Optimal redshift weighting for baryon acoustic oscillations

Fangzhou Zhu,† Nikhil Padmanabhan and Martin White

†Department of Physics, Yale University, New Haven, CT 06511, USA
2Department of Physics and Astronomy, U.C. Berkeley, Berkeley, CA, USA

Accepted 2015 April 29. Received 2015 April 16; in original form 2014 November 10

ABSTRACT
Future baryon acoustic oscillation (BAO) surveys will survey very large volumes, covering wide ranges in redshift. We derive a set of redshift weights to compress the information in the redshift direction to a small number of modes. We suggest that such a compression preserves almost all of the signal for most cosmologies, while giving high signal-to-noise measurements for each combination. We present some toy models and simple worked examples. As an intermediate step, we give a precise meaning to the ‘effective redshift’ of a BAO measurement.

Key words: cosmological parameters – dark energy – distance scale.

1 INTRODUCTION
The baryon acoustic oscillation (BAO) method has now become an essential component of the dark energy toolbox by mapping the expansion history of the Universe (see Weinberg et al. 2013; Mortonson, Weinberg & White 2014, for recent reviews). Current sets of surveys (Blake et al. 2012; Anderson et al. 2014) have achieved distance accuracies at the per cent level when averaged across relatively wide redshift bins; the next generation of surveys (Laureijs et al. 2011; Sugai et al. 2012; Levi et al. 2013; Spergel et al. 2013) aim to improve both the redshift resolution and the redshift range of these measurements. This paper aims to develop techniques to robustly map the distance–redshift relation for these surveys.

Conceptually, one can improve the resolution of the reconstructed distance–redshift relation by splitting samples into narrower redshift bins and repeating the traditional analysis on these. However, such a procedure has a significant disadvantage for BAO measurements: the robustness of the BAO technique derives from focusing only on the BAO feature in the correlation function or power spectrum, marginalizing over the broad-band shape. A reliable measurement of distances therefore requires a significant detection of the BAO feature in the correlation function and low signal-to-noise detection can lead to very non-Gaussian likelihood functions for the inferred distance (e.g. Blake et al. 2012; Xu et al. 2013) Splitting samples into multiple bins exacerbates this problem. Binning also has other disadvantages: the choice of bins is often arbitrary and splitting a sample into completely disjoint bins loses signal across bin boundaries unless a large number of cross-spectra are included. While some of these disadvantages may be overcome, they can add significantly to the complexity of the analysis.

Rather than breaking the sample into multiple redshift slices, we instead seek a compression of the information in the redshift direction to a small number of ‘modes’, or weighted integrals. This allows us to work with relatively wide redshift bins while retaining sensitivity to evolution of the distance–redshift relation and gives a firm meaning to the ‘effective redshift’ of the sample. Our approach is reminiscent of principal component analysis or optimal sub-space filtering, though being technically somewhat different.

To do so, we start by parametrizing the distance–redshift relation relative to a fiducial choice of cosmology (see Percival et al. 2007, for related efforts to fit a distance–redshift relation using BAO measurements). We demonstrate that simple parametrizations capture a broad range of possible distance–redshift relations over wide redshift intervals. Given a particular choice of parametrization, we then derive appropriate weighted integrals (in redshift) of the correlation function that are (under particular assumptions) optimal estimates of the parameters in the distance–redshift relation. These integrals may be thought of as generalizations of simple redshift bins, designed to efficiently combine the information at different redshifts.

While we argue in this paper for a model-agnostic parametrization, different choices may be more appropriate for more specific applications. For instance, one may choose the equation of state of dark energy to parametrize the distance–redshift relation. The modes would then be different from the ones we focus on for much of the presentation, but the overall approach would remain unchanged. We give a simple example later to illustrate this flexibility.

The outline of this paper is as follows. We introduce our parametrization of the distance–redshift relation in Section 2, showing that it fits a wide range of cosmologies to the sub-per cent level. Section 3 introduces the weighted correlation functions and our optimal weights, including a toy model to demonstrate the improvements that can be expected by weighting. Section 4 specializes our discussion to the case of BAO and describes a possible implementation of the method. We conclude in Section 5 with a discussion of our results and directions for future development. Some technical details are relegated to an appendix.

* E-mail: fangzhou.zhu@yale.edu
2 PARAMETRIZING THE DISTANCE–REDSHIFT RELATION

We begin by parametrizing the distance–redshift relation as a small fluctuation about a fiducial relation. This scheme is standard in BAO studies, as one typically uses a fiducial cosmology to convert the angular positions and redshifts of the galaxies into 3D positions before measuring the clustering signal. Motivated by inflation and supported by recent observations (Planck Collaboration et al. 2013; Anderson et al. 2014) we shall assume throughout that the spatial hypersurfaces are flat, and we shall denote the comoving radial distance as $\chi$.

We choose a pivot redshift $z_0$ within the redshift range of a planned survey and denote the fiducial comoving distance at $z_0$ by $\chi_{1.0}$. We defer the reader to the end of this section for a discussion about the choice of $z_0$, while merely pointing out the analysis below does not depend on a particular choice of $z_0$. We model the true distance–redshift relation $\chi(z)$ as

$$\frac{\chi(z)}{\chi(z_0)} = \alpha_0 \left(1 + \alpha_1 x + \frac{1}{2} \alpha_2 x^2\right),$$

(1)

where $1 + x \equiv \chi(z)/\chi_{1.0}$. The parameters $\alpha_0$, $\alpha_1$, and $\alpha_2$ are specified by matching the distance–redshift relation and its derivatives at $z_0$:

$$\alpha_0 = \frac{\chi_0}{\chi_{1.0}}$$

(2)

$$\alpha_1 = \frac{H_{1.0}\chi_{1.0}}{H_0\chi_0} - 1$$

(3)

$$\alpha_2 = (1 + \alpha_1)\chi_{1.0} \left[H_{1.0}' - \alpha_0(1 + \alpha_1)H_0\right] - 2\alpha_1,$$

(4)

where $H_{1.0}' = H'(z_0)$ and $H_0' = H'(z_0)$. An explicit calculation of $\alpha_0$ and $\alpha_2$ is in the appendix.

This parametrization to the second order can recover the distance–redshift relation to sub-per cent levels across a wide range of redshifts and cosmologies. Fig. 1 displays the fractional error in the parametrization between $z = 0.5$ and 1.5 for three example cosmologies. These span a wide range of cosmological parameters and are not meant to represent currently viable models. Even for the extreme $\Omega_m = 1$ case, the error is less than 0.5 per cent over the entire redshift range.

Fig. 2 shows a similar plot of the fractional error in the Hubble parameter $H(z) = 1/\chi'(z)$. A second-order approximation in distance translates to a first order one in $H$ and explains the larger errors. Despite that, for cosmologies close to the fiducial model, the accuracy is still at the sub-per cent level.

It is worth noting at this point that $z_0$ is a metaparameter which can be chosen for convenience, rather than a property of the sample under consideration. For best results the fiducial redshift should be chosen close to the centre of the redshift range, for example at the ‘effective’ or pair-weighted redshift of the sample. Fig. 3 highlights this freedom in the choice of $z_0$, demonstrating that the reconstructed distance–redshift relations in each of these cases agree to better than 0.02 per cent, much smaller than the expected statistical uncertainty in these measurements.

3 OPTIMAL WEIGHTS

3.1 Derivation

We start by deriving the general formalism for optimal linear weights. The derivation is modelled on Tegmark, Taylor & Heavens (1997); we summarize it here for completeness, while specializing to the particular problem at hand. Imagine measuring the correlation function in $n$ $z$-bins. The uncompressed correlation function data vector is

$$\Xi(r) = \begin{pmatrix} \xi(r, z_1) \\ \xi(r, z_2) \\ \vdots \\ \xi(r, z_n) \end{pmatrix}.$$

(5)

Weighting the redshift slices by the n-dimensional vector $w = (w_1, w_2, \ldots, w_n)$, we linearly compress the data into one single $z$-bin according to these weights.

$$\xi_w(r) = w^T \Xi(r)$$

(6)

where $\xi_w$ is the weighted correlation function. The mean and covariance $C_w$ of the compressed data are

$$\langle \xi_w(r) \rangle = w^T \langle \Xi(r) \rangle$$

(7)
The impact of changing $\zeta_0$ on the distance–redshift relation. The fiducial cosmology here is the same as the previous two figures, while the true cosmology has $\Omega_M = 0.3$, $h = 0.7$ and $w = -1$. The dashed (blue) line assumes $\zeta_0 = 1.08$ (equal to the effective redshift), the solid (red) line corresponds to $\zeta_0 = 0.75$, while the dotted (green) line assumes $\zeta_0 = 1.25$. The reconstructed distance–redshift relations all agree to better than 0.02 per cent over the entire redshift range.

\[
\langle \xi_w(r) \xi_w^*(r) \rangle - \langle \xi_w(r) \rangle \langle \xi_w^*(r) \rangle = w^T C w, \tag{8}
\]

where $C$ is an $n \times n$ matrix and $C_w = w^T C w$ is a scalar. We write our model correlation functions at distance $r$ in different $z$-bins $M(r)$ in the same format as $C$. Using the redshift weights $w$, we linearly compress the model correlation function data in the same fashion,

\[
m(r) = w^T M(r). \tag{9}
\]

The model correlation function depends on parameters $\Theta = \{\theta_1, \theta_2, \ldots\}$. Restricting to a single unknown parameter, the error on $\theta_i$ is $(F_{ii})^{-1/2}$, where $F_{ii}$ is the $i$th diagonal element of the Fisher information matrix. This is given by

\[
F_{ii} = m_{ii} = \frac{\partial}{\partial \alpha_i} m(\alpha_i) \tag{10}
\]

where $m_{ii}$ is the derivative of the model correlation function $m$ with respect to parameter $\theta_i$. Since $C_w$ is a scalar,

\[
F_{ii} = m_{ii} = w^T M_{ii} w = m_{ii} \tag{11}
\]

where $M_{ii} = m \partial M/\partial \alpha_i$ is an $n \times n$ matrix. Our goal is to find the weights that minimize the error of $\theta_i$, and therefore maximize $F_{ii}$.

Since $F_{ii}$ is unchanged under a rescaling of $w$, we constrain $w^T C w = 1$ for convenience. Introducing a Lagrange multiplier $\lambda$, we maximize

\[
G = w^T M_{ii} w - \lambda (w^T C w - 1). \tag{12}
\]

Setting the gradient of $G$ with respect to $w$ to zero yields

\[
M_{ii} w = \lambda C w. \tag{13}
\]

Using $M_{ii} w = M_{ii} M_{ii} w$ and as $M_{ii}$ is a scalar, we have

\[
C w \propto M_{ii}, \tag{14}
\]

so that the weights are (up to normalizations)

\[
w = C^{-1} M_{ii}. \tag{15}
\]

For these weights, the Fisher matrix is

\[
F_{ii} = M_{ii} C^{-1} M_{ii}, \tag{16}
\]

while the error on $\theta_i$ is

\[
\Delta \theta_i = (M_{ii} C^{-1} M_{ii})^{-1/2}. \tag{17}
\]

It is worth noting that the Fisher matrix is identical to what we would have without any compression. Our compression is therefore theoretically lossless, and has extracted all the information in the original data.

### 3.2 A toy model

In order to build intuition for the results of the previous section, we first apply this to a toy model, which captures the essential aspects of the full BAO problem. We assume that the correlation function at a separation $s$ and redshift $z$ is a Gaussian peaked at $s_0$:

\[
\xi(s, z) = \exp\left[-(s - s_0)^2\right]. \tag{18}
\]

We also assume that choosing an incorrect cosmology (here parametrized simply by $\alpha_1$) changes the separation $s \rightarrow s - \alpha_1 z$,

\[
\xi(s) \rightarrow \xi(s - \alpha_1 z). \tag{19}
\]

Our goal is therefore to estimate $\alpha_1$ from observations of $\xi$.

There are three cases we consider. The first (‘Full’) is where we imagine fitting to the correlation function in both $s$ and $z$; this captures all the information and should yield the most precise measurements. This would be equivalent to splitting the galaxy sample into infinitesimal redshift bins and fitting a full model to it, which is not possible for the real correlation function for the reasons described in Section 1. The second (‘Unweighted’) is to integrate the two-dimensional correlation function $\xi(s, z)$ along the $z$-axis to collapse it down to a one-dimensional function $\xi_w(s) \equiv \int dz \xi(s, z)$.

The third case (‘Weighted’) is to construct the weighted combination $\xi_w(s) \equiv \int dz w(z) \xi(s, z)$ where $w(z)$ are the weights computed in the previous section. We note that we only combine the data in the $z$ direction here, even though the formalism of the previous section would have allowed us to collapse in both $s$ and $z$. We do this because the shape of the true galaxy correlation function is uncertain (hence the marginalization over nuisance shape parameters). Furthermore, as written, the weights do not have any $s$ dependence. While not generic, this is explicitly true for our toy model (as described below). We also construct the weights for the actual correlation function (described in the next section) to have this property.

We can now derive the optimal redshift weights for $\alpha_1$. For simplicity, we will assume that the measurements of $\xi(s, z)$ have constant, uncorrelated errors. In this case, the weights are simply proportional to $d^2 \xi / dz^2$, giving

\[
w(z) = -2(z - \alpha_1 z) \exp\left[-(s - s_0 - \alpha_1 z)^2\right]. \tag{20}
\]

The above derivation does still depend on $\alpha_1$. We make the usual approximation that we are making small perturbations to our fiducial cosmology and set $\alpha_1 = 0$. If this were not true (i.e. the measured values of $\alpha_1$ were significantly different from 0), one would need to iterate using an improved estimate of the true cosmology. Substituting into the above equation, and using the fact that the normalization of the weights is arbitrary, we find the simple relation

\[
w(z) = z. \tag{21}
\]

This weighting makes physical sense; one wants to downweight measurements close to $z = 0$, where the correlation function has little sensitivity to $\alpha_1$, and upweight the measurements as $z$ increases, where the sensitivity is greater. The optimal weight is nothing but a matched filter, tracing the redshift dependence of the signal.
Optimal redshift weighting for baryon acoustic oscillations

4 REDSHIFT WEIGHTING FOR BAO

4.1 Preliminaries

Motivated by the results of the previous section, we turn to the construction of similar weights for BAO measurements. We lay out some basic preliminaries here, with the derivation in the next section. We will assume that the distance–redshift relation is parametrized as in Section 2.

In order to construct a redshift-dependent model for the correlation function, we work in the plane parallel approximation and parametrize the effects of an incorrect choice of a distance–redshift relationship by an isotropic dilation ($\alpha(z)$) and an anisotropic warping ($\epsilon(z)$) parameter (Padmanabhan & White 2008). However, unlike previous analyses which have used this parametrization, we will explicitly assume that these are redshift dependent (although we will often suppress this dependence for notational brevity). These distortion parameters can be related to stretching in the perpendicular and parallel directions by

\[ r_\parallel = \alpha(1 + \epsilon)^2 r_\parallel^0 \]
\[ r_\perp = \alpha(1 + \epsilon)^{-1} r_\perp^0 \]

and the inverse relations by

\[ \alpha = \left( \frac{r_\parallel}{r_\parallel^0} \right) \left( \frac{r_\perp}{r_\perp^0} \right)^{2/3} \]
\[ \epsilon = \left( \frac{r_\parallel}{r_\parallel^0} \right)^{1/3} - 1. \]

In the plane parallel limit, we have

\[ \frac{r_\parallel}{r_\parallel^0} = \frac{H(z)}{H(z)} \]
\[ \frac{r_\perp}{r_\perp^0} = \frac{\chi(z)}{\chi(z)} \]

giving us

\[ \alpha(z) = \left( \frac{H(z)\chi^2(z)}{H(z)\chi^2(z)} \right)^{1/3} \]
\[ \epsilon(z) = \left( \frac{H(z)\chi(z)}{H(z)\chi(z)} \right)^{1/3} - 1. \]

Using the definitions above, and the ratios $H(z)/H(z)$ and $\chi(z)/\chi(z)$ from Section 2 and the appendix, we now relate $\alpha(z)$ and $\epsilon(z)$ as a function of $x \equiv \chi(z)/\chi(z) - 1$ (defined in Section 2) to $\alpha_0$, $\alpha_1$ and $\alpha_2$. We get

\[ \alpha = \left[ \alpha_0 \left( 1 + \alpha_1 + (2\alpha_1 + \alpha_2) x + \frac{3}{2} \alpha_2 x^2 \right) \right]^{1/3} \]
\[ \epsilon = \left[ \alpha_0 \left( 1 + \alpha_1 x + \frac{1}{2} \alpha_2 x^2 \right) \right]^{2/3} \]

\footnote{Note that an analogous derivation could be carried out for a stretching in the perpendicular ($\alpha_\perp$) and parallel ($\alpha_\parallel$) directions as in e.g. Kazin, Sánchez & Blanton (2012).}
\[ a_{\alpha} \left[ 1 + \frac{1}{3} \bar{\alpha}_{\alpha} + \frac{1}{3} (4 \bar{\alpha}_{\alpha} + \bar{\alpha}_{\beta}) x + \frac{5}{6} \bar{\alpha}_{\alpha} x^2 \right] \] (32)

and

\[ 1 + \epsilon = \left[ a_{\alpha} \left( 1 + \alpha_{\alpha} + (2 \alpha_{\alpha} + \bar{\alpha}_{\beta}) x + \frac{3}{2} \bar{\alpha}_{\alpha} x^2 \right) \right]^{1/3} \] (33)

\[ \cdot \left[ a_{\alpha} \left( 1 + \alpha_{\alpha} x + \frac{1}{2} \bar{\alpha}_{\alpha} x^2 \right) \right]^{-1/3} \] (34)

\[ \approx 1 + \frac{1}{3} \alpha_{\alpha} + \frac{1}{3} (\alpha_{\alpha} + \bar{\alpha}_{\beta}) x + \frac{1}{3} \bar{\alpha}_{\alpha} x^2. \] (35)

where the approximations in the second line are obtained by only working to linear order in \( \alpha_{\alpha} \) and \( \bar{\alpha}_{\alpha} \). We will also require the following partial derivatives (all evaluated at the fiducial value of \( \alpha_{\alpha} = 1, \bar{\alpha}_{\alpha} = \bar{\alpha}_{\beta} = 0 \) and denoted by subscript 0):

\[ \frac{\partial \alpha}{\partial \alpha_{\alpha}} \bigg|_0 = 1, \quad \frac{\partial \epsilon}{\partial \alpha_{\alpha}} \bigg|_0 = 0 \] (36)

\[ \frac{\partial \alpha}{\partial \bar{\alpha}_{\alpha}} \bigg|_0 = \frac{1}{3} (1 + 4 x), \quad \frac{\partial \epsilon}{\partial \bar{\alpha}_{\alpha}} \bigg|_0 = \frac{1}{3} (1 + x) \] (37)

\[ \frac{\partial \alpha}{\partial \bar{\alpha}_{\beta}} \bigg|_0 = \frac{1}{6} (2 x + 5 x^2), \quad \frac{\partial \epsilon}{\partial \bar{\alpha}_{\beta}} \bigg|_0 = \frac{1}{3} (x + x^2). \] (38)

### 4.2 Derivation

As was laid out in Section 3.1, we compress the correlation multipoles \( \xi_i(\ell, r) \) along the \( z \) direction. The optimal weights are proportional to \( C \left[ \xi_i(\ell, r, z), \xi_i(\ell, r, z) \right] \) \( \xi_{\ell,i} \), where \( i \) runs over the parameters \( \alpha_{\alpha}, \bar{\alpha}_{\alpha} \) and \( \bar{\alpha}_{\beta} \). Note that we do not consider mixing in the \( r \) or \( \ell \) directions, to avoid building in sensitivity to the model of the correlation function.

To compute \( \xi_{\ell,i} \), we need the dependence of the correlation function on \( \alpha \) and \( \epsilon \) (Padmanabhan & White 2008; Xu et al. 2013):

\[ \xi_0(r) \rightarrow \xi_0(r) + (\alpha - 1) \frac{\partial \xi_0(r)}{\partial \log r} \] (39)

\[ \xi_{\ell}(r) \rightarrow 2 \epsilon \frac{\partial \xi_{\ell}(r \log r)}{\partial \log r} \] (40)

Note that in the above, we set the intrinsic quadrupole correlation function \( \bar{\xi}_2 \) to zero. We make this approximation motivated by the fact that we are interested in distance measurements from the BAO feature, which is very weak in the quadrupole. The derivatives of the correlation function \( \xi_{\ell,i} \equiv \partial \xi_{\ell,i} / \partial \alpha_i \) are then given by (using equation 41)

\[ \xi_{\ell,\alpha} = \frac{\partial \xi_{\ell,0}}{\partial \alpha} \frac{\partial \alpha}{\partial \alpha_i} \propto \alpha_{i,\ell} \] (42)

\[ \xi_{\ell,\bar{\alpha}_{\alpha}} = \frac{\partial \xi_{\ell,0}}{\partial \bar{\alpha}_{\alpha}} \frac{\partial \alpha}{\partial \alpha_i} + \frac{\partial \xi_{\ell,\bar{\alpha}_{\beta}}}{\partial \alpha} \frac{\partial \epsilon}{\partial \alpha_i} \propto \epsilon_{i,\ell}. \] (43)

where we have dropped all non-redshift-dependent proportionality constants. We also note that the derivatives are all evaluated at the fiducial values of \( \alpha = 1 \) and \( \epsilon = 0 \).

Turning to the inverse covariance matrix, we note that the variance of a correlation function bin from a slice at redshift \( z \) is schematically given by

\[ C \sim \left( P + \frac{1}{\bar{\eta}} \right)^2 \frac{1}{dV}. \] (44)

where

\[ dV = \frac{\chi^2(z)}{\bar{H}(z)} \frac{dz d\Omega}{n} \] (45)

is the volume of slice being considered, \( \bar{\eta} \) is the (redshift dependent) number density and \( P \) is a measure of the power on scales corresponding to the bin in question. Again, since we are focusing on the BAO feature, we will ignore the scale dependence in \( P \) and simply treat it as a single number determined by the approximate scale corresponding to the BAO feature (\( k \sim 0.1 \ h \ Mpc^{-1} \)). We make the further simplifying assumption that the different redshift slices are independent, making the covariance matrix diagonal and trivial to invert:

\[ dW(z) \equiv C^{-1} = \left( \frac{\bar{\eta}}{n P + 1} \right)^2 dV. \] (46)

This assumption simply means the weights are not strictly optimal, but will not bias the results. We observe this weighting is nothing but the commonly used FKP weights (Feldman, Kaiser & Peacock 1994) in galaxy surveys, a correspondence we make more explicit in Section 4.5. While the normalization of the weights is arbitrary, it is convenient to normalize the weights by the integral of the inverse covariance matrix over the survey

\[ N = \int dW. \] (47)

We can now plug together the different pieces to get our final results. We define the weighted correlation functions by \( \xi_{w,\ell,i} \):

\[ \xi_{w,\ell,i} \equiv \frac{1}{N} \int dW(z) w_{\ell,i}(z) \xi_{i}(z), \] (48)

where, as before, \( i \) runs over the distance–redshift parameters. The \( w_{\ell,i}(z) \) are given by

\[ w_{0,\eta_{\alpha}} = 1 \quad w_{2,\eta_{\alpha}} = 0 \] (49)

\[ w_{0,\bar{\alpha}_{\alpha}} = \frac{1}{3} (1 + 4 x) \quad w_{2,\bar{\alpha}_{\alpha}} = \frac{1}{3} (1 + x) \] (50)

\[ w_{0,\bar{\alpha}_{\beta}} = \frac{1}{6} (2 x + 5 x^2) \quad w_{2,\bar{\alpha}_{\beta}} = \frac{1}{3} (x + x^2). \] (51)

Equation (48) makes explicit that the overall redshift weights are the product of \( w_{\ell,i}(z) \) and \( dW(z) \). We show two examples in Fig. 5. The top panel plots the weights for a constant \( \eta \) survey. The bottom panel is the same plot assuming a Gaussian \( \bar{\eta} \) centred around \( z = 1 \). In a constant \( \eta \) survey, the weights for \( \alpha_{\alpha} \) are simply the inverse variance weights. The upward sloping weights curve suggest an upweight in the high-redshift region of the survey. This is consistent with the intuition that slices at higher redshifts contain more volume and more galaxies, and thus smaller error bars, than those of lower redshifts.

### 4.3 Interpretation and examples

In order to develop some intuition for these weighted correlation functions, we consider the correlation function measured perpendicular to the line of sight; the case parallel to the line of sight is analogous. This case is simpler than the multipoles presented above,
Overall redshift weights for $A\alpha$ $\equiv -1993 + \alpha = \alpha \bar{w}$ and $w$ and $\Omega_1\alpha(z + \bar{w})$ is flattening out (or $-57$ $\bar{w}$ fixed; a full treatment for dark $B\alpha$ are integrals depending on the particular weights and $\alpha_0$, $\alpha_1$ and $\alpha_2$,

$$B_{\alpha_i} = \int \alpha_i(z)w_{\alpha_i}(z)dz.$$ (58)

The derivative term in equation (57) is equivalent to a shift of the BAO feature. The different redshift weights result in different shifts for a given set of distance parameters. From these measured shifts, one can then infer the underlying parameters.

To make this example more concrete, we consider a redshift range defined by $-1 \leq z \leq 1$ and assume the inverse variance weights are simply constant. The integrals are then trivial and are given by

$$B_{\alpha_0} \propto 2\alpha_0 + \frac{\alpha_0\alpha_2}{3}.$$ (59)

$$B_{\alpha_1} \propto \frac{2}{3}\alpha_0\alpha_1.$$ (60)

$$B_{\alpha_2} \propto \frac{1}{3}\alpha_0 + \frac{1}{10}\alpha_0\alpha_2.$$ (61)

The normalizations of the weights is arbitrary. We have boosted the amplitudes of the weights for better visualization.

4.4 Alternate parametrizations

Although we recommend our parametrization of the distance–redshift relation, the idea of optimal redshift weights is more general. To illustrate this, we consider a standard $w$CDM cosmology, with the Hubble parameter (and therefore, the distance–redshift relation) parametrized by a constant equation of state $w$ for dark energy:

$$H^2(z) = H_0^2(1 + z)^3 + (1 - \Omega_m)(1 + z)^{3(1+w)}.$$ (62)

Since this is illustrative, we hold $H_0$ and $\Omega_m$ fixed; a full treatment would vary multiple parameters.

Proceeding as before, the weights will depend on derivatives of the monopole and quadrupole with $w$, which in turn, are proportional to derivatives of $\alpha(z)$ and $\epsilon(z)$ with $w$. Fig. 6 plots these weights for a survey with a constant number density of objects. The figure highlights the relative importance of different redshifts to measuring $w$. At low redshifts, the weights are suppressed both due to the small volume, as well as the fact that all distance–redshift relations tend to their Hubble law form $\chi = c/H_0 z$ independent of the equation of state. The impact of the volume factor with increasing redshift is also clearly evident, upweighting the higher redshift regions even when the raw sensitivity to $w$ is flattening out (or decreasing).

4.5 Implementation notes

We briefly discuss a possible implementation of this weighting scheme. We do not claim that this is a final implementation and we defer tests with mock catalogues to future work. Our goal here is to provide a proof-of-principle that such an implementation is both possible and straightforward.

We start with equation (48), and write the integral as a sum over infinitesimal redshift slices indexed by $z$. We further write the correlation function as a combination of data and random pairs, using the Landy & Szalay (1993) estimator:
Applying a similar argument to the $N\Phi$ term in the denominator by using equation (65),

$$N\Phi = \Phi \sum \Delta W_{z} = \sum w_{\text{FKP}}^{2}(z)RR_{z}.$$  

(69)

We see that this is nothing but the usual random–random pairs $RR$, with every random weighted by $(1 + \bar{n}P)^{-1}$.

The weighted correlation function estimator can now be written succinctly as

$$\xi_{\ell,i} = \frac{\bar{DD} - 2\bar{DR} + \bar{RR}}{\bar{RR}}.$$  

(70)

Our proposed algorithm is as follows:

(i) Weight every galaxy/random by $w_{\text{FKP}} = (1 + \bar{n}P)^{-1}$.

(ii) Compute $DD$, $DR$ and $RR$ as a weighted pair sum, including $w_{\ell,i}(z)$ as an additional pair weight. Since the change in redshift is small for the scales of interest, one could simply use e.g. the mean redshift of the pair in computing $w_{\ell,i}$.

(iii) Compute $RR$ as a weighted pair sum including $w_{\text{FKP}}$, but not the $w_{\ell,i}(z)$ weight.

(iv) Compute the correlation function estimator according to equation (70). Note that in the absence of the $w_{\ell,i}$, $DD \rightarrow DD$, $\bar{DR} \rightarrow DR$, $RR \rightarrow RR$. We recover the usual algorithm for computing correlation functions.

(v) A final practical simplification: since the weights are simply linear combinations of $1$, $x$ and $x^{2}$, it is more efficient to compute correlation functions weighted by them instead of the original weights. As an aside, the above discussion makes it clear how the usual ‘weighted correlation functions’. We find that a small number of linear combinations of $1$, $x$ and $x^{2}$, it is more efficient to compute correlation functions weighted by them instead of the original weights.

As an aside, the above discussion makes it clear how the usual ‘weighted correlation functions’. We find that a small number of linear combinations of $1$, $x$ and $x^{2}$, it is more efficient to compute correlation functions weighted by them instead of the original weights.

5 DISCUSSION

Large redshift surveys capable of measuring the BAO signal have proven to be an effective way of measuring the distance–redshift relation in cosmology. Future BAO surveys will survey very large volumes, covering wide ranges in redshift. The question arises how best to analyse and interpret the information from such a wide range of $z$. One approach is to split the sample into a number of smaller redshift shells, with the BAO signal in each shell providing a measurement of the distance to a fiducial redshift within the shell. The drawback of such a procedure is that the BAO signal in each shell is of lower signal-to-noise, and this makes the analysis more sensitive to the tails of the distribution. Since variations in the distance–redshift relation are very smooth in most cosmologies, we advocate a different procedure. Specifically we suggest compressing the information in the redshift direction into a small number of ‘weighted correlation functions’. We find that a small number of redshift weights preserves almost all of the information, with little dilution in signal-to-noise per measurement.

We demonstrate that the distance–redshift relation, relative to a fiducial model, can be approximated by a quadratic function to sub-per cent accuracy over a broad range of cosmologies and redshifts. The coefficients can be straightforwardly related to the distance–redshift relation and its first two derivatives at a reference redshift $z_{0}$. This reference redshift is now an explicit choice and resolves ambiguities in the interpretation of the ‘effective’ redshift of a BAO measurement. For this particular choice of parameters, we then construct weights to optimally constrain their values from the data.
We explicitly demonstrate the process with toy models, and outline a possible implementation of this approach. Our parametrization of the distance–redshift relation is not unique. For instance, one could choose a parametrization based on the dark energy equation of state. Different choices would result in different weights, although their construction would be identical to what we present here. These differences are however superficial. The key message of this paper remains the same – that the parameters of the expansion history may be constrained by a relatively small number of modes.

This paper has made the simplifying assumption that the true correlation function does not have any redshift dependence. This is clearly not true, especially for samples selected using different techniques and for wide redshift ranges. This problem is not unique to this approach though. We expect it can be solved in part constructing a model for this redshift evolution from e.g. small-scale clustering measurements, and in part by introducing additional nuisance parameters to absorb remaining discrepancies. The exact details will depend on the particular sample and we defer further investigations to future work. We do point out that an interesting open question is how well this redshift evolution needs to be known to not compromise the fidelity and precision of future BAO measurements.

This paper lays the framework for a new approach to fitting BAO measurements. We plan on continuing to develop this approach in future work by testing possible implementations on mock catalogues, and ultimately applying it to existing surveys. These ideas of optimal weighting are likely to have broader applications elsewhere in large-scale structure analyses.

ACKNOWLEDGEMENTS

We thank Chris Blake, Daniel Eisenstein and Will Percival for useful discussions. We also thank Gongbo Zhao for pointing out an error in an earlier version of Fig. 6. FZ, NP and MW acknowledge the Santa Fe Cosmology Workshop where part of this work was completed. This work was supported in part by the National Science Foundation under Grant no. PHYS-1066293 and the hospitality of the Aspen Center for Physics. NP is supported in part by DOE, NASA and a Sloan Research Fellowship.

REFERENCES

Anderson L. et al., 2014, MNRAS, 441, 24
Blake C. et al., 2012, MNRAS, 425, 405
Feldman H. A., Kaiser N., Peacock J. A., 1994, ApJ, 426, 23
Hamilton A. J. S., 1993, ApJ, 417, 19
Kazin E. A., Sánchez A. G., Blanton M. R., 2012, MNRAS, 419, 3223
Landy S. D., Szalay A. S., 1993, ApJ, 412, 64
Laureijs R. et al., 2011, preprint (arXiv)
Levi M. et al., 2013, preprint (arXiv)
Mortonson M. J., Weinberg D. H., White M., 2014, preprint (arXiv)
Padmanabhan N., White M., 2008, Phys. Rev. D, 77, 123540
Padmanabhan N., White M., Eisenstein D. J., 2007, MNRAS, 376, 1702
Percival W. J., Cole S., Eisenstein D. J., Nichol R. C., Peacock J. A., Pope A. C., Szalay A. S., 2007, MNRAS, 381, 1053
Planck Collaboration et al., 2013, preprint (arXiv)
Spergel D. et al., 2013, preprint (arXiv)
Sugai H. et al., 2012, in McLean I. S., Ramsay S. K., Takami H., eds, Proc. SPIE Conf. Ser. Vol. 8446, Ground-based and Airborne Instrumentation for Astronomy IV. SPIE, Bellingham, p. 84460Y
Tegmark M., Taylor A. N., Heavens A. F., 1997, ApJ, 480, 22
Weinberg D. H., Mortonson M. J., Eisenstein D. J., Hirata C., Riess A. G., Rozo E., 2013, Phys. Rep., 530, 87
Xu X., Cuesta A. J., Padmanabhan N., Eisenstein D. J., McBride C. K., 2013, MNRAS, 431, 2834

APPENDIX A: DERIVATION OF THE DISTANCE–REDSHIFT PARAMETERS

Our distance–redshift relation in Section 2 can be written as

\[ \chi(z) = a_0 \left( 1 + a_1 x + \frac{1}{2} a_2 x^2 \right) \chi(z). \]  

(A1)

Taking the derivative of both sides with respect to \( z \) by using

\[ \frac{dx}{dz} = \frac{1}{H(z) \chi(z)} \]  

(A2)

we have

\[ \frac{1}{H(z)} = a_0 \left[ 1 + a_1 x + \frac{1}{2} a_2 x^2 \right] \frac{1}{H(z)} + (a_1 + a_2 x) \frac{\chi(z)}{H(z) \chi(z)} . \]  

(A3)

Evaluating at \( z = z_0 \) (\( x \) vanishes), we have

\[ \frac{1}{H_0} = a_0 \left( \frac{1}{H_{0,0}} + \frac{a_1}{H_{1,0}} \right) . \]  

(A4)

Solving for \( a_1 \) yields

\[ a_1 = \frac{H_{1,0} X_{1,0}}{H_{0,0} X_0} - 1 . \]  

(A5)

Another useful quantity is \( H_0(z)/H(z) \), which can also be derived from equation (A3). Together with the definition of \( x \),

\[ \frac{H_0(z)}{H(z)} = a_0 \left[ 1 + a_1 x + \frac{1}{2} a_2 x^2 + (a_1 + a_2 x)(1 + x) \right] \]  

(A6)

\[ = a_0 \left[ 1 + a_1 + (2a_1 + a_2) x + \frac{3}{2} a_2 x^2 \right] . \]  

(A7)

To calculate \( a_2 \), we take the second derivative of equation (A1) and evaluate at \( z = z_0 \).

\[ \chi''(z_0) = a_0 \left[ \chi''(z_0) + 2 a_1 x_0' + \chi_0 \left( a_1 x_0'' + a_2 x_0' \right) \right] , \]  

(A8)

where \( x' = dx/dz \) and \( x'' = d^2x/dz^2 \). The subscript 0 indicates the derivatives are all evaluated at \( z = z_0 \). Solving for \( a_2 \) by plugging in equation (A2), and using \( \chi(z) = 1/H(z) \), we obtain

\[ a_2 = (1 + a_1) X_{1,0} \left[ H_{1,0}' - a_0(1 + a_1) H_{0,0}' \right] - 2 a_1 . \]  

(A9)

This paper has been typeset from a TeX/XeLaTeX file prepared by the author.