1) \mathcal{W}(r_1, \eta) n_p(r_1) \int dr'_1 n_s(r'_1) \mathcal{W}(r'_1, \eta) n_p(r'_1) \int d^2 k_\perp P(k_\perp) P(r_1 \mathbf{k}_\perp / r'_1) e^{-i \mathbf{k}_\perp \cdot \mathbf{R}} e^{-i(r_1/r'_1) \mathbf{k}_\perp \cdot \mathbf{S}}. \tag{A29}$$

We set  $\mathbf{u} = r_1 \mathbf{k}_\perp$  again, so that

$$\langle B \rangle = \frac{b_s^2 b_p^2}{(2\pi)^2} \mathcal{A} \int dr_1 r_1^2 n_s(r_1) \mathcal{W}(r_1, \eta) n_p(r_1) \int dr'_1 n_s(r'_1) \mathcal{W}(r'_1, \eta) n_p(r'_1) \int d^2 u \frac{1}{r_1^2} P\left(\frac{u}{r_1}\right) P\left(\frac{u}{r'_1}\right) e^{-i\mathbf{u} \cdot \mathbf{R}/r_1} e^{-i\mathbf{u} \cdot \mathbf{S}/r'_1}, \quad (\text{A30})$$

$$= \frac{b_s^2 b_p^2}{(2\pi)^2} \mathcal{A} \int dr_1 n_s(r_1) \mathcal{W}(r_1, \eta) n_p(r_1) \int dr'_1 n_s(r'_1) \mathcal{W}(r'_1, \eta) n_p(r'_1) \int d^2 u P\left(\frac{u}{r_1}\right) P\left(\frac{u}{r'_1}\right) e^{-i\mathbf{u} \cdot (\mathbf{R}/r_1 + \mathbf{S}/r'_1)} \quad (\text{A31})$$

Lastly, the denominator of the covariance is given by

$$D_C = \int dr_1 n_s(r_1) \mathcal{W}(r_1, \eta) \int dr'_1 r_1^2 n_p(r'_1) \int dr_2 n_s(r_2) \mathcal{W}(r_2, \eta) \int dr'_2 r_2^2 n_p(r'_2) \int d^2 \theta_1 \int d^2 \theta_2, \quad (\text{A32})$$

$$= \mathcal{A}^2 \int dr_1 n_s(r_1) \mathcal{W}(r_1, \eta) \int dr'_1 r_1^2 n_p(r'_1) \int dr_2 n_s(r_2) \mathcal{W}(r_2, \eta) \int dr'_2 r_2^2 n_p(r'_2). \quad (\text{A33})$$

Accordingly, with  $c(\mathbf{R}, \mathbf{S}) = \langle N_C \rangle / D_C$ , we obtain

$$\begin{aligned} c(\mathbf{R}, \mathbf{S}) &= \frac{b_s^2 b_p^2}{4\pi^2 \mathcal{A}} \left( \int \frac{dr_1}{r_1^2} n_s(r_1)^2 \mathcal{W}(r_1, \eta)^2 \int dr'_1 r_1^2 n_p(r'_1)^2 \int d^2 u P\left(\frac{u}{r_1}\right) P\left(\frac{u}{r'_1}\right) e^{i\mathbf{u} \cdot (\mathbf{S} - \mathbf{R})/r_1} \right. \\ &\quad \left. + \int dr_1 n_s(r_1) \mathcal{W}(r_1, \eta) n_p(r_1) \int dr'_1 n_s(r'_1) \mathcal{W}(r'_1, \eta) n_p(r'_1) \int d^2 u P\left(\frac{u}{r_1}\right) P\left(\frac{u}{r'_1}\right) e^{-i\mathbf{u} \cdot (\mathbf{R}/r_1 + \mathbf{S}/r'_1)} \right) \\ &\quad \times \left[ \int dr_1 n_s(r_1) \mathcal{W}(r_1, \eta) \int dr'_1 r_1^2 n_p(r'_1) \int dr_2 n_s(r_2) \mathcal{W}(r_2, \eta) \int dr'_2 r_2^2 n_p(r'_2) \right]^{-1}. \end{aligned} \quad (\text{A34})$$

We see that the covariance matrix has two distinct terms: the first integrates over the spectroscopic and photometric samples separately, whereas the second mixes the two samples.

## APPENDIX B: BINNED CORRELATION FUNCTION

We produce a binned correlation function as follows:

$$w(R_1, R_2) = 2 \int_{R_1}^{R_2} \frac{R dR}{(R_2^2 - R_1^2)} \int \frac{d\phi}{2\pi} \langle W(\mathbf{R}) \rangle. \quad (\text{B1})$$

Inserting the form in equation (40) gives

$$\langle W(\mathbf{R}) \rangle = \frac{b_s b_p}{2\pi} w_p(\mathbf{R}) \frac{\int dr r^2 n_s(r) \mathcal{W}(r, \eta) n_p(r)}{\int dr n_s(r) \mathcal{W}(r, \eta) \int dr' r'^2 n_p(r')} = \frac{b_s b_p}{(2\pi)^2} \int d^2 k'_\perp P(k'_\perp) e^{i\mathbf{k}'_\perp \cdot \mathbf{R}} \frac{\int dr r^2 n_s(r) \mathcal{W}(r, \eta) n_p(r)}{\int dr n_s(r) \mathcal{W}(r, \eta) \int dr' r'^2 n_p(r')}, \quad (\text{B2})$$

and writing  $\mathbf{k}'_\perp \cdot \mathbf{R} = k'_\perp R \cos(\phi)$ , we obtain the following representation in terms of the Bessel function  $J_0$ :

$$w(R_1, R_2) = \frac{b_s b_p}{4\pi^3 (R_2^2 - R_1^2)} \frac{\int dr r^2 n_s(r) \mathcal{W}(r, \eta) n_p(r)}{\int dr n_s(r) \mathcal{W}(r, \eta) \int dr' r'^2 n_p(r')} \int d^2 k'_\perp P(k'_\perp) \int_{R_1}^{R_2} R dR \int d\phi e^{i k'_\perp R \cos(\phi)}, \quad (\text{B3})$$

$$w(R_1, R_2) = \frac{b_s b_p}{4\pi^3 (R_2^2 - R_1^2)} \frac{\int dr r^2 n_s(r) \mathcal{W}(r, \eta) n_p(r)}{\int dr n_s(r) \mathcal{W}(r, \eta) \int dr' r'^2 n_p(r')} \int d^2 k'_\perp P(k'_\perp) \int_{R_1}^{R_2} R dR 2\pi J_0(k'_\perp R). \quad (\text{B4})$$

The integral over  $R$  can now be done by making use of the properties of Bessel functions, leading to a final result of

$$w(R_1, R_2) = \frac{b_s b_p}{2\pi^2 (R_2^2 - R_1^2)} \frac{\int dr r^2 n_s(r) \mathcal{W}(r, \eta) n_p(r)}{\int dr n_s(r) \mathcal{W}(r, \eta) \int dr' r'^2 n_p(r')} \int d^2 k'_\perp P(k'_\perp) \left[ \frac{R_2 J_1(k'_\perp R_2) - R_1 J_1(k'_\perp R_1)}{k'_\perp} \right]. \quad (\text{B5})$$

## APPENDIX C: BINNED COVARIANCE MATRIX

Lastly, we compute the binned covariance matrix. In this case, we have two terms to consider (temporarily omitting constant values to focus on the derivation):

$$c_1 \sim \int d^2 u P\left(\frac{u}{r_1}\right) P\left(\frac{u}{r'_1}\right) e^{i\mathbf{u} \cdot (\mathbf{S} - \mathbf{R})/r_1}, \quad (\text{C1})$$

$$c_2 \sim \int d^2 u P\left(\frac{u}{r_1}\right) P\left(\frac{u}{r'_1}\right) e^{-i\mathbf{u} \cdot (\mathbf{R}/r_1 + \mathbf{S}/r'_1)}. \quad (\text{C2})$$

Accordingly, for the first term, we find

$$C_1 \sim 4 \int_{R_1}^{R_2} \frac{R dR}{(R_2^2 - R_1^2)} \int \frac{d\phi_R}{2\pi} \int_{S_1}^{S_2} \frac{S dS}{(S_2^2 - S_1^2)} \int \frac{d\phi_S}{2\pi} \int d^2 u P\left(\frac{u}{r_1}\right) P\left(\frac{u}{r'_1}\right) e^{i\mathbf{u} \cdot (\mathbf{S} - \mathbf{R})/r_1}, \quad (\text{C3})$$

which, with  $\mathbf{u} \cdot \mathbf{R} = uR \cos(\phi_R)$  and  $\mathbf{u} \cdot \mathbf{S} = uS \cos(\phi_S)$ , becomes

$$C_1 \sim \frac{1}{\pi^2 (R_2^2 - R_1^2) (S_2^2 - S_1^2)} \int d^2u P\left(\frac{u}{r_1}\right) P\left(\frac{u}{r'_1}\right) \int R dR \int d\phi_R e^{iuR \cos(\phi_R)/r_1} \int S dS \int d\phi_S e^{-iuS \cos(\phi_S)/r_1}, \quad (C4)$$

$$\sim \frac{4}{(R_2^2 - R_1^2) (S_2^2 - S_1^2)} \int d^2u P\left(\frac{u}{r_1}\right) P\left(\frac{u}{r'_1}\right) \left[\frac{r_1^2}{u^2}\right] \left[ R_2 J_1\left(\frac{uR_2}{r_1}\right) - R_1 J_1\left(\frac{uR_1}{r_1}\right) \right] \left[ S_2 J_1\left(\frac{uS_2}{r_1}\right) - S_1 J_1\left(\frac{uS_1}{r_1}\right) \right]. \quad (C5)$$

For the second term,

$$C_2 \sim 4 \int_{R_1}^{R_2} \frac{R dR}{(R_2^2 - R_1^2)} \int \frac{d\phi_R}{2\pi} \int_{S_1}^{S_2} \frac{S dS}{(S_2^2 - S_1^2)} \int \frac{d\phi_S}{2\pi} \int d^2u P\left(\frac{u}{r_1}\right) P\left(\frac{u}{r'_1}\right) e^{-iu \cdot (\mathbf{R}/r_1 + \mathbf{S}/r'_1)}, \quad (C6)$$

$$\sim \frac{1}{\pi^2 (R_2^2 - R_1^2) (S_2^2 - S_1^2)} \int d^2u P\left(\frac{u}{r_1}\right) P\left(\frac{u}{r'_1}\right) \int R dR \int d\phi_R e^{-iuR \cos(\phi_R)/r_1} \int S dS \int d\phi_S e^{-iuS \cos(\phi_S)/r'_1}, \quad (C7)$$

$$\sim \frac{4}{(R_2^2 - R_1^2) (S_2^2 - S_1^2)} \int d^2u P\left(\frac{u}{r_1}\right) P\left(\frac{u}{r'_1}\right) \left[\frac{r_1 r'_1}{u^2}\right] \left[ R_2 J_1\left(\frac{uR_2}{r_1}\right) - R_1 J_1\left(\frac{uR_1}{r_1}\right) \right] \left[ S_2 J_1\left(\frac{uS_2}{r'_1}\right) - S_1 J_1\left(\frac{uS_1}{r'_1}\right) \right]. \quad (C8)$$

Thus, the binned covariance becomes (cf. equation A34)

$$\begin{aligned} C(R_1, R_2, S_1, S_2) &= \frac{b_s^2 b_p^2 \left[ \int dr_1 n_s(r_1) \mathcal{W}(r_1, \eta) \int dr'_1 r_1'^2 n_p(r'_1) \int dr_2 n_s(r_2) \mathcal{W}(r_2, \eta) \int dr'_2 r_2'^2 n_p(r'_2) \right]^{-1}}{\pi^2 \mathcal{A} (R_2^2 - R_1^2) (S_2^2 - S_1^2)} \\ &\quad \times \left( \int dr_1 n_s(r_1)^2 \mathcal{W}(r_1, \eta)^2 \int dr'_1 r_1'^2 n_p(r'_1)^2 \int \frac{d^2u}{u^2} P\left(\frac{u}{r_1}\right) P\left(\frac{u}{r'_1}\right) \right. \\ &\quad \times \left[ R_2 J_1\left(\frac{uR_2}{r_1}\right) - R_1 J_1\left(\frac{uR_1}{r_1}\right) \right] \left[ S_2 J_1\left(\frac{uS_2}{r_1}\right) - S_1 J_1\left(\frac{uS_1}{r_1}\right) \right] \\ &\quad + \int dr_1 r_1 n_s(r_1) \mathcal{W}(r_1, \eta) n_p(r_1) \int dr'_1 r_1' n_s(r'_1) \mathcal{W}(r'_1, \eta) n_p(r'_1) \int \frac{d^2u}{u^2} P\left(\frac{u}{r_1}\right) P\left(\frac{u}{r'_1}\right) \\ &\quad \times \left[ R_2 J_1\left(\frac{uR_2}{r_1}\right) - R_1 J_1\left(\frac{uR_1}{r_1}\right) \right] \left[ S_2 J_1\left(\frac{uS_2}{r'_1}\right) - S_1 J_1\left(\frac{uS_1}{r'_1}\right) \right] \left. \right). \quad (C9) \end{aligned}$$

This paper has been typeset from a  $\text{\TeX}/\text{\LaTeX}$  file prepared by the author.