Finding the hidden signature of the Baryonic Acoustic Oscillations in the Galaxy Temperature distribution

Pankaj Meghani

Podar International High School (Santacruz), Mumbai, India

**Summary (Abstract)**

Baryonic Acoustic Oscillations have been known to show their effect on the clustering of galaxies (1) and as temperature anisotropies in the cosmic microwave background (2). This study aims to see whether this clustering can also be found in the mass distribution of the galaxies. Five hundred galaxies between redshifts 0.45 and 1.0 and between -90 to 90, ra and dec taken from the Sloan Digital Sky Survey DR15 (3) are evaluated using three separate procedures (galaxy clustering, temperature distribution, and temperature difference distribution) in a Jupyter notebook (Python3). Then, galaxy distribution data was estimated from (4) to generate 500 galaxies, which were then analyzed in the same manner. The two were then compared. It is hypothesized that the mass distribution in the galaxies would be affected due to dissimilarities between the concentration of observable matter, dark matter, and radiation pressure in the primordial universe. The results support this hypothesis.

**Introduction**

After the big bang, the mean energy of the photons was stronger than the electrostatic attraction between electrons and the nucleus. This created a single tightly coupled fluid called the Photon-Baryon plasma through which sound waves (density fluctuations) traveled at around half the speed of light. This fluid flowed towards regions of higher density (due to random quantum fluctuations before inflation) and dragged the non-interacting dark matter along by its gravitational force. This fluid repeatedly underwent compression and rarefaction to create a sound wave of a radius ~115 Mpc with matter concentrated towards the center and the shell, the sound horizon. The density waves were frozen in place at recombination (decoupling of matter and light) leaving its signature in the spatial distribution of matter. Galaxies were then formed at a higher frequency at regions with higher density. Previous studies have looked at the clustering of galaxies (1), the temperature distribution of the Cosmic Microwave Background (a relic radiation field that we observe in all directions at a uniform temperature of 3 Kelvin) (2) , and its polarization. All such studies have been focused on finding the curvature of spacetime which has been often shown to be zero, withen error bounds. These methods are more prone to the effects of noise because of the motion of galaxies and dust particles swamp signals with B mode polarization on the map of the CMB (5).

This study aims to investigate whether the amount matter in galaxies is distributed according to their position in the density wave. The mass distribution of the galaxies can be used to constrain the percentage of mass composition of dark energy, cold matter, and Baryon matter as given by the simplified Lambda-CDM model modeled by:



The methods used in this study is extremely simillar to those used in the previous ones with the sole difference that instead of the galaxies postion (1), the galaxies temperatures are used. It is hypothesized that the histogram of variation of the temperature of each galaxy will contain two peaks corresponding to the shell and the center of the ripple while the histogram of the temperature difference between galaxy pairs will have one peak and one bump corresponding to the galaxies at the center and on the surface of the spheres. It is also theorized that this method is advantageous over galaxy distances as the position of the matter changes while the amount of matter remains relatively more constant as major galaxy merger events are rare and over CMB fluctuations as intergalactic gas can only have minor variations on the measured mass distribution.

**Results**

The galaxies are chosen to have the highest possible redshifts as in these galaxies, the effect from merger events would be the least. The most straightforward measure of the amount of matter would be the mass of the galaxy, but this is unsuitable for multiple reasons. As there are a fewer number of galaxies with redshift more than z=0.3, this eliminates the standard methods of gravitational lensing (bending of light around matter) and Virial theorem (the dirstribution of galaxies is neither stable nor self gravitating, both conditions required to estimate the kinetic energy). Because Sloan Digital Sky Survey does not have the measurements of the 21 cm HI line and other surveys do not measure this for galaxies at such redshifts, this method is also not feasible. This means that a different approach is needed to *estimate* the mass. The emission spectra of stars depend on the amount of gas available to it during formation; stars with a higher concentration of hydrogen are bluer than others per Plank's law. Also, as galaxies are clusters of stars, it can be assumed that the spectra of a galaxy are the same as the average spectra of its stars. This means that the spectra of a galaxy depend on the amount of matter available to it, which in turn depends on its position in the acoustic ripples.

The data is obtained from the Sloan Digital Sky Survey DR15, looking for galaxies whose redshift is higher than 0.4. The three graphs, as stated in the introduction, are plotted, and a curve of best fit is plotted through it; this curve is used for further analysis. It is seen that the temperature, and therefore, the mass of the galaxy distribution follows the same pattern as theorized. As for in the temperature difference graphs, the two principal peaks are visible, but surprisingly, a third less prominent peak is also seen at ~250 K. Following this, a number of galaxies were also generated multiple times. As it would not be feasible to store the plots for each of them, just the areas under each of the graphs were stored in the array which was averaged at the end. This result is used for further analysis.

The results, rounded to 4 significant figures, are shown in the given table.

|  |  |  |
| --- | --- | --- |
| **Features** | **Real Data** | **Generated Data** |
| **Temperature Distribution** | 0.2150 | 0.1176 |
| **Temperature Difference Separation** | 0.3437 | 0.4219 |
| **Galaxy Clustering Distribution** | 0.3082 | 0.6686 |

**Discussion**

The original question posed was whether the distribution of galaxy masses correspond to their positions on the sound wave. In figure 1, the temperature distribution of galaxies, it is seen that there are two peaks. This is expected as the area of the surface of the sphere is less than the area of its center, thereby creating a small region of high-temperature stars and a greater region of low-temperature stars. Therefore, there are a more significant number of low-temperature stars than high-temperature stars. This corresponds to M and K type stars, respectively. Note that the temperatures of galaxies containing these stars are higher than the actual temperatures of these stars. This can be attributed to the accretion disk around their central black-hole and the presence of other hotter stars which would be found in greater number than today as more amount of available hydrogen was available for stellar formation then.

The graph of galaxy temperature difference distribution (figure 2) is seen to have three significant 'bumps.' The first and the third bumps were predicted and expected, but the second bump has not been seen before. Upon further investigation, it is found that this peak corresponds to the temperature differences at locations where multiple shells overlap. From my research, this feature has not been seen for galaxy separation distribution because of the dynamic nature of a galaxy’s position, that is the bump is blurred out by the galaxy’s evolution. This supports the original hypothesis that this method offers significant advantages to other methods.

The galaxy distribution plots for the actual data (Figure 3) or the simulated data does not contain a peak at 115 Mpc. This can be explained by the dearth of galaxies available for comparison. To test this hypothesis, 5000 galaxies were created, and their galaxy distribution plots were formed. These galaxies show the pattern in them, which validates the hypothesis

Upon further inspection, it is seen that majorly galaxies discarded are of lower temperature. It is possible that this could create a bias in the galaxies that remain which could be reflected in the results. However, considering that the galaxy clustering of the remaining galaxies show that there is no difference, it is concluded that this does not affect the results significantly. This issue has not been investigated further though better curve fitting could mitigate this. Also, the spectra of the obtained galaxies could be affected by the opacity of the atmosphere for different wavelengths as the observatory of SDSS is in New Mexico. However, it was not found on the official website whether the data has been corrected for this, this issue is not addressed while programming. Another assumption that mush hold true for the analysis is that that the mass is proportional to the stellar temperature, which holds only for stars in the main sequence . As a galaxy contains all types of stars, in and out of the main series, these results hold true solely for an ideal state. Still, the number of stars in the main series outweighs stars that are not in the main sequence by a ratio of 90:10 which is sufficient for this analysis.

Another important note, the effects of Special Relativity are ignored because they do not have a noticeable impact on the observed frequency (the impact is on the order of  ). This could have some effect on the observed frequency which could be investigated in a future experiment, but because of the relatively low speeds, it remains doubtful if doing so would affect the results a little, if at all.

The correlation between the simulated and the actual galaxies is very high as the p-value for 0.01 confidence interval is  and the chi-square statistic of dependence is 1165.2967. This is important as it shows that the research does not contradict any past research. A high statistical significance is also important as this data can be used to constrain the proportions of matter in the universe further and help trace its evolution.

## **Materials and Methods**

The data is obtained from the Sloan Digital Sky Survey DR15, by looking for galaxies whose redshift is higher than 0.4 and storing their position, redshift, and intensity in the ugriz bands whose wavelengths are  respectively. The redshifts of each galaxy are then converted to distance from Earth using Astropy for Python. Following this, a black body curve is fitted throughout the bands to find the peak wavelength. All the galaxies of whose deviation of the data points from the curve are over the upper quartile are discarded as they do not fit the model well enough. This peak wavelength is corrected for redshift as given by , and for this wavelength, the temperature is calculated. The velocity of the galaxy is ignored as the change in frequency given by  . As the velocity of the galaxy is much lesser than the speed of light, effects of Special Relativity are negligible. Then, each of the galaxies is paired with every other galaxy, and their separation and temperature difference are found. From this, the area under the two graphs is calculated.

Following this, galaxies were generated with the centers of the spheres being random and galaxies around it following an approximate probability distribution function of distance from (1). Their temperature distribution was also modeled by this pdf. These 500 galaxies were run through 100 different generations. At the end of each generation, the same procedures were implemented, i.e., temperature histogram, temperature difference histogram, and galaxy distance plot and their areas calculated.

From this, the degree of correlation between the real data and galaxy simulation is found by calculating the Chi-Squared Correlation statistic.

**References**

1. Detection of the Baryon Acoustic Peak in the Large-Scale Correlation Function of SDSS Luminous Red Galaxies; Eisenstein et al. (2005);arXiv:astro-ph/0501171
2. Structure in the COBE Differential Microwave Radiometer First-Year Maps; Smoot et al. (1992); DOI: 10.1086/186504
3. The Fifteenth Data Release of the Sloan Digital Sky Surveys: First Release of MaNGA Derived Quantities, Data Visualization Tools and Stellar Library;Aguado et al. 2018;arXiv:1812.02759
4. On the Robustness of the Acoustic Scale in the Low-Redshift Clustering of Matter; Daniel J. Eisenstein, Hee-Jong Seo, Martin White; 17 Apr 2006;arXiv:astro-ph/0604361
5. BICEP2 I: Detection Of B-mode Polarization at Degree Angular Scales; Ade et al. 2014; [arXiv:1403.3985](https://arxiv.org/abs/1403.3985)

**Data Figures**

**Figure 1: Distribution of the real galaxy temperatures are found after curve fitting**

**Figure 2: Distribution of temperature difference between each pairs of galaxies. First peak: 50K. Second Peak (of interest): 250K. Third peak: 600K**

**Figure 3: Galaxy seperation distribution after calculating the difference in 3 dimensions.**