

University of Groningen

Large-scale Structure in CHILES Using DisPerSE

Luber, Nicholas; van Gorkom, J. H.; Hess, Kelley M.; Pisano, D. J.; Fernández, Ximena; Momjian, Emmanuel

Published in:
The Astronomical Journal

DOI:
[10.3847/1538-3881/ab1b6e](https://doi.org/10.3847/1538-3881/ab1b6e)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2019

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Luber, N., van Gorkom, J. H., Hess, K. M., Pisano, D. J., Fernández, X., & Momjian, E. (2019). Large-scale Structure in CHILES Using DisPerSE. *The Astronomical Journal*, 157(6), [254].
<https://doi.org/10.3847/1538-3881/ab1b6e>

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.



Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.



Large-scale Structure in CHILES Using DisPerSE

Nicholas Luber^{1,2,3}, J. H. van Gorkom³, Kelley M. Hess^{4,5}, D. J. Pisano^{1,2,6}, Ximena Fernández^{7,9} , and Emmanuel Momjian⁸ 

¹Department of Physics and Astronomy, West Virginia University, P.O. Box 6315, Morgantown, WV 26506, USA; nicholas.m.luber@gmail.com

²Center for Gravitational Waves and Cosmology, West Virginia University, Chestnut Ridge Research Building, Morgantown, WV 26505, USA

³Department of Astronomy, Columbia University, New York, NY 10027, USA

⁴ASTRON, the Netherlands Institute for Radio Astronomy, Postbus 2, 7990 AA, Dwingeloo, The Netherlands

⁵Kapteyn Astronomical Institute, University of Groningen, Landleven 12, 9747 AD, Groningen, The Netherlands

⁶Adjunct Astronomer at Green Bank Observatory, Green Bank, WV, USA

⁷Department of Physics and Astronomy, Rutgers, The State University of New Jersey, Piscataway, NJ 08854-8019, USA

⁸National Radio Astronomy Observatory, P.O. Box 0, Socorro, NM 87801, USA

Received 2017 October 31; revised 2019 April 16; accepted 2019 April 21; published 2019 June 4

Abstract

We demonstrate that the Discrete Persistent Source Extractor (DisPerSE) can be used with spectroscopic redshifts to define the cosmic web and its distance to galaxies in small-area deep fields. Here we analyze the use of DisPerSE to identify structure in observational data. We apply DisPerSE to the distribution of galaxies in the Cosmic Evolution Survey (COSMOS) field and find the best parameters to identify filaments. We compile a catalog of 11,500 spectroscopic redshifts from the Galaxy and Mass Assembly (GAMA) G10 data release. We analyze two-dimensional slices, extract filaments, and calculate the distance for each galaxy to its nearest filament. We find that redder and more massive galaxies are closer to filaments. To study the growth of galaxies across cosmic time, and environment, we are carrying out an HI survey covering redshifts of $z = 0\text{--}0.45$, the COSMOS HI Large Extragalactic Survey (CHILES). In addition we present the predicted HI mass fraction as a function of distance to filaments for the spectroscopically known galaxies in CHILES. Lastly we discuss the cold gas morphology of a few individual galaxies and their positions with respect to the cosmic web. The identification of the cosmic web, and the ability of CHILES to study the resolved neutral hydrogen morphologies and kinematics of galaxies, will allow future studies of the properties of neutral hydrogen in different cosmic web environments across the redshift range of $z = 0.1\text{--}0.45$.

Key words: Galaxy: evolution – large-scale structure of universe – galaxies: ISM

1. Introduction

1.1. Cosmic Web

We have long known, both in observations and simulations, that the matter in the universe is situated in an intricate system of filaments, walls, voids, and clusters, appropriately referred to as the cosmic web. The first large-scale redshift surveys undertaken revealed large underdense regions with sharp boundaries traced by string-like filaments of galaxies (de Lapparent et al. 1986). Since then, numerous surveys have been conducted with large fields of view that probe fainter and fainter magnitudes. Results from surveys such as the Sloan Digital Sky Survey (SDSS) and the 6 Degree Field Galaxy survey (6dFGS) have provided detailed insights into the large-scale matter distributions in the local universe (York et al. 2000; Jones et al. 2004). The SDSS has been used to map the distribution of galaxies, at low redshifts, with thorough completeness, and there have been successful implementations of many different methods to distinguish individual features of the cosmic web (Kitaura et al. 2009; Tempel et al. 2014).

The Cosmic Evolution Survey (COSMOS) is a galaxy survey with a much narrower field of view than 6dFGS, or SDSS that has been carried out to probe the properties of galaxies from the near universe out to redshifts of $z \approx 6$ (Scoville et al. 2007b). Both photometric and spectroscopic redshifts from the COSMOS data releases have been used extensively to identify the structures and the properties of galaxies in relation to these structures. Techniques to identify

structures, such as adaptive smoothing, have been used to obtain scale-free measures of densities from which persistent structures can be extrapolated (Scoville et al. 2007a). In this paper we explore another algorithm, the Discrete Persistence Source Extractor (DisPerSE; Sousbie 2011), for use on a subfield of COSMOS. There is some difficulty in applying structure-identifying techniques to observational data sets due to the increasing incompleteness as higher redshifts are approached, and due to errors in the measurements. However, work has been done in the comparison of different algorithms' abilities to identify the cosmic web in these large-scale observational data sets. Libeskind et al. (2018) have recently explored these different structure-identifying techniques and analyzed their abilities to recover mass distributions in the universe across different environments. They show the strengths of these different techniques in recovering different types of three-dimensional structure and nicely summarize the ability of different algorithms and the different scientific goals they can help achieve. They found that DisPerSE is well able to extract the filamentary structure in simulated data and can recover volumes and masses of the different components of the cosmic web that are in good agreement with the corresponding predictions from independent simulations. Recently, DisPerSE has also been successfully used in observational data (Kleiner et al. 2017; Kraljic et al. 2018; Laigle et al. 2018). This paper differs from the aforementioned papers because we are using a much smaller area than Kraljic et al. (2018) and Kleiner et al. (2017), and we use spectroscopic redshifts from the Galaxy and Mass Assembly (GAMA) as opposed to the photometric redshifts used in Laigle et al. (2018). In CHILES we probe a

⁹ NSF Astronomy and Astrophysics Postdoctoral Fellow.

limited redshift range with many galaxies at lower redshifts, where the uncertainties in photometric redshifts produce relatively large uncertainties in the derived large-scale structure.

While the dark matter properties of galaxies as a function of distance in the cosmic web have been well studied, with currently conflicting results, studies of the baryonic properties as a function of distance from the cosmic web have started only recently (Goh et al. 2019). Statistical studies of the local universe have shown a tight correlation between the gas content and local galaxy density. Specifically galaxies in clusters are highly HI deficient. Detailed HI imaging studies of small samples of galaxies have identified possible mechanisms of gas accretion, and removal, in voids and clusters, respectively (Chung et al. 2009; Kreckel et al. 2012; Beygu et al. 2017). Moreover, HI is being studied in galaxies residing on the filamentary network of the cosmic web (Kleiner et al. 2017) as possible observational evidence to support the cosmic web detachment (CWD) model (Aragon-Calvo et al. 2016). The CWD model theorizes that when galaxies detach from the cosmic web, they are cut off from their supply of cold gas, and the efficient phase of star formation ends. In addition, the problem of the missing baryons can be explored by observations of the gas content of the cosmic web. Although these cold gas signals would be extremely weak, it is theorized that with higher resolution telescopes and longer integration times, the gas will become detectable (Horii et al. 2017).

1.2. CHILES

The COSMOS HI Large Extragalactic Survey, CHILES, is a deep spectral line radio survey of a 40×40 arcmin area of the COSMOS field, undertaken with the Karl G. Jansky Very Large Array (VLA), to study the atomic neutral hydrogen (HI) properties of galaxies across cosmic time. CHILES is the first survey to study the properties of neutral hydrogen in individual galaxies over a continuous redshift range from $z = 0$ to $z = 0.45$. The CHILES pilot, undertaken in 2011, covering a bit less than half the redshift range from $z = 0$ to $z = 0.193$ showed that with a single pointing, and 50+ hours of integration, we could detect the atomic gas in galaxies at its highest redshift. Thus demonstrating that the large data volume and abundant presence of man-made radio frequency interference (RFI) could be adequately dealt with demonstrating that the VLA is ready to do a deep HI survey (Fernández et al. 2013). At the completion of the 1000 hr of observation, the survey will provide HI masses, and resolved morphology and kinematics for over 300 galaxies in a variety of different environments, stellar masses, and redshifts. We have reduced and imaged the first observing epoch, 180 hr. We successfully characterized HI emission at the highest redshift to date in Fernández et al. (2016), and we showed that we can derive meaningful results in the most challenging RFI environment in Hess et al. (2019). The survey will also allow for statistical measures of HI properties as a function of redshift and various properties of galaxies via stacking, as well as provide for an estimate on the measurement of Ω_{HI} as a function of redshift. In this paper, we test the use of DisPerSE, a topological algorithm that calculates density ridges in a distribution of points, which can be used to identify the large-scale structure in the CHILES volume in order to measure the gaseous properties of the cosmic web.

In Section 2 we discuss DisPerSE, the structure defining package on which this work is based. In Section 3, we discuss the

development of the catalog we use throughout the rest of the paper in both the implementation of DisPerSE, and the subsequent scientific analysis, and how we slice up the data for further calculation and visualization. In Section 4, we present the testing we did on the parameters of DisPerSE and conclude what the best parameters are. In Section 5, we present as a sanity check results on known properties of galaxies as a function of distance from filaments in the cosmic web, and the predicted HI content for galaxies in the CHILES survey. Lastly we discuss a few examples of the HI morphology found for galaxies at different locations in the cosmic web. In Section 6, we present our conclusions. Throughout this paper, we adopt a flat Λ CDM cosmology with $H_0 = 71 \text{ km s}^{-1} \text{ Mpc}^{-1}$, $\Omega_{\text{M}} = 0.27$, and $\Omega_{\Lambda} = 0.73$.

2. Methods

2.1. DisPerSE

We have chosen to use DisPerSE for defining the cosmic web because of its easy implementation on an observational data set, and the fact that it is a scale-free topological algorithm, which means that it can operate on different distance scales. There are many different ways and algorithms that people use to define large-scale structure. They are discussed at length in Libeskind et al. (2018) where it is concluded that DisPerSE sufficiently recovers the filaments of the cosmic web. DisPerSE is an algorithm that uses discrete Morse theory to calculate large-scale structure from a distribution of points, with particular focus on extracting the cosmic web in either simulations or large redshift surveys (Sousbie 2011). DisPerSE relies on Morse theory, a mathematical framework in which Morse functions, a type of twice-differentiable smooth scalar functions, which can be used to describe the topology of a region. This is possible due to the intricate relationship between a point distributions geometry and topology. The resulting description of the topology can then be used to define the structures in the point distributions. Discrete Morse theory is implemented as an alternative to Morse theory due to the fact that galaxy catalogs cannot be used to construct Morse functions with the necessary criterion. Instead, discrete Morse theory is an adaption that defines discrete functions over simplicial complexes allowing for discrete data sets, i.e., data from redshift surveys, to be properly analyzed.

A Delaunay tessellation, a technique that outlines triangles over every point in a distribution such that the area of the triangle is inversely proportional to the density of the region where that point is located, is performed on the data set and then used to calculate the density field using the Delaunay Tessellation Field Extractor (DTFE; Schaap & van de Weygaert 2000; van de Weygaert & Schaap 2009). Then, the discrete Morse–Smale complex, the intersection of the ascending and descending manifolds of discrete Morse–Smale functions, is derived from this density field to ascertain topological features with different indices corresponding to the various components of the cosmic web. The intersection is defined by a series of critical points, these points are where the gradients of the manifolds are equal to zero, and thus define the maxima or minima of that topological feature in the point distribution. The series of critical points that make up the discrete Morse–Smale complex can then be connected by a series of arcs to trace out the topological features. These features are visualized by returning the ascending 0, 1, 2, and 3

manifolds that correspond to voids, walls, filaments, and clusters, respectively.

3. Catalog Development

3.1. Compilation

In order to provide for the most complete and accurate catalog of galaxies to utilize in the implementation of DisPerSE, we construct a catalog that combines data from different collaborations. The spectroscopic redshifts come from the GAMA G10 region, a GAMA target field coincident with COSMOS (Davies et al. 2015). The stellar properties come from the COSMOS 2015 photometric data release (Laigle et al. 2016). GAMA is a collaboration that is undertaking deep spectroscopic redshift surveys in different small angular area fields. In particular, in the G10 region, GAMA is reprocessing zCOSMOS data to produce higher confidence spectroscopic redshifts (Driver et al. 2011). The COSMOS region is a 2 degree² area of the sky chosen for its lack of stellar density from our own Galaxy, and the GAMA G10 region was chosen to be coincident with this region in order that GAMA may use the vast photometry available through COSMOS. The much smaller CHILES region is similarly coincident with the COSMOS region for the same reason of vast ancillary data. However, the CHILES region is slightly off center from COSMOS to avoid any bright radio sources. We select every galaxy with a spectroscopic redshift from GAMA. This is done because the spectroscopic redshifts calculated by GAMA have been done with more sophisticated fitting techniques, and accuracy checks then what was done by zCOSMOS in their most recent release of spectroscopic redshifts (Davies et al. 2015). We select only those galaxies that GAMA has determined to have a high-quality redshift determination, for which the mean uncertainty in redshift is 0.005. We then match these redshifts with the galaxies as they appear in the COSMOS 2015 catalog. We then, if available, add ancillary data from the COSMOS 2015 photometric catalog for each galaxy (Laigle et al. 2016). This includes the stellar mass, the quiescent or star-forming classification, and the K-corrected NUV-*r* color. The stellar mass is calculated using the BC03 template (Bruzual & Charlot 2003), in which the mass is derived from a fitted stellar spectrum. The quiescent and star-forming classification was derived using the NUV-*r*/*r*-*J* relationship by the COSMOS team and the method is further elaborated in Laigle et al. (2016). There is a factor of 10 more photometric redshifts from COSMOS than there are spectroscopic ones. However, due to the increased accuracy of spectroscopic redshifts, we find the loss of data to be an acceptable trade off. The final catalog that we synthesized contains 11,500 galaxies in the COSMOS field, each with a redshift from GAMA and stellar properties from COSMOS2015.

In Figure 1, we present a comparison between the contents of the photometric catalog compiled from COSMOS2015 photometry and the catalog compiled from GAMA redshifts over the COSMOS field. We see that the photometric catalog has significantly more coverage, with redshifts for roughly a factor 5 more galaxies. Especially at the lower redshifts the increased accuracy in using spectroscopic redshifts allows for a more accurate description of the large-scale structure in the CHILES field. We show in the following sections that we are able to use DisPerSE to define the large-scale structure using spectroscopic redshifts, despite the smaller number of galaxies with spectroscopic redshifts.

3.2. Two-dimensional Slicing

To assess the quality of DisPerSE in the extraction of the cosmic web, an examination by eye is necessary to confirm if the identified filaments trace the visual patterns in the images. However, it is impossible to determine the validity of the results due to our inability to visualize three-dimensional structures on a two-dimensional surface. Although condensing the galaxies into a two-dimensional slice eliminates the calculation of physical distances from a galaxy to a filament, we are still able to calculate relative separation and maintain the ability to verify structure by eye. We use a redshift thickness of $z = 0.01$ to ensure that we are not forcing galaxies to be part of a shared geometry that is too far separated in redshift. Furthermore, we implement DisPerSE on these two-dimensional slices, rather than isolating the best-fit parameters and doing the analysis in three-dimensions, for the aforementioned reason of visual analysis. Lastly, although this work is primarily concerned with applicability to CHILES, we implement DisPerSE over the full COSMOS region for two reasons. First, a larger sky area allows for topological structures to be more easily identified. Second, this allows for future work done with COSMOS data to use the large-scale structure as defined by this paper. In future work, we plan to further test DisPerSE in three-dimensional space.

The use of spectroscopic redshifts in our study is a key difference in our definition of the cosmic web, as previous studies only use photometric redshifts. An important consideration is the difference in relative uncertainty between photometric and spectroscopic redshifts. The mean uncertainty on a spectroscopic redshift is 0.005, while for our photometric redshift the mean uncertainty is 0.01. We are working in redshift slices of a thickness of $z = 0.01$, and to use photometric redshifts that have a similar uncertainty would clearly cast a cloud of doubt over the calculated definition of the cosmic web. However, incompleteness of spectroscopic redshifts affects our results by not allowing us to accurately measure the density in areas that may be dominated by low surface brightness galaxies. Incompleteness in the spectroscopic data is a concern. The GAMA collaboration quotes a completeness down to 26 magnitudes. Figure 1(D) shows that at low to intermediate redshifts, the range probed by CHILES and future HI deep fields, the average number density of galaxies with spectroscopic galaxies goes down by only a factor 2, allowing us to be able to use spectroscopic redshifts across the entire range. For these reasons, we use spectroscopic redshifts in order to ensure that we recover meaningful structures, suitable for physical interpretation.

In Figure 2, we illustrate the inaccessible nature of the visualization of a three-dimensional display of the cosmic web. Although there are clearly defined filaments in the image, it is impossible to distinguish by eye which galaxies do or do not belong to these filaments. For an example of a two-dimensional rendering of this same redshift slice with the same parameters, see Figure 5. In Figures 2, 3, 5, and 6 we choose to display the redshift slice of 0.21–0.22 because of its variety of large-scale structures while also lacking large over- or underdensities. We believe the results in this slice to be an accurate representation of DisPerSE’s abilities. Additionally, it falls near the middle of the CHILES redshift coverage, which is important as it underscores the success of DisPerSE to describe the data set that we, as the CHILES collaboration, are interested in. The

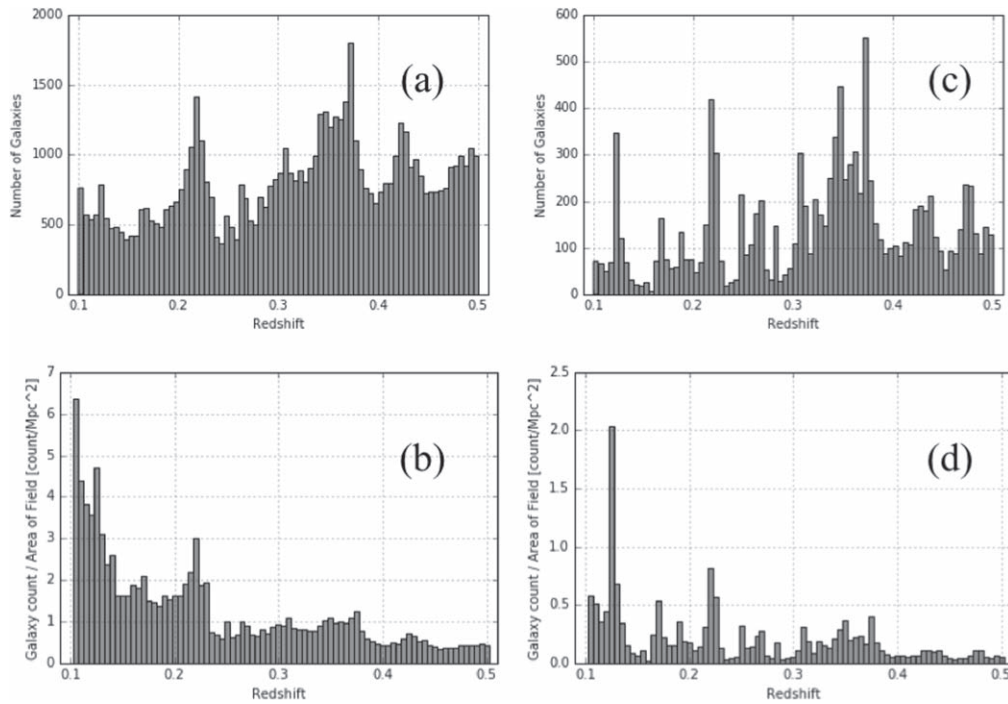


Figure 1. Quantitative description of the photometric and spectroscopic galaxy catalogs across redshifts in bins of thickness $z = 0.01$. (a) The distribution of the population of galaxies with photometric redshifts. (b) The number density of galaxies in physical units with photometric redshifts. (c) The distribution of the population of galaxies with spectroscopic redshifts. (d) The number density of galaxies in physical units with spectroscopic redshifts.

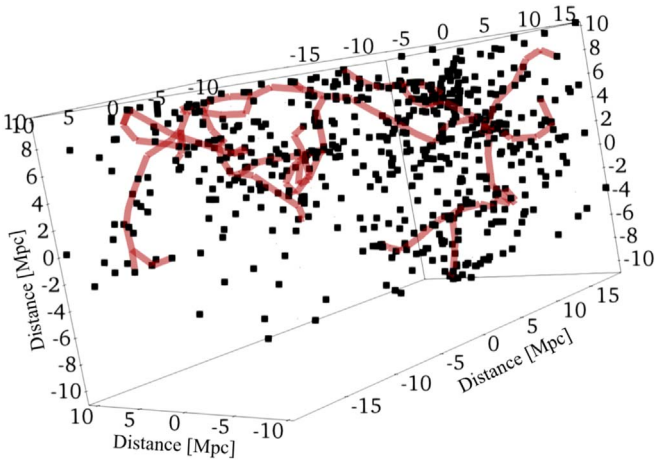


Figure 2. Illustration of the three-dimensional large-scale structure that DisPerSE extracts in the redshift range of $z = 0.21$ – 0.22 . Each point is at the location of a galaxy and the dark magenta lines indicate the extracted filaments. The axes are the physical size of the field at each redshift measured in megaparsecs.

slice shown is the entirety of the COSMOS field, because that is what was used to derive the filamentary network.

4. Using DisPerSE

4.1. Boundary Type

One challenging aspect of defining structure in the COSMOS field is the relatively small field of view. At redshifts of $z = 0.1$ – 0.2 the field of view, in physical size, is between 5 and 10 Mpc. This small scale greatly affects the identification of structure near the boundaries, the physical

edges of the field, while also having a significant impact on our understanding of the cosmic web throughout the entire field. There are four boundary types, referred to as *btype*, possible in DisPerSE, and they treat how boundary particles are added to the galaxy catalog when running the internal DisPerSE function, *delaunay_2D*, to properly calculate the DTFE around the boundaries of the field. These four boundary types are described as follows.

- (i) *Smooth*: particles are added to the boundaries based off of an interpolated density estimation.
- (ii) *Void*: there are no boundary particles added.
- (iii) *Mirror*: the particles outside the field mirror the particles on the edge.
- (iv) *Periodic*: normal periodic conditions.

Smooth was eliminated as a choice for efficiency reasons. The Delaunay tessellation under that condition, after 10 hr, showed no sign of completion and added 20+ gigabytes worth of boundary points. This is in opposition to a run time of 10 s for the other boundary conditions.

In Figures 3(a)–(c), we illustrate the effect that a change in boundary has on the filaments that DisPerSE identifies in the same redshift slice. The void boundary condition clearly fails to return logical results on the boundaries of the field, with the filaments being jagged lines that connect galaxies on opposite ends of the field. Toward the center of the image, it is clear that too many spurious filaments are created. Toward the center of the image many lines are drawn, which seem not to follow a physical filament, but rather connect two regions of enhanced density. There is a less noticeable difference between the effects that mirror and periodic boundary conditions have on the resultant filaments. However, we believe that mirror has extraneously extracted filaments in underdense regions. We

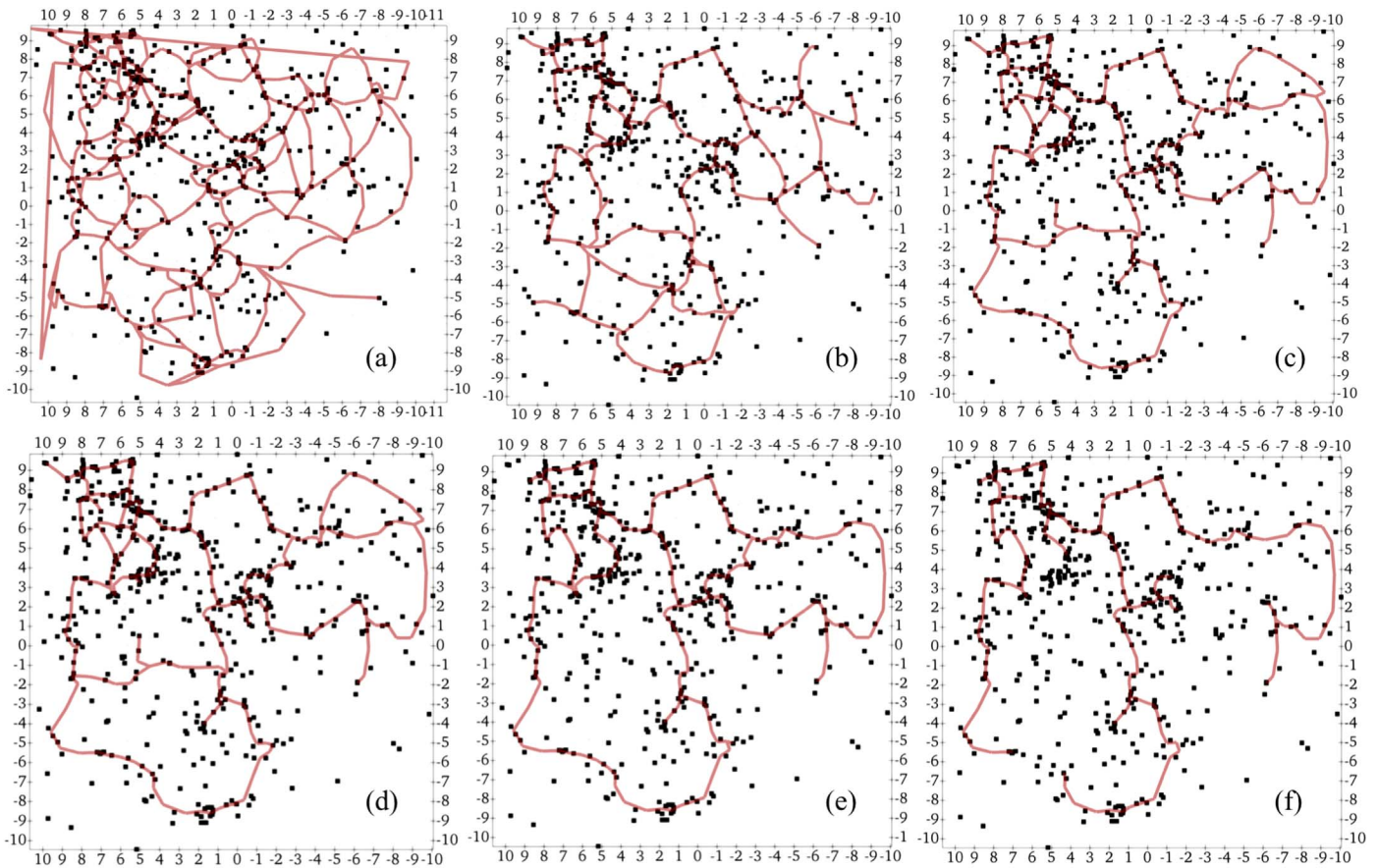


Figure 3. Extracted filamentary network for the redshift range of $z = 0.21\text{--}0.22$ overlaid on the galaxy distribution. The axes are physical size of the field at each redshift measured in megaparsecs. (a)–(c) The plots illustrate the effect that the following boundary types used for the Delaunay tessellation have on the resultant filaments: (a) void, (b) periodic, and (c) mirror. Each of these plots has an $nsig$ of 1. (d)–(f) The plots illustrate the effect that varying $nsig$ has on the filaments. (d) The filaments when $nsig$ is equal to 1. (e) The filaments when $nsig$ is equal to 3. (f) The filaments when $nsig$ is equal to 5. In each of these plots we use a mirror boundary type.

also find that when periodic is used there are increased numbers of filaments in the overdense regions of the distribution, which are useful when designating walls of galaxies. In our work we will use the periodic boundary condition.

4.2. Significance Level

The boundary condition is not the only critical parameter in DisPerSE, at least equally important is the threshold that is used to decide whether structures are significant. In the main function of DisPerSE, *mse*, there is a critical parameter referred to as $nsig$. In *mse* the user inputs the delaunay tessellation and the function returns the cosmic web. The parameter $nsig$ determines the persistence ratio threshold to use when determining whether a persistence pair is eliminated or not. A persistence ratio is the ratio of two points with consecutive indices in which a larger number indicates a stronger topological feature. A persistence pair are the two adjacent points on a topological structure used to derive the persistence ratio. Any persistence pair that has a persistence ratio less than the probability of the pair appearing in a random field will be canceled. Raising this parameter will eliminate extracted structures that may be spurious or less clearly defined by the galaxies in the field.

In Figures 3(d)–(f), we show the effects as the value for $nsig$ is varied over a galaxy distribution with all other parameters unchanged. As the parameter is raised, filaments that were previously extracted with a lower significance level are not returned. After an examination of this effect over different redshift slices, we determined that the best value for $nsig$ is 2. This is a result of the relative incompleteness of the survey at higher redshifts, which forces us to choose a lower value in order to identify any present structures. A lower value is also necessary to recover structures that may be better defined in three-dimensions but that are harder to identify in two-dimensions, such as a filament that is oriented mostly in the z -direction. Such a filament would have extension in a compressed plane and therefore the density ridges would not be as clearly defined in the two-dimensional projection.

Although the above selection process may be slightly subjective, we reached our conclusions by inspecting many different combinations of boundary conditions and thresholds. A brief summary of this selection process is simply put that the significance level of 2 both returns reliable filaments while allowing for incompleteness at higher redshifts, and that periodic boundary conditions visually trace dense regions in the images. We decided to use a single value of $nsig$ across the entire redshift range to simplify the comparison of different

redshift slices. That way any analysis of trends across cosmic time are not obfuscated by a varying n_{sig} parameter.

4.3. Identifying Filaments

The ultimate goal of our use of DisPerSE is to identify where a galaxy is located with respect to the cosmic web. We must find a way to determine the distance to a filament. Henceforth we define the cosmic web as the network of filaments, defined in DisPerSE by the ascending two manifolds. In order to understand a galaxy’s placement in the cosmic web, we developed two methods to measure a galaxy’s relative separation from the cosmic web.

Critical Point Density. In this method, we build a spheroid, in two-dimensions this is a circle and in three-dimensions a sphere, of a specified diameter, which grows larger as a function of redshift, around each critical point that comprises the filamentary structure. The higher the redshift the larger the diameter to account for the greater distance between critical points. Then, for each galaxy, we add up how many of these spheroids the galaxy falls within. One strength of this method is that a threshold for the number of critical points near a galaxy can be further used to assess the large-scale structure. For example, a galaxy that is a neighbor of 10+ critical points is likely to be a member of a wall or a cluster. However, one weakness of this method is that it does not return any information about the physical distances between an individual galaxy and the cosmic web on the two-dimensional projection.

Distance to Filament. In this method, we take a galaxy and calculate the distance between the galaxy and the nearest critical point. By distance, we refer to the angular distance between a galaxy and the closest filament in the cosmic web. This is calculated using the scale factor as it would be for the central redshift of each slice. However, this method introduces an uncertainty in the distance measured, such that it assumes that a galaxy is closest to the cosmic web at the point of the nearest critical point, and not simply any point along the filament. Fortunately, this error is bounded by half the distance between the critical point used and the next closest one, which is a small distance, as the critical points sample the filamentary network well. The strength of this method is that despite the small errors it introduces, it produces meaningful distances between a galaxy and the cosmic web. One weakness is that the distances between galaxies and the cosmic web grow systematically larger at higher redshift due to the fact that the critical points have larger separation. This is a result of increasing incompleteness in the COSMOS data set at larger redshifts. The incompleteness creates lower spatial resolution in the output of DisPerSE. As the redshift slices become under-sampled, the filamentary network extracted is not indicative of the entirety of the cosmic web. The relative error on this measurement ranges from 0.1% to 5%.

In Figure 4 we present the distances of the galaxies with redshifts between 0.1 and 0.5 to filaments in the COSMOS field. We find a clear logarithmic relationship between number of galaxies and distance from the cosmic web. Moreover, considering that recent simulation work has shown that typical filament width varies from 0.1 to 1 Mpc (Gheller et al. 2016), we calculate the number of galaxies that are within 1 Mpc of the cosmic web and find that 80% of the galaxies in the volume are within these bounds. This is reasonable, insofar as we

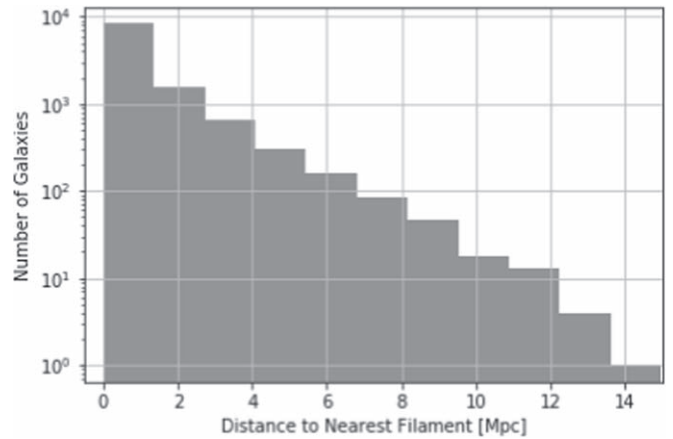


Figure 4. Histogram of the galaxies in the redshift range of $z = 0.1-0.5$ as a function of distance from the cosmic web.

expect most galaxies to reside within the filamentary structures of the cosmic web as opposed to being within voids.

When calculating distances we return both the critical point density and distance to the filament for each galaxy. The combination of these will be used in future work to create further distinctions between different components of the cosmic web at different redshifts. We produce images that have both these quantities as attributes for each galaxy in the redshift slice. In Figure 5, we illustrate these two methods and the features they identify on the same data set. The result of both methods shows a distinction in different components of the cosmic web. The critical point density method, as seen in the upper panel of Figure 5, clearly highlights overdense regions as opposed to the entire filamentary network. This is illustrated by the steep color gradient seen only in overdense regions, where multiple filaments meet. All other points are of almost uniform in color. The distance method, as seen in the bottom panel of Figure 5, more clearly traces the individual filaments and properly illustrates the transition from filaments to the underdense regions. This is evident by the smooth color gradient from the filamentary network into the underdense regions.

4.4. Final Procedure Optimized for COSMOS

To identify the spine of the cosmic web, we use the following procedure.

- (i) *Split the Data.* We take our master galaxy catalog, provide lower and upper redshifts with a separation of $z = 0.01$, and put all of the galaxies in a given redshift bin in a separate subcatalog.
- (ii) *Use DisPerSE Methods.* We run a Delaunay tessellation on each redshift slice in the data set with the boundary condition set to periodic. The output tessellation is then used as an input in the *mse* function, where we set $n_{\text{sig}} = 2$ and return the filamentary network of the cosmic web in sky coordinates.
- (iii) *Convert Coordinates.* We then translate these original galaxy catalogs and the returned filaments to the center of the COSMOS field and convert the coordinates into physical distances using the scale factor for angular distance—physical distance that is determined by the median redshift in the slice.

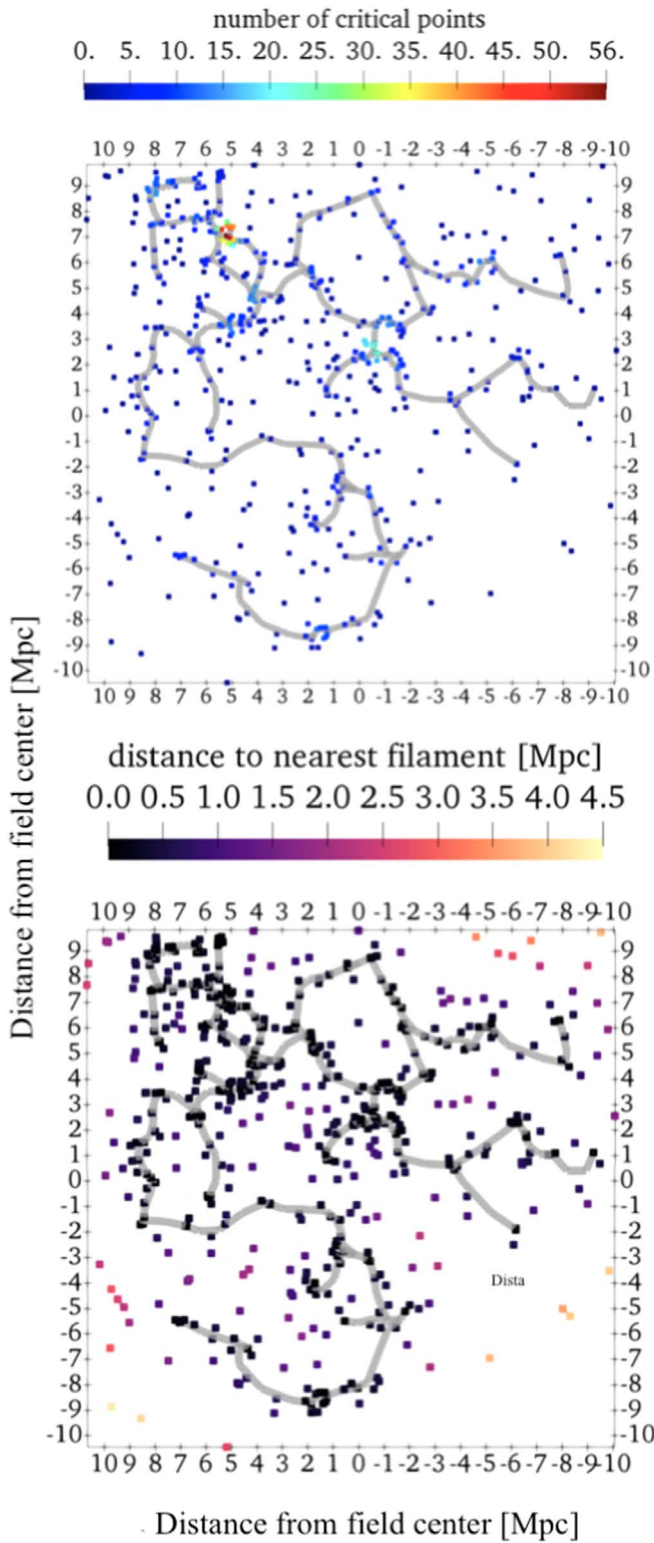


Figure 5. Extracted filamentary network for the redshift range of 0.21–0.22 overlaid on the galaxy distribution. Top: the color scale for the galaxies corresponds to a number of neighboring critical points as described in the method the *Number of Critical Points* in Section 4.3. Bottom: the color scale for the galaxies corresponds to the distance from the filament as measured by the method described in the *Distance to Filament* in Section 4.3.

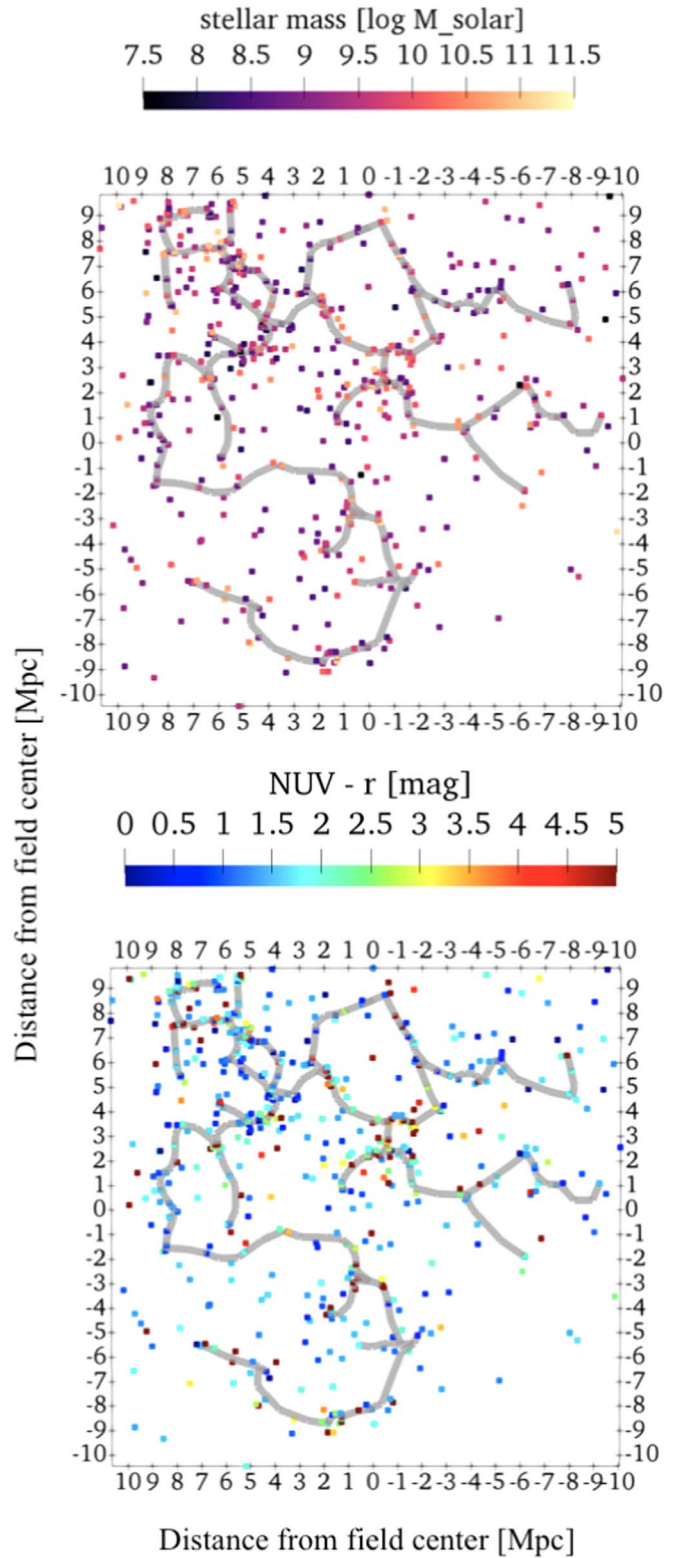


Figure 6. Extracted filamentary network for the redshift range of 0.21–0.22 overlaid on the galaxy distribution. Top: the color scale for the galaxies corresponds to the log stellar mass of each galaxy. Bottom: the color scale for the galaxies corresponds to the $NUV - r$ color for each galaxy.

(iv) *Calculate a Galaxy’s Position.* Then, we proceed to run the filament distance calculators and append both the number of critical points as well as the returned physical distances.

(v) *Final Product.* We then return an image file for visualization for each galaxy that includes the log stellar mass, color, star-forming classification, and cosmic web strength.

5. Applications

5.1. Procedure for Analysis

With the procedure for defining large-scale structure defined in the previous sections, we now examine the physical properties of galaxies as a function of their placement within the large-scale structure. Several large galaxy surveys have analyzed the observed properties of galaxies as a function of position within the cosmic web (Guo et al. 2015; Alpaslan et al. 2016). In this section we do the same for the galaxies in the COSMOS field to verify that our method provides sensible results, while also identifying any interesting trends. We also then discuss the predicted HI properties of the galaxies within the CHILES volume, and how they vary with placement within the cosmic web.

In order to display relationships with distance to the cosmic web, we had to present the distance axis in such a way that it reflects the changing scales of the extracted filaments and the increased distance between the critical points that define the cosmic web. This increased distance between critical points is a result of incompleteness increasing at higher redshifts. The filamentary network is sampled by the same number of critical points, but their intrinsic spacing increases as the physical field of view increases with redshift. This is as a result of the lower number of galaxies due to the magnitude limits of a survey. We address this by normalizing each distance by the average separation of all the critical points and their nearest neighbors. This allows all of the galaxies' relative distances to be compared to one another in a physically more meaningful way.

When analyzing the cosmic web across the whole redshift range, we proceed along the following method. We run DisPerSE, with all of the preceding specifications, and then calculate the distance of a galaxy to the filament, as described in Section 4.3. However, we only save the filament distance data for the galaxies in the inner half of the redshift range, for example, for the redshift slice of $z = 0.21\text{--}0.22$ we only append the galaxies with $z \geq 0.2125$ and $z < 0.2175$. We do this because the geometry of galaxies at the edge of the slice may be more effected by the distribution of galaxies that are not included within the slice. For example, a galaxy at redshift 0.2105 will be more affected by the galaxies at 0.2095 than those at 0.2175. The incrementing of slices gives a better description of the geometry around a galaxy.

5.2. Color and Mass versus Environment

We will now show some results on the optically derived properties of spectroscopically known galaxies in COSMOS as a function of location in the cosmic web, derived using DisPerSE. In Figure 6, we present images of the cosmic web at a redshift of $z = 0.21\text{--}0.22$ in order to display a typical output of our method for any redshift slice.

Color versus Stellar Mass. Figures 7(a) shows the color–mass relationship for the COSMOS field galaxies. The plot shows a strong correlation between the average of these two properties and it also shows that above a log stellar mass of 9.5 the redder galaxies dominate the densities of the bins, thus causing the clear upward trend. Similar analysis of the properties of galaxies in large redshift surveys have shown this. For example, the VISIBLE Multi-Object Spectrograph (VIMOS) Public Extragalactic Redshift Survey studied massive galaxies with intermediate redshifts and found a similar significant trend indicating that galaxies with

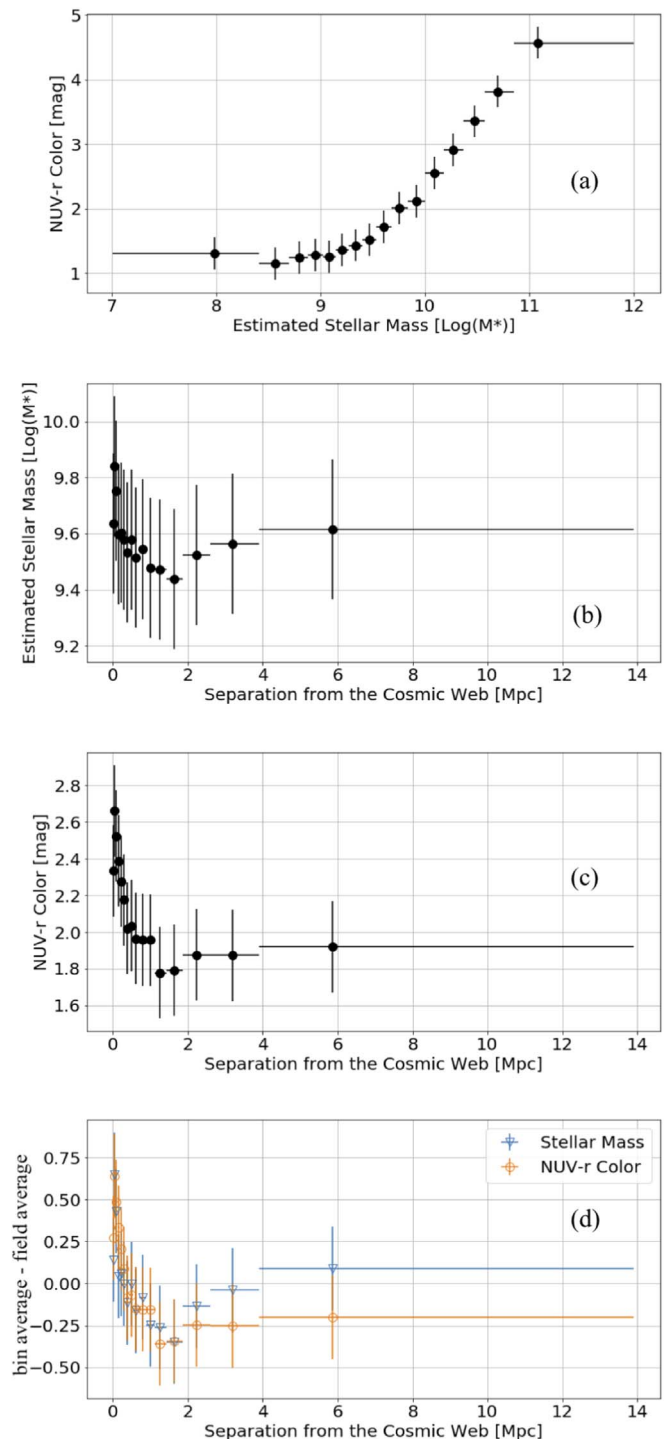


Figure 7. (a) Here we illustrate the color vs. mass relationship by sorting the galaxies by stellar mass, binning them into equal bins, and then finding the average NUV $- r$ color and log stellar mass. (b)–(c) Here we sorted the galaxies by separation from the cosmic web, binned them in bins of equal number of galaxies, and found in (b) the average log stellar mass, and in (c) the average NUV $- r$ color. (d) Here we rescaled the plots in (b) and (c) and superimposed them.

higher masses are, on average, redder than the lower mass galaxies (Cucciati et al. 2017).

Distance versus Stellar Mass. Figure 7(b) shows the relation between stellar mass and distance from the cosmic web. The average mass of galaxies decreases as the distance of the galaxy from the cosmic web increases from 0 to 1.5 Mpc. However,

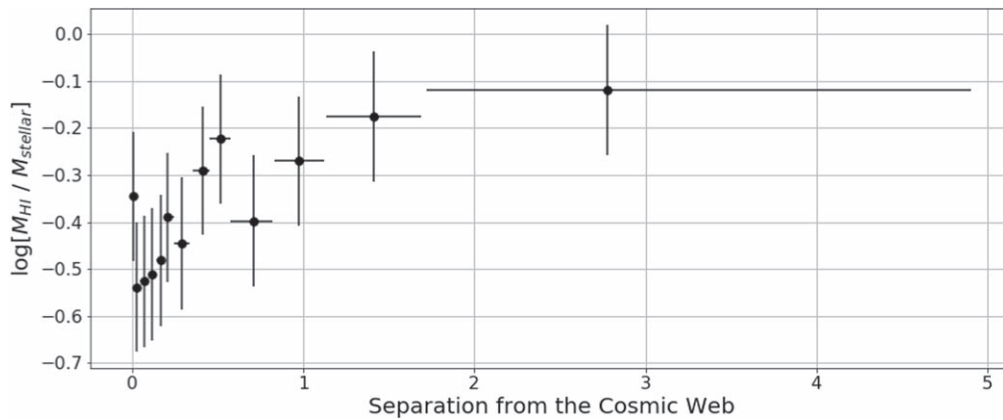


Figure 8. Here we present the relationship between neutral hydrogen to the stellar mass ratio and separation from the cosmic web. Each point represents the average distance from the cosmic web for the 28 galaxies that are considered detectable by CHILES.

after a separation of 1.5 Mpc, a galaxy’s average mass plateaus. Similar results are shown in a variety of low to intermediate redshifts in SDSS, where, independent of local density of galaxies, the galaxies along density ridges or filaments are, on average, more massive (Chen et al. 2017). Another illustration of the effects that the large-scale environment has is shown in Ricciardelli et al. (2017), where stellar mass, elliptical to the spiral ratio and star formation rate, are examined for galaxies as a function of distance from the center of a void. They find that as distance increases from the center of a void the elliptical to the spiral ratio and mass increase, while the star formation rate decreases. In Kraljic et al. (2018), a similar trend in mass and color as a function of distance from the cosmic web is found at lower redshifts, $z \approx 0.2$ in the G12 region.

Distance versus Color. Figures 7(c) illustrates a correlation between a galaxy’s color and its distance from the cosmic web. The effect of distance on $\log \text{NUV} - r$ color is similar to the effect distance has on \log stellar mass. We see a turn at a distance of 1 Mpc from a gradual approach toward bluer galaxies to a plateau. This turn in slope indicates that the distance bins are no longer being affected by redder galaxies on the same scale that they were before the turn. Moreover, the gradual slope before 1 Mpc indicates a smooth decrease in the fraction of red galaxies as the distance from the cosmic web increases. Similar results have been seen in filaments leading to a cluster environment, where it was shown that the fraction of blue to red galaxies increases dramatically as the distance to the filament increases (Braglia et al. 2007).

In Figure 7(d) we superimpose plots 7(b) and 7(c) with the scales changed to a relative scale. We do this to illustrate the similarities between the trends in the two properties. Both have a similar drop off in slope between 0 and 1.5 Mpc from the cosmic web and then both have a similar upturn at 2 Mpc from the cosmic web. The mechanism causing this upkick is unclear, but the fact that we see it in both properties could indicate either an observational bias or a physical mechanism at work.

In summary, a synthesis of previous work and our work with DisPerSE and the COSMOS data set, indicate that redder galaxies are more massive and have less separation from the cosmic web while blue galaxies tend toward the less massive side and are observed across both filaments and underdensities. However, this statement relies on averages of galaxy properties and the detailed study of individual galaxies, and the cosmic web environment is crucial to understand the role large-scale

structure plays in galaxy evolution. In the future we will track these properties across cosmic time.

5.3. Preliminary HI Results

The first 178 hr of CHILES data have been reduced and imaged. In these first 178 hr, we have detected and analyzed the properties of 16 galaxies at redshifts of $z \approx 0.12$ and $z \approx 0.17$, a redshift range which is strongly affected by RFI to demonstrate what we will be able to do with the data (Hess et al. 2019). In Hess et al. (2019), we identified the cosmic web in the procedure outlined in this paper and examined their properties with respect to the cosmic web. The majority of the HI detections appear to lie within 100 kpc of a filament. Some of the detections show interesting alignment of the HI distribution of the galaxies and the orientation of the filaments. The most massive group of five galaxies is in the vicinity of an intersection of filaments, which can be characterized as a wall of galaxies. This is in good agreement with cosmological simulations that predict the most massive galaxies to be in dense environments. The two most isolated galaxies have very extended HI and a relatively low star formation rate, in good agreement with what has been found in surveys of void galaxies (Kreckel et al. 2012).

The CHILES collaboration recently published the detection of neutral hydrogen in a star-bursting galaxy at redshift $z = 0.376$, COSMOS J100054.83+023126.2, the highest redshift HI emission detection to date (Fernández et al. 2016). Figure 1(b) of Fernández et al. (2016) shows that the HI morphology is asymmetric within the galaxy and very extended to the south. This extension is aligned with the cosmic web as identified by DisPerSE. The galaxy lies on a north–south filament and the gas distribution is reflective of this. Additionally, the high concentration of gas to the south of the galaxy is pointing toward an intersection of filaments and high density region of galaxies perhaps indicative of some interaction between the detection and the wall.

The small sample of galaxies currently detected in CHILES, the 16 galaxies in Hess et al. (2019), and the one in Fernández et al. (2016), are too small to provide significant insights, but as HI is detected in more galaxies across cosmic time, these observations of HI morphology and placement in the cosmic web could allow for a deeper understanding of the evolution of galaxies. Further work will be done to characterize HI morphology quantitatively to more deeply analyze the effects

of large-scale structures on HI perturbation and galaxy evolution.

In Figure 8 we show what we can expect to find in the CHILES survey. It shows the average neutral hydrogen mass to the stellar mass ratio of galaxies detectable by CHILES as a function of separation from the cosmic web. For this figure we calculated photometric gas fractions taking galaxies that have both SDSS spectra and *GALEX* colors, calculating stellar masses using Bell et al. (2003) and *g*, *r*, and *i* magnitudes from SDSS. We then adapted results from Catinella et al. (2010) by taking a color cut at $NUV - 3 = 3.4$ mag and assuming an HI to stellar mass fraction of 28% for blue galaxies and 1.6% for red galaxies. We use a 5σ detection limit which is calculated assuming a profile width of 150 km s^{-1} , a 1000 hr on source, and perfect Gaussian noise. We only consider galaxies within the full width half power of the primary beam and we correct for the primary beam response. With this criterion we consider 364 galaxies to be detectable by CHILES, and are therefore used in the creation of this plot. We show that we expect to find that on average, as galaxies become farther from the cosmic web, their HI to stellar mass ratio increases. Local surveys have found conflicting results on this (see, for example, Kleiner et al. 2017; Odekon et al. 2016).

6. Conclusions

We have shown that DisPerSE can be utilized to define the cosmic web in deep fields of a small angular area with the use of spectroscopic redshifts. This will be useful for the many upcoming small-area HI deep fields being carried out by the Meer Karoo Array Telescope (MeerKAT), the Giant Metrewave Radio Telescope (GMRT), and the Australian Square Kilometer Array Pathfinder (ASKAP).

Knowing the details of the cosmic web has many implications for our understanding of galaxies and galaxy evolution. We have developed a method to identify the cosmic web that utilizes DisPerSE and from this produced a schematic of the large-scale environment of the COSMOS field across the CHILES redshift range. A few choice observations on some trends in the COSMOS data were presented. We demonstrated the correlation between a galaxy's color, stellar mass, and displacement from the cosmic web. The correlation of red, high-mass galaxies with low cosmic web displacement and blue, low-mass galaxies with high cosmic web displacement may begin to be further investigated, and its role in galaxy evolution understood. These results are generalized and are only undertaken over a small field of view but their implications in the role of large-scale structures in galaxy evolution are critical in directing further investigations to tackle more specific issues.



As the remaining CHILES data is processed and imaged, we plan to use the filaments extracted with DisPerSE to understand the dependence of neutral hydrogen on large-scale structure. We will examine morphologies, kinematics, and HI mass to the stellar mass ratio in the different environments and identify any dependence on large-scale structure for these properties. At the completion of our 1000 hr of observation we will be able to identify statistical properties of HI in galaxies that are in different components of the cosmic web at redshifts that have never before been probed. These resolved high redshifts detections will allow for key insights in the role of large-scale

structures in galaxy evolution. Furthermore, through the efforts of stacking and image smoothing we will attempt to identify cold gas in the filaments in between galaxies, a vital step necessary to understand the role of gas and the cosmic web in the evolution of galaxies.

We acknowledge useful discussions with the entire CHILES collaboration. This work was in part supported by the National Science Foundation under grant number AST-1413102 to Columbia University. D.J.P. acknowledges partial support from National Science Foundation under grant number AST-1412578.

The National Radio Astronomy Observatory is a facility of the National Science Foundation operated under cooperative agreement by Associated Universities, Inc. X.F. is supported by an NSF-AAPF under award AST-1501342.

ORCID iDs

Ximena Fernández  <https://orcid.org/0000-0002-9259-3776>
Emmanuel Momjian  <https://orcid.org/0000-0003-3168-5922>

References

- Alpaslan, M., Grootes, M., Marcum, P. M., et al. 2016, *MNRAS*, 457, 2287
Aragon-Calvo, M. A., Neyrinck, M. C., & Silk, J. 2016, arXiv:1607.07881
Bell, E. F., McIntosh, D. H., Katz, N., & Weinberg, M. D. 2003, *ApJS*, 149, 289
Beygu, B., Peletier, R. F., van der Hulst, J. M., et al. 2017, *MNRAS*, 464, 666
Braglia, F., Pierini, D., & Böhringer, H. 2007, *A&A*, 470, 425
Bruzual, G., & Charlot, S. 2003, *MNRAS*, 344, 1000
Catinella, B., Schiminovich, D., Kauffmann, G., et al. 2010, *MNRAS*, 403, 683
Chen, Y.-C., Ho, S., Mandelbaum, R., et al. 2017, *MNRAS*, 466, 1880
Chung, A., van Gorkom, J. H., Kenney, J. D. P., Crowl, H., & Vollmer, B. 2009, *AJ*, 138, 1741
Cucciati, O., Davidzon, I., Bolzonella, M., et al. 2017, *A&A*, 602, A15
Davies, L. J. M., Driver, S. P., Robotham, A. S. G., et al. 2015, *MNRAS*, 447, 1014
de Lapparent, V., Geller, M. J., & Huchra, J. P. 1986, *ApJL*, 302, L1
Driver, S. P., Hill, D. T., Kelvin, L. S., et al. 2011, *MNRAS*, 413, 971
Fernández, X., Gim, H. B., van Gorkom, J. H., et al. 2016, *ApJL*, 824, L1
Fernández, X., van Gorkom, J. H., Hess, K. M., et al. 2013, *ApJL*, 770, L29
Gheller, C., Vazza, F., Brüngen, M., et al. 2016, *MNRAS*, 462, 448
Goh, T., Primack, J., Lee, C. T., et al. 2019, *MNRAS*, 483, 2101
Guo, Q., Tempel, E., & Libeskind, N. I. 2015, *ApJ*, 800, 112
Hess, K. M., Luber, N. M., Fernández, X., et al. 2019, *MNRAS*, 484, 2234
Horii, T., Asaba, S., Hasegawa, K., & Tashiro, H. 2017, *PASJ*, 69, 73
Jones, D. H., Saunders, W., Colless, M., et al. 2004, *MNRAS*, 355, 747
Kitaura, F. S., Jasche, J., Li, C., et al. 2009, *MNRAS*, 400, 183
Kleiner, D., Pimblett, K. A., Jones, D. H., Koribalski, B. S., & Serra, P. 2017, *MNRAS*, 466, 4692
Kraljic, K., Arnouts, S., Pichon, C., et al. 2018, *MNRAS*, 474, 547
Kreckel, K., Platen, E., Aragón-Calvo, M. A., et al. 2012, *AJ*, 144, 16
Laigle, C., McCracken, H. J., Ilbert, O., et al. 2016, *ApJS*, 224, 24
Laigle, C., Pichon, C., Arnouts, S., et al. 2018, *MNRAS*, 474, 5437
Libeskind, N. I., van de Weygaert, R., Cautun, M., et al. 2018, *MNRAS*, 473, 1195
Odekon, M. C., Koopmann, R. A., Haynes, M. P., et al. 2016, *ApJ*, 824, 110
Ricciardelli, E., Cava, A., Varela, J., & Tamone, A. 2017, *ApJL*, 846, L4
Schaap, W. E., & van de Weygaert, R. 2000, *A&A*, 363, L29
Scoville, N., Aussel, H., Benson, A., et al. 2007a, *ApJS*, 172, 150
Scoville, N., Aussel, H., Brusa, M., et al. 2007b, *ApJS*, 172, 1
Sousbie, T. 2011, *MNRAS*, 414, 350
Tempel, E., Stoica, R. S., Martínez, V. J., et al. 2014, *MNRAS*, 438, 3465
van de Weygaert, R., & Schaap, W. 2009, *LNP*, 665, 291
York, D. G., Adelman, J., Anderson, J. E., Jr., et al. 2000, *AJ*, 120, 1579