

The Convergence of 21cm Power Spectra within a 1 Gpc/h Box Simulation during the Epoch of Reionization

Andrew Caruso & Christopher Cain

The University of California Riverside

June 15, 2023

Abstract

We performed a convergence test to ascertain the maximal simulation box size necessary to encapsulate a sufficient the 21cm radiation power in a simulation box of length 1 Gpc/h during the EoR. Using the data from two intense models of ionization sources, oligarchic and extremely oligarchic, we found that a subvolume size of 500 Mpc/h is adequate, given its mean in violin plots of 0.0190 ± 0.0204 and 0.0221 ± 0.0203 , respectively. Although a maximal size was determined, further investigation is needed to ascertain the box size needed for less intense models of ionizing sources during the EoR.

1 Introduction

The Epoch of Reionization (EoR) is a stage in the thermal history of the early universe in which the first sources of radiation are able to ionize the prior neutral hydrogen (HI) of the intergalactic medium (IGM). However, several aspects of the EoR remain unresolved including the morphology, evolution, and sources of ionizing radiation. The investigation of the EoR was recently boosted by several observations constraining the end of the EoR, including limits on the reionization optical depth observed by the Planck Collaboration et. al. [1]. Likewise, [2] used Lyman Alpha damping wing data to find that the reionization of the IGM is incomplete at $z \sim 7$. This was also observed through massive Lyman Alpha emitter surveys conducted by [3]. On the other hand, [4] found that the reionization of the IGM may be complete at $z \lesssim 6$.

With these new constraints and the recent James Webb Space telescope observations, which have clashed with theoretical predictions [?], more attention is drawn towards understanding the EoR and the Cosmic Dawn (the formation of the first stars, prior to the EoR) [?].

One way to potentially resolve these uncertainties is by directly observing the HI regions of the EoR. This can be done by detecting the 21cm radiation during the EoR. In particular, measuring the 21cm power spectrum will reveal the drivers or sources of ionization, such as the first stars and galaxies, including their properties and impact on the IGM and structure formation thereafter [?]. Likewise, dark matter halos hosting these first sources can also be studied [?]. Ultimately,

understanding the evolution of the universe during the EoR using 21cm radiation may provide new constraints on the first ionizing sources and potential dark matter models and their roles during the early universe [?].

However, the task has been experimentally challenging due to the large foreground of later 21cm radiation [?]. Nevertheless, radio telescope interferometers such as GMRT [?], LOFAR [?], LWA [?], MWA [?], PAPER [?], and now HERA [?] have set constraints on the upper limits of the 21cm radiation power spectrum. The most recent limit from HERA on the brightness temperature spectrum is less than 946 mK^2 at redshift 7.9 with wavenumber 0.19 Mpc^{-1} and less than $9,166 \text{ mK}^2$ at redshift 10.4 with wavenumber 0.26 Mpc^{-1} [5]. Moreover, HERA [?] and SKA [?] are predicted to potentially observe the 21cm power spectrum for the first time [?].

During this search for the 21cm power from the EoR, theoretical models have become further constrained and advanced in predicting the aspects of EoR using simulations in conjunction with machine learning [?]. But one major concern remains as to how large the simulation box size must be in order to capture all the 21cm radiation power spectrum for optimal simulation environments [?]. A few collaborations have explored potentially a minimum box size of approximately 250 Mpc/h [6] or 100 Mpc/h [7], however, the question still remains open regarding the maximum size.

One way to address this question is by studying the convergence of the 21cm power at different length scales. In order to test the convergence, we performed several radiative transfer simulations using a variety of models of ionizing sources and a very large simulation volume of 1 Gpc/h , which has not been previously explored. However, we began the convergence test using a smaller box size of 300 Mpc/h as a forerunner to the larger volume analysis. We also pushed the convergence test to an upper limit by utilizing oligarchic and extremely oligarchic models of ionizing sources.

2 Procedure

2.1 Models of Ionizing Sources

A key parameter in performing the radiative transfer simulation is the model of the ionizing sources. Currently, it is unknown how the EoR depends exactly on the ionizing sources (corresponding to different sizes of dark matter halos), such as quasars and galaxies [?], and which type of ionizing source contributes predominantly [?]. Consequently, the simulated 21cm power spectrum may depend on the choice of the ionizing sources and thus requires the simulation to be run multiple times with different sources.

Interestingly, when the the 21cm power spectrum is observed in future experiments [?], it will be helpful to have simulated expectations of the 21cm power spectrum to narrow down the range of possible models of first ionizing sources of the universe [?].

We utilized three distinct ionization models: fiducial, oligarchic, and democratic. Essentially, each model differs by the choice of dark matter halo size and thus the source of ionization. The fiducial model involves dark matter halos with a minimum mass of $2 \times 10^9 M_\odot$, wherein the emissivity of ionizing radiation is proportional to each halo's luminosity. Typical simulations utilize this type of model. In contrast, the democratic model removes the emissivity dependency on luminosity for all dark matter halos. However, the democratic model's sources are currently undetectable. Lastly, the oligarchic model involves even more massive dark matter halos with a minimum mass of $2 \times 10^{10} M_\odot$. Unlike both models, the oligarchic sources have already been detected [?].

As an illustration of each model, Figure 1 displays the geometric cross sections of the 21cm brightness temperature fields at a chosen positional cut of 10 Mpc/h , with a fixed redshift of

7.0445. Each plot obeys the same color scheme as denoted by the spectral color bar below. The very cold regions of dark blue (less than about 10 mK) correspond to the lack of 21cm radiation and are, therefore, HII (ionized hydrogen) regions rather than HI (neutral hydrogen) regions. Note that these contour maps are from the simulation with a maximum box length of 300 mpc/h.

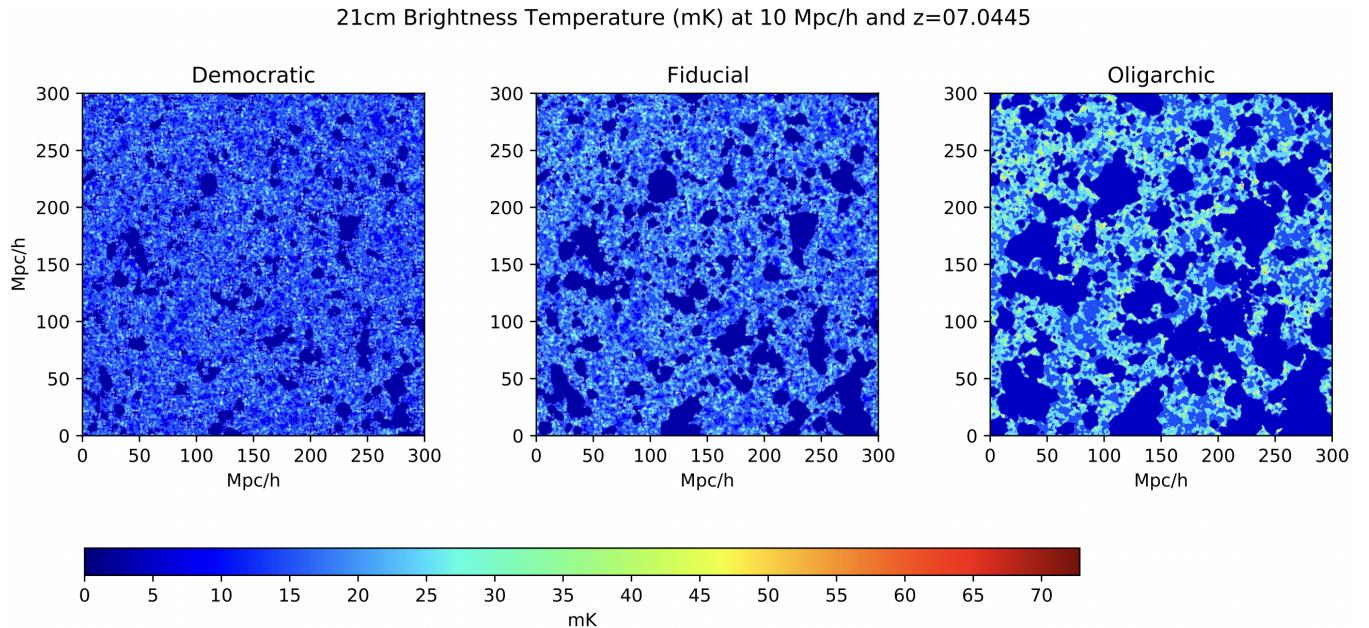


Figure 1: Brightness temperature contour maps (in units of mK) for the democratic, fiducial, and oligarchic models at redshift of 7.0445 with a chosen positional cut of 10 Mpc/h for the 300 Mpc/h length box simulation. **This is a small thing but can you center the color-bar here like you did in the other plot?**

At redshift ~ 7 , the 300 Mpc/h box is expected to be about 50% ionized. Clearly, from Figure 1, the democratic model has few and smaller HII regions. The effect increases in the fiducial model, which has more and bigger HII regions, but also a greater difference in temperature between HI and HII regions. This effect intensifies further in the oligarchic model, which has much greater HII regions and a more clear distinction between HI and HII regions. This illustrates that depending on the source model, the ionization occurs at different rates, which ultimately effects the reionization evolution. Overall, the democratic sources are slower than fiducial sources, while oligarchic sources are faster at ionizing than fiducial sources.

We also implemented a fourth model, the extreme oligarchic model, which is simply an amplification of the oligarchic model, as seen in Figure 2. As expected, the extreme model has greater and more frequent HII regions than the oligarchic model, with a larger temperature gradient between the HII and HI regions. However, we only used both of these oligarchic models for the 1 Gpc/h box simulation, given our interest in the extreme limit of the ionizing sources. Nevertheless, the three previous models were implemented for the 300 Mpc/h box simulation as a starting point of the data analysis.

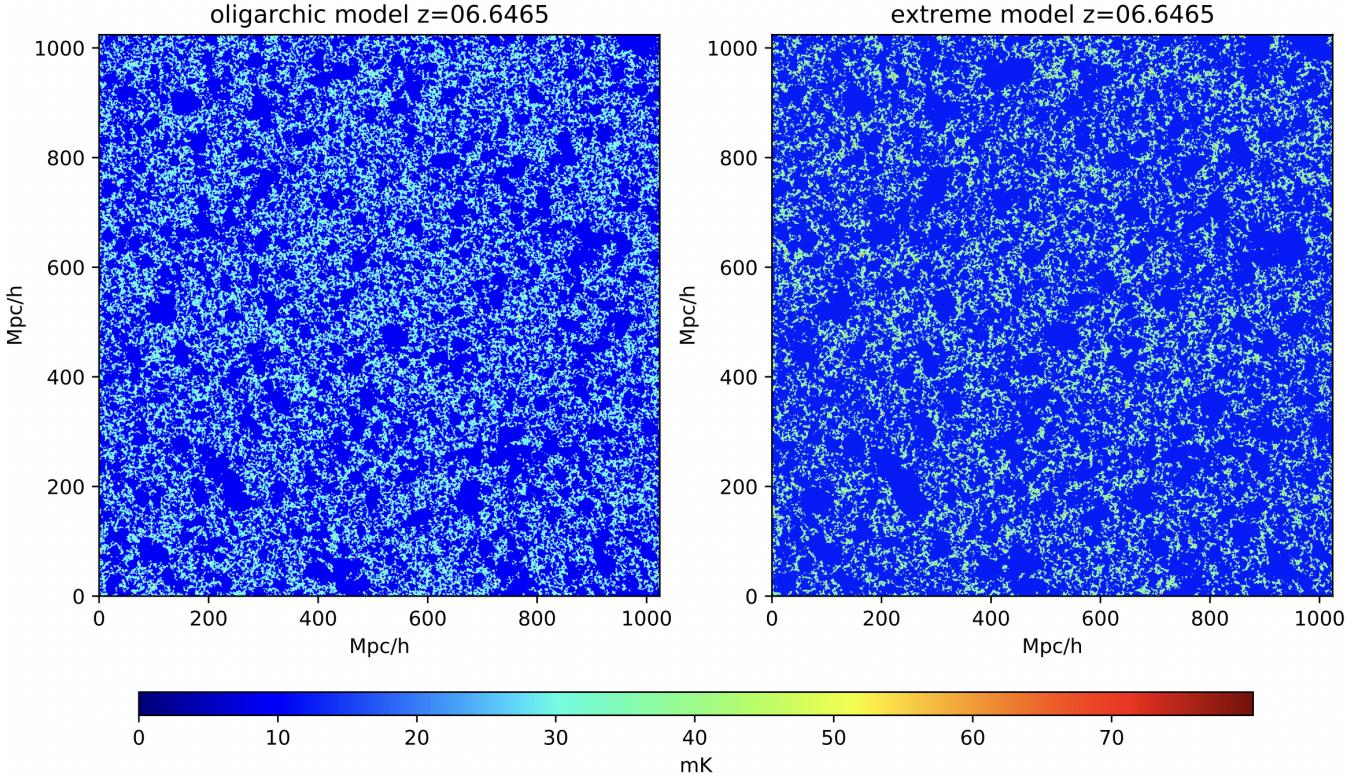


Figure 2: Brightness temperature contour maps (in units of mK) for the oligarchic and extremely oligarchic models at redshift of 6.6465 with a chosen positional cut of 10 Mpc/h for the 1 Gpc/h length box simulation.

2.2 Subvolume Selection

The simulation was performed using each model to generate the 21cm radiation brightness temperature fields for redshifts z between 4.5 and 11.8 for the 300 Mpc/h box and 6 to 10.9 for the 1 Gpc/h box. Thereafter, the power spectra was generated from the 21cm fields by performing Fast Fourier Transforms (FFT).

Since the simulation output data was large and the data analysis required fast computational power, we utilized Expanse supercomputer cluster, which is part of the ACCESS project at the San Diego Supercomputer Center.

In order to chose the subvolumes aptly, such that their average densities are comparable to the average cosmic density, the density fields corresponding to the 21cm brightness temperature fields at each redshift were utilized. Note that the density field data were in units of the average cosmic density (which corresponded to unity). Incriminating through the cell scales of the simulation box at a fixed subvolume size, the average densities of these subvolume cuts were compared to the average cosmic density 1 ± 0.99 (with a chosen weight of 0.1 on the standard deviation). Subvolumes within this range of the average cosmic density were stored for analysis. The weighted deviation was determined by calculating the standard deviation of the average density fields and then adjusting the weight on the standard deviation to capture a sufficient number of subvolumes. The fixed box sizes included 100, 200, 300, 400 and 500 Mpc/h for the 1 Gpc/h box and 75, 100, and 150 Mpc/h for the 300 Mpc/h box. Additionally, the step sizes were 60 Mpc/h for the 300 Mpc/h box and 128 Mpc/h for the 1 Gpc/h box.

Duly note that for the smaller box simulation, the number of cells were equal to the physical distance scale, however, this was not the case the larger simulation. The 1 Gpc/h box had 1024 cells in comparison to the 1000 Mpc/h physical length. Furthermore, the collection of subvolumes at a fixed box size were kept separate from the other box sizes, in order to perform the convergence test, as explained in the next section.

Without the comparison to the average cosmic density, the convergence test on random subvolumes would overestimate or underestimate unexpectedly at large redshifts or wavenumbers, which would effect the power spectra plots accordingly. Additionally, the 21cm brightness temperature fields and density fields were not necessarily generated in the same direction. Thus, to determine a matched orientation, contour maps were created at a chosen positional cut of 10 Mpc/h along the z axis. The orientation test was performed at a high redshift $z = 12$ since this precedes the EoR such that most of the space is not ionized and therefore the 21cm fields and corresponding density fields will clearly exhibit similar features if aligned. By swapping the 3D positional axes of the 21cm brightness temperature fields, a matched orientation was achieved for the 300 Mpc/h box and the 1 Gpc/h box.

Although the density and brightness temperature contours were on different scales of temperature in mK, they both displayed similar high and low brightness temperature features which indicated a correct orientation. This was essential when corresponding the fields to densities when selecting appropriate subvolumes to generate power spectra.

2.3 Convergence Test

After selecting the appropriate subvolumes for each model separately, the power spectra were generated from the 21cm brightness temperatures via the FFT. The resultant power was generated per wavenumber k at different redshifts and subvolume sizes for each model. Using all the power spectra for a single model and fixed subvolume size, the average power was determined at each wavenumber or redshift, alongside the standard deviation. Likewise, the power spectrum of the entire box was found and then compared to the average power spectrum from the subvolume and its corresponding deviation. This process was repeated for all the chosen subvolumes and again reiterated for each model.

The convergence test essentially compares the average power spectrum to the total power spectrum at different subvolumes cuts and models. We created ratio plots by taking the average power, standard deviation from the average, and the total power, and dividing by the total power. For the 300 Mpc/h simulation, the smallest subvolume cut was 75 Mpc/h and the ratio plots in respect to varying wavenumber and fixed redshifts is seen in Figure 3.

Each subplot in Figure 3 is the ratio of the average power to the total power (dash green line) with its corresponding deviation (the shaded green region), at a fixed redshift or wave number for each model. The red line denotes unity (the ratio of the total power to itself). Ultimately, the closer the average power is to unity and the more narrow the standard deviation from the average power is, the more converged the average power is to the total, thus capturing most of the 21cm power spectra at a given redshift, wavenumber, and model.

As seen in Figure 3, at very short wave numbers or very small redshifts, the average power seems to diverge rapidly from the total power. Clearly the convergence depends on the choice of redshift and wavenumber, as well as the model of the ionizing sources.

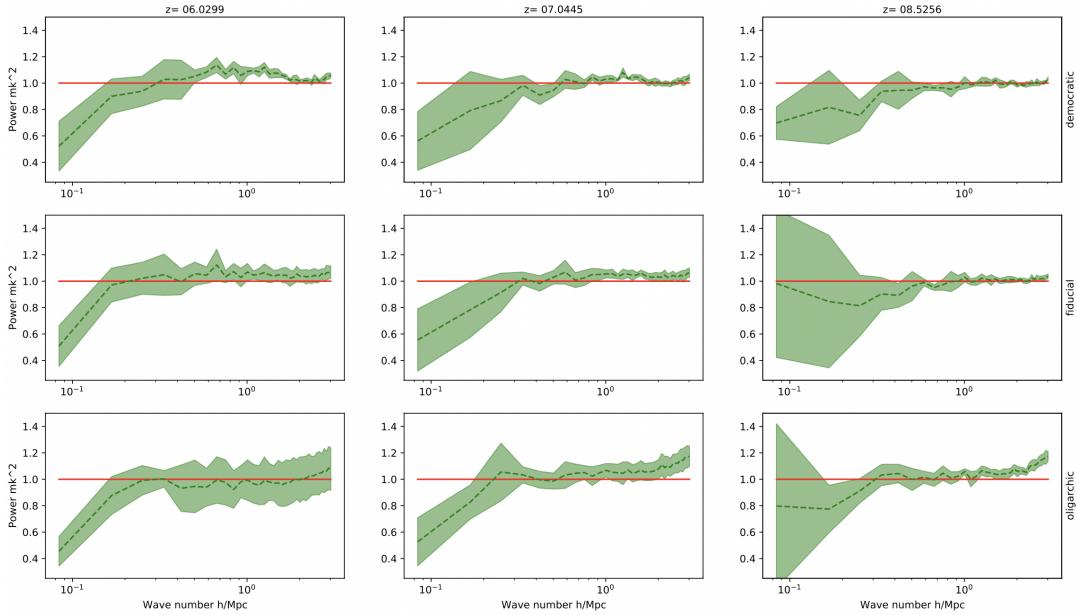


Figure 3: Ratio plots for subvolume cut of 75 Mpc/h for 300 Mpc/h box simulation. Each subplot is power in mk^{-2} versus wavenumber in h/Mpc , with each column being at a fixed redshift and each row corresponding to a model.

On the other hand, Figure 4 displays the largest subvolume cut of 150 Mpc/h with respect to wavenumber. In comparison to Figure 3, the larger subvolume has a smaller standard deviation than the smaller subvolume, as expected. Furthermore, the average power seems to be more aligned with unity for the larger subvolume, thus suggesting more convergence to the total power of the 300 Mpc/h box at each corresponding model and redshift.

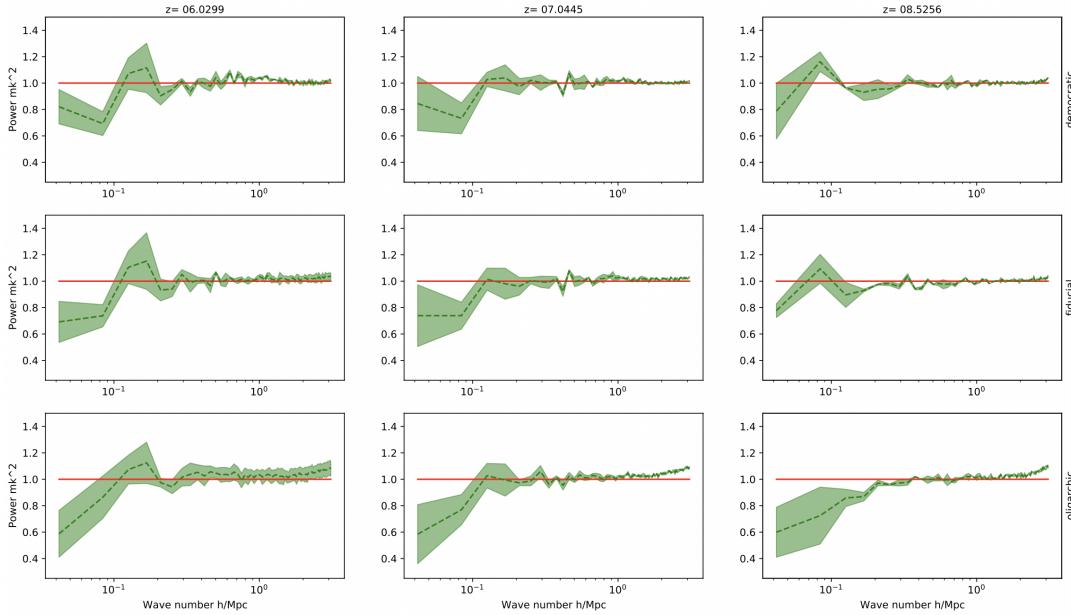


Figure 4: Ratio plots for subvolume cut of 150 Mpc/h for 300 Mpc/h box simulation. Each subplot is power in mk^2 versus wavenumber in h/Mpc , with each column being at a fixed redshift and each row corresponding to a model.

Figure 5 displays the smallest subvolume of 75 mpc/h, however, with varying redshift instead at fixed wavenumbers. The subplots exhibit similar behavior to that of Figure 3.

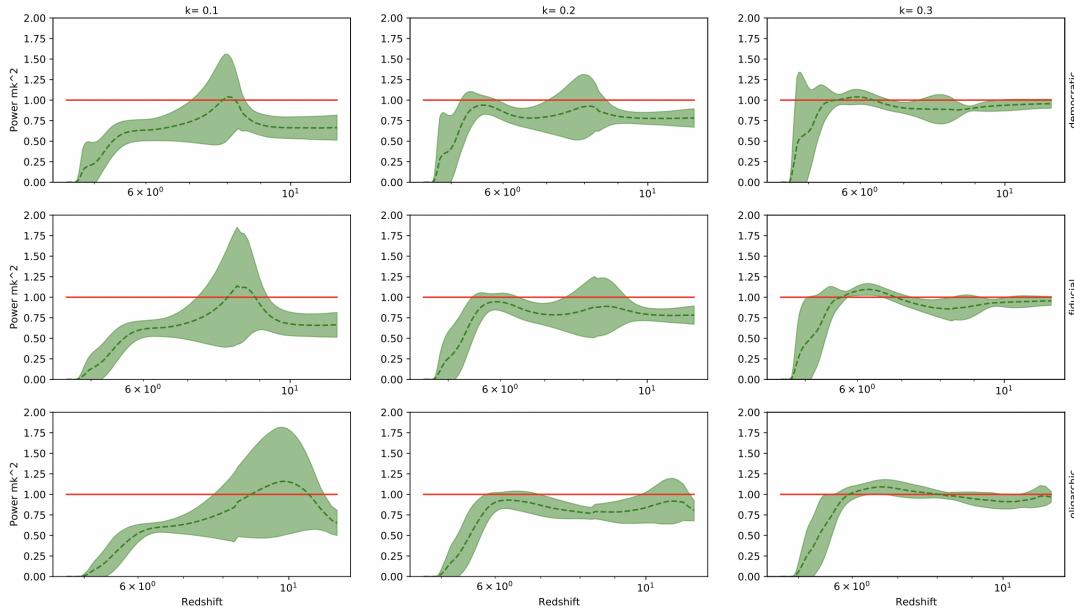


Figure 5: Ratio plots for subvolume cut of 75 Mpc/h for 300 Mpc/h box simulation. Each subplot is power in mk^2 versus redshift, with each column being at a fixed wavenumber in Mpc/h and each row corresponding to a model.

Besides the ratio plots for the 300 Mpc/h box simulation, we also acquired similar plots for the 1 Gpc/h box. Figure 7 and 6 display the ratio plots for the 300 Mpc/h subvolume and 500 Mpc/h

subvolume cuts at fixed redshifts of 6.6465 and 6.9746. Furthermore, the dependence on redshift at fixed wavenumbers of

Both exhibit similar characteristics described previously for the smaller box simulation, however, it is less clear which one is most converged given the trade off between a more narrow standard deviation for the 300 Mpc/h subvolume in Figure 7 and a closer to unity average power to total power for the 500 Mpc/h subvolume in Figure 6.

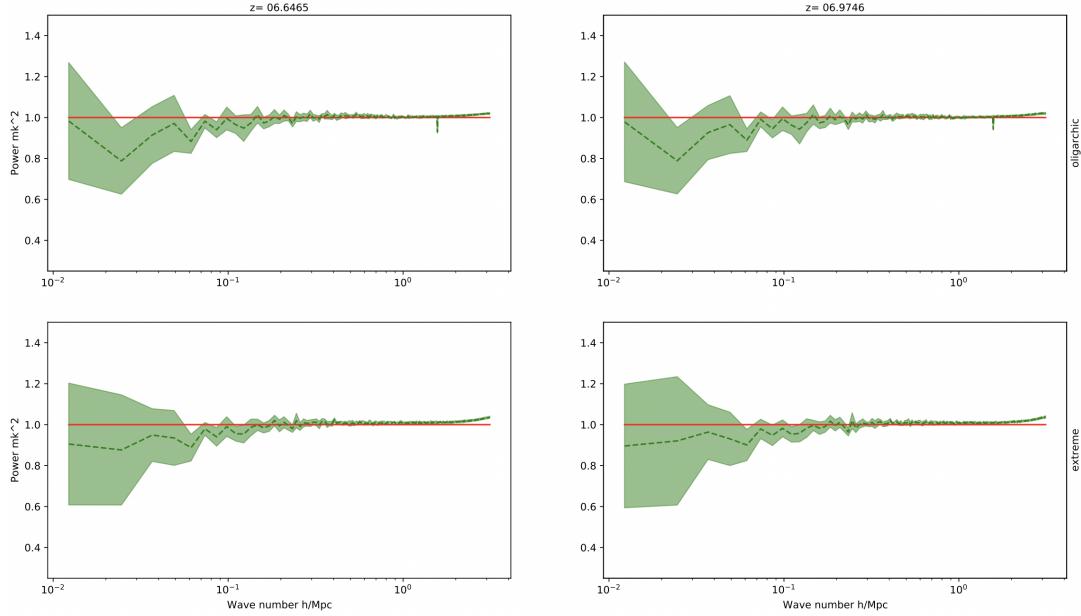


Figure 6: Ratio plots for subvolume cut of 500 Mpc/h for 1 Gpc/h box simulation. Each subplot is power in mk^2 versus wavenumber in h/Mpc , with each column being at a fixed redshift and each row corresponding to a model.

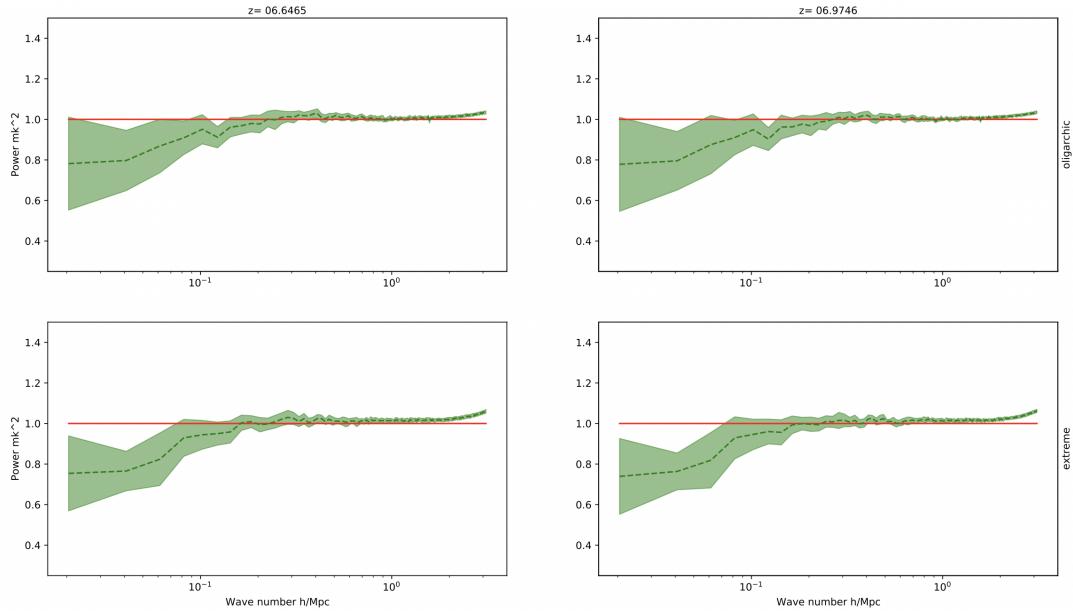


Figure 7: Ratio plots for subvolume cut of 300 Mpc/h for 1 Gpc/h box simulation. Each subplot is power in mk^2 versus wavenumber in h/Mpc , with each column being at a fixed redshift and each row corresponding to a model.

An alternative view contrasting the two cuts can be found in Figure 9 and Figure 8, which vary in respect to redshift in lieu of wavenumber. Both appear to be converged well for wavenumbers of 0.1, 0.2, and 0.3 h/Mpc , however, the 500 Mpc/h cut appears to have a more tighter adhesion to unity with less spread than the 300 Mpc/h cut.

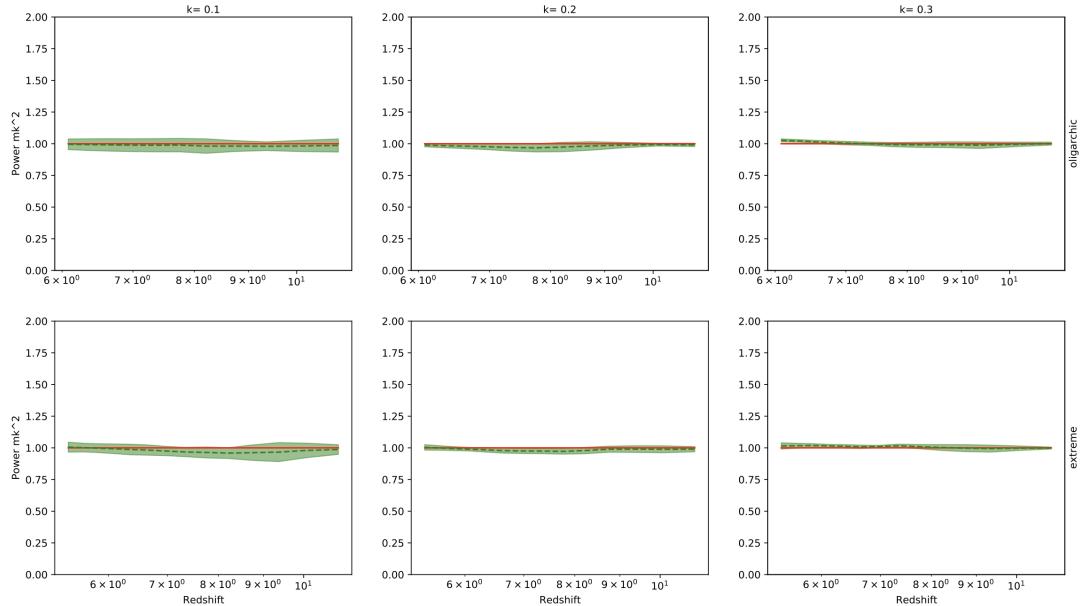


Figure 8: Ratio plots for subvolume cut of 500 Mpc/h for 1 Gpc/h box simulation. Each subplot is power in mk^2 versus redshift, with each column being at a fixed wavenumber in h/Mpc and each row corresponding to a model.

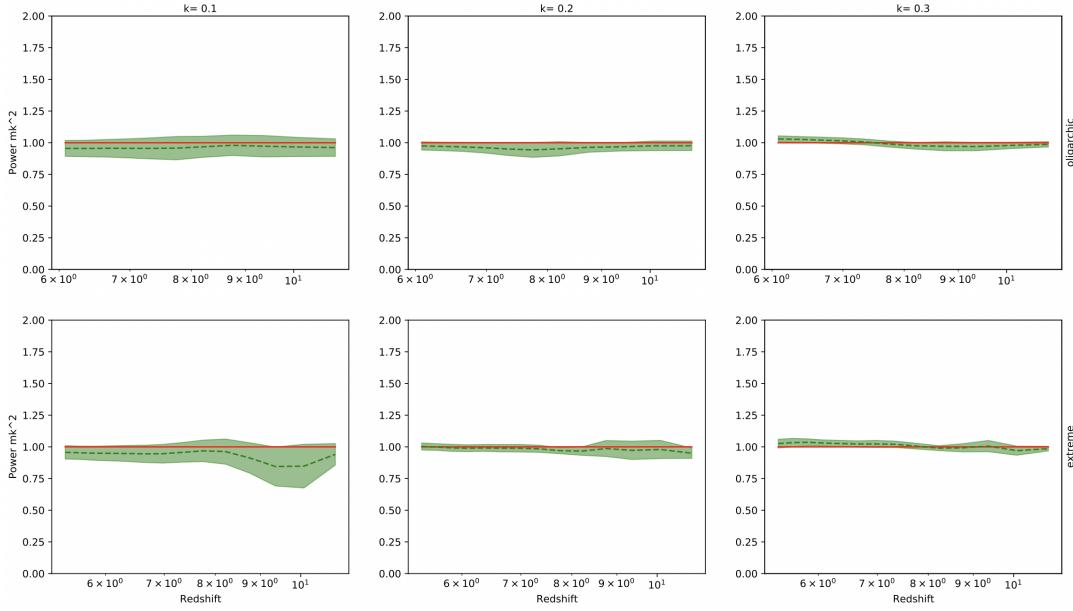


Figure 9: Ratio plots for subvolume cut of 300 Mpc/h for 1 Gpc/h box simulation. Each subplot is power in mk^2 versus redshift , with each column being at a fixed wavenumber in h/Mpc and each row corresponding to a model.

3 Results and Analysis

In order to illustrate the optimal subvolume cut size more readily for each model in the 1 Gpc/h box simulation, we created violin plots for the oligarchic and extreme models. The chosen subvolume cuts included 100 to 500 Mpc/h with a step of 100 Mpc/h. The oligarchic model violin plot is shown in Figure 10 for fixed redshift of 6.6465 and wavenumber 0.2 Mpc/h, relevant to the EoR.

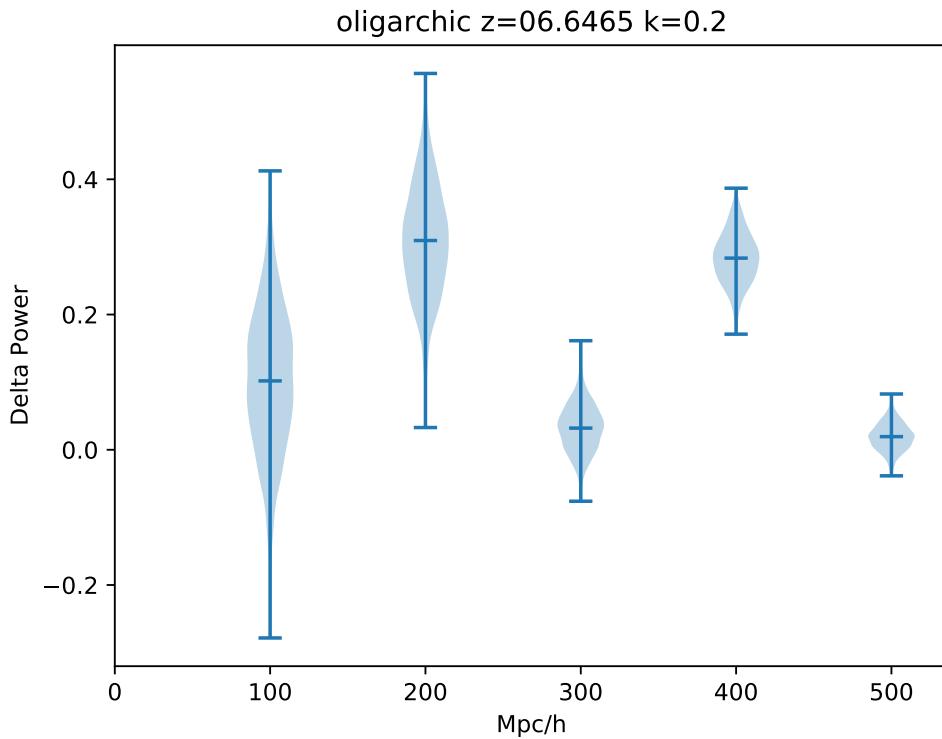


Figure 10: Violin plots for the oligarchic model in the 1 Gpc/h box simulation at fixed redshift and wavenumber with subvolume cuts of 100, 200, 300, 400, and 500 Mpc/h as labeled on the x axis.

Each violin plot in Figure 10 represents the Gaussian distribution for each cut in Mpc/h in terms of ΔP . The Gaussian distributions utilized the standard deviation of all the power ratios (power value to the total power) and a mean ΔP as expressed in equation 1. Note that the offset of 1 was merely for plotting convenience.

$$\Delta P = \left| \frac{P_{avg}}{P_{tot}} - 1 \right| \quad (1)$$

For the oligarchic model, the 500 Mpc/h cut is most converged with a smaller standard deviation and more narrow distribution than other subvolume cuts. Moreover, the 500 Mpc/h mean is the closest to zero, implying the average power is approximately equal to the total power of the simulation box. To be precise, the 500 Mpc/h cut had a mean of 0.0190 ± 0.0204 , in comparison to the 300 Mpc/h cut with a mean of 0.0327 ± 0.0369 . A similar trend appears also for the extremely oligarchic model, in Figure 11, which has means of 0.0221 ± 0.0203 and 0.0104 ± 0.0291 for the 500 Mpc/h and 300 Mpc/h cuts, respectively. However, in this model, the 500 Mpc/h cut is not the closest, but is still very close relative to the competing 300 Mpc/h cut and has a clearly the smallest deviation and narrowest distribution.

Duly note, however, that the violin plots do not contain subvolume cuts greater than 500 Mpc/h. Therefore, we can only infer that box sizes greater are sufficiently or far more converged to the total power of the 1 Gpc/h box. Nevertheless, the subvolumes do not converge sufficiently until, in this case, a subvolume of 500.

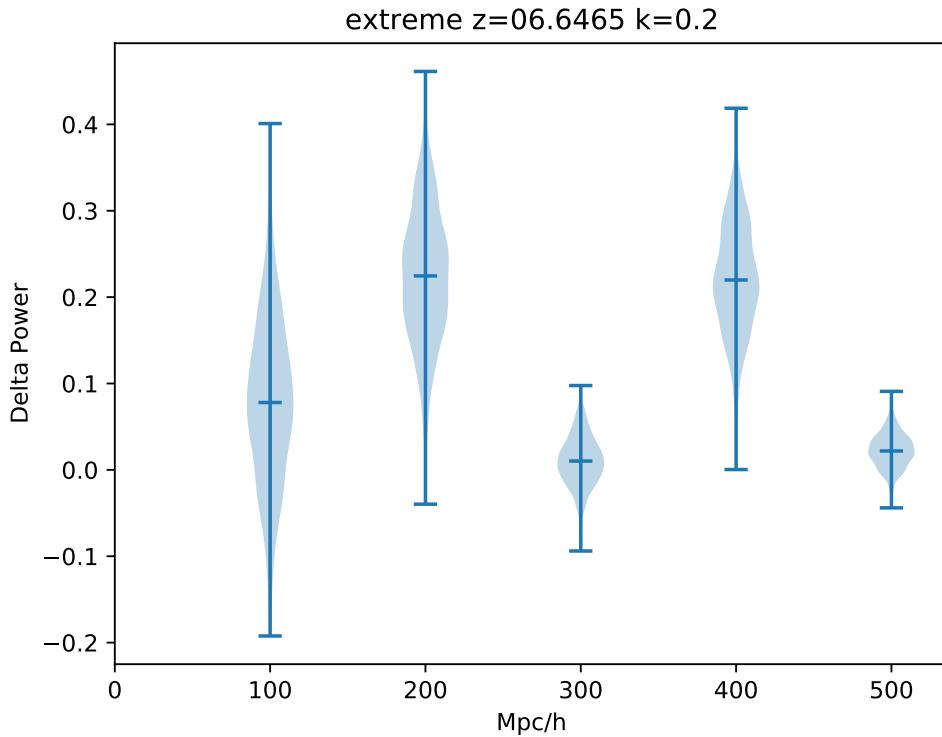


Figure 11

4 Conclusion

we found

Consequently, it was found that for the 1 Gpc/h simulation box, the convergence test for both the oligarchic and extremely oligarchic models yielded a converged 500 Mpc/h cut, which has the shortest standard deviation and narrowest distribution compare to other smaller subvolume cuts, with a mean of 0.0190 ± 0.0204 and 0.0221 ± 0.0203 , respectively. Id est, subvolumes below 500 Mpc/h do not converge enough and subvolumes greater than 500 Mpc/h sufficiently converge, thus placing an upper threshold on simulation box size.

Since these models of ionization sources are most intense in driving the EoR, the choice of a 500 Mpc/h simulation box to capture a significant amount of 21cm radiation power is only valid in that regime.

Therefore, further analysis of a large simulation box of the 21cm radiation using less intense models, to better constrain the maximum box size needed to encapsulate a significant amount of the total 21cm radiation power. Likewise, more subvolume cut sizes can be selected, such that the window of the threshold is much smaller and more precise.

Nevertheless, the analysis herein, has provided some illumination as to how large a simulation box must be to adequately analyze the 21cm radiation from the EoR. Furthermore, the general methods may also be applicable in looking at other parameters associated with the EoR, to further constrain simulation size. The ultimate effort will be to create better simulations prior to the first detection of the 21cm radiation from the EoR in order to easily and rapidly pinpoint theoretical models regarding the EoR.

5 Acknowledgements

This paper served as the senior thesis for Andrew Caruso for his undergraduate senior thesis at the University of California Riverside (UCR). Thus, a few important acknowledgements ought to be expressed in gratitude of undertaking this project.

Recognition goes foremost to the supervisor, Chris Cain, the graduate researcher who ran the simulations per se and over-viewed the data analysis performed by Andrew Caruso.

Furthermore, the project was completed under the principal investigator, Professor Anson D'Aloisio at the University of California Riverside, of whom this project would not have transpired without. Andrew would like to recognize Professor D'Aloisio for the opportunity to partaken in this theoretical cosmological research with Chris and for inspiring curiosity in astrophysics and cosmology in Andrew's future field of interest: theoretical astroparticle physics.

References

- [1] N. Aghanim *et al.*, “Planck 2018 results. VI. Cosmological parameters,” *Astron. Astrophys.*, vol. 641, p. A6, 2020. [Erratum: *Astron.Astrophys.* 652, C4 (2021)].
- [2] B. Greig, A. Mesinger, Z. Haiman, and R. A. Simcoe, “Are we witnessing the epoch of reionization at $z=7.1$ from the spectrum of J1120+0641?,” *Mon. Not. Roy. Astron. Soc.*, vol. 466, no. 4, pp. 4239–4249, 2017.
- [3] W. Hu, J. Wang, Z.-Y. Zheng, S. Malhotra, J. E. Rhoads, L. Infante, L. F. Barrientos, H. Yang, C. Jiang, W. Kang, L. A. Perez, I. Wold, P. Hibon, L. Jiang, A. A. Khostovan, F. Valdes, A. R. Walker, G. Galaz, A. Coughlin, S. Harish, X. Kong, J. Pharo, and X. Zheng, “The $\text{Ly}\alpha$ luminosity function and comsic reionization at $z \sim 7.0$: A tale of two lager fields,” *The Astrophysical Journal*, vol. 886, p. 90, nov 2019.
- [4] Y. Zhu *et al.*, “Chasing the Tail of Cosmic Reionization with Dark Gap Statistics in the $\text{Ly}\alpha$ Forest over $5 < z < 6$,” *Astrophys. J.*, vol. 923, no. 2, p. 223, 2021.
- [5] Z. Abdurashidova *et al.*, “Improved Constraints on the 21 cm EoR Power Spectrum and the X-Ray Heating of the IGM with HERA Phase I Observations,” *Astrophys. J.*, vol. 945, no. 2, p. 124, 2023.
- [6] H. D. Kaur, N. Gillet, and A. Mesinger, “Minimum size of 21-cm simulations,” 4 2020.
- [7] I. T. Iliev, G. Mellema, K. Ahn, P. R. Shapiro, Y. Mao, and U.-L. Pen, “Simulating cosmic reionization: How large a volume is large enough?,” *Mon. Not. Roy. Astron. Soc.*, vol. 439, no. 1, pp. 725–743, 2014.