RNAAS RESEARCH NOTES OF THE AAS

The Study of Quasar Clustering at Low Redshifts

To cite this article: Jose Ordonez and Carolyn Stripling 2022 *Res. Notes AAS* **6** 90

Draft version April 25, 2022
Typeset using LaTeX default style in AASTeX631

The Study of Quasar Clustering at Low Redshifts

Jose Ordonez[1] and Carolyn Stripling[1]

[1] The University of Texas at Austin, Austin, Texas

ABSTRACT

We use a nearest neighbor algorithm combined with the machine learning clustering function DBSCAN (Density-Based Spatial Clustering of Applications with Noise) to analyze a subset of quasars from the Sloan Digital Sky Survey. Our analysis shows evidence of clustering up to $z \sim 2$, with evidence of an increase in clustering at lower redshifts ($z < 0.5$). Our findings may suggest that, over time, the gravitational interaction between quasars has led to more clustering at low redshifts.

1. INTRODUCTION

Due to quasars' high luminosity and location at the center of massive, distant galaxies, analyzing their properties is crucial in understanding galaxy formation and the large scale structure of the universe (Bahcall 1988; Shen et al. 2009). Attempts have been made to determine the extent of quasar clustering over the past 40 years, for example by using the nearest neighbor test along with a Monte Carlo routine (Chu & Zhu 1983) or by using the redshift-space correlation function (Ross et al. 2009). However, even in 2009, results from research on quasar clustering were not yet comparable to those of local galaxy clustering (Shen et al. 2009). In this paper, we revisit the problem of quasar clustering using a similar approach as Chu & Zhu, but using a more modern value of $H_0 = 72$ km s^{-1} Mpc^{-1} instead of their value of $H_0 = 50$ km s^{-1} Mpc^{-1}, and combining a nearest neighbor algorithm with DBSCAN (Density-Based Spatial Clustering of Applications with Noise), a machine learning function used to determine cluster membership (Pedregosa et al. 2011).

2. DATA

Data for our quasars were obtained from the First Data Release of the Sloan Digital Sky Survey (SDSS) (Schneider et al. 2003), from which a total of 16,826 quasars were sampled. This was done by using an SQL query to select all objects with a `specclass=3` (the quasar class) from the `SpecPhoto` table. The obtained quasars range in redshift from $0.1 < z < 5.4$, although the vast majority ($> 96\%$) have a $z < 2.3$. The given right ascension and declination were transformed into galactic longitude and latitude as well as Cartesian x, y, and z coordinates. We computed estimated distances to each quasar using their redshifts along with Hubble's constant.

3. RESULTS

We first wrote our own nearest neighbor algorithm. By only considering quasars which are close to a given quasar's x, y, and z coordinates instead of all the quasars in the sample, our algorithm can efficiently determine the distance to a given quasar's nearest neighbor. Using our nearest neighbor algorithm, we find that the mean distance to a quasar's nearest neighbor in our data is 41.8 Mpc. This value is then used as one of the two parameters for the DBSCAN function. Specifically, we set `eps` (the maximum distance a quasar can be from another to be considered part of the same cluster) to 41.8. For the parameter `n` (the minimum number of quasars in a dense region required to be considered a cluster), we use a value of 5. This value is chosen to be small enough to maximize the number of clusters obtained from running DBSCAN while being large enough to allow the estimation of cluster radii using a convex hull without all hulls becoming irregular tetrahedrons.

Running the DBSCAN function with these parameters finds a total of 228 clusters in our data. An analysis of each cluster was done to obtain its number of members, approximated radius (idealizing each cluster to be a sphere with the volume of the convex hull encompassing the members), and mean redshift. On average, our clusters had 7.5 members and a radius of 15.5 Mpc. The distribution of the clusters' redshifts is compared to the distribution of the original quasars in Figure 1. One can see that, apart from a gap in $0.5 < z < 1$, there exists evidence of clustering up to $z \sim 2$. The 4% of quasars in our data with a $z > 2.3$ were too few to result in any significant clustering at that redshift range.

Furthermore, the amount of clustering peaks at a redshift of ~ 0.25. The largest cluster that was found in our data also resides in this area, with a $z = 0.126$ and a total of 52 members.

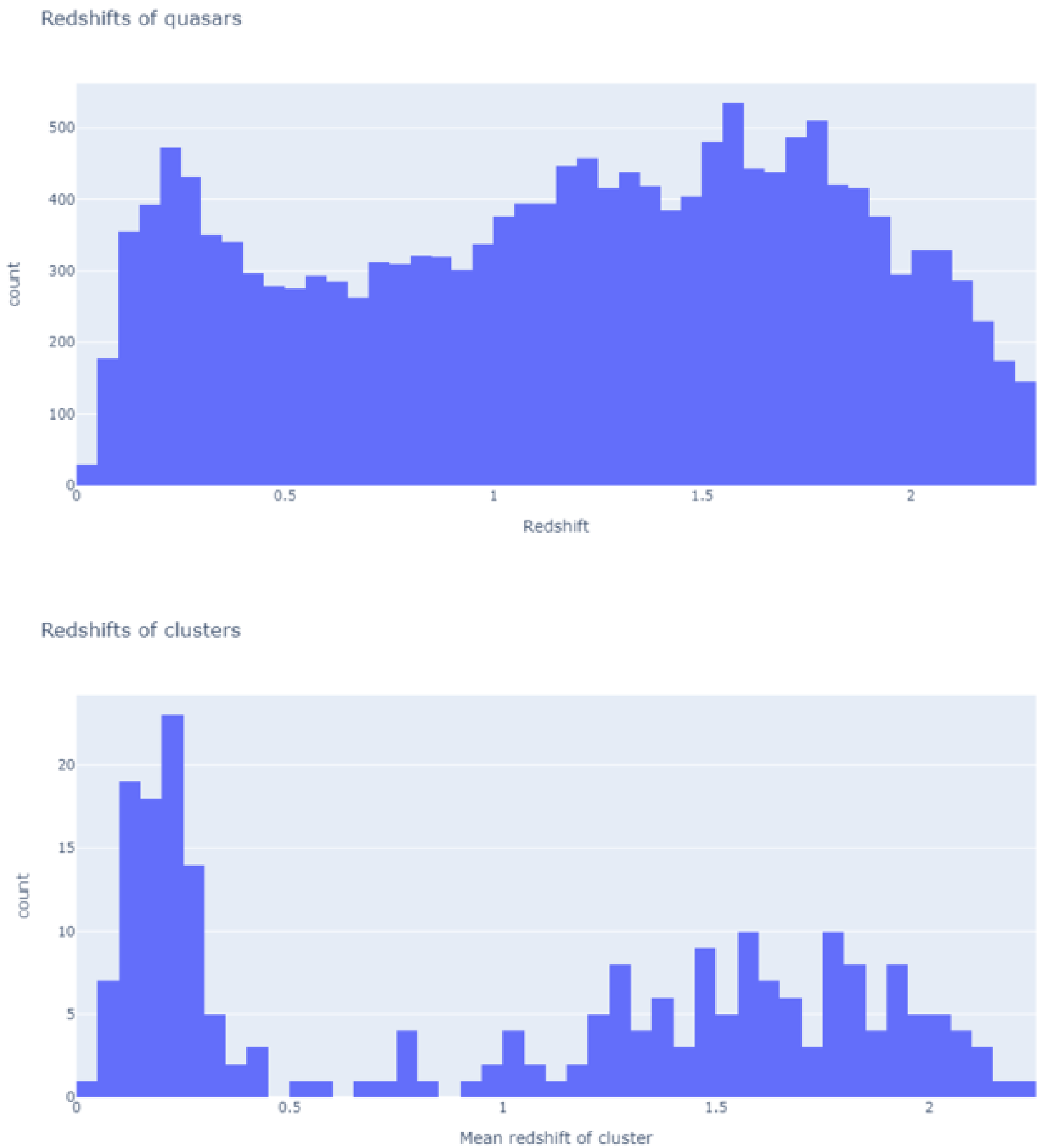

Figure 1. Histograms of the distribution of redshifts of quasars and redshifts of the found clusters.

4. CONCLUSIONS

Our value for the mean distance to a quasar's nearest neighbor of 41.8 Mpc is substantially smaller than Chu & Zhu's value of 166.7 Mpc in their fields. We attribute this to the fact that Chu & Zhu analyzed two fields of ~120 quasars and used a smaller value for Hubble's constant. Our more recently updated catalogue of thousands of quasars along with a more modern value of $H_0 = 72$ km s^{-1} Mpc^{-1} help explain this discrepancy, and suggest that our value is more accurate.

By definition, DBSCAN finds areas of clustering by determining whether a quasar is in a dense region of space (here defined as a region of space where distances between quasars are less than the average distance to a quasar's nearest neighbor) or not. The distribution seen in Figure 1 implies that, although there exists evidence of clustering up to

$z \sim 2.2$ (apart from a gap at $0.5 < z < 1$), there is significantly more clustering at lower redshifts ($z < 0.5$) than higher ones. This finding is especially significant when taking into account the fact that, due to the survey gathering data from runs of the sky, more total volume of space is considered as redshift increases. We would therefore expect an equal if not larger amount of clustering at higher redshifts, yet we see the opposite. A possible explanation for this trend is that quasars along with their host galaxies need time to be able to interact gravitationally in order to cluster in higher-density areas, explaining the greater amount of clustering in the closest redshifts.

We would like to thank Shyamal Mitra, our mentor in the Geometry of Space research group at the University of Texas at Austin, for his support and guidance in developing our research.

REFERENCES

Bahcall, N. A. 1988, Annual Review of Astronomy and Astrophysics, 26, 631, doi: 10.1146/annurev.aa.26.090188.003215

Chu, Y., & Zhu, X. 1983, The Astrophysical Journal, 267, 4, doi: 10.1086/160839

Pedregosa, F., Varoquaux, G., Gramfort, A., et al. 2011, Journal of Machine Learning Research, 12, 2825

Ross, N. P., Shen, Y., Strauss, M. A., et al. 2009, The Astrophysical Journal, 697, 1634, doi: 10.1088/0004-637X/697/2/1634

Schneider, D. P., Fan, X., Hall, P. B., et al. 2003, The Astronomical Journal, 126, 2579, doi: 10.1086/379174

Shen, Y., Strauss, M. A., Ross, N. P., et al. 2009, The Astrophysical Journal, 697, 1656, doi: 10.1088/0004-637X/697/2/1656