DePaul University
UNIVERSITY LIBRARIES

DePaul Discoveries

Volume 3 | Issue 1 Article 7

2014

Cosmic Sound in the Lyman Alpha Forest

Spencer Everett
spencerweverett@gmail.com

Matt Kwiecien
mkwieci2@mail.depaul.edu

Cordell Newmiller
cjnewmil@gmail.com

Cosmic Sound in the Lyman Alpha Forest

Acknowledgements

Advisor: Dr. Jesus Pando, DePaul Physics Department Research conducted in summer of 2013 Spencer Everett: spencerweverett@gmail.com Matt Kwiecien: mkwieci2@mail.depaul.edu Cordell Newmiller: cjnewmil@gmail.com

Cosmic Sound in the Lyman Alpha Forest

Spencer Everett, Matt Kwicien, and Cordell Newmiller
Department of Physics

ABSTRACT Using the Baryon Oscillation Spectroscopic Survey (BOSS) from the Sloan Digital Sky Survey (SDSS), the authors attempt to detect the baryonic acoustic oscillations (BAOs) using the discrete wavelet transform. The wavelet transform is used to construct the power spectrum of intergalactic clouds of matter at large (Mpc) distance scales. It was found that the wavelet transform used here does not have high enough resolution to detect the BAOs. However, the techniques used in this study allow for future improvements in the transform that could potentially resolve the expected peak in the power spectrum and indicate the existence of BAOs.

1. Introduction

380,000 years after the big bang, the universe was not completely homogeneous [1]. This can be seen as small anisotropic variations in the cosmic microwave background (CMB), or the imprint left when photons decoupled from matter at that time [2]. This decoupling event, known as recombination, left a distribution imprint on the photons comprising the CMB that traced the matter it decoupled from. Standard cosmological models hold that these anisotropies were initially present in dark matter shortly after the Big Bang [3]. Dark matter is different from the normal matter of everyday experience in that it only interacts with normal matter, called baryonic matter, through the gravitational force. While elusive to detect, the existence of dark matter is inferred from how baryonic matter moves through space as if there is an unseen source of gravity. Recent evidence suggests that 85% of all matter in the universe is dark rather than baryonic [4].

Faculty Advisor: Dr. Jesus Pando, Department of Physics
Research completed Summer 2013
Author contact: spencerweverett@gmail.com, mkwieci2@mail.depaul.edu, cjnewmil@gmail.com

These models also predict that the anisotropies gave rise to certain pressure oscillations between light and matter in the early universe; waves that were frozen in place when recombination occurred. Because these are pressure waves, the waves are analogous to sound. Thus the imprints they left on the distribution of matter in the universe are called baryonic acoustic oscillations (BAOs). The exact form, amplitude, and scale of these oscillations provide key clues to the fundamental makeup of the universe. Eisenstein *et al.* first conclusively observed baryonic acoustic oscillations a decade ago [5]. The oscillations were detected at a scale of about 100 mega-parsecs (Mpc), which is thousands of times the size scale of the Milky Way. The research goal was to corroborate Eisenstein's findings using publicly available data, simple discrete algorithms, and affordable computational resources.

2. The Baryon Oscillation Spectroscopic Survey

Much of the baryonic matter in the universe is in the form of massive intergalactic clouds of hydrogen. Most of these clouds are remnants of the first neutral hydrogen atoms that formed once the universe had cooled enough for neutral atoms to form. This can be assumed as the clouds have very few traces of the heavier elements that are formed after stars are born and die. These clouds are believed to be held in place by the underlying dark matter, and thus serve to trace the matter distribution in the universe. These clouds of hydrogen are observable because very bright and distant galaxies, called quasars, illuminate them in a measurable way [6].

As shown in figures 1 and 2, the mechanism by which this works is that high-energy light from an active quasar excites the electron in a hydrogen atom. A specific wavelength is now missing, or absorbed, from the observed light of the quasar. As only light of a specific wavelength is absorbed by different kinds of molecular species, these objects can be directly identified as neutral hydrogen [7].

While absorption spectra allow the identification of the hydrogen clouds, their distance away also needs to be determined. Similar to how the pitch of a siren is distorted as an ambulance moves towards and away from an observer, the wavelength of light is red shifted when its source is moving away from the earth. The magnitude of the shift is related to the distance of the source from the observer. Therefore, the cosmological redshift can be used to determine the position of these clouds along the one-dimensional line of sight between the distant quasar and the earth. More practically, only the most prominent hydrogen absorption line, Lyman-alpha, is used. The term Lyman-alpha forest refers to this collection of individual Lyman-alpha lines with various redshifts seen in a quasar spectrum [8].

The data used were from the Sloan Digital Sky Survey (SDSS). SDSS uses a 2.5m telescope located at Apache Point Observatory in New Mexico to observe astronomical phenomenon in the northern sky. SDSS records the position in the sky of an object along with its spectrum. The spectrum is used to both identify what the object is and, using the red shift, to determine the distance of the object. The ninth data release of the Sloan Digital Sky Survey (SDSS) included a component called the Baryon Oscillation Spectroscopic Survey (BOSS). BOSS is a compilation of the spectra of over 50,000 distant quasars. The quasars are grouped together in about 800 plates,

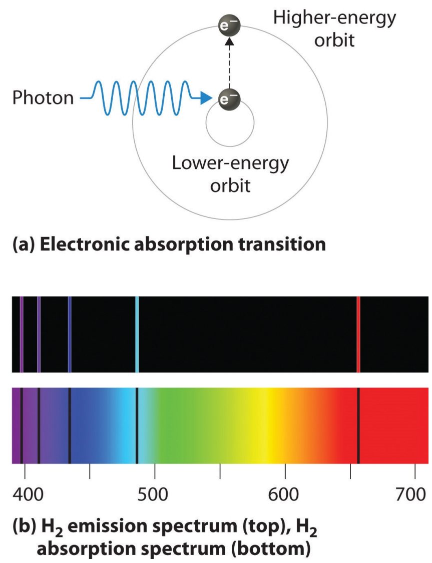

Figure 1. A representation of photon absorption in a hydrogen atom [9]. Part (a) of the figure shows a photon in blue intercepting an electron in its lowest energy state, called its ground state. The energy of the photon is absorbed by the electron and it moves to a higher energy state. These different energy states occur at specific discrete values, and so only photons with the same energy as the difference between two energy states will be absorbed. Photons of a particular energy have a corresponding wavelength, and these wavelengths will be missing in the absorption spectrum shown in part (b) of the figure.

where one plate covers a sector of sky approximately 3 astronomical degrees in diameter. The publicly available data supplied by BOSS is raw and only marginally processed, but the release does include flags on questionably reliable data points, as well as a paper describing how to normalize and clean the data of background noise [6].

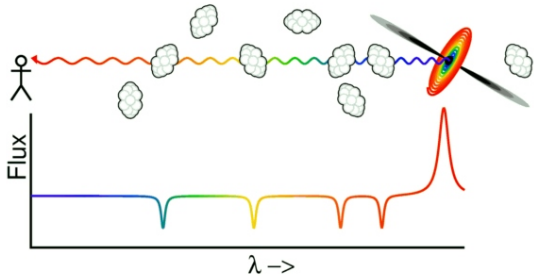

FIGURE 2. A simple representation of the Lyman-alpha forest [10]. The large object at the top right is a quasar, which is emitting light towards an observer on the far left. Light passing through hydrogen clouds causes a decrease in flux at the Lyman-alpha wavelength due to photon absorption, which is shown on the juxtaposed graph. The dips in flux occur at different wavelengths as the light is redshifted by different factors at different distances.

3. Technique

The MATLAB computing environment was used to analyze the data. The first steps were to import the data from BOSS and then to remove flagged data. Parts of the spectrum not near the Lyman-alpha forest are discarded as it does not pertain to the intended analysis, and the rest is cleaned and normalized using the function described in the BOSS paper [6]. Figure 3 shows an example of the raw spectrum of quasar 4498-55615-0410. The horizontal axis is the redshift while the vertical axis is the photon flux. Note that some flux is negative due to the flux of the quasar being subtracted away from the flux of the absorption clouds.

Each spectrum needed to be converted from redshift to distance. To do this, the distance-redshift relation was used. This is a general relationship between commoving distance and redshift, where commoving distance is defined such that it does not increase as the universe expands. At the large red shifts of quasars, it is safe to assume that the red shift is entirely cosmological in origin. The relationship is given by

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

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 pressureless mass (baryonic and dark matter), Ω_r is the density parameter of relativistic particles, and R_0 is radial distance from the observation point, which in this case is Earth [11]. The values used during analysis were $H_0 = 70$ $kms^{-1}Mpc^{-1}$, $\Omega = 1.02$, $\Omega_\Lambda = 0.73$, $\Omega_m = 0.27$, and $\Omega_r = 8.24 \times 10^{-5}$, which are well accepted values by the concordance model of cosmology [12].

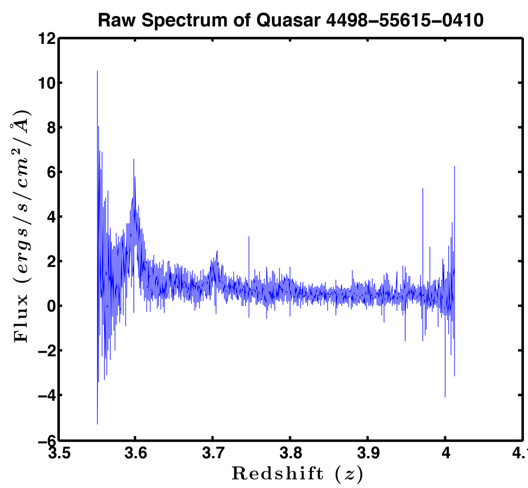

FIGURE 3. The raw spectrum of quasar 4498-55615-0410, one of over 50,000 quasar spectra supplied by the Sloan Digital Sky Survey (SDSS). This spectrum measures flux as a function of redshift and contains the Lyman-alpha Forest described in figure 1. These data were cleaned according to the procedures outlined in [6].

Published by Via Sapientiae, 2014

To reduce computational requirements, the individual Lyman-alpha forests were grouped together by plate number to form one long forest for each plate, recalling that each plate covers about 3 degrees of the sky. The data for each plate were binned into 512 bins, or, equivalently, a distance scale of 2.92 Mpc. If there were not enough quasars in a particular plate, there was a chance that a small gap of zeroed bins would be created in the combined spectra. Data with gaps that were too big, where many of the bins were empty, would distort correlation statistics of the whole data set and needed to be replaced in a way that would keep the correct statistics unaltered. For the case of a small number of empty bins an algorithm was derived to fill the gaps using a weighted average of data mirrored from both sides of the gap. The algorithm is

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

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. This algorithm weights the mirrored data so that any linear progression is conserved: applied to a straight line with a gap, the line will be reconstructed perfectly. Figure 4 shows the results of the algorithm on a generated data set, where the green line is the original data and the blue line shows the reconstruction made by the algorithm. An interval of data was purposely zeroed in this known set and the gap filler algorithm approximated the missing data. Comparing the results with known data sets, it was found that, for small gaps, the algorithm performed well with typically less than a 20% error.

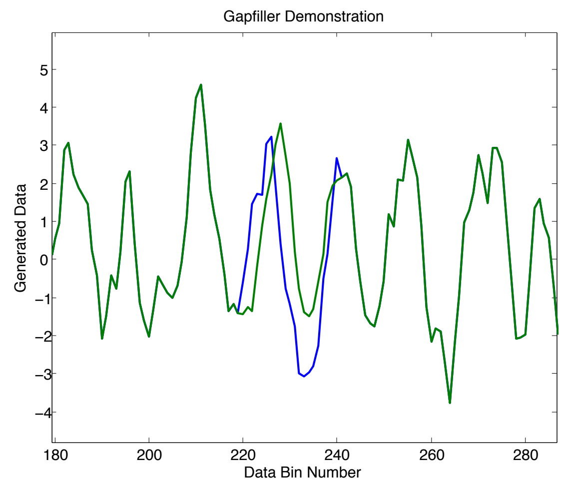

FIGURE 4. This graph shows a reconstruction of a data set using the gap filler algorithm to fill zeroed bins of data. A generated data set with known power spectrum is graphed in green. An interval of data was removed from the data set, and the algorithm used data on either side of the gap to approximate the missing data which is graphed in blue. While the reconstruction is only an approximation, it retains the statistics of the entire data set far better than leaving the bins zeroed.

After the pre-processing, the two-point correlation function of the data needed to be calculated. The correlation function is a measure of the excess probability of encountering another object at some distance compared to what would be expected from a random probability. This is useful in detecting BAOs because their effects on the early universe can be seen in the matter distribution in the universe today. Thus if the level of structure in the universe can be correlated as a function of distance, a bump of high correlation would indicate the existence of BAOs [5]. However, this function is computationally intensive to compute directly. The power spectrum, the Fourier transform of the correlation function, can be computed much more quickly. Because of this ease of computation, the power spectrum is the quantity predicted directly by theories of structure formation and it is the quantity measured in this work.

To calculate the power spectrum a mathematical technique known as a discrete wavelet transform was used [13]. The transform breaks a signal down into localized averages and localized fluctuations about those averages. After the transformation, two signals emerge. The first is a signal that is representative of the original signal with the high frequency noise smoothed away, while the other is a signal that captures the local fluctuations. The latter signal is effectively a measure of the variance. The coarse grained signal can then be passed again to the transform, which again results in to two new signals. In this way, the original distribution can be analyzed on a scale-by-scale basis and the variance as a function of scale is obtained. This variance is the power spectrum.

The wavelet transform also effectively compresses the data into more manageable data sets without loss of information. BAOs are a localized phenomenon, so using a transform that preserves localized information is advantageous. However one drawback of the wavelet transform used in this research is that it decomposes the data in powers of two, which reduces the resolution at large separation distances. Even with this optimized transform and a computation of the power spectrum instead of the correlation directly, the computational analysis took many days to complete.

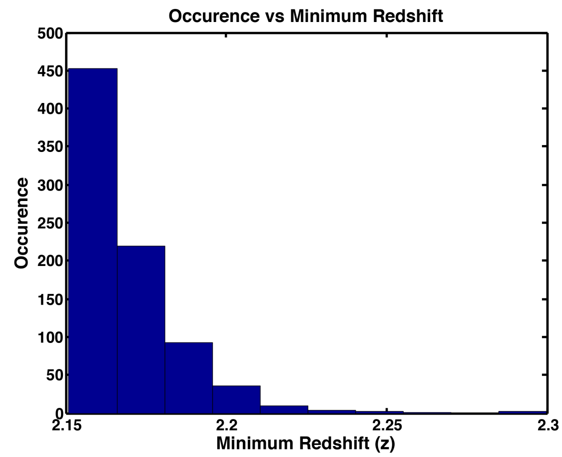

FIGURE 5. A histogram of the minimum redshift of each combined spectra plate. Finding a common minimum redshift value confined the analysis to distances where the most data were present and required less use of the gap filler algorithm.

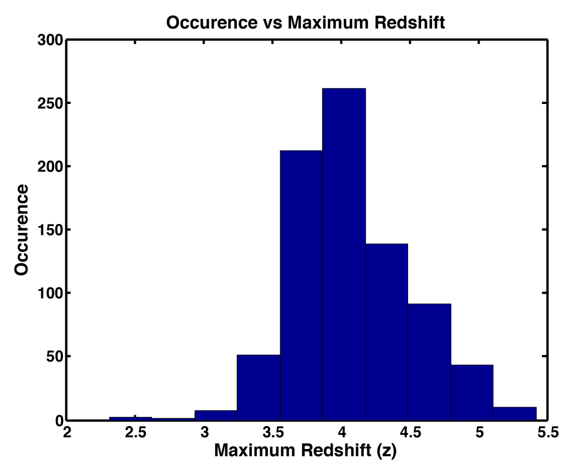

FIGURE 6. A histogram of the maximum redshift of each combined spectra plate. Finding a common maximum redshift value confined the analysis to distances where the most data were present and required less use of the gap filler algorithm.

4. Results

The first result was two histograms of the maximum and minimum redshift values among the entire data set, which are given in figures 5 and 6 respectively. By bounding the data with a minimum and maximum redshift value that was common to most spectra, fewer zeroed bins were created and less use of the gap-filler was required.

After data cleaning and applying the discrete wavelet transform to the BOSS data, the average power spectrum of the 800 plates was found at eight separation distances: 12, 23, 47, 94, 188, 375, 750, and 1,500 Mpc. These data are shown in figure 7. The error bars are one standard deviation from the mean of the variance computed at each scale. Overall, the power spectrum that was computed, as shown in figure 7, has the generally accepted features as predicted by theory [5].

However, this trial only had a tenth of the resolution that Eisenstein *et al.* had and is not sufficient to capture the BAO feature in the power spectrum. The resolution used by Eisenstein *et al.* was close to the minimum required to resolve the BAOs, and they would have manifested as a slight bump in correlation at slightly over 100 Mpc of separation with a width of about 20 Mpc. The correlation points in figure 7, nearest of which are at 47, 94, and 188 Mpc, do not preclude the existence of the bump, but clearly cannot indicate it. While the measured scales do not capture the effects of BAOs, the results are still promising as they closely match the expected theoretical power spectrum for the same scales. This makes the authors optimistic that the techniques used here can be adapted to use a modified transform with higher resolution to observe the existence of BAOs on the required scale.

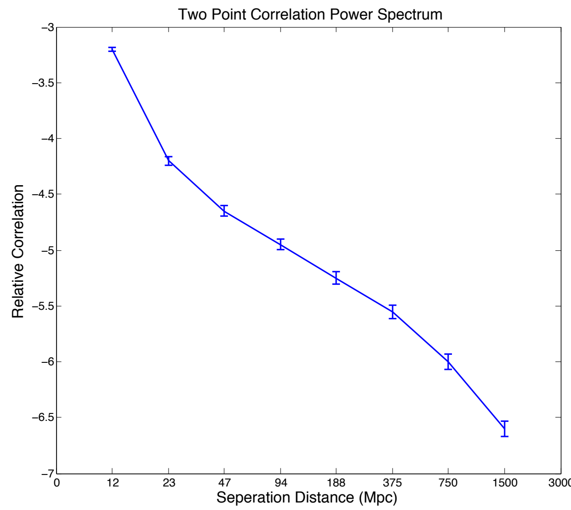

FIGURE 7. The power spectrum of all combined spectra plates. The power spectrum was graphed as a function of separation distance, with error bars of one standard deviation from the mean of the variance at each scale. A high relative amplitude at a particular separation distance signifies that the hydrogen clouds are more likely to be found at this separation than at a separation of lower relative amplitude. Previous results from Eisenstein *et al.* suggested that there should be a noticeable increase in amplitude at slightly over 100 Mpc due to BAOs [5]. The BAO peak is not evident in the processed data due to the low resolution of the discrete wavelet transform.

5. Conclusion

The techniques described in this research did yield the power spectrum of the Lyman-alpha forest from the SDSS BOSS survey. However, the resolution was inadequate to confirm the existence of BAOs at the expected separation distance. If a different wavelet transform is used that can refine the resolution at which the distribution is decomposed, this technique may be capable of detecting signatures of the BAOs. The wavelet has several advantages over traditional techniques that calculate the power spectrum, so refining the technique is well worth the effort. As it stands, the conclusion is that the power-of-two discrete transform used in this work does not have the required resolution to detect the BAOs. While no evidence of BAOs was detected, a computational foundation for future research to build upon was created. The gap-filling algorithm is a particularly versatile accomplishment and should be applicable to any almost-complete data set that needs statistical parameters to be preserved.

Future research will include applying other wavelet transforms to this data that have better resolution. In particular, the use of wavelet packets that have much higher scale resolution will be tried. As the foundation has been laid with this work, the extension to wavelet packets is greatly simplified. The authors also hope to use the data to generate a three-dimensional model of the locations of the hydrogen clouds in the regions examined by the SDSS.

6. Acknowledgments

The authors would like to thank the Sloan Digital Sky Survey for the public availability of their data, the DePaul College of Science and Health for providing funding for this research, and to Dr. Jesus Pando for his invaluable guidance and assistance throughout the project.

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://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 Extraterrestial Physics, New Mexico State University, New York University, Ohio State University, Pennsylvania State University, University of Portsmouth, Princeton University, the Spanish Participation Group, University of Tokyo, University of Utah, Vanderbilt University, University of Virginia University of Washington, and Yale University.

References

[1] Jones, M. H., & Lambourne, R. J. (2004). Big Bang Cosmology - The Evolving Universe. An introduction to galaxies and cosmology (). Cambridge: Open University.

[2] Bennett, C. L., Halpern, M., Hinshaw, G., Jarosik, N., Kogut, A., Limon, M., Weiland, J. L. (2003). First-year Wilkinson Microwave Anisotropy Probe (WMAP) observations: preliminary maps and basic results. The Astrophysical Journal Supplement Series, 148(1), 1.

[3] Ostriker, J. P., & Steinhardt, P. J. (1995). Cosmic concordance. arXiv preprint astro-ph/9505066.

[4] Ade, P. A. R., Aghanim, N., Armitage-Caplan, C., Arnaud, M., Ashdown, M., Atrio-Barandela, F., Davies, R. D. (2013). Planck 2013 results. XVI. Cosmological parameters. arXiv preprint arXiv:1303.5076.

[5] D. J. Eisenstein, I. Zehavi, D. W. Hogg, R. Scoccimarro, M. R. Blanton, R. C. Nichol, R. Scranton, H. Seo, M. Tegmark, Z. Zheng, S. Anderson, J. Annis, N. Bahcall, J. Brinkmann, S. Burles, F. J. Castander, A. Connolly, I. Csabai, M. Doi, M. Fukugita, J. A. Frieman, K. Glazebrook, J. E. Gunn, J. S. Hendry, G. Hennessy, Z. Ivezic, S. Kent, G. R. Knapp, H. Lin, Y. Loh, R. H. Lupton, B. Margon, T. McKay, A. Meiksin, J. A. Munn, A. Pope, M. Richmond, D. Schlegel, D. Schneider, K. Shimasaku, C. Stoughton, M. Strauss, M. SubbaRao, A. S. Szalay, I. Szapudi, D. Tucker, B. Yanny, D. York, 2005 Detection of the baryon acoustic peak in the large-scale correlation function of SDSS luminous red galaxies, Astrophys. J. 633 560.

[6] K.G. Lee, S. Bailey, L.E. Bartsch, W. Carithers, K. S. Dawson, D. Kirkby, B. Lundgren, D. Margala, N. Palanque-Delabrouille, M.M. Pieri, D.J. Schlegel, D.H. Weinberg, C. Yeche, E. Aubourg, J. Bautista, D. Bizyaev, M. Blomqvist, A.S. Bolton, A. Borde, H. Brewington, N.G. Busca, R.C. Croft, T. Delubac, G. Ebelke, D.J. Eisenstein, A. Font-Ribera, J. Ge, J.C. Hamilton, J.F. Hennawi, S. Ho, K. Honscheid, J.M. Le Goff, E. Malanushenko, V. Malanushenko, J. Miralda-Escude, A.D. Myers, P. Noterdaeme, D. Oravetz, K. Pan, I. Paris, P. Petitjean, J. Rich, E. Rollinde, N.P. Ross, G. Rossi, D.P. Schneider, A. Simmons, S. Snedden, A. Slosar, D.N. Spergel, N. Suzuki, M. Viel, B.A. Weaver. 2012 The BOSS Lyman- Forest Sample from SDSS Data Release 9.

[7] Averill, B., Eldredge, P. (2013). General chemistry: Principles, patterns, and applications. (1st ed., Vol. 1.0, Chapter 6). Flat World knowledge, Inc. Retrieved from http://catalog.flatworldknowledge.com/bookhub

[8] Ahn, C. P., Alexandroff, R., Prieto, C. A., Anderson, S. F., Anderton, T., Andrews, B. H., Edmondson, E. M. (2012). The ninth data release of the Sloan Digital Sky Survey: first spectroscopic data from the SDSS-III Baryon Oscillation Spectroscopic Survey. The Astrophysical Journal Supplement Series, 203(2), 21.

[9] Averill, B., Eldredge, P. (2013). General chemistry: Principles, patterns, and applications. (1st ed., Vol. 1.0, Chapter 6). Flat World knowledge, Inc. Retrieved from http://catalog.flatworldknowledge.com/

[10] http://www.astro.ucla.edu/wright/Lyman-alpha-forest.html

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

[12] Lange , A.E, Ade, P.R., Block, J.J., Bond, J.R., Borrill, J., Boscaleri, A., Coble, K., Crill, B.P., de Bernardis, P., Farese, P., Ferreira, P., Ganga, K., Giacometti, M., Hivon, E., Hristov, V.V., Lacoangeli, A., Jaffe, A.H., Martinis, L., Masi, S., Mauskopf,

P.D., Melchiorri, A., Montroy, T., Netterfield, C.B., Pascale, E., Piacentini, F., Pogosyan, D., Prunet, S., Rao, S., Romeo, G., Ruhl, J.E., Scaramuzzi, F., Sforna, D. "Cosmological Parameters from the first results of Boomerang." Physical Review D 63. January 2001 (2001): 1-8.

[13] Daubechies, Ingrid,1992, Ten Lectures on Wavelets, Society for Industrial and Applied Mathematics, Philadelphia.