An Empirical Analysis of Baryon Acoustic Oscillations in Galaxy and Quasar Clustering

BEHZAD ANSARINEJAD

An Empirical Analysis of Baryon Acoustic Oscillations in Galaxy and Quasar Clustering

Behzad Ansarinejad

Abstract: We present our analysis of the baryon acoustic oscillation (BAO) signal in the correlation functions of LOWZ and CMASS galaxy samples from Data Release 12 of the SDSS BOSS survey. We draw a comparison between our results and the findings of Cuesta et al. (2016) who analysed the BAO feature in the same datasets. Upon using subsets of the data to obtain an empirical estimate of the uncertainties on the correlation functions, we find our results to be in general agreement with the uncertainties presented by Cuesta et al. (2016), obtained from simulated mocks. We detect the BAO peak at $\approx 4\sigma$ and $\approx 7\sigma$ in the correlation function of the LOWZ and CMASS samples respectively. We demonstrate that when fitting the correlation functions with our fiducial ΛCDM model, the $A(s)$ nuisance fitting parameters play a significant role in providing a good fit. Based on a F-ratio test, we find that in our primary fitting range, the simple ΛCDM model without the nuisance fitting terms is rejected in favour of the full model as indicated by the small p-values of $p = 0.018$ and $p = 0.00084$, for the CMASS and LOWZ samples respectively. We demonstrate that the significance of the detection of the BAO peak is the quantity most sensitive to the choice of the fitting range. Based on isotropic fitting to the correlation functions, we obtain a distance of $D_V(z = 0.32)r_{d,fid}/r_d = 1226 \pm 32$ Mpc and $D_V(z = 0.57)r_{d,fid}/r_d = 1988 \pm 24$ Mpc, a 2.6 per cent and 1.2 percent measurement respectively, assuming a fiducial sound horizon $r_{d,fid} = 147.10$ Mpc. We extend our BAO analysis to higher redshifts by performing fitting to the mean of the correlation functions obtained by Chehade et al. (2016) for the 2QDESp, SDSS DR5 UNIFORM, 2QZ and 2SLAQ quasar samples. This gives a distance constraint of $D_V(z = 1.49)r_{d,fid}/r_d = 3583 \pm 249$ Mpc (assuming $r_{d,fid} = 147.70$ Mpc), a 6.9 per cent measurement to $z = 1.49$, the mean redshift of the combined QSO sample.

An Empirical Analysis of Baryon Acoustic Oscillations in Galaxy and Quasar Clustering

Behzad Ansarinejad

A Thesis presented for the degree of
Master of Science by Research

Centre for Extragalactic Astronomy
Department of Physics
Durham University
United Kingdom

December 2016

Contents

List of Figures

List of Tables

Declaration

The work in this thesis is based on research carried out in the Centre for Extragalactic Astronomy, Department of Physics, Durham University, United Kingdom. No part of this thesis has been submitted elsewhere for any other degree or qualification and it is all my own work unless referenced to the contrary in the text.

Acknowledgements

I would like to thank Prof. Tom Shanks for providing guidance and supervision throughout the project, as well as Ruari Mackenzie, Paddy Alton and Ben Chehade for their valuable advice and discussions. Additional thanks goes to Antonio J. Cuesta for kindly providing the measured correlation functions as well as the BOSS DR12 covariance matrices used in the analysis of Cuesta et al. (2016). I would also like to thank David Alonso for providing public access to the CUTE code used in obtaining the correlation functions in this study. Finally, I thank Kate Moloney for proofreading this work.

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 Astrofisica 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.

Chapter 1

Introduction

The determination of the expansion history of the universe is currently one of the primary goals of observational cosmology. The late-time transition of the expansion rate of the universe from a deceleration to a phase of acceleration (e.g. based on observational evidence from supernovae; Riess et al. 1998; Perlmutter et al. 1999) in particular, remains one of the most puzzling problems in modern physics. Investigating this problem and exploring the nature of Dark Energy (a hypothetical cause of the accelerated expansion rate of the universe (Peebles & Ratra, 2003), within the framework of ΛCDM, the current standard cosmological model), have driven efforts to obtain robust and high precision measurements of the cosmological expansion rate. To this end, a great interest was sparked in exploiting large galaxy redshift surveys in order to constrain the distance-redshift relation across a wide range of redshifts, making use of the Baryon Acoustic Oscillation (BAO) feature in the clustering of galaxies (Shanks 1985; Blake & Glazebrook 2003; Linder 2003; Seo & Eisenstein 2003; Matsubara 2004; Glazebrook & Blake 2005; Dolney et al. 2006; Sánchez et al. 2008).

The early universe consisted of a hot, dense plasma in which photons and baryons were tightly coupled due to Thomson scattering. Primordial density fluctuations lead to formation of overdense regions which gravitationally attracted matter while the increase in the matter density led to an increase in the outward photon radiation pressure. These contracting forces of gravity and radiation pressure in turn led to the formation of oscillations which could be thought of as relativistic sound waves consisting of photons and baryons, propagating away from the overdense region (Peebles & Yu 1970; Sunyaev & Zeldovich 1970; Doroshkevich et al. 1978). This propagation continued until the universe had cooled down due to expansion, reaching a temperature low enough to form neutral atoms at an epoch referred to as recombination. At this point photons decoupled from baryons, leaving behind a spherical shell of baryonic

matter at a scale related to the characteristic comoving radius known as the acoustic scale (also referred to as the sound horizon at the drag epoch) $r_d \approx 150$ Mpc (a more detailed description of the physics giving rise to the BAO feature can be found in Eisenstein & Hu 1998). The oscillations that lead to BAO can be seen as anisotropies in the Cosmic Microwave Background (CMB) radiation (Hu & White, 1996), and the characteristic comoving scale of the BAO feature imprinted in the distribution of galaxies can be used as a "standard ruler" in order to constrain the distance-redshift relation. As the BAO feature is imprinted on such large scales, BAO experiments are quite insensitive to systematics due to non-linear and astrophysical processes that occur on smaller scales (for a comparison between BAO and various other methods of measuring the distance-redshift relation see the review by Weinberg et al. 2013).

Although the BAO feature has been detected in the distribution of galaxy clusters (Veropalumbo et al., 2014), and high redshift BAO measurements have been obtained using the Lyman-α forest in quasar spectra (e.g. Slosar et al. 2013 and Delubac et al. 2015), galaxies have by far been the most commonly used tracers of the BAO feature. Notable studies among early efforts in constraining cosmological distances using the BAO feature in various galaxy redshift surveys include the analysis of the final 2dF Galaxy Redshift Survey (2dFGRS; Colless et al. 2003) by Cole et al. (2005), as well as various Luminous Red Galaxy (LRG) samples in early Sloan Digital Sky Survey (SDSS; York et al. 2000) data releases by Eisenstein et al. (2005), Hütsi (2006), Padmanabhan et al. (2007), Gaztañaga et al. (2009) and Kazin et al. (2010).

Constraints on the distance-redshift relations were obtained with increasing precision in later studies such as Percival et al. (2010) in which a 2.7 per cent distance measurement at redshift $z = 0.275$ was obtained by combining samples from SDSS and 2dFGRS surveys. This was followed by a 7.2, 4.5 and 5.0 per cent measurements to $z = 0.44, 0.6$ and 0.73 obtained by Blake et al. (2011) based on the WiggleZ survey (Drinkwater et al., 2010), and a 4.5 per cent measurement at $z = 0.1$ by Beutler et al. (2011) using the 6dF Galaxy Redshift Survey (6dFGRS; Jones et al. 2009).

The precision of the distance scale measurements have been further improved in more recent studies as a technique known as reconstruction began to be implemented. Due to non-linear structure growth, in practice the BAO peak appears smoother in the data than predictions from linear theory alone (see e.g. Jeong & Komatsu 2006; Guzik et al. 2007; Eisenstein et al. 2007b; Crocce & Scoccimarro 2008; Angulo et al. 2008; Matsubara 2008; Padmanabhan & White 2009). Additionally large-scale redshift-space distortions due to the peculiar velocities

of galaxies (Kaiser Effect; Kaiser 1987) also distort the clustering pattern in redshift-space. The effect of these phenomena can however be reversed using reconstruction techniques such as those introduced by Eisenstein et al. (2007a) which sharpen the BAO peak and could improve the precision of the measurement of the BAO scale by a factor of 2 or higher, with larger improvements expected at lower redshifts where the effect of non-linearities are worse and the dark energy density is highest. In more recent works, reconstruction has been used by Xu et al. (2012) to obtain a 2 per cent distance measurement to $z = 0.35$ using the SDSS DR7 LRG catalogue, followed by a 1.7 per cent measurement at $z = 0.57$ by Anderson et al. (2012) in the first analysis of the Data Release 9 (DR9) of the Baryon Oscillation Spectroscopic Survey (BOSS; Dawson et al. 2013) from the Sloan Digital Sky Survey III (SDSS III; Eisenstein et al. 2011). In their analysis of the BOSS DR11 sample, Anderson et al. (2014) obtained a 2 and 1 per cent measurement to $z = 0.32$ and $z = 0.57$, with the latter being the most precise distance constraint ever obtained at the time. Cuesta et al. (2016) further improved the precision of the measurement at $z = 0.32$ to 1.7 per cent in their analysis of an extended version of the same samples from the final BOSS data release (DR12).

Improvements and extension of the BAO analysis to higher redshifts will be made possible in the future making use of quasars with the Extended BOSS survey (eBOSS; Dawson et al. 2016), as well as HET-DEX (Hill et al., 2008), making use of Lyman-α emitting galaxies. Furthermore, substantial improvement in the precision of cosmological distance measurements across a wide range of redshifts bins are expected from the upcoming generation of ground and space-based surveys including DESI (Levi et al., 2013), EUCLID (Laureijs et al., 2011), LSST (LSST Dark Energy Science Collaboration, 2012) and WFIRST (Spergel et al., 2015).

Chapter 2

BOSS DR12 BAO Analysis

2.1 Introduction

A measurement of the BAO signature in the monopole two point correlation function of the "Constant Stellar Mass" (CMASS) and the low-redshift (LOWZ) galaxy samples from the Data Release 12 (DR12; Alam et al. 2015) of the SDSS BOSS survey was presented by Cuesta et al. (2016). The CMASS and LOWZ samples are extensions to previous SDSS LRG samples. Ross et al. (2015) also performed a measurement of the BAO feature at a lower redshift using the magnitude limited SDSS DR7 MAIN galaxy sample (Abazajian et al., 2009).

In this chapter we present the results of our independent measurement of the BAO feature in the DR12 CMASS and LOWZ samples, followed by a comparison of the results with the findings of the two previous studies. Specifically, the goal of this study is to attempt to replicate the results of Cuesta et al. (2016) as an independent verification of the applied methodology, placing particular focus on the uncertainties on the correlation functions. Cuesta et al. (2016) obtained an estimate of the uncertainties based on the covariance matrix of 1000 BOSS DR12 simulated QPM mocks (White et al., 2014). In this study, we divide the data into subsamples upon which measurements of the correlation function are performed, giving an empirical estimate of the uncertainty on the mean correlation function. We also investigate the possibility of obtaining an empirical estimation of the covariance matrix from the data for the CMASS sample, comparing the results of fitting the correlation function using this empirical covariance matrix to those obtained using the DR12 covariance matrix. Furthermore, we investigate certain aspects of the fitting procedure commonly implemented in BAO analysis studies, including potential effects of the choice of the fitting range on the results, as well as

the extent of the role played by the nuisance fitting parameters. As obtaining a measurement of the position of the BAO peak to a high degree of precision is beyond the immediate scope of this work, we do not attempt to perform reconstruction and hence we draw comparison with the pre-reconstruction results from the other studies throughout.

The layout of this chapter is as follows: Section 2.2 contains a brief description of the samples along with the basic properties of the selected subsamples. In section 2.3 we present a description of the relevant methodology involved in measuring the correlation function, error analysis and the fitting procedure. This is followed by a presentation and discussion of our results and a comparison of our findings with those of Cuesta et al. (2016) and Ross et al. (2015), in Section 2.4.

2.2 Datasets

In this study we use a set of 777,202 galaxies in the redshift range $0.43 < z < 0.7$ from the BOSS DR12 CMASS sample, with an effective redshift of 0.57, and 361,762 galaxies in the redshift range $0.15 < z < 0.43$ from the DR12 LOWZ sample, with an effective redshift of 0.32. The CMASS and LOWZ samples have been limited to magnitudes of $17.5 < i_{cmod} < 19.9$ and $16 < r_{cmod} < 19.6$ respectively. Full details of the target selection criteria can be found in Reid et al. (2016) and the treatment of systematics and the relevant corrections is discussed in Ross et al. (2016). In accordance with Cuesta et al. (2016), the samples, mocks and random datasets were obtained from the DR12 database[1]. The DR7 MAIN sample (used by Ross et al. 2015) is limited to $14.5 < r_{petr} < 17.6$, covering a lower redshift range of $z < 0.2$. The redshift distributions $n(z)$, of the galaxies in the DR12 CMASS and LOWZ samples, along with that of the DR7 MAIN sample are displayed in Fig. 2.1.

In order to obtain an empirical estimate of the uncertainties on the correlation functions, the CMASS sample is subsetted into five fields (subsamples) of equal size covering an overall area of 8487.77 deg^2, about 90.5 per cent of the total effective sample area (9376.09 deg^2). The LOWZ sample is similarly divided into five equally sized fields covering 7294.87 deg^2, roughly 87.5 per cent of the total sample area (8337.47 deg^2). Within the initial scope of this work, dividing the samples into five fields was deemed sufficient in order to produce an estimate of the uncertainties to a reasonable degree of accuracy. However, as demonstrated in later sections, the precision of the empirical estimate of uncertainties can be further improved

[1] `https://data.sdss.org/sas/dr12/boss/lss/`

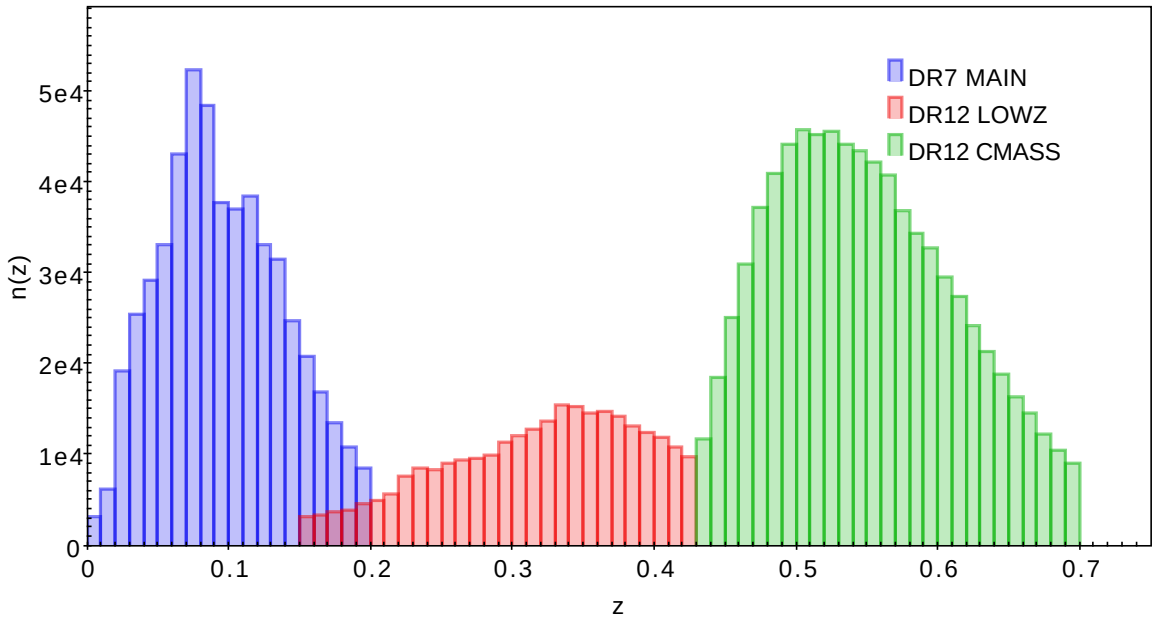


Figure 2.1: Histogram showing the number of galaxies as a function of redshift in the BOSS DR12 LOWZ (red) and CMASS (green) samples analysed in this study, as well as in Cuesta et al. (2016). The blue bars display the redshift distribution of the SDSS DR7 MAIN sample analysed by Ross et al. (2016). Bins are $\Delta z = 0.01$ in width.

by using a larger number of subsamples. The positions of all selected fields are illustrated in Fig. 2.2, with Table 2.1 providing a description of the basic properties of the selected fields. Once the correlation function for each field is obtained, a mean correlation function is calculated and is taken to represent the correlation function of the sample, using the standard error on the mean as an estimate of the uncertainty.

In this study we assume the same fiducial cosmology as Cuesta et al. (2016) with $\Omega_m = 0.29$, $\Omega_b h^2 = 0.02247$, $\Omega_\Lambda = 0.71$, $\Omega_k = 0$, $\Omega_\nu = 0$, $h = 0.7$, $w = -1$, $n_s = 0.97$ and $\sigma_8 = 0.8$. The fiducial distances to $z = 0.32$ and 0.57 (the effective redshifts of our samples), based on our assumed cosmology are presented in Table 2.2.

2.3 Methodology

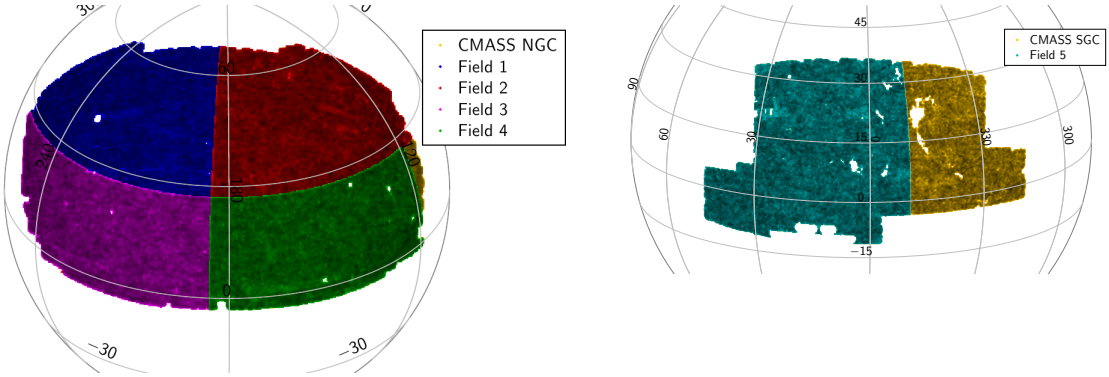
2.3.1 Measuring the Correlation Function

The monopole two-point correlation function (in redshift-space), $\xi(s)$, is calculated for each individual field using the CUTE[2] algorithm described by Alonso (2012).

[2] http://members.ift.uam-csic.es/dmonge/CUTE.html

(a) CMASS Northern Galactic Cap

(b) CMASS Southern Galactic Cap



(c) LOWZ Northern Galactic Cap

(d) LOWZ Southern Galactic Cap

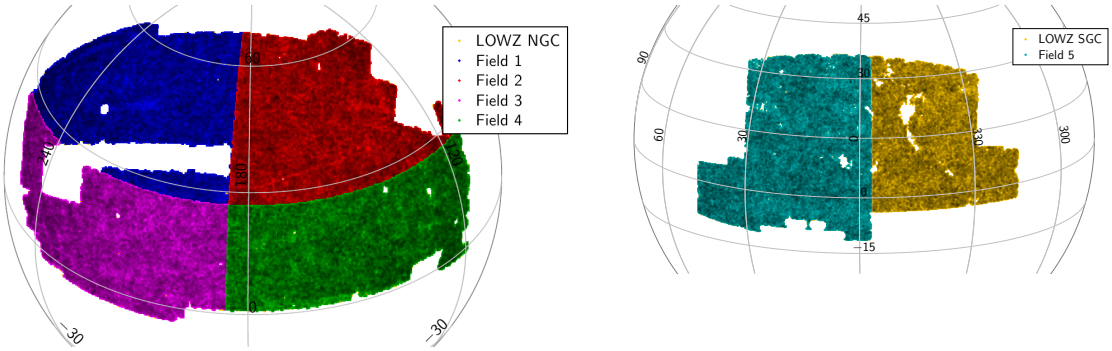


Figure 2.2: The coverage of the 5 selected fields in the Northern and Southern Galactic caps of the CMASS and LOWZ samples. The sample areas not selected are shown in yellow. (The basic properties of these fields can be found in Table 2.1).

Table 2.1: The basic properties of the 5 chosen fields (shown in Fig. 2.2) in the CMASS and LOWZ samples.

CMASS

Field	Ra°	Dec°	Area (deg^2)	Number of galaxies
1	>185	>27	1703	142,636
2	<185	>27	1686	141,706
3	>185	<27	1699	141,847
4	119-185	<27	1698	137,891
5	350-45.5	>-11	1701	144,820

LOWZ

Field	Ra°	Dec°	Area (deg^2)	Number of galaxies
1	>185	>27	1447	61,319
2	<185	>27	1453	63,109
3	>185	<27	1474	61,605
4	<185	<27	1463	63,431
5	357-45.5	>-11	1459	68,057

Table 2.2: A summary of the fiducial distances and values of the Hubble parameter used in this work and by Cuesta et al. (2016), computed at the effective redshifts of the LOWZ ($z = 0.32$) and CMASS ($z = 0.57$) samples, based on our assumed flat ΛCDM cosmological model.

r_d (Mpc)	$D_A(z = 0.32)$ (Mpc)	$H(z = 0.32)$ (km s^{-1} Mpc^{-1})	$D_V(z = 0.32)$ (Mpc)	$D_A(z = 0.57)$ (Mpc)	$H(z = 0.57)$ (km s^{-1} Mpc^{-1})	$D_V(z = 0.57)$ (Mpc)
147.10	962.43	82.142	1235.28	1351.13	94.753	2009.55

To perform the estimation of the correlation function we make use of the Landy-Szalay estimator (Landy & Szalay, 1993),

$$\xi(s) = \frac{DD(s) - 2DR(s) + RR(s)}{RR(s)}, \tag{2.3.1}$$

where $DD(s)$, $DR(s)$ and $RR(s)$ are data-data, data-random and random-random pair-counts respectively.

In our analysis we make use of the BOSS DR12 FKP-weighted (Feldman et al., 1994) randoms, and in accordance with Reid et al. (2016), apply a weighting of $w_{tot}w_{FKP}$ to the galaxies. A full description of the constituents of w_{tot} is presented in Reid et al. (2016); in short, this weight consists of three terms which account for effects of angular systematics, fibre collisions and redshift failures. In order to facilitate direct comparison with the findings of Cuesta et al. (2016), we sum our pair counts into 25 bins of width 8 h^{-1}Mpc in our calculation of the correlation functions, covering the range of $s \leqslant 200h^{-1}$Mpc in redshift space.

2.3.2 Error Analysis

Following the procedure proposed by Norberg et al. (2009), the bootstrap resampling method is used to provide an estimate of the errors on the mean correlation functions of our CMASS and LOWZ samples. In total we generate $N = 100$ resamplings and obtain the mean correlation function $\bar{\xi}(s)$ of these resamplings. As demonstrated by Norberg et al. (2009), an oversampling factor of 3 appears to be optimal in improving the bootstrap recipe. Hence we calculate the mean correlation function of each resampling, ξ_n, based on the correlation functions of $N_r = 3 \times N_{sub}$ randomly selected subvolumes (with replacement), from the original 5 subvolumes defined in Section 2.2 for the CMASS and LOWZ samples. The covariance matrix is then calculated using,

$$C_{boot}(\xi_i, \xi_j) = \frac{3}{N-1} \sum_{n=1}^{N} (\xi_n(s_i) - \bar{\xi}(s_i))(\xi_n(s_j) - \bar{\xi}(s_j)), \tag{2.3.2}$$

where subscripts i and j correspond to the ith and jth bins in redshift-space and the factor of 3 is added to account for our use of an oversampling factor of 3. The errors on the mean correlation functions of the CMASS and LOWZ samples are then given by the square root of the diagonal elements of their respective covariance matrices. (Note that as specified in Section 2.3.3, we do not make use of the covariance matrices calculated here when performing fitting to the correlation functions (equation 2.3.10), and instead use the BOSS DR12 covariance matrices obtained from QPM mocks in sections focused on investigating different aspects of the fitting procedure. This is with the exception of Section 2.4.9 in which we compare our bootstrap covariance matrix obtained from the data for the CMASS sample with the DR12 covariance matrix and compare the fitting results using the two matrices.)

A second set of errors are determined for the mean correlation functions of our samples, simply based on obtaining the standard errors on the mean. This is done using,

$$\sigma_{mean} = \frac{\sigma_{N_{sub}-1}}{\sqrt{N_{sub}}} = \sqrt{\frac{\sum(\xi_i - \bar{\xi})^2}{N_{sub}^2 - N_{sub}}}, \tag{2.3.3}$$

where $\sigma_{N_{sub}-1}$ is the standard deviation normalized to $N_{sub} - 1$ (as σ_{mean} is obtained from the same dataset reducing the number of degrees of freedom by one); N_{sub} is the number of subvolumes in each sample (i.e. 5); ξ_i is the correlation function of the ith subvolume, and $\bar{\xi}$ is the mean correlation function of the sample.

A comparison of the estimated errors from these two different methods and the errors found by Cuesta et al. (2016) based on the covariance matrix of the DR12 QPM mocks is presented in Section 2.4.2.

2.3.3 Fitting the Correlation Function

To fit the correlation functions we follow a procedure based on the methods described in Xu et al. (2012) and Anderson et al. (2012). We present a brief description of these techniques in this section.

We use a fitting model of the form

$$\xi^{fit}(s) = B^2 \xi_m(\alpha s) + A(s), \tag{2.3.4}$$

where ξ_m is defined in equation 2.3.8, B^2 is a constant term allowing for any unknown large-scale bias and $A(s)$ is given by

$$A(s) = \frac{a_1}{s^2} + \frac{a_2}{s} + a_3, \tag{2.3.5}$$

where $a_{1,2,3}$ are nuisance parameters. The $A(s)$ term is included in order to marginalise over broad-band effects due to redshift-space distortions and scale-dependent bias as well as any errors made in our assumption of the fiducial cosmology. The form of the $A(s)$ term was chosen by Xu et al. (2012) due to its simplicity and was further justified in that work by comparing it to various alternatives and demonstrating that it performs optimally in providing a good fit. We can obtain distance constraints by finding the optimum value of the scale dilation parameter α. This parameter provides a measure of any isotropic shifts in the position of the BAO peak in the data compared to the fiducial model, due to non-linear structure growth. An $\alpha > 1$ shifts the model towards larger scales while an $\alpha < 1$ shifts the model towards smaller scales. This term is defined as

$$\alpha = \frac{D_V(z)}{r_d} \frac{r_{d,fid}}{D_{V,fid}(z)}, \tag{2.3.6}$$

where z is the redshift, r_d is the sound horizon at the drag epoch and fid denotes the fiducial values (given in Table 2.2). The approximate volume-averaged distance to redshift z is

$$D_V(z) \equiv \left[cz(1+z)^2 \frac{D_A(z)^2}{H(z)} \right]^{1/3}, \tag{2.3.7}$$

where $D_A(z)$ is the angular diameter distance and $H(z)$ is the Hubble parameter at redshift z. This "distance" is proportional to the volume-averaged dilation factors (Ballinger et al., 1996) in the redshift and angular directions at a redshift z.

The model correlation function in equation (2.3.4), ξ_m, is given by

$$\xi_m(s) = \int \frac{k^2 dk}{2\pi^2} P_m(k) j_0(ks) e^{-k^2 a^2}, \tag{2.3.8}$$

where the Gaussian term is added to damp the oscillatory transform kernel $j_0(ks) = sin(ks)/ks$ at high-k. Here we set $a = 2h^{-1}$Mpc, which is small enough as to not cause significant damping effects at our scales of interest.

The template power spectrum is given by

$$P_m(k) = [P_{lin}(k) - P_{noBAO}(k)] e^{-k^2 \sum_{nl}^2 /2} + P_{noBAO}(k), \tag{2.3.9}$$

where P_{lin} is the linear power spectrum at $z = 0$ (generated using CAMB[3]; Lewis et al. 2000) and P_{noBAO} is the power spectrum with the BAO feature removed as described in Eisenstein & Hu (1998). The $\sum_{nl}^2 /2$ term damps the BAO features in P_{lin}, accounting for the effects of

[3] `http://cosmologist.info/camb/`

non-linear structure evolution. Here we set $\sum_{nl} = 8h^{-1}$Mpc.

The best fit values of the B^2, a_1, a_2 and a_3 fitting parameters in equation 2.3.4 are determined using the scipy.optimize.curve_fit module in Python which makes use of the Levenberg-Marquardt algorithm. To obtain the optimum value of α we compute the χ^2 goodness-of-fit indicator for fits obtained from shifting the model in the range $0.8 < \alpha < 1.2$ with intervals of $\Delta\alpha = 0.0001$, taking the value of α which corresponds to the minimum χ^2, (χ^2_{min}).

The χ^2 function is given by

$$\chi^2(\alpha) = [\xi^{obs} - \xi^{fit}(\alpha)]^T C^{-1} [\xi^{obs} - \xi^{fit}(\alpha)], \tag{2.3.10}$$

where ξ^{obs} is the observed correlation function, $\xi^{fit}(\alpha)$ is the best fit model at each α and C is the BOSS DR12 covariance matrix obtained from 1000 simulated QPM mocks.

In this study we investigate potential effects on the measured value of α and its uncertainty based on fitting the data across various ranges, using the complete ξ^{fit} model with and without the $A(s)$ nuisance parameters. Furthermore, by comparing the $\Delta\chi^2$ vs. α curves from fitting the ξ^{fit} and ξ^{noBAO} models (the latter is obtained by setting the term $P_m = P_{noBAO}$ in the model correlation function ξ_m), we obtain a measure of the significance at which the BAO signature is detected in the data. Here $\Delta\chi^2 = \chi^2(\alpha) - \chi^2_{min}$.

To obtain an estimate of the uncertainty in α, we assume a Gaussian form for the probability distribution of α,

$$p(\alpha_i) = \frac{e^{-\chi^2(\alpha_i)/2}}{\sum_j e^{-\chi^2(\alpha_j)/2}\Delta\alpha}, \tag{2.3.11}$$

where the denominator is a normalization factor ensuring the distribution integrates to unity. In effect $p(\alpha_i)$ is the probability that the acoustic scale $\alpha = \alpha_i$, based on the χ^2 distribution obtained from comparing the model ξ^{fit} (equation (2.3.4) with α_i), to our observed correlation function ξ^{obs}. We then calculate the standard deviation of our probability distribution which serves as an estimate of the uncertainty in α:

$$\sigma_\alpha = \sqrt{\langle\alpha^2\rangle - \langle\alpha\rangle^2}; \tag{2.3.12}$$

here $\langle\alpha\rangle$ represents the mean of the $p(\alpha_i)$ distribution given by:

$$\langle\alpha\rangle = \sum_i \alpha_i p(\alpha_i)\Delta\alpha, \tag{2.3.13}$$

and

$$\langle \alpha^2 \rangle = \sum_i \alpha_i^2 p(\alpha_i) \Delta\alpha. \tag{2.3.14}$$

The estimated uncertainty obtained from this method is equivalent to the value given by the $\Delta\chi^2$ curve at the 1σ level (see Fig. 2.8).

2.4 Results and Discussion

The correlation functions of the individual fields for the LOWZ and CMASS samples along with the corresponding mean correlation functions are displayed in Fig. A.1 in Appendix A. In the following sections we compare our measurement of the mean correlation functions with the measurements of Cuesta et al. (2016) and Ross et al. (2015), perform fitting to the mean correlation functions and analyse various aspects of the fitting procedure. Furthermore, we obtain measurements of $D_V(z)$ based on our measured position of the BAO peak and briefly investigate the possibility of obtaining an estimate of the covariance matrix from the data.

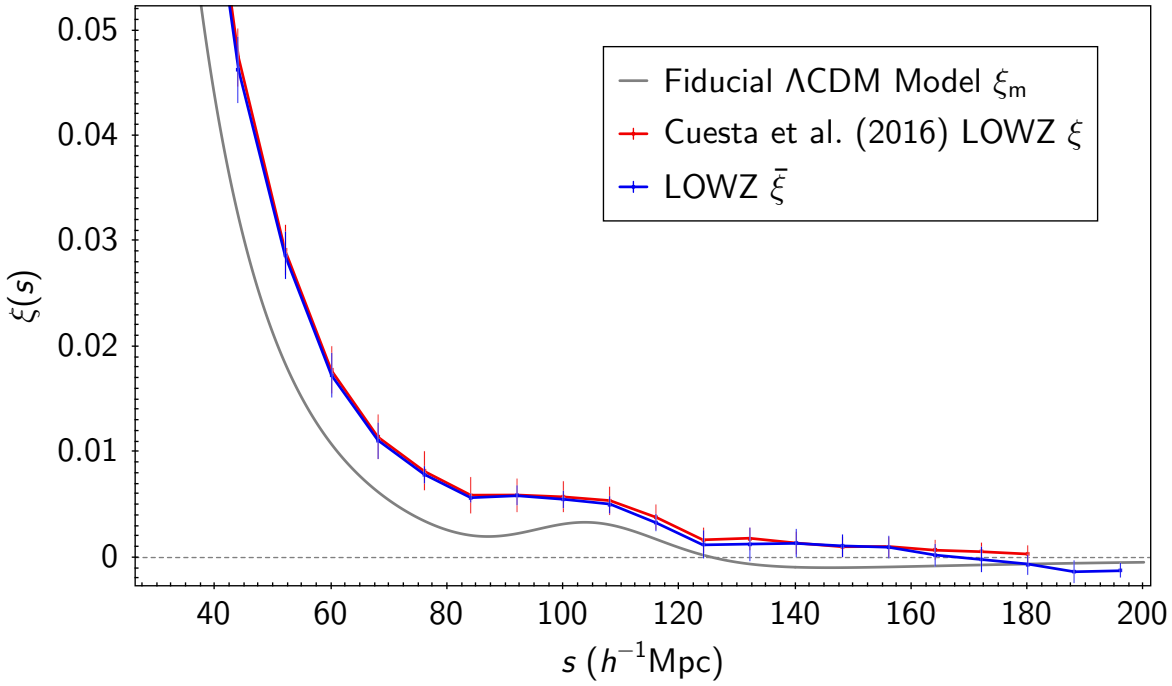
2.4.1 Comparison with Cuesta et al. (2016)

Fig. 2.3 shows a comparison between our mean correlation functions and the correlation functions obtained by Cuesta et al. (2016) for the DR12 LOWZ and CMASS samples. We find that our measured correlation functions are in excellent agreement with those presented in Cuesta et al. (2016) and we observe no significant changes when we replace the BOSS DR12 randoms with randoms generated by CUTE. Furthermore, we observe no significant variations when we do not apply any weights to the data or randoms. This outcome is however expected due to the high completeness of 98.8 per cent and 97.2 per cent for the CMASS and LOWZ samples respectively (see Fig. 8 of Reid et al., 2016).

2.4.2 Error Analysis Results

This section contains a comparison between our two measures of uncertainties (standard error and bootstrap resampling) on the mean correlation functions of the LOWZ and CMASS samples. Here we also include the bootstrap uncertainties obtained in Section 2.4.9 for the CMASS sample, based on dividing the data into 30 subsamples. We distinguish between the two bootstrap uncertainties using the labels 'CMASS 5' and 'CMASS 30'. More importantly comparisons are drawn between our measured empirical errors and errors obtained from simulations presented in Cuesta et al. (2016) for the correlation functions of the LOWZ and

(a)



(b)

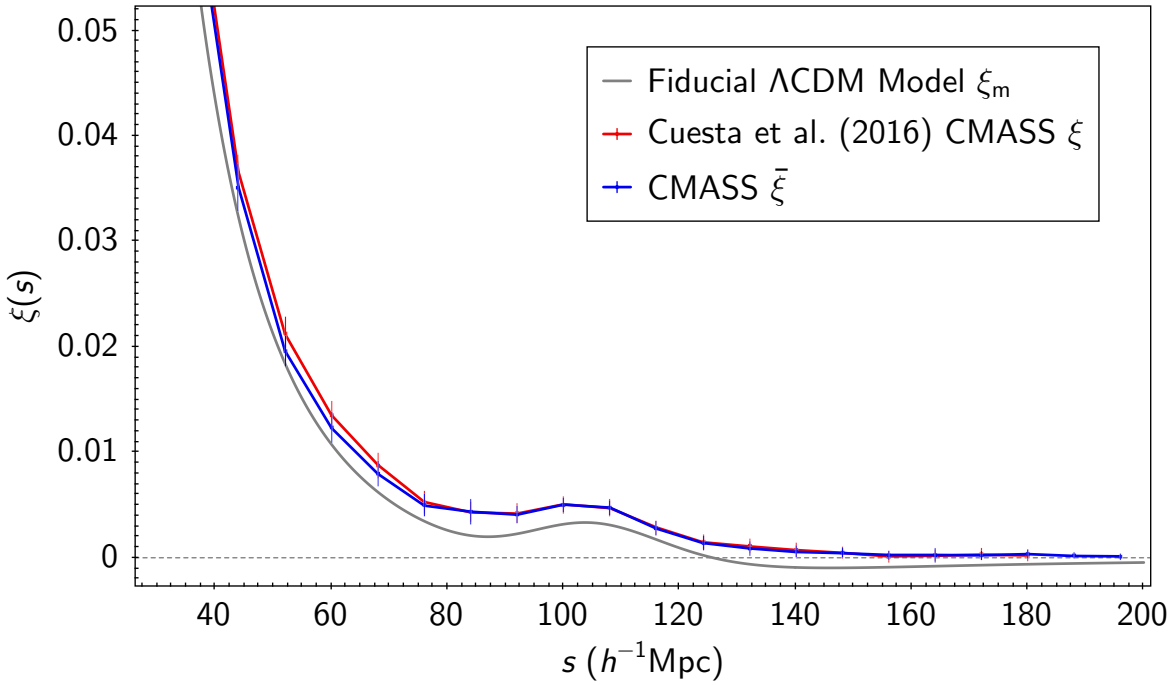


Figure 2.3: A comparison of our mean monopole correlation functions (blue curves) for (a) DR12 LOWZ and (b) CMASS samples, and the pre-reconstruction correlation functions presented in Fig. 1 of Cuesta et al. (2016) (red curves) for these samples. Error bars represent the standard error on the mean. ξ_m, the ΛCDM model based on our fiducial cosmology (equation 2.3.8; solid grey curve) is added for comparison.

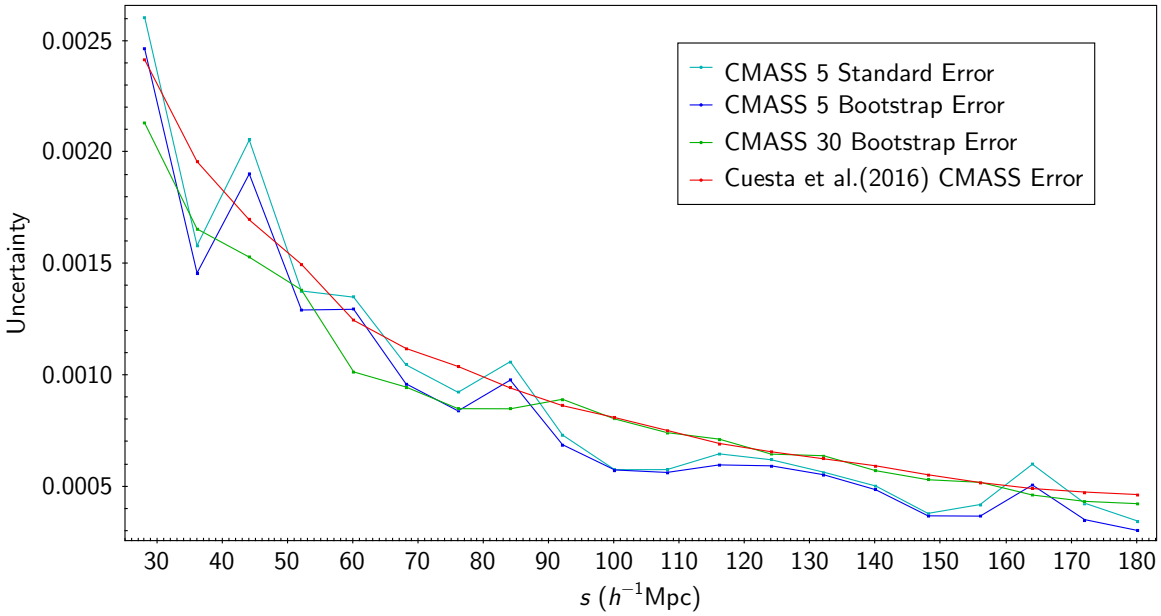
CMASS samples. In order to account for the fact that our selected fields do not cover the entire sample area, when comparing our results with those from Cuesta et al. (2016) we scale our measured errors by the square root of the ratio of the total coverage area of our fields to the total sample area.

As shown in Figs. 2.4a and 2.5a we find a good agreement between the standard error and bootstrap estimates of error for both samples. Furthermore, at our main scale of interest (in the vicinity of the 108 h^{-1}Mpc bin where the BAO peak lies), the results also appear to be in reasonable agreement with the errors presented by Cuesta et al. (2016) for the 5 fields CMASS sample, while the 30 fields bootstrap uncertainties appear to be in excellent agreement with those from Cuesta et al. (2016), at scales larger than 90 h^{-1}Mpc. To provide a quantitative demonstration of the level agreement between the errors from Cuesta et al. (2016) and the simple case of standard errors obtained from 5 fields, we make use of the fractional error in the error, given by $1/\sqrt{2N-2}$ (Squires, 2001). Here N is the number of measurements (in our case 5), giving a fractional error in the error of ≈ 35 per cent. Fig. 2.4b shows the ratio of our measured standard error to the errors presented by Cuesta et al. (2016) for the CMASS sample with the error bars being the error on our measured standard error. We can see that at the 108 h^{-1}Mpc bin this ratio is 0.8 which is consistent with unity within the error bars, and the general agreement between the errors is an indication that the QPM mocks reproduce an accurate representation of the data. As shown in Fig. 2.5b however, in the case of the LOWZ sample the ratio between the two errors varies to a greater extent with the scale, with the discrepancy between the two errors being larger around the BAO scale. This indicates that the errors presented by Cuesta et al. (2016) do not appear to be underestimated in this region. (Note that due to time constraints and the presence of large gaps in the LOWZ data, we did not attempt to divide the sample into 30 fields in a similar manner to the CMASS sample. However, obtaining errors from a larger number of fields is likely to account for some of the observed discrepancy between our errors and the findings of Cuesta et al. (2016), for the LOWZ sample.)

2.4.3 Comparison with Ross et al. (2015)

Fig. 2.6a displays a comparison between our mean correlation functions for the LOWZ and CMASS samples, with the correlation function presented by Ross et al. (2015) for the DR7 MAIN sample. The difference in amplitude is caused by the difference in the galaxy clustering bias in each sample, with LRGs being more biased. Upon normalizing the data to the ΛCDM

(a) CMASS Uncertainties



(b) CMASS Uncertainties Ratio

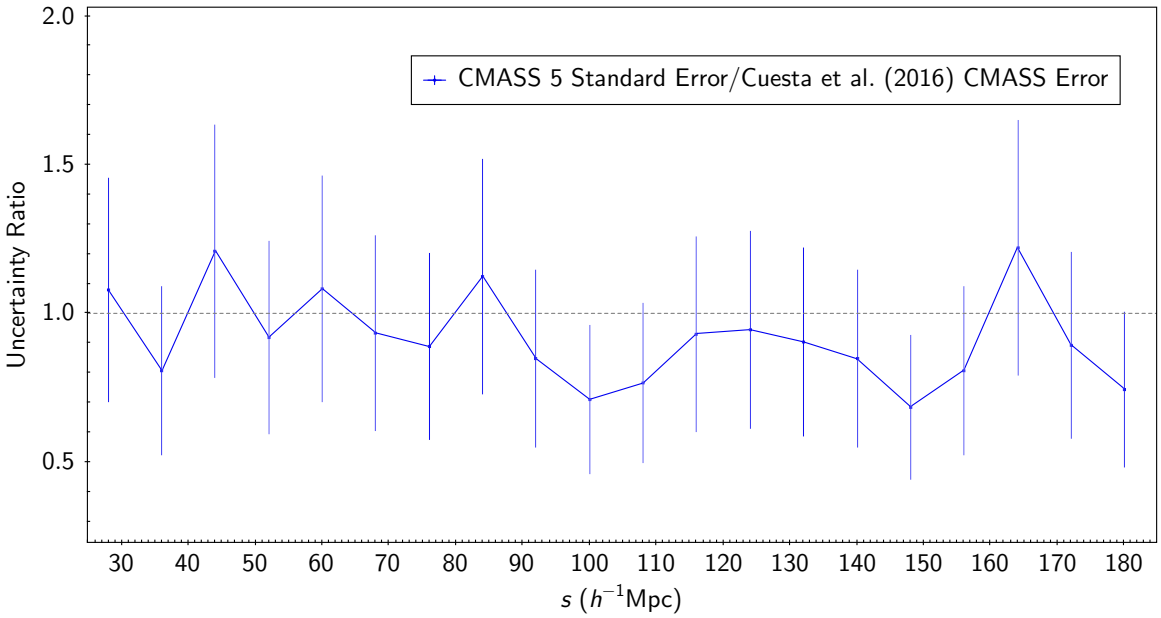
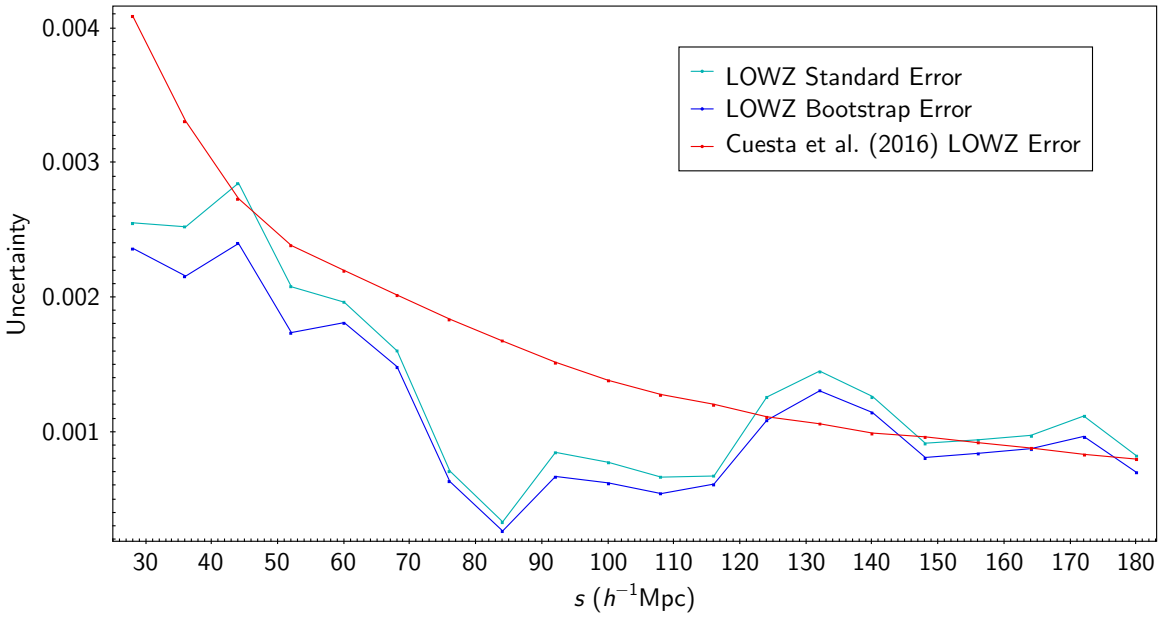


Figure 2.4: A comparison of the uncertainties on our measured mean correlation function of the CMASS sample, at our primary fitting range $28 \leqslant s \leqslant 180 h^{-1}$Mpc. The standard error on the mean (light blue line) and bootstrap (dark blue line) estimates of error for the 5 fields appear to be in good agreement. The bootstrap error from the 30 fields (green line) and the uncertainties on the measured correlation functions of Cuesta et al. (2016) (Fig.1) (red line) are also plotted, showing excellent agreement between the two at scales larger than 90 h^{-1}Mpc. Here all our measured errors are scaled by the square root of the ratio of the area covered by our selected fields, to the total sample area (e.g. in the case of 5 fields CMASS by $\sqrt{0.905}$). Subplot (b) shows the ratio of our standard error to the errors presented by Cuesta et al. (2016) for the CMASS sample. Here the error bars represent the error in the error (see the discussion in Section 2.4.2).

(a) LOWZ Uncertainties



(b) LOWZ Uncertainties Ratio

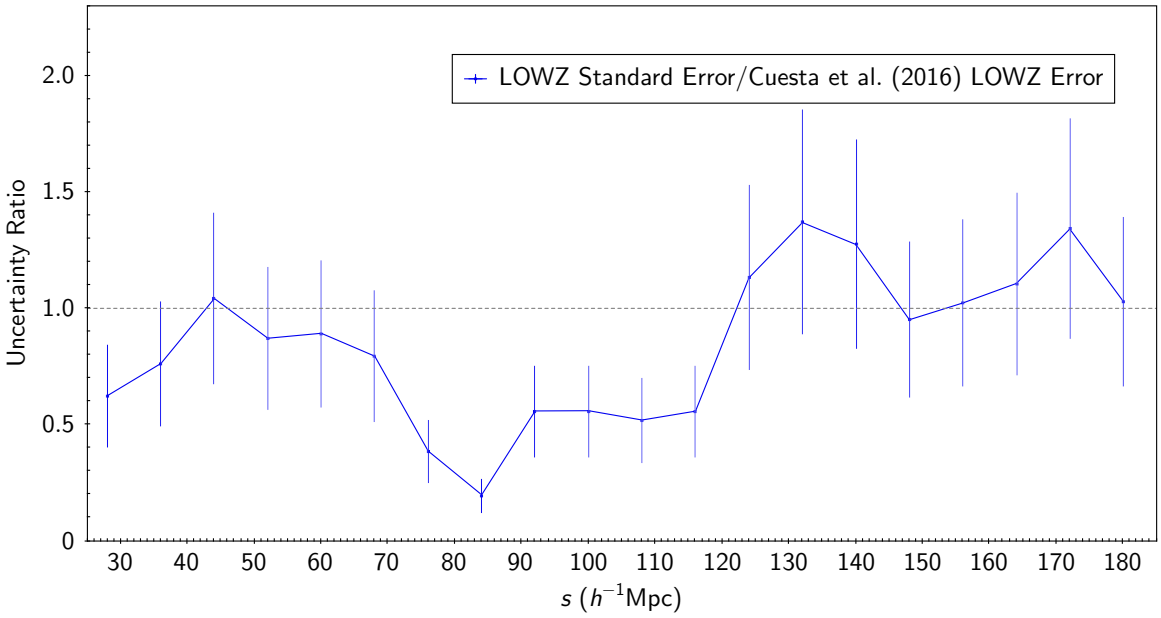


Figure 2.5: This plot is the equivalent of Fig. 2.4, showing a comparison of the uncertainties on the correlation function of the LOWZ sample.

model at 50 h^{-1}Mpc (Fig. 2.6b), we find a reasonable agreement between the correlation functions and the model.

2.4.4 Data Fitting Results

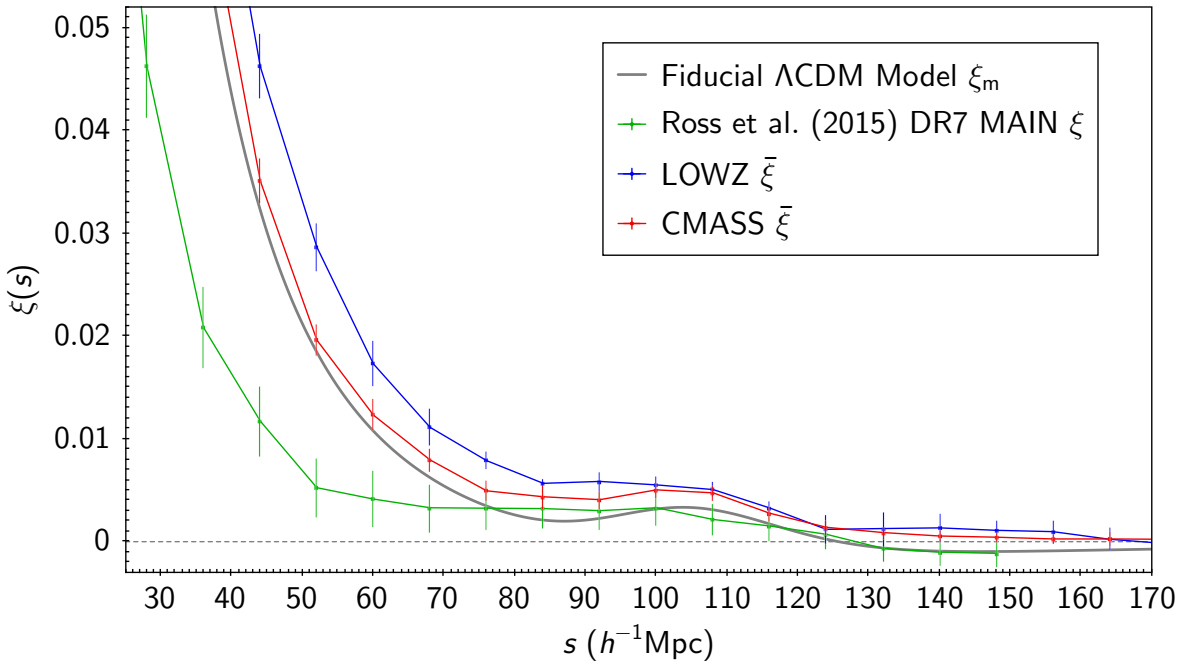
The best-fit values of α obtained from fitting the data with various models, across the range $28 \leqslant s \leqslant 180h^{-1}$Mpc are summarised in Table 2.3[4]. The pre-reconstruction best fit values of α from Cuesta et al. (2016) are included in this table for comparison. Here, 'α' refers to values obtained from fitting to the mean correlation functions of the LOWZ and CMASS samples, with errors given by the procedure described in Section 2.3.3. The '5-fields $\bar{\alpha}$' values in this table are obtained by fitting to the correlation functions of each field individually resulting in 5 measurements of α (these are presented in Table 2.4), and calculating the mean and standard error of these measurements. (Note that in line with the discussion in Section 2.4.2, as the '5-fields $\bar{\alpha}$' values are based on 5 measurements and hence contain a large error in the error (35 per cent), here one should round the error to 1 significant figure and round the mean to the appropriate number of decimal places accordingly). When fitting to correlation functions of individual fields we scale the BOSS DR12 covariance matrix by a factor of 5.

As shown in Table 2.3, we find that our measured '5-fields $\bar{\alpha}$' values are in good agreement with our values of α. This demonstrates the robustness of the implemented fitting procedure in producing an accurate measurement of the position of the BAO peak. Furthermore, when comparing the results corresponding to fits with the complete model, we find that for the LOWZ sample, our measured value of α is consistent with the measurement presented by Cuesta et al. (2016) within the uncertainties, while we find an $\approx 1.5\sigma$ tension between the two measurements for the CMASS sample; (note that for both samples the size of our measured errors are also comparable to those found by Cuesta et al. 2016). The reason behind the mild tension between our measurement of α and that of Cuesta et al. (2016) for the CMASS sample remains unknown.

We find the values of α measured for the individual fields in Table 2.4 to be in general agreement with the measurements of α from Cuesta et al. 2016. In cases where there appears to be tension between the measurements, (for instance the result of fitting the correlation function of field 4 in the CMASS sample with the complete model appears to be in an $\approx 1.7\sigma$

[4]Note that we place the main focus of our analysis on the results corresponding to this fitting range in order to match the fitting range chosen in Cuesta et al. (2016), allowing for direct comparison of the results. As discussed in Section 2.3.3, when fitting the correlation functions we use the BOSS DR12 covariance matrix used in the analysis of Cuesta et al. (2016).

(a)



(b)

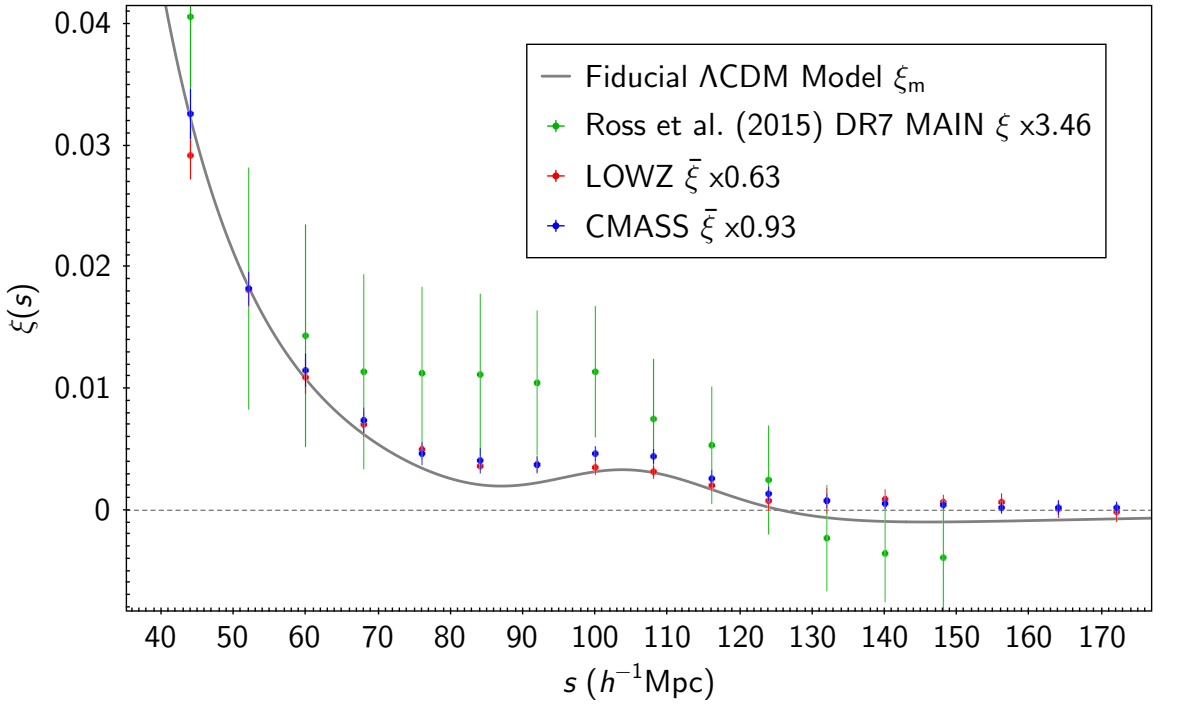


Figure 2.6: (a) Our measured mean correlation functions for the DR12 LOWZ (blue line) and CMASS (red line) samples, in comparison with the pre-reconstruction correlation function presented in Fig. 4 of Ross et al. (2015) (green line) for the DR7 Main sample. The fiducial ΛCDM model (equation 2.3.8; solid grey line) is added for comparison. (b) The MAIN, LOWZ and CMASS correlation functions normalised by factors of 3.46, 0.63 and 0.93 in order to match the ΛCDM model at 50 h^{-1}Mpc.

Table 2.3: Results of fitting the correlation functions of the LOWZ and CMASS samples using the complete ξ^{fit} model described in Eq. 2.3.4 and the same model without the $A(s)$ nuisance fitting parameters. In line with Cuesta et al. (2016) the fitting is performed in the range $28 \leqslant s \leqslant 180h^{-1}$Mpc. Here the '$\alpha$' values are obtained from fitting to the mean correlation function $\bar{\xi}(s)$ of each sample, presenting the corresponding χ^2_{min} over the number of degrees of freedom and 'Significance' refers to the significance of the detection of the BAO peak, using the complete fitting model (see Section 2.4.6). The '5-fields $\bar{\alpha}$' values are based on taking the mean and standard error of the individual αs, measured from fits to correlation functions of the 5 fields in the LOWZ and CMASS samples (see Table 2.4). We have used the BOSS DR12 covariance matrices in our fits scaling them by a factor of 5 when fitting to the 5 fields individually. For comparison, the best-fit values of α from Cuesta et al. (2016) (Table 10), for the pre-reconstruction LOWZ and CMASS sample are also included.

This Work	Model	α	χ^2_{min}/dof	Significance	5-fields $\bar{\alpha}$
CMASS	$B^2\xi_m + A(s)$	0.9892 ± 0.0118	14.9/15	7.2σ	0.9879 ± 0.0172
	$B^2\xi_m$	0.9991 ± 0.0116	28.5/18		0.9979 ± 0.0101
LOWZ	$B^2\xi_m + A(s)$	0.9927 ± 0.0260	15.5/15	4.0σ	0.9950 ± 0.0421
	$B^2\xi_m$	1.0311 ± 0.0629	45.5/18		0.9940 ± 0.0195
Cuesta et al. (2016)	Model	α	χ^2_{min}/dof		
CMASS	$B^2\xi_m + A(s)$	1.0153 ± 0.0134	12/15		
LOWZ	$B^2\xi_m + A(s)$	1.0085 ± 0.0300	13/15		

tension with the measurement from Cuesta et al. 2016), the dependency appears to be due to the shape of the BAO peak which in this case is extremely skewed to one side as seen in Fig. A.1 and as a result appears to be relatively flat. However, as the '5-fields $\bar{\alpha}$' values are in agreement with the measurements of α from the mean correlation functions, these effects seem to cancel out when we take the average over the 5 fields, even given our relatively small number of subsamples.

The performance of the two models in fitting the correlation functions (given by the χ^2_{min}/dof goodness of fit indicator) also appear to vary largely depending on the shape of the correlation function. However, with the exception of certain fields (for instance field 1 of the LOWZ sample), the complete model appears to perform better overall in providing good fits. It is important to note however, that the performance of a model in providing a good fit is not necessarily indicative that the correlation function has provided a representative and accurate measurement of α, and one should also consider the shape and prominence of the BAO peak in the correlation function itself[5]. This is exemplified by field 4 in the CMASS sample where the χ^2_{min}/dof value indicates that the complete model has provided a reasonably good fit to the data but due to the shape of the correlation function, an accurate determination of the position of the peak has not been possible. Finally we find that the significance of detection

[5]In Section 2.4.6 we discuss how the shape of the $\Delta\chi^2$ curve could also provide a measure of the degree to which we could be confident in our measurement of α.

Table 2.4: Results of fitting the correlation functions of the 5 individual fields in the LOWZ and CMASS samples using two different models, over the range $28 \leqslant s \leqslant 180h^{-1}$Mpc. Here we have used the BOSS DR12 covariance matrices, scaled by a factor of 5 and 'Significance' refers to the significance of the detection of the BAO peak, using the complete fitting model (see Section 2.4.6). The mean α and its standard error obtained based on the values of α in this table are presented under the '5-fields $\bar{\alpha}$ ' column in Table 2.3.

This Work	Field	Model	α	χ^2_{min}/dof	Significance
CMASS	1	$B^2\xi_m + A(s)$	0.9930 ± 0.0204	14.5/15	4.2 σ
		$B^2\xi_m$	0.9966 ± 0.0219	21.7/18	
	2	$B^2\xi_m + A(s)$	1.0356 ± 0.0266	16.2/15	3.3 σ
		$B^2\xi_m$	1.0255 ± 0.0297	21.1/18	
	3	$B^2\xi_m + A(s)$	1.0077 ± 0.0413	12.6/15	2.9 σ
		$B^2\xi_m$	1.0154 ± 0.0287	12.2/18	
	4	$B^2\xi_m + A(s)$	0.9343 ± 0.0461	13.5/15	2.0 σ
		$B^2\xi_m$	0.9734 ± 0.0393	29.7/18	
	5	$B^2\xi_m + A(s)$	0.9691 ± 0.0244	10.8/15	3.3 σ
		$B^2\xi_m$	0.9784 ± 0.0256	12.2/18	
LOWZ	1	$B^2\xi_m + A(s)$	0.9492 ± 0.0639	34.6/15	1.3 σ
		$B^2\xi_m$	0.9613 ± 0.0775	35.3/18	
	2	$B^2\xi_m + A(s)$	0.9051 ± 0.1132	23.7/15	1.6 σ
		$B^2\xi_m$	0.9413 ± 0.0425	33.8/18	
	3	$B^2\xi_m + A(s)$	0.9884 ± 0.0736	17.0/15	1.8 σ
		$B^2\xi_m$	1.0186 ± 0.0360	16.3/18	
	4	$B^2\xi_m + A(s)$	0.9794 ± 0.0345	22.3/15	2.3 σ
		$B^2\xi_m$	0.9991 ± 0.0370	26.0/18	
	5	$B^2\xi_m + A(s)$	1.1529 ± 0.0907	18.2/15	1.7 σ
		$B^2\xi_m$	1.0496 ± 0.0105	43.5/18	

of the peak in the individual fields to be generally lower then the significance of the detection of the peaks in the mean correlation functions of the two samples (as shown in Table 2.3). This is a further indication of the lack of prominent and well defined peaks in the correlation functions of the individual fields and as shown once again by field 4 in the CMASS sample, a low significance of detection of the peak could also hint towards the potential unreliability of the measured α.

2.4.5 Model Comparison

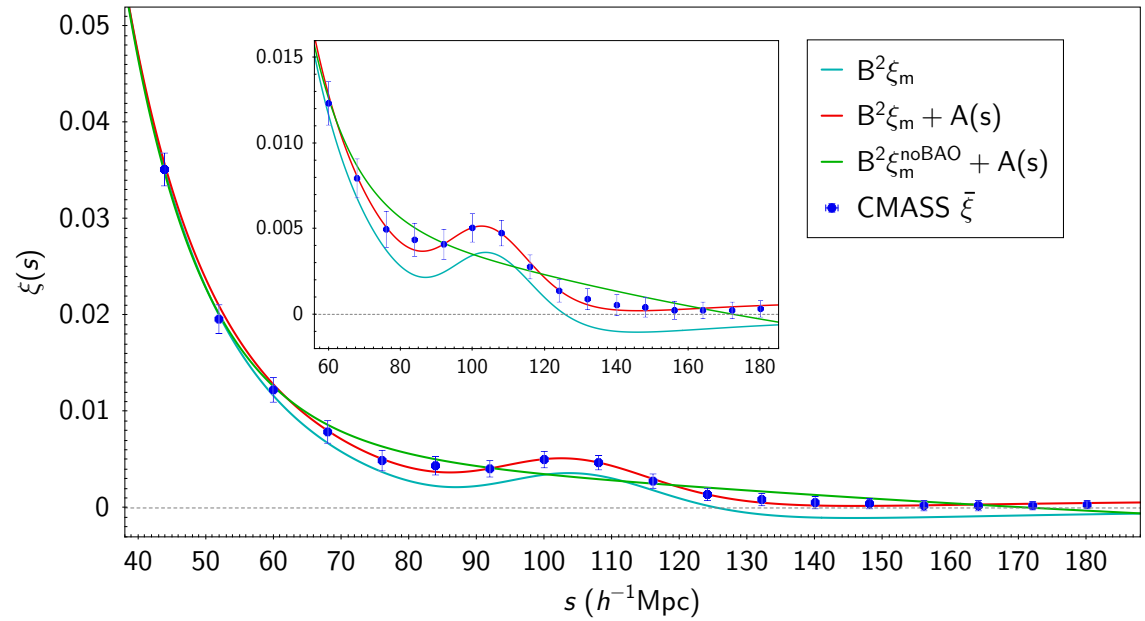
Fig. 2.7 shows the results of fitting the mean correlation functions of the CMASS and LOWZ samples with the ξ^{fit} model, fitted with and without the $A(s)$ nuisance parameters, and the ξ^{noBAO} model fitted with both B and $A(s)$ fitting terms. The important role played by the $A(s)$ nuisance fitting terms in producing a good fit is highlighted in these plots. This is also demonstrated numerically in Table 2.3, with the fits without the $A(s)$ having increased χ^2_{min}/dof values indicating the lower quality of fits. The χ^2_{min}/dof statistic gives an indication of the strength of evidence provided by the data against a null hypothesis, which in our case is that the data is consistent with the model in question (i.e. the probability of obtaining the experimental outcome, under the assumption that the model is correct). We assess the χ^2_{min}/dof statistic based on the corresponding p-value, which is defined as the probability of obtaining a χ^2 value at least as extreme as the value obtained, for the given number of degrees of freedom under the assumption that the model is consistent with the data. In the discussion that follows we refer to this quantity as the significance at which the model is rejected by the data.

We note that the visual impression given in Fig 2.7 is that the ΛCDM model without nuisance parameters is rejected at a higher significance than by the 28.5/18 ($p = 0.055$) indicated in Table 2.3, for the CMASS sample. Indeed, when only the diagonal terms of the covariance matrix are used, the significance of rejection rises to 64.9/18 ($p = 3.23 \times 10^{-7}$) (see Table 2.5), thus in this case the inclusion of the full covariance matrix causes a large reduction in χ^2_{min}/dof[6]. It is not immediately clear why this is the case.

We therefore take a more detailed look at how significant the nuisance parameters are in achieving a good fit for the ΛCDM model. Given our two nested fit models, we can make use

[6] As in some cases fitting with the full covariance matrix and the diagonal elements only could result in different best fit models, in order to ensure the fairness of the comparison in Table 2.5, in all cases when calculating the χ^2_{min}/dof values we use the best fit models obtained using the full covariance matrix.

(a)



(b)

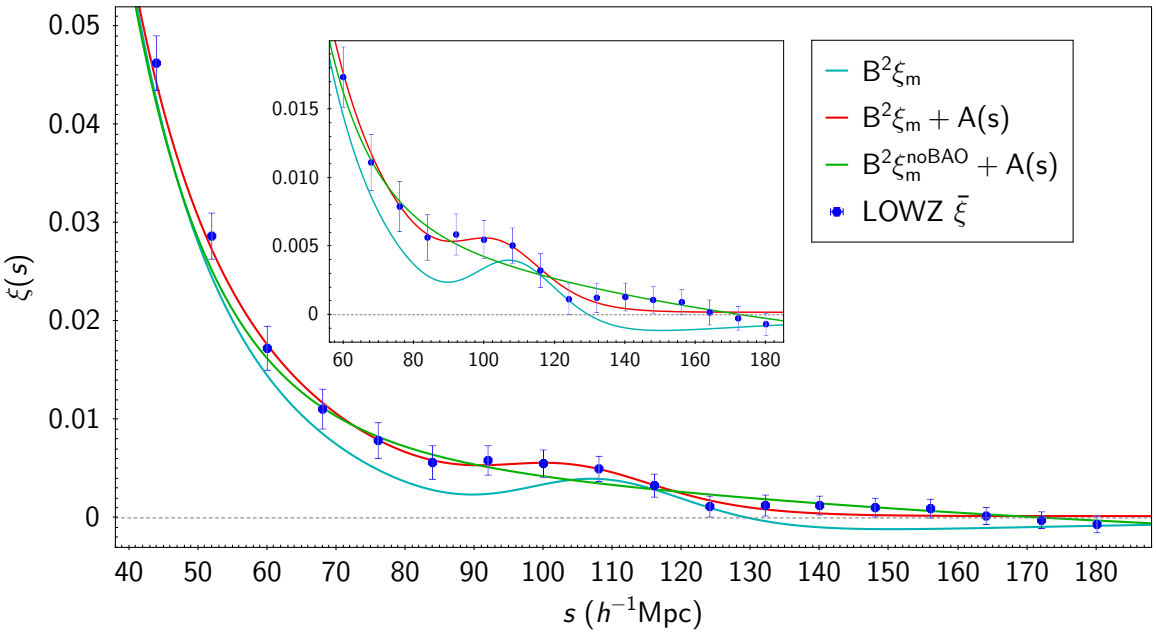


Figure 2.7: The results of fitting the mean correlation function of (a) CMASS and (b) LOWZ samples with various fitting models in the range $28 \leqslant s \leqslant 180h^{-1}$Mpc. The light blue curve is the ξ^{fit} model (equation 2.3.4) with the B fitting parameter only, while the red curve shows the same model fitted with both B and $A(s)$ fitting terms. The green curve is the ξ^{noBAO} model fitted with the B and $A(s)$ fitting terms. The error bars shown are the square root of the diagonal elements of the BOSS DR12 covariance matrices.

Table 2.5: The significance of rejection of various models, based on χ^2_{min}/dof and corresponding p-values obtained from fitting the mean correlation function of the CMASS sample in the range $28 \leqslant s \leqslant 180h^{-1}$Mpc, using the full covariance matrix and the diagonal elements of the matrix only.

		Full Matrix		Diagonal Elements
Model	χ^2_{min}/dof	p-value	χ^2_{min}/dof	p-value
$B^2\xi_m + A(s)$	14.9/15	4.59×10^{-1}	2.9/15	9.99×10^{-1}
$B^2\xi_m$	28.5/18	5.48×10^{-2}	64.9/18	3.23×10^{-7}
$B^2\xi_m^{noBAO} + A(s)$	55.5/15	1.47×10^{-6}	21.0/15	1.37×10^{-1}
$B^2\xi_m^{noBAO}$	63.4/18	5.71×10^{-7}	61.2/18	1.31×10^{-6}

of the F-ratio in order to determine whether the use of the more complex model results in a statistically significant improvement in fit quality. The F-ratio is given by

$$F = \frac{(\chi^2_{simple} - \chi^2_{complex})/(dof_{simple} - dof_{complex})}{\chi^2_{complex}/dof_{complex}}. \tag{2.4.1}$$

Here χ^2_{simple} and $\chi^2_{complex}$ refer to the χ^2_{min} values obtained from fitting the ξ^{fit} model without the $A(s)$ nuisance fitting terms, and by the complete ξ^{fit} model respectively, and dof are the degrees of freedom associated with each model. Once the F value is obtained we can test the validity of our null hypothesis that the complex model does not provide a significantly better fit than the simple model. Similar to the χ^2 analysis above, we asses the validity of the null hypothesis based on the p-value associated with the resulting F value.

Based on the χ^2_{min}/dof values presented in Table 2.3, for the fitting range $28 \leqslant s \leqslant 180h^{-1}$Mpc, we obtain F values of 4.56 ($p = 0.018$) and 9.68 ($p = 0.00084$) for the CMASS and LOWZ samples respectively. In other words our simple model is rejected in favour of the full ξ^{fit} model by the data, (given that assuming the null hypothesis is correct, i.e. that there is no significant difference between the two models, the probability of obtaining an F statistic at least as extreme as the values here by chance are ≈ 1.8 and 0.1 per cent for the CMASS and LOWZ samples respectively). This means that the inclusion of the nuisance parameters results in a significant improvement to the fit. This is specially true in the case of the LOWZ sample, where as seen in Fig. 2.7b, the BAO peak in the correlation function appears to have an extension in the $\approx 80 - 100h^{-1}$Mpc range, explaining the strong need for the nuisance parameters at the level of significance indicated by the F test.

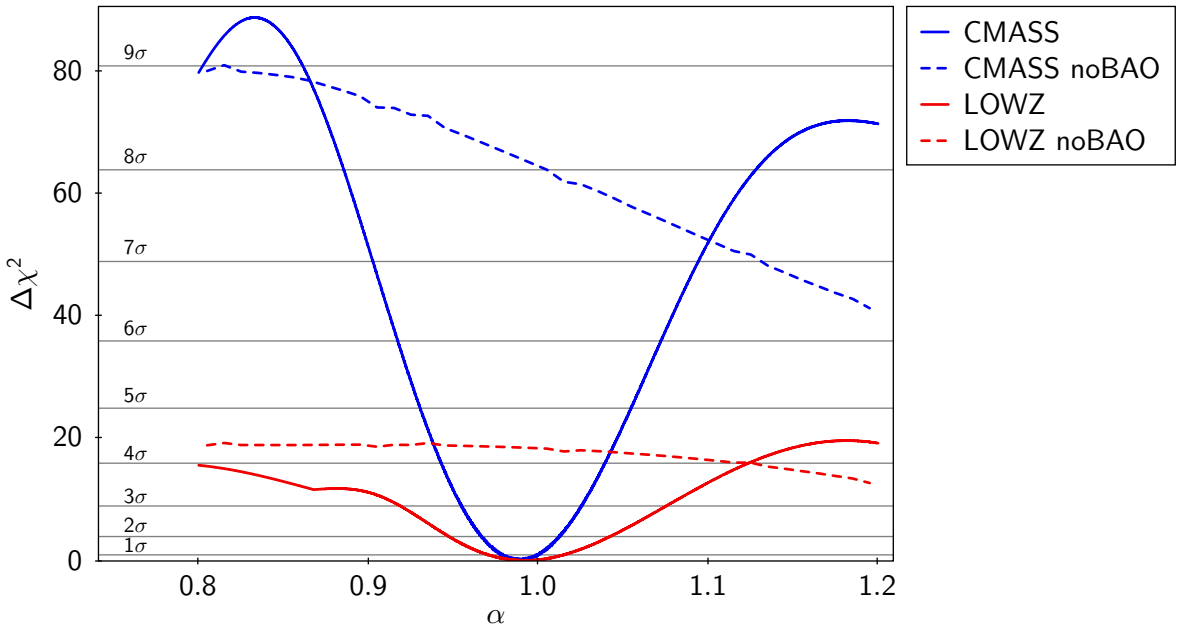


Figure 2.8: Significance of the detection of the BAO feature based on fitting the $\bar{\xi}$ for the LOWZ (red) and CMASS (blue) samples in the range $28 \leqslant s \leqslant 180h^{-1}$Mpc. The solid lines correspond to fits to the data based on the ξ^{fit} model which contains BAO, while the dashed lines correspond to fits based on the ξ^{noBAO} model with no BAO feature. In all cases the complete models including the $A(s)$ fitting terms are used. Here $\Delta\chi^2 = \chi^2(\alpha) - \chi^2_{min}$, where χ^2_{min} is the minimum χ^2 value using the model containing BAO. Comparing the dashed and solid lines provides a measure of our level of confidence that the BAO feature exists in the data. Here the BAO peak is detected at $\approx 4\sigma$ for the LOWZ sample and $\approx 7\sigma$ for the CMASS sample.

2.4.6 Significance of BAO Peak Detection

The $\Delta\chi^2$ curves based on fitting the mean correlation functions of the CMASS and LOWZ samples, with the ξ^{fit} and ξ^{noBAO} models are presented in Fig. 2.8. Here the complete fitting models including the $A(s)$ fitting terms are used and $\Delta\chi^2 = \chi^2(\alpha) - \chi^2_{min}$, where χ^2_{min} is the minimum χ^2 value using the model containing BAO. A comparison of the two models shows that we detect the BAO peak in the data at an $\approx 4\sigma$ level for the LOWZ sample and at greater than 7σ for the CMASS sample.

A second test of BAO significance is also captured in Fig. 2.8. For the CMASS sample, it can be seen from the plateau height of the $\Delta\chi^2$ curve that local maximum lies at a value of ≈ 72 above the minimum, meaning that we can be confident in our measured best-fit value of α at $\approx 8.5\sigma$. For the LOWZ sample, the maximum lies at ≈ 20, indicating that our best fit value of α is preferred at $\approx 4.5\sigma$ by the data. These values indicate that we have obtained well-constrained measurements of α in both cases.

2.4.7 The Choice of Fitting Range

In order to investigate the effects of the choice of fitting range on our measured value of α and the significance of the detection of the BAO peak, we perform our fitting across 7 different ranges using the ξ^{fit} model with and without the $A(s)$ nuisance fitting terms. We summarise the results in Table 2.6. It can be seen that the value of α and the magnitude of its error are largely insensitive to the choice of the fitting range for the CMASS sample. Slight variations in the value of α are observed as the fitting range is varied in the case of the LOWZ sample, however, these values remain consistent within the uncertainties. It can be seen that the quality of the fits produced by the ξ^{fit} model without the $A(s)$ nuisance fitting terms are consistently lower than the fits produced by the complete model across various ranges as shown by the χ^2_{min}/dof values. The quantity that appears to be most sensitive to the choice of the fitting range is the significance of the detection of the BAO peak in the data. At the two extremes, the significance of the detection of the peak varies from 7.2σ to 5.3σ for the CMASS sample and from 4.3σ to 3.0σ for the LOWZ sample, depending on the choice of the fitting range. This level of variation highlights the importance of providing appropriate justification for the choice of fitting range in studies performing analysis of the BAO feature.

2.4.8 Cosmological Distance Constraints

Using our measured values of α and 5-fields $\bar{\alpha}$ presented in Table 2.3 (for the complete ξ^{fit} model), and our fiducial distances presented in Table 2.2, we calculate the volume-averaged distance to redshift z, $D_V(z)$ for the LOWZ and CMASS samples. A comparison of our results and the findings of Cuesta et al. (2016) is given in Table 2.7. As expected given our measurements of α, we find our results to be in agreement with those from Cuesta et al. (2016) for the LOWZ sample, while we find an $\approx 1.4\sigma$ tension between the two results, for the CMASS sample. Furthermore, it can be seen that the magnitude of the errors are comparable between the two studies in the case of $D_V(z)$ which is based on the errors on α (giving a 2.6 and 1.2 percent distance measurement for the LOWZ and CMASS samples respectively), while the '5-fields $D_V(z)$' errors are larger due to the larger errors on the 5-fields $\bar{\alpha}$ values.

2.4.9 Covariance Matrix Comparison

In this section we draw comparisons between our bootstrap estimated covariance matrix obtained from the data for the CMASS sample, and the DR12 covariance matrix obtained

Table 2.6: Results of fitting the correlation functions of the LOWZ and CMASS samples using two different models and over various fitting ranges. In performing these fits the BOSS DR12 covariance matrices were used, and as before, 'Significance' refers to the significance of the detection of the BAO peak using the complete fitting model.

This Work	Range (h^{-1}Mpc)	Model	α	χ^2_{min}/dof	Significance
CMASS	$28 \leqslant s \leqslant 180$	$B^2\xi_m + A(s)$	0.9892 ± 0.0118	14.9/15	7.2 σ
		$B^2\xi_m$	0.9991 ± 0.0116	28.5/18	
	$36 \leqslant s \leqslant 172$	$B^2\xi_m + A(s)$	0.9868 ± 0.0114	11.9/13	6.5 σ
		$B^2\xi_m$	0.9917 ± 0.0114	19.0/16	
	$44 \leqslant s \leqslant 164$	$B^2\xi_m + A(s)$	0.9849 ± 0.0115	11.0/11	6.2 σ
		$B^2\xi_m$	0.9930 ± 0.0116	20.6/14	
	$52 \leqslant s \leqslant 156$	$B^2\xi_m + A(s)$	0.9859 ± 0.0122	6.6/9	6.5 σ
		$B^2\xi_m$	0.9930 ± 0.0116	23.2/12	
	$60 \leqslant s \leqslant 148$	$B^2\xi_m + A(s)$	0.9869 ± 0.0120	6.3/7	7.0 σ
		$B^2\xi_m$	0.9892 ± 0.0114	24.7/10	
	$68 \leqslant s \leqslant 140$	$B^2\xi_m + A(s)$	0.9854 ± 0.0115	6.2/5	7.2 σ
		$B^2\xi_m$	0.9860 ± 0.0112	28.3/8	
	$76 \leqslant s \leqslant 132$	$B^2\xi_m + A(s)$	0.9887 ± 0.0126	5.6/3	5.3 σ
		$B^2\xi_m$	0.9860 ± 0.0105	29.9/6	
LOWZ	$28 \leqslant s \leqslant 180$	$B^2\xi_m + A(s)$	0.9927 ± 0.0260	15.5/15	4.0 σ
		$B^2\xi_m$	1.0311 ± 0.0629	45.5/18	
	$36 \leqslant s \leqslant 172$	$B^2\xi_m + A(s)$	0.9880 ± 0.0253	13.9/13	4.3 σ
		$B^2\xi_m$	1.0284 ± 0.0179	49.0/16	
	$44 \leqslant s \leqslant 164$	$B^2\xi_m + A(s)$	0.9844 ± 0.0242	11.7/11	3.3 σ
		$B^2\xi_m$	1.0210 ± 0.0179	48.5/14	
	$52 \leqslant s \leqslant 156$	$B^2\xi_m + A(s)$	0.9774 ± 0.0243	7.5/9	3.0 σ
		$B^2\xi_m$	1.0043 ± 0.0189	39.0/12	
	$60 \leqslant s \leqslant 148$	$B^2\xi_m + A(s)$	0.9787 ± 0.0248	6.9/7	3.2 σ
		$B^2\xi_m$	1.0051 ± 0.0194	43.2/10	
	$68 \leqslant s \leqslant 140$	$B^2\xi_m + A(s)$	0.9787 ± 0.0248	6.7/5	3.2 σ
		$B^2\xi_m$	1.0002 ± 0.0189	42.3/8	
	$76 \leqslant s \leqslant 132$	$B^2\xi_m + A(s)$	0.9706 ± 0.0222	5.6/3	3.3 σ
		$B^2\xi_m$	1.0031 ± 0.0185	31.3/6	

Table 2.7: Distance constrains obtained from the analysis of the BAO feature in the correlation function of CMASS and LOWZ samples in this work and by Cuesta et al. (2016) (Table 11). Here $D_V(z)$ is calculated based on the value of α obtained from fitting to the mean correlation function of the samples, while the '5-fields $D_V(z)$' values are calculated based on $\bar{\alpha}$, which is obtained by taking the mean of the values of α attained from individually fitting to the 5 fields in the LOWZ and CMASS samples. In both cases the αs correspond to fitting to the range $28 \leqslant s \leqslant 180h^{-1}$Mpc using the complete ξ^{fit} model described in Eq. 2.3.4. We assume a fiducial sound horizon value of $r_{d,fid} = 147.10$ Mpc. The distance constrains are quoted at the effective redshifts of $z = 0.57$ and $z = 0.32$ for the CMASS and LOWZ samples respectively.

Study, Sample	$D_V(z)r_{d,fid}/r_d$ (Mpc)	5-fields $D_V(z)r_{d,fid}/r_d$ (Mpc)
This work, CMASS	1988 ± 24	1989 ± 40
Cuesta et al. (2016), CMASS Pre-Recon	2040 ± 28	—
This work, LOWZ	1226 ± 32	1235 ± 49
Cuesta et al. (2016), LOWZ Pre-Recon	1246 ± 37	—

from 1000 QPM mocks. As demonstrated by Hartlap et al. (2007) the resulting covariance matrix will be singular if $p > n - 1$, where p is the number of bins (i.e. 20 in our primary fitting range of the correlation function), and n is the number of independent observations (i.e. the number of subsamples). This was also observed by Pan & Szapudi (2005) who proposed the use of singular value decomposition (SVD) to calculate a pseudo-inverse of the covariance matrix which could then be used to obtain χ^2 (see also Gaztañaga & Scoccimarro 2005). Hartlap et al. (2007) also note that even when the resulting covariance matrix C is not singular, simple matrix inversion results in a biased estimation of the inverse which we denote by C_B^{-1}. The bias is linked to ratio of p and n, and if unaccounted for could have a significant effect on the fitting results and their estimated uncertainties, as the two numbers become comparable. As shown by Hartlap et al. (2007) for $p < n-2$ the bias can be corrected for to obtain the unbiased inverse C^{-1} by:

$$C^{-1} = \frac{n-p-2}{n-1} C_B^{-1}; \tag{2.4.2}$$

however, they advise against the use of SVD to invert the estimated covariance matrix beyond $p = n - 1$, as the bias of the pseudo-inverse is no longer controllable by the above correction. Furthermore, as pointed out by Norberg et al. (2009) even when using the bootstrap recipe where n is the number of resamplings and can be chosen to be much larger than p, the data must initially be divided into a large enough number of subsamples in order to satisfy the criteria discussed in Hartlap et al. (2007) (i.e. $p < n - 2$).

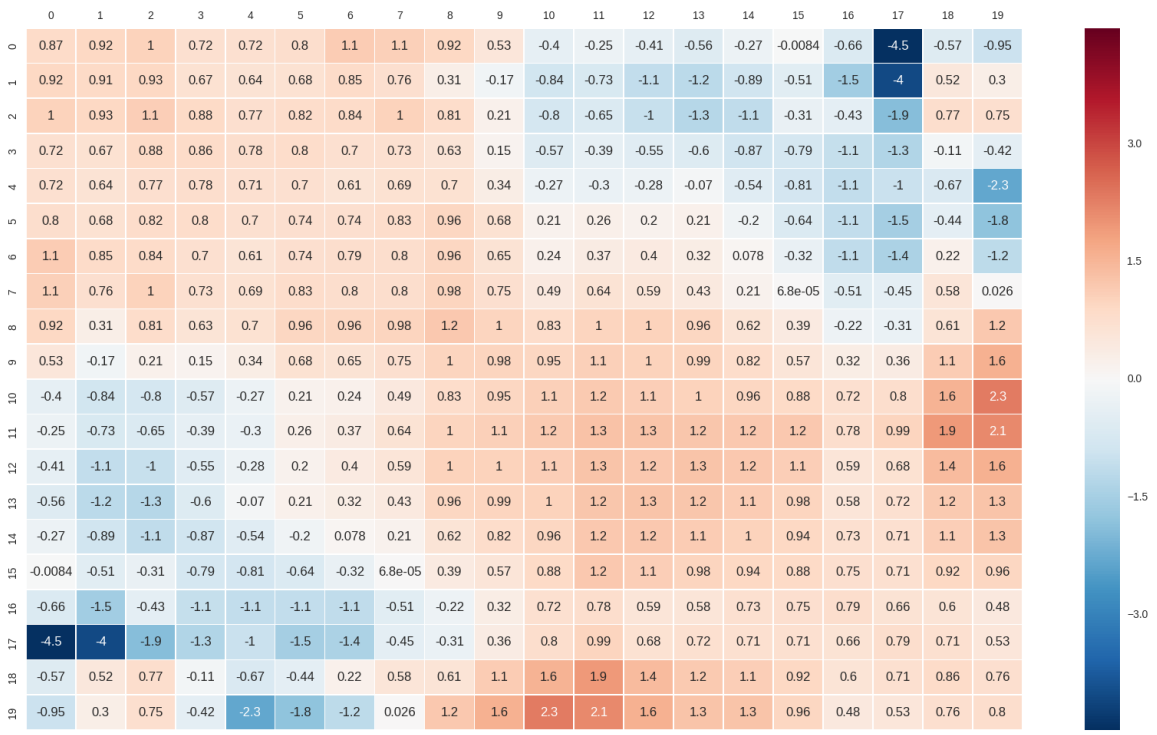


Figure 2.9: The ratio of the CMASS covariance matrix estimated using the bootstrap method based on $n = 500$ resamplings of our 30 subsamples, to the DR12 CMASS covariance matrix obtained from 1000 simulated QPM mocks. The matrix corresponds to the 20 bins in our primary fitting range $28 \leqslant s \leqslant 180h^{-1}$Mpc.

Considering the above requirements it is not possible to fit the correlation function using the inverse of the covariance matrix estimated from our original set up where the data is divided into 5 subsamples. As a result we divide the CMASS sample into 30 subsamples and estimate a covariance matrix using the bootstrap recipe described in Section 2.3.2, based on the new dataset. The position of the 30 selected fields are shown in Fig. A.2 in Appendix A. Each field contains about 23,500 galaxies and has an area of $\simeq 275\,\mathrm{deg}^2$, with the selected fields covering 88 per cent of the total sample area. The corresponding correlation functions for the 30 subsamples are shown in Fig. A.3. We find the mean correlation function to be in a good agreement with the mean correlation function from our 5 fields as well as the CMASS correlation function from Cuesta et al. (2016), once integral constraint is accounted for.

Fig. 2.9 shows the ratio of our bootstrap covariance matrix based on the 30 subsamples and 500 bootstrap resamplings, to the DR12 covariance matrix. The first 5 off-diagonals of the two matrices appear to be consistent to within 50 per cent while the diagonal elements are recovered within 30 per cent in our estimation of the covariance matrix.

It is important to note that the bootstrap estimate of the covariance matrix contains an intrinsic uncertainty that along with numerical round-off errors propagate in the calculation

Table 2.8: Results of fitting the correlation function of CMASS samples using the complete ξ^{fit} model in the range $28 \leqslant s \leqslant 180h^{-1}$Mpc with our bootstrap covariance matrix in comparison with the DR12 covariance matrix. Here n is the number of resamplings used in the bootstrap recipe and *rcond* determines the cutoff point for small singular values (see the discussion in the text).

Covariance Matrix	*rcond*	α	χ^2_{min}/dof
DR12		0.9892 ± 0.0118	14.9/15
Bootstrap ($n = 100$)	0.001	0.9819 ± 0.0111	13.0/15
Bootstrap ($n = 500$)	0.0015	0.9878 ± 0.0098	14.1/15

of the eigenvalues and eigenvectors. It is therefore reasonable to only consider the principal components corresponding to large eigenvalues which represent the most stable linear combinations of the data; however there is no unique and objective method to determine the number of principal components that need to be considered (see the discussions in Section 4.2 of Porciani & Giavalisco 2002 as well as Section 3.3 of Porciani & Norberg 2006 for further details). When inverting our covariance matrix we implement a method similar to Pan & Szapudi (2005) and Gaztañaga & Scoccimarro 2005 where SVD is used to create a pseudo inverse of the matrix, setting singular values corresponding to small eigenvalues to zero. To that end, we use the scipy.linalg.pinv function[7] in Python, which allows the user to define a cut-off for small singular values by considering singular values smaller than a constant ('*rcond*') multiplied by the largest singular value as zero[8]. We determine the value of this constant using trial and error until the size of the χ^2_{min}/dof reaches the correct scale and the fit value (α and its uncertainty) are in agreement with the results form fitting with the DR12 covariance matrix. We summarize the results of our fits with the bootstrap covariance matrix in Table 2.8, comparing the results to the values obtained from the DR12 covariance shown previously in Table 2.3. We show the results of fitting with two bootstrap covariance matrices calculated based on $n = 100$ and $n = 500$ resamplings, in each case applying the correction factor given by equation 2.4.2 to correct for the bias in the inverse covariance matrix.

It can be seen from Table 2.8 that once the correction for small singular values is applied our bootstrap covariance matrix produces fit results that are in good agreement with the results of fitting with the DR12 covariance matrix, with the corresponding χ^2_{min} values being at the right scale. The bootstrap covariance matrix calculated based on $n = 500$ resamplings performs slightly better than the one based on $n = 100$ resamplings, however we do not observe any major further improvements by increasing the number of resamplings. At a basic

[7] `http://docs.scipy.org/doc/scipy-0.16.0/reference/generated/scipy.linalg.pinv.html`
[8] In this comparison the modulus of the values are considered.

level, the results demonstrate that it is possible to empirically estimate the covariance matrix from the data using the bootstrap method, producing fit results comparable to those given by the covariance matrix from mocks. However, given that our method of selecting the cutoff point for small singular values relies on our previous knowledge of the fit results and there is currently no objective and universally applicable recipe for determining the appropriate number of principal components to include (for instance the techniques implemented by Pan & Szapudi 2005, Gaztañaga & Scoccimarro 2005 and Porciani & Norberg 2006 results in the inclusion of either too few or two many principal components when applied to our covariance matrix), the applicability of our method remains limited and further investigation is required to determine if variations of this technique could have more general applications.

Chapter 3

QSO BAO Analysis

3.1 Introduction

In this chapter we extend our BAO analysis to higher redshifts by performing isotropic fitting to the combined monopole correlation function as measured by Chehade et al. (2016), based on quasar samples from the 2dF QSO Redshift Survey (2QZ; Smith et al. 2005), SDSS Data Release 5 (SDSS DR5; Adelman-McCarthy et al. 2007), 2dF-SDSS LRG and QSO survey (2SLAQ; Richards et al. 2005) and the 2dF Quasar Dark Energy Survey pilot (2QDESp; Chehade et al. 2016).

Although a number of previous efforts have been made in measuring the BAO feature in the Lyman-α forest of various BOSS quasar samples (Busca et al. 2013; Kirkby et al. 2013; Slosar et al. 2013; Delubac et al. 2015), as well as using the quasar-Lyman α forest cross-correlation (Font-Ribera et al., 2014); to our knowledge, this work is among the very first attempts to obtain a direct high redshift distance constraint, using large-scale quasar clustering as the tracer of the BAO signal.

3.2 Datasets

Here we provide a brief summary of the relevant properties of the quasar samples used in our BAO analysis. A more detailed description of these samples can be found in the referenced papers.

The 2QZ sample (Croom et al., 2004) covers a total area of 721.6 deg^2, containing 22,655 QSOs (≈ 31 quasars deg^{-2}) up to $z \approx 3$ in the magnitude range $18.25 < b_j < 20.85$.

The SDSS DR5 "UNIFORM" sample was constructed by Ross et al. (2009) by taking a subsample of the DR5 quasar catalogue (Schneider et al., 2007). This sample covers an area of $\approx$ 4000 deg^2, containing 30,239 QSOs ($\approx$ 8 quasars deg^{-2}), in the redshift range $0.3 \leqslant z \leqslant 2.2$ with a magnitude limit of $i_{SDSS} \leqslant 19.1$.

The 2SLAQ sample (Croom et al., 2009) covers an area of $\approx$ 192 deg^2 containing $\approx 9,000$ QSOs ($\approx$ 47 quasars deg^{-2}) in the redshift range $z \lesssim 3$ and magnitude range $20.5 < g_{SDSS} < 21.85$.

The 2QDESp sample (Chehade et al., 2016) covers an area of $\approx$ 150 deg^2 in the southern sky, containing $\approx 10,000$ QSOs ($\approx$ 67 quasars deg^{-2}) with magnitudes $g \leqslant 22.5$. The quasars in the sample have a mean redshift of $z = 1.55$ and with 80 per cent of the objects in the sample lying in the range $0.8 < z < 2.5$.

In order to ensure a good agreement between the redshift distributions of the four samples Chehade et al. (2016) restricted their analysis to objects in the redshift range $0.3 < z < 2.9$. This leads to a total number of quasars N_q, of 22211, 32560, 6374 and 9705 for the 2QZ, SDSS, 2SLAQ and 2QDESp samples respectively. The mean of the correlation functions of these samples is taken to represent the correlation function of the combined quasar sample (henceforth referred to as the combined QSO sample), containing 70,940 quasars with a mean redshift of $\bar{z} = 1.49$ and an effective volume of $\approx 0.6h^{-3}\text{Gpc}^3$. For comparison the original SDSS LRG survey analysed by Eisenstein et al. (2005) covered an effective volume of $\approx 0.55h^{-3}\text{Gpc}^3$, while the BOSS DR12 LOWZ and CMASS samples analysed by Cuesta et al. (2016) cover effective volumes of $\approx 0.67h^{-3}\text{Gpc}^3$ and $\approx 1.58h^{-3}\text{Gpc}^3$ respectively (in all cases P_0 the amplitude of the power spectrum at the BAO scale is assumed to be $10,000h^{-3}\text{Mpc}^3$).

In line with Chehade et al. (2016), in this study we assume a flat ΛCDM cosmology with parameters given by the *Planck+lensing* results from Table 2 of Planck Collaboration et al. (2014). In this cosmology, $\Omega_m = 0.307$, $\Omega_b h^2 = 0.02217$, $\Omega_\Lambda = 0.693$, $\Omega_k = 0$, $\Omega_\nu = 0$, $h = 0.679$, $w = -1$, $n_s = 0.9635$ and $\sigma_8 = 0.823$. Based on this cosmology, the fiducial distances to $z = 1.49$ (the mean redshift of the combined QSO sample) are $D_{A,fid}(1.49) = 1789.50$ Mpc and $D_{V,fid}(1.49) = 3826.40$ Mpc, $H(1.49) = 158.31$ km s^{-1}Mpc^{-1} and $r_{d,fid} = 147.70$ Mpc.

3.3 Methodology

In this section we present a summary of the methodology applied by Chehade et al. (2016) in their measurements of the correlation function, as well as in our analysis of the BAO feature in the combined QSO sample.

The Landy-Szalay estimator (described in Section 2.3.1), was used by Chehade et al. (2016) in order to measure the correlation functions of the four samples, using random catalogues 20 times larger than the data in each sample. To account for effects of photometric and spectroscopic incompleteness, appropriate normalization has been applied to these randoms on a field to field basis.

We combine the four measured correlation functions by taking the mean and use the standard error on the mean as an estimate of the error. We then perform fitting to the resultant mean correlation function following the procedure described in Section 2.3.3.

As obtaining an accurate estimation of the covariance matrix for the combined QSO sample requires the generation of a large set of realistic mocks, a large task which lies beyond the scope of this work, when performing the fittings, we simply make use of the calculated standard error on the mean. Furthermore, due to time constraints we did not attempt to perform independent measurements of the correlation functions of the quasar samples, and are therefore limited to the choice of binning selected by Chehade et al. (2016). The fitting could however, benefit from a reduction in bin size in the measured correlation functions, providing more data points in the fitting range and allowing for a more accurate and precise determination of the position of the BAO peak. Due to these two constraints, we present our current measurement of the BAO scale in the combined QSO sample only as an initial estimate, with room for further improvement in future works.

3.4 Results and Discussion

In this section we present the results of our BAO analysis in the correlation function of the combined QSO sample.

3.4.1 Fitting the Combined QSO Sample

The results of fitting to the correlation function of the combined QSO sample with the complete ξ^{fit} model (equation 2.3.4), the ξ^{fit} model without the $A(s)$ nuisance fitting terms,

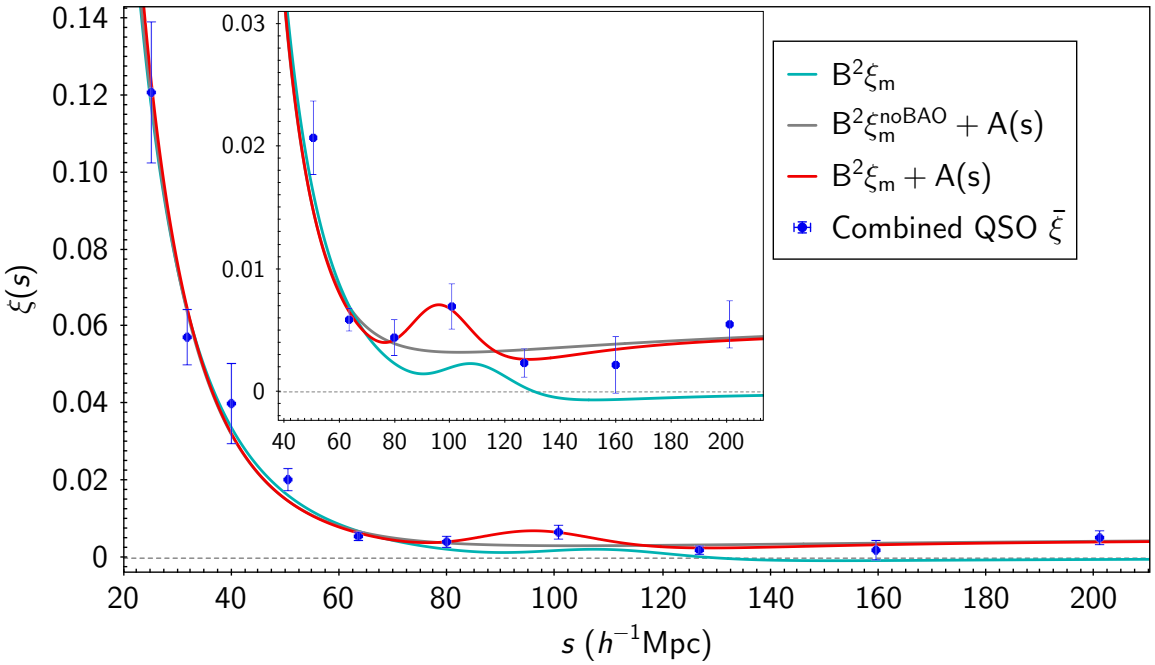


Figure 3.1: The results of fitting the mean of the 2QZ, SDSS, 2SLAQ and 2QDESp quasar samples correlation functions. The error bars are the standard error on the mean. The fitting is performed using various models for bins in the range $25 \leqslant s \leqslant 201h^{-1}$Mpc. The light blue curve is the ξ^{fit} model (equation 2.3.4) with the B fitting parameter only, while the red curve shows the complete ξ^{fit} model. The grey curve is the ξ^{noBAO} model fitted with the B and $A(s)$ fitting terms.

and a complete ξ^{noBAO} model in the range $25 \leqslant s \leqslant 201h^{-1}$Mpc, are presented in Fig. 3.1. As shown in this figure and in line with the discussion in the previous section, in the current choice of binning, the position of the fitted BAO peak is largely determined based on the position of the single bin located at $\approx 100h^{-1}$Mpc. The fitting can therefore benefit from having more bins in the region containing the BAO peak, allowing for a more accurate determination of the peak position. The values of α corresponding to the two variations of the ξ^{fit} model and the value of $D_V(z)$ corresponding to the full model, are presented in Table 3.1. Similar to fits performed in the previous chapter, upon performing an F-ratio test it can be seen that the complete model provides a better fit in comparison to the simple model ($F = 5.27$, $p = 0.05238$), althought in this case the complex model is preffered at a lower level of significane in comparison to fits discussed in the previous chapter.

3.4.2 Significance of BAO Peak Detection

The $\Delta\chi^2$ curves, from fitting the correlation function of the combined QSO sample, with the ξ^{fit} and ξ^{noBAO} models is presented in Fig. 3.2. A comparison of the two curves shows that the BAO peak is detected at $\approx 2.4\sigma$ in the data, and the plateau height of the curve

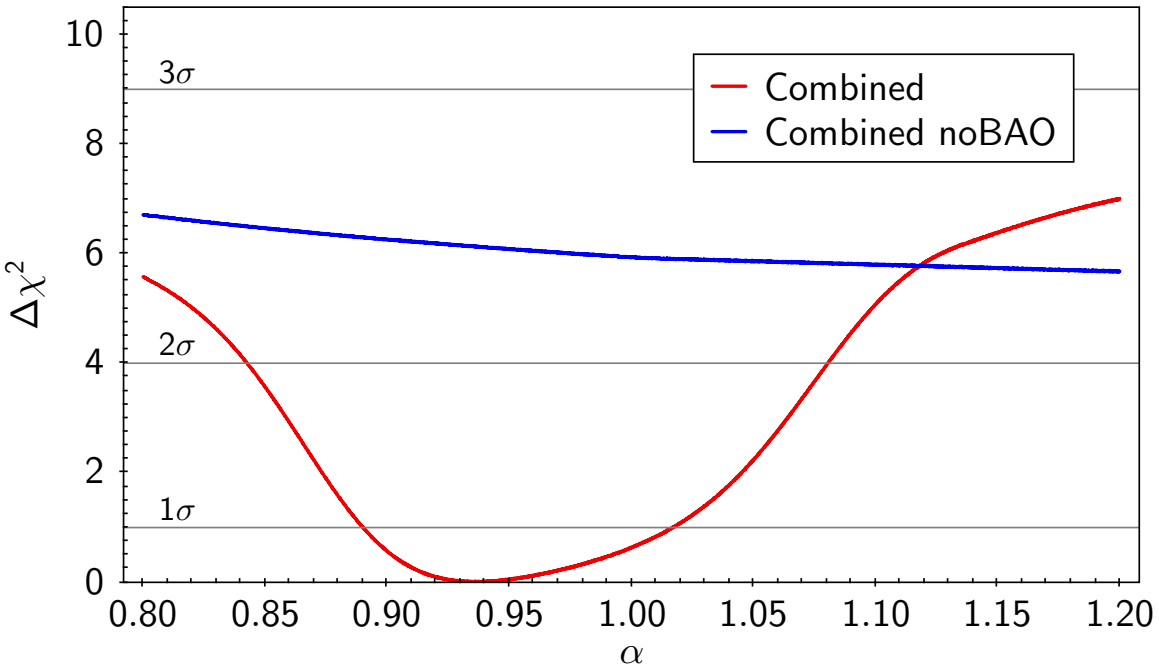


Figure 3.2: The significance of the detection of the BAO peak based on fitting to the correlation function of the combined QSO sample in the range $25 \leqslant s \leqslant 201h^{-1}$Mpc. The red curve correspond to a fits to the data based on the ξ^{fit} model which contains BAO, while the blue curve correspond to the fits based on the ξ^{noBAO} model with the BAO feature removed. In both cases the complete models including the $A(s)$ fitting terms are used. Here $\Delta\chi^2 = \chi^2(\alpha) - \chi^2_{min}$, where χ^2_{min} is the minimum χ^2 value using the model containing BAO. Comparing the two curves indicates that the BAO peak is detected at an $\approx 2.4\sigma$ level in the data.

corresponding to the ξ^{fit} model indicates that we can be confident in our measured value of α at an $\approx 2.3\sigma$ level.

3.4.3 BAO Distance Constraints on $D_V(z)$

We present our measured values of D_V for the combined QSO sample (Table 3.1), as well as for the BOSS DR12 CMASS and LOWZ samples (presented in Table 2.7 of Chapter 2), in Fig. 3.3. The pre-reconstruction measurements of D_V by Cuesta et al. (2016) based on the DR12 CMASS and LOWZ samples, as well as those from Beutler et al. (2011) for the 6dFGS sample, Ross et al. (2015) for the SDSS DR7 Main sample and Kazin et al. (2014) for the WiggleZ galaxy sample are also included for comparison. (Note that in the case of the WiggleZ distance measurements, we are plotting the post-reconstruction results due to much poorer constraints on the pre-reconstruction values). The flat ΛCDM prediction based on the Planck 2015 cosmology (TT, TE, EE+lowP+lensing+ext parameters from Table 4 of Planck Collaboration et al. 2015) is added for comparison. The grey region represents the

Table 3.1: Results of fitting the correlation functions of the combined quasar sample using the complete ξ^{fit} model described in Eq. 2.3.4 and the same model without the $A(s)$ nuisance fitting parameters, in the range $25 \leqslant s \leqslant 201h^{-1}$Mpc. The distance constraint on D_V calculated based on the measured value of α for the complete model is also included. Based on our fiducial cosmology, we assume a fiducial sound horizon value of $r_{d,fid} = 147.70$ Mpc. Here the value of D_V is quoted at the redshift of $z = 1.49$, corresponding to the mean redshift of the combined QSO sample.

Model	α	χ^2_{min}/dof	$D_V(z)r_{d,fid}/r_d$ (Mpc)
$B^2\xi_m + A(s)$	0.9363 ± 0.0652	6.7/5	3583 ± 249
$B^2\xi_m$	1.0641 ± 0.0671	27.9/8	—

1σ variation on the Planck prediction of $D_V(z)$. As these variations are dominated by the uncertainties in $\Omega_m h^2$ (see e.g. Anderson et al. 2014), this region is determined via sampling $\Omega_m h^2$ under the assumption that it follows a Gaussian distribution given by the Planck 2015 measurement and its 68 per cent confidence limit.

For the LOWZ sample, we find our measurement of D_V to be only in a mild tension with the Planck 2015 prediction, while there appears to be an $\approx 2\sigma$ tension between our measurement and the prediction for the CMASS sample. Although the measured D_V for the combined QSO sample also appears lower than the Planck prediction, the two results are in agreement within uncertainties.

The data in Fig. 3.3a can also be used to evaluate the plausibility of alternative cosmological models against observations. Here we generate a Einstein-de Sitter model by setting $\Omega_m = 1$ and using a value of $H_0 = 67.7$ km s^{-1}Mpc^{-1}. We obtain the best fit to the data by performing χ^2 minimisation, treating $(r_{d,fid}/r_d)$ as a free scaling parameter, with $r_{d,fid} = 147.50$ Mpc being the Planck 2015 value and r_d corresponding to the Einstein-de Sitter value (which in effect is our fitting parameter). We vary $(r_{d,fid}/r_d)$ in the range 1.0 to 1.5 in intervals of 0.001, taking the value corresponding to χ^2_{min} as the best fit. Here, χ^2 is given by

$$\chi^2(r_{d,fid}/r_d) = \sum_{i=1}^{N} \frac{(D_V^{obs}(i) - D_V^{fit}(i))^2}{\sigma^2_{D_V^{obs}}(i)}, \tag{3.4.1}$$

where i denotes the ith data point, N is the number of data points, D_V^{obs} is the value of the observed $D_V(z)(r_{d,fid}/r_d)$, with $\sigma_{D_V^{obs}}$ being the uncertainty on this value and D_V^{fit} is the corresponding model value at that redshift. Using the same approach as in Section 2.4.5, we assess the quality of the fits provided by the ΛCDM and Einstein-de Sitter models to the data based on their corresponding χ^2_{min}/dof values.

As there are two sets of measurements for the LOWZ and CMASS samples we compare the model to three subsets of the data with (i) our measurements, (ii) Cuesta et al. (2016)

Table 3.2: Results of fitting the ΛCDM and Einstein-de Sitter models to the $D_V(z)(r_{d,fid}/r_d)$ values plotted in Fig. 3.3a. As there are two sets of measurements for LOWZ and CMASS, we fit to three subsets of the data with (i) our measurements, (ii) Cuesta et al. (2016) measurements, and (iii) our CMASS and Cuesta et al. (2016)'s LOWZ measurements. Here $r_{d,fid}/r_d$ is our fitting (scaling) parameter, and the p-values indicate the significance of rejection of each model by the data.

		ΛCDM		Einstein-de Sitter	
Subset	$r_{d,fid}/r_d$	χ^2_{min}/dof	p-value	χ^2_{min}/dof	p-value
(i)	1.220	11.3/7	0.13	49.7/7	1.65×10^{-8}
(ii)	1.239	3.8/7	0.80	51.9/7	6.10×10^{-9}
(iii)	1.225	10.1/7	0.18	45.4/7	1.14×10^{-7}

measurements, and (iii) our CMASS and Cuesta et al. (2016)'s LOWZ data points. The yellow curve shown in Fig. 3.3a corresponds to the best fit Einstein-de Sitter model, using the Cuesta et al. (2016) measurements. The results of fitting the ΛCDM and Einstein-de Sitter models to the data are presented in Table 3.2. We find that the ΛCDM model provides a good fit to the data, while for the three different subsets, the Einstein-de Sitter model is rejected by the observations to a high level of significance as indicated by the very small p-values. We note however that if we use the pre-reconstruction WiggleZ values with larger errors and do not include the 6dFGS data point in the fit, we obtain an Einstein-de Sitter best fit model with the $r_{d,fid}/r_d = 1.232$ and $\chi^2_{min}/dof = 14.0/6$ (p=0.03).

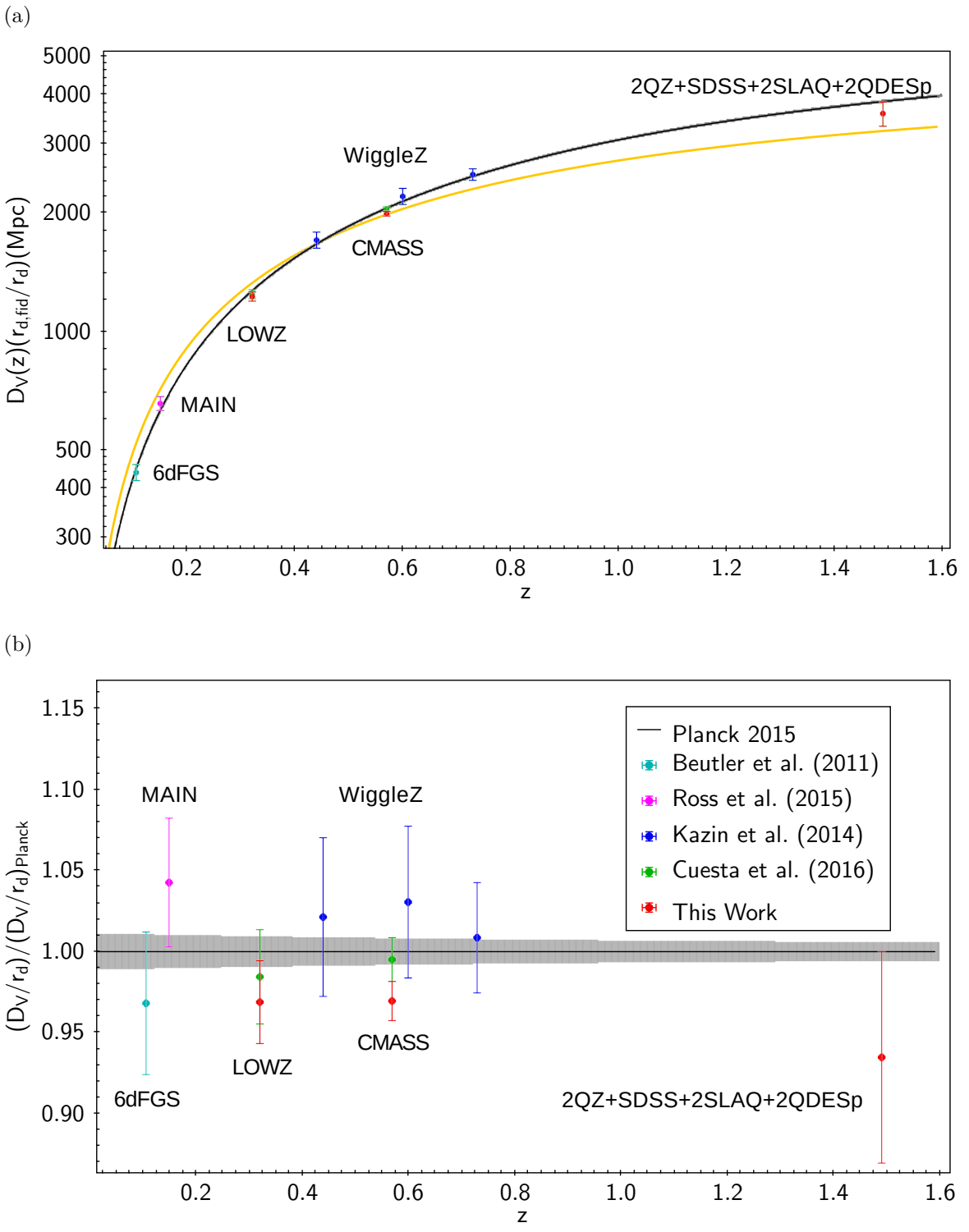


Figure 3.3: A comparison of our measured values of D_V for the LOWZ, CMASS and combined QSO samples (red data points) with the predictions based on a flat ΛCDM model with the Planck 2015 parameters (black line). The grey region represents the 1σ variation on the Planck prediction of $D_V(z)$. The measurements of D_V from Cuesta et al. (2016) for the CMASS and LOWZ samples (green data points), Beutler et al. (2011) for the 6dFGS sample (light blue data point), Ross et al. (2015) for the SDSS DR7 Main sample (pink data point) and Kazin et al. (2014) for the WiggleZ galaxy sample (dark blue data points), are also included for comparison. The yellow curve in subplot (a) is the best-fit Einstein-de Sitter model (i.e. scaled by $r_{d,fid}/r_d = 1.239$), fitted to a subset of the data with the LOWZ and CMASS measurements from Cuesta et al. (2016).

Chapter 4

Conclusions

In this study we have obtained an independent empirical measurement and estimation of errors on the correlation function of the BOSS DR12 LOWZ and CMASS samples. This was done by dividing each sample into 5 subsamples, measuring the correlation functions for these fields individually and taking the mean to represent the correlation function of the entire sample. We found an excellent agreement between our mean correlation functions and the correlation functions measured by Cuesta et al. (2016) for both samples, detecting the BAO peak at $\approx 4\sigma$ and $\approx 7\sigma$ for the LOWZ and CMASS samples respectively.

Based on our 5 subsamples, we obtained two estimates of uncertainties on the mean correlation function, using bootstrap resampling and the standard error on the mean and found good agreement between the two sets of errors. We also found our empirical estimates of uncertainties to be in general agreement with the uncertainties presented by Cuesta et al. (2016) which were calculated using covariance matrices from 1000 simulated DR12 QPM mocks. With the exception of the LOWZ sample where our uncertainties were lower than those from Cuesta et al. (2016) in the vicinity of the BAO peak, overall our findings serve as a verification that these simulated mocks are accurate representations of the data.

Based on performing isotropic fitting to the mean correlation functions we have obtained measurements of the position of the BAO peak. We found our results to be in agreement with those from Cuesta et al. (2016), finding only a small tension between the two measurements in the case of the CMASS sample. Furthermore, we have provided an empirical verification of the robustness of the fitting technique, by performing fits to the correlation functions of our 5 subsamples and taking the mean of the results as a second measurement of the BAO peak position, obtaining consistent results with the two approaches.

We have also demonstrated that the $A(s)$ nuisance fitting parameters play a significant role in producing a good fit, when fitting the correlation functions with a fiducial ΛCDM model. At our primary fitting range, based on the p-values corresponding to an F-ratio test ($p = 0.018$ and $p = 0.00084$ for the CMASS and LOWZ samples respectively), we have shown that the simple ΛCDM model without the $A(s)$ nuisance parameters is a significantly worse fit to the data compared to the full model, specially in the case of the LOWZ sample where the shape of the BAO peak is skewed to one side.

By testing the effect of the choice of fitting range on our measurements we have demonstrated that the measured position of the BAO peak and its uncertainty are largely insensitive to the choice of fitting range, while the significance of detection of the peak could vary considerably (by up to 1.9σ for the CMASS sample and 1.3σ for the LOWZ sample), depending on this choice.

We have investigated the possibility of obtaining an estimate of the covariance matrix from the data using the bootstrap method. We found our fitting results obtained using the bootstrap covariance matrix to be in good agreement with those obtained using the DR12 covariance matrix which is based on simulated QPM mocks. However, due to certain limitations in our methodology further investigation is required to determine the wider applicability of this technique for general use.

Based on our measurements of the position of the BAO peak in the LOWZ and CMASS samples, we have measured the volume-averaged distance to $z = 0.32$ and $z = 0.57$ to be $D_V(0.32)r_{d,fid}/r_d = 1226 \pm 32$ Mpc and $D_V(0.57)r_{d,fid}/r_d = 1988 \pm 24$ Mpc, a 2.6 per cent and 1.2 percent measurement respectively, assuming $r_{d,fid} = 147.10$ Mpc. We found our LOWZ measurement to be in agreement with the findings of Cuesta et al. (2016) within the uncertainties. However, we found a lower value for our measurement for the CMASS sample in comparison to the result from Cuesta et al. (2016) (an $\approx 1.4\sigma$ tension), as well as the prediction from Planck Collaboration et al. (2015) (an $\approx 2\sigma$ tension).

In Chapter 3 we extend our analysis to higher redshifts by performing fitting to the mean of the correlation functions obtained by Chehade et al. (2016), for the 2QZ, SDSS DR5, 2SLAQ and 2QDESp quasar samples. Here the BAO feature was detected at $\approx 2.4\sigma$ in the data. Fitting the correlation function resulted in a distance constraint of $D_V(z = 1.49)r_{d,fid}/r_d = 3583 \pm 249$ Mpc (assuming $r_{d,fid} = 147.70$ Mpc), a 6.9 per cent measurement to $z = 1.49$, the mean redshift of the combined QSO sample. This value is in agreement with the prediction from Planck Collaboration et al. (2015). As a covariance matrix based on simulated mocks is

yet to be constructed for the combined QSO sample, and the current binning of the correlation function is not ideal for the purpose of obtaining an accurate measurement of the BAO peak position, we present this distance measurement as an initial estimate which could be improved upon in future works.

Additionally, we made use of our measurements of $D_V(z)$ along with those from Cuesta et al. (2016) for the CMASS and LOWZ samples, Beutler et al. (2011) for the 6dFGS sample, Ross et al. (2015) for the SDSS DR7 Main sample and Kazin et al. (2014) for the WiggleZ galaxy sample, to test the plausibility of an Einstein-de Sitter model with $\Omega_m = 1$ and $H_0 = 67.7$ km s^{-1}Mpc^{-1}. We found that while a flat ΛCDM model based on Planck 2015 parameters is in good agreement with the data, a normalised Einstein-de Sitter model is ruled out by the observations at a high level of significance ($\chi^2_{min}/dof = 49.7/7$, p-value$= 1.65 \times 10^{-8}$).

Appendix A

Appendices

A.1 Correlation Functions of Individual Fields

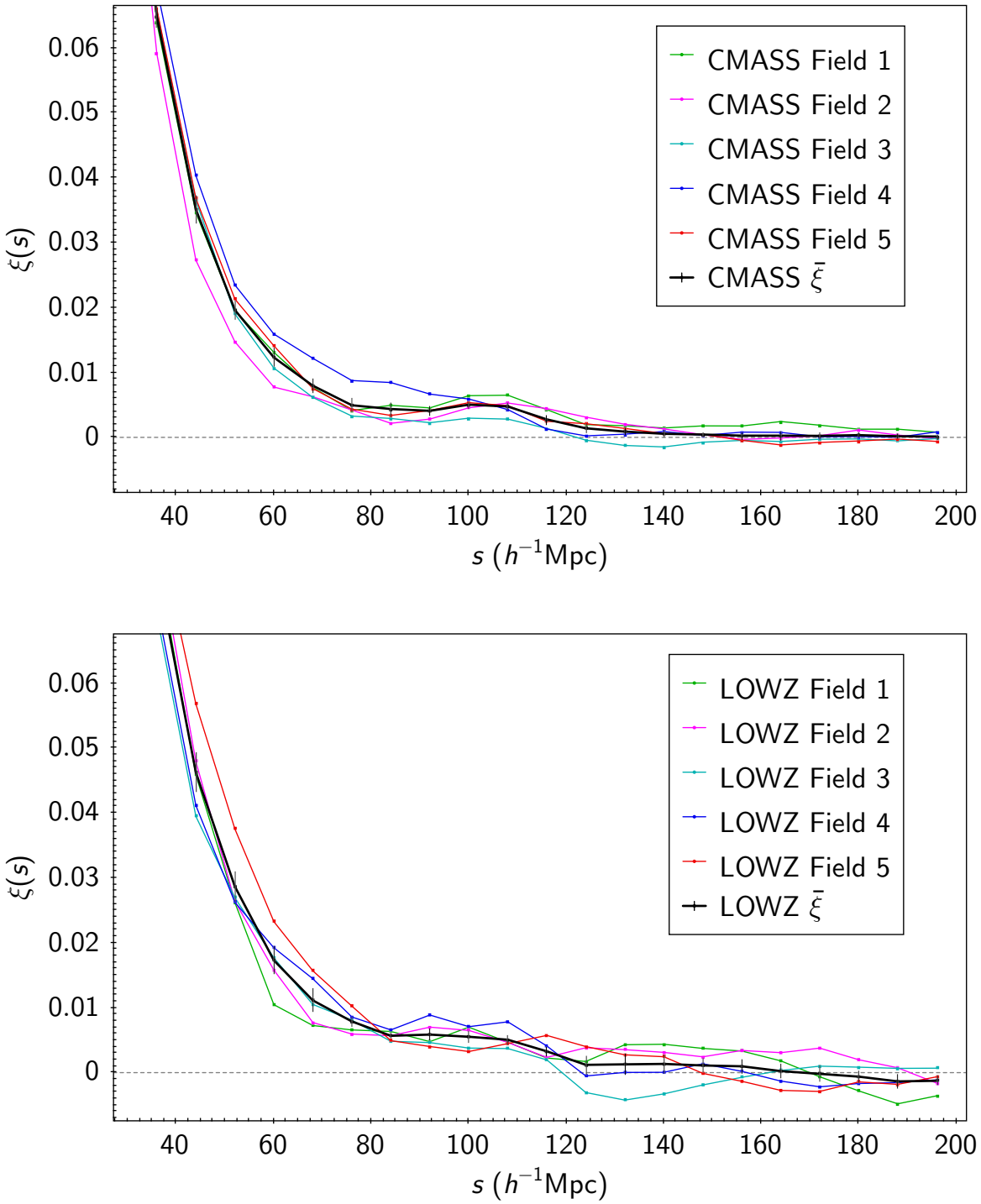


Figure A.1: The monopole correlation functions for the individual fields (coloured lines) and the corresponding mean correlation function (black line), of CMASS and LOWZ samples. The error bars on the mean correlation functions are the standard error on the mean. These correlation functions correspond to fields with the higher coverage and are the primary focus of our analysis in this study.

A.2 CMASS 30 Fields

(a) CMASS Northern Galactic Cap (b) CMASS Southern Galactic Cap

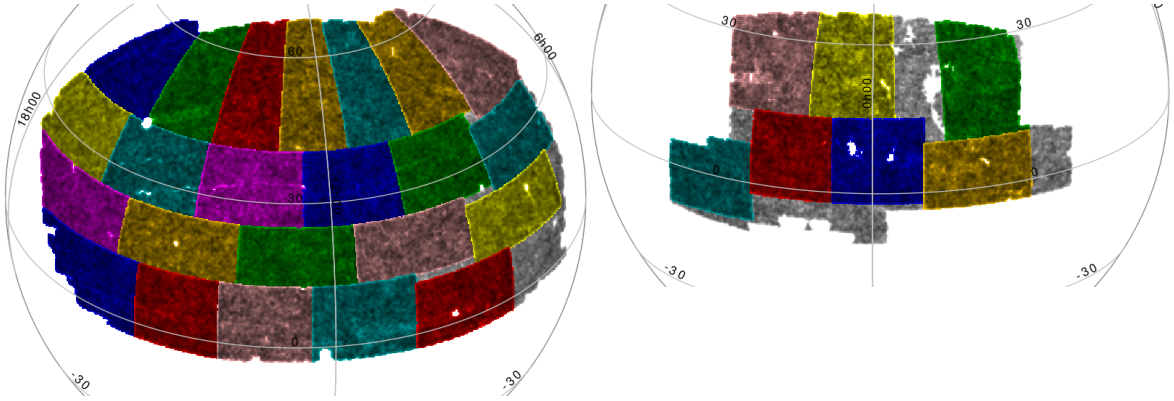


Figure A.2: The coverage of the chosen 30 fields in the Northern and Southern Galactic caps of the CMASS sample. The selected fields are highlighted by various colours while the unselected areas are shown in grey. These subsamples are used in Section 2.4.9 in order to obtain a bootstrap estimation of the covariance matrix from the data.

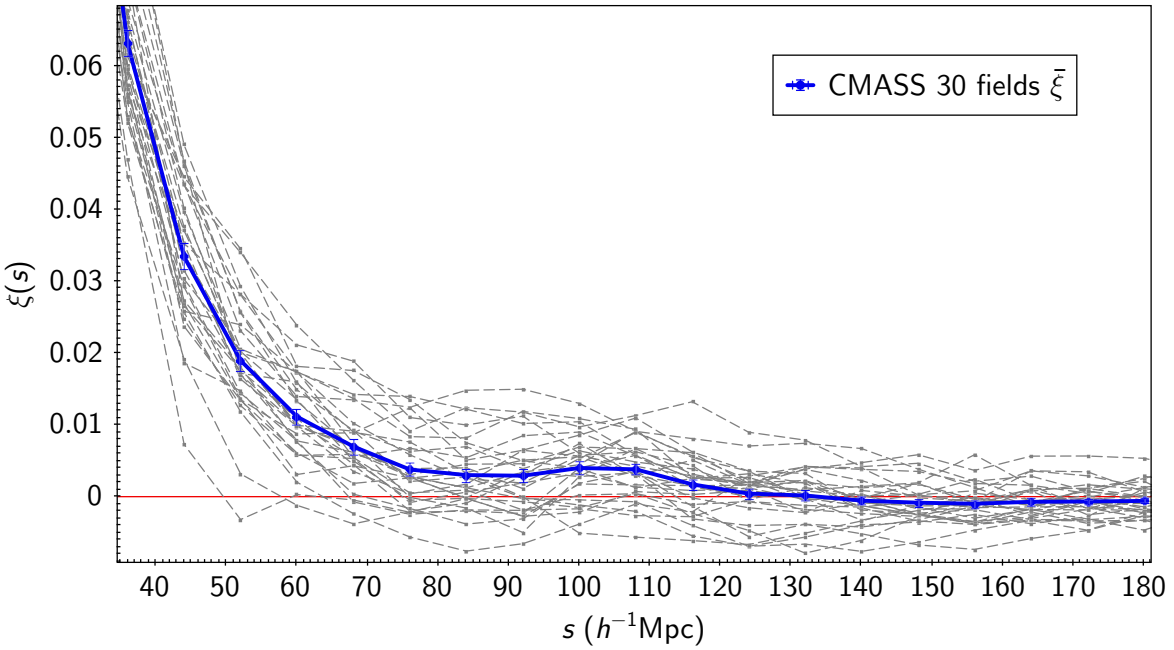


Figure A.3: Correlation functions of the 30 fields in the CMASS sample (grey dashed lines) and the mean correlation function (blue line). The error bars on the mean correlation function are the standard error on the mean.

Bibliography

Abazajian, K.N., Adelman-McCarthy, J.K., Agüeros, M.A., et al. *The Seventh Data Release of the Sloan Digital Sky Survey.* ApJS, **182** (2009), 543-558. 0812.0649.

Adelman-McCarthy, J.K., Agüeros, M.A., Allam, S.S., et al. *The Fifth Data Release of the Sloan Digital Sky Survey.* ApJS, **172** (2007), 634. 0707.3380.

Alam, S., Albareti, F.D., Allende Prieto, C., et al. *The Eleventh and Twelfth Data Releases of the Sloan Digital Sky Survey: Final Data from SDSS-III.* ApJS, **219** (2015), 12. 1501.00963.

Alonso, D. *CUTE solutions for two-point correlation functions from large cosmological datasets.* ArXiv e-prints (2012). 1210.1833.

Anderson, L., Aubourg, E., Bailey, S., et al. *The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: baryon acoustic oscillations in the Data Release 9 spectroscopic galaxy sample.* MNRAS, **427** (2012), 3435. 1203.6594.

Anderson, L., Aubourg, É., Bailey, S., et al. *The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: baryon acoustic oscillations in the Data Releases 10 and 11 Galaxy samples.* MNRAS, **441** (2014), 24. 1312.4877.

Angulo, R.E., Baugh, C.M., Frenk, C.S., et al. *The detectability of baryonic acoustic oscillations in future galaxy surveys.* MNRAS, **383** (2008), 755. astro-ph/0702543.

Ballinger, W.E., Peacock, J.A., Heavens, A.F. *Measuring the cosmological constant with redshift surveys.* MNRAS, **282** (1996), 877. astro-ph/9605017.

Beutler, F., Blake, C., Colless, M., et al. *The 6dF Galaxy Survey: baryon acoustic oscillations and the local Hubble constant.* MNRAS, **416** (2011), 3017. 1106.3366.

Blake, C., Glazebrook, K. *Probing Dark Energy Using Baryonic Oscillations in the Galaxy Power Spectrum as a Cosmological Ruler.* ApJ, **594** (2003), 665. astro-ph/0301632.

Blake, C., Kazin, E.A., Beutler, F., et al. *The WiggleZ Dark Energy Survey: mapping the distance-redshift relation with baryon acoustic oscillations.* MNRAS, **418** (2011), 1707. 1108.2635.

Busca, N.G., Delubac, T., Rich, J., et al. *Baryon acoustic oscillations in the Lyα forest of BOSS quasars.* A&A, **552** (2013), A96. 1211.2616.

Chehade, B., Shanks, T., Findlay, J., et al. *The 2QDES Pilot: the luminosity and redshift dependence of quasar clustering.* MNRAS, **459** (2016), 1179. 1603.04849.

Cole, S., Percival, W.J., Peacock, J.A., et al. *The 2dF Galaxy Redshift Survey: power-spectrum analysis of the final data set and cosmological implications.* MNRAS, **362** (2005), 505. astro-ph/0501174.

Colless, M., Peterson, B.A., Jackson, C., et al. *The 2dF Galaxy Redshift Survey: Final Data Release.* ArXiv Astrophysics e-prints (2003). astro-ph/0306581.

Crocce, M., Scoccimarro, R. *Nonlinear evolution of baryon acoustic oscillations.* Phys. Rev. D, **77**(2) (2008), 023533. 0704.2783.

Croom, S.M., Richards, G.T., Shanks, T., et al. *The 2dF-SDSS LRG and QSO Survey: the spectroscopic QSO catalogue.* MNRAS, **392** (2009), 19. 0810.4955.

Croom, S.M., Smith, R.J., Boyle, B.J., et al. *The 2dF QSO Redshift Survey - XII. The spectroscopic catalogue and luminosity function.* MNRAS, **349** (2004), 1397. astro-ph/0403040.

Cuesta, A.J., Vargas-Magaña, M., Beutler, F., et al. *The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: baryon acoustic oscillations in the correlation function of LOWZ and CMASS galaxies in Data Release 12.* MNRAS, **457** (2016), 1770. 1509.06371.

Dawson, K.S., Kneib, J.P., Percival, W.J., et al. *The SDSS-IV Extended Baryon Oscillation Spectroscopic Survey: Overview and Early Data.* AJ, **151** (2016), 44. 1508.04473.

Dawson, K.S., Schlegel, D.J., Ahn, C.P., et al. *The Baryon Oscillation Spectroscopic Survey of SDSS-III.* AJ, **145** (2013), 10. 1208.0022.

Delubac, T., Bautista, J.E., Busca, N.G., et al. *Baryon acoustic oscillations in the Lyα forest of BOSS DR11 quasars.* A&A, **574** (2015), A59. 1404.1801.

Dolney, D., Jain, B., Takada, M. *Baryon oscillations and dark-energy constraints from imaging surveys.* MNRAS, **366** (2006), 884. astro-ph/0409445.

Doroshkevich, A.G., Zel'dovich, Y.B., Syunyaev, R.A. *Fluctuations of the microwave background radiation in the adiabatic and entropic theories of galaxy formation.* Soviet Ast., **22** (1978), 523.

Drinkwater, M.J., Jurek, R.J., Blake, C., et al. *The WiggleZ Dark Energy Survey: survey design and first data release.* MNRAS, **401** (2010), 1429. 0911.4246.

Eisenstein, D.J., Hu, W. *Baryonic Features in the Matter Transfer Function.* ApJ, **496** (1998), 605. astro-ph/9709112.

Eisenstein, D.J., Seo, H.J., Sirko, E., et al. *Improving Cosmological Distance Measurements by Reconstruction of the Baryon Acoustic Peak.* ApJ, **664** (2007a), 675. astro-ph/0604362.

Eisenstein, D.J., Seo, H.J., White, M. *On the Robustness of the Acoustic Scale in the Low-Redshift Clustering of Matter.* ApJ, **664** (2007b), 660. astro-ph/0604361.

Eisenstein, D.J., Weinberg, D.H., Agol, E., et al. *SDSS-III: Massive Spectroscopic Surveys of the Distant Universe, the Milky Way, and Extra-Solar Planetary Systems.* AJ, **142** (2011), 72. 1101.1529.

Eisenstein, D.J., Zehavi, I., Hogg, D.W., et al. *Detection of the Baryon Acoustic Peak in the Large-Scale Correlation Function of SDSS Luminous Red Galaxies.* ApJ, **633** (2005), 560. astro-ph/0501171.

Feldman, H.A., Kaiser, N., Peacock, J.A. *Power-spectrum analysis of three-dimensional redshift surveys.* ApJ, **426** (1994), 23. astro-ph/9304022.

Font-Ribera, A., Kirkby, D., Busca, N., et al. *Quasar-Lyman α forest cross-correlation from BOSS DR11: Baryon Acoustic Oscillations.* J. Cosmology Astropart. Phys., **5** (2014), 027. 1311.1767.

Gaztañaga, E., Cabré, A., Hui, L. *Clustering of luminous red galaxies - IV. Baryon acoustic peak in the line-of-sight direction and a direct measurement of H(z).* MNRAS, **399** (2009), 1663. 0807.3551.

Gaztañaga, E., Scoccimarro, R. *The three-point function in large-scale structure: redshift distortions and galaxy bias.* MNRAS, **361** (2005), 824. astro-ph/0501637.

Glazebrook, K., Blake, C. *Measuring the Cosmic Evolution of Dark Energy with Baryonic Oscillations in the Galaxy Power Spectrum.* ApJ, **631** (2005), 1. astro-ph/0505608.

Guzik, J., Bernstein, G., Smith, R.E. *Systematic effects in the sound horizon scale measurements.* MNRAS, **375** (2007), 1329. astro-ph/0605594.

Hartlap, J., Simon, P., Schneider, P. *Why your model parameter confidences might be too optimistic. Unbiased estimation of the inverse covariance matrix.* A&A, **464** (2007), 399. astro-ph/0608064.

Hill, G.J., Gebhardt, K., Komatsu, E., et al. *The Hobby-Eberly Telescope Dark Energy Experiment (HETDEX): Description and Early Pilot Survey Results.* In T. Kodama, T. Yamada, K. Aoki, editors, *Panoramic Views of Galaxy Formation and Evolution*, volume 399 of *Astronomical Society of the Pacific Conference Series* (2008), page 115. 0806.0183.

Hu, W., White, M. *Acoustic Signatures in the Cosmic Microwave Background.* ApJ, **471** (1996), 30. astro-ph/9602019.

Hütsi, G. *Acoustic oscillations in the SDSS DR4 luminous red galaxy sample power spectrum.* AAP, **449** (2006), 891. astro-ph/0512201.

Jeong, D., Komatsu, E. *Perturbation Theory Reloaded: Analytical Calculation of Nonlinearity in Baryonic Oscillations in the Real-Space Matter Power Spectrum.* ApJ, **651** (2006), 619. astro-ph/0604075.

Jones, D.H., Read, M.A., Saunders, W., et al. *The 6dF Galaxy Survey: final redshift release (DR3) and southern large-scale structures.* MNRAS, **399** (2009), 683. 0903.5451.

Kaiser, N. *Clustering in real space and in redshift space.* MNRAS, **227** (1987), 1.

Kazin, E.A., Blanton, M.R., Scoccimarro, R., et al. *The Baryonic Acoustic Feature and Large-Scale Clustering in the Sloan Digital Sky Survey Luminous Red Galaxy Sample.* ApJ, **710** (2010), 1444. 0908.2598.

Kazin, E.A., Koda, J., Blake, C., et al. *The WiggleZ Dark Energy Survey: improved distance measurements to $z = 1$ with reconstruction of the baryonic acoustic feature.* MNRAS, **441** (2014), 3524. 1401.0358.

Kirkby, D., Margala, D., Slosar, A., et al. *Fitting methods for baryon acoustic oscillations in the Lyman-α forest fluctuations in BOSS data release 9.* J. Cosmology Astropart. Phys., **3** (2013), 024. 1301.3456.

Landy, S.D., Szalay, A.S. *Bias and variance of angular correlation functions.* ApJ, **412** (1993), 64.

Laureijs, R., Amiaux, J., Arduini, S., et al. *Euclid Definition Study Report.* ArXiv e-prints (2011). 1110.3193.

Levi, M., Bebek, C., Beers, T., et al. *The DESI Experiment, a whitepaper for Snowmass 2013.* ArXiv e-prints (2013). 1308.0847.

Lewis, A., Challinor, A., Lasenby, A. *Efficient Computation of Cosmic Microwave Background Anisotropies in Closed Friedmann-Robertson-Walker Models.* ApJ, **538** (2000), 473. astro-ph/9911177.

Linder, E.V. *Baryon oscillations as a cosmological probe.* Phys. Rev. D, **68**(8) (2003), 083504. astro-ph/0304001.

LSST Dark Energy Science Collaboration. *Large Synoptic Survey Telescope: Dark Energy Science Collaboration.* ArXiv e-prints (2012). 1211.0310.

Matsubara, T. *Correlation Function in Deep Redshift Space as a Cosmological Probe.* The Astrophysical Journal, **615**(2) (2004), 573. URL http://stacks.iop.org/0004-637X/615/i=2/a=573.

Matsubara, T. *Resumming cosmological perturbations via the Lagrangian picture: One-loop results in real space and in redshift space.* Phys. Rev. D, **77**(6) (2008), 063530. 0711.2521.

Norberg, P., Baugh, C.M., Gaztañaga, E., et al. *Statistical analysis of galaxy surveys - I. Robust error estimation for two-point clustering statistics.* MNRAS, **396** (2009), 19. 0810.1885.

Padmanabhan, N., Schlegel, D.J., Seljak, U., et al. *The clustering of luminous red galaxies in the Sloan Digital Sky Survey imaging data.* MNRAS, **378** (2007), 852. astro-ph/0605302.

Padmanabhan, N., White, M. *Calibrating the baryon oscillation ruler for matter and halos.* Phys. Rev. D, **80**(6) (2009), 063508. 0906.1198.

Pan, J., Szapudi, I. *The monopole moment of the three-point correlation function of the two-degree Field Galaxy Redshift Survey.* MNRAS, **362** (2005), 1363. astro-ph/0505422.

Peebles, P.J., Ratra, B. *The cosmological constant and dark energy.* Reviews of Modern Physics, **75** (2003), 559. astro-ph/0207347.

Peebles, P.J.E., Yu, J.T. *Primeval Adiabatic Perturbation in an Expanding Universe.* ApJ, **162** (1970), 815.

Percival, W.J., Reid, B.A., Eisenstein, D.J., et al. *Baryon acoustic oscillations in the Sloan Digital Sky Survey Data Release 7 galaxy sample.* MNRAS, **401** (2010), 2148. 0907.1660.

Perlmutter, S., Aldering, G., Goldhaber, G., et al. *Measurements of Ω and Λ from 42 High-Redshift Supernovae.* ApJ, **517** (1999), 565. astro-ph/9812133.

Planck Collaboration, Ade, P.A.R., Aghanim, N., et al. *Planck 2013 results. XVI. Cosmological parameters.* A&A, **571** (2014), A16. 1303.5076.

Planck Collaboration, Ade, P.A.R., Aghanim, N., et al. *Planck 2015 results. XIII. Cosmological parameters.* ArXiv e-prints (2015). 1502.01589.

Porciani, C., Giavalisco, M. *The Clustering Properties of Lyman Break Galaxies at Redshift z˜3.* ApJ, **565** (2002), 24. astro-ph/0107447.

Porciani, C., Norberg, P. *Luminosity- and redshift-dependent quasar clustering.* MNRAS, **371** (2006), 1824. astro-ph/0607348.

Reid, B., Ho, S., Padmanabhan, N., et al. *SDSS-III Baryon Oscillation Spectroscopic Survey Data Release 12: galaxy target selection and large-scale structure catalogues.* MNRAS, **455** (2016), 1553. 1509.06529.

Richards, G.T., Croom, S.M., Anderson, S.F., et al. *The 2dF-SDSS LRG and QSO (2SLAQ) Survey: the $z < 2.1$ quasar luminosity function from 5645 quasars to $g = 21.85$.* MNRAS, **360** (2005), 839. astro-ph/0504300.

Riess, A.G., Filippenko, A.V., Challis, P., et al. *Observational Evidence from Supernovae for an Accelerating Universe and a Cosmological Constant.* AJ, **116** (1998), 1009. astro-ph/9805201.

Ross, A.J., Beutler, F., Chuang, C.H., et al. *The clustering of galaxies in the completed SDSS-III Baryon Oscillation Spectroscopic Survey: Observational systematics and baryon acoustic oscillations in the correlation function.* ArXiv e-prints (2016). 1607.03145.

Ross, A.J., Samushia, L., Howlett, C., et al. *The clustering of the SDSS DR7 main Galaxy sample - I. A 4 per cent distance measure at $z = 0.15$.* MNRAS, **449** (2015), 835. 1409.3242.

Ross, N.P., Shen, Y., Strauss, M.A., et al. *Clustering of Low-redshift (z <= 2.2) Quasars from the Sloan Digital Sky Survey.* ApJ, **697** (2009), 1634. 0903.3230.

Sánchez, A.G., Baugh, C.M., Angulo, R.E. *What is the best way to measure baryonic acoustic oscillations?* MNRAS, **390** (2008), 1470. 0804.0233.

Schneider, D.P., Hall, P.B., Richards, G.T., et al. *The Sloan Digital Sky Survey Quasar Catalog. IV. Fifth Data Release.* AJ, **134** (2007), 102. 0704.0806.

Seo, H.J., Eisenstein, D.J. *Probing Dark Energy with Baryonic Acoustic Oscillations from Future Large Galaxy Redshift Surveys.* ApJ, **598** (2003), 720. astro-ph/0307460.

Shanks, T. *Arguments for an Omega = 1, low H0 baryon dominated universe.* Vistas in Astronomy, **28** (1985), 595.

Slosar, A., Iršič, V., Kirkby, D., et al. *Measurement of baryon acoustic oscillations in the Lyman-α forest fluctuations in BOSS data release 9.* J. Cosmology Astropart. Phys., **4** (2013), 026. 1301.3459.

Smith, R.J., Croom, S.M., Boyle, B.J., et al. *The 2dF QSO Redshift Survey - III. The input catalogue.* MNRAS, **359** (2005), 57.

Spergel, D., Gehrels, N., Baltay, C., et al. *Wide-Field InfrarRed Survey Telescope-Astrophysics Focused Telescope Assets WFIRST-AFTA 2015 Report.* ArXiv e-prints (2015). 1503.03757.

Squires, G.L. *Practical Physics.* Cambridge University Press, Cambridge, 4th edition (2001).

Sunyaev, R.A., Zeldovich, Y.B. *Small-Scale Fluctuations of Relic Radiation.* Ap&SS, **7** (1970), 3.

Veropalumbo, A., Marulli, F., Moscardini, L., et al. *An improved measurement of baryon acoustic oscillations from the correlation function of galaxy clusters at $z \sim 0.3$.* MNRAS, **442** (2014), 3275. 1311.5895.

Weinberg, D.H., Mortonson, M.J., Eisenstein, D.J., et al. *Observational probes of cosmic acceleration.* Phys. Rep., **530** (2013), 87. 1201.2434.

White, M., Tinker, J.L., McBride, C.K. *Mock galaxy catalogues using the quick particle mesh method.* MNRAS, **437** (2014), 2594. 1309.5532.

Xu, X., Padmanabhan, N., Eisenstein, D.J., et al. *A 2 per cent distance to z = 0.35 by reconstructing baryon acoustic oscillations - II. Fitting techniques.* MNRAS, **427** (2012), 2146. 1202.0091.

York, D.G., Adelman, J., Anderson, Jr., J.E., et al. *The Sloan Digital Sky Survey: Technical Summary.* AJ, **120** (2000), 1579. astro-ph/0006396.