

Felix: A Topology based Framework for Visual Exploration of Cosmic Filaments

Nithin Shivashankar, *Student Member, IEEE*, Pratyush Pranav, Vijay Natarajan, *Member, IEEE*, Rien van de Weygaert, E G Patrick Bos, and Steven Rieder.

Abstract—The large-scale structure of the universe is comprised of virialized blob-like clusters, linear filaments, sheet-like walls and huge near empty three-dimensional voids. Characterizing the large scale universe is essential to our understanding of the formation and evolution of galaxies. The density range of clusters, walls and voids are relatively well separated, when compared to filaments, which span a relatively larger range. The large scale filamentary network thus forms an intricate part of the cosmic web.

In this paper, we describe Felix, a topology based framework for visual exploration of filaments in the cosmic web. The filamentary structure is represented by the ascending manifold geometry of the 2-saddles in the Morse-Smale complex of the density field. We generate a hierarchy of Morse-Smale complexes and query for filaments based on the density ranges at the end points of the filaments. The query is processed efficiently over the entire hierarchical Morse-Smale complex, allowing for interactive visualization.

We apply Felix to computer simulations based on the heuristic Voronoi kinematic model and the standard Λ CDM cosmology, and demonstrate its usefulness through two case studies. First, we extract cosmic filaments within and across cluster like regions in Voronoi kinematic simulation datasets. We demonstrate that we produce similar results to existing structure finders. Second, we extract different classes of filaments based on their density characteristics from the Λ CDM simulation datasets. Filaments that form the spine of the cosmic web, which exist in high density regions in the current epoch, are isolated using Felix. Also, filaments present in void-like regions are isolated and visualized. These filamentary structures are often over shadowed by higher density range filaments and are not easily characterizable and extractable using other filament extraction methodologies.

Index Terms—Morse-Smale complexes, tessellations, cosmology theory, cosmic web, large-scale structure of the universe.



1 INTRODUCTION

AT scales from a megaparsec to a few hundred megaparsecs¹, the universe has a web-like appearance. In the *cosmic web* [1], [2], galaxies, intergalactic gas, and dark matter have aggregated in an intricate wispy spatial pattern marked by dense compact *clusters*, elongated *filaments* and sheetlike *walls*, and large near-empty void regions. The filaments, stretching out as giant tentacles from the dense cluster nodes, serve as transport channels along which mass flows towards the clusters. They surround the flattened walls, which are tenuous, membrane-like features in the cosmic mass distribution.

All structures and objects in the universe emerged out of primordial fluctuations that were generated during the

inflationary era, moments after its birth, as the universe underwent a rapid phase of expansion [3], [4], [5]. The quantum fluctuations generated during this phase manifest themselves as fluctuations in the temperature of the cosmic microwave background [6], [7], [8]. The gravitational growth of these density and velocity perturbations has resulted in the wealth of structure that we see in the Universe. The web-like patterns mark the transition phase from the primordial Gaussian random field to highly nonlinear structures that have fully collapsed into halos and galaxies. As our insight into the complex structural pattern of the cosmic web has increased rapidly over the past years, it has become clear that the cosmic-web contains a wealth of information on a range of cosmological and astronomical aspects and processes.

An important illustration of the cosmological significance of the cosmic web concerns its dependence on the nature of dark energy and matter, the dominant but as yet unidentified forms of energy and matter in the Universe. One telling example of this is the recent realization that cosmic voids are sensitive and useful probes of the nature of dark energy and dark matter and testing grounds for modified gravity theories [9], [10], [11], [12], [13], [14], [15], [16]. As the cosmic web is first and foremost defined and shaped by the gravitationally dominant dark matter, it would be of considerable importance to be able to obtain detailed maps of dark matter distribution. In recent years, great strides have been made towards this goal as gravitational lensing of distant galaxies and objects by the dark matter have enabled an increasingly accurate view of its spatial distribution [17], [18]. Initial efforts concentrated on the detection and mapping of the deep potential wells

- N. Shivashankar is with the Department of Computer Science and Automation, Indian Institute of Science, Bangalore, India, 560038. Email: nithin@csa.iisc.ernet.in
- P. Pranav is with Kapteyn Astronomical Institute, University of Groningen, Netherlands. Email: pratyuze@gmail.com
- V. Natarajan is with the Department of Computer Science and Automation and Supercomputer Education and Research Centre, Indian Institute of Science, Bangalore, India, 560038. Email: vijayn@csa.iisc.ernet.in
- R. van de Weygaert is with Kapteyn Astronomical Institute, University of Groningen, Netherlands. Email: weygaert@astro.rug.nl
- E G P Bos is with Kapteyn Astronomical Institute, University of Groningen, Netherlands. Email: pbos@astro.rug.nl
- S. Rieder is with RIKEN Advanced Institute for Computational Science, Japan. Email: steven.rieder@riken.jp

1. A parsec is the standard unit of measurement of distances in the cosmos. A parsec is 3.26 times the *light-year*, the distance light covers in a year. A megaparsec is a million parsecs, the typical scale of measurement of size of the large scale structures in the universe.

of the nodes in the cosmic web, *i.e.*, of galaxy clusters. Recent results have opened the path towards the mapping of filaments via their lensing effect on background sources [19]. The identification of the structural components of the cosmic web is also important for our understanding of the relation between the formation, evolution, and properties of galaxies and the structural environment of the cosmic web. A direct manifestation of this is the generation of the angular momentum of galaxies. This is a product of the torquing by the large-scale tidal force field [20], [21], [22]. While these force fields are also the agents for the formation and shaping of filaments, we would expect that this results in the alignment of the spin axis of galaxies with respect to cosmic filaments [23], [24], [25].

The identification, description, and characterization of the elements of the cosmic web is a non-trivial problem. Several characteristics of the mass distribution in the cosmic web have made it an extremely challenging task to devise an appropriate recipe for identifying them:

a) The cosmic web is a complex spatial pattern of connected structures displaying a rich geometry with multiple morphologies and shapes.

b) There are no well-defined structural objects at a single spatial scale or within a specific density range. Instead, elements of cosmic web are found at a wide range of densities and spatial scales. This is a consequence of the hierarchical evolution of structure formation in the universe, such that smaller high-density structures merge to form larger objects.

c) There is a clear anisotropy in the structures of the cosmic web, a consequence of gravitational instability. The structures in the cosmic web exhibit elongated and flattened characteristics.

The attempts to analyze the structure of the cosmic web have a long history. The absence of an objective and quantitatively accurate procedure for identifying and isolating the components of the cosmic web has been a major obstacle in describing it. In recent years, more elaborate and advanced techniques have been developed to analyze and describe the structural patterns in the cosmic web. Nonetheless, a consensus on the proper definition of filaments is yet to be achieved. In the subsequent subsection 1.1 we present a short account of the available techniques and the definitions on which they are based.

1.1 Related Work

Statistical measures such as the auto-correlation function [26] of the matter distribution in the web have been the mainstay of cosmological studies over many decades. However, while this second-order measure of clustering does not contain any phase information (one may e.g. always reproduce a distribution with the same 2nd order moments and random Fourier phases), the auto-correlation function is not sensitive to the existence of complex spatial patterns. Higher order correlation functions only contain a very limited amount of such structural information, while in practical observational circumstances, it quickly becomes cumbersome to measure them as the magnitude of the error increases drastically with increasing order.

The first attempts towards characterizing complex geometric patterns in the galaxy distribution mainly involved heuristic measures. Early examples of techniques addressing the global connectivity of structure in the Universe are percolation analysis [27] and the minimum spanning tree of the spatial galaxy distribution [28], [29]. While these are useful global descriptions, they do not capture and describe local characteristics of the mass distribution.

More elaborate and advanced techniques have been developed in recent years. Several of these methods apply sophisticated mathematical and visualization techniques, involving geometric and topological properties of the cosmic mass distribution. There are a multitude of different methods for detecting filaments, based on a range of different techniques. We may recognize several categories of techniques.

One class of methods seeks to describe the local geometry on the basis of the Hessian of the density field [30], [31], [32], [57] or closely related quantities such as the tidal force field [33], [34] or the velocity shear field [35], [57]. The Hessian provides direct information on the local shape and dynamical impact of the corresponding field. The morphological elements of the cosmic web are identified by connecting the areas within which a specific range of anisotropies is registered.

These studies concentrate on a single scale by appropriately smoothing the field, and do not consider the multi-scale nature of the cosmic mass distribution. The Hessian based Nexus/MMF technique [31], which was perfected into a versatile and parameter-free method [57], implicitly takes into account the multi-scale nature of the web-related fields. It accomplishes this by a scale-space analysis of the fields. At each location the optimal morphological signal is extracted via the application of a sophisticated filter bank applied to the Hessian of the corresponding fields in scale space. The application of this machinery has enabled thorough studies of the hierarchical evolution and buildup of the cosmic web [36].

A promising and highly interesting recent development has opened up the path towards dynamical analysis of the evolving mass distribution in full six-dimensional phase-space (in which the position of each mass element is specified by its space coordinates and velocity/momentum). In the 6D phase space, the cosmic mass distributions defines a 3D sheet. Independently, three groups arrived at tessellation based formalisms that exploit the evolving structure and folding of the *phase space sheet* in phase space [37], [38], [39] (also see e.g. [40]). The number of folds of the phase space sheet at a given location indicates the number of local velocity streams, and forms a direct indication of the morphology of the local structure. Interestingly, the resulting characterization of the web-like distribution, the Origami formalism of Neyrinck [39] for example, appears to resemble that of the Nexus/MMF formalism [36].

An entirely different class of techniques, based on statistical methods, have also been used to recover the filamentary patterns in the Universe. The key idea behind these techniques is to treat the galaxy distribution as a Markov point process. Within this class, particularly worth mentioning is the Bisous or the Candy model, which has been used by Stoica *et al.* [41]

to detect the filamentary network². In this model, one places a random configuration of interacting geometric cylindrical objects on the point process to detect filamentary structure. It has been developed into a versatile, statistically solid yet computationally challenging formalism for the identification of filaments in a spatial point distribution, such as N-body simulations and galaxy redshift surveys [43]. An additional example of a method involving statistical analysis of a geometric model is that of Genovese *et al.* [44], which seeks to describe the filamentary patterns of the cosmic web in a non-parametric way by recovering the medial axis [45] of the point-set of galaxies.

The fourth major class of methods, the one which we will also pursue in this paper, exploits the topological structure of the cosmic mass distribution. The fundamental basis of these methods is Morse theory [46]. The geometric structure of the *Morse-Smale complex* [47] naturally delineates the various morphological components on the basis of the connections between the critical points of the density fields and the higher dimensional cells that are incident on the critical points. Various Morse theory based formalisms have been applied to the identification of components of the cosmic web. One of the first applications concerned the detection of voids in the cosmic density field. The Watershed Void Finder [48] identifies these with the watershed basins around the density minima. The SpineWeb procedure [49] extended the watershed transform towards the detection of the full array of structural components, filaments, walls and voids. These techniques use a user-defined filter to incorporate the multi-scale structure of the cosmic density field.

A natural topological means to address the multi-scale topological structure emanates from the concept of persistence [50]. It provides a natural recipe for detecting and quantifying the components of the cosmic web in a truly hierarchical fashion. Sousbie [51], [52] has exploited and framed this in an elegant and impressive framework, the DisPerSE formalism. Following the construction of the Morse-Smale complex, they proceed to simplify it. The simplification proceeds by canceling pairs of critical points iteratively, where each pair represents a structure in the cosmic web. Topological persistence is invoked to order the critical point pairs. However, this measure of importance is not unique, and one may consider alternatives, dependent on the specific interest and purpose.

In effect, to tackle similar issues in other visualization areas, a range of variations have been proposed in other studies [53], [54], [55]. Weinkauff *et al.* [55] describe the concept of separatrix persistence, where they compute the strength of separation of points on a separatrix curve (in 2D) connected to a saddle as the sum of the absolute differences of function values of the saddle and the extrema connected to it. This concept is extended to 3D separating sheets by Gunther *et al.* [53]. Reininghaus *et al.* [54] develop the concept of scale-space persistence where they accumulate the absolute difference in function value measure of critical points across a hierarchy of derived functions. The set of derived functions are generated

by smoothing the function using a family of Gaussian kernels of increasing variances. This is similar to the Multi-scale Morphology Filter Nexus/MMF [31], [57] described above. Both methods adopt the scale space formalism as the first step to detect features at multiple scales. However, scale space persistence and separatrix persistence, disregard specific density regimes of interest and are potentially inappropriate when small scale features with specific density characteristics are of interest.

1.2 Present study: contributions

In the present study, we describe and introduce a technique for the identification of filaments based on the topological characteristics of the density field. A key aspect of the proposed technique is its interactive nature, involving a tunable density parameter. Specifically, we describe the following contributions:

a) We describe Felix³: a topology based framework for visual exploration of filaments in the cosmic web. In particular, we develop a query framework to extract filamentary structures from a hierarchy of Morse-Smale complexes of the density field. The filaments in Felix are parameterized by the density values of the maxima and the 2-saddles that define them.

b) Using Felix, we develop a semi-automatic structure finder that classifies galaxies as cluster/filamentary or not. We demonstrate its efficiency through two tests. First, using the Voronoi Kinematic model as a benchmark, we demonstrate that we are able to recover the classification with high efficiency. Second, we show that the classifications are quantitatively comparable to, and in several cases better than, existing classifiers.

c) We investigate the nature of filaments in two different density regimes from the Λ CDM simulations. The first concerns filaments in the high density regions around compact dense clusters, which are known to function as the transport channels along which matter moves into the clusters. A second regime concerns the tenuous low-density filaments found in low-density void regions. In the supplemental document, we present an additional experiment, where we investigate the nature of three classes of filaments in a relatively cleaner region of a Λ CDM dataset.

d) In the supplemental document, we describe an efficient structure based volume rendering enhancement routine that allows us to highlight the density distribution in regions that are close to the selected features.

The distinction between noise and significant structures is often ill-defined, and at occasions noise may be confused with genuine structures in the hierarchically evolved mass distribution (see Figure 4 and the caption thereof for an illustration). This problem is more pronounced when one studies the properties of tenuous filaments and walls in low density void-like regions. For the understanding of the formation and evolution of galaxies in such regions, we need to assess the possible dependence of galaxy and halo properties on the morphology and density of the local environment. This must be based on the successful extraction of filaments in

2. See Stoica *et al.*[42], for a detailed description of this object point process which is used to characterize an observed point distribution by a fitting procedure using global optimization techniques.

3. The name Felix is formed from an abbreviation of *Filament explorer*.

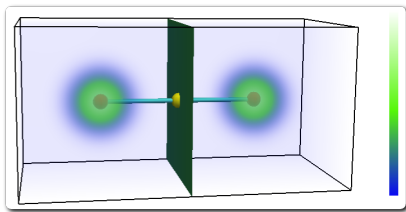


Fig. 1: Ascending manifolds of a 2-saddle (yellow sphere). The scalar function is a sum of two 3D Gaussians centered on either side of the volume. The two arcs incident on the 2-saddle constitute the ascending manifold and terminate at the two maxima (red spheres) of the scalar function.

low density regions and the correct identification of galaxies associated with them. In view of this, we include an interactive handle on the density regimes so that one can concentrate on and probe structures in specific density regimes.

The remainder of the paper is organized as follows. Section 2 introduces necessary background material. Section 3 describes Felix. Section 4 introduces the cosmological datasets used in the experiments. Section 5 discusses the application of Felix as a structure finder and as a tool for exploring filaments in different density regimes. Section 6 concludes the paper by summarizing the main results and possible future directions.

2 BACKGROUND

This section reviews relevant background on Morse functions, the Morse-Smale complex, and topological simplification. This is a necessary prerequisite for understanding the definition of filaments and extraction methods described in the subsequent sections.

2.1 Morse theory and the Morse-Smale complex

Let $f: \mathbb{M} \rightarrow \mathbb{R}$ be a real-valued scalar function defined on a manifold \mathbb{M} . Critical points of f are points of f where the gradient of f vanishes *i.e.*, $\nabla f = 0$. Morse theory is the study of the relationship between the topology of level sets of scalar functions and the critical points of the function. The function f is said to be a *Morse function* if all of its critical points are non-degenerate *i.e.*, the *Hessian* of f , equal to the matrix of second order partial derivatives, is non-singular. The non-degeneracy condition imposes a locally quadratic form for f within a small neighborhood of its critical points. In other words, using a coordinate transformation, the function near a critical point p of the n -dimensional manifold \mathbb{M} can be written as $f_p(x) = f(p) \pm x_1^2 \pm x_2^2 \pm \dots \pm x_n^2$. The *index* of p is equal to the number of negative quadratic terms in the above expression. In 3D, the index 0 corresponds to minima, the index 1 corresponds to 1-saddles, the index 2 corresponds to 2-saddles, and the index 3 corresponds to maxima. An *integral line* is a maximal curve in the domain, whose tangent aligns with the gradient of f at every point. The function f increases along the integral line and its limit points are the critical points of f .

The set of all integral lines that originate at the critical point p together with p is called the *ascending manifold* of p . Similarly, the set of all integral lines that terminate at

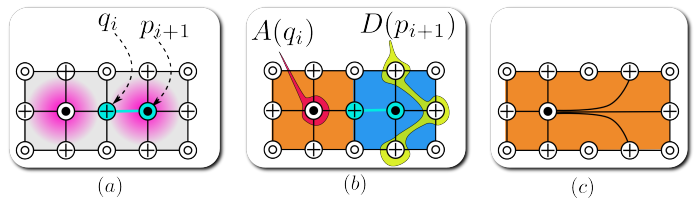


Fig. 2: Topological cancellation of a pair of critical points in a 2D Morse-Smale complex. (a) Morse-Smale complex of the function shown in Figure 1 restricted to a 2D slice. Maxima are denoted by \ominus , saddles by \oplus , and minima by \odot . A pair (p_{i+1}, q_i) of critical points connected by a single arc is scheduled to be canceled. (b) $D(p_{i+1})$ is the set of surviving index i critical points connected to p_{i+1} and $A(q_i)$ is the set of surviving index $i+1$ critical points connected to q_i . (c) Combinatorial realization: connect all critical points $D(p_{i+1})$ to those in $A(q_i)$. Geometric realization: merge the descending manifold of p_{i+1} with those of critical points in $A(q_i)$. Merge the ascending manifold of q_i with those of critical points in $D(p_{i+1})$.

the critical point p together with p is called the *descending manifold* of p . The ascending manifolds (similarly, the descending manifolds) of all critical points partition the domain. The ascending manifold of a critical point with index i has dimension $n-i$, where n is the dimension of the domain. Thus, the ascending manifold of a minimum is a three dimensional cell, the ascending manifold of a 1-saddle is a two dimensional sheet, the ascending manifold of a 2-saddle is a one dimensional arc (see Figure 1), and the ascending manifold of a maximum is equal to the maximum. The converse is true for the descending manifold *i.e.*, the descending manifold of a critical point with index i has dimension i . The *Morse-Smale complex* is a partition of the domain into cells formed by the collection of integral lines that share a common source and a common destination. The function f is called a *Morse-Smale function* if the ascending and descending manifolds of all pairs of critical points intersect only transversally *i.e.*, if the ascending and descending manifolds of two critical points intersect, then the intersection has dimension exactly equal to the difference in the indices of the two critical points. The critical points, referred to as *nodes*, together with the 1-manifolds that connect them, referred to as *arcs*, form the 1-skeleton of the Morse-Smale complex, which is referred to as the *combinatorial structure* of the Morse-Smale complex.

2.2 Morse-Smale complex simplification

The Morse-Smale complex may be simplified by repeated application of the *topological cancellation* procedure. Topological cancellation eliminates a pair of critical points in the Morse-Smale complex connected by a single arc. The resulting complex is representative of a smoother version of f , which may be obtained by a local smoothing operation within an infinitesimal neighborhood of the arc. The canceled pair of points are no longer critical in the smoother version of f . By repeatedly applying topological cancellations, one obtains a simpler Morse-Smale complex, where undesirable pairs of critical points are eliminated. This simplified Morse-Smale complex is representative of the function f with several

smoothing operations applied to it. Pairs of critical points that are multiply connected in the Morse-Smale complex, referred to as *strangulations*, cannot be simplified by topological cancellation.

The cancellation procedure mandates changes to the combinatorial structure and ascending/descending manifolds of critical points that survive the cancellation. These changes are respectively referred to as the combinatorial and geometric realization of the cancellation. Let p_{i+1} and q_i denote the pair of critical points to be eliminated with $i+1$ and i being the respective Morse indices. The combinatorial realization proceeds by first removing p_{i+1} and q_i as well as all arcs incident upon them. Next, new arcs are introduced between every surviving index i critical point that was connected to p_{i+1} and every surviving index $i+1$ critical point that was connected to q_i . The geometric realization merges the descending manifold of p_{i+1} with the descending manifold of each surviving index $i+1$ critical point connected to q_i . Analogously, the ascending manifold of q_i is merged with the ascending manifold of each surviving index i critical point connected to p_{i+1} . Figure 2 illustrates both realizations of a topological cancellation procedure applied to the Morse-Smale complex of a two-dimensional slice of the function shown in Figure 1.

The ordering of cancellation pairs plays a crucial role in determining the resulting structure of the Morse-Smale complex and its geometry. To simplify the Morse-Smale complex, pairs of singularly connected critical points having the least absolute difference in function value are iteratively canceled. This approach is equivalent to the notion of topological persistence [47], [50] for 2D Morse-Smale complexes, but not necessarily for 3D Morse-Smale complexes [56].

2.3 The hierarchical Morse-Smale complex

A sequence of cancellations results in a hierarchical sequence of Morse-Smale complexes $MSC_0, MSC_1, \dots, MSC_n$, where each Morse-Smale complex is a simpler version of the preceding Morse-Smale complex containing fewer critical points. Morse-Smale complex MSC_i is said to be coarser than MSC_j if $i > j$ and finer if $i < j$. The version index i enumerates the Morse-Smale complexes in the hierarchy. Each non-zero version of the Morse-Smale complex, MSC_i , is associated with the absolute difference in function value, t_i , of the pair of critical points canceled in the preceding version, MSC_{i-1} . As each iteration selects the pair of critical points with the least absolute difference in function value, the sequence of t_i 's is monotonically increasing i.e. $(t_0 = 0) \leq t_1 \leq t_2 \dots \leq t_n$. For completeness of the sequence, t_0 is set to zero. Figure 3 illustrates a hierarchy of Morse-Smale complexes of a 2D equivalent of the function shown in Figure 1.

It is not necessary to explicitly store all versions of the hierarchy of Morse-Smale complexes. Instead, the combinatorial representation of only MSC_0 is computed initially. Subsequently, MSC_i can be obtained from a finer MSC_{i-1} by performing topological cancellation (see Section 2.2). Analogously, MSC_i is obtained from a coarser MSC_{i+1} by applying the inverse operation of a topological cancellation.

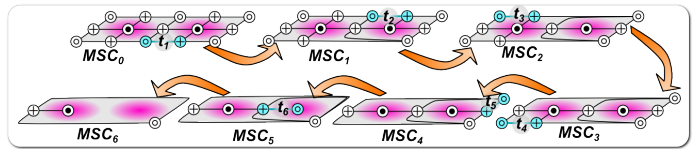


Fig. 3: Hierarchical Morse-Smale complex. A family of Morse-Smale complexes generated by iteratively canceling pairs of critical points. MSC_0 is the Morse-Smale complex of a 2D equivalent of the function shown in Figure 1. It is simplified to generate a coarser version, MSC_1 , by canceling a pair of critical points (cyan) connected by a single arc and having least absolute difference in function value t_1 . Successive versions MSC_i are computed similarly by selecting arcs so that $t_0 = 0 \leq t_1 \leq \dots \leq t_6$.

3 METHODOLOGY

Exploring the filamentary patterns of the cosmic web is challenging because of the large range of the spatial scales and density range it exhibits. A proper characterization should also account for the hierarchical nature of structures, which adds considerable challenges to the task. Though there exist different notions of filaments, the primary evidence relied upon for extraction and analysis is most often visual. It is therefore not surprising that structure finding methods often visually verify results by superimposing the extracted structures upon visualizations of the density field or the particle distribution. However the visualization plays a role only after structure extraction process in these methods. We differ in this respect by providing the capability to interact with the structure finding procedure and extract structures that are visually relevant. To accomplish such a visual exploration framework, a succinct model of filament definition, an efficient representation of hierarchical structures, and an appropriate query mechanism that supports the extraction of these structures are paramount. The following exposition details our framework on these terms.

3.1 Density estimation and filament modeling

Cosmological simulations are N-body particle experiments that simulate structure formation and evolution by tracing positions of the particles under the influence of physical laws. In the observational reality, the information about structures in the cosmos comes through observing the galaxies. The galaxies can be treated as particles also for the purpose of analysis in the context of large scale structures.

The input to Felix is the logarithm of the density scalar field on the given 3D domain of interest. The domain could be 3D structured grids or tetrahedral meshes, with the density specified on the vertices of the grid/mesh. We find that the logarithm of the density field, instead of the density field itself, resolves the structures with more visual clarity. This has also been independently established in an earlier work [57]. Additionally, the input may be specified as a distribution of particles within a 3D region of interest. This could be a snapshot from a cosmological simulation, or galaxies in real observational data. We use the Delaunay tessellation field estimator (DTFE) [58], [59] to estimate the density of the input

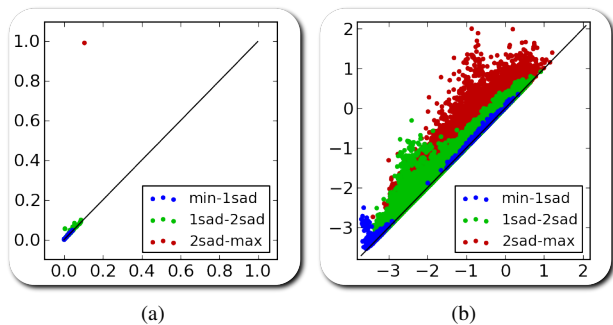


Fig. 4: (a) A scatter plot of the function values of the canceled critical point pairs for the function shown in Figure 1. A 2-saddle-maximum pair is the only pair that is far removed from the diagonal. This corresponds to cancellation of the 2-saddle with a maximum that represents one of the Gaussians in Figure 1. Other pairs close to the diagonal represent insignificant features that manifest due to the added Gaussian noise as well as sampling noise. (b) A scatter plot of the function values of the canceled critical point pairs for the Voronoi-Kinematic dataset B (see Section 4.1). No discernible separation of points is seen, though there are many points that are far removed from the diagonal. Thus, no clear global simplification threshold may be used for filament extraction.

particles in the 3D region of interest. This procedure begins by computing the periodic Delaunay triangulation[60] on the points (simulation particles or galaxies). Next, the density at each vertex of the triangulation is estimated by the inverse of the volume of the tetrahedra incident upon it. Finally, the density is linearly interpolated onto the edges, faces, and tetrahedra of the Delaunay triangulation to yield a piecewise linear density function on the domain. The periodic Delaunay triangulation, computed by the DTFE procedure, is used to represent the domain.

The Morse-Smale complex of the logarithm of the density field is computed. Filaments are modeled as the ascending manifolds of 2-saddles of the Morse-Smale complex. These arcs represent paths of steepest descent from the two maxima merging at the 2-saddle. This 2-saddle represents the lowest density point along the arcs connecting the two maxima. A schematic illustration of this is presented in Figure 1. There are many algorithms available in the literature to compute the 3D Morse-Smale complex. The algorithms are primarily based on either the quasi Morse-Smale complex formulation [61], [62] or Forman’s [63] discrete Morse theory[64], [65], [66], [67]. We use a parallel algorithm based on the latter approach [67] resulting in fast computation even for large datasets.

The density field is rarely smooth and several local maxima obscure a view of the larger scale behavior of the density field. This is especially true if the density field is computed on the raw particle distribution, where the density field tends to be spiky and with a lot of fluctuations in the high density cluster-like regions. The Morse-Smale complex is simplified by iteratively canceling pairs of singularly connected critical points with least absolute difference in function value to generate a hierarchy of Morse-Smale complexes.

In most applications, a specific version of the Morse-Smale

complex from the hierarchy is chosen based on a perceptibly clear separation of noise and features. One way to choose such a threshold separating noise and feature is by using a scatter plot of the function values of canceled critical point pairs (see Figure 4) where the lower function value among the pair corresponds to the x-coordinate and the higher function value corresponds to the y-coordinate. In datasets where topological features are well separated (see Figure 4a), pairs representing significant features appear far away and isolated from the diagonal. In such cases, the coarsest Morse-Smale complex wherein the insignificant pairs are removed is selected for feature analysis/extraction. However, this strategy is not easily applicable to cosmology datasets (see Figure 4b). A well defined separation is rarely discernible, though there are many scatter points that are far removed from the diagonal. Hence, we drop the assumption that we must work with a specific version of the Morse-Smale complex. Instead, we query for features across all Morse-Smale complexes in the hierarchy, as discussed in the following sub-section.

3.2 Density range based filament selection

Cosmic filaments exhibit a large range of variation in their density characteristics. Indeed, one expects filaments to be present both in void like regions and between cluster like regions. While strong dense filaments in between clusters define the spine of the Cosmic Web, in the hierarchically evolving mass distribution we encounter a wide spectrum of ever more tenuous filaments on smaller mass scales. Small filaments define the directions of mass inflow into galaxies, and form a crucial component in the formation of galaxies [68]. Even more tenuous are the systems of filaments stretching over the hollows of voids, often conspicuously aligned along the direction defined by neighbouring superstructures. The understanding of this network is tightly related to the issue of the “missing” dwarf galaxies in voids [69]. While this illustrates the complexity of the multiscale filigree of filaments in the Cosmic Web, we follow a strategy in which we focus our attention on specific aspects and details of the cosmic web. Dependent on the identity of objects and structures of interest, we wish to be able to zoom in on to the corresponding filamentary network. This is largely dependent on the mass scales of the objects involved, and the density values of the corresponding filament generating density peaks [70], [36].

Following this rationale, we translate this strategy into the use of queries that depend on the density properties of interest. Specifically, we query for filaments by specifying the density range $[M_b, M_e]$ of the clusters they connect (the maxima at the end points), as well as the density range $[S_b, S_e]$ of the lowest point along the connecting path (the density range of the 2-saddles). Figure 5 conceptually illustrates the characterization of filaments using density ranges, where density along filaments varies significantly necessitating simplification.

Algorithm SELECT2SADDLES lists the algorithm to process such a query. The algorithm accepts, together with the combinatorial Morse-Smale complex MSC , the density ranges of 2-saddles $[S_b, S_e]$ and maxima $[M_b, M_e]$ as input (the subscripts $_b$ and $_e$ denote the beginning and ending of each density range

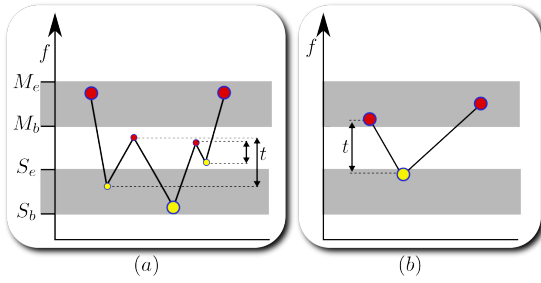


Fig. 5: (a) Filaments are modeled as the ascending paths of 2-saddles connecting two extrema. The 2-saddles are filtered based on the range constraints $[M_b, M_e]$ and $[S_b, S_e]$ on the highest and lowest values respectively along the ascending paths. The highest values along the 2-saddle's ascending manifold are at extrema and the lowest value is at the 2-saddle. The function along the paths needs to be simplified as it is rarely smooth. In the illustration, a simplification threshold of t reveals a filament with appropriate density characteristics. However, imposing such a threshold uniformly will cause another filament (b) having the required density characteristics to be destroyed. It is therefore necessary to extract filaments by querying all Morse-Smale complexes within a given hierarchy.

respectively). The algorithm returns a list of 2-saddles that satisfy the above criteria together with the maximal Morse-Smale complex version in which they do so.

Algorithm SELECT2SADDLES ($[S_b, S_e]$, $[M_b, M_e]$)

```

1:  $S_{ver} = \text{Empty Map}$ 
2:  $S = \{s | s \text{ is a 2-saddle, } S_b < f(s) < S_e\}$ 
3: for all  $s \in S$  do
4:    $v_a = \text{Max}\{i | s \text{ is not simplified in } MSC_i\}$ 
5:    $v_b = \text{Max}\{i | t_i < M_b - f(s)\}$ 
6:    $v_c = \text{Max}\{i | s \text{ connects distinct Maxima in } MSC_i\}$ 
7:    $S_{ver}[s] = \text{Min}(v_a, v_b, v_c)$ 
8: Sort  $S$  by  $S_{ver}$ 
9:  $S_{sel} = \text{Empty Set}$ 
10: for all  $s \in S$  do
11:    $i = S_{ver}[s]$ 
12:    $m_a, m_b = \text{Maxima connected to } s \text{ in } MSC_i$ 
13:   if  $M_b < f(m_a), f(m_b) < M_e$  then
14:     insert  $(s, S_{ver}[s])$  in  $S_{sel}$ 
15: return  $S_{sel}$ 

```

The algorithm begins by creating a list S of 2-saddles that have their function value in the given 2-saddle range $[S_b, S_e]$. Then, for each 2-saddle in S , a Morse-Smale complex version in which it possibly connects two maxima within $[M_b, M_e]$ is computed. The appropriate version is given by the minimum of three version indices v_a, v_b and v_c .

The version index v_a is the finest Morse-Smale complex version in which the 2-saddle s survives. In other words, s is canceled in MSC_{v_a+1} but not in MSC_{v_a} . This is pre-computed by examining the cancellation sequence. The version index v_b corresponds to the last Morse-Smale complex version at which the 2-saddle s connects two maxima, both with function value less than M_e . This is possible because in successive

versions of the Morse-Smale complex, the maxima connected to a 2-saddle via the same arc form an increasing sequence in terms of their function value. Thus, in the version where the absolute difference in function value of the last canceled pair is less than $M_b - f(s)$, the 2-saddle s still possibly connects two maxima with function value less than M_e . The version index v_c is the last Morse-Smale complex version at which the 2-saddle s separates distinct maxima. In other words, it is not a strangulation in MSC_{v_c} . As a consequence of the cancellation preconditions, once a strangulation is created by a 2-saddle, it may be destroyed only by canceling the 2-saddle with a 1-saddle. Thus there exists a maximal version index v_c after which the 2-saddle remains connected to a single maximum. The version index v_c is -1 when the 2-saddle is a strangulation in the initial Morse-Smale complex. In this case, the 2-saddle is not considered in further steps and is removed from S . Again, this is easily pre-computed for each 2-saddle by examining the cancellation sequence.

The 2-saddles in the set S are sorted based on their version indices. This is done to optimize switching between the required Morse-Smale complex versions. Next, each 2-saddle s in S is checked to see if it separates two maxima within the maxima density range $[M_b, M_e]$. The list of 2-saddles that fulfill all of the above criteria is returned together with the associated version number of each 2-saddle. This above list of 2-saddles is used to extract the filament geometry. Specifically, the ascending manifold of each 2-saddle is extracted from the corresponding version of the Morse-Smale complex. This may be done efficiently using the *cancellation merge DAG* data structure discussed by Gyulassy *et al.* [71].

In some situations, it is desirable to perform some simplification to eliminate Poisson noise introduced due to meshing the domain. In these cases, a global simplification specifically for noise elimination, can be optionally introduced. Specifically, Algorithm SELECT2SADDLES returns only those 2-saddles that survive in hierarchical Morse-Smale complex versions above a specified threshold T_s , where T_s is specified as a normalized fraction of the range of log-density values (normalized to $[0, 1]$). T_s is set to 0.0 unless specifically mentioned. Similar to the inputs of Algorithm SELECT2SADDLES, T_s may be updated during run-time.

3.3 Volume Rendering

We use volume visualizations of the density field to aid selection of parameters for Algorithm SELECT2SADDLES. The geometry of the selected filaments using Algorithm SELECT2SADDLES is superimposed upon a volume rendering of the density field. Based on the visualization of the extracted filaments and the density volume rendering, the parameters may be adjusted so that the structures correspond with the density volume rendering. Figure 6 shows an example of the overlay of the volume visualization with the selected structures. Direct volume rendering of the density is often not effective for visualization because of the formation of clusters at multiple scales. Furthermore, these clusters are often spatially far removed from the features of interest. In the supplemental document, we describe an application of

our framework to suppress the opacity of regions spatially far removed from filamentary features of interest, which leads to a feature based volume rendering enhancement.

4 MODEL DESCRIPTION

In this section, we briefly describe the models used to test our filament detection routine. These are the Voronoi evolution models and Λ CDM cosmological simulations.

The Voronoi models provide us with vital quantitative information on the sensitivity of Felix to anisotropic filamentary patterns in the galaxy distribution. To this end it is of key importance that the Voronoi models have an a-priori known population fraction in different morphological elements: clusters, filaments, walls and voids. This makes them perfect test models for evaluating success and failure rates of the various identification methods.

Although they involve filaments with a broad distribution of densities, the Voronoi models do not incorporate the multi-scale web-like patterns seen in realistic cosmological scenarios. To assess this aspect of the cosmic mass distribution, we turn to simulations of structure formation in the standard Λ CDM cosmology. Implicitly these include all relevant physical and dynamical processes of the evolving cosmic dark matter distribution. However, as we have no control over all aspects of the emerging mass distribution in Λ CDM simulations, for testing purposes they are not as informative as Voronoi models ⁴

4.1 Voronoi evolution models

The Voronoi evolution models are a class of heuristic models for cellular distributions of galaxies that mimics the evolution of the Megaparsec universe towards a weblike pattern. They use Voronoi tessellations as a template for distribution of matter and related galaxy population [76], [77], [74], and its subsequent evolution.

In these models, one begins by fixing an underlying Voronoi skeleton, defined by a small set of randomly distributed nuclei in the simulation box. One then superposes a set of N randomly distributed particles on this skeleton. The resulting spatial distribution of particles in the model is obtained by projecting the initially random distribution of particles on to the faces, edges, and nodes of the Voronoi tessellation. This results in a pattern in which one can distinguish four structural components: *field* particles located in the interior of Voronoi cells, *wall* particles within and around the Voronoi faces, *filament* particles within and around the Voronoi edges and *cluster* particles within and around the Voronoi nodes.

One particular class of Voronoi clustering models are the Voronoi kinematic models, which seek to approximate the dynamical evolution of the large scale cosmic mass distribution. These models involve a continuous flow of galaxies towards the nearest wall, along a filament at the wall's edge, and subsequently towards the final destination, a vertex of the

4. In the literature, several studies use simplistic models using Voronoi tessellations. The models we use here are considerably more sophisticated, and represent a rather realistic depiction of the cosmic web in void-dominated cosmologies, see e.g. [72], [73], [74], [75].

	cell	wall	filament	cluster
A	29.88%	43.57%	22.20%	4.33%
B	9.82%	32.13%	38.62%	19.42%
C	3.5%	16.50%	28.70%	51.30%

TABLE 1: The relative abundance of particles in each structural element throughout the course of evolution.

Voronoi tessellation. This motion is regulated by the increase of mean distance between the galaxies, an expression of void expansion and evacuation as a function of time.

The Voronoi models used in our experiments have 262,144 particles distributed along the vertices, edges, faces and cells of the Voronoi skeleton in a box of side-length $200h^{-1}$ Mpc. The skeleton is generated by 32 randomly placed nuclei in the box. For the least evolved stage, most of the particles are in the cells, while for the most evolved stage, most of the particles are located in and around clusters. Table 1 presents the percentage distribution of particles in the various structural elements, as it changes with time. Stage 1 (dataset A) is the least evolved, while Stage 3 (dataset C) is the most evolved. The particles in and around the nodes, edges, and walls are Gaussian distributed around these elements, characterized by a thickness scale R_f which is the standard deviation of the distribution. For our models, $R_f = 2h^{-1}$ Mpc. In the supplemental document, we present the volume renderings of the density distribution.

4.2 Λ CDM cosmological simulations

The Λ CDM simulations are fully physical models that trace the distribution and evolution of dark matter in the universe based on current understanding of real physical laws. Dark matter is the gravitationally dominant matter component in the Universe and constitutes the major fraction of matter. As it is known to only interact gravitationally, modeling the behavior of dark matter is computationally fast and efficient. Such dark matter simulations form one of the principal tools towards understanding the evolution of the matter distribution in the Universe.

The cosmological simulations that we used follow the standard Λ CDM cosmology. In this model, the matter content of this Universe is dominated by collisionless *cold dark matter* (CDM) particles. The biggest contribution to the energy content of this Universe comes from dark energy, in the form of the cosmological constant Λ (see [78]), which drives its accelerated expansion at the current epoch.

To present the results of our visual exploration framework, we use the Cosmogrid simulations [79]. It is a suite of simulations in a box of size $21h^{-1}$ Mpc, each differing in the number of particles. The particular simulation we use for our study comprises of 512^3 particles. This is a relatively small scale in the context of the cosmic web. The mass resolution achieved is $8.21 \times 10^6 M_\odot$. The initial conditions are setup at $z = 65$ using the Zel'dovich approximation [80]. The log-density field is available on a $128 \times 128 \times 128$ structured grid.

Particularly characteristic in the evolving mass distribution of the Cosmogrid simulation is the large central under-density, surrounded by a range of smaller voids near its outer edge. In combination with its extremely high spatial resolution and state-of-the-art dynamic range, this renders the Cosmogrid

simulation uniquely suited as a testbed for a case study of the internal structure of voids. It was precisely this circumstance that formed the rationale behind its exploitation in a previous study of the formation of dark halos along tenuous void filaments [81].

5 RESULTS AND DISCUSSION

In this section, we demonstrate and discuss the salient features and potential applications of Felix. First, we evaluate the filaments extracted using Felix and compare with those extracted using MMF, SpineWeb, and DisPerSE using the Voronoi kinematic datasets. Next, we present a visual exploration of the filaments in the Cosmogrid simulation. The methods described above were implemented and tested on a computer system with an Intel Xeon(R) 2.0 GHz CPU and 8GB of RAM. In the supplemental document, we briefly discuss the structure finders against which we compare Felix.

5.1 Filaments in the Voronoi model: a comparison

Here we present an analysis of the filamentary structures extracted using Felix, and compare the results with the techniques detailed above. The comparison study concerns the analysis results obtained for the set of heuristic Voronoi evolution models described above. Since they are input parameters, in these models the classification of galaxies as void, wall, filament, and cluster are known a-priori. Following the application of one of the detection techniques we may then examine the validity and authenticity of the extracted structures by direct comparison with the true identity of a galaxy.

For the comparison study, we define two measures. One quantifies the true detection rate of a method, the other the false identifications. We classify all galaxies within a distance d from the extracted structures to be filament and cluster particles and the others to be void and wall galaxies. For a given set of structures extracted from a given dataset and a distance d , the true positive classification rate TP_d is defined as

$$TP_d = \frac{\# \text{ filament and cluster galaxies correctly classified}}{\# \text{ filament and cluster galaxies}}.$$

Similarly, the false positive classification rate FP_d is defined as

$$FP_d = \frac{\# \text{ filament and cluster galaxies incorrectly classified}}{\# \text{ filament and cluster galaxies}}.$$

A large separation between these two measures indicates good discriminatory power of the classifier, and thus the proximity of relevant galaxies to the extracted structures.

As we discuss in more detail below, the Felix's true and false detection rates are comparable, and in some situations better, than those obtained by DisPerSE, SpineWeb and Nexus/MMF. A brief description of SpineWeb, and Nexus/MMF is provided in the supplemental document. As the DisPerSE methodology is closely related to Felix, in the following paragraphs, we briefly describe it and contrast it against Felix.

Felix: Figure 6 shows the extracted filaments for the Voronoi kinematic datasets A , B , and C using Felix. The input density range parameters for Algorithm SELECT2SADDLES are selected interactively, using the visualization information of the procedure. Each update is accomplished within 2-3 seconds. This enables interactive visual feedback so that parameters may be adjusted further in subsequent iterative steps.

Comparison A: Felix and DisPerSE: DisPerSE [51], [52] is a closely related structure finder that also uses the Morse-Smale complex of the logarithm of the density field. It simplifies the Morse-Smale complex using Topological Persistence. The function values are normalized by the rms of the density field fluctuation with respect to the mean, and the significance level for simplification is quoted in this unit. Felix is closely related to DisPerSE as both use the Morse-Smale complex of the log-density field and involve feature extraction from it. DisPerSE defines significant features as only those that remain unsimplified using the user defined significance threshold. It ignores the density range characteristics of the extracted features. A significant consequence is that filaments within void-like regions and cluster like regions are ignored/simplified away. If they are retained, then the mixing of features causes visual clutter. Furthermore, the significance parameter selection is a fixed constant and visual interaction plays no role in its selection. In contrast, given the ubiquity of filaments in various density regimes, Felix allows for density ranged based probes into filaments, within clusters and voids. Furthermore, the visual interactive aspect allows for user engagement in parameter selection, which is crucial for the set of features identified. Another difference is that Felix uses simplification only for noise removal and not feature identification.

In this experiment, we demonstrate the consequences of not correlating the density characteristics for filament extraction. Specifically, we demonstrate that the filaments extracted using Felix are more spatially proximal to filament and cluster particles in the Voronoi Kinematic simulation. Furthermore, we show that tuning the significance parameter is not a sufficient mechanism to extract the desired filaments in this dataset. In the next experiment, we demonstrate the exploration of filaments within high-density cluster like regions and low-density void like regions. Such a delineation, coupled with the visual exploration process, is not possible using DisPerSE.

The recovery and failure rates for the Voronoi kinematic models A , B , and C are shown in Figure 8. The TP_d and FP_d of the Felix filaments are plotted in the top-left panel, the ones for DisPerSE can be found in the top right-hand panel (1σ significance threshold), bottom left-hand panel (3σ significance threshold) and bottom-right panel (5σ significance threshold).

Felix shows good recovery rates for all datasets, particularly around $d = 3h^{-1}\text{Mpc}$. For the least evolved configuration A , and for locations where structures are least distinct, Felix still obtains moderately good recovery rates. The results for DisPerSE with simplification thresholds 1σ and 3σ are comparable: at short distances the true detection rate is slightly lower than that of Felix, while at larger distances it performs

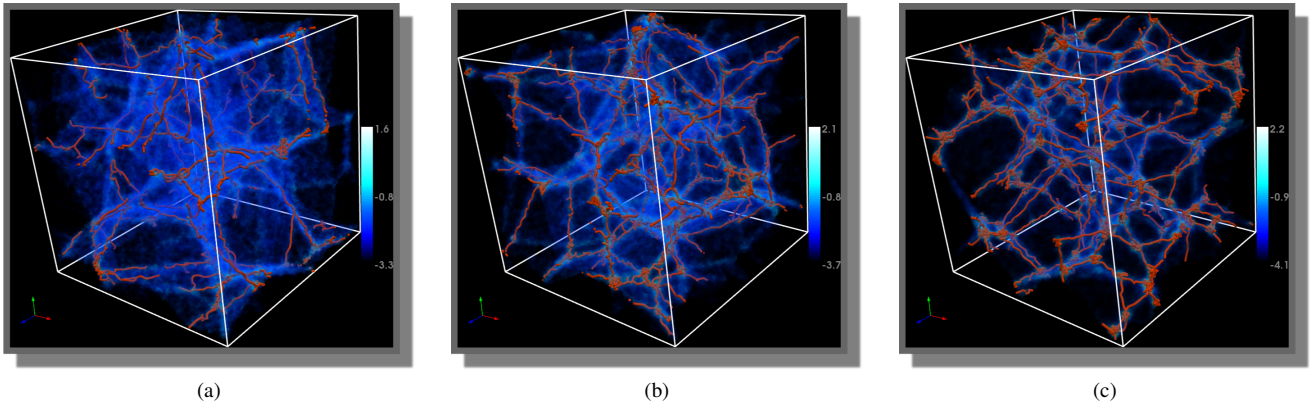


Fig. 6: Filamentary structures extracted from datasets A, B, and C of the Voronoi evolution time-series using Felix with parameters for Algorithm SELECT2SADDLES set as follows: (a) $[S_b, S_e] = [10^{-1.6}, \infty]$ and $[M_b, M_e] = [10^{-0.1}, \infty]$, (b,c) $[S_b, S_e] = [10^{-1.6}, \infty]$ and $[M_b, M_e] = [10^0, \infty]$. Filaments are shown as orange tubes along with a volume rendering of the log-density field. The dense knot like structures show filaments within cluster-like regions.

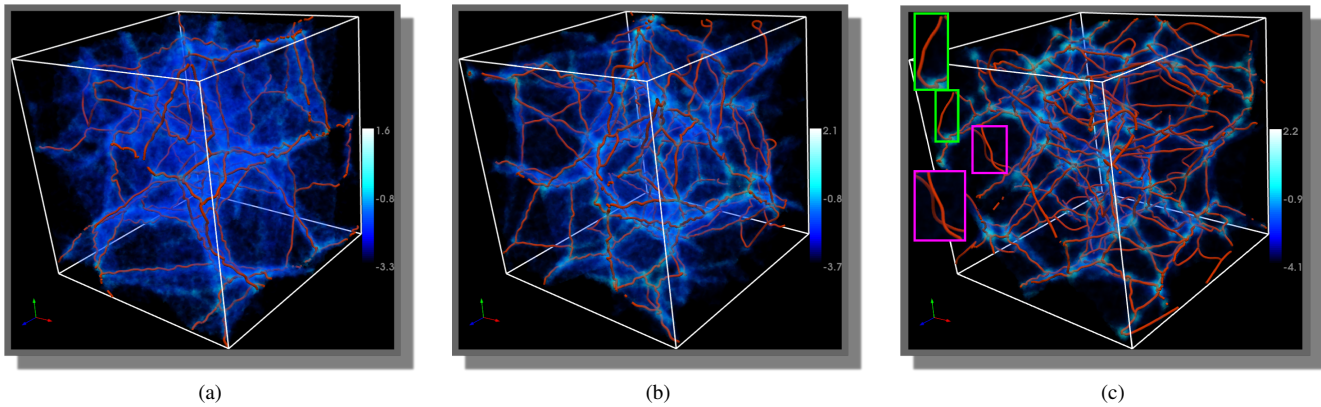


Fig. 7: Filamentary structures extracted from datasets A, B and C of the Voronoi evolution time-series using DisPerSE with significance level of 5σ . Filaments are shown as orange tubes along with a volume rendering of the log-density field. The inset pictures show identified filaments that are within wall-like and void-like regions of the Voronoi kinematic datasets.

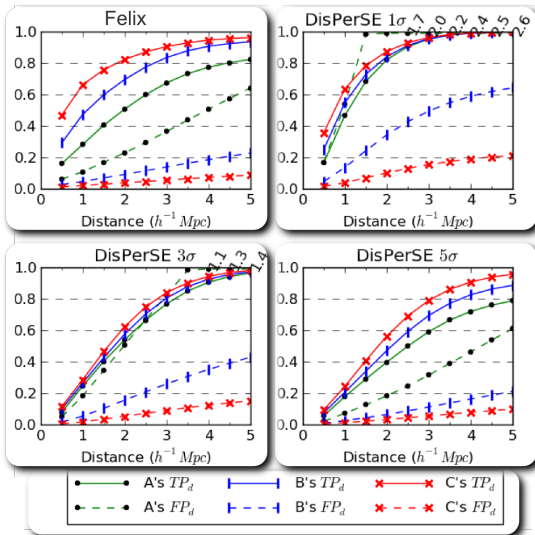


Fig. 8: Classification recovery rates for Voronoi kinematic datasets A, B, and C. Top left: Felix; top right: DisPerSE, 1σ significance level; bottom left: DisPerSE, 3σ significance level, and (bottom right) DisPerSE with a 5σ significance level. False positive rates greater than 1.0 are clipped and respective values are shown.

marginally better. The situation is slightly different in the case of DisPerSE with a 5σ simplification threshold. In the case of the more strongly evolved *B* and *C* datasets, DisPerSE and Felix have similar false detection rates FP_d , while the true detection rates TP_d of Felix are consistently higher.

For the lesser evolved datasets *A* and *B*, the false detection rates FP_d for DisPerSE quickly increase to rather large values. For simplification thresholds of 1σ and 3σ it even surpasses values of unity. This may indicate that in certain circumstances an automatic detection of filaments from the Morse-Smale complex runs the risk of over-determining the population of filaments, even after considerable simplification. While the problem is not so acute in the most evolved stage *C*, where the morphologies are well separated, direct simplification strategies may not always succeed in properly classifying all filament, wall and void regions in the more moderately evolved stages *A* and *B*.

Figure 7 shows the filaments in the three Voronoi models detected by DisPerSE, with a simplification threshold of 5σ . In comparison with the structures in Figure 6, the knot like structures present in clusters are absent. This leads to the cluster particles being far away from the filament end points, and thus the reduced TP_d rates of DisPerSE. In contrast,

Dataset		TP_d	FP_d
A	<i>Felix</i> ($d = 3h^{-1}\text{Mpc}$)	0.67	0.36
A	<i>DisPerSE</i> ($5\sigma, d = 3h^{-1}\text{Mpc}$)	0.59	0.31
B	<i>Felix</i> ($d = 3h^{-1}\text{Mpc}$)	0.84	0.13
B	<i>DisPerSE</i> ($5\sigma, d = 3h^{-1}\text{Mpc}$)	0.69	0.11
	<i>Nexus/MMF</i>	0.85	0.13
C	<i>Felix</i> ($d = 3h^{-1}\text{Mpc}$)	0.90	0.05
C	<i>DisPerSE</i> ($5\sigma, d = 3h^{-1}\text{Mpc}$)	0.78	0.05
	<i>SpineWeb</i>	0.87	0.10

TABLE 2: Recovery rates of galaxies within $3h^{-1}\text{Mpc}$ of structures extracted using Felix compared with Nexus/MMF, SpineWeb, and DisPerSE (5σ significance level).

Felix’s ranged query allows us to retain only the filaments in cluster like regions and those that connect these cluster like regions, leading to better TP_d rates. Also, many filaments found by DisPerSE are within the wall-like and void-like regions of the Voronoi Kinematic simulation. Some examples are highlighted using insets in Figure 7. Again these are filtered out by Felix’s ranged query, which is not directly possible in DisPerSE. The inclusion of such structures in DisPerSE leads to its higher FP_d rates.

These findings suggest that a structure identification strategy based on a direct simplification procedure of the Morse-Smale complex should be applied with care to the density regimes being studied. Specifically, the superior classification rate profiles confirm that, using Felix, we can easily extract filaments that are spatially more proximal to the cluster and filament particles in the Voronoi Kinematic datasets. This issue is not easily addressed by the significance threshold of DisPerSE. In contrast, Felix provides an intuitive density based handle to extract the desired features. Also, the possibility of having FP_d value larger than 1 in extreme situations is indicative of over-detection of filaments. This is potentially cumbersome for the analysis of genuine cosmological simulations and observational surveys. In more complex realistic circumstances, cosmic structure involves features over a wide range of densities and scales and structural morphologies that are not as well separated as in the simpler Voronoi models.

Comparison B: Felix and SpineWeb: For the comparison of the Felix and SpineWeb [70], we use a Voronoi Evolution model realization that is comparable to the advanced state of dataset C. We use the test result reported in [70] with respect to the model that has a similar percentage of particles in the four morphological features. In these model realizations, the clusters, filaments and walls have a Gaussian density profile with a scale of $R_g = 1h^{-1}\text{Mpc}$. The spine has an effective width $d = 2R_g$ as the identified structures are thickened by 1 voxel, with a size of $R_g = 1h^{-1}\text{Mpc}$.

For this configuration, Aragon-Calvo *et al.* [70] report true and false detection rates of $TP_d = 0.87$ and $FP_d = 0.10$ (see Table 2). Felix attains the same recovery rate of $TP_d = 0.87$ at a smaller distance $d = 3h^{-1}\text{Mpc} = 1.5R_g$. For the same configuration, the failure rate parameter, $FP_d = 0.05$, is comparable to that reported for SpineWeb. By comparison, at $d = 2R_g$, for Felix the recovery rates are $TP_d = 0.93$ and $FP_d = 0.07$.

In summary, these results appear to suggest that Felix performs as well as SpineWeb.

Comparison C: Felix and Nexus/MMF: Since Nexus/MMF concerns a sophisticated formalism based on a scale space analysis, the parameters of detection do not have a direct correspondence with topology based techniques like DisPerSE, SpineWeb, and Felix.

We use dataset B for a comparison with Nexus/MMF. This dataset is similar to the least evolved dataset used by [31] in an evaluation of the MMF, the original density field based Nexus/MMF implementation. For similar values of the detection rate TP_d , both Felix and MMF have identical failure rates FP_d . This indicates that both procedures have a comparable detection behavior.

5.2 Filament Exploration

In this section we discuss the application of Felix to explore different classes of filaments from cosmological simulations. The ability to filter filaments on a combination of morphological and density properties is helpful in situations where we wish to focus on, for example, the properties of galaxies residing in filaments in low-density void regions or in the high-density outskirts of clusters. This ability of Felix to identify a specified population of intravoids filaments or cluster inflow channels provides us with a microscopic instrument that allows a detailed and systematic exploration of the fine structure in the hierarchy of cosmic structure. Felix is able to zoom in on such regions and delineate their detailed infrastructure.

As the criteria for the identification of filaments and other web-like features still differ substantially between the various available techniques, the visual interaction aspect of Felix is a major practical asset in obtaining a proper user-defined selection of filaments. To this end, we may also point out that available automatic detection techniques may produce significant spurious results, which may substantially influence the results of targeted studies as the one illustrated here. A telling example of this has been discussed in the previous section.

Figures 9a and 9c present volume renderings of a 3D region of the Cosmogrid dataset ranging from z-coordinates 69 to 105. The bounding box of the dataset is $128 \times 128 \times 128$. This region is selected as it contains a large void like region surrounded by a large number of high density regions. The transfer function opacities have been adjusted to highlight the filament like structures in cluster-like and void-like regions respectively. The hierarchical Morse-Smale complex computation and the filament selection are executed on the entire dataset. The resulting filaments and volume rendering are clipped to the above mentioned region of interest. The filament selection parameters of algorithm SELECT2SADDLES are adjusted interactively, via a visualization step (see accompanying video). Each selection query takes approximately 1 second to process. The subsequent extraction of filament geometry depends on the number of selected 2-saddles. This takes approximately 4 seconds. Thus, the query framework may be used to interactively change parameters and visually correlate the set of extracted features with the underlying density distribution.

The filament selections obtained following the application of the interactive procedure are illustrated Figures 9b and

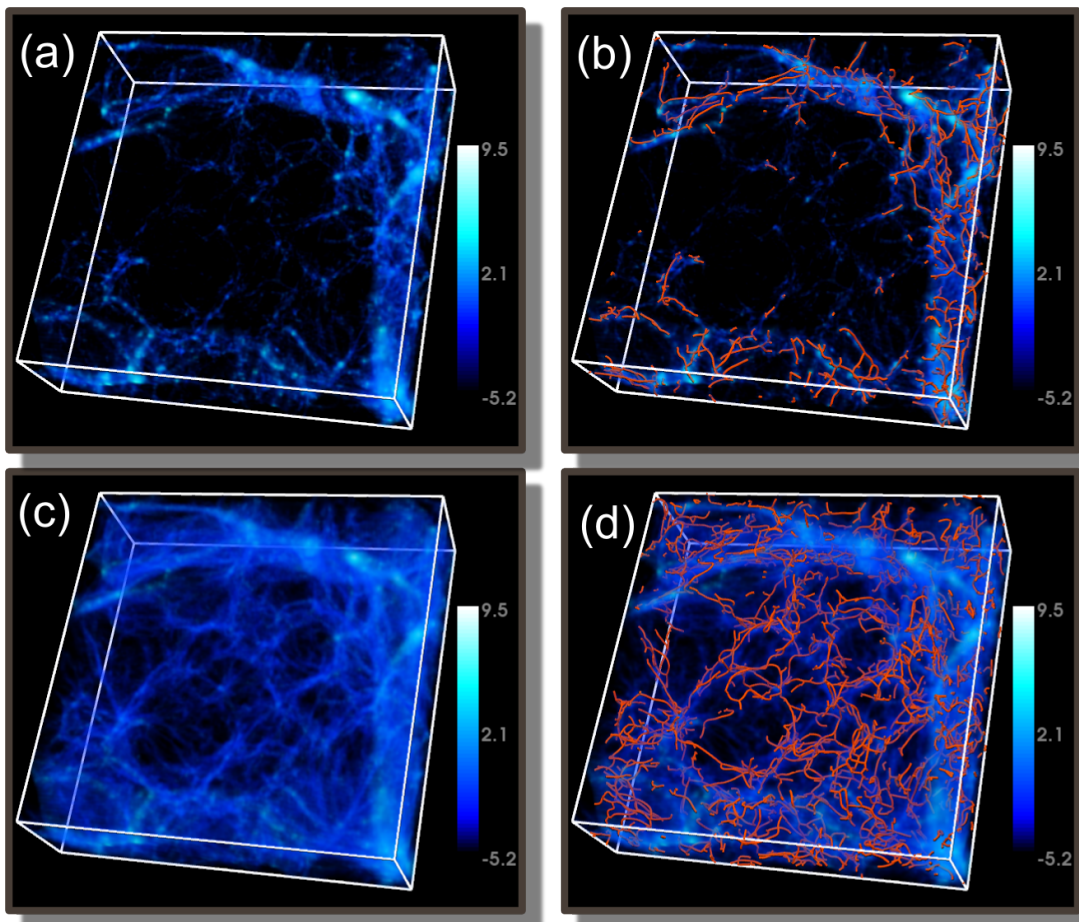


Fig. 9: Exploring filaments in high-density cluster-like environments and low-density void-like regions in the Cosmogrid dataset. The selected region, comprising of a 3D region of z -slices from 69 through 105, contains filamentary structures in both types of environments. Volume rendering of the density field with opacity adjusted to highlight filaments in (a) high density cluster-like regions, and (c) low density void-like regions. (b) Filaments within high density cluster-like regions extracted with parameters $[S_b, S_e] = [10^0, 10^{9.6}]$ and $[M_b, M_e] = [10^{2.3}, 10^{9.6}]$. (d) Filaments within low density void-like regions extracted with parameters $[S_b, S_e] = [10^{-2.5}, 10^{0.5}]$ and $[M_b, M_e] = [10^{0.5}, 10^{3.5}]$. For both sets of filaments, the value of T_s is set to 0.05.

9d. Each shows filaments in a different environmental density regime. Figure 9b shows filaments that exist near and within the high density cluster like regions. These are the filaments that form the spine of the cosmic web. Figure 9d shows filaments within void-like regions. The combination of density criteria and interactive visualization enables us to zoom in on this system of intra-void filaments. They are the faint residuals of the smaller-scale filaments that constituted the spine of the cosmic web at earlier cosmic epochs, and as such represent a direct manifestation of the hierarchical buildup of cosmic structure. At the current epoch, the intra-void filaments appear to define a different pattern than the prominent filamentary bridges between clusters of galaxies. As a result of the tidal influence of surrounding large-scale mass concentrations they are conspicuously aligned along a direction correlated with the main axis of the embedding void.

6 CONCLUSION

We have presented a Topology based Framework, named Felix, to probe filament structure in the large scale universe. The framework is particularly designed to probe filamentary

structures in different density regimes, and optimally preserve structural detail in regimes of interest. While other cosmic structure analysis tools do not include a facility to select web-like features according to tailor-made aspects and characteristics, this is precisely the mission of the Felix procedure. We directed Felix towards a case study of the the filamentary infrastructure and architecture of cosmic voids and demonstrated that it successfully extracts the network of tenuous filaments pervading their interior [72], [73], [74], [75].

In an accompanying study, we plan to exploit the Felix facility to study the physical characteristics of the extracted samples of intra-void filaments. This also involves their halo and subhalo population, their gas content, and the relation of these with the embedding voids and surrounding large-scale mass distribution. This will be of key importance towards understanding the formation and evolution of void galaxies [82], [83], [84] and specifically that of the issue of the missing dwarf galaxies [69]. In addition, following the recognition that void architecture represents a potentially sensitive probe of dark energy and dark matter and a keen test of modified gravity theories [9], [10], [11], [12], [13], [14], [15], [16], the filament samples extracted by Felix will be subjected to a systematic

study of their dependence on cosmological parameters.

As an immediate extension, we plan to use Felix with other scalar fields such as the tidal force field. Another possible direction is the visualization and analysis of the hierarchy of voids, walls, and filaments in cosmological datasets. Interactive visual exploration of these intricate structural networks remains a challenging and largely unexplored problem of major significance.

ACKNOWLEDGMENT

V.N. and N.S. acknowledge support by the Department of Science and Technology, India (grant no. SR/S3/EECE/0086/2012). R.v.d.W., P.P., S.R., and P.B. acknowledge support by the John Templeton Foundation (grant no. FP5136-O). The authors also thank the anonymous reviewers for their valuable review comments.

REFERENCES

- [1] J. R. Bond, L. Kofman, and D. Pogosyan, "How filaments of galaxies are woven into the cosmic web," *Nature*, vol. 380, pp. 603–606, 1996.
- [2] R. van de Weygaert and J. R. Bond, "Clusters and the Theory of the Cosmic Web," in *A Pan-Chromatic View of Clusters of Galaxies and the Large-Scale Structure*, ser. Lecture Notes in Physics, Berlin Springer Verlag, M. Plionis, O. López-Cruz, and D. Hughes, Eds., vol. 740, 2008, p. 335.
- [3] A. H. Guth, "Inflationary universe: A possible solution to the horizon and flatness problems," *Physical Review D*, vol. 23, pp. 347–356, 1981.
- [4] A. D. Linde, "Coleman-Weinberg theory and the new inflationary universe scenario," *Physics Letters B*, vol. 114, pp. 431–435, 1982.
- [5] —, "Scalar field fluctuations in the expanding universe and the new inflationary universe scenario," *Physics Letters B*, vol. 116, pp. 335–339, 1982.
- [6] D. N. Spergel *et al.*, "Three-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Implications for Cosmology," *Astrophysical Journal Supplement Series*, 2007.
- [7] E. Komatsu *et al.*, "Seven-year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Interpretation," *Astrophysical Journal Supplement Series*, 2011.
- [8] P. Collaboration *et al.*, "Planck 2013 results. XVI. Cosmological parameters," *arXiv preprint arXiv:1303.5076*, 2013.
- [9] B. S. Ryden and A. L. Melott, "Voids in Real Space and in Redshift Space," *Astrophys. J.*, vol. 470, p. 160, 1996.
- [10] J. D. Schmidt, B. S. Ryden, and A. L. Melott, "The Size and Shape of Voids in Three-dimensional Galaxy Surveys," *Astrophys. J.*, vol. 546, pp. 609–619, 2001.
- [11] D. Park and J. Lee, "Void Ellipticity Distribution as a Probe of Cosmology," *Physical Review Letters*, vol. 98, no. 8, 2007.
- [12] G. Lavaux and B. D. Wandelt, "Precision cosmology with voids: definition, methods, dynamics," *MNRAS*, vol. 403, pp. 1392–1408, 2010.
- [13] E. G. P. Bos, R. van de Weygaert, K. Dolag, and V. Pettorino, "The darkness that shaped the void: dark energy and cosmic voids," *MNRAS*, vol. 426, pp. 440–461, 2012.
- [14] G. Lavaux and B. D. Wandelt, "Precision Cosmography with Stacked Voids," *Astrophys. J. Lett.*, vol. 754, p. 109, 2012.
- [15] J. Clampitt, Y.-C. Cai, and B. Li, "Voids in modified gravity: excursion set predictions," *MNRAS*, vol. 431, pp. 749–766, 2013.
- [16] P. M. Sutter, A. Pisani, B. D. Wandelt, and D. H. Weinberg, "A measurement of the alcock-paczyński effect using cosmic voids in the sdss," *MNRAS*, vol. 443, pp. 2983–2990, 2014.
- [17] J. Tyson, "Gravitational Lensing by Clusters of Galaxies," in *Encyclopedia of Astronomy and Astrophysics*, P. Murdin, Ed., 2000.
- [18] J.-P. Kneib and P. Natarajan, "Cluster lenses," *Astronomy & Astrophysics*, vol. 19, p. 47, 2011.
- [19] J. P. Dietrich, N. Werner *et al.*, "A filament of dark matter between two clusters of galaxies," *Nature*, vol. 487, pp. 202–204, 2012.
- [20] F. Hoyle, "The Origin of the Rotations of the Galaxies," in *Problems of Cosmical Aerodynamics*, J. M. Burgers and H. C. van de Hulst, Eds., 1951, p. 195.
- [21] P. J. E. Peebles, "Origin of the Angular Momentum of Galaxies," *Astrophys. J.*, vol. 155, p. 393, 1969.
- [22] S. D. M. White, "Angular momentum growth in protogalaxies," *Astrophys. J.*, vol. 286, pp. 38–41, Nov. 1984.
- [23] J. Lee and U.-L. Pen, "Cosmic Shear from Galaxy Spins," *Astrophys. J. Lett.*, vol. 532, pp. L5–L8, 2000.
- [24] B. J. T. Jones, R. van de Weygaert, and M. A. Aragón-Calvo, "Fossil evidence for spin alignment of Sloan Digital Sky Survey galaxies in filaments," *MNRAS*, vol. 408, pp. 897–918, 2010.
- [25] E. Tempel, R. S. Stoica, and E. Saar, "Evidence for spin alignment of spiral and elliptical/S0 galaxies in filaments," *MNRAS*, vol. 428, pp. 1827–1836, 2013.
- [26] P. Peebles, *The Large Scale Structure of the Universe*. Princeton Univ. Press, Princeton, New Jersey, 1980.
- [27] S. F. Shandarin and I. B. Zeldovich, "Topology of the large-scale structure of the universe," *Comments on Astrophysics*, vol. 10, pp. 33–45, 1983.
- [28] J. D. Barrow, S. P. Bhavsar, and D. H. Sonoda, "Minimal spanning trees, filaments and galaxy clustering," *MNRAS*, vol. 216, 1985.
- [29] M. Alpaslan, A. S. G. Robotham *et al.*, "Galaxy and Mass Assembly (GAMA): fine filaments of galaxies detected within voids," *MNRAS*, vol. 440, pp. L106–L110, 2014.
- [30] M. A. Aragón-Calvo, B. J. T. Jones, R. van de Weygaert, and J. M. van der Hulst, "Spin Alignment of Dark Matter Halos in Filaments and Walls," *Astrophys. J. Lett.*, vol. 655, pp. L5–L8, 2007.
- [31] —, "The multiscale morphology filter: identifying and extracting spatial patterns in the galaxy distribution," *Astron. Astrophys.*, vol. 474, pp. 315–338, 2007.
- [32] N. A. Bond, M. A. Strauss, and R. Cen, "Crawling the cosmic network: identifying and quantifying filamentary structure," *MNRAS*, vol. 409, pp. 156–168, 2010.
- [33] O. Hahn, C. Porciani *et al.*, "Properties of dark matter haloes in clusters, filaments, sheets and voids," *MNRAS*, vol. 375, pp. 489–499, 2007.
- [34] J. E. Forero-Romero, Y. Hoffman, S. Gottlöber, A. Klypin, and G. Yepes, "A dynamical classification of the cosmic web," *MNRAS*, vol. 396, pp. 1815–1824, 2009.
- [35] N. I. Libeskind, Y. Hoffman, A. Knebe, M. Steinmetz, S. Gottlöber, O. Metuki, and G. Yepes, "The cosmic web and the orientation of angular momenta," *MNRAS*, vol. 421, pp. L137–L141, 2012.
- [36] M. Cautun, R. van de Weygaert, B. J. T. Jones, and C. S. Frenk, "Evolution of the cosmic web," *ArXiv e-prints*, 2014.
- [37] S. F. Shandarin, "The multi-stream flows and the dynamics of the cosmic web," *JCAP*, vol. 5, pp. 15–, 2011.
- [38] T. Abel, O. Hahn, and R. Kaehler, "Tracing the dark matter sheet in phase space," *MNRAS*, vol. 427, pp. 61–76, 2012.
- [39] M. C. Neyrinck, "Origami constraints on the initial-conditions arrangement of dark-matter caustics and streams," *MNRAS*, vol. 427, pp. 494–501, 2012.
- [40] S. Shandarin, S. Habib, and K. Heitmann, "Cosmic web, multistream flows, and tessellations," *Phys. Rev. D*, vol. 85, 2012.
- [41] R. S. Stoica, V. J. Martínez, and E. Saar, "A three-dimensional object point process for detection of cosmic filaments," *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 56 (4), 459–477, vol. 56, 2007.
- [42] R. Stoica, P. Gregori, and J. Mateu, "Simulated annealing and object point processes: Tools for analysis of spatial patterns," *Stochastic Processes and their Applications*, 2005.
- [43] E. Tempel, R. S. Stoica, E. Saar, V. J. Martínez, L. J. Liivamägi, and G. Castellan, "Detecting filamentary pattern in the cosmic web: a catalogue of filaments for the SDSS," *preprint arXiv:1308.2533*, 2013.
- [44] C. Genovese, M. Perone-Pacífico, I. Verdini, and L. Wasserman, "Minimax Manifold Estimation," *ArXiv e-prints*, 2010.
- [45] H. Blum, "A Transformation for Extracting New Descriptors of Shape," *Models for the Perception of Speech and Visual Form*, 1967.
- [46] J. Milnor, "Morse theory," *Journal of Mathematical Physics*, 1963.
- [47] H. Edelsbrunner, J. Harer, and A. Zomorodian, "Hierarchical Morse Complexes for Piecewise Linear 2-manifolds," in *Proc. of Symp. on Comp. Geom.*, 2001, pp. 70–79.
- [48] E. Platen, R. van de Weygaert, and B. J. T. Jones, "A cosmic watershed: the VWF void detection technique," *MNRAS*, vol. 380, pp. 551–570, 2007.
- [49] M. A. Aragón-Calvo, E. Platen *et al.*, "The Spine of the Cosmic Web," *Astrophys. J.*, vol. 723, pp. 364–382, 2010.
- [50] H. Edelsbrunner, D. Letscher, and A. Zomorodian, "Topological persistence and simplification," *Disc. & Comp. Geom.*, vol. 28, no. 4, pp. 511–533, 2002.
- [51] T. Sousbie, "The persistent cosmic web and its filamentary structure. I. Theory and implementation," *MNRAS*, vol. 414, pp. 350–383, 2011.

- [52] T. Sousbie, C. Pichon, and H. Kawahara, "The persistent cosmic web and its filamentary structure II. Illustrations," *MNRAS*, vol. 414, no. 1, pp. 384–403, 2011.
- [53] D. Günther, H.-P. Seidel, and T. Weinkauff, "Extraction of dominant extremal structures in volumetric data using separatrix persistence," *Comp. Graph. Forum*, vol. 31, no. 8, pp. 2554–2566, 2012.
- [54] J. Reininghaus, N. Kotava *et al.*, "A scale space based persistence measure for critical points in 2d scalar fields," *IEEE Trans. Vis. & Comp. Graphics*, vol. 17, no. 12, pp. 2045–2052, 2011.
- [55] T. Weinkauff and D. Günther, "Separatrix persistence: extraction of salient edges on surfaces using topological methods," in *Proc. of Symp. on Geom. Proc.*, 2009, pp. 1519–1528.
- [56] D. Günther, J. Reininghaus, H. P. Seidel, and T. Weinkauff, "Notes on the simplification of the Morse-Smale complex," in *Topological Methods in Data Analysis and Visualization III*, ser. Mathematics and Visualization, P.-T. Bremer, I. Hotz, V. Pascucci, and R. Peikert, Eds. Springer International Publishing, 2014.
- [57] M. Cautun, R. van de Weygaert, and B. J. T. Jones, "NEXUS: tracing the cosmic web connection," *MNRAS*, vol. 429, pp. 1286–1308, 2013.
- [58] W. E. Schaap and R. van de Weygaert, "Continuous fields and discrete samples: reconstruction through delaunay tessellations," *arXiv*, vol. 363, pp. L29–L32, 2000.
- [59] R. van de Weygaert and W. Schaap, "The Cosmic Web: Geometric Analysis," in *Data Analysis in Cosmology*, ser. Lecture Notes in Physics, Berlin Springer Verlag, V. J. Martínez, E. Saar, E. Martínez-González, and M.-J. Pons-Bordería, Eds., vol. 665, 2009, pp. 291–413.
- [60] "Computational Geometry Algorithms Library," <http://www.cgal.org>.
- [61] A. Gyulassy, V. Natarajan *et al.*, "Efficient computation of Morse-Smale complexes for three-dimensional scalar functions," *IEEE Trans. Vis. & Comp. Graphics*, vol. 13, no. 6, pp. 1440–1447, 2007.
- [62] H. Edelsbrunner, J. Harer, V. Natarajan, and V. Pascucci, "Morse-Smale complexes for piecewise linear 3-manifolds," in *Proc. of Symp. on Comp. Geom.*, 2003, pp. 361–370.
- [63] R. Forman, "A user's guide to discrete Morse theory," *Séminaire Lotharingien de Combinatoire*, vol. 48, 2002.
- [64] A. Gyulassy, P.-T. Bremer *et al.*, "A practical approach to Morse-Smale complex computation: Scalability and generality," *IEEE Trans. Vis. & Comp. Graphics*, vol. 14, no. 6, pp. 1619–1626, 2008.
- [65] V. Robins, P. Wood, and A. Sheppard, "Theory and algorithms for constructing discrete Morse complexes from grayscale digital images," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 33, no. 8, pp. 1646–1658, 2011.
- [66] N. Shivashankar, M. Senthilnathan, and V. Natarajan, "Parallel computation of 2D Morse-Smale complexes," *IEEE Trans. Vis. & Comp. Graphics*, vol. 18, no. 10, pp. 1757–1770, 2012.
- [67] N. Shivashankar and V. Natarajan, "Parallel computation of 3D Morse-Smale complexes," *Computer Graphics Forum*, vol. 31, no. 3pt1, pp. 965–974, 2012.
- [68] S. Codis, C. Pichon *et al.*, "Connecting the cosmic web to the spin of dark haloes: implications for galaxy formation," *MNRAS*, vol. 427, pp. 3320–3336, 2012.
- [69] P. J. E. Peebles, "The Void Phenomenon," *Astrophys. J.*, vol. 557, pp. 495–504, 2001.
- [70] M. A. Aragón-Calvo, R. van de Weygaert, and B. J. T. Jones, "Multiscale phenomenology of the cosmic web," *MNRAS*, vol. 408, pp. 2163–2187, 2010.
- [71] A. Gyulassy, N. Kotava *et al.*, "Direct feature visualization using Morse-Smale complexes," *IEEE Trans. Vis. & Comp. Graphics*, vol. 18, no. 9, pp. 1549–1562, 2012.
- [72] R. van de Weygaert and E. van Kampen, "Voids in Gravitational Instability Scenarios - Part One - Global Density and Velocity Fields in an Einstein - De-Sitter Universe," *MNRAS*, vol. 263, 1993.
- [73] R. K. Sheth and R. van de Weygaert, "A hierarchy of voids: much ado about nothing," *MNRAS*, vol. 350, pp. 517–538, 2004.
- [74] R. v. d. Weygaert, "Voronoi tessellations and the cosmic web: Spatial patterns and clustering across the universe," in *ISVD '07: Proc. of Symp. on Voronoi Diagrams in Science and Engineering*. Washington, DC, USA: IEEE Computer Society, 2007, pp. 230–239.
- [75] M. A. Aragón-Calvo and A. S. Szalay, "The hierarchical structure and dynamics of voids," *MNRAS*, vol. 428, pp. 3409–3424, 2013.
- [76] R. van de Weygaert and V. Icke, "Fragmenting the universe. II - Voronoi vertices as Abell clusters," *Astron. Astrophys.*, vol. 213, pp. 1–9, 1989.
- [77] R. van de Weygaert, *Voids and the geometry of large scale structure*. PhD thesis, University of Leiden, 1991.
- [78] E. Komatsu *et al.*, "Five-Year Wilkinson Microwave Anisotropy Probe Observations: Cosmological Interpretation," *Astrophysical Journal Supplement Series*, vol. 180, no. 2, 2009.
- [79] T. Ishiyama, S. Rieder *et al.*, "The Cosmogrid Simulation: Statistical Properties of Small Dark Matter Halos," *Astrophys. J.*, vol. 767, 2013.
- [80] Y. B. Zel'dovich, "Gravitational instability: An approximate theory for large density perturbations," *Astronomy & astrophysics*, vol. 5, pp. 84–89, 1970.
- [81] S. Rieder, R. van de Weygaert, M. Cautun, B. Beygu, and S. Portegies Zwart, "Assembly of filamentary void galaxy configurations," *MNRAS*, vol. 435, pp. 222–241, 2013.
- [82] R. van de Weygaert, K. Kreckel *et al.*, "The Void Galaxy Survey," in *Environment and the Formation of Galaxies: 30 Years Later*, I. Ferreras and A. Pasquali, Eds., 2011, p. 17.
- [83] K. Kreckel, E. Platen *et al.*, "Only the Lonely: H I Imaging of Void Galaxies," *Astron.J.*, vol. 141, p. 4, 2011.
- [84] —, "The Void Galaxy Survey: Optical Properties and H I Morphology and Kinematics," *Astron.J.*, vol. 144, p. 16, 2012.



Nithin Shivashankar is a doctoral candidate at the Department of Computer Science and Automation, Indian Institute of Science, Bangalore. He holds a bachelors degree in computer science from the Manipal Institute of Technology. His research interests include scientific visualization and computational topology.



Pratyush Pranav is working for his PhD degree at the Kapteyn Astronomical Institute, Groningen, the Netherlands. He received a bachelors degree from St. Stephens college, Delhi, India, and holds a Masters degree from the Department of Physics at the Indian Institute of Science, Bangalore, India. His research interests include computational topology, cosmology, large scale structures in the Universe and Scientific visualization.



Vijay Natarajan is an associate professor at the Indian Institute of Science, Bangalore. He received the Ph.D. degree in computer science from Duke University in 2004 and holds the B.E. degree in computer science and M.Sc. degree in mathematics from Birla Institute of Technology and Science, Pilani, India. His research interests include scientific visualization, computational geometry, computational topology, and meshing.



Rien van de Weygaert is professor of cosmic structure formation at the Kapteyn Astronomical Institute of the University of Groningen. He received his Ph.D. in astrophysics at Leiden University in 1991 on the role of cosmic voids in the large scale mass distribution in the Universe. His research interests are cosmology, the large scale structure of the Universe, cosmic structure formation, galaxy formation and evolution, computational geometry and topology, and the history of astronomy.



Patrick Bos is a PhD student at the Kapteyn Astronomical Institute of the University of Groningen in The Netherlands. He studies alternative models of dark energy through simulations and tries to find new probes to disentangle the nature of dark energy. He is also involved in building high-dimensional statistical models and techniques for inferring the initial conditions of the universe.



Steven Rieder is a postdoctoral researcher at RIKEN Advanced Institute for Computational Science, Japan. He received his Ph.D. degree in astrophysics at Leiden University in 2013, on the topic of multi-scale N-body simulations. His research interests include computational astrophysics, the large scale structure of the Universe, galaxy formation and evolution and the formation and disruption of star clusters.