Detection of Baryonic Acoustic Oscillations in the Matter Power Spectrum

Spencer Everett
DePaul University, spencerweverett@gmail.com

Ian Johnson
DePaul University, johnson_ian@me.com

Jon Murphy
DePaul University, jmurphy41@gmail.com

Mary Tarpley
DePaul University, tarpleymk@gmail.com

Detection of Baryonic Acoustic Oscillations in the Matter Power Spectrum

Acknowledgements

Detection of Baryonic Acoustic Oscillations in the Matter Power Spectrum

Spencer Everett
Department of Physics
Ian Johnson
Department of Physics
Jon Murphy
Department of Physics
Mary Tarpley
Department of Physics

ABSTRACT Using the spectra of 22,923 high-redshift quasars from the Baryon Oscillation Spectrosocpic Survey (BOSS) subset of the Sloan Digital Sky Survey (SDSS), the authors detect evidence of the primordial baryonic acoustic oscillations (BAOs) in the matter power spectrum. The detection further endorses the currently accepted Λ-CDM model of cosmology based upon the existence of dark energy (Λ) and cold dark matter (CDM). Additionally, the use of the continuous wavelet transform to calculate the power spectrum has many advantages over traditional Fourier methods and independently corroborates previous detections.

1. INTRODUCTION

The early universe was a largely homogenous plasma of ionized gas in which light was coupled to baryonic, or ordinary, matter. However, small inhomogeneities arose from matter density perturbations left over from cosmic inflation: a very short period of massive expansion immediately after the Big Bang [1]. Over time the slightly unequal distribution of matter caused baryons to collapse into overdense regions of dark matter. As the density increased, photons coupled to the baryons exerted radiation pressure, which drove the baryonic matter outwards from the perturbation. Gravity again became dominant as the density decreased and a cosmic tug-of-war ensued for the first 380,000 years of the universe resulting in matter waves known as baryonic acoustic oscillations (BAOs).

Once the universe cooled to near $3000°$ K, neutral hydrogen atoms formed and light decoupled from matter in an event known as recombination. With the absence of photonic pressure, the baryonic oscillations ceased. This suppression led to the formation of baryonic structures, which over time led to the creation of galaxies [2]. The imprint left by photons during recombination was first detected in the 1960s as the cosmic microwave background radiation (CMB) [3]. This has been thoroughly researched, most recently by the Wilkinson Microwave Anisotropy Probe (WMAP) [4] and the Planck satellite [5].

— 1 —

However, it is more difficult to measure the imprint left on baryonic matter than on photons because the signal has been diminished by billions of years of nonlinear structure formation. If the matter distribution of a large section of the universe is known, then evidence of BAOs should be detectable as a characteristic scale of matter separation that is statistically significant when compared to a random sample. Historically, this has been measured by examining the large-scale structure of matter in astronomical surveys. The largest data set of this kind comes from the Sloan Digital Sky Survey (SDSS), an ongoing spectroscopic redshift survey that uses a 2.5 m wide-angle optical telescope at the Apache Point Observatory in Sunspot, New Mexico. SDSS has imaged over 500 million objects and recorded the spectra of more than 1 million objects in the Northern sky [6]. The Baryon Oscillation Spectroscopic Survey (BOSS) is a subset of SDSS that captures the spectra and spatial distribution of quasars and luminous red galaxies. One of its goals is to detect the characteristic scale of the primeval oscillations [7]. Using spectroscopic data of 22,923 quasars from BOSS and improving upon the methodology of Everett *et al.* in 2014 [8], the authors calculate the matter power spectrum using the continuous wavelet transform and detect features indicating the existence of BAOs not present in [8].

2. THEORY

BAOs were first theorized by Peebles and Yu in 1970 as "primeval adiabatic perturbations" left over from the Big Bang that catalyzed the formation of galaxies [9]. Evidence of the oscillations was first detected by Eisenstein *et al.* in 2005 [10] and has since been corroborated by Martinez *et al.* in 2009 [11] and Labatie *et al.* in 2012 [12]. Research into BAOs has increased greatly in the past decade as the size of the oscillations at recombination, called the sound horizon, may be used as a standard ruler to study the acceleration of the universe independent of Type 1a supernovas [13]. Additionally, the exact size and features of the oscillations help constrain the currently accepted concordance model of cosmology known as the Λ-CDM model, which is based upon

the theorized existence of dark energy (Λ) and cold dark matter (CDM) [14].

Previous attempts to detect BAOs by Eisenstein *et al.* and others ([11], [12]) have used galaxies as their data sample as they are numerous and easy to detect. However, these galaxies were created relatively recently in the cosmic timescale making the BAO signature difficult to detect. Instead of galaxies, this research uses distant intergalactic clouds of hydrogen that were present long before most galaxies formed. These clouds are illuminated by very old and bright galaxies, called quasars, whose light is absorbed and scattered by the clouds at specific wavelengths. The most common of which is at 1216 $\mathring{A}$ and is called the Lyman-alpha line. As the universe itself expands, the wavelength of the absorbed light is stretched proportionally to how far it travels. This distortion is called redshift. The effects of redshift are analogous to the Doppler Effect, in which the frequency of the siren of an ambulance is shifted as it moves toward or away an observer. Due to redshift, the individual absorption lines are stretched into a collection of absorption lines known as the Lyman-alpha forest [15]. If it is assumed that the missing flux from a quasar spectrum is proportional to the amount of baryonic matter present, then one is able to create a map of the matter distribution in the early universe as a function of redshift.

If the redshifting of the Lyman-alpha lines is assumed to be entirely cosmological in origin (*i.e.* from the expansion of the universe), then it is ideal to compare the distances of objects by the difference in their redshifts. Unfortunately, small peculiar velocities of the hydrogen clouds not associated with the expansion of space can cause distortions in redshift space. Specifically, these random deviations in velocity cause objects with similar physical distances to have slightly different redshifts. This elongates the objects radially towards the observer and is known as the Fingers-of-God effect [16].

As this research searches for structure at specific scales, redshift distortions can lead to detection of artificial structures which is a large source of potential error. Instead spectra are often converted from redshift space to comoving physical

— 2 —

space. To do this, one must assume a cosmology which requires many assumptions about the makeup of the universe such as its curvature and distribution of energy density. Once in physical space, the most direct method to detect BAOs is to use a two-point correlation function which measures the excess probability of finding objects separated at a certain distance compared to a random distribution. However, this is computationally very expensive to calculate, and it is more common to calculate the power spectrum which measures matter density as a function of scale [17]. If the acoustic oscillations were present in the early universe, then the amplitude of the power spectrum will have a small, but distinct, peak at scales corresponding to the size of the sound horizon.

While this is traditionally calculated using a Fourier transform, instead a continuous wavelet transform was chosen for this research as it has many advantages over Fourier methods. Firstly, wavelets contain localized information of the data. In the context of this research a large Fourier coefficient at a given scale would suggest a large number of objects are separated by that distance, but a wavelet coefficient would *also* indicate specific locations in the spectra where the object separations at this scale occurred. Localization also reduces the effect of outliers as only wavelet coefficients near the unwanted data point will be affected, while in a Fourier transform all coefficients would be altered. Wavelets also come in many different types that can be used to serve different purposes. For example, Haar wavelets are excellent at detecting discontinuities [18] while Morlet wavelets are related to human perceptions of hearing and vision [19]. As a final advantage, detection of the oscillations using wavelets will independently verify the results of previous detections that used Fourier methods.

3. **METHODS**

The data set used in this research was comprised of the Lyman-alpha Forest Catalog and Sample from the ninth data release (DR9) of the BOSS component of SDSS-III. In this catalog are the spectra of 54,468 quasars with redshift $z > 2.15$, covering an area of 3,275 deg^2 in the
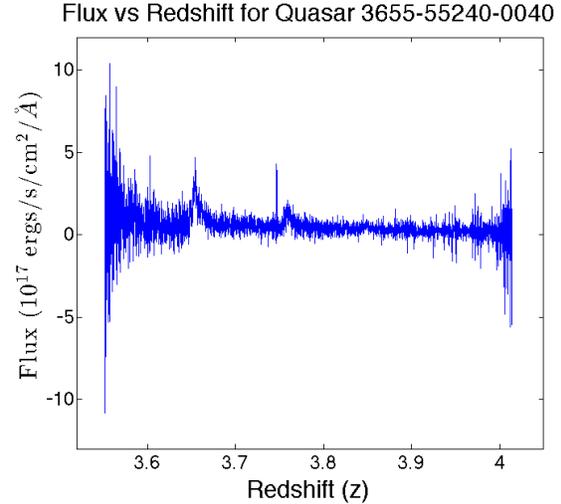

FIGURE 1. The raw spectrum of quasar 3655-55240-0040, one of 54,468 spectra included in the DR9 release of BOSS subset of SDSS. The flux density subtracted from the mean flux density of the universe is plotted against redshift z. Each dip in flux density is assumed to be the result of scattering by an intergalactic cloud of hydrogen at the corresponding redshift. The collection of these dips is known as the Lyman-alpha forest.

northern sky [20]. These spectra are distributed unequally across 817 plates where each plate covers 3.1 deg^2 of the sky. A single spectrum from quasar 3655-55240-0040 is shown in Figure 1 where the spectral flux density is plotted against the redshift. The large spike in flux on the right corresponds to the quasar and the collection of dips in the flux density is the Lyman-alpha forest.

The quasars vary in redshift from $z = 2.15$ to $z = 5.7$, which signifies nearly two billion years of expansion and structure formation. It is difficult to detect the oscillation signature over such a large time period. At low redshifts, nonlinear structure formation begins and diminishes the BAO imprint on the matter density. While the imprint should be clearer at high redshifts, there

— 3 —

are far fewer quasars at redshifts above $z = 4$. Thus any statistical analysis at these large redshift values will be less reliable as the sample size is too small. Because of these constraints the analysis was limited to a narrower Δz range. To do this, the mean minimum redshift $\bar{z}_{min}$ and mean maximum redshift $\bar{z}_{max}$ of all plates was found and only spectrum flux within this range was considered. It was found that $\bar{z}_{min} = 2.2$ and $\bar{z}_{max} = 3.6$ which gives a $\Delta z = 1.4$. A histogram of this redshift data is shown in Figure 2.

All spectra in the catalog have a signal-to-noise (SNR) ratio of at least 0.5. However, this means unwanted noise could potentially constitute two-thirds of a spectrum's data. To improve the quality of the sample, the analysis was restricted to spectra with SNR > 2. This reduced the sample size to 22,923 quasars. Once this sample was chosen the Lyman-alpha forest transmission field F_i was extracted from quasar spectrum i using

$$(1) \quad F_i(z) = f_{p,i}(\lambda) \cdot \left(\frac{\epsilon_{dla,i}(\lambda)}{\epsilon_{flux}(\lambda) \cdot C_{MF,i}(\lambda)} \right)$$

where $f_{p,i}$ is the flux of spectrum i, $\epsilon_{dla,i}$ is the damped Lyman-alpha (DLA) wings correction for spectrum i, ϵ_{flux} is the flux calibration corrections, and $C_{MF,i}$ is the quasar continuum in the rest frame 1040-1600 $\mathring{A}$ region for spectrum i. While the main purpose of this equation is to separate the Lyman-alpha forest from the background quasar flux (C_{MF}), it also calibrates each plate fiber (ϵ_{flux}) and accounts for the damping of photon flux when a photon is incident on the edge of a pixel (ϵ_{dla}) [20].

To avoid the effects of redshift distortion, the data were converted from redshift space to comoving physical space using the relation

$$(2) \quad R_0 \, dr = \frac{c}{H_0}[(1 - \Omega)(1 + z)^2 + \Omega_\Lambda$$
$$+ \Omega_m(1 + z)^3 + \Omega_r(1 + z)^4]^{-1/2} \, dz$$

where H_0 is Hubble's constant, c is the speed of light, z is the redshift, Ω is the total density parameter, Ω_Λ is the density parameter of vacuum energy (dark energy), Ω_m is the density parameter of baryonic and dark matter, Ω_r is the density parameter of relativistic particles, and R_0 is the
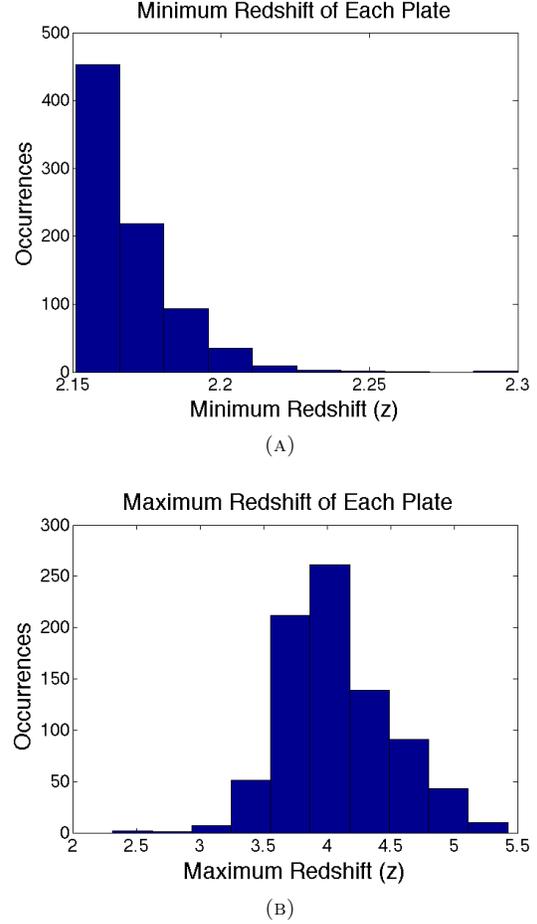

FIGURE 2. Two histograms showing the minimum redshift (A) and maximum redshift (B) of all 817 plates. It was found that $\bar{z}_{min} = 2.2$ and $\bar{z}_{max} = 3.6$. Finding a common minimum and maximum redshift across all plates restricted the analysis to a smaller period of structure formation and to distances which had the most data.

radial distance from Earth [21]. The most recent Planck satellite data provided parameter values of $H_0 = 68 \, \text{km·s}^{-1} \cdot \text{Mpc}^{-1}$, $\Omega = 1.01$, $\Omega_\Lambda = 0.691$, $\Omega_m = 0.318$, and $\Omega_r = 8.24 \times 10^{-5}$ [5] where Mpc

— 4 —

is a megaparsec which is approximately 3.3 million light years. These values are in agreement with the concordance model of cosmology [14].

To make the analysis computationally feasible, all of the spectra from a given plate were combined into a single spectrum spanning the entire Δz range. This is a reasonable approximation as all quasars of a given plate are within a 3.1 deg^2 area. Thus the data sample collapsed from 22,923 single quasar spectra into 817 plate spectra that contain all Lyman-alpha forest absorptions in its corresponding 3.1 deg^2 cone in the sky. These data were binned into 512 bins of size 2.92 Mpc creating a matter density map spanning 1,500 Mpc. This bin size was chosen as a compromise between higher resolution (small bin size) and fewer empty bins (large bin size).

While combining the spectra of a plate into one larger spectrum is far more efficient for computational purposes, this often leads to gaps of no flux data. This is not due to a physical absence of matter but from an incomplete data set. To account for this, all plates with over 10% of bins empty were discarded. However for smaller gaps of 10 or less empty bins an algorithm was derived that interpolated the missing data using a weighted average of data mirrored from both sides of the gap. The algorithm used is

$$(3) \qquad F_x = \frac{(2b - 2x - 1)F_{2a-x+1}}{2(b - a - 1)}$$
$$+ \frac{(2x - 2a - 1)F_{2b-x-1}}{2(b - a - 1)}$$

where F is the flux of a given unknown bin x, and a and b are the positions of the nearest known bins on the left and right respectively. When used on simulated data over gaps of 5 to 10 bins the algorithm averaged slightly under 20% error and retained the general features of the missing data. Figure 3 shows the results of the gap-filling algorithm graphed in blue on one of the simulated flux spectra graphed in green. In total, 76 plates were discarded during gap-filling leaving 741 plate spectra to be analyzed.

The continuous wavelet transform was then applied to each plate spectrum using a Morlet wavelet at 32 distinct scales. For each plate, the
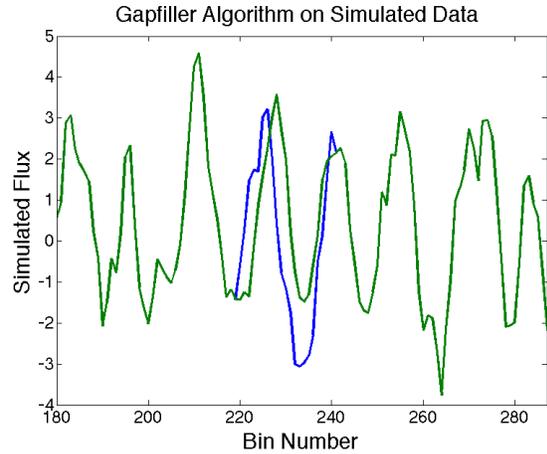

FIGURE 3. This shows the reconstruction of a simulated flux density sample using the gapfiller algorithm. The simulated data plotted in green had 20 bins emptied and then interpolated by the gapfiller which is plotted in blue. While only an approximation, the interpolated data is a more accurate reflection of the missing matter density.

mean wavelet coefficient at each scale was calculated and the power spectrum was found by taking the square of these averaged coefficients. Finally, the total matter power spectrum was calculated by averaging each of the 741 individual plate power spectra weighted by the variance of the wavelet coefficients at each scale. The MATLAB computing environment was used for all computational work in this research[22].

4. RESULTS AND DISCUSSION

The combined power spectrum of all plates is shown in Figure 4 where the matter density variance, or the difference between local matter density and the mean matter density, is graphed versus scale in units of $\log_{10}$ Mpc^{-1}. Each error bar represents the variance of the power spectral density across all 741 plates.

— 5 —

The matter power spectrum calculated with the new methodology has six times the resolution of the power spectrum calculated in [8] with the discrete wavelet transform. This is critical, as the previous power spectrum failed to detect evidence of BAOs due to low resolution at the scales of the expected sound horizon. Not only does this power spectrum have the expected size and shape from Eisenstein and others who used Fourier techniques ([10], [11], [12]), but a distinct peak in amplitude is detected near -1.7 $\log_{10}$ Mpc^{-1}, enlarged in the inset, that corresponds to the sound horizon at 160 Mpc. While the peak appears to be small, it shows a statistically significant increase in amplitude that would be absent in the power spectrum if BAOs were not present in the early universe. As the acoustic oscillations are a direct prediction of the Λ-CDM model of cosmology, this research further corroborates a universe dominated by dark energy and cold dark matter.

While the described feature does suggest the existence of BAOs, its amplitude is likely diminished by the large Δz range used as the BAO imprint undoubtedly changed over the nearly two billion year period of structure evolution. A better method would be to repeat the presented analysis in much smaller Δz ranges and track how the power spectrum evolves with redshift (or equivalently, time) from $z = 2.2$ to $z = 3.6$. Another potential source of error is the combination of all spectra in a given plate into a single larger spectrum. This approximation gets progressively worse at larger redshifts as the data is more susceptible to redshift distortions. Either the spectral data should be weighted by its distance from Earth or the analysis needs to be applied to all individual spectra.

5. CONCLUSION

The authors presented an improved methodology for detecting the primordial baryonic acoustic oscillations using spectral data from the DR9 BOSS component of the SDSS and the continuous wavelet transform. The appearance of the BAO peak in the matter power spectrum is further confirmation of the Λ-CDM concordance model of cosmology and encourages further analysis.
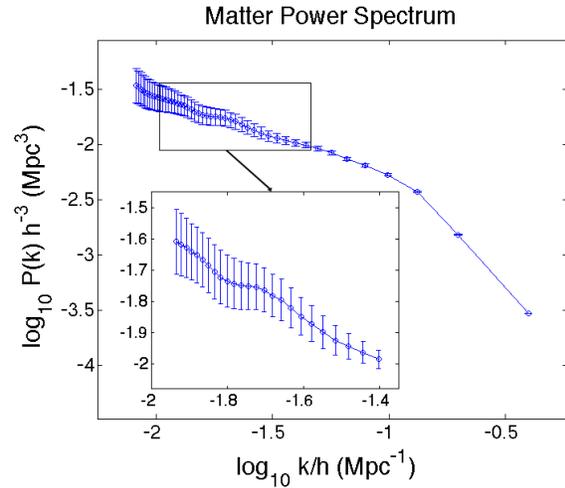

FIGURE 4. The mean matter power spectrum of all 741 plate spectra combined, weighted by the variance of the wavelet coefficients of each plate at each scale. The error bars are equal to the variance of the power across all plates at each scale. The peak enlarged in the inset at -1.7 $\log_{10}$ Mpc^{-1} is consistent with findings by Percival *et al.* in 2006 [23] and detects the sound horizon at 160 Mpc. This is direct evidence of BAOs and supports a universe based on the Λ-CDM model.

In particular, the authors hope to extend this methodology to track the evolution of cosmological parameters over time and calculate the correlation function from the matter power spectrum.

References

[1] Guth, A. 1981, Phys. Rev. D 23, 347.

[2] Peebles, P., ApJ, 153,1.

[3] Hu, W., Sugiyama, N., Silk, J., 1997, Nature, 386, 37-43.

[4] Hinshaw, G., Larson, D., Komatsu E., Spergel, D., Bennett, C., Dunkley, J., Nolta, M., Halpern, M., Hill, R., Odegard, N., Page, L., Smith, K., Weiland, J., Gold, B., Jarosik, N., Kogut, A., Limon, M., Meyer, S., Tucker, G., Wollack, E., Wright, E. 2013, ApJS, 208, 19.

[5] Planck Collaboration XVI. 2013, Astronomy and Astrophysics, 1303, 507, 3.

[6] Ahn, C., et al., 2012, ApJ, S, 203, 21.

[7] Stoughton,C., Astronomical Journal, 2002, 123, 485548.

[8] Everett, S., Kwiecien, M., Newmiller, C. 2014, DePaul Discoveries: Vol. 3, Iss. 1, Article 7.

[9] Peebles, P., Yu, J. 1970, ApJ, 162, 815-836.

[10] Eisenstein, D., Zehavi, I., Hogg, D., Scoccimarro, R., Blanton, M., Nichol, R., Scranton, R., Seo, H., Tegmark, M., Zheng, Z., Anderson, S., Annis, J., Bahcall, N., Brinkmann, J., Burles, S., Castander, F., Connolly, A., Csabai, I., Doi, M., Fukugita, M., Frieman, J., Glazebrook, K., Gunn, J., Hendry, J., Hennessy, G., Ivezi, Z., Kent, S., Knapp, G., Lin, H., Loh, Y., Lupton, R., Margon, B., McKay, T., Meiksin, A., Munn, J., Pope, A., Richmond, W., Schlegel, D., Schneider, D., Shimasaku, K., Stoughton, C., Strauss, M., SubbaRao, M., Szalay, A., Szapudi, I., Tucker, D., Yanny, B., York, D. 2005 ApJ, 501, 171, 11.

[11] Martinez, V., Arnalte-Mur, P., Saar, E., de la Cruz, P., Pons-Bordera, M., Paredes, S., Fernandez-Soto, A., Tempel, E. 2009, ApJ, 696, L93.

[12] Labatie, A., Starck, J., Lachize-Rey, M. 2012, ApJ, 746, 172.

[13] Branch, D., Tammann, G. 1992, Annual Review of Astronomy and Astrophysics, 30, 359-89.

[14] Liddle, A. An Introduction to Modern Cosmology (2nd ed.). London: Wiley, 2003.

[15] Weinberg, D., Dav'e, R., Katz, N., Kollmeier J. 2003, AIP Conf. Proc. 666, 157-169.

[16] Percival, W., White, M., 2009, MNRAS, 393: 297-308.

[17] Magira, H., Jing, Y., Suto, Y. 2000, ApJ, 528, 30-50.

[18] Stankovic, R., Falkowski, B., (2003), 'The Haar wavelet transform: its status and achievements.', Computers & Electrical Engineering 29 (1), 25-44.

[19] Kirby, J., Swain, C. Geophys. J. Int. (2013) 194 (1): 78-99.

[20] Lee, K., Bailey, S., Bartsch, L., Carithers, W., Dawson, K., Kirkby, D., Lundgren, B., Margala, D., Palanque-Delabrouille, N., Pieri, M., Schlegel, D., Weinberg, D., Yeche, C., Aubourg, E., Bautista, J., Bizyaev, D., Blomqvist, M., Bolton, A., Borde, A., Brewington, H., Busca, N., Croft, R., Delubac, T., Ebelke, G., Eisenstein, D., Font-Ribera, A., Ge, A., Hamilton, J., Hennawi, J., Ho, S., Honscheid, K., Le Goff, J., Malanushenko, E., Malanushenko, V., Miralda-Escud, J., Myers, E., Noterdaeme, P., Oravetz, D., Pan, K., Paris, I., Petitjean, P., Rich, J., Rollinde, E., Ross, N., Rossi, G., Schneider, D., Simmons, A., Snedden, S., Slosar, A., Spergel, D., Suzuki, N., Viel, M., Weaver, B. 2012, ApJ, 1211, 5146, 2.

[21] Peacock, A. 1999, Cosmological Physics, Cambridge University Press, Cambridge.

[22] MATLAB 8.5, The MathWorks Inc., Natick, MA, 2000.

[23] Percival, W., Nichol, R., Eisenstein, D., Frieman, J., Fukugita, M., Loveday, J., Pope, A., Schneider, D., Szalay, A., Tegmark, M., Vogeley, M., Weinberg, D., Zehavi, I., Bahcall, N., Brinkmann, J., Connolly, A., Meiksin, A. 2007, ApJ, 657, 645-663.

— 7 —