Protocluster Confirmation at Redshifts 3-4

Abigail Moreno[a,1], Lawrence Nhan[a,1], Grey Murphree[a]

[a]*Institute for Astronomy, University of Hawai'i, 2680 Woodlawn Drive, Honolulu, HI 96822, USA*
[b]*Department of Physics, University of California, Riverside, CA, USA*

Abstract

Using the Hawaii Two-0 (H20) Survey we succeeded to find and analyze protoclusters at redshifts 3-4 while also creating results and coding packages that ensure repeatability as well as scalability. We analyzed the current H20 spectroscopy catalog and selected protocluster candidates using the spatial and redshift separation criteria in Toshikawa et al. (2016). The results yielded 5 confirmed protoclusters out of 50 photometrically determined mask candidates. We also found that stellar mass, UV magnitude, and Star Formation Rates (SFR) proved to show no significant distinction within a protocluster versus local field galaxies.

Keywords: redshift, protoclusters, H20

1. Introduction

According to the Lambda Cold Dark Matter (ΛCDM) model for cosmology, at the beginning of the universe matter was spread in a mostly homogeneous manner with some slight over densities. Following the Big Bang, matter and radiation were coupled, opposing density fluctuations by radiation pressure. Dark matter did not interact with radiation, leading to the development of gravitationally bound potential wells. Once atoms cooled, and matter and radiation decoupled, matter was attracted to these wells, leading to the distinct large scale structure (LSS) we observe today at $z \sim 0$ (Peebles, 1980). As the universe has aged, these slight over densities caused the development web like regions refereed to as a *cosmic web* (hierarchical theory of structure) of regions with high density, and other void like regions (Ramakrishnan et al., 2024; Eisenhardt et al., 2008; Rykoff et al., 2014; Tempel et al., 2014; Bleem et al., 2015; Rykoff et al., 2016; Malavasi et al., 2017).

The importance of these webbed structures can be seen in the effects and differences in field galaxies and galaxies that are part of larger galaxy clusters. This difference in environment affects properties such as stellar mass, star formation rates (SFR), morphology, velocity dispersion, and metallicity (Forrest et al., 2024).

Higher densities regions typically house clusters or groups of galaxies which tend to be larger, older, and often no longer star forming. At high redshift, we can observe the progenitors of these galaxy clusters, known as protoclusters.

Protoclusters are groups of galaxies observed at high redshift that, over time, will evolve into galaxy clusters. Though many definitions are used scientifically, galaxy clusters can be described as 'clusters' of many gravitationally bound galaxies, which are virialized with masses $\geq 10^{14}$ M/ $M_{\odot}$ at $z \sim 0$ (Overzier, 2016).

Identifying and measuring protoclusters is crucial for advancing our understanding of galaxy cluster formation and the early universe. Our research builds on Toshikawa et al. (2016) presenting data from a large and deeper survey in hopes of confirming and observing a larger number of protoclusters.

Finding and understanding the properties of protoclusters is important to the development of our understanding of the universe, and aids our understanding in the tools and observational techniques involved in measuring objects at higher redshifts. Among the methods of obtaining the observational data needed to confirm protoclusters, spectral data is the most accurate but also the most costly.

The data for our study comes from the H20 spectroscopy catalog, which contain redshift measurements and associated confidence levels. Although spectroscopy is more time-consuming and costly compared to photometry, it provides significantly more accurate

redshift measurements. Consequently, the data points selected for our analysis were chosen based on photometric evidence suggesting the presence of protoclusters. When taking the high-accuracy spectroscopic data, the data is organized in masks derived from the photometric data, representing groups of galaxies that are hypothesized to contain protoclusters. Our research group's work differs from Toshikawa et al. (2016) by focusing on a larger scale and performing a more detailed investigation of confirmed protoclusters.

2. Method

2.1. Data / Mask Selection

Data was chosen from the H20 photometric catalogs (EDF-N and EDF-F). We focused on g-dropout galaxies, which tend to be star-forming galaxies at $z \sim 3-4$. From the g-dropout galaxies in this catalog, candidate galaxies were chosen to be examined with spectra in groups called masks using DEIMOS. Each mask targets an overdensity of g-dropouts, estimated using the methodology in Chartab et al. (2020). Among the masks other low redshift galaxies ($z \sim 0-1$) were chosen to be examined as filler sources as shown in 1.

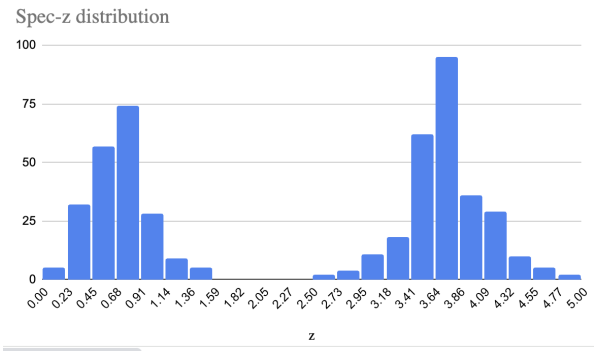


Figure 1: Spectroscopic redshift distribution for the galaxies examined through DEIMOS spectroscopy. Distribution shows a bimodal distribution showing low redshift filler galaxies at $z \sim 0$ to 2 and g-dropouts at $z \sim 3-4$.

Within the data selection, we have included three masks which were previously confirmed with DEIMOS spectra to have protoclusters, and followed up with MOSFIRE spectroscopy to measure their metallicities. These mask stand as controls to confirm our spectra observations and redshift calculations.

2.2. Redshift Calculation

Once we identified the masks and observed the data using spectra, we calculated for redshift using the Lyman-alpha (Lyα) emission line.

$$z = \frac{\lambda_{\text{obs}} - \lambda_{\text{rest}}}{\lambda_{\text{rest}}} \tag{1}$$

The quality of the results were rated based on how clear the line was on a scaled from 1-4, with 4 being the most clear candidates.

Using the H20 photometric catalogs we visualized the distribution of g-dropout galaxies redshifts to be bimodal, as we have both galaxies in masks at redshifts $\sim$ 3-4 and low redshift filler galaxies at redshifts $\sim$0-2. This distribution can be seen in figure 1.

2.3. Code Development

The logic of the code has many different functions to display and understand the results of our spectral data. For this data we developed an object oriented code that utilized a string based data entry and divided the data into different data sets by mask ID.

2.3.1. Grouping by Mask

Utilizing the grouping function, we have developed multiple sub-functions to display and compare spectral and photometric data. The code provides several visualization options for the spectral data, including tabular formats, histograms, and scatter plots. Furthermore, by employing the Python package SkyCoord, we identified the nearest Right Ascension (RA) and Declination (Dec) coordinates between the spectral and photometric datasets. These coordinates were then plotted on a scatter plot to highlight the discrepancies between the datasets. Small coordinate offsets are expected in this comparison, as the galaxies were selected from a previous version of the H2O photometry catalogs.

2.3.2. Finding Protoclusters

The second primary function of the code was to determine whether each grouped mask contained a protocluster. This was achieved by iterating through the masks and appending rows of data with sufficiently close redshifts (differences ≤ 0.01) to tables within a dictionary of mask IDs (Toshikawa et al., 2016). Subsequently, we provided visualizations of this data in the form of histograms to facilitate the analysis of redshift distributions.

2.3.3. *Confirming Protoclusters*

Next, we sought to confirm the protocluster candidates identified as sufficiently close in spatial coordinates using the methods outlined in Toshikawa et al. (2016), specifically by examining the galaxy density surrounding the protocluster. This was accomplished using the kdeplot function from the Seaborn library. We identified the g-dropout photometric data within a 0.25-degree radius of each protocluster's center and generated a kernel density estimate (KDE) plot, scaled by levels of σ from the mean. Over this KDE plot, we overlaid a scatter plot representing the protocluster.

To facilitate the interpretation of these results, we employed the Python package Minisom to train on H20's photometric data and project COSMOS data onto a self-organizing map (SOM). This visualization helps illustrate the distribution of our data. An example of this distribution significance can be seen in 4.

Furthermore, to assess the significance of protocluster occurrence in this region, the code performs 100,000 iterations, each selecting random samples from the SOM distribution. The code then calculates and returns the probability and significance of each protocluster's occurrence.

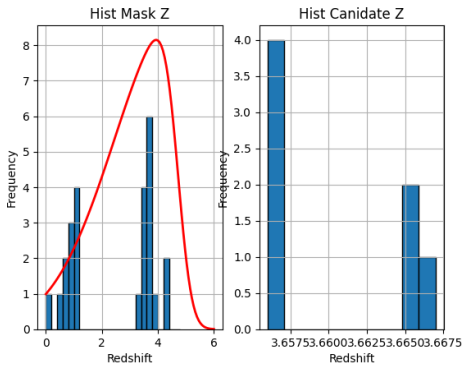


Figure 4: SOM distribution result shown overlaid onto mask distribution on mask Fgc601

3. Results

From the development of the code we were able to confirm 5 protoclusters from a set of 50 masks, or candidate protoclusters. The confirmed protoclusters then were analyzed to learn more of the properties of protoclusters.

3.1. *Identified Protoclusters*

As further confirmation beyond redshift clustering is needed to confirm protoclusters, we defined necessary minimum overdensity in the protocluster region of 3σ to confirm the possible candidates that passed the redshift threshold test. Of the 5 protoclusters identified using a 0.01 redshift difference threshold Toshikawa et al., 2016, we confirmed them all to be above the baseline over density necessary ($\sim 3\sigma$)(2). Overdensity averages in for each protocluster ranged from $\sim$3.44 to 4.48 σ, with this highest max density at $\sim$7.58 σ.

Through the data observed we found some protoclusters candidates that were identified with only two candidate galaxies. Our definition of protoclusters extends to candidates with 3 or more, therefore 'protocluster pairs' must be further observed in order confirm whether they are protoclusters.

3.2. *Testing Significance*

Using our SOM-generated model to estimate the redshift distribution of the sample space, we conducted tests to estimate the chance of observing the redshift clustering of the protoclusters that we observed. This was done by randomly sampling redshifts from the SOM-generated distribution by the number of galaxies found within 10 arcmin of the center of each identified protocluster and finding the number of times these randomly selected values meet the threshold test over a defined number of tests. Out of 100,000 tests of each protocluster, the significance of the protoclusters' redshift clustering ranged from $\sim$1.5 to 3.25 σ. This distribution can be seen in reference to specific candidate redshifts (4).

3.3. *Analysing Protoclusters*

Confirmed protoclusters and 'protocluster pairs' were then analyzed by plotting mass, UV magnitude, and star formation rates (SFR) on a histogram, comparing these qualities within the protocluster (10 arcmin) and outside the protocluster (10-15 arcmin) using data from the photometric g-dropout galaxies within these distances from the center of the protocluster as specified by the spectral data (5). We found that these qualities do not seem to be significantly different within and around a protocluster. In future work, we will continue exploring the properties of protocluster galaxies to their nearby field neighbors.

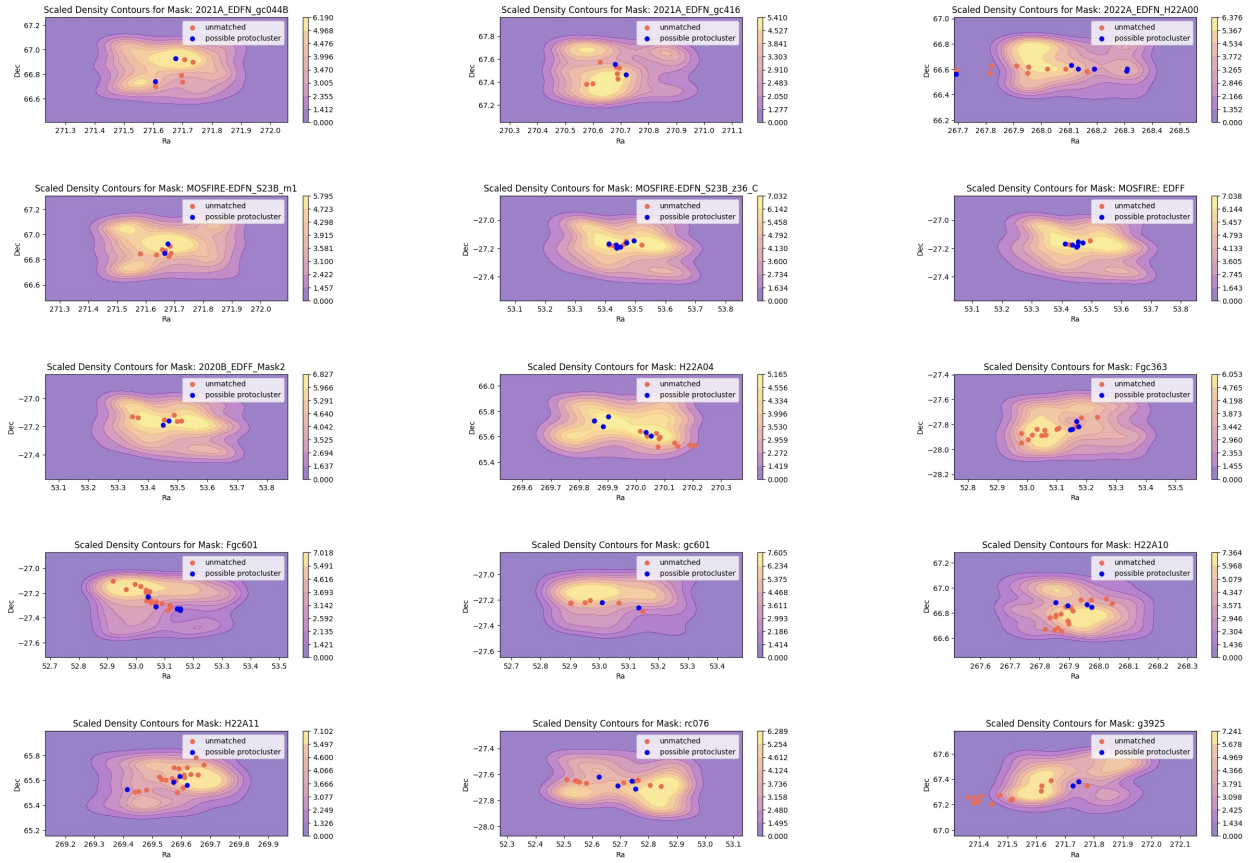
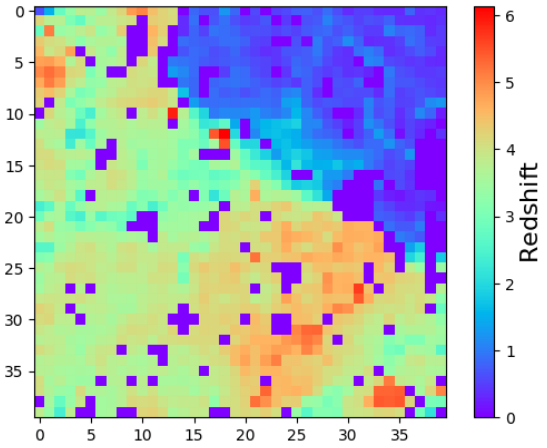


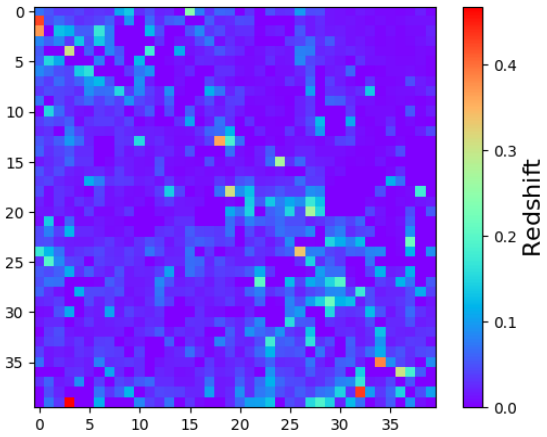
Figure 2: KDE plots for protoclusters that pass the redshift threshold defined. This plot includes 'protocluster pairs.'

Table 1: Measurements for Confirmed Protoclusters and Protocluster Pairs.

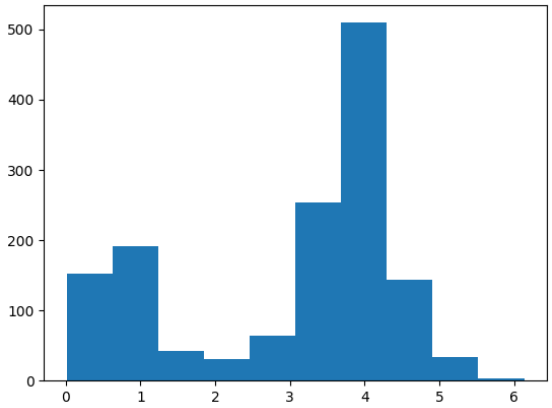
Mask	Num Observed	Num Spectroscopically Observed	Protocluster?	Protocluster members	Mean Redshift
2021A_EDFN_gc044B	67.00	7.00	maybe	2.00	3.69
2021A_EDFN_gc416	0.00	9.00	maybe	2.00	3.56
2022A_EDFN_H22A00	66.00	11.00	maybe	2.00	3.71
2022A_EDFN_H22A00_2	0.00	11.00	maybe	2.00	4.26
2022A_EDFN_H22A00_3	0.00	11.00	maybe	2.00	3.96
MOSFIRE-EDFN_S23B_m1	0.00	10.00	maybe	2.00	3.83
MOSFIRE-EDFN_S23B_z36_C	0.00	10.00	Yes	4.00	3.74
MOSFIRE-EDFN_S23B_z36_C_2	0.00	8.00	maybe	2.00	3.60
MOSFIRE: EDFF	0.00	6.00	maybe	2.00	3.73
MOSFIRE: EDFF_2	0.00	8.00	Yes	4.00	3.60
H22A04	62.00	13.00	Yes	3.00	3.80
H22A04_2	0.00	12.00	maybe	2.00	3.67
Fgc363	71.00	15.00	maybe	2.00	3.70
Fgc363_2	0.00	15.00	maybe	2.00	3.60
Fgc601	68.00	23.00	Yes	3.00	3.66
Fgc601_2	0.00	24.00	Yes	4.00	3.66
gc601	57.00	11.00	maybe	2.00	3.79
H22A10	65.00	19.00	maybe	2.00	3.57
H22A10_2	0.00	19.00	maybe	2.00	3.55
H22A11	64.00	21.00	maybe	2.00	3.62
H22A11_2	0.00	21.00	maybe	2.00	3.83
rc076	55.00	12.00	maybe	2.00	3.92
rc076_2	0.00	12.00	maybe	2.00	3.77
g3925	70.00	15.00	maybe	2.00	3.00



(a) SOM generated using the python package minisom. This SOM was trained using g-dropouts of the H20 photometric catalog

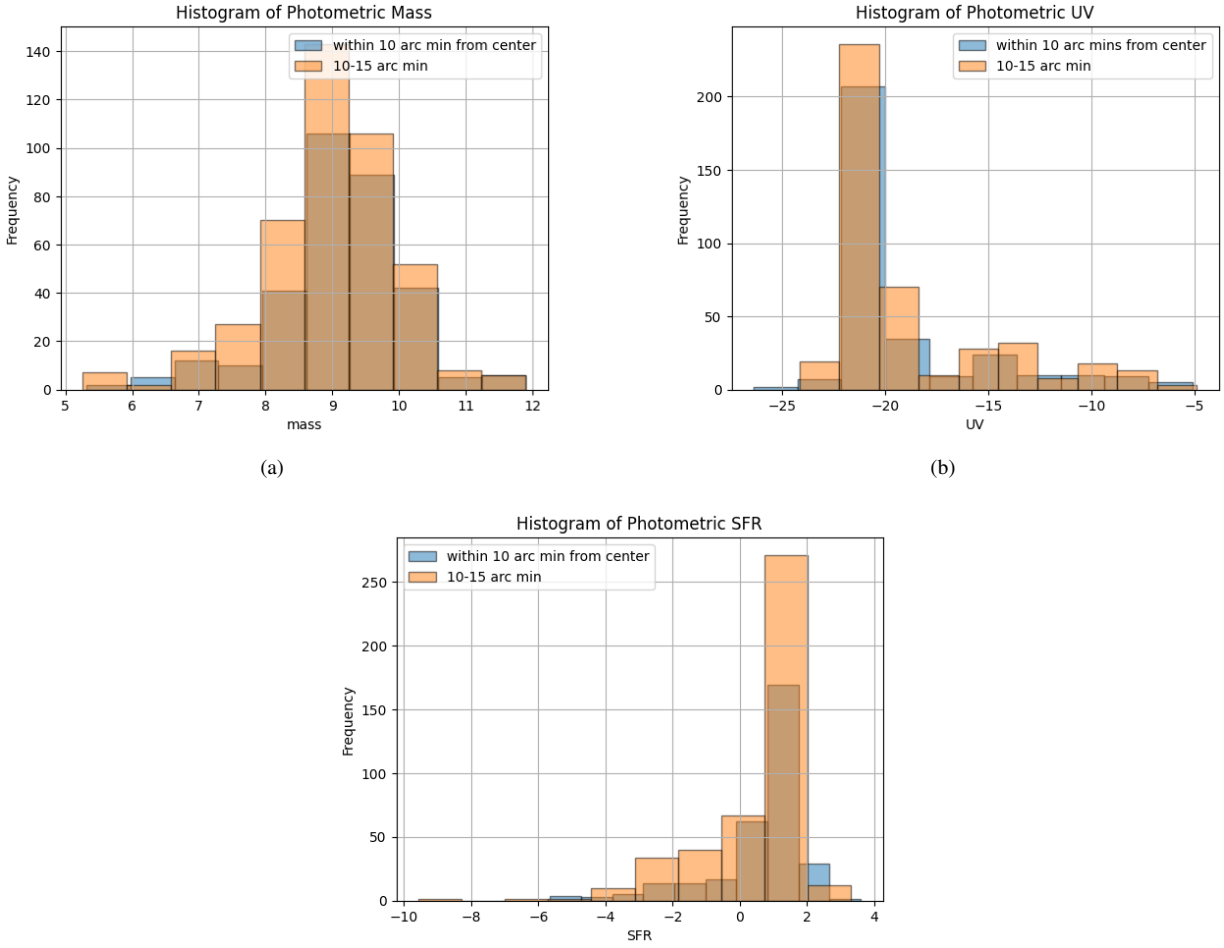


(b) Error on 3(a)



(c) Histogram created through the median redshift in each best matching unit (BMU) coordinate of the SOM in 3(a)

Figure 3



(a) (b) (c)

Figure 5: Mass 5(a), UV magnitude 5(b), and star formation rates (SFR) 5(c) plotted on histogram with photometric galaxies within 10 arcmin of the center of the protocluster found through spectra plotted in blue and 10-15 arcmin plotted in orange. Plot shows similar trends in data for each quality suggesting these qualities do not prove to be unique within protoclusters.

4. Discussion

In section 2.3 we discuss the parameters we chose to assume when confirming protoclusters throughout this paper. We assume the same two parameters for confirmation as Toshikawa et al. (2016): a 0.01 redshift threshold and minimum over density of 3σ. This is important to state when comparing protocluster classifications in different papers as not all assume these same assignment parameters which can confound similarities measurement between protoclusters. Given these constraints we display our results in comparison to other papers protoclusters masses following Casey (2016).

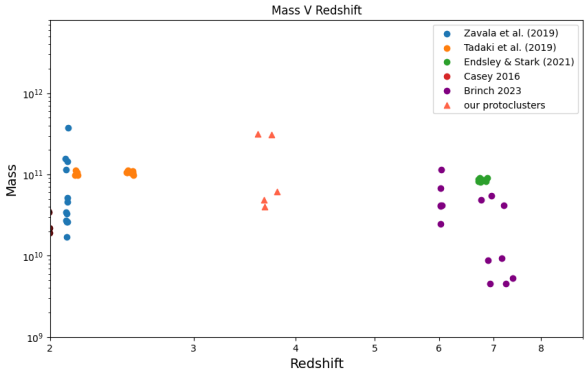


Figure 6: This figure compares the protoclusters found in this paper to others in terms of total stellar mass, modeling the same comparison done in Casey (2016). The protoclusters found in paper are shown in the orange triangles and seem to fit the trend between redshift and total stellar mass at redshifts 3-4.

In comparison to Toshikawa et al. (2020) we found protoclusters at a slightly lower redshift with less confirmed protocluster members (ranging from 3-4). We believe that with a more complete spectroscopic sample, we will be able to find more protocluster members. Overdensity however seems to be higher as Toshikawa et al. (2020) seems to have a find overdensity of $\sim 4\sigma$ while the confirmed protoclusters in this paper share an over density of $\sim 7\sigma$.

5. Conclusion

In this paper we analyzed protocluster candidates in the H20 (Hawaii Two-0) spectroscopy survey and developed an object oriented code to achieve protocluster confirmation and visualization. We confirmed candidate through a redshift threshold test then imposed a minimum over density of 3σ.

Of the 50 mask candidates, we found 24 candidates that were within a 0.01 redshift difference threshold Toshikawa et al., 2016. 19 of these were found to be protocluster pairs, and 5 were confirmed as protoclusters.

After analyzing the protocluster properties: mass, UV magnitude, and star formation rates, we found that these properties did not differ significantly to surrounding field galaxies.

To improve the the study we suggest carrying out more spectra observations over a larger degrees of the sky. In the future work we plan on expanding our spectroscopic sample with the upcoming Prime Focus Spectrograph (PFS) and follow up our confirmed protoclusters with MOSFIRE.

References

L. E. Bleem, B. Stalder, T. de Haan, K. A. Aird, S. W. Allen, D. E. Applegate, M. L. N. Ashby, M. Bautz, M. Bayliss, B. A. Benson, S. Bocquet, M. Brodwin, J. E. Carlstrom, C. L. Chang, I. Chiu, H. M. Cho, A. Clocchiatti, T. M. Crawford, A. T. Crites, S. Desai, J. P. Dietrich, M. A. Dobbs, R. J. Foley, W. R. Forman, E. M. George, M. D. Gladders, A. H. Gonzalez, N. W. Halverson, C. Hennig, H. Hoekstra, G. P. Holder, W. L. Holzapfel, J. D. Hrubes, C. Jones, R. Keisler, L. Knox, A. T. Lee, E. M. Leitch, J. Liu, M. Lueker, D. Luong-Van, A. Mantz, D. P. Marrone, M. McDonald, J. J. McMahon, S. S. Meyer, L. Mocanu, J. J. Mohr, S. S. Murray, S. Padin, C. Pryke, C. L. Reichardt, A. Rest, J. Ruel, J. E. Ruhl, B. R. Saliwanchik, A. Saro, J. T. Sayre, K. K. Schaffer, T. Schrabback, E. Shirokoff, J. Song, H. G. Spieler, S. A. Stanford, Z. Staniszewski, A. A. Stark, K. T. Story, C. W. Stubbs, K. Vanderlinde, J. D. Vieira, A. Vikhlinin, R. Williamson, O. Zahn, and A. Zenteno. Galaxy Clusters Discovered via the Sunyaev-Zel'dovich Effect in the 2500-Square-Degree SPT-SZ Survey. ApJS, 216(2):27, February 2015. doi: 10.1088/0067-0049/216/2/27.

Caitlin M. Casey. The Ubiquity of Coeval Starbursts in Massive Galaxy Cluster Progenitors. ApJ, 824(1):36, June 2016. doi: 10.3847/0004-637X/824/1/36.

Nima Chartab, Bahram Mobasher, Behnam Darvish, Steve Finkelstein, Yicheng Guo, Dritan Kodra, Kyoung-Soo Lee, Jeffrey A. Newman, Camilla Pacifici, Casey Papovich, Zahra Sattari, Abtin Shahidi, Mark E. Dickinson, Sandra M. Faber, Henry C. Ferguson, Mauro Giavalisco, and Marziye Jafariyazani. Large-scale Structures in the CANDELS Fields: The Role of the Environment in Star Formation Activity. ApJ, 890(1):7, February 2020. doi: 10.3847/1538-4357/ab61fd.

Peter R. M. Eisenhardt, Mark Brodwin, Anthony H. Gonzalez, S. Adam Stanford, Daniel Stern, Pauline Barmby, Michael J. I. Brown, Kyle Dawson, Arjun Dey, Mamoru Doi, Audrey Galametz, B. T. Jannuzi, C. S. Kochanek, Joshua Meyers, Tomoki Morokuma, and Leonidas A. Moustakas. Clusters of Galaxies in the First Half of the Universe from the IRAC Shallow Survey. ApJ, 684(2):905–932, September 2008. doi: 10.1086/590105.

Ben Forrest, Brian C. Lemaux, Ekta A. Shah, Priti Staab, Roy R. Gal, Lori M. Lubin, M. C. Cooper, Olga Cucciati, Denise Hung, Ian McConachie, Adam Muzzin, Gillian Wilson, Sandro Bardelli, Letizia P. Cassarà, Wenjun Chang, Finn Giddings, Emmet Golden-Marx, Nimish Hathi, Stephanie M. Urbano Stawinski, and Elena Zucca. Environmental Effects on the Stellar Mass Function in a z~3.3 Overdensity of Galaxies in the COSMOS Field. *arXiv e-prints*, art. arXiv:2405.18491, May 2024. doi: 10.48550/arXiv.2405.18491.

N. Malavasi, S. Arnouts, D. Vibert, S. de la Torre, T. Moutard, C. Pichon, I. Davidzon, K. Kraljic, M. Bolzonella, L. Guzzo, B. Garilli, M. Scodeggio, B. R. Granett, U. Abbas, C. Adami, D. Bottini, A. Cappi, O. Cucciati, P. Franzetti, A. Fritz, A. Iovino, J. Krywult, V. Le Brun, O. Le Fèvre, D. Maccagni, K. Małek, F. Marulli, M. Polletta, A. Pollo, L. Tasca, R. Tojeiro, D. Vergani, A. Zanichelli, J. Bel, E. Branchini, J. Coupon, G. De Lucia, Y. Dubois, A. Hawken, O. Ilbert, C. Laigle, L. Moscardini, T. Sousbie, M. Treyer, and G. Zamorani. The VIMOS Public Extragalactic Redshift Survey (VIPERS): galaxy segregation inside filaments at z = 0.7. MNRAS, 465(4):3817–3822, March 2017. doi: 10.1093/mnras/stw2864.

Roderik A. Overzier. The realm of the galaxy protoclusters. A review. A&A Rev., 24(1):14, November 2016. doi: 10.1007/s00159-016-0100-3.

P. J. E. Peebles. *The large-scale structure of the universe*. 1980.

Vandana Ramakrishnan, Kyoung-Soo Lee, Maria Celeste Artale, Eric Gawiser. Yujin Yang, Changbom Park, Robin Ciardullo, Lucia Guaita, Sang Hyeok Im, Seongjae Kim, Ankit Kumar, Jaehyun Lee, Seong-Kook Lee, Byeongha Moon, Nelson Padilla, Alexandra Pope, Roxana Popescu, Hyunmi Song, Paulina Troncoso, Francisco Valdes, and Ann Zabludoff. ODIN: Identifying Protoclusters and Cosmic Filaments Traced by Lyα-emitting Galaxies. *arXiv e-prints*, art. arXiv:2406.08645, June 2024. doi: 10.48550/arXiv.2406.08645.

E. S. Rykoff, E. Rozo, M. T. Busha, C. E. Cunha, A. Finoguenov, A. Evrard, J. Hao, B. P. Koester, A. Leauthaud, B. Nord, M. Pierre, R. Reddick, T. Sadibekova, E. S. Sheldon, and R. H. Wechsler. redMaPPer. I. Algorithm and SDSS DR8 Catalog. ApJ, 785(2):104, April 2014. doi: 10.1088/0004-637X/785/2/104.

E. S. Rykoff, E. Rozo, D. Hollowood, A. Bermeo-Hernandez, T. Jeltema, J. Mayers, A. K. Romer, P. Rooney, A. Saro, C. Vergara Cervantes, R. H. Wechsler, H. Wilcox, T. M. C. Abbott, F. B. Abdalla, S. Allam, J. Annis, A. Benoit-Lévy, G. M. Bernstein, E. Bertin,

D. Brooks, D. L. Burke, D. Capozzi, A. Carnero Rosell, M. Carrasco Kind, F. J. Castander, M. Childress, C. A. Collins, C. E. Cunha, C. B. D'Andrea, L. N. da Costa, T. M. Davis, S. Desai, H. T. Diehl, J. P. Dietrich, P. Doel, A. E. Evrard, D. A. Finley, B. Flaugher, P. Fosalba, J. Frieman, K. Glazebrook, D. A. Goldstein, D. Gruen, R. A. Gruendl, G. Gutierrez, M. Hilton, K. Honscheid, B. Hoyle, D. J. James, S. T. Kay, K. Kuehn, N. Kuropatkin, O. Lahav, G. F. Lewis, C. Lidman, M. Lima, M. A. G. Maia, R. G. Mann, J. L. Marshall, P. Martini, P. Melchior, C. J. Miller, R. Miquel, J. J. Mohr, R. C. Nichol, B. Nord, R. Ogando, A. A. Plazas, K. Reil, M. Sahlén, E. Sanchez, B. Santiago, V. Scarpine, M. Schubnell, I. Sevilla-Noarbe, R. C. Smith, M. Soares-Santos, F. Sobreira, J. P. Stott, E. Suchyta, M. E. C. Swanson, G. Tarle, D. Thomas, D. Tucker, S. Uddin, P. T. P. Viana, V. Vikram, A. R. Walker, Y. Zhang, and DES Collaboration. The RedMaPPer Galaxy Cluster Catalog From DES Science Verification Data. ApJS, 224(1):1, May 2016. doi: 10.3847/0067-0049/224/1/1.

E. Tempel, R. S. Stoica, V. J. Martínez, L. J. Liivamägi, G. Castellan, and E. Saar. Detecting filamentary pattern in the cosmic web: a catalogue of filaments for the SDSS. MNRAS, 438(4):3465–3482, March 2014. doi: 10.1093/mnras/stt2454.

Jun Toshikawa, Nobunari Kashikawa, Roderik Overzier, Matthew A. Malkan, Hisanori Furusawa, Shogo Ishikawa, Masafusa Onoue, Kazuaki Ota, Masayuki Tanaka, Yuu Niino, and Hisakazu Uchiyama. A Systematic Survey of Protoclusters at z ~3-6 in the CFHTLS Deep Fields. ApJ, 826(2):114, August 2016. doi: 10.3847/0004-637X/826/2/114.

Jun Toshikawa, Matthew A. Malkan, Nobunari Kashikawa, Roderik Overzier, Hisakazu Uchiyama, Kazuaki Ota, Shogo Ishikawa, and Kei Ito. Discovery of Protoclusters at z ~ 3.7 and 4.9: Embedded in Primordial Superclusters. ApJ, 888(2):89, January 2020. doi: 10.3847/1538-4357/ab5e85.