THE UNIVERSITY OF MICHIGAN
DEPARTMENT OF PHYSICS

Enhancement of Intercluster Filament Signals Through Stacking GMBCG Galaxy Cluster Pairs

By
Alex Thuc-Quyen Tran Nguyen

Supervised by
Professor Timothy McKay
&
Post-Doctoral Fellow Joerg Dietrich

A thesis submitted in partial fulfillment of the requirement for the concentration in
HONORS INTERDISCIPLINARY PHYSICS
with a focus on
ASTROPHYSICS

ANN ARBOR, MICHIGAN
APRIL 26, 2011

In Honor of Ngoc-Tu Tran

Mom, your endless sacrifices to provide me with an educational opportunity will never be forgotten.

Acknowledgments

First of all, I would like to express my gratitude toward Dr. Timothy McKay who has given me the opportunity to work on this thesis project. He exposed me to my first undergraduate research experience in astrophysics and has continued to guide me throughout my undergraduate career. His kind words and support have been invaluable to my success in both academics and extracurriculars. Without his guidance, I would never have considered pursuing an Honors concentration in Physics.

Next, I would like to thank Post-Doctoral researcher Joerg Dietrich for continuously motivating me to think critically about many challenging questions that I was attempting to answer in my thesis. He guided me through each step of my project and made sure I moved forward in the right direction. His expertise was invaluable in helping me understand many aspects about cosmology, especially galaxy clusters and dark matter filaments.

I am grateful for the tremendous help I have received from Post-Doctoral researcher Brian Nord during the editing process of my thesis. He also spent time to help me figure out a number of IDL coding issues. Brian was always there to listen to my ranting about the difficulties I encountered during the research process. I also want to thank my lab-mate Kate Miller for her endless support and encouragement, which has motivated me to complete this thesis. She was always willing to listen to me and help me think through the many problems I came across while working on the data analysis.

I would like to acknowledge Jiangang Hao for generously providing me with the GMBCG cluster catalog and other research data. Also, without the Sloan Digital Sky Survey collaboration, I would not have been able to gather the necessary information for my research. In addition, the visualization package TARA used in part of my stacking methodology was made available by the Astronomy Department of Pennsylvania State University. I am thankful for all of the scientific resources from which I have benefited.

Last but not least, I want to thank Cynthia Yoon who is near and dear to my heart for always staying by my side not only as a friend, but also as family. Her endless emotional support has made the thesis writing process much more bearable when I needed it the most.

Abstract

One of the prominent features of the large-scale structure is the formation of galaxy clusters at the intersections of dark matter filaments. Given more than 55,000 galaxy clusters in the GMBCG catalog, we devise an algorithm that select cluster pairs with high probability of possessing filament connections based on constraints of redshift and the number of closest neighbors. The pair–finder algorithm yields more than 16,000 cluster pairs with intercluster separations up to 50 Mpc. To enhance the filament signals, we stack the cluster pairs and their filament galaxies based on various criteria such as separation, redshift, and cluster richness. Significance maps of these stacks are constructed by using bootstrapping techniques and comparing the stacked signals with respect to the background galaxy density. Preliminary results of the stacking method show that intercluster filaments have overdensity at $\sim 2\sigma_s$ above the mean background density. The presence of filament signals is also verified by the cross–stacking method. Correlations between intercluster separation, or cluster richness, and the strength of filament signals are not obvious based on the constructed significance maps. We report the results from these preliminary studies and provide suggestions for future analyses which might enhance the filament contrast with respect to the background.

I. Introduction

A. Formation of Large-Scale Structure

In order to understand matter density fluctuations, one must look back in time when the universe was much denser and hotter than it is today. Immediately following the Big Bang, high-energy photons had very short mean free path, continuously scattering off of electrons. As the early universe expanded and cooled, fundamental particles in this primordial soup, such as quarks and gluons, began to form the first stable subatomic particles—protons and neutrons. While the universe continued to expand, the temperature dropped, and formations of deuterium, hydrogen and helium nuclei started taking place. During the *recombination* period, the positive ions combined with free electrons to form neutral atoms, and the photons now could travel more freely (Mukhanov 2005). The density fluctuations due to recombination are reflected by the temperature anisotropies (**Figure 1**) observed in the Cosmic Microwave Background (CMB, Komatsu et al. 2009). The photons that were released at the end of the recombination period are the same photons that we observe in the CMB radiation.

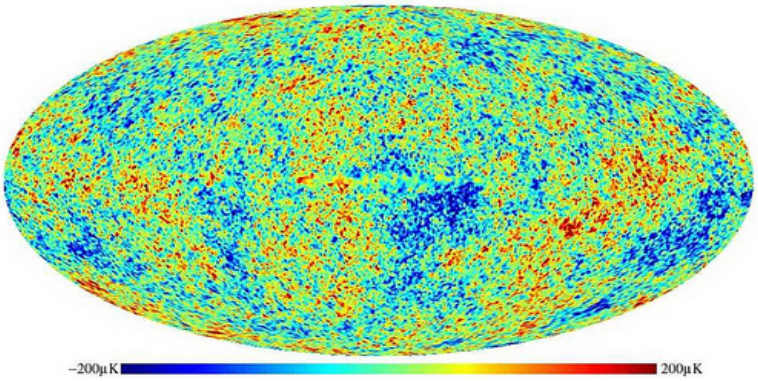

Figure 1. CMB temperature fluctuations

Studies of CMB temperature perturbations have yielded observational insights on the primordial density fluctuations in the distribution of early baryonic matter. These inhomogeneities were amplified by the presence of *nonbaryonic dark matter*, which does not interact with baryonic matter through any other means but gravity (White et al. 1987) and weak interactions (Ryden 2003). The density fluctuations of dark matter give rise to spatially varying gravitational potential that attacts matter into overdense regions. Post recombination, clumps of matter that collapse under their own gravity again attract even more matter from the less dense surroundings. As stable gravitationally bound systems form and evolve, gas collapses into stars, stars bind into galaxies, and galaxies group into clusters. Thus, the evolution of the large-scale structure that we observe today can be traced back to the gravitational instability of dark matter in the early universe (Blumenthal et al. 1984).

The assumption of the Cosmological Principle, one of the foundations of the Big Bang model, states that there is nothing special about our location in or our view of the universe. This principle implies that the universe on scales larger than ∼100 Mpc is both isotropic and homogeneous (Ryden 2003). On smaller scales, the hierarchically structured universe is comprised of gravitationally bound systems: from planets orbiting a star, to hundreds of billions of stars revolving around the center of a galaxy, and even larger collections of galaxies forming clusters. These clumps of matter, which are separated by voids, embody the inhomogeneous and anisotropic features of the large-scale structure (LSS). The snapshots of a LSS formation simulation based on the ΛCDM cosmological model (Kravtsov & National Center for Supercomputer Applications) are shown in **Figure 2**.

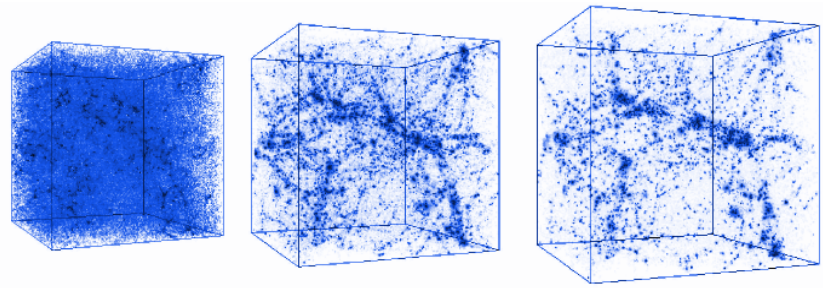

Figure 2. Simulation of large-scale structure formation over cosmic timespan

Advances in observational astronomy such as the Sloan Digital Sky Survey (SDSS) and the 2dF Galaxy Redshift Survey (2dFGRS) have allowed scientists to reconstruct a spatial 3D map of distant galaxies, which revealed walls, filaments, and voids (**Figure 3**). Several of these prominent structures include the "Great Wall" (Geller & Huchra 1989), the Sloan Great Wall (Gott et al. 2005)—one of the largest structures in the universe known to date, and a possible supervoid in the constellation Eridanus (Rudnick et al. 2007) that coincides with the Wilkinson Microwave Anisotropy Probe (WMAP) Cold Spot. In between giant bubble-like voids lie the sheets and filaments of galaxies, with galaxy clusters appearing as dense nodes. This visible network of structures is commonly referred to as the "cosmic web."

Within this web, the matter density of clusters is greater than that of filaments and much greater than that of the average density of the universe. Let the density contrast be defined as $\delta = \frac{\rho - \langle\rho\rangle}{\langle\rho\rangle}$ where $\langle\rho\rangle$ is the mean density of the background. Filaments have a density constrast of $\delta_{\text{filament}} \sim 10$, while typical galaxy clusters have a value of $\delta_{\text{cluster}} \sim 200$ within $R \approx 1\,\text{Mpc}$ (Bothun 1998). Galaxy clusters are structures that have collapsed under their own self-gravity, and that have come to equilibirium. Most of the masses of galaxies are contained in their dark matter halos (Schneider 2006). They are the largest of the density fluctuations in the universe, and gravitational instability is the mechanism responsible for their assembly.

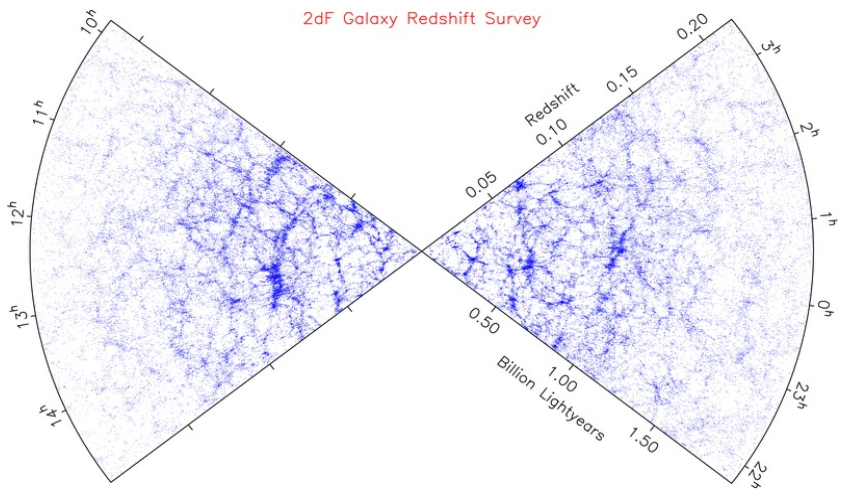

Figure 3. A slice of the universe can be reconstructed by combining redshift data and angular positions of galaxies on the sky. Clusters and voids are visible on this map.

Inside the cosmic web, the densest regions lie at the intersections of dark matter filaments (Liddle & Lyth 2000). Even though dark matter cannot be observed directly, the galaxies that form along these filaments constitute the visible thread-like structures on large scales. Since dark matter interacts with baryonic matter through gravitational attraction, at the junctions of these cosmic threads, many galaxies clump together to form clusters. Intuitively, one can imagine a vast number of cluster pairs that are connected by dark matter filaments. Observationally, the filamentary large scale structure has been revealed mostly by means of sky surveys such as SDSS and 2DFGRS. Detecting the filaments individually is still a challenge for scientists due to the filaments' low density contrast. Several other methods which have been employed to detect such filament signals include weak-leansing studies (Dietrich et al. 2005) and X-ray emission detections of the Warm-Hot Intergalatic Medium (WHIM) by the XMM–Newton observatory (Werner et al. 2008).

Yet, the interesting question remains: how does one go about using the galaxy distribution to detect filaments if individually they are difficult to observe? One of the methods that has the potential to enhance filament signals with respect to the background is the stacking of multiple filaments, provided certain constraints on their lengths and thicknesses.

This stacking method may allow us to see the overdensity of these filaments and likely reveal other structural features pertinent to their member galaxies. In this research project, computational methods are devised based on basic stacking techniques and applied to the entire GMBCG cluster catalog (Hao et al. 2010). First, cluster pairs are found based on an algorithm that takes into account constraints such as angular positions, redshifts, and numbers of closest neighbors. Following this initial step is the stacking algorithm that takes in different parameters that define the framework for the desired stacks, e.g. rotation angle, translation, number of cluster pair samples, redshift ranges, etc. Ultimately, our goal is not only to devise a robust computational method for stacking filament signals that is applicable to any cluster catalog, but also to understand the galaxy population within stacked filaments.

Together with previously established studies on the cosmic microwave background, galaxy clusters, and the 3D galaxy distribution, the study of intercluster filaments can provide more insights into our understanding of the ΛCDM cosmological model.

B. Cosmological Redshift & Distance

Due to the expansion of the universe, what astronomers are interested in measuring is the cosmological redshift of galaxy clusters, which depends solely on the *Hubble flow*, even though the peculiar velocity of a cluster can make it look closer or more distant to the observer. The cosmological redshift z of an object is directly related to the size, $R(t)$, of the universe:

$$1+z=\frac{R(t_0)}{R(t_e)}$$

where t_e is the time the light from the object was emitted, and t_0 is the time it is observed (Sparke & Gallagher 2007). Since the homogeneous universe is expanding in all directions, both the observer on Earth and a distant galaxy cluster are moving along with the Hubble flow. In cosmology, one of the distances that we are interested in calculating is the *comoving distance* that factors out the cosmic expansion (Hogg 2000). Large gravitationally bounded systems such as galaxy clusters are comoving in the Hubble flow at the same rate. By factoring out this rate of expansion, we can treat these large lumps of matter as "static" in the comoving frame of reference. Since the laws of physics remain the same regardless of coordinate choices, the comoving frame is easier to work with, and the comoving distance can tell us where the clusters are "now," even though we look back in time into the distant universe.

For a galaxy cluster with cosmological redshift z, its total comoving distance, D_c, measured from an observer on Earth is the sum of all the infinitesimal $\delta D_c(z')$ contributions between nearby events along the line of sight from $z'=0$ to $z'=z$. Let us define the function

$$E(z)\equiv\sqrt{\Omega_m(1+z)^3+\Omega_k(1+z)^2+\Omega_\Lambda}$$

where Ω_m, Ω_k, and Ω_Λ are the three density parameters of the universe at the present epoch. Ω_m is the matter (dark and baryonic) density, Ω_k, is the spatial curvature density, and Ω_Λ is the cosmological constant that represents dark energy density. In this research, we let $\Omega_m\sim 0.3$, $\Omega_\Lambda\sim 0.7$, and $\Omega_k\sim 0$. D_c is then calculated by the following integration:

$$D_c=D_H\int_0^z\frac{dz'}{E(z')}$$

where D_H, the *Hubble distance*, is defined as

$$D_H\equiv\frac{c}{H_0}=3000\,h^{-1}\,\mathrm{Mpc}.$$

The Hubble constant H_0 can be expressed with uncertainty as

$$H_0=h\times 100\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$$

where $h \sim 0.7$ according to recent measurements (Sparke & Gallagher 2007).

Another way to look at this comoving distance is to relate it to the *proper distance*, which would be measured with a ruler by the observer at the time the object is being observed. Dividing this proper distance by the ratio of the scale factor of the Universe or multiply it by $(1+z)$, we get the comoving distance.

II. Galaxy Cluster Pairs

A. Cluster Catalog

The GMBCG cluster catalog used in the following cluster pair–finder algorithm has been provided by Hao et al. (2010). It was compiled using SDSS Data Release 7 (DR7). The catalog contains more than 55,000 galaxy clusters with photometric redshifts, $z_{\rm photo}$, in the range $0.1 \lesssim z_{\rm photo} \lesssim 0.55$. **Figure 4** shows a 2D projection of the cluster distribution based on the angular positions of the clusters on the sky, i.e. right ascension (RA) and declination (DEC).

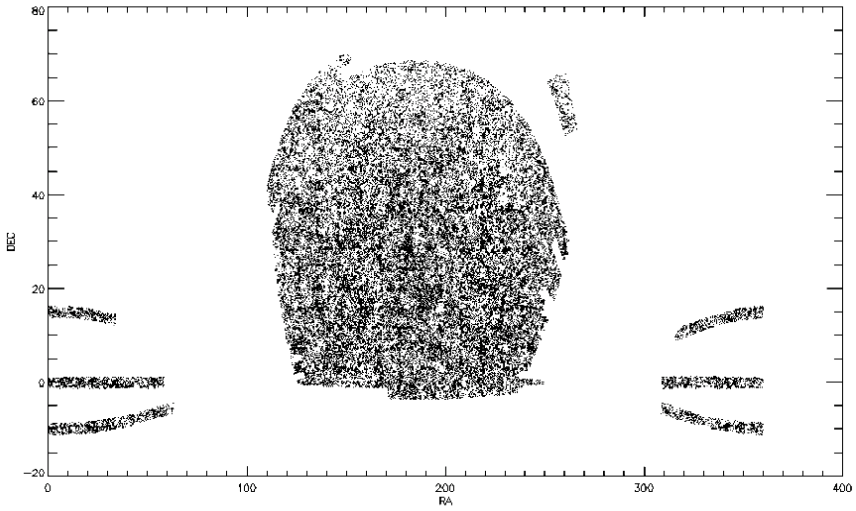

Figure 4. In the equatorial coordinate system, the cluster catalog covers a large patch and several smaller strips on the sky due to SDSS Legacy Survey's limits of coverage area.

Beside angular positions and redshifts, another property of the clusters in this catalog is their scaled richness. It is determined by the number of member galaxies brighter than $0.4L^*$ within a scaled radius measured from the brightest cluster galaxy (BCG) where where L^* is the characteristic luminosity in the Schechter luminosity function (Hao et al., 2010). The distribution of cluster richness spans the range $10 < N_{\rm gals}^{\rm scaled} < 150$, as seen in **Figure 5**. This information will have some implications in analyzing stacked filament signals, but at the moment, it is not one of the parameters crucial to the search for cluster pairs.

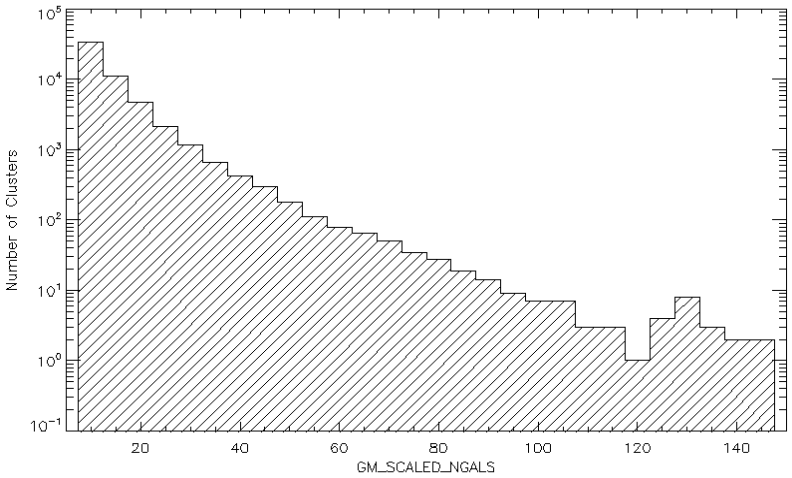

Figure 5. Distribution of cluster richness plotted in log scales.

B. Redshift Distribution

The photometric redshift distribution of the clusters in this catalog peaks at $z_{\rm photo} \sim 0.35$ (**Figure 6**). In general, the photometric redshift of an object is estimated by comparing its apparent brightness in several bandpasses to expectations for galaxies of a given type at each possible redshift. Calculations of the $z_{\rm photo}$ values for the clusters are based on a nearest neighbor polynomial algorithm (Abazajian et al. 2009). Inevitably, there are uncertainties in redshift measurements. The member galaxies of each cluster have a broad redshift probability distribution $p(z)$ (Cunha et al. 2009), but in our study, we only represent each cluster's redshift as a single average value of its BCG's $z_{\rm photo}$.

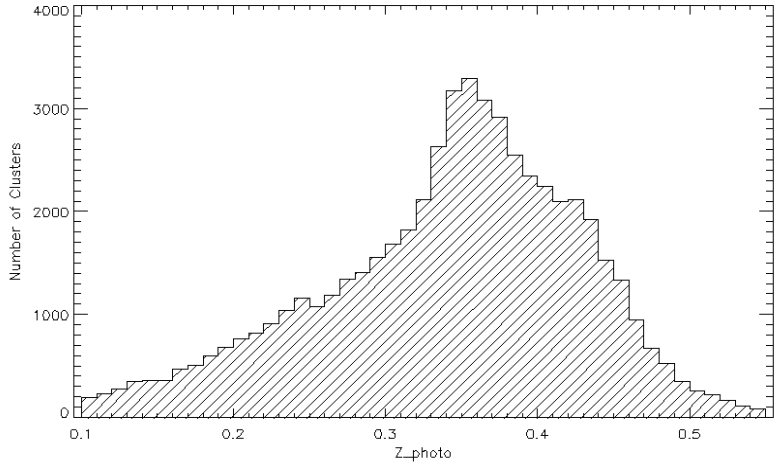

Figure 6. Redshift distribution of all the clusters in the catalog

To estimate redshift errors, we use the spectroscopic redshift, $z_{\rm spec}$, of a cluster to compare with the $z_{\rm photo}$ value. Spectroscopic redshifts are more accurate because by using spectroscopy to observe wavelengths of characteristic spectral lines, one can measure how far they are shifted from the actual positions. In this specific catalog, $z_{\rm spec}$ measurements are

available for the BCGs of 16,262 clusters. Their redshift errors, defined as $\Delta z = z_{\rm photo} - z_{\rm spec}$, fall in the range $-1.9 < \Delta z < 0.51$. However, after taking away a small number of outliers in the error distribution, 16,230 of the clusters have errors in the range $[-0.2, 0.2]$. Their error distribution is shown in **Figure 7** and superimposed by a Gaussian curve with standard deviation $\sigma_s \sim 0.023$; the Gaussian curve has broader tails than the core of the histogram. The outliers in this error distribution could have resulted from BCG misidentifications, which led to a large difference between $z_{\rm photo}$ and $z_{\rm spec}$ for a number of clusters. The value of σ_s is taken into consideration in the following algorithm that searches for cluster pairs with possible filamentary connections.

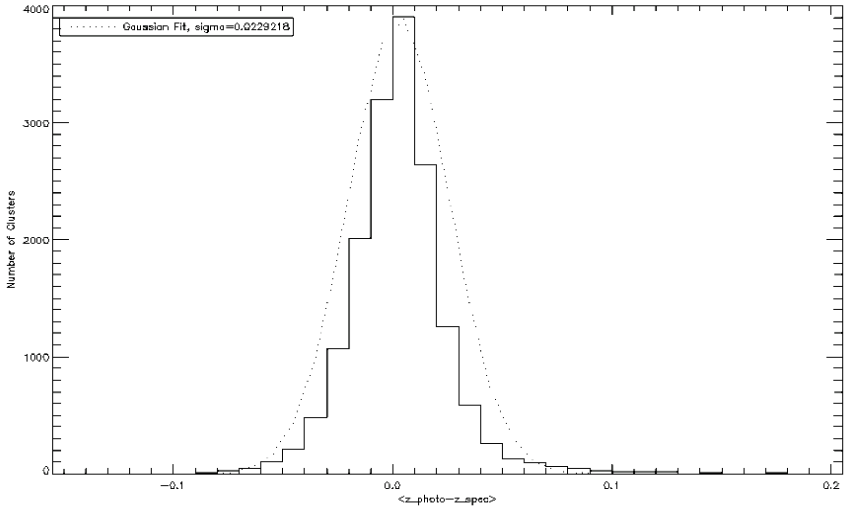

Figure 7. Distribution of redshift errors plotted against a Gaussian curve. The Gaussian curve is broader due to the dominating tails of the histogram distribution.

C. Search Algorithm

1) Redshift Slice

Since filaments have a finite length and redshift width, the process of searching for cluster pairs can be optimized by dividing the redshift distribution into smaller slices. The selection criterion for the thickness of these slices depends on the redshift error Δz. Suppose we take into consideration any error that falls within one standard deviation of the redshift error distribution, $\sigma_s \sim 0.023$. We choose the thickness of each redshift slice to be $\sim$0.02. If two clusters of a pair fall into the same slice, detecting such pairs becomes a trivial problem. However, the algorithm must also be able to take into account when two member clusters of a pair fall into different slices; we accomplish this by staggering the redshift slices. For example, let us define the first set of slices in the range $0.10 \lesssim z_{\rm photo} \lesssim 0.54$ with increments of 0.02. The boundaries of these slices are thus marked at 0.12, 0.14, 0.16, and so on. After searching for all the pairs in this first set, we redefine each slice's boundaries, for example, by creating another set of boundaries marked at 0.13, 0.15, 0.17, etc. This second set then yields another sample of cluster pairs that overlaps with the outcomes of the first set. In order to find all the pairs without repetitions, we need to find the union of both sets to get the final results.

2) Pair Separation

With the given photometric redshift data of galaxy clusters and the tools to calculate their distances, our next step is to select cluster pairs based on expectations for the lengths of filaments observed in the ΛCDM simulation. Colberg et al. (2005) theoretically found that about a third of all cluster pairs whose separation is between $15h^{-1}$ Mpc and $20h^{-1}$ Mpc are connected by a filament. Their analysis of filament abundance also shows that as the length of cluster-cluster connections becomes larger than $25h^{-1}$ Mpc, the fractional abundance of filaments fall off quickly, and beyond $50h^{-1}$ Mpc, the abundance of filaments is very minimal. In addition, close cluster pairs with separation $\lesssim 5h^{-1}$ Mpc are always connected by a filament. This is because two very close clusters tend to have overlapping infall regions (Diaferio & Geller, 1997) and they may eventually merge due to gravity. These results agree with observational data from 2dFGRS (Pimbblet et al. 2004). In this algorithm, we will select all cluster pairs with separation up to 50 Mpc. Technically, this separation is related to the transverse as well as the line-of-sight comoving distance. However, we also know that cluster pair separation is much smaller compared to other large-scale distances. Thus, the calculation of such separation can be estimated using simple trigonometry. Suppose Cluster 1 is at a distance D_1 from the observer, Cluster 2 at D_2, and the two of them separated by an angle θ on the sky (θ is computed based on their angular positions, i.e. ra and dec). D_1 and D_2 are the comoving distances of the clusters measured along our line of sight. Their physical separation S can be estimated using the law of cosines (**Figure 8**), $S^2 = D_1^2 + D_2^2 - 2D_1D_2\cos\theta$.

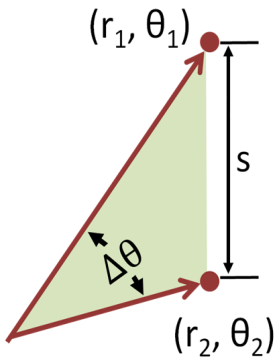

Figure 8. Law of cosines: $S^2 = r_1^2 + r_2^2 - 2r_1r_2\cos(\Delta\theta)$. Image Courtesy of Brews Ohare.

3) Closest Neighbors

Another parameter that helps optimize the search algorithm is to set an upper limit on the number of filaments each cluster is connected to. In the same study by Colbert et al. (2005), the majority of all clusters possesses between one and four filaments. This result puts a constraint on how many neighbors a cluster can pair up with. In our algorithm, instead of randomly searching for all the permutation of pairings within a certain redshift slice, we apply a criterion that limits the number of closest neighbors of each cluster to five. Of course, not all of these neighbors will be connected to the cluster through a filament. However, we would like to include all the possibilities instead of setting a lower limit that might take away possible filamentary connections.

4) Field galaxies

At the end of the cluster-pair search algorithm, we would like to learn about the galaxy distribution of these pairs and extract the information about their field galaxies such as angular positions and redshifts. The properties of these galaxies are contained in the SDSS DR7 catalog. From the angular positions of individual cluster pairs on the sky, we can select the galaxies that fall into the sky areas that extend to 15 arcmin measured from the intercluster axis and each cluster center. This selection allows a large enough area for future weak-lensing studies, but the area will be cut down in the stacking procedure depending on the separations of cluster pairs. Since each cluster pair was selected within a slice of $\Delta z = 0.02$, we use the same redshift thickness to eliminate the background and foreground objects, and only select the field galaxies that fall into that slice. The narrow thickness of each slice is a conservative choice in our search for galaxies associated with each pair. This constraint may be changed in future studies by dividing the redshift distribution into less, but wider slices. Due to different filament alignments, only a fraction of the galaxies observed between two clusters of the same pair truly lie along the dark matter filament. This source of noise in our data will contaminate individual filament signals, but through stacking procedures, the signal-to-noise ratio will be enhanced.

D. Results of Search Algorithm

Of more than 55,000 clusters from the GMBCG catalog (Hao et al. 2010), the search algorithm selects 16,643 cluster pairs with physical separation up to 50 Mpc. A 2D map of field galaxies of several pairs are shown in **Figure 9** with their intercluster axis drawn between two cluster centers. In two dimensions, we can only observe the different orientations of the cluster pairs as projected onto the sky and their angular separations. Included in these maps are not only the clusters themselves but also the field galaxies within 15 arcmin from the intercluster axis after the background and foreground have been eliminated.

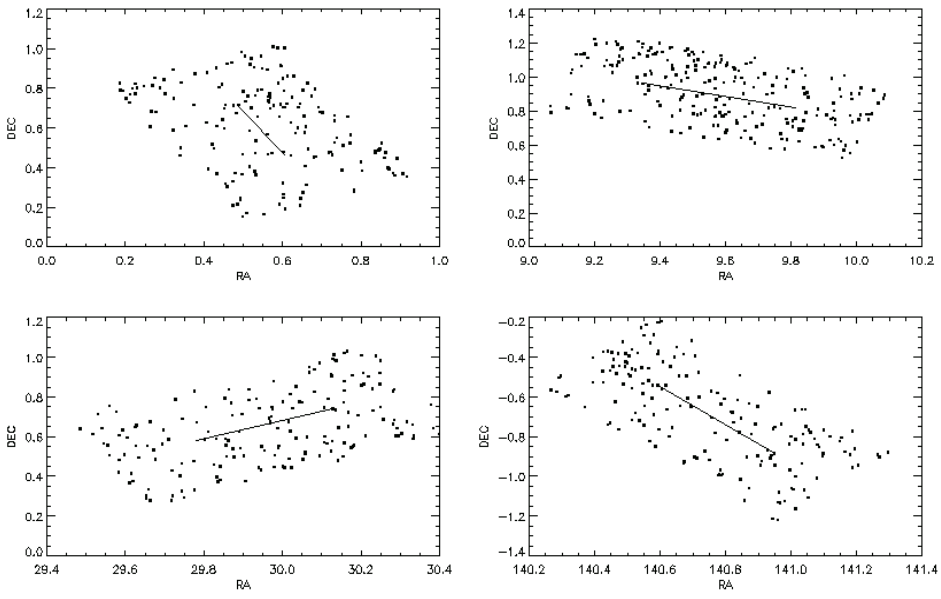

Figure 9. Examples of cluster pairs found by the search algorithm. An intercluster axis is drawn on each 2D map constructed based on angular positions of the galaxies.

Since ∼64% of the redshift distribution of the clusters in the GMBCG catalog falls in the range $0.30 < z_{\rm photo} < 0.45$, many cluster pairs that we have found by applying the search algorithm also have redshifts in this range (**Figure 10**). The distribution of intercluster separations is within the range $1\,{\rm Mpc} \lesssim d_{\rm sep} \lesssim 50\,{\rm Mpc}$, since we have set the upper limit at 50 Mpc (**Figure 11**). About ∼97% of all the cluster pairs found have $d_{\rm sep} > 5\,{\rm Mpc}$. We are interested in these pairs because the ones with $d_{\rm sep} < 5$ Mpc are likely to have overlapping infall regions (Colberg et al. 2005), and we try to eliminate the noise caused by these regions to enhance the filament signal in the stacking procedure. Also, it is worth noting that the separation distribution falls off quickly as the clusters are separated by $< 7\,{\rm Mpc}$. This is due to the fact that any two close clusters could be merging and thus have been detected as one single cluster in the GMBCG catalog.

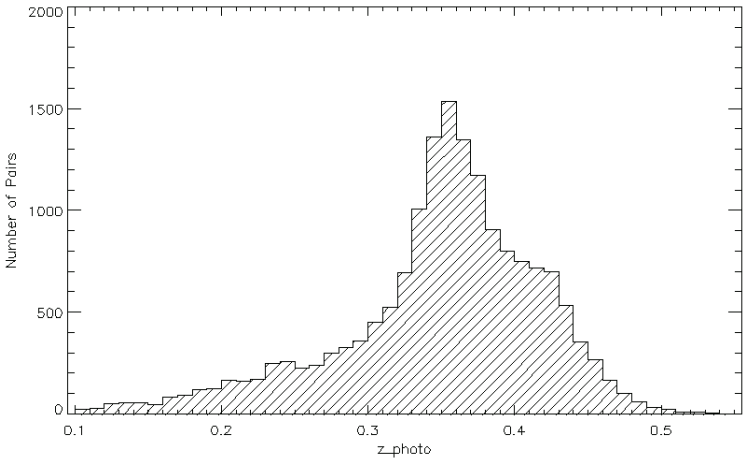

Figure 10. Redshift distribution of the cluster pairs

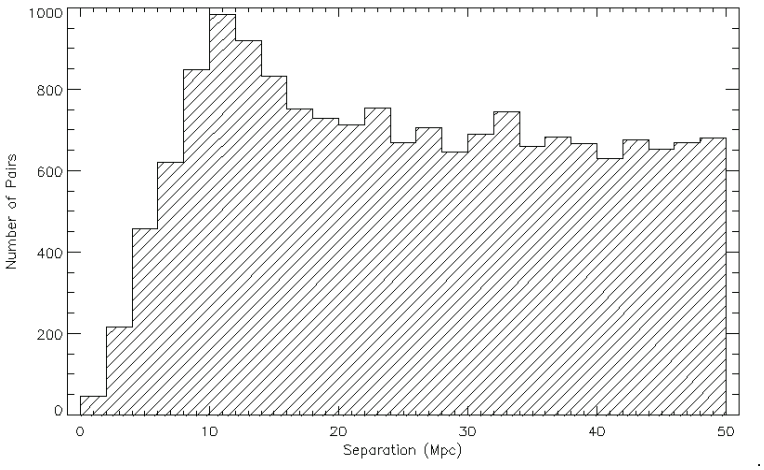

Figure 11. Distribution of cluster pairs' separations in Mpc.

Another property of invidual cluster pairs that is correlated with the strength of their filament signals is cluster richness. **Figure 12** shows the scaled richness of each cluster (Hao et al. 2010) in our catalog of cluster pairs as a function of its redshift. Filaments can form

between two clusters with comparable richness, or one with much higher richness than the other. Richer clusters accumulate more mass, and Colberg et al. (2005) have shown in their ΛCDM simulations that more massive clusters are connected to more filaments, forming the large-scale structure's backbone.

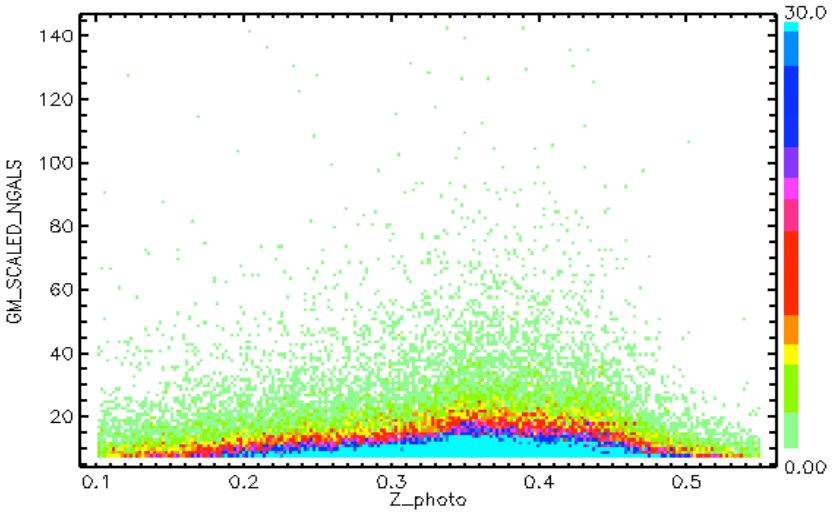

Figure 12. Color-coded 2D histogram of scaled richness vs. redshift.

For each cluster pair, let us define the total richness of its two clusters as

$$\Sigma N_{\rm gals} = N^{\rm scaled}_{\rm gals,1} + N^{\rm scaled}_{\rm gals,2}$$

where $N^{\rm scaled}_{{\rm gals},i}$ is the scaled richness of each cluster. **Figure 13** shows the distribution of $\Sigma N_{\rm gals}$ for all cluster pairs found in the search algorithm. The distribution of total richness spans the range $16 \leq \Sigma N_{\rm gals} \leq 272$. However, about 70% of all cluster pairs have total richness $\Sigma N_{\rm gals} \lesssim 30$.

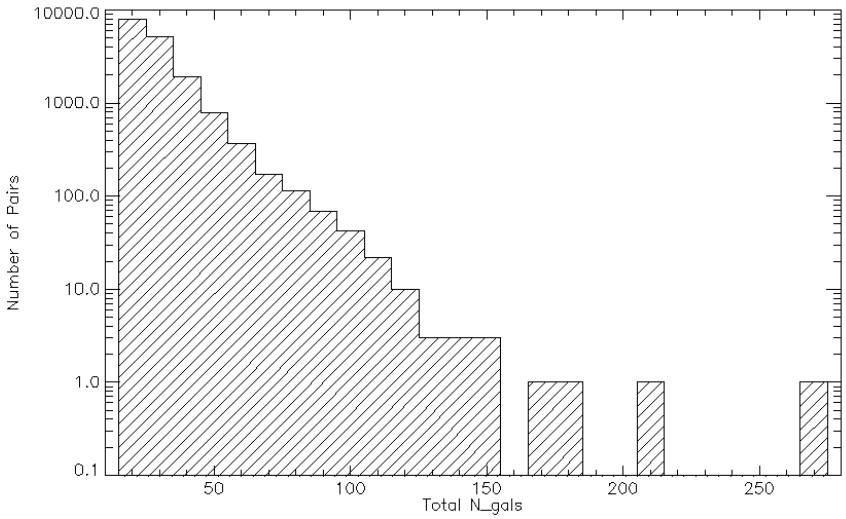

Figure 13. Distribution of total richness of invidual cluster pairs.

With our knowledge of redshifts, separations, and richness for more than 16,000 cluster pairs, we are interested in devising a stacking method to study different properties and correlations associated with the population of galaxies that lie along the filaments between clusters. The stacking method will be a generic algorithm that can be applied to any set of cluster pairs and take into consideration different input parameters of interests. For example, we can stack cluster pairs that fall into different ranges of intercluster separation, which will allow us to study how the strength of filament signals is correlated with the distance between clusters.

II. Stacking Procedure

The cluster pairs that we found in the search algorithm have different orientations in the three-dimensional space and varied intercluster distances. In stacking cluster pairs, we treat each pair as a two-dimensional map with their filament galaxies included. The mathematical approach undertaken involves translation and rotation of coordinate systems. In addition, due to various pair separations, the maps of some clusters will be either stretched or compressed depending on the relative distance of their galaxies with respect to the intercluster separation.

A. Rotation & Translation

Consider a 2D map of a cluster pair consisting of n galaxies in the equatorial coordinate system. Let us treat the galaxies as n points on the map. We define this set of points as $S = \{g_1, g_2, g_3, ..., g_n\}$. The spatial location of each galaxy g_k $(1 \leq k \leq n)$ is defined by RA and DEC coordinates on the sky, so each point g_k is located at (φ_k, θ_k). To rotate each of these points clockwise through an angle α, we apply the typical transformation matrix

$$R = \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix}$$

to get the transformed coordinates (x_k, y_k) for the point g_k,

$$\begin{pmatrix} x_k \\ y_k \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \varphi_k \\ \theta_k \end{pmatrix}.$$

In the stacking algorithm, we want to rotate each of the cluster pairs about the midpoint of its intercluster axis such that in the transformed coordinate system, all the intercluster axes lie in the horizontal direction. An example is shown in **Figure 14**. Before rotating each galaxy on the map of a cluster pair about the midpoint (φ_0, θ_0), we first need another transformation in which the location of each point is defined relative to (φ_0, θ_0). Let us translate each point (φ_k, θ_k) into a coordinate system where the intercluster midpoint is the origin:

$$(x_k^*, y_k^*) = (\varphi_k - \varphi_0, \theta_k - \theta_0),$$

and let α be the angle between the unit vector of the intercluster axis and the horizontal. The rotational matrix R can be constructed as above. Depending on the orientation of the intercluster axis, the rotation can be either clockwise or counterclockwise. Rotating clockwise by α is the same as rotating counterclockwise by $2\pi - \alpha$. After applying the rotation, each galaxy on the cluster pair map has the final transformed coordinates (x_k, y_k) where

$$\begin{pmatrix} x_k \\ y_k \end{pmatrix} = \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix} \begin{pmatrix} x_k^* \\ y_k^* \end{pmatrix},$$

and the intercluster axis aligns with the line $y=0$ with its midpoint at $(0,0)$.

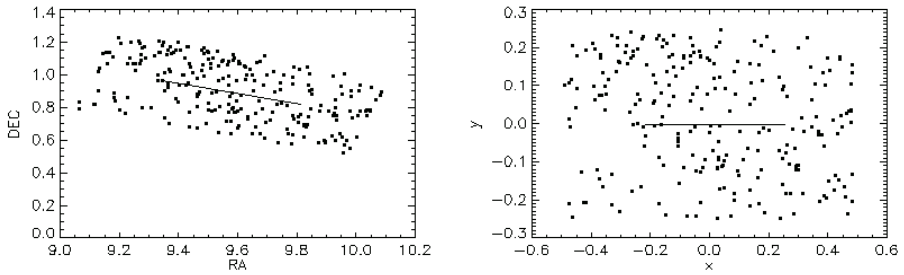

Figure 14. The map of a cluster pair before (left) and after (right) rotational transformation about the center of its intercluster axis.

B. Distance Rescaling

After rotating all of the cluster pairs to have their intercluster axes align in one uniform orientation, our next step is to rescale their maps such that the two cluster centers of each pair will lie exactly on top of the other pairs' during stacking. Suppose we rescale the distance of each galaxy from the intercluster midpoint (0, 0) with respect to the separation of the respective cluster pair. Let $(x_1, 0)$ and $(x_2, 0)$ be the two cluster centers of such a pair, and $\Delta x = |x_2 - x_1|$ be their separation in arbitraty units of the transformed coordinate system. For each galaxy located at (x_k, y_k) on the map, its rescaled coordinates (x'_k, y'_k) are defined as

$$(x'_k, y'_k) = \tfrac{1}{\Delta x}(x_k, y_k).$$

This method of rescaling results in redefining the separation to become unity in the scaled coordinate system,

$$\begin{aligned}\Delta x' &= |x'_2 - x'_1| \\ &= \left|\tfrac{x_2}{\Delta x} - \tfrac{x_1}{\Delta x}\right| \\ &= \tfrac{1}{\Delta x}|x_2 - x_1| \\ &= 1\end{aligned}$$

with $(x'_1, 0)$ and $(x'_2, 0)$ as the cluster centers. Since the origin of each map is the midpoint between two clusters, $x'_1 = -0.5$ and $x'_2 = 0.5$.

Applying the same algorithm to all cluster pairs, every intercluster separation is rescaled to unity, and each intercluster axis has the same horizontal orientation. This allows us to stack the cluster pairs such that the cluster centers of each pair always superimposed on those of the others at (-0.5,0) and (0.5,0) in the arbitrary transformed coordinate system. The size of each cluster pair could be stretched or compressed during this process. We avoid the edge effects by cutting out the area that does not include galaxies from all stacked cluster pairs. Since we are only interested in studying the population of filament galaxies within the intercluster region, as long as the location of each filament galaxy is rescaled appropriately, the filament signal will be enhanced in the stacked map.

C. Gaussian Smoothing

When many cluster pairs are stacked on top of each other, the noise increases along with the filament signal. The density map of any stack of cluster pairs can be divided into smaller square grids called pixels. Each pixel contains a certain number of galaxies. The pixelated map contains the details of finite small areas and looks rather noisy. In order to reduce this noise, we apply Gaussian smoothing to the pixelated map to gain a better visualization of the larger structures and not the small details (**Figure 15**). For each pixel, the galaxy density is convolved with with a Gaussian kernel, taking into account the weighted average density of its neighbors. A detailed explanation of how to implement Gaussian smoothing can be found in Ritter & Wilson (2001).

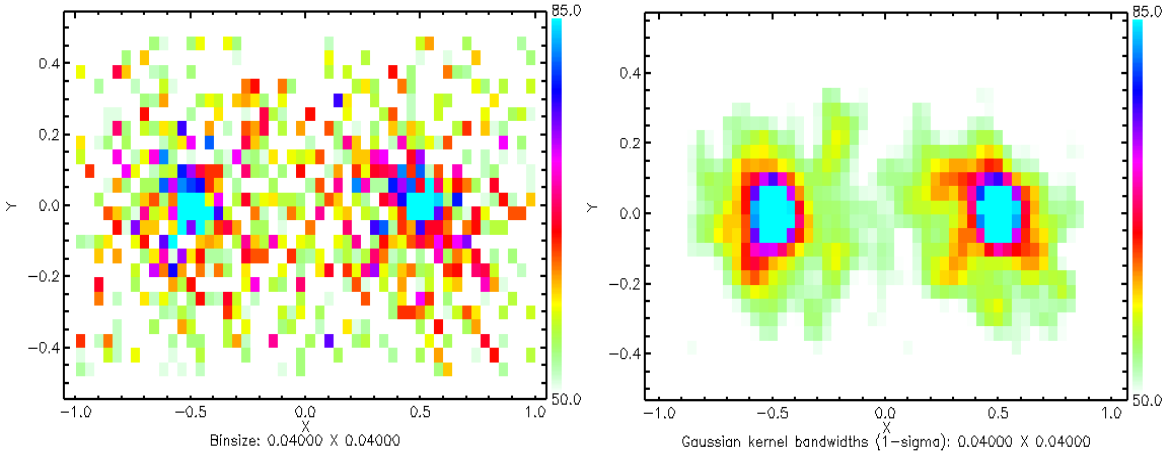

Figure 15. A stack of 1,000 cluster pairs selected randomly from the cluster pair catalog. The left image shows a pixelated 2D density map of the stack. The right image shows the density map after applying Gaussian smoothing.

D. Bootstrap Statistics

In order to learn more about the statistical significance of each pixel in a stack of cluster pairs, we apply bootstrap methods (Davison et al. 1997) to that stack by treating all of its pairs as a set of independent objects. The most basic bootstrapping technique is resampling with replacement. Suppose we want to create bootstrap samples for a stack of N cluster pairs. Each sample must consist of the same number of cluster pairs as the original stack. To construct a bootstrap sample, we randomly select N pairs from the original set. Note that a pair can be included more than once in the new sample. A large number of boostrap samples allows us to study the distribution of galaxies that fall into a specific pixel or a finite area on the stacked map. Similarly, we can also apply bootstrapping techniques to the population of galaxies within a stack. Instead of resampling the pairs, we resample the galaxies by randomly selecting them from the original galaxy population. While boostrapping cluster pairs allows us to study the significance of each pair in a stack, boostrapping the galaxies is important in determining the significance of the filament signal with contrast to the mean galaxy density of the background.

III. Stacking Analysis

The stacking procedure outlined above can take in a number of different input parameters such as redshift, separation, and richness of cluster pairs. This gives us the flexibility to stack cluster pairs that share similar properties and allow us to compare their filament signals in response to the change in those parameters.

A. Stacking Based On Redshift

Given the cluster catalog spanning the range $0.10 \leq z_{\rm photo} \leq 0.55$, we divide the redshift space into smaller slices of width $\delta z = 0.05$. For each slice, we stack all of the cluster pairs with separation $d_{\rm sep} > 5$Mpc. Recall that the galaxies of each pair lie within a redshift thickness of $\Delta z = 0.02$, thus each of the stacks constructed actually falls into a slice wider than the above mentioned δz. **Figure 16** shows four of the nine stacks created by such selection of cluster pairs based on redshifts. The maps of these stacks have been smoothed using the Gaussian method mentioned in the previous section. The color scale for each map varies because some stacks have more galaxies than others. Since more than 60% of the number of cluster pairs have redshifts within $0.30 \lesssim z_{\rm photo} \lesssim 0.45$, it is expected that the two stacks in this range will have a larger overall population of galaxies compared with the other stacks.

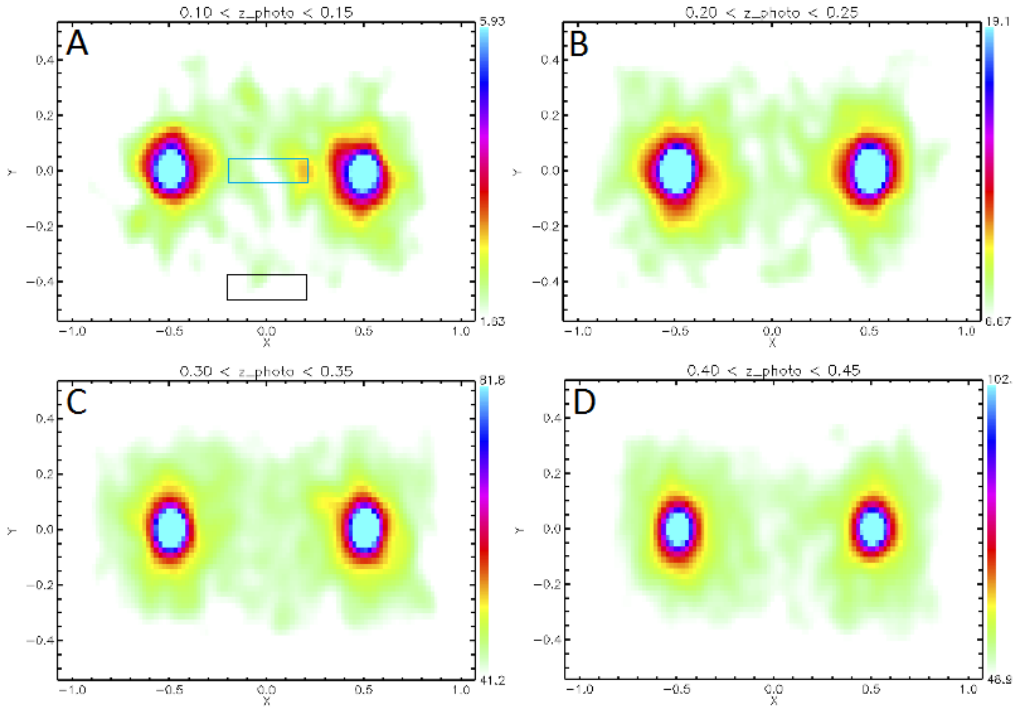

Figure 16. Stacks of cluster pairs in four different redshift slices. The total number of galaxies in each stack is as follow: A)$\sim$9,000; B)$\sim$38,000; C)$\sim$219,000; D)$\sim$255,000.

For each of the stacks in **Figure 16**, we apply the bootstrap method to find the distribution of galaxy overdensity in the filament and background. Suppose the filament region is defined

within the following boundaries: $-0.2 \leq x \leq 0.2$, $-0.05 \leq y \leq 0.05$; and the background region is within $-0.2 \leq x \leq 0.2$, $-0.46 \leq y \leq -0.36$; both have the same area as outlined in **Figure 16 A**. For each stack, two Gaussian distributions are constructed based on 1,000 bootstrap samples of its cluster pairs: one for the galaxy density within the defined filament region, and the other for the density in the background. The results are shown in **Figure 17**. For slice $0.10 \leq z_{\text{photo}} \leq 0.15$, there is much overlapping between the two distributions. This stack has a smaller number of cluster pairs and thus less filament galaxies, which makes the distinction between the filament signal and background less obvious. For the other stacks at higher redshifts, the distribution of filament galaxy density is more separated from that of the background due to a larger number of cluster pairs, hence a larger population of filament galaxies that enhances the filament signal. Looking at slice $0.40 \leq z_{\text{photo}} \leq 0.45$, the galaxy density distribution of the filament is $\sim 4\sigma_s$ away from the mean of the background's density.

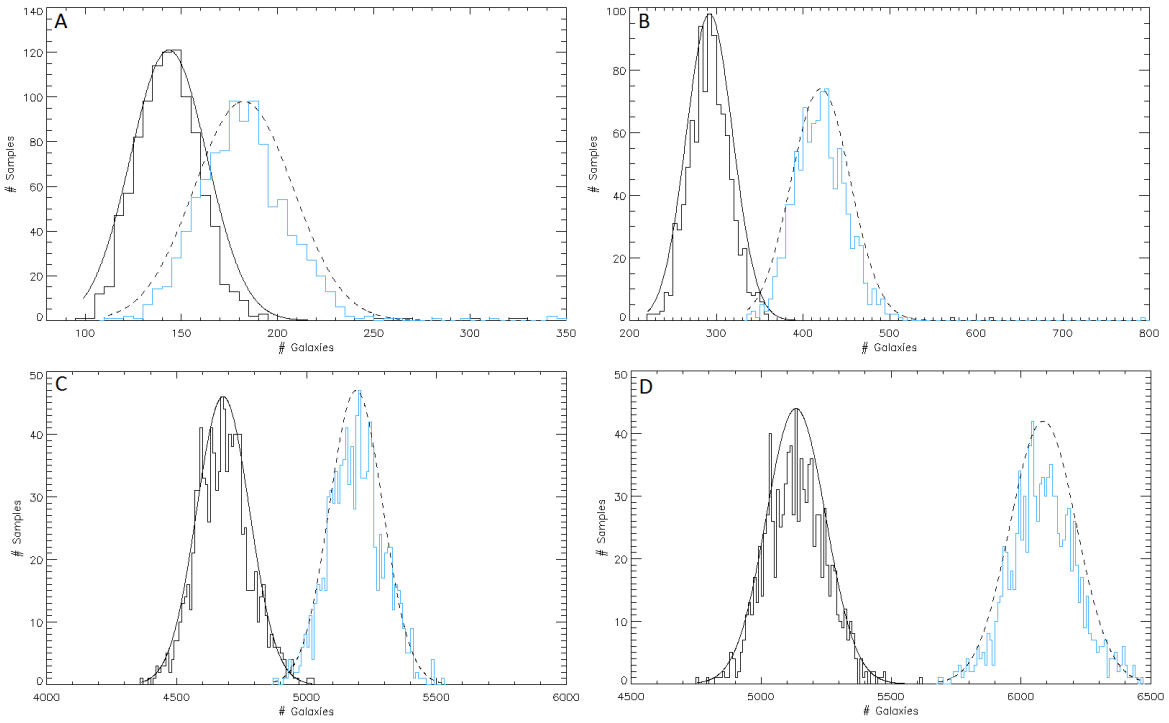

Figure 17. Distributions of filament and background galaxy overdensities, shown in blue and black, respectively, for each redshift stack based on 1,000 bootstrap samples of cluster pairs.

A) $0.10 \leq z_{\text{photo}} \leq 0.15$; B) $0.20 \leq z_{\text{photo}} \leq 0.25$; C) $0.30 \leq z_{\text{photo}} \leq 0.35$; D) $0.40 \leq z_{\text{photo}} \leq 0.45$

Instead of using the bootstrap samples constructed from the set of cluster pairs of each stack, we can also use the boostrap samples from the galaxy population to study the density within the filament and background regions. **Figure 18** shows galaxy density distributions similar to those in **Figure 17**, except the bootstrap samples for each stack come from its galaxies instead of cluster pairs. The filament density distribution is more separated from the background density distribution in those stacks with larger populations of galaxies.

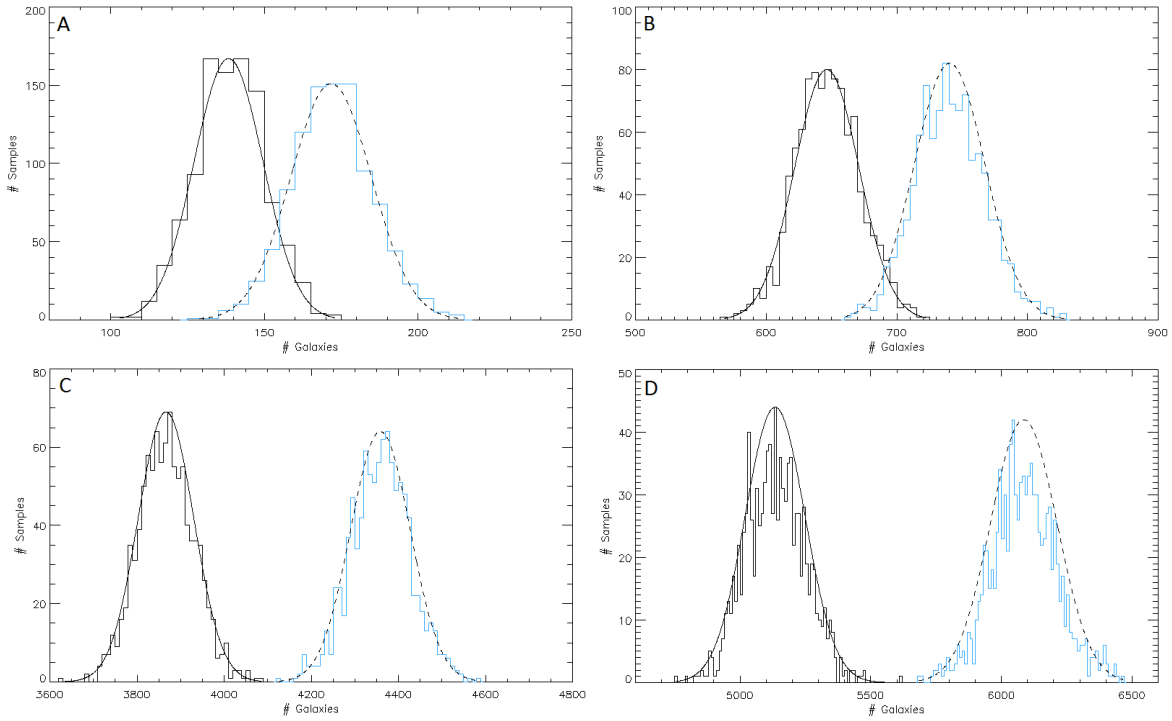

Figure 18. Distributions of galaxy overdensities based on 1,000 bootstrap samples from the galaxy populations in the filament and background regions.

Besides using bootstrap statistics to account for the random noise inherent to the data, we also need to subtract the background galaxies from the filament signals and find the significance of those signals. Assuming Gaussian statistics, the significance for each of N pixels in a density map is defined as follow,

$$s_k = \frac{\rho_k - \bar{\rho_b}}{\sigma_s}$$

where σ_s is the standard deviation of the background density distribution, $\bar{\rho_b}$ is the mean background density, and ρ_k is the mean density of each pixel determined from bootstrapping with k denoting the pixel index such that $1 \leq k \leq N$. Assuming that this distribution follows Poisson statistics, the uncertainty σ_s is defined as $\sigma_s = \sqrt{\bar{\rho_b}}$.

Taking the stack $0.25 \leq z_{\rm photo} \leq 0.30$ as an example, we use its bootstrap samples of cluster pairs to find the mean galaxy density for each pixel on the map where the pixel size is 0.04×0.04 in the aforementioned arbitrary units. The map's x and y ranges are $[-1.0, 1.0]$ and $[-0.5, 0.5]$, respectively, so there are 50 pixels in the x direction and 25 pixels in the y direction. The significance map of this stack is shown in **Figure 19** with contour lines drawn up to the $5\sigma_s$ level. The filament signal is $\sim 2\sigma_s$ above the background level.

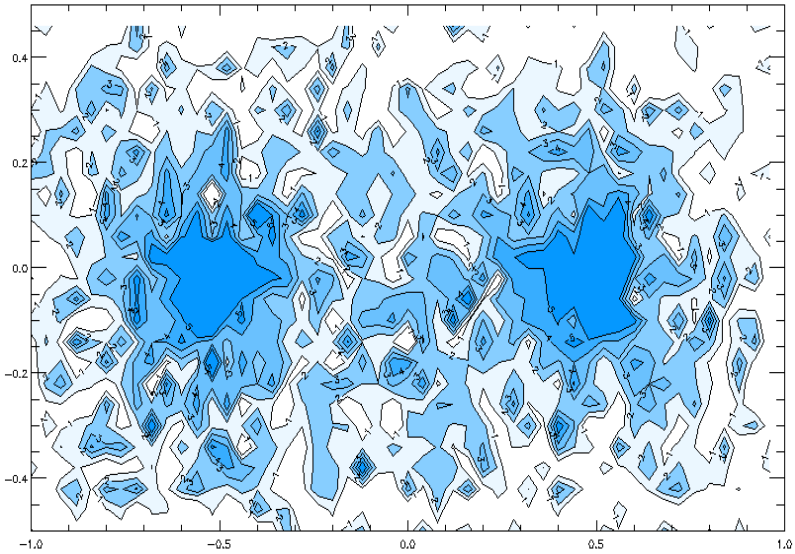

Figure 19. Significance contour map of stack $0.25 \leq z_{\text{photo}} \leq 0.30$.

B. Stacking Based On Separation

Intercluster separation has been shown in simulations to affect the strength of filament signals (Colberg et al. 2005). In addition, Pimbblet et al. (2004) has observed from 2dFGRS catalog that very close cluster pairs will always possess a filament. Before analyzing how filament strength varies with separation between clusters, we stack the cluster pairs based on the following ranges of separation:

$$[5, 10], [10, 15], [15, 20], [20, 30], [30, 40], \text{ and } [40, 50]$$

in units of Mpc. The 5 Mpc and 50 Mpc are the lower and upper limits that have been set earlier in selecting cluster pairs for our catalog. Four of the stacks are shown in **Figure 20**; each consists of cluster pairs from all redshifts that have the desired separation.

Similar to how the significance maps have been constructed for the stacks of different redshifts, a significance map for each of the stacks above allows us to take a closer look at the density of galaxies along the filaments (**Figure 21**). The correlation between the strength of filament signals and separation are not obvious in these maps due to the following reasons: first, the maps still have noise, and second, we do not yet have a good method to quantify the signal. In order to study the filament strength, we will need to construct significance maps based on the same background level. However, each of the maps in **Figure 21** has a different background. Therefore, higher significance does not necessarily correspond with stronger filament signal. On average, the filaments have a mean density at $\sim 2\sigma_s$ above the background with occurrences of small "blobs" at $\sim 3\sigma_s$.

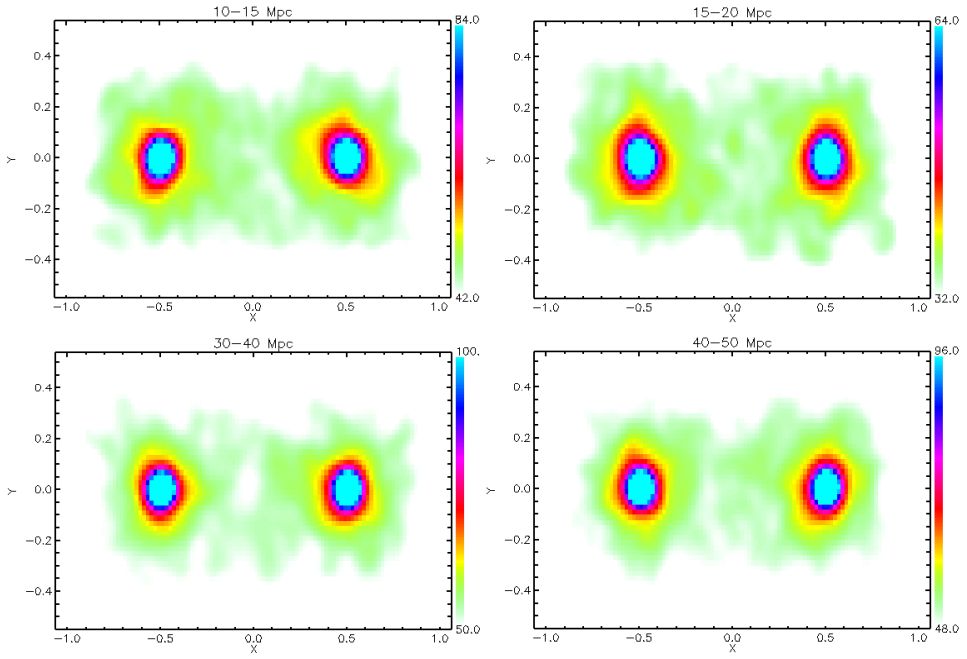

Figure 20. Gaussian-smoothed maps of stacks in four separation ranges.

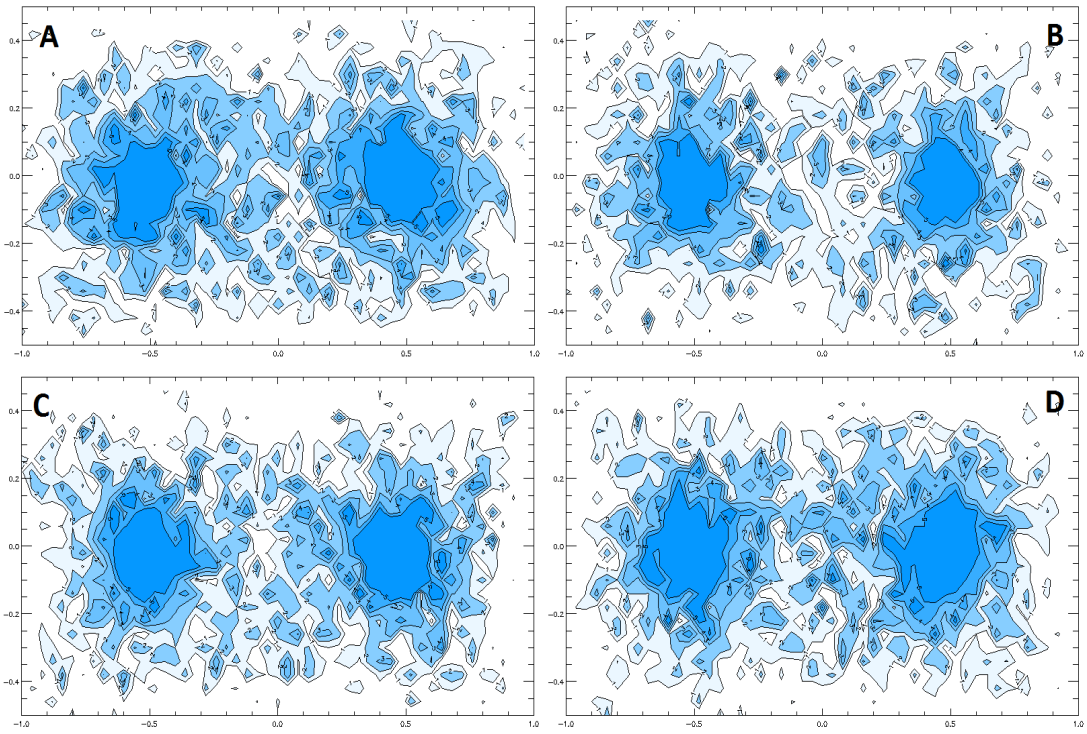

Figure 21. Significance maps constructed using 1,000 bootstrap samples from the set of cluster pairs of each stack. A) $10 \lesssim d_{\rm sep} \lesssim 15$; B) $15 \lesssim d_{\rm sep} \lesssim 20$; C) $30 \lesssim d_{\rm sep} \lesssim 40$; D) $40 \lesssim d_{\rm sep} \lesssim 50$.

C. Cross–Stacking

Another test that may be used to confirm the presense of filament signals observed in a stacked map is the cross–stacking method. We cut the map into two halves with respect to the midpoint of the intercluster axis, each consisted of a cluster center. Next, the position of each half–map is swapped with the other, and similar Gaussian smoothing and bootstrap routines are applied to the new map. However, instead of seeing a filament between two cluster centers as before, we now expect the filament to project outward toward the map's boundaries. If the filament signal was not merely just an artifact of cluster infall regions, the center region of the cross–stacked map should not possess any significant galaxy overdensity. Taking the stack $40\,\mathrm{Mpc} \lesssim d_{\mathrm{sep}} \lesssim 50\,\mathrm{Mpc}$ as an example, we apply the cross-stacking routine mentioned above to the cluster pairs in the stack. **Figure 22** shows the Gaussian smoothed version of the resulted cross–stack along with its significance map, which is constructed based on 1,000 bootstrap samples from the galaxy population. Two regions are outlined on the smooth map, and their galaxy density distributions are compared in **Figure 23**. Since the two distributions almost completely overlap, we are convinced that the center region of the stack has no filament signal and consists of only background galaxies.

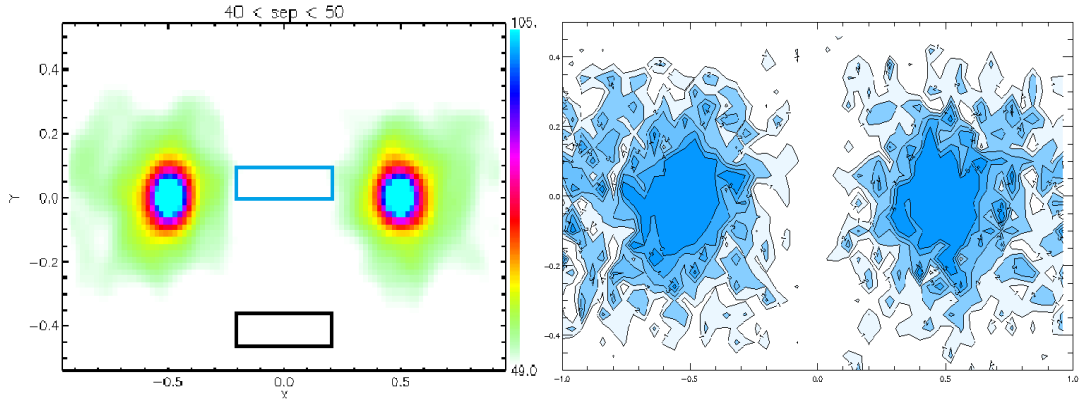

Figure 22. Results of cross-stacking cluster pairs in the range $40\,\mathrm{Mpc} \lesssim d_{\mathrm{sep}} \lesssim 50\,\mathrm{Mpc}$. As expected, the filament projects toward the boundaries of the map while the center region has a mean density comparable to the background level.

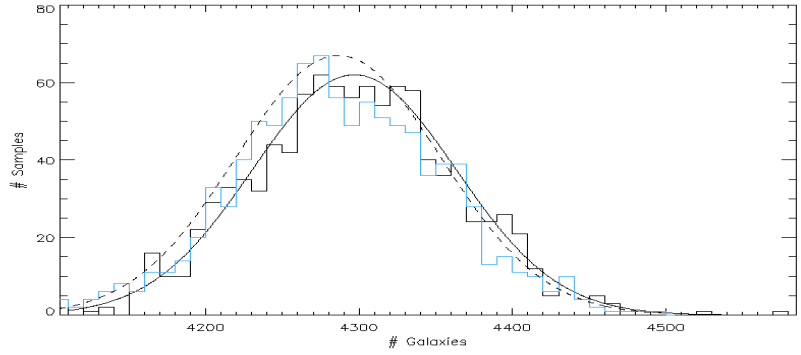

Figure 23. Distributions of galaxy density in two color-coded, outlined areas on the stacked map. The overlapping suggests that the center region of the map only contains background galaxies.

D. Stacking Based On Cluster Richness

After sorting all cluster pairs based on their total richness, $\Sigma N_{\rm gals}$, two stacks are constructed as shown in **Figure 24**, each consisting of 2,000 pairs. The first stack include cluster pairs with lowest total richness, $16 < \Sigma N_{\rm gals} < 18$, and the second stack with greatest total richness, $41 \leq \Sigma N_{\rm gals} \leq 272$.

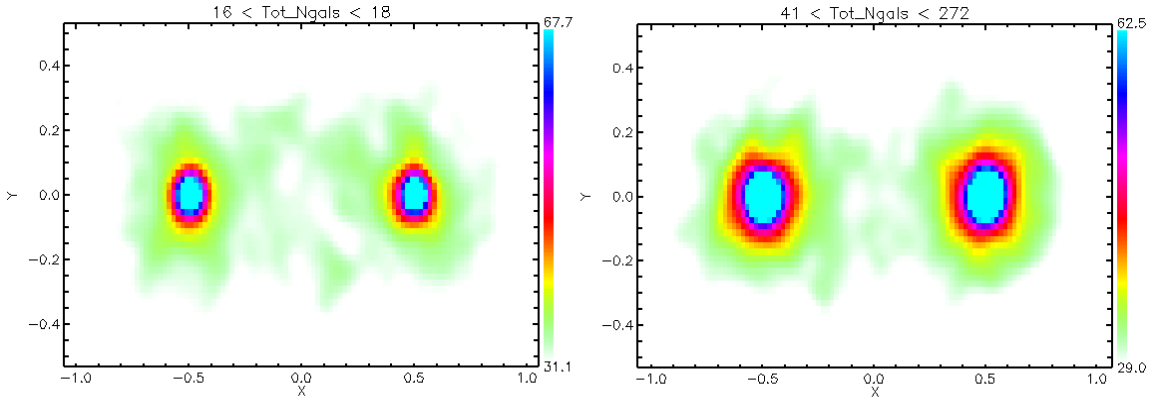

Figure 24. Two stacks with different total–richness ranges. Each stack has $\sim$150, 000 galaxies.

The siginificance maps of the two stacks above are shown in **Figure 25**. They are constructed using 1,000 bootstrap samples from the galaxy populations. It is obvious that the stack with the greater total cluster richness has greater galaxy overdensities around its two cluster centers. However, the filament signal, which is $\sim 2\sigma_s$ above the background level, is about the same in both of them. We cannot examine the filament strength because each stack has a different background level. The significance of the filament signal in one stack does not really correspond with its strength relative to the other stack.

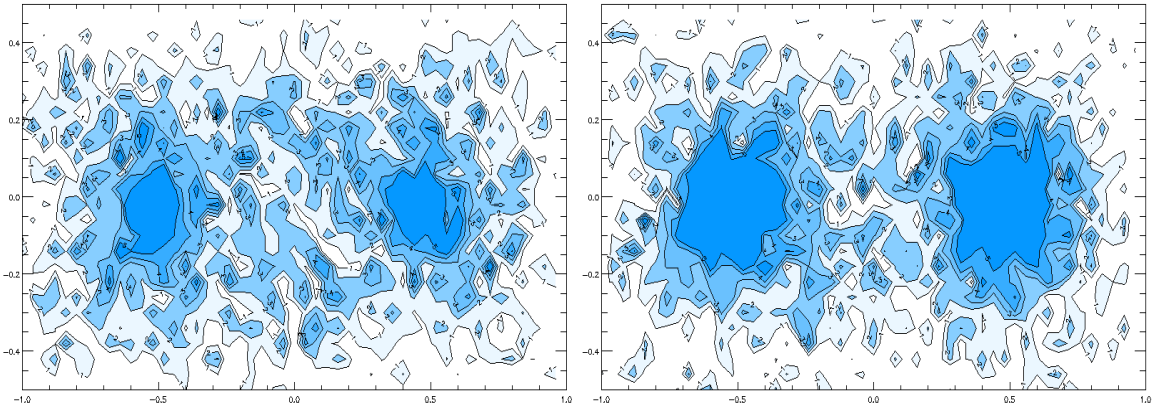

Figure 25. Significance maps of two stacks based on $\Sigma N_{\rm gals}$.

Left: $16 < \Sigma N_{\rm gals} < 18$; Right: $41 < \Sigma N_{\rm gals} < 272$.

IV. Conclusions

We have devised a generic algorithm that searches for cluster pairs with possible filamentary connections based on redshift, intercluster separation, and the number of closest neighbors. As a result, approximately 16,000 cluster pairs have been selected from the GMBCG cluster catalog provided by Hao et al. (2010). These cluster pairs have separations ranging from 5 Mpc to 50 Mpc and BCG photometric redshifts within $0.10 \lesssim z_{\rm photo} \lesssim 0.55$. The field galaxies of each pair were extracted from SDSS DR7 catalog; they lie within a redshift slice of $\Delta z \sim 0.02$ centered on the midpoint of the intercluster axis. We eliminated the foreground and background galaxies that fall inside the field of each pair to reduce the noise in detecting filament signals.

Since intercluster filaments are known to have low density contrast, we have created a stacking procedure to enhance the filament signal of cluster pairs. We aligned the intercluster axes to get the overdensity of filament galaxies within each stacked map. While the stacking method can enhance the filament signal, the background noise increases as well. Therefore, bootstrap statistics have been used to determine the distribution of galaxy overdensity in the filament and background regions. We also constructed significance maps for a number of stacks by assigning a significance value to each pixel in the stacked maps with respect to the mean density of background galaxies. The filament signals in these maps have a mean density at $\sim 2\sigma_s$ above the background with occurrences of pixels at the$\sim 3\sigma_s$ level. No strong correlation between filament strength and intercluster separation was observed. Although the significance maps are not completely rid of noise, they have shown that our stacking method can be potentially useful in detecting intercluster filaments. In this research, we only stacked cluster pairs based on redshift, separation, and cluster richness. Yet, the approach can be generalized to include other parameters such as luminosity, color, and the type of galaxies that lie along the filaments.

In addition, the pair-finder algorithm that we devised can be applied to any cluster catalog from redshift surveys and modified to include more constraints on pair selection. In the future, it can also be run on mock cluster catalogs from N-body simulations. If one is interested in applying weak lensing techniques for the purpose of mass reconstruction of the filaments, a larger area for the field of each cluster pair is needed along with the background galaxies. This can be easily modified within our original framework, thus making it convenient to obtain data that are useful for both the stacking method and weak lensing studies.

The challenge of using 3D galaxy distributions to detect dark matter filaments may not be completely solved, but the preliminary results of our research have shown some promise on using various stacking techniques to study the galaxy populations that lie along these filaments. Future improvements are necessary in order to further amplify the filament signal while at the same time reducing the background noise. Looking into details at the properties of filament galaxies will also help us gain a better understanding of large-scale structure formation, and allow us to have a more complete picture of the ΛCDM model of the universe.

References

Abazajian, K. et al. 2009, ApJS, 182, 543

Blumenthal, G. R., Faber, S. M., Primack, J. R., & Rees, M. J. 1984, Nature, 311, 517

Bothun, G. 1998, Modern Cosmological Observations and Problems (1st ed.; Padstow, UK: TF International Ltd.)

Colberg, J. M., Krughoff, K. S., & Connolly, A. J. 2005, MNRAS, 359, 272

Cunha, C. E., Lima, M., Oyaizu, H., Frieman, J., & Lin, H. 2009, MNRAS, 396, 2379

Davison, A. C., Hinkley, D. V., & Canty, A. J. 1997, Bootstrap Methods and their Application (New York, NY: Cambridge University Press)

Diaferio, A., & Geller, M. J. 1997, ApJ, 481, 633

Dietrich, J. P., Schneider, P., Clowe, D., Romano-Diaz, E., & Kerp, J. 2005, A&A 440, 453

Geller, M., & Huchra, J. 1989, Science, 246, 897

Gott J. R. I., Juric M., Schlegel D., Hoyle F., Vogeley M., Tegmark M., Bahcall N., & Brinkmann J. 2005, ApJ, 624, 463

Hao, J. et al. 2010, ApJS, 191, 254

Hogg, D. W. 2000, arXiv:astro-ph/9905116v4

Komatsu, E. et al. 2009, ApJS, 180, 330

Kravtsov, A., & Klypin, A. Simulations of Cold Dark Matter Universe at the National Center for Supercomputer Applications. *Center for Cosmological Physics*, National Science Foundation, http://cosmicweb.uchicago.edu/sims.html

Liddle, A. R., & Lyth, D. H. 2000, Cosmological Inflation and Large-Scale Structure (New York, NY: Cambridge University Press)

Mukhanov, V. F. 2005, Physical Foundations of Cosmology (New York, NY: Cambridge University Press)

Pimbblet, K. A., Drinkwater, M. J., & Hawkrigg, M. C. 2004, MNRAS, 354, L61

Ritter, G. X., & Wilson, J. N. 2001, Handbook of Computer Vision Algorithms in Image Algebra (2nd ed.; Boca Raton, FL: CRC Press)

Rudnick L., Brown S., & Williams L. 2007, ApJ, 671, 40

Ryden, B. 2003, Introduction to Cosmology (San Francisco, CA: Addison Wesley)

Schneider, P. 2006, Extragalatic Astronomy and Cosmology: An Introduction (New York, NY: Springer)

Sparke, L. S., & Gallagher, J. S. 2007, Galaxies in the Universe: An Introduction (2nd ed.; New York, NY: Cambridge University Press)

Werner, N., Finoguenov, A., Kaastra, J. S., Simionescu, A., Dietrich, J. P., Vink, J., & Bohringer, H. 2008, A&A, 482, L29

White, S. D. M., Frenk, C. S., Davis, M., & Efstathiou, G. 1987, ApJ, 313, 505