Excavating the Baryon Acoustic Oscillation Signature from Cosmic Voids

Ishani Kulkarni

Christina Kreisch

Princeton University, Department of Astrophysical Sciences

August 6, 2019

Abstract

The Baryon Acoustic Oscillation (BAO) signal has been found from the correlation function of halos. Since voids are formed from the clustering of halos, at large scales, we expect to find that the void-void correlation function is the same as the halo-halo correlation function. Therefore, we will try to figure out if we can isolate the BAO signal from the clustering pattern of voids with high significance. We also aim to find out if there is a resonant void size that emphasizes the BAO signal. In this paper, we were able to find that the BAO can be extracted from the void-void correlation function; however, we did not find it within high significance. We also found that the resonant void size is near 40 because that is the void cut that most enhanced the signal, though the statistics for this conclusion are weak.

Keywords: Baryon Acoustic Oscillation, Correlation Function, Halo, Cosmic Voids, Quijote simulations

1 Introduction

Baryon Acoustic Oscillations are the fluctuations in density of normal baryonic matter. In the early universe, photons were trapped by plasma of electrons and baryons. As the universe cooled, the plasma cooled to a low enough energy so that the electrons and protons in the plasma combined to form neutral hydrogen atoms. This recombination allowed photons to decouple from the normal matter and free-stream through the universe. After recombination, the baryons were frozen in a spherical shell around the initial excess density of dark matter. These regions of high-density matter gave rise to the galaxies in the universe today. By looking at points of high density in baryonic and dark matter, we can find the correlation of the distances between galaxies, producing the galaxy-galaxy correlation function. A correlation function shows the probability of finding a galaxy at a particular distance from a given galaxy. In simulations, this correlation can be found using halos, which are galaxies made of dark matter particles. Both the galaxy and halo correlation functions show a peak somewhere between 100 and 150 megaparsecs/h, known as the "BAO distance". Therefore, we expect to see a larger number of galaxy pairs separated by 100-150 megaparsecs. This means that for a given halo, the probability of finding another halo at the BAO distance is higher than it would be if the BAO signal was not there. This is seen in the halo correlation function in Figure 1. There is a significant bump in the difference function at the BAO distance.

Cosmic voids are under-dense regions of the universe created from the clustering pattern of galaxies. Since voids are defined as the underdense space inside surrounding halos, we should expect that at large scales, the void-void correlation function is the same as the halo-halo correlation function.

A 2015 paper[1] performed a similar experiment in which they extracted the BAO signal from cosmic voids. However, this paper's statistics were not entirely accurate because they included internal voids, which are voids inside other voids. This technique may result in double-counting data if not properly taken into account.

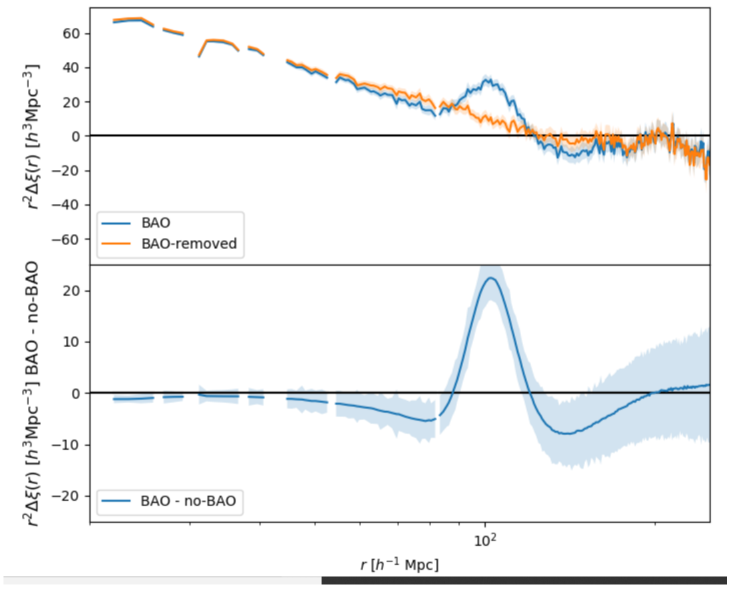

Figure 1: **Halo-Halo Correlation Function**

Top panel: Correlation function. *Bottom panel:* BAO - no-BAO Correlation Difference Function

2 Methodology:

To answer our research questions, we use the n-body simulations *Quijote*[2], which contain the halo data for the normal simulation as well as a "BAO removed" simulation. The "BAO removed" method removed the BAO perturbation when the simulation was initialized so that once galaxies form in the simulated universe, they did not trace

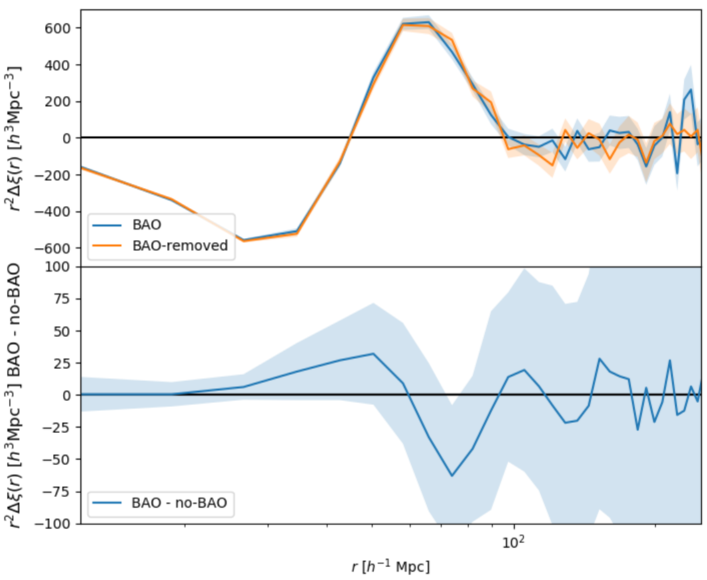

Figure 2: **Void- Void Correlation Function**

Top panel: Correlation function. *Bottom panel:* BAO - no-BAO Correlation Difference Function

any BAO signal. We subsampled both sets of data by randomly removing galaxies such that the resulting galaxy density of the simulation produces voids of a particular average size. For each sample, we created galaxy correlation functions to ensure our method was accurate. This confirms that the bump in the halo correlation function happens at the BAO distance. It also shows us if the halo correlation function shifted to the right or left, so that we can compare it to the void correlation function to see if the same shift occurred. We then applied the void finder *VIDE*[3] on each of these halo samples to produce a list of voids (denoted as a void catalog) and their properties, such as their positions and velocities. We subsampled the data in various ways to see how this impacts the extraction of the BAO signal. One method we used was stratified subsampling, where we used a smaller subset of the data that was distributed proportionally by mass according to the original dataset (as shown in Figure 3). We also performed mass cuts, where we took halos of masses greater or less than a certain mass. Furthermore, we did a radius cut on the original void catalog by taking voids of greater or less than a certain size (Figure 3).

Next, we analyzed the void catalogs from the void finder. We created the void correlation function, which tells us information about how voids cluster at a given distance. We also produced a void-halo cross-correlation function because seeing a

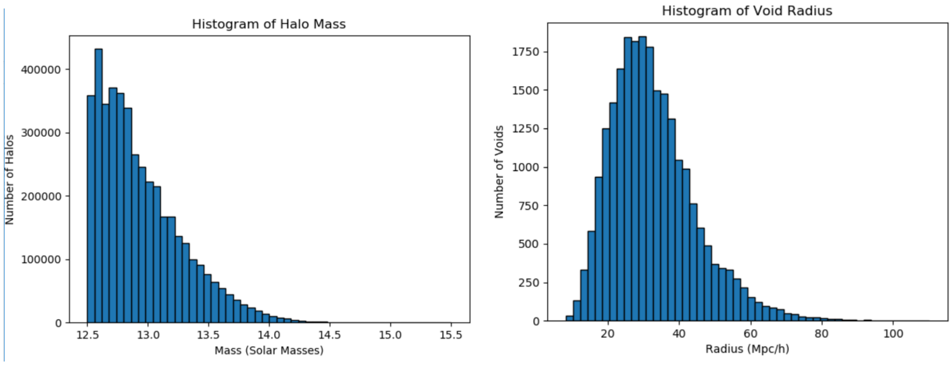

Figure 3: *left:*Histogram of Halo Masses (Solar Masses). *right:* Histogram of Void Radii (Mpc/h)

significant effect from the BAO in the cross-correlation function would confirm that voids are a useful tool in extracting the BAO signal. Finally, we used these analyses to create a conclusion and answer our research questions.

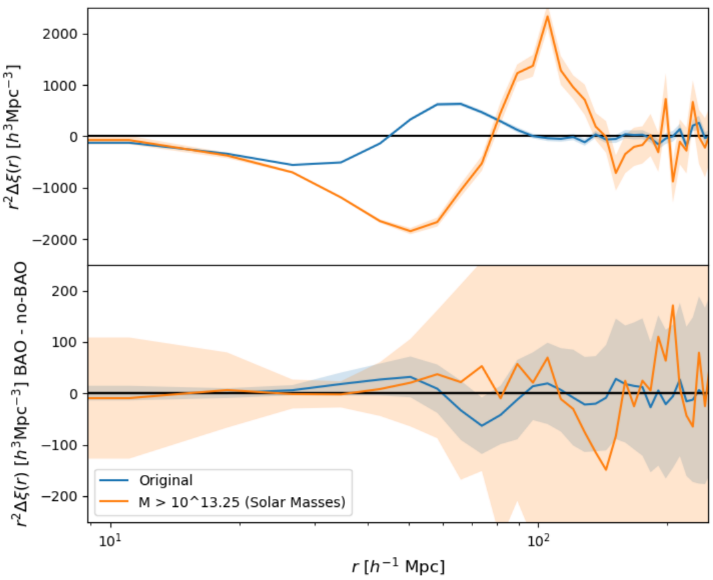

Figure 4: **Void- Void Correlation Function: Mass Cut** *Top panel:* Correlation function. *Bottom panel:* BAO - no-BAO Correlation Difference Function

3 Results:

Primarily, we found that we could extract the BAO signal from cosmic voids because when we subtracted the BAO correlation function by the removed-BAO correlation

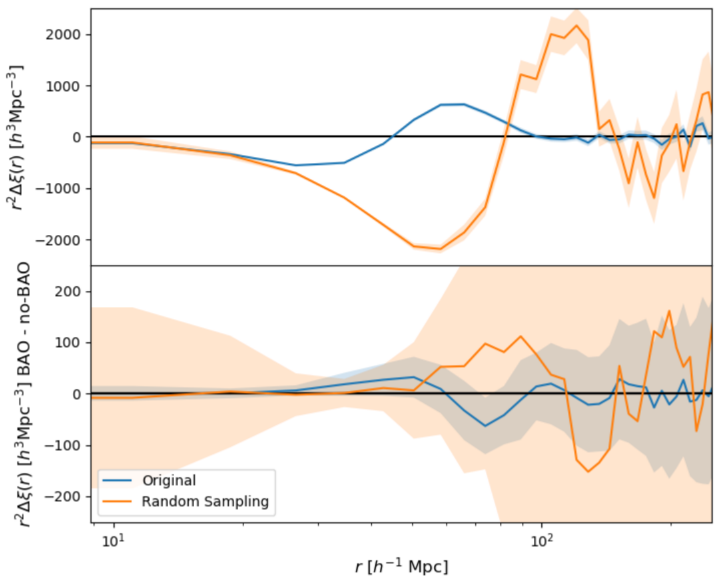

Figure 5: **Void- Void Correlation Function: Random Sampling**

Top panel: Correlation function. *Bottom panel:* BAO - no-BAO Correlation Difference Function

function, we saw a plot with wiggles and a bump at the BAO distance. However, we were unable to extract the signal with high significance, as shown by the large error bars. Therefore, we were unable to completely isolate the signal.

After plotting the difference of correlation functions, we noticed non-linear effects. In the original BAO - no-BAO void correlation function (as shown in Figure 2), the function has a small peak and trough in addition to the bump at the BAO distance. This is because in the normal void correlation functions, the non-BAO function overlaps the BAO function right after the main peak. These non-linear effects prove that voids are an interesting and useful tool in finding the BAO. The voids correlation function has its characteristic peak at the void exclusion scale, which is the location at which voids are most closely packed together. This occurs at 2 * average void radius because voids cannot overlap, and that is the closest distance two voids can be from each other.

Subsampling the data also affected the correlation function in various ways. Halo bias plays a large role in this. Halo bias refers to the fact that halos are biased tracers of dark matter, since dark matter exists outside of halos. It signifies that more massive halos will be found in higher-density regions. Therefore, if we do a mass cut by taking halos greater than a certain mass, when we run the void finder on them, we will only get the voids that are within the high-density regions or low-density regions. This would increase the average void size because it eliminates the smaller voids created

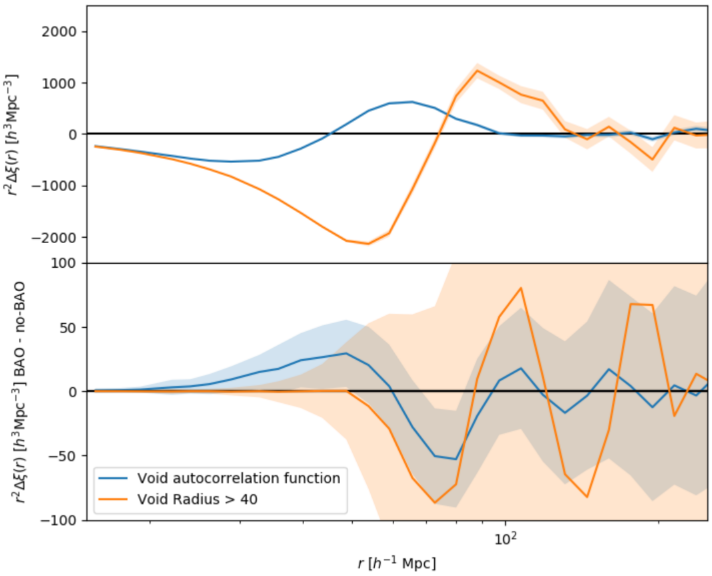

Figure 6: **Void- Void Correlation Function**

Top panel: Correlation function. *Bottom panel:* BAO - no-BAO Correlation Difference Function

from halos in low-density regions.

We decided to subsample by mass because although void densities vary, in general, mass correlates to void size, and sampling by void sizes affects the correlation function. When we did a mass cut by taking halos greater than a certain mass and then running the void finder on them, the peak shifted to the right and up (as seen in Figure 4). It shifted to the right because by removing halos, the average void size increased. Since the main peak occurs at the average void radius times 2, the correlation function peak shifted to the right. It shifted up because when we only take large halos, we are concentrating the halos in the high-density regions. This means that the average void size increases, so the void exclusion scale shifts to the right, where larger voids are more correlated.

Randomly subsampling the halo data gives us a smaller subset of the data that is distributed proportionally according to the masses of the original dataset. We found that randomly subsampling the data shifted the peak to the right and up (Figure 5). By removing many galaxies, the average void size increased, which shifted the peak to the right.

We also performed radius cuts on the void catalog from the original uncut halo data. When we took voids greater than a certain radius, the peak shifted to the right and up. When we took voids of only a smaller radius, the peak shifted to the left and up.

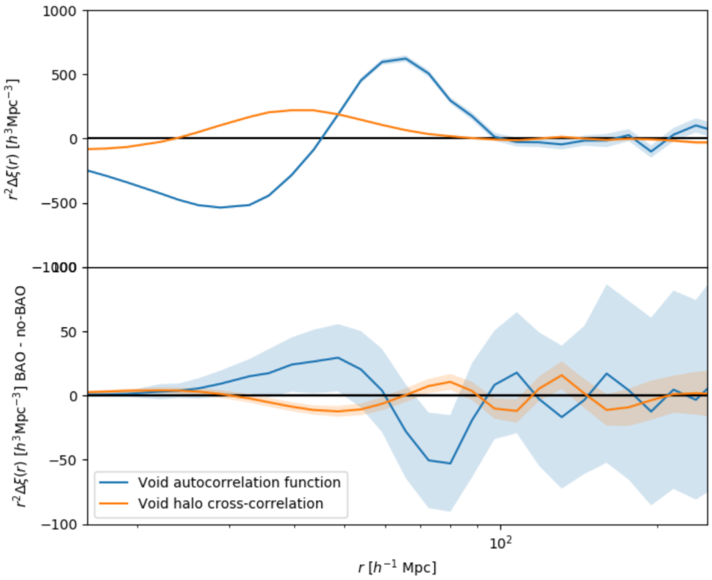

Figure 7: **Void- Halo Correlation Function**

Top panel: Correlation function. *Bottom panel:* BAO - no-BAO Correlation Difference Function

This is because the peak occurs at 2*radius, and radius cuts would increase or decrease the average radius.

Performing these same sampling techniques on the halo correlation function only affected the height of the peak. It did not shift the peak horizontally. Taking the halos of greater than a certain mass increased the correlation function because of halo bias: the larger halos will be clustered in high density regions of space, so there will be a higher correlation among larger halos. Taking the smaller halos decreased it because they tend to be in a low-density area and are more spread out.

The fraction of voids we used for the void cut (Figure 6) is 5529/24081, or about 23%. Because we used a small subset of the data, the plot was noisy and the statistics were weak.

As shown in the bottom panel in Figures 7 and 8, the void-halo cross-correlation difference functions have at least two peaks and three troughs. On the other hand, the halo correlation function (Figure 1) has only one peak and one trough. Having more wiggles shows that voids provide additional information in studying the BAO. The location of the peaks and troughs are shifted as compared to the voids correlation function because the peak was shifted to the left. The cross-correlation functions have significantly smaller error bars than the voids correlation function because halos have

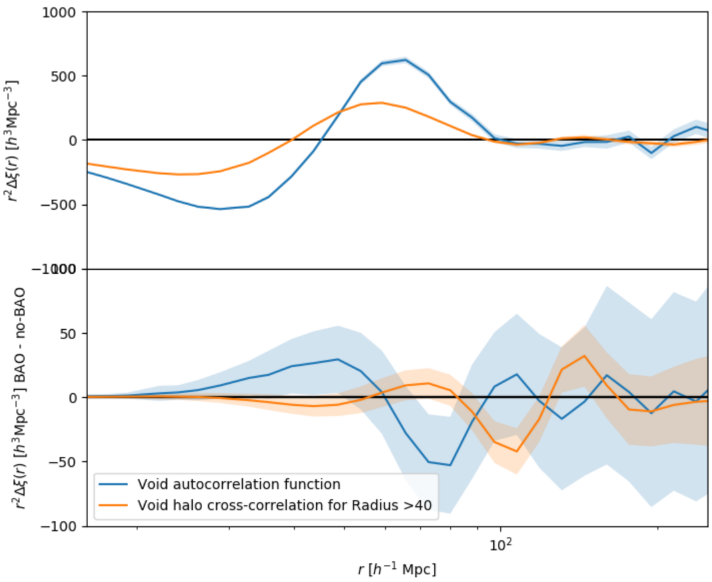

Figure 8: **Void- Halo Correlation Function Void Cut**

Top panel: Correlation function. *Bottom panel:* BAO - no-BAO Correlation Difference Function

better statistics (Figures 7 and 8).

4 Conclusion:

The BAO can be sufficiently proven by using a void-void correlation function. The bump in the correlation difference function occurs at the BAO distance, as expected. However, this particular result does not have a high significance, so we did not entirely isolate the BAO signal. For future research, having more simulations could reduce the noise and thus give a higher significance. Since we saw non-linear effects in the correlation difference functions, this shows that voids can capture higher order statistics that halos cannot capture. The cross-correlation functions (Figures 7 and 8) show the BAO signal with high significance, which proves that voids are a useful tool in extracting the BAO signal.

We found that the void cut of radii larger than 40 Mpc/h enhanced the peak the most because it shifted the original peak to the right, above the BAO distance. Therefore, the resonant void size that emphasizes the BAO signal is 40 Mpc/h. However, this result has weak statistics.

5 Implications:

Since the first claimed Baryon Acoustic Oscillation (BAO) signal in voids only occurred in 2015, this is a relatively new and unexplored research topic. This research will have implications for large-scale galaxy surveys in the 2020s, including the Large Synoptic Survey Telescope (LSST), in which Princeton University is heavily involved. These upcoming surveys have a goal of measuring the BAO. Since large scale surveys have very good statistical capabilities, these surveys could utilize cosmic voids as a valuable tool.

This project could have implications on future survey designs. Once the error bars are narrowed down, we could design a survey with a specific luminosity cut to get a better void catalog that enhances the BAO signal.

Additionally, combining BAO measurements with cosmic microwave background measurements could provide some of the most powerful constraints on our cosmological model. Thus, better constraining the BAO could significantly impact our understanding of cosmology.

6 Acknowledgements

I would like to thank my mentor Christina Kreisch for guiding me through the research process and ReMatch+ for funding the research. I would also like to thank Princeton Unviersity Department of Astrophysical Sciences for hosting me this summer.

7 References

(1) arXiv:1511.04391v1 [astro-ph.CO] 13 Nov 2015

(2) https://github.com/franciscovillaescusa/Quijote-simulations

(3) VIDE: arXiv:1406.1191 [astro-ph.CO]

Appendix

Quijote Simulation

The Quijote simulations consist of a set of 34,500 N-body simulations. They mainly serve to quantify the information content on cosmological observables and provide the statistics to train machine learning algorithms. These simulations include 15,500 simulations for a fiducial Planck cosmology (based on observational data) and 4000 simulations in a latin hypercube expanding 5 cosmological parameters. Each simulation consists of 512^3 cold dark matter particles and the same number of neutrino particles. A single redshift has 5 trillion particles. The box size is 2 Gpc/h, with a spatial resolution of 50 kpc/h. The initial conditions have outputs at redshifts 0, 0.5, 1, 2, 3, and 127. The halo catalogues and void catalogues both have length 172,500. The power spectra and bispectra are both more than 1 million. There is 750 Tb of data and the simulations have 18 million cpu hours. The simulations run with the TreePM code Gadget-III.