Alpha, Betti and the Megaparsec Universe:
on the Topology of the Cosmic Web

Rien van de Weygaert*, Gert Vegter†, Herbert Edelsbrunner‡, Bernard J.T. Jones*, Pratyush Pranav*, Changbom Park§,
Wojciech A. Hellwing¶, Bob Eldering†, Nico Kruithof†, E.G.P. (Patrick) Bos*, Johan Hidding*,
Job Feldbrugge*, Eline ten Have‖, Matti van Engelen†, Manuel Caroli**, Monique Teillaud**

*Kapteyn Astronomical Institute, University of Groningen, P.O. Box 800, 9700 AV Groningen, the Netherlands
†Johann Bernoulli Institute for Mathematics and Computer Science, University of Groningen,
P.O. Box 407, 9700 AK Groningen, the Netherlands
‡IST Austria, Am Campus 1, 3400 Klosterneuburg, Austria
§School of Physics, Korea Institute for Advanced Study, Seoul 130-722, Korea
¶Interdisciplinary Centre for Mathematical and Computational Modeling, University of Warsaw,
ul. Pawinskiego 5a, 02-106 Warsaw, Poland
‖ Stratingh Institute for Chemistry, University of Groningen, Nijenborgh 4, 9747 AG Groningen, the Netherlands
**Géométrica, INRIA Sophia Antipolis-Méditerranée, route des Lucioles, BP 93, 06902 Sophia Antipolis Cedex, France

Abstract—We study the topology of the Megaparsec Cosmic Web in terms of the scale-dependent Betti numbers, which formalize the topological information content of the cosmic mass distribution. While the Betti numbers do not fully quantify topology, they extend the information beyond conventional cosmological studies of topology in terms of genus and Euler characteristic. The richer information content of Betti numbers goes along the availability of fast algorithms to compute them.

For continuous density fields, we determine the scale-dependence of Betti numbers by invoking the cosmologically familiar filtration of sublevel or superlevel sets defined by density thresholds. For the discrete galaxy distribution, however, the analysis is based on the alpha shapes of the particles. These simplicial complexes constitute an ordered sequence of nested subsets of the Delaunay tessellation, a filtration defined by the scale parameter, α. As they are homotopy equivalent to the sublevel sets of the distance field, they are an excellent tool for assessing the topological structure of a discrete point distribution. In order to develop an intuitive understanding for the behavior of Betti numbers as a function of α, and their relation to the morphological patterns in the Cosmic Web, we first study them within the context of simple heuristic Voronoi clustering models. These can be tuned to consist of specific morphological elements of the Cosmic Web, i.e. clusters, filaments, or sheets. To elucidate the relative prominence of the various Betti numbers in different stages of morphological evolution, we introduce the concept of alpha tracks.

Subsequently, we address the topology of structures emerging in the standard LCDM scenario and in cosmological scenarios with alternative dark energy content. The evolution of the Betti numbers is shown to reflect the hierarchical evolution of the Cosmic Web. We also demonstrate that the scale-dependence of the Betti numbers yields a promising measure of cosmological parameters, with a potential to help in determining the nature of dark energy and to probe primordial non-Gaussianities. We also discuss the expected Betti numbers as a function of the density threshold for superlevel sets of a Gaussian random field.

Finally, we introduce the concept of persistent homology. It measures scale levels of the mass distribution and allows us to separate small from large scale features. Within the context of the hierarchical cosmic structure formation, persistence provides a natural formalism for a multiscale topology study of the Cosmic Web.

Keywords-Cosmology: theory - large-scale structure of Universe - Methods: data analysis - techniques: Image Processing - Computational Geometry: tessellations - Computational Topology: homology.

I. INTRODUCTION: THE COSMIC WEB

The large scale distribution of matter revealed by galaxy surveys features a complex network of interconnected filamentary galaxy associations. This network, which has become known as the *Cosmic Web* [1], contains structures from a few megaparsecs[1] up to tens and even hundreds of megaparsecs of size. Galaxies and mass exist in a wispy web-like spatial arrangement consisting of dense compact clusters, elongated filaments, and sheet-like walls, amidst large near-empty voids, with similar patterns existing at earlier epochs, albeit over smaller scales; see Figure 1 for an illustration of a simulated cosmic mass distribution in the

---

[1]The main measure of length in astronomy is the parsec. Technically, 1pc is the distance at which we would see the distance Earth-Sun at an angle of 1 arcsec. It is equal to 3.262 lightyears $= 3.086 \times 10^{13}$km. Cosmological distances are substantially larger, so that a megaparsec, with $1\mathrm{Mpc} = 10^6$ pc, is the common unit of distance.

standard LCDM cosmology[2] in a box of size $80h^{-1}\mathrm{Mpc}$ [3]. The hierarchical nature of this mass distribution, marked by substructure over a wide range of scales and densities, has been clearly demonstrated [2]. Its appearance has been most dramatically illustrated by the recently produced maps of the nearby cosmos, the 2dFGRS, the SDSS and the 2MASS redshift surveys [3], [4], [5] [4].

The vast Megaparsec Cosmic Web is one of the most striking examples of complex geometric patterns found in nature, and certainly the largest in terms of sheer size. Computer simulations suggest that the observed cellular patterns are a prominent and natural aspect of cosmic structure formation through gravitational instability [6], the standard paradigm for the emergence of structure in our Universe [7], [8]. According to the *gravitational instability scenario*, cosmic structure grows from primordial density and velocity perturbations. These tiny primordial perturbations define a Gaussian density field and are fully characterized by the power spectrum; see Section VIII.

A. Web Analysis

Over the past decades, we have seen many measures for characterizing different aspects of the large scale cosmic structure: correlation functions (describing the n-point distribution), Minkowski functionals and genus (characterizing the local and global curvature of isodensity surfaces), multifractals (summarizing the statistical moments on various scales), and so on. Despite the multitude of descriptions, it has remained a major challenge to characterize the structure, geometry and topology of the Cosmic Web. Many attempts to describe, let alone identify, the features and components of the Cosmic Web have been of a rather heuristic nature. The overwhelming complexity of both the individual structures

as well as their connectivity, the lack of structural symmetries, its intrinsic multiscale nature and the wide range of densities in the cosmic matter distribution has prevented the use of simple and straightforward instruments.

In the observational reality, galaxies are the main tracers of the Cosmic Web, and it is mainly through the measurement of the redshift distribution of galaxies that we have been able to map its structure. Likewise, simulations of the evolving cosmic matter distribution are almost exclusively based upon N-body particle computer calculations, involving a discrete representation of the features we seek to study. Both the galaxy distribution as well as the particles in an N-body simulation are examples of *spatial point processes* in that they are *discretely sampled* and have an *irregular spatial distribution*.

For furthering our understanding of the Cosmic Web, and to investigate its structure and dynamics, it is of prime importance to have access to a set of proper and objective analysis tools. In this contribution, we address the topological and morphological analysis of the large scale galaxy distribution. In particular, we introduce a new measure that is particularly suited to differentiating web-like structures: the homology of the distribution as measured by Betti numbers and their persistence.

B. Genus and Minkowski Functionals

Gott and collaborators were the first to propose the use of the topology of the megaparsec structure of the Universe [9], [10], to examine whether or not the initial fluctuations had been Gaussian. They introduced the genus of isodensity surfaces of the cosmic mass distribution as a quantitative characterization of cosmic topology in terms of the connectivity of these surfaces. Being an intrinsic topology measure, the genus is relatively insensitive to systematic effects such as non-linear gravitational evolution, galaxy biasing, and redshift-space distortions [11]. This, in fact, is one of the most important advantages of using the genus in studying the patterns in the Megaparsec Universe. This allowed [9], along with a series of subsequent studies, to conclude that the measured genus of the observed large scale structure is consistent with the predictions of cosmological scenarios whose initially Gaussian density and velocity fluctuations correspond to that of the observed power spectrum [12], [13], [14], [15], [16], [17], [18]. Nonetheless, recent work has shown that on small scales, the observed genus is inconsistent with theoretical models [19]. On these scales the - not fully known - details of the galaxy formation process become important.

Additional quantitative information on the morphology of the large scale distribution of galaxies are the Minkowski functionals. Mecke et al. [20] introduced them within a cosmological context as a more extensive characterization of the morphology and geometry of spatial patterns, after which Schmalzing et al. [21], [22] and others developed
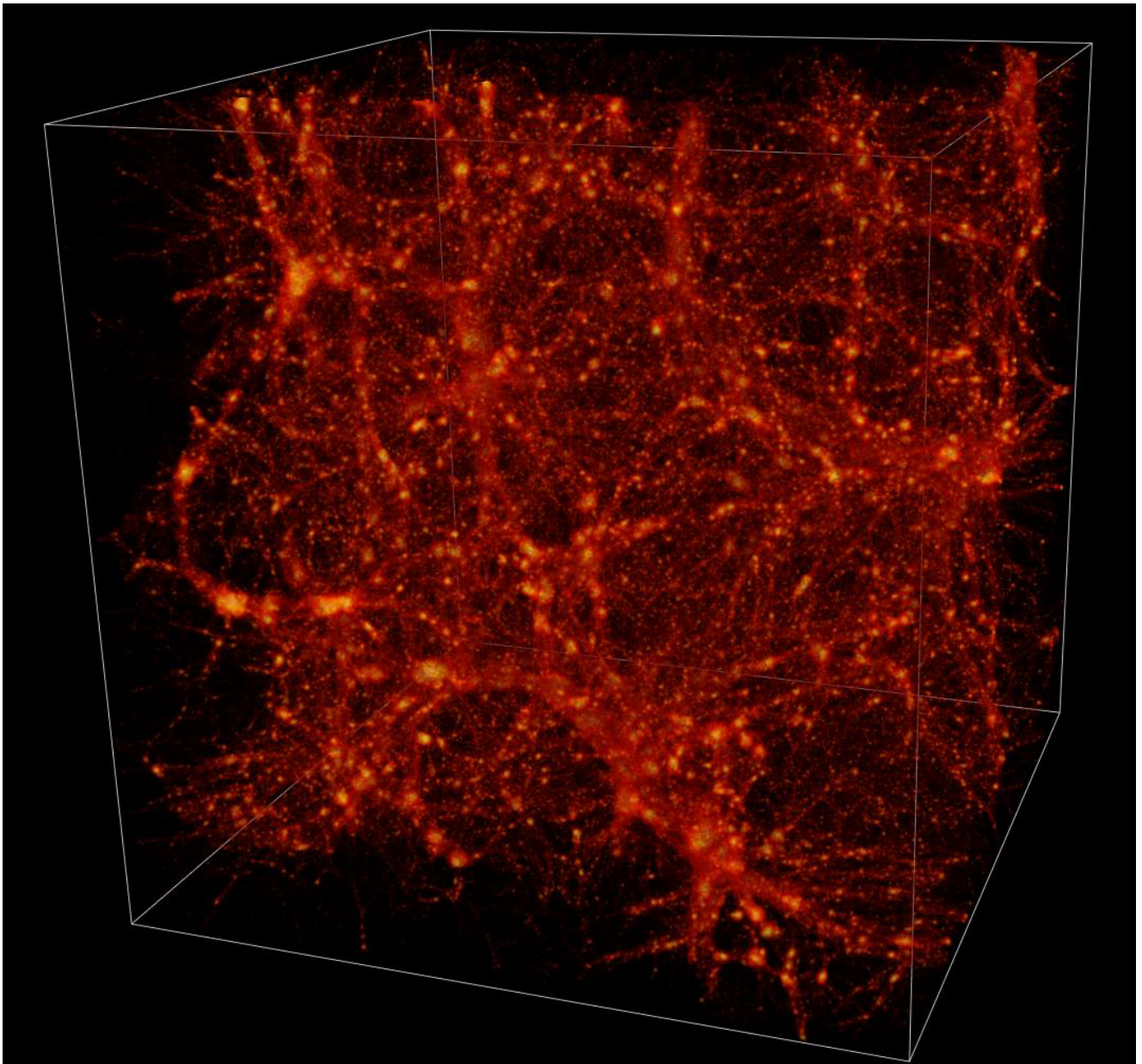
---

[2]Currently, the LCDM cosmological scenario is the standard - or, "concordance" - cosmological scenario. It seems to be in agreement with a truly impressive amount of observational evidence, although there are also some minor deficiencies. The "L" stands for Λ, the cosmological constant that dominates the dynamics of our Universe, causes its expansion to accelerate and represents in the order of about 73% of its energy content. "CDM" indicates the Cold Dark Matter content of the Universe. Invisible, it appears to form the major fraction of gravitating matter in the Universe, representing some 23% of the cosmic energy density as opposed to the mere 4.4% that we find in the normal baryonic matter we, and the planets and stars, consist of.

[3]The Universe expands according to Hubble's law, stating that the recession velocity of a galaxy is linearly proportional to its distance: $v = Hr$. The Hubble parameter, H, quantifies the expansion rate of the Universe, and is commonly expressed in units of km/s/Mpc. Its present value, the "Hubble constant", is estimated to be $H_0 \approx 71$ km/s/Mpc. Quite often, its value is expressed by a dimensionless number, h, that specifies the Hubble parameter in units of 100 km/s/Mpc.

[4]Because of the expansion of the Universe, any observed cosmic object will have its light shifted redward : its *redshift*, z. According to Hubble's law, the redshift z is directly proportional to the distance r of the object, for $z \ll 1$: $cz = Hr$, in which the constant H is the Hubble parameter. Because it is extremely cumbersome to measure distances r directly, cosmologists resort to the expansion of the Universe and use z as a distance measure. Because of the vast distances in the Universe, and the finite velocity of light, the redshift z of an object may also be seen as a measure of the time at which it emitted the observed radiation.

Figure 1. The Cosmic Web in an LCDM simulation. Shown is the dark matter distribution at the current cosmological epoch in a 384^3 particles N-body simulation of structure formation in a universe with a cosmological constant Λ, accounting for an energy density $\Omega_\Lambda = 0.73$, and cold dark matter, accounting for a total mass density of $\Omega_m = 0.23$. The box has a co-moving size of $80h^{-1}\mathrm{Mpc}$ (i.e. a box co-expanding with the universe and having a current size of $80h^{-1}\mathrm{Mpc}$). Clearly visible is the intriguing network of filaments and high-density cluster nodes surrounding low-density voids. The mass distribution has a distinct multiscale character, marked by clumps over a wide range of mass scales, reflecting the hierarchical evolution of the cosmic mass distribution. Figure courtesy of Bartosz Borucki.

them into a useful tool for studying cosmological datasets. The Minkowski functionals measure the geometry and shape of manifolds, providing a quantitative characterization of the level sets (the isodensity surfaces) of the matter distribution at a given level of smoothing of the original data. In a d-dimensional space, there are $d + 1$ Minkowski functionals, and the Euler characteristic of isodensity surfaces is one of them. In 3 dimensions, there are four Minkowski functionals: the volume enclosed by an isodensity surface,

the surface area, the integrated mean curvature, and the Euler characteristic (as in (2)). From these, it is possible to generate nonparameteric shape descriptors such as the "typical" thickness, breadth and length of the structures involved at that level of smoothing [23]. There are also algorithms for computing these quantities from discrete point sets [24], and the technique has found wide applications outside of astronomy and physics.

What the Minkowski functionals do not tell about is the

topology of the structure as described by the number of blobs, tunnels and voids that are present with changing scale. In that sense, the Betti numbers and their persistence will provide a substantial extension of available topological information. Also, the related use of alpha shapes will provide effective methods of evaluating the Minkowski functionals.

C. Homology Groups

Here we advance the topological characterization of the Cosmic Web to a more complete description, the homology of the distribution as measured by the scale-dependent Betti numbers of the sample. While a full quantitative characterization of the topology of the cosmic mass distribution is not feasible, its homology is an attractive compromise between detail and speed, providing a useful summary measurement of topology. The ranks of the homology groups, known as the *Betti numbers*, completely characterize the topology of orientable 2-manifolds, such as the isodensity surfaces used in earlier topological analyzes of the Cosmic Web. For a d-dimensional manifold, we have $d+1$ homology groups, H_p, and correspondingly $d+1$ Betti numbers, β_p, for $0 \le p \le d$. The Betti numbers may be considered as the number of p-dimensional holes. In fact, homology groups can be seen as a useful definition of holes in the cosmic mass distribution. They are fundamental quantities from which the Euler characteristic and genus are derived, including those of the isodensity 2-manifolds. In that sense, they are a powerful generalization of genus analysis.

While the Euler characteristic and the Betti numbers give information about the connectivity of a manifold, two of the other three Minkowski functionals are sensitive to local manifold deformations, and their topological information is limited. The Betti numbers represent a more succinct as well as more informative characterization of the topology.

Following this observation, we will argue and demonstrate that the scale-dependence of Betti numbers makes them particularly suited to analyze and differentiate web-like structures. The Betti numbers measured as a function of scale are a sensitive discriminator of cosmic structure and can effectively reveal differences arising in alternative cosmological models. Potentially, they may be exploited to infer information on important issues such as the nature of dark energy or even possible non-Gaussianities in the primordial density field, as will be corroborated by [25]; see Section VII.

D. Alpha Shapes

Most of the topological studies in cosmology depend on some sort of user-controlled smoothing and related threshold to specify surfaces of which the topology may be determined. In cosmological studies, this usually concerns isodensity surfaces and their superlevel and sublevel sets defined on a Gaussian filter scale.

A preferred alternative would be avoid filters and to invoke a density field reconstruction that adapts itself to the galaxy distribution. For our purpose, the optimal technique would be that of the Delaunay Tessellation Field Estimator (DTFE) formalism [26], [27], [28], which uses the local density and shape sensitivity of the Delaunay tessellation to recover a density field that retains the multiscale nature as well as the anisotropic nature of the sampled particle or galaxy distribution. This results in an adaptive and highly flexible representation of the underlying density field, which may then be used to assess its topological and singularity structure; see e.g. [18], [24], [29].

A closely related philosophy is to focus exclusively on the shape of the particle or galaxy distributions itself, and seek to analyze its topology without resorting to the isodensity surfaces and level sets of the corresponding density field. This is precisely where *alpha shapes* enter the stage. They are subsets of a Delaunay triangulation that describe the intuitive notion of the shape of a discrete point set. Introduced by Edelsbrunner and collaborators[30], they are one of the principal concepts from the field of Computational Topology [31], [32], [33].

E. Outline

This review contains a report on our project to study the homology and topology of the Cosmic Web. It is a major upgrade of an earlier report [34].

Sections II and III provide the mathematical background for this review, consisting of a short presentation of homology and Betti numbers, followed by a discussion of alpha shapes and the formalism to infer Betti numbers.

One of the intentions of our study is to understand the information content of scale-dependent Betti numbers with respect to the corresponding geometric patterns observed in the Cosmic Web, i.e. the relative prominence of clusters, filaments and walls. To this end, we first investigate a number of Voronoi clustering models. These models use the geometric structure of Voronoi tessellations as a skeleton of the cosmic mass distribution. In Section IV, we describe these models, and specify the class of Voronoi Element Models and Voronoi Kinematic Models. Section V contains an extensive description of the analysis of a set of Voronoi Element and Voronoi Kinematic Models.

Following the analysis of Voronoi clustering models, we turn to the homology analysis in a few cosmological situations of current interest. Section VI discusses the scale-dependent Betti numbers inferred for the current standard theory of cosmic structure formation, LCDM. The results are obtained from a computer simulation of structure formation in this cosmological scenario. The discussion in Section VI connects the resulting homological characteristics to the hierarchically evolving mass distribution. In an attempt to investigate and exploit the sensitivity of Betti numbers to key cosmological parameters, we apply our analysis to a

set of cosmological scenarios with a different content of dark energy. Section VII demonstrates that indeed homology may be a promising tool towards inferring information on the dark energy content that dominates the dynamics of our Universe. Equally compelling is the issue of the Betti numbers of superlevel or sublevel sets of Gaussian random fields. To high accuracy, the initial density field out of which all structure in our Universe arose has had a Gaussian character. As reference point for any further assessment of the homology of the evolving cosmic mass distribution, we therefore need to evaluate the homology of Gaussian random fields. This is the subject of Section VIII, which focusses on the expected values of the Betti numbers. In addition, it will be a starting point for any related study looking for possible primordial non-Gaussian deviations.

Having shown the potential of homology studies for cosmological purposes, we discuss the prospects of assessing the persistence of cosmological density fields in Section IX. Persistence measures scale levels of the mass distribution and allows us to systematically separate small from large scale features in the mass distribution. It provides us with a rich language to study intrinsically multiscale distributions as those resulting from the hierarchically evolving structure in the Universe. Finally, Section X addresses future prospects and relates this review to other work.

II. HOMOLOGY AND BETTI NUMBERS

Homology groups and Betti numbers are concepts from algebraic topology, designed to quantify and compare topological spaces. They characterize the topology of a space in terms of the relationship between the cycles and boundaries we find in the space[5]. For example, if the space is a d-dimensional manifold, $\mathbb{M}$, we have cycles and boundaries of dimension p from 0 to d. Correspondingly, $\mathbb{M}$ has one homology group $H_p(\mathbb{M})$ for each of $d + 1$ dimensions, $0 \leq p \leq d$. By taking into account that two cycles should be considered identical if they differ by a boundary, one ends up with a group $H_p(\mathbb{M})$ whose elements are the equivalence classes of p-cycles[6]. The *rank* of the homology group $H_p(\mathbb{M})$ is the p-th *Betti number*, $\beta_p = \beta_p(\mathbb{M})$.

---

[5]Assuming the space is given as a simplicial complex, a *p-cycle* is a p-chain with empty boundary, where a *p-chain*, γ, is a sum of p-simplices. The standard notation is $\gamma = \sum a_i \sigma_i$, where the σ_i are the p-simplices and the a_i are the *coefficients*. For example, a 1-cycle is a closed loop of edges, or a finite union of such loops, and a 2-cycle is a closed surface, or a finite union of such surfaces. Adding two p-cycles, we get another p-cycle, and similar for the p-boundaries. Hence, we have a group of p-cycles and a group of p-boundaries.

[6]Correctly defined, the *p-th homology group* is the p-th cycle group modulo the p-th boundary group. In algebraic terms, this is a quotient group. In intuitive terms, this says that two p-cycles are considered the same, or *homologous*, if together they bound a $(p+1)$-chain or, equivalently, if they differ by a p-boundary. Indeed, we do not want to distinguish between two 1-cycles of, say, the torus, if they both go around the hole, differing only in the geometric paths they take to do so.

In heuristic - and practical - terms, the Betti numbers count topological features and can be considered as the number of p-dimensional holes. When talking about a surface in 3-dimensional space, the zeroth Betti number, β_0, counts the components, the first Betti number, β_1, counts the tunnels, and the second Betti number, β_2, counts the enclosed voids. All other Betti numbers are zero. Examples of spaces with Betti numbers which are of interest to us are the sublevel and superlevel density set of the cosmic mass distribution, i.e. the regions whose density is smaller or greater than a specified threshold level.

A. Genus and Euler Characteristic

Numerous cosmological studies have considered the *genus* of the isodensity surfaces defined by the megaparsec galaxy distribution [9], [10], [35], which specifies the number of handles defining the surface. More formally, it is the maximum number of disjoint closed curves such that cutting along these curves does not increase the number of components. The genus has a simple relation to the Euler characteristic, χ, of the isodensity surface. Consider a 3-manifold subset $\mathbb{M}$ of the Universe and its boundary, $\partial\mathbb{M}$, which is a 2-manifold. With $\partial\mathbb{M}$ consisting of $c = \beta_0(\partial\mathbb{M})$ components, the Gauss-Bonnet Theorem states that the genus of the surface is given by

$$G = c - \frac{1}{2}\chi(\partial\mathbb{M}), \qquad (1)$$

where the Euler characteristic $\chi(\partial\mathbb{M})$ is the integrated Gaussian curvature of the surface

$$\chi(\partial\mathbb{M}) = \frac{1}{2\pi} \oint_x \frac{\mathrm{d}x}{R_1(x)R_2(x)}. \qquad (2)$$

Here $R_1(x)$ and $R_2(x)$ are the principal radii of curvature at the point x of the surface. The integral of the Gaussian curvature is invariant under continuous deformation of the surface: perhaps one of the most surprising results in differential geometry.

The Euler characteristic of the surface can also be expressed in terms of its Betti numbers, namely, $\chi(\partial\mathbb{M})$ is equal to $\beta_0(\partial\mathbb{M})$ minus $\beta_1(\partial\mathbb{M})$ plus $\beta_2(\partial\mathbb{M})$. This is implied by the Euler-Poincaré Formula, which we will discuss shortly, after introducing triangulation of spaces. Combining these two equations for the Euler characteristic, we get a fundamental relationship between differential geometry and algebraic topology. Returning to the 3-dimensional subset, $\mathbb{M}$, of the Universe, its Euler characteristic is

$$\chi(\mathbb{M}) = \beta_0(\mathbb{M}) - \beta_1(\mathbb{M}) + \beta_2(\mathbb{M}) - \beta_3(\mathbb{M}). \qquad (3)$$

Whenever $\mathbb{M}$ is a non-exhaustive subset of the connected Universe, its third Betti number vanishes, $\beta_3(\mathbb{M}) = 0$, and its boundary is necessarily non-empty, namely a 2-manifold without boundary. As mentioned above, the Euler characteristic of $\partial\mathbb{M}$ is the alternating sum of Betti numbers, where $\beta_0(\partial\mathbb{M})$ is the number of surface components, $\beta_1(\partial\mathbb{M})$ is

twice the genus, and $\beta_2(\partial\mathbb{M}) = \beta_0(\partial\mathbb{M})$. Assuming the Universe is connected like the 3-sphere, we can use Alexander duality and the Mayer-Vietoris sequence to establish a direct relation between the Betti numbers of the 3-manifold with boundary, $\mathbb{M}$, and those of the 2-manifold without boundary, $\partial\mathbb{M}$; see e.g. [36]:

$$\beta_0(\partial\mathbb{M}) = \beta_2(\partial\mathbb{M}) = \beta_0(\mathbb{M}) + \beta_2(\mathbb{M}), \quad (4)$$
$$\beta_1(\partial\mathbb{M}) = 2\beta_1(\mathbb{M}). \quad (5)$$

From this, we infer that the Euler characteristic of the boundary is directly proportional to the Euler characteristic of the 3-manifold:

$$\chi(\partial\mathbb{M}) = \beta_0(\partial\mathbb{M}) - \beta_1(\partial\mathbb{M}) + \beta_2(\partial\mathbb{M}) \quad (6)$$
$$= 2\chi(\mathbb{M}). \quad (7)$$

In the cosmologically interesting situation in which $\partial\mathbb{M}$ is the isodensity surface of either density superlevel or sublevel sets, we find a relation between the genus[7] of the surface and the Betti numbers of the enclosed manifold:

$$G = c - \frac{1}{2}\chi(\partial\mathbb{M}) \quad (8)$$
$$= c - \chi(\mathbb{M}) \quad (9)$$
$$= c - (\beta_0(\mathbb{M}) - \beta_1(\mathbb{M}) + \beta_2(\mathbb{M})). \quad (10)$$

In the analysis described in this paper, we will restrict ourselves to the three Betti numbers, β_0, β_1, and β_2, of the 3-manifolds with boundary defined by the cosmic mass distribution.

B. Triangulated Spaces

A practical simplification occurs when we represent a space by a *triangulation*, which is a simplicial complex that retains the topological properties of the space. These are topological spaces assembled from vertices, edges, triangles, tetrahedra, and possibly higher-dimensional simplices[8]. In this situation, homology is defined as described earlier, by comparing chains that are cycles with chains that are boundaries. The availability of simplices has a number of advantages, including the existence of fast algorithms to compute homology. Here, we focus on the connection between the number of simplices in the simplicial complex and the Betti numbers of the space.

Suppose $\mathbb{M}$ is a d-dimensional manifold, and K is a triangulation of $\mathbb{M}$. Write n_p for the number of p-dimensional

[7]For consistency, it is important to note that the definition in previous topology studies in cosmology [9], [10] slightly differs from this. The genus, g, in these studies has been defined as the number of holes minus the number of connected regions: $g = G - c$. Here, we will refer to g as the *reduced genus*.

[8]Technically, we require that the union of simplices in the simplicial complex is *homeomorphic* to the space it represents, which means that there is a bijective map between the two sets that is continuous and whose inverse is continuous. As an example, consider the boundary of the octahedron, consisting of 6 vertices, 12 edges, and 8 triangles, which forms a triangulation of the 2-dimensional sphere.

simplices in K. For example, n_0 is the number of vertices, n_1 is the number of edges, and so on. The *Euler characteristic* of K is defined as the alternating sum of these numbers:

$$\chi(K) = \sum_{p=0}^{d}(-1)^p n_p. \quad (11)$$

This is the d-dimensional generalization of the classical Euler characteristic of a polytope:

$$\chi = \#\text{vertices} - \#\text{edges} + \#\text{faces}. \quad (12)$$

For a convex polytope, Euler's Formula states that $\chi = 2$. The *Euler-Poincaré Formula* is a far-reaching generalization of this relation. To state this generalization, we first note that homology is independent of the choice of triangulation, and so are the Betti numbers and the Euler characteristic. The Euler-Poincaré Formula says that the Euler characteristic of a triangulation, which is the alternating sum of simplex numbers, is equal to the alternating sum of Betti numbers of the triangulated space:

$$\chi(K) = \sum_{p=0}^{d}(-1)^p \beta_p(\mathbb{M}). \quad (13)$$

Coming back to the case of a convex polytope in 3-dimensional space, its faces decompose the boundary, which is homeomorphic to the 2-dimensional sphere. We may further decompose the faces into triangles, if necessary, but this makes no difference here. Since this is true for all convex polytopes, their faces are but different triangulations of this same sphere, so the alternating sums must be the same. They are all equal to 2 because this is the Euler characteristic of the 2-dimensional sphere.

C. Homology of a Filtration

For the assessment of the topology of a mass or point distribution, a rich source of information is the topological structure of a filtration. Given a space $\mathbb{M}$, a *filtration* is a nested sequence of subspaces:

$$\emptyset = \mathbb{M}_0 \subseteq \mathbb{M}_1 \subseteq \ldots \subseteq \mathbb{M}_m = \mathbb{M}. \quad (14)$$

The nature of the filtrations depends, amongst others, on the representation of the mass distribution. When representing the mass distribution by a continuous density field, $f(\mathbf{x})$, a common practice is to study the sublevel or superlevel sets of the field smoothed on a scale R_s:

$$f_s(\mathbf{x}) = \int f(\mathbf{y})W_s(\mathbf{y} - \mathbf{x})\,d\mathbf{y}, \quad (15)$$

where $W_s(\mathbf{x} - \mathbf{y})$ is the smoothing kernel. The *sublevel sets* of this field are defined as the regions

$$\mathbb{M}_\nu = \{\mathbf{x} \in \mathbb{M} \mid f_s(\mathbf{x}) \in (-\infty, f_\nu]\} \quad (16)$$
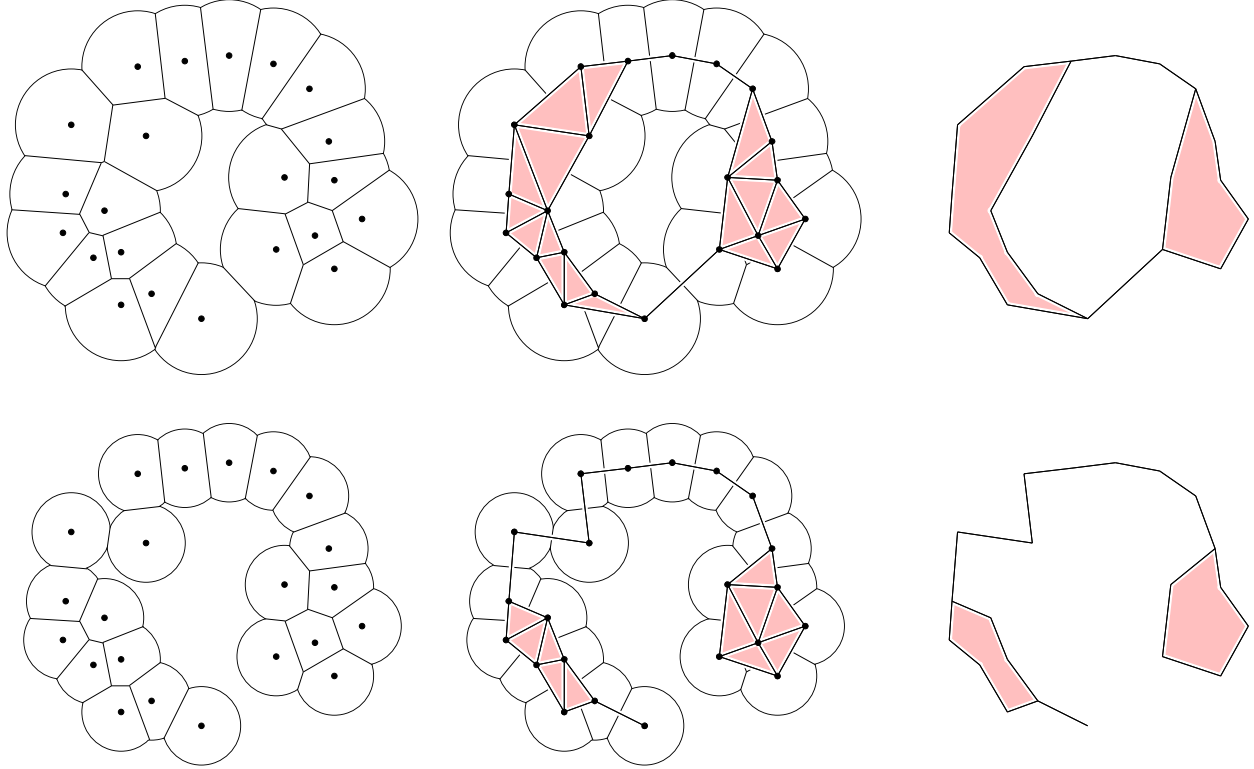$$= f_s^{-1}(-\infty, f_\nu]. \quad (17)$$

Figure 2. Illustration of alpha shapes. For two different values of α, we show the relation between the 2-dimensional point distribution, the value of α, and the resulting alpha shape. Around each point in the sample, we draw a circle of radius α. The outline of the corresponding Voronoi tessellation is indicated by the edges (left). All Delaunay simplices *dual* to the decomposition of the union of disks by the Voronoi polygons are shown in black (center). The final resulting alpha shape is shown on the right. Top: large α value. Bottom: small α value.

In other words, they are the regions where the smoothed density is less than or equal to the threshold value $f_\nu = \nu\sigma_0$, with σ_0 the dispersion of the density field. When addressing the topology of the primordial Gaussian density field, as in Section VIII, the analysis will be based on the filtration consisting of superlevel sets.

The representation of the mass by a discrete particle or galaxy distribution leads to an alternative strategy of filtration, this time in terms of simplicial complexes generated by the particle distribution. In the case of simplicial complexes that are homotopy equivalent to the sublevel sets of either the density or the distance function field, the computation of the homological characteristics of the field is considerably facilitated. Often it is also much easier to visualize the somewhat abstract notion of homology by means of these simplicial complexes.

In our study, we will concentrate on *alpha shapes* of a point set, which are subsets of the corresponding Delaunay triangulation; see Section III. The alpha shapes are homotopy equivalent to the sublevel sets of the distance field defined by the point distribution, and they constitute a nested sequence of simplicial complexes that forms a topologically useful filtration of the Delaunay triangulation.

In the following sections, we will extensively discuss the use of alpha shapes to assess the homology of cosmological particle and galaxy distributions.

III. ALPHA SHAPES

One of the key concepts in the field of Computational Topology are *alpha shapes*, as introduced by Edelsbrunner and collaborators [30], [37], [38]; see [36] for a recent review. They generalize the convex hull of a point set and are concrete geometric objects that are uniquely defined for a particular point set and a real value α. For their definition, we look at the union of balls of radius α centered on the points in the set, and its decomposition by the corresponding Voronoi tessellation; see the left diagrams in Figure 2. The *alpha complex* consists of all Delaunay simplices that record the subsets of Voronoi cells that have a non-empty common intersection within this union of balls; see the center diagrams of Figure 2. The *alpha shape* is the union of simplices in the alpha complex; see the right diagrams of Figure 2.

Alpha shapes reflect the topological structure of a point distribution on a scale parameterized by the real number α. The ordered set of alpha shapes constitute a filtration of the
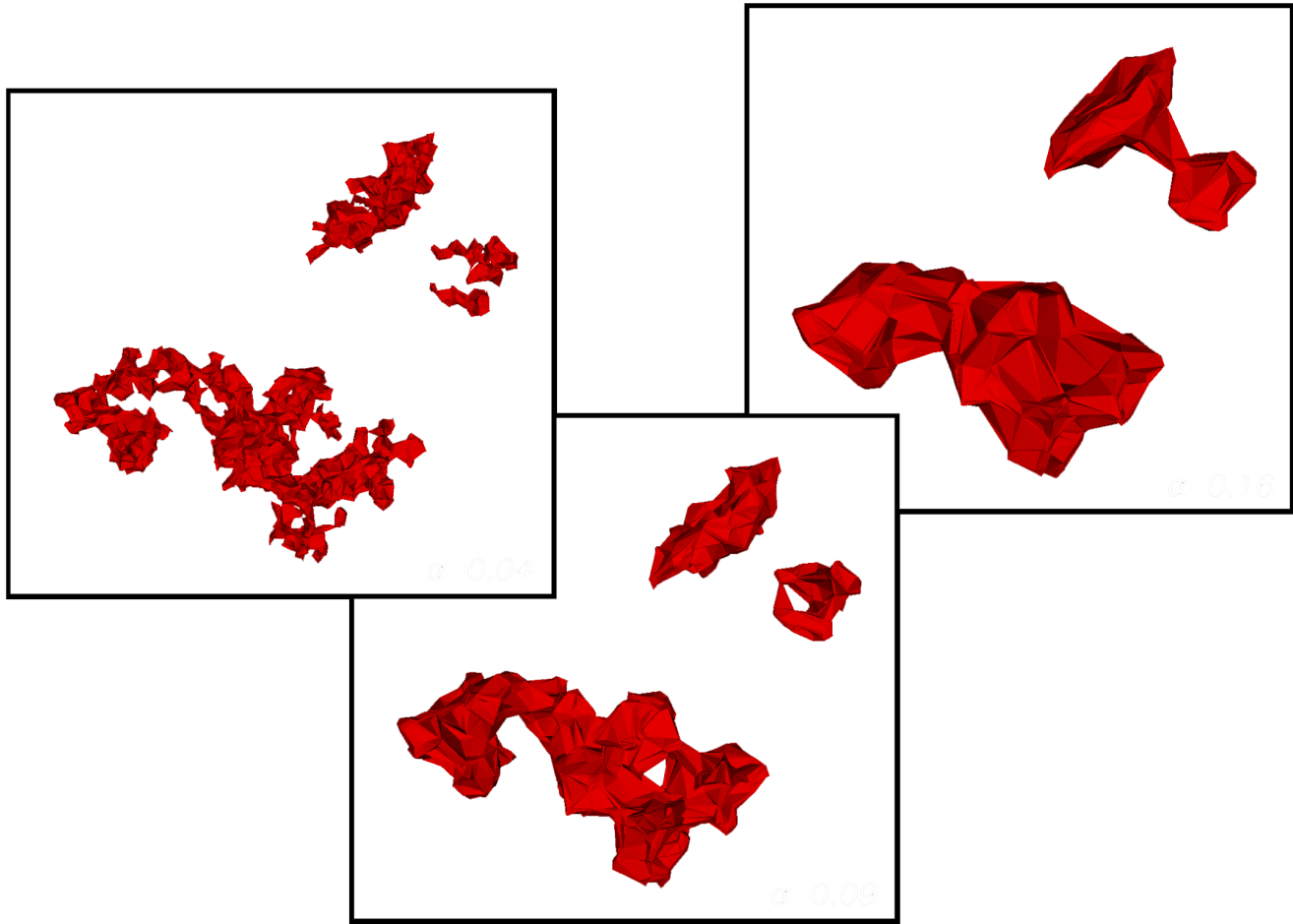
Figure 3. Illustration of 3-dimensional alpha shapes for three different values of $\alpha = 0.001, 0.005, 0.01$. Clearly visible are the differences in connectivity and topological structure of the different alpha shapes.

Delaunay tessellation. The link between alpha shapes and the homology of a point distribution can be appreciated from the fact that tunnels will be formed when, at a certain value of α, an edge is added between two vertices that were already connected, which increases the first Betti number. When new triangles are added, the tunnel may be filled, which decreases the first Betti number. More about this process of growing and shrinking Betti numbers in Section IX, where we discuss the persistence of tunnels and other topological features.

Connections of alpha shapes to diverse areas in the sciences and engineering have developed, including to pattern recognition, digital shape sampling and processing, and structural molecular biology [36]. Applications of alpha shapes have as yet focussed on biological systems, where they have been used in the characterization of the topology and structure of macromolecules. The work by Liang and collaborators [39], [40], [41], [42] uses alpha shapes and Betti numbers to assess the voids and pockets in an effort to classify complex protein structures, a highly challenging task given the tens of thousands of protein families involving thousands of different folds. Given the interest in the topology of the cosmic mass distribution [9], [20], [21], [22], it

is evident that alpha shapes also provide a highly interesting tool for studying the topology of the galaxy distribution and the particles in N-body simulations of cosmic structure formation. Directly connected to the topology of the point distribution itself, it avoids the need of any user-defined filter kernels.

A. Definition

Figure 2 provides an impression of the concept by illustrating the process of defining the alpha shape, for two different values of α. If we have a finite point set, S, in 3-dimensional space and its Delaunay triangulation, we may identify all simplices – vertices, edges, triangles, tetrahedra – in the triangulation. For a given non-negative value of α, the *alpha complex* consists of all simplices in the Delaunay triangulation that have an empty circumsphere with radius less than or equal to α. Here "empty" means that the open ball bounded by the sphere does not include any points of S. For an extreme value $\alpha = 0$, the alpha complex merely consists of the vertices of the point set. There is also a smallest value, α_{max}, such that for $\alpha \geq \alpha_{max}$, the alpha complex is the Delaunay triangulation and the alpha shape
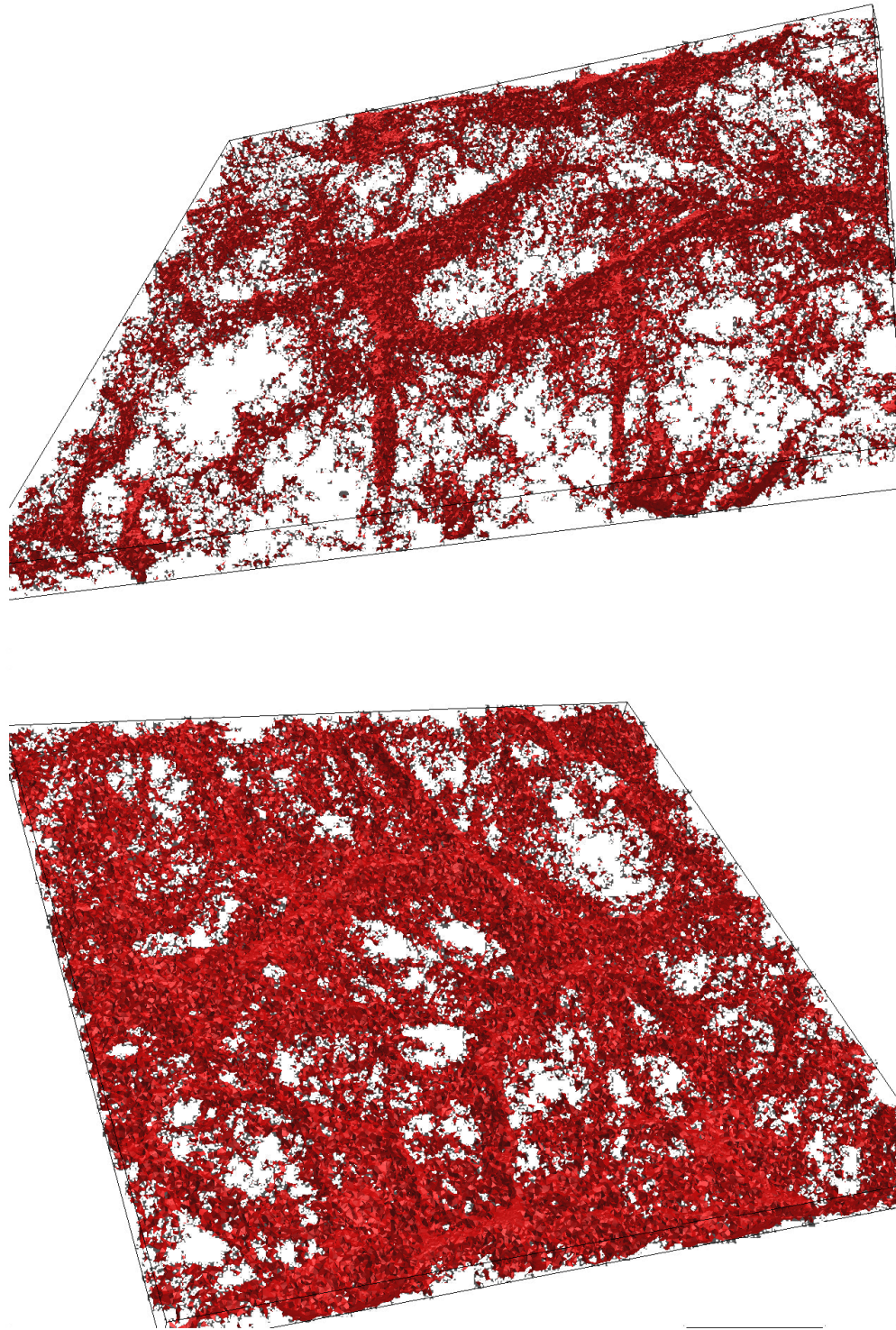
Figure 4. Examples of alpha shapes of the LCDM GIF simulation. Shown are central slices through the complete alpha shape for two different values of α, viewed from different angles. The sensitivity to the structure and topology of the matter distribution in the Cosmic Web is clearly visible when comparing the lower value of α in the top panel with the higher value of α in the bottom panel.

is the convex hull of the point set.

As mentioned earlier, the alpha shape is the union of all simplices in the alpha complex. It is a polytope in a fairly general sense: it can be concave and even disconnected. Its components can be three-dimensional clumps of tetrahedra, two-dimensional patches of triangles, one-dimensional strings of edges, and collections of isolated points, as well as combinations of these four types. The set of real numbers leads to a family of shapes capturing the intuitive notion of the overall versus fine shape of a point set. Starting from the convex hull gradually decreasing α, the shape of the point set gradually shrinks and starts to develop enclosed voids. These voids may join to form tunnels and larger voids. For negative α, the alpha shape is empty. An intuitive feel for the evolution of the topological structure may be obtained from the three different alpha shapes of the same point set in 3-dimensional space shown in Figure 3.

It is important to realize that alpha shapes are not triangulations of the union of balls in the technical sense. Instead, they are simplicial complexes that are homotopy equivalent to the corresponding union of balls with radius α. While this is a nontrivial observation that follows from the Nerve Theorem [43], it is weaker than being homeomorphic, which would be necessary for being a triangulation. Nevertheless, the homotopy equivalence implies that the alpha shape and the corresponding union of balls have the same Betti numbers.

Although the alpha shape is defined for all real numbers α, there are only a finite number of different alpha shapes for any finite point set. In other words, the alpha shape process is never continuous: it proceeds discretely with increasing α, marked by the addition of new Delaunay simplices once α exceeds the corresponding threshold.

B. Computing Betti Numbers

Following the description above, one may find that alpha shapes are intimately related to the topology of a point set. Indeed, they form a direct way of characterizing the topology of a point distribution. The complete description of its homology in terms of Betti numbers may therefore be inferred from the alpha shapes.

For simplicial complexes, like Delaunay tessellations and alpha complexes, the Betti numbers can be defined on the basis of the p-simplices. We illustrate this for a simplicial complex in 3-dimensional space. Cycling through all its simplices, we base the calculation on the following straightforward considerations. When a vertex is added to the alpha complex, a new component is created and β_0 is increased by 1. Similarly, if an edge is added, it connects two vertices, which either belong to the same or to different components of the current complex. In the former case, the edge creates a new tunnel, so β_1 is increased by 1. In the latter case, two components get connected into one, so β_0 is decreased by 1. If a triangle is added, it either completes a void or

it closes a tunnel. In the former case, β_2 is increased by 1, and in the latter case, β_1 is decreased by 1. Finally, when a tetrahedron is added, a void is filled, so β_2 is lowered by 1. Following this procedure, the algorithm has to include a technique for determining whether a p-simplex belongs to a p-cycle. For vertices and tetrahedra, this is rather trivial. On the other hand, for edges and triangles, we use a somewhat more elaborate procedure, involving the classical computer science concept of a union-find data structure [44].

Turning our attention to software that implements these algorithmic ideas, we resort to the Computational Geometry Algorithms Library, CGAL [9]. Within this context, Caroli and Teillaud recently developed an efficient code for the calculation of two-dimensional and three-dimensional alpha shapes in periodic spaces. We use their software for the computation of the alpha shapes of our cosmological models. In the first stage of our project, which concerns the analysis of Voronoi clustering models, we computed the Betti numbers of alpha shapes with a code developed within our own project. Later, for the analysis of the cosmological LCDM models, we were provided with an optimized code written by Manuel Caroli.

C. Alpha Shapes of the Cosmic Web

In a recent study, Vegter et al. computed the alpha shapes for a set of GIF simulations of cosmic structure formation [45]. It concerns a 256^3 particles GIF N-body simulation, encompassing a LCDM ($\Omega_m = 0.3$, $\Omega_\Lambda = 0.7$, $H_0 = 70$ km/s/Mpc) density field within a (periodic) cubic box with length $141h^{-1}$Mpc and produced by means of an adaptive P^3M N-body code [46].

Figure 4 illustrates the alpha shapes for two different values of α, by showing sections through the GIF simulation. The top panel uses a small value of α, the bottom uses a high value. The intricacy of the web-like patterns is very nicely followed. The top configuration highlights the interior of filamentary and sheet-like features, and reveals the interconnection between these structural elements. The bottom configuration covers an evidently larger volume, which it does by connecting finer features in the Cosmic Web. Noteworthy are the tenuous filamentary and planar extensions into the interior of the voids.

These images testify of the potential power of alpha shapes in analyzing the web-like cosmic matter distribution, in identifying its morphological elements, their connections and, in particular, their hierarchical character. However, to understand and properly interpret the topological information contained in these images, we need first to assess their behavior in simpler yet similar circumstances. To this end, we introduce a set of heuristic spatial matter distributions, Voronoi clustering models.

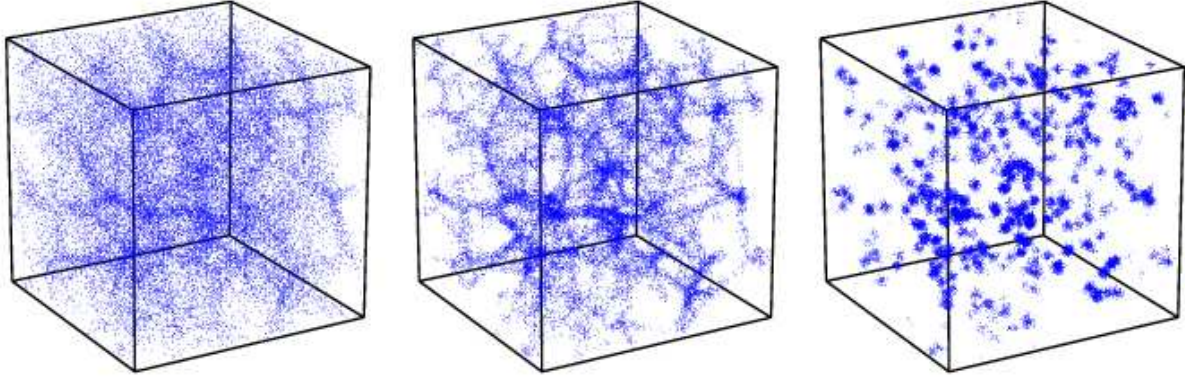[9]CGAL is a C++ library of algorithms and data structures for computational geometry, see www.cgal.org.

Figure 5. Three different Voronoi element models, shown in a cubic setting. From left to right: a wall-dominated Voronoi universe, a filamentary Voronoi universe, and a cluster-dominated Voronoi universe.

IV. VORONOI CLUSTERING MODELS

In this section, we introduce a class of heuristic models for cellular distributions of matter using Voronoi tessellations as scaffolds, know as *Voronoi clustering models* [47], [48], [49]. They are well suited to model the large scale clustering of the morphological elements of the Cosmic Web, defined by the stochastic yet non-Poissonian geometrical distribution of the matter and the related galaxy population forming *walls*, *filaments*, and *clusters*.

The small-scale distribution of galaxies within the various components of the cosmic skeleton involves the complicated details of highly nonlinear interactions of the gravitating matter. This aspect may be provided by elaborate physical models and/or N-body computer simulations, but this would distract from the purpose of the model and destroy its conceptual simplicity. In the Voronoi models, we complement a geometrically fixed Voronoi tessellation defined by a small set of nuclei with a heuristic prescription for the location of particles or model galaxies within the tessellation. We distinguish two complementary approaches: the *Voronoi element* and the *Voronoi evolution models*. Both are obtained by moving an initially random distribution of N particles toward the faces, lines, and nodes of the Voronoi tessellation. The Voronoi element models do this by a heuristic and user-specified mixture of projections onto the various geometric elements of the tessellation. The Voronoi evolution models accomplish this via a gradual motion of the galaxies from their initial, random locations towards the boundaries of the cells.

The Voronoi clustering models identify the geometric elements of a 3-dimensional Voronoi tessellations with the morphological component of the Cosmic Web. To describe this relationship, we will adhere to the terminology listed in Table I, which also lists the terminology to the corresponding types of holes in the alpha shape of the particle

Cosmic Web	Voronoi	Alpha
voids, field	cell	void
walls, sheets, superclusters	face	tunnel
filaments, superclusters	line	gap
clusters	node	

Table I
IDENTIFICATION OF MORPHOLOGICAL COMPONENTS OF THE COSMIC WEB (LEFTHAND COLUMN) WITH GEOMETRIC VORONOI TESSELLATION ELEMENTS (CENTRAL COLUMN). THE RIGHTHAND COLUMN LIST THE TERMINOLOGY FOR THE CORRESPONDING TYPES OF HOLES IN THE ALPHA SHAPE OF THE PARTICLE DISTRIBUTION.

distribution[10].

A. Voronoi Element Models

These are fully heuristic, user-specified spatial galaxy distributions within the cells, faces, lines, and nodes of a Voronoi tessellation. They are obtained by projecting the initially randomly distributed N model galaxies onto the relevant Voronoi face, line, or node. We can also retain a galaxy within the Voronoi cell in which it is located. The Voronoi element models are particularly apt for studying systematic properties of spatial galaxy distributions confined to one or more structural elements of nontrivial geometric spatial patterns.

Pure Voronoi element models place their galaxies exclusively inside or near either faces, lines, or nodes. In contrast,

---

[10]For consistency and clarity, throughout this paper we adopt the following nomenclature for the spatial components of the Cosmic Web, of Voronoi tessellations and of Delaunay tessellations. The cosmic web consists of *voids, walls, filaments, and clusters*. Voronoi tessellations consist of *cells, faces, lines, and nodes*. Finally, Delaunay tessellations (and alpha shapes) involve *tetrahedra, triangles, edges, and vertices*.

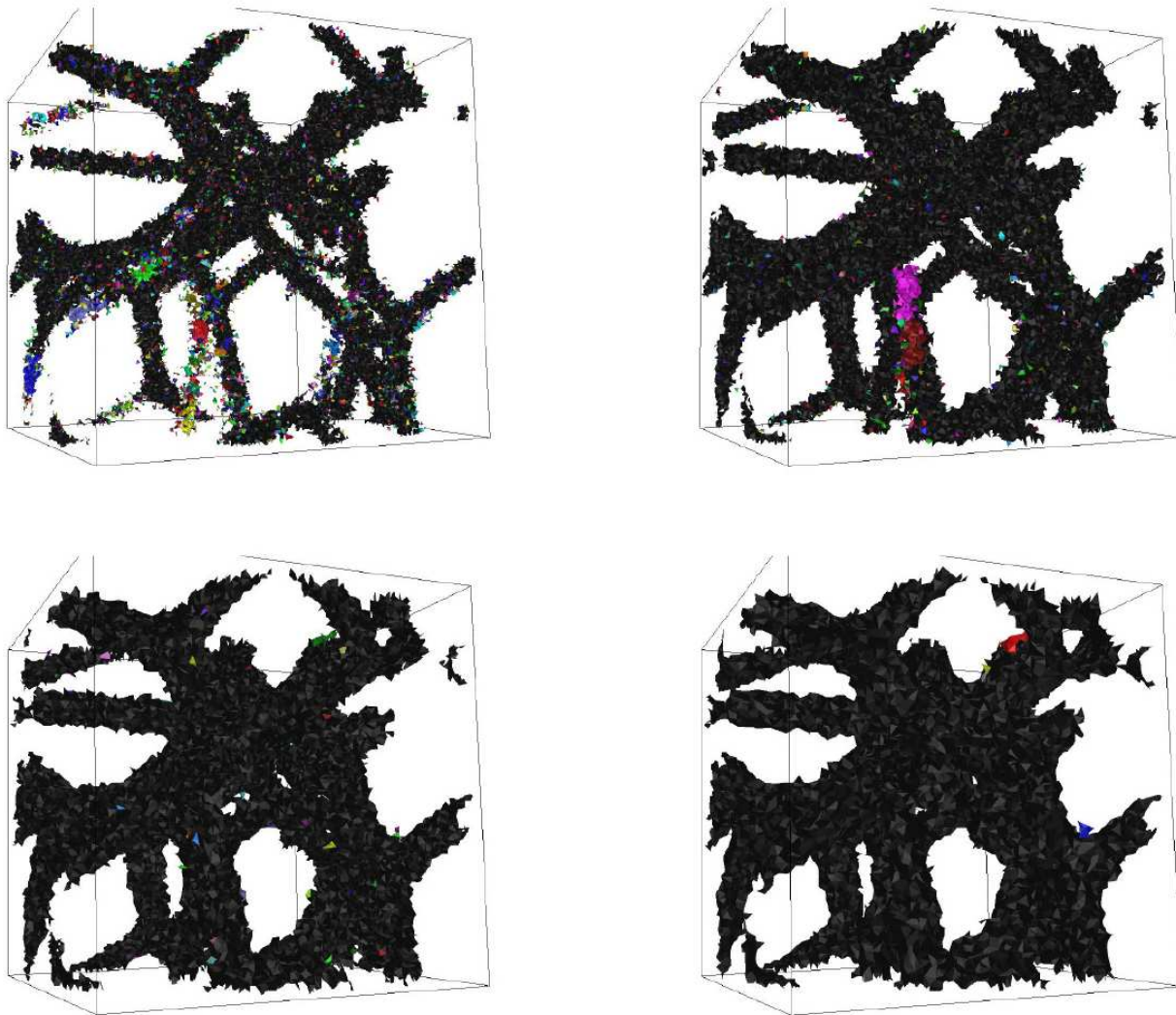Figure 6. Four alpha shapes of a Voronoi filament model consisting of $200,000$ particles in a periodic box of size $50h^{-1}\mathrm{Mpc}$ with 8 Voronoi cells. We use colors to highlight different components. From top left to bottom right: $10^4 \cdot \alpha = 0.5, 1.0, 2.0, 4.0$.

mixed models allow combinations in which distributions inside or near cells, faces, lines, and nodes are superimposed. These include particles located in four distinct 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 lines,
- *cluster* particles within and around the Voronoi nodes.

The characteristics of the spatial distributions in the mixed models can be varied and tuned according to desired fractions of galaxies of each type. These fractions are free parameters that can be specified by the user; see Figure 5.

B. Voronoi Evolution Models

The second class we consider are the *Voronoi evolution models*. They provide web-like galaxy distributions mimicking the outcome of realistic cosmic structure formation scenarios. They are based upon the notion that voids play a key organizational role in the development of structure, causing the universe to resemble a soapsud of expanding bubbles [50]. While the galaxies move away from the void centers, and stream towards the walls, filaments, and clusters, the fractions of galaxies in or near the cells, faces, lines, and nodes evolve continuously. The details of the model realization depends on the specified time evolution.

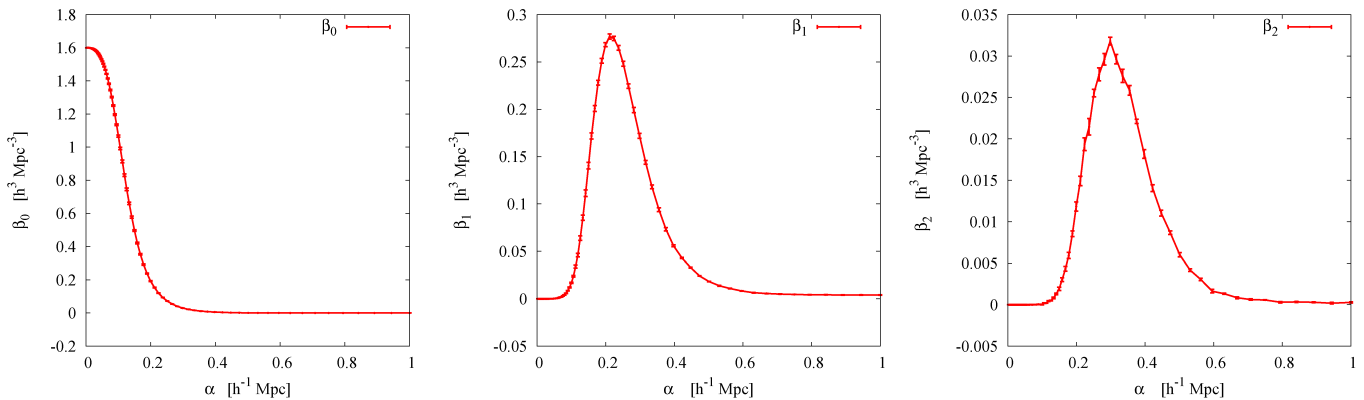Within the class of Voronoi evolution models, the most

Figure 7. Topological analysis of the Voronoi filament model illustrated by its alpha shape shown in the upper left panel of Figure 6. From left to right: β_0, β_1, and β_2 counting the components, tunnels, and voids of the alpha shape. The realization contains 64 nuclei or cells.

representative and most frequently used are the *Voronoi kinematic models*. Forming the idealized and asymptotic description of the outcome of a hierarchical gravitational structure formation process, they simulate the asymptotic web-like galaxy distribution implied by the hierarchical void formation process by assuming a single-size dominated void population. Within a void, the mean distance between galaxies increases with time. Before a galaxy enters an adjacent cell, the velocity component perpendicular to the otherwise crossed face disappears. Thereafter, the galaxy continues to move within the face. Before it enters the next cell, the velocity component perpendicular to the otherwise crossed edge disappears. The galaxy continues along a filament and, finally, comes to rest at a node. The resulting evolutionary progression within the Voronoi kinematic model proceeds from an almost featureless random distribution towards a distribution in which matter ultimately aggregates into conspicuous compact cluster-like clumps.

The steadily increasing contrast of the various structural features is accompanied by a gradual shift in the topology of the distribution. The virtually uniform and featureless particle distribution at the beginning ultimately unfolds into a highly clumped distribution of clusters. This evolution involves a gradual progression via a wall-like through a filamentary towards an ultimate cluster-dominated matter distribution.

V. TOPOLOGICAL ANALYSIS OF VORONOI UNIVERSES

In this section, we study the systematic behavior of the Betti numbers of the alpha shapes of Voronoi clustering models; see [45], [51]. For each point sample, we investigate the alpha shape for the full range of the α parameter. We generate six Voronoi clustering models, each consisting of $200,000$ particles within a periodic box of size $50h^{-1}\mathrm{Mpc}$. Of each model, we make two realizations, one with 8 and

the other with 64 nuclei. We start with two pure Voronoi element models: a wall model, and a filament model. In addition, we study four Voronoi kinematic models, ranging from a mildly evolved to a strongly evolved configuration. In each case, the clusters, filaments, and walls have a finite Gaussian width of $1.0h^{-1}\mathrm{Mpc}$.

An impression of the sequence of alpha shapes may be gained from the four panels in Figure 6. For the smallest value of α, we see that the simplices delineate nearly all the filaments in the particle distribution. As α increases, going from the top-left panel down to the bottom right panel, we find that the alpha shape fills in the walls. For even larger values of α, the alpha shape includes the large Delaunay simplices that cover the interior of the Voronoi cells. It is a beautiful illustration of the way in which alpha shapes define, as it were, naturally evolving surfaces that are sensitive to every detail of the morphological structure of the cosmic matter distribution.

A. Filament Model Topology

We take the Voronoi filament model as a case study, investigating its topology by following the behavior of the three Betti numbers as functions of the parameter α.

Figure 7 shows the relation between the Betti numbers of the alpha shape and the value of α. The zeroth Betti number, β_0, counts the components. Equivalently, $\beta_0 - 1$ counts the gaps between the components. In the current context, the latter interpretation is preferred as we focus on the holes left by the alpha shape. Starting with $\beta_0 = 200,000$ at $\alpha = 0$, the zeroth Betti number gradually decreases to 1 as the components merge into progressively larger entities.

The first Betti number, β_1, counts the tunnels. At first, it increases steeply when edges are added to the alpha complex, some of which bridge gaps while others form tunnels. After β_1 reaches its maximum for α roughly equal
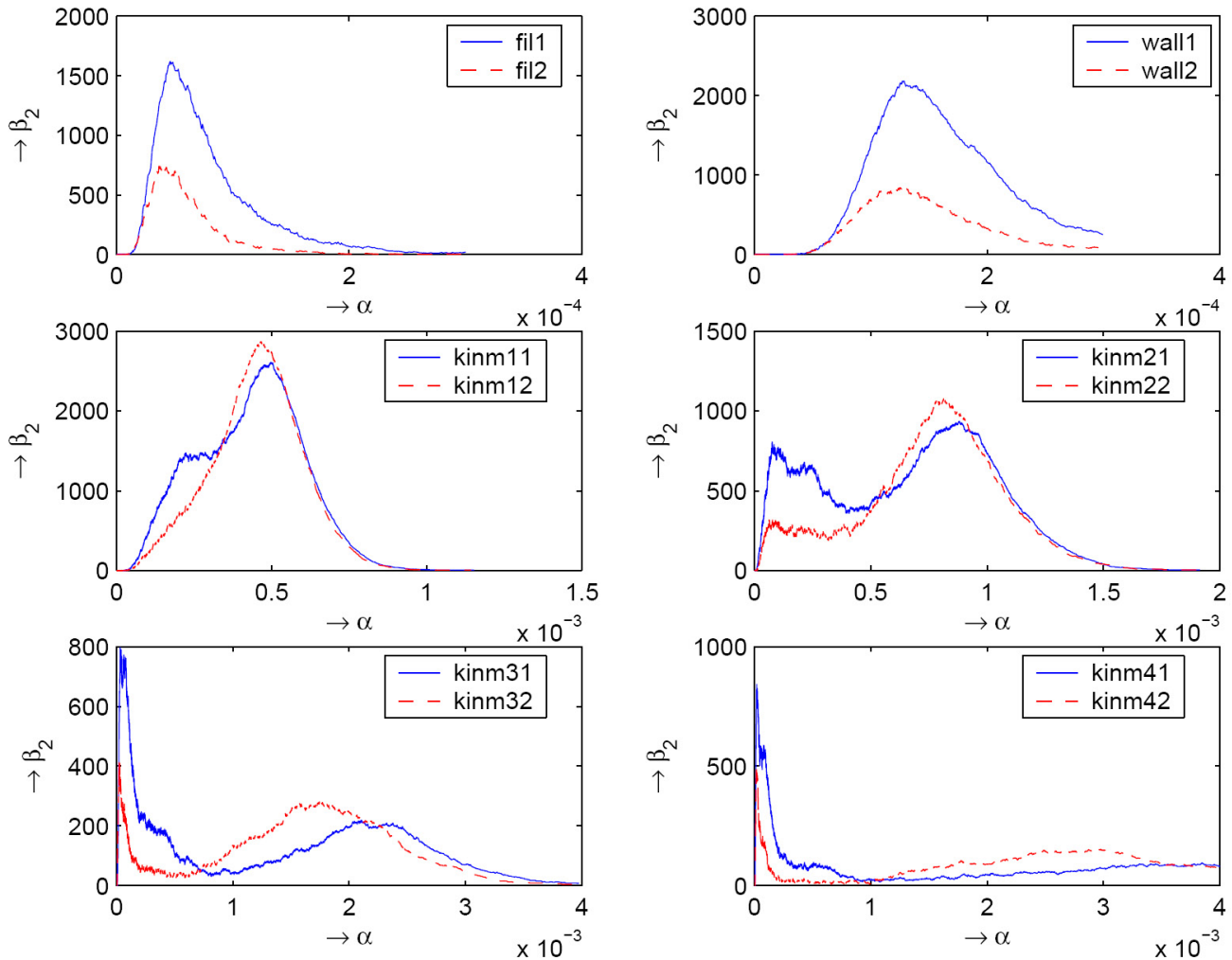
Figure 8. The dependence of the second Betti number, β_2, on α, for six different Voronoi clustering models. Top left: Voronoi filament model. Top right: Voronoi wall model. Centre left to bottom right: four stages of the Voronoi kinematic model, going from a moderately evolved model dominated by walls (center left) to a highly evolved model dominated by filaments and clusters (bottom right). Blue lines: realizations with 8 nuclei or cells. Red lines: realizations with 64 nuclei or cells.

to 0.25 times 10^{-4}, the number of tunnels decreases sharply as triangles enter the alpha complex in large numbers.

The second Betti number, β_2, counts the voids in the alpha shape. In the case of the Voronoi filament model, its behavior resembles that of β_1: a peaked distribution around a moderate value of α. It increases when the entering triangles are done closing tunnels and start creating voids. However, when α is large enough to add tetrahedra, these voids start to fill up and β_2 decreases again, reaching zero eventually. Notice that in the given example of the Voronoi filament model, β_2 reaches its maximum for α roughly equal to 0.45 times 10^{-4}, which lies substantially beyond the peak in the β_1 distribution; see also Figure 8.

B. Void Evolution

Having assessed one particular Voronoi clustering model in detail, we turn our attention to the differences between the models. While this is still the subject of ongoing research, we find substantial differences on a few particular aspects. For example, β_0 decreases monotonically with α for all models, but the range over which the number of gaps is substantially positive, and the rate with which it decrease are highly sensitive to the underlying distribution. In fact, the approximate derivative, $\partial \beta_0 / \partial \alpha$, contains interesting features. Examples are a minimum and a varying width, both potentially interesting for discriminating between the underlying topologies.

Most interesting is the difference in behavior of the second Betti number. As one may infer from Figure 8, substantial
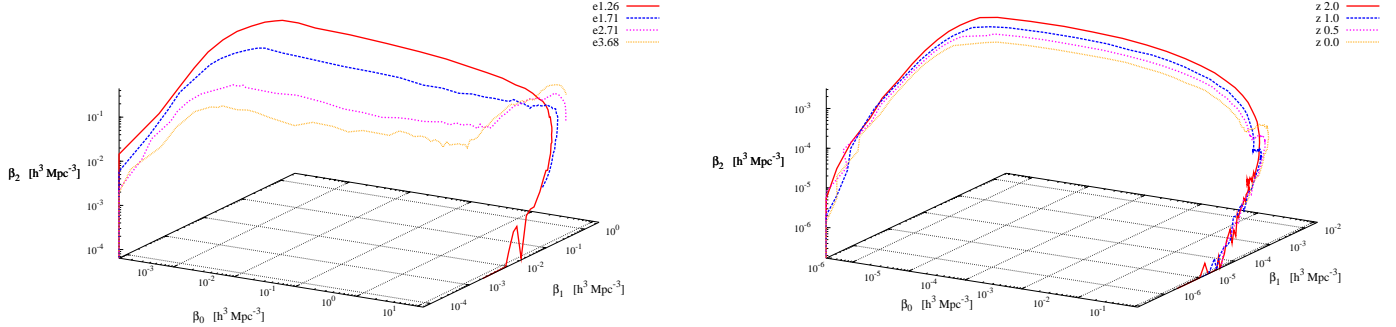
Figure 9. Left: the alpha tracks for four stages of the kinematic Voronoi clustering model, evolving from almost uniform (red) to a distribution in which most matter resides in cluster nodes (orange). The wall-like (blue) and filamentary (magenta) distributions form intermediate cases. Right: the alpha tracks at different cosmic epochs in an evolving LCDM mass distribution. The redshift, z, is indicated in the top right corner, with the track at an early epoch ($a = 0.17$, $z = 5.0$) bearing resemblance to the early Voronoi kinematic model, and the track at the current epoch ($a = 1.0$, $z = 0.0$) resembling the Voronoi kinematic model at an intermediate distribution.

differences between models can be observed. This concerns the values and range over which β_2 reaches maxima, as well as new systematic behavior. For kinematic models, we find two or more peaks, each corresponding to different morphological components of the particle distribution. It is revealing to follow the changes in the dependence of β_2 on α, as we look at different evolutionary stages. The panels from center left to bottom right in Figure 8 correspond to four different stages of evolution. The center left panel shows the dependence of β_2 on α for a moderately evolved matter distribution, which is dominated by walls. The center right panel shows the dependence for a stage at which walls and filaments are approximately equally prominent. In the bottom left panel, the filaments represent more than 40% of the mass, while the walls and the gradually more prominent clusters each represent around 25% of the particles. The final, bottom right panel shows the dependence for a highly evolved mass distribution, with clusters and filaments each representing around 40% of the particles.

The different morphological patterns of the Voronoi kinematic models are reflected in the behavior of β_2. In the center left panel, we find a strong peak at α roughly equal to 5 times 10^{-4}, with a shoulder at lower values. The peak reflects the voids inside the walls in the distribution, while the shoulder finds its origin in the somewhat smaller voids inside the filaments. The identity of the peaks becomes more clear when we turn to the two-peak distribution in the center right panel. The strong peak at α roughly equal to 10^{-4} is a manifestation of the strongly emerged filaments in the matter distribution. As the shift to filaments and clusters continues, we see the rise of a third peak at a much smaller value of α; see the bottom two panels. This clearly corresponds to the voids in the high density cluster regions.

C. Alpha Tracks

In Figure 9, we synthesize the homology information by tracing curves in the 3-dimensional *Betti space*, whose

coordinates are the three Betti numbers. Each curve is parametrized by α, with points $(\beta_0(\alpha), \beta_1(\alpha), \beta_2(\alpha))$ characterizing the homology of the alpha shape at this given scale. We call each curve an *alpha track*, using it to visualize the combined evolution of the Betti numbers for a given Voronoi cluster model. Note that the evolution of the Euler characteristic, $\chi(\alpha)$, is the projection of an alpha track onto the normal line of the plane $\beta_0 - \beta_1 + \beta_2 = 0$. The richer structure of the full alpha track illustrates why patterns with similar genus may still have a substantially different topology.

The left frame of Figure 9 shows the alpha tracks for the four Voronoi kinematic models. We notice that β_0 varies over a substantially larger range of values than β_1 and β_2. The tracks reveal that the few components left after an extensive process of merging still contain a substantial number of tunnels and voids. More generally, we see that the Betti numbers diminish to small values in sequence: first β_0, then β_1, and finally β_2.

As the matter distribution evolves, the tracks shift in position. The four tracks in the left frame of Figure 9 form a good illustration of this effect. The red dashed track corresponds to an early phase of the kinematic model, at which the particle distribution is still almost uniform, while the orange track represents the final time-step marked by a pattern in which nearly all particles reside in the clusters. In the intermediate stages, most of the particles are located in the walls (blue) and in the filaments (magenta). The decreasing number of tunnels and voids at intermediate scales is a manifestation of the hierarchical formation process, in which small scale structures merge into ever larger ones; see e.g. [52].

VI. TOPOLOGICAL ANALYSIS OF THE LCDM UNIVERSE

Having discussed the scale-dependent Betti numbers in heuristic Voronoi clustering models, we turn to the analysis
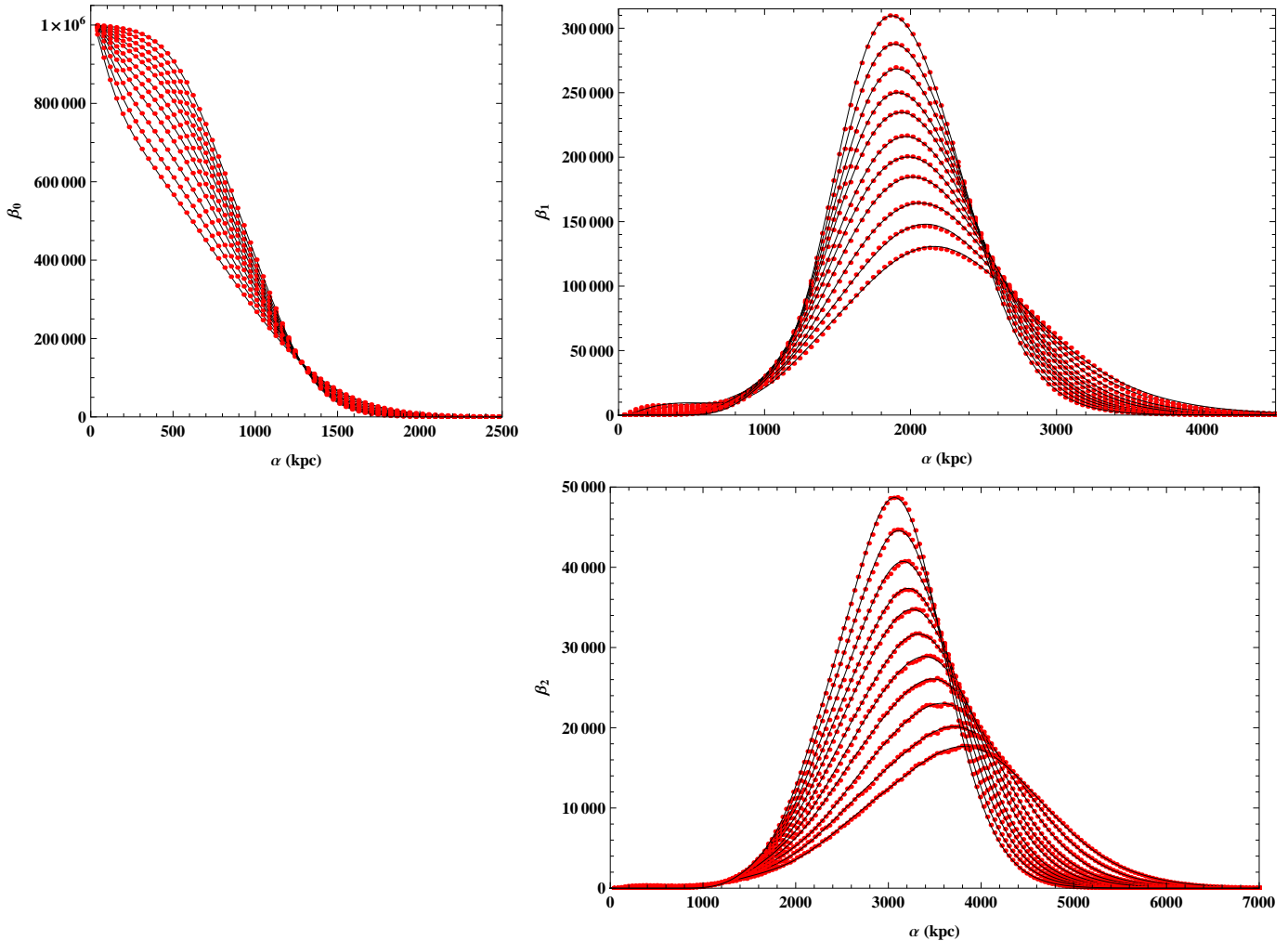
Figure 10. The evolution of the Betti numbers in an LCDM universe. We show the dependence of β_p on α for dimensions $p = 0, 1, 2$ and for eleven expansion factors running from $a = 0.25$ ($z = 3.0$) to the current epoch $a = 1.0$ ($z = 0.0$). From upper left to lower right: $\beta_0(\alpha)$, $\beta_1(\alpha)$, $\beta_2(\alpha)$. Note that all three curves gradually shift from smaller to larger scales.

of more realistic megaparsec cosmic mass distributions. These are characterized by an intricate multiscale configuration of anisotropic web-like patterns. The crucial question is whether we can exploit the topological information toward determining crucial cosmological parameters, such as the nature of dark energy.

We concentrate on the analysis of computer simulations of cosmic structure formation in the Universe, leaving the analysis of the observed distribution in galaxy redshift surveys for the future. The computer simulations follow the nonlinear evolution of structure as it emerges from the near uniform early Universe. Once the gravitational clustering process has progressed beyond the initial linear growth phase, we see the emergence of intricate patterns in the density field; see Figure 1. The near homogeneous initial conditions evolve into an increasingly pronounced clustering pattern.

A. N-body Simulations

In cosmological N-body simulations, the cosmic mass distribution is represented by a large number of particles, which move under the influence of the combined gravitational force of all particles. The initial conditions (location and velocities of the particles) are a realization of the mass distribution expected in the cosmological scenario at hand. Since the majority of the matter in the Universe is non-dissipative dark matter, a major aspect of cosmological N-body simulations concerns itself with dark matter particles that only interact via gravity. State-of-the-art computer simulations, such as the Millennium simulation [8], count in the order of 10^{10} particles and are run on the most powerful supercomputers available to the scientific community.

Figure 1 illustrates the matter distribution in the standard, or "concordance" LCDM cosmological model. It shows the dark matter distribution in a box of $80h^{-1}\mathrm{Mpc}$ co-moving

size, based on a 384^3 particle N-body simulation. The LCDM cosmological scenario assumes that most gravitating matter in the Universe consists of as yet undetected and unidentified cold dark matter, accounting for approximately 23% of the energy content of the Universe, i.e. $\Omega_{dm} = 0.23$. The normal, baryonic matter, which consists mostly of protons and neutrons, represents only 4.4% of the energy content of the Universe, i.e. $\Omega_b = 0.044$. Most importantly, the model assumes the presence of a cosmological constant, Λ, or equivalent dark energy component, representing 73% of the density of the Universe, i.e. $\Omega_\Lambda = 0.73$; see Section VII. The Hubble parameter, which specifies the expansion rate of the Universe, is taken to be $\mathrm{H}_0 = 70$ km/s/Mpc. The amplitude of the initial fluctuation has been normalized to a level at which the current density fluctuation on a scale of $8h^{-1}\mathrm{Mpc}$ is equal to $\sigma_8 = 0.8$.

The most prominent aspect of the cosmic mass distribution is the intriguing network of filaments and high-density clusters, which surround low-density voids. The mass distribution has a distinct multiscale character, marked by clumps over a wide range of scales, reflecting the hierarchical evolution of the distribution.

B. Homological Evolution

We follow the developing structure in the LCDM scenario in eleven time-steps that run from an expansion factor $a = 0.25$ ($z = 3.0$) to $a = 1.0$ ($z = 0.0$), i.e. from around 3.4 Gigayears after the Big Bang to the current epoch. At each time-step, we compute the alpha shapes of a randomly sampled subset of the particle distribution for α from 0.0 to $10.0h^{-1}\mathrm{Mpc}$, and their Betti numbers, from which we extrapolate the scale-dependent behavior of homology. The resulting curves of the Betti numbers for dimensions $p = 0, 1, 2$ are shown in Figure 10. The curves in the upper left frame show that the value of β_0 decreases for $\alpha < 1.2h^{-1}\mathrm{Mpc}$, while it increases for larger values of α. This is a manifestation of the hierarchical structure formation process. It reflects the progressively earlier merging of clumps at small scale and the progressively later merging of massive components at large scale.

The curves of β_1 and β_2 have a slightly different appearance. Both start as highly peaked distributions. Most prominent in the near homogeneous initial conditions, the peaks represent the imprint of a Poisson distribution with a similar particle density. In principle, one should remove this imprint via a persistence procedure; see Section IX. Here we keep to the unfiltered version. As the mass distribution evolves under the influence of gravity, the particles get more and more clustered. This leads to a decrease in the number of holes on small scales. As these merge into ever larger supercluster complexes, their formation goes along with the evacuation of ever larger tunnels and voids enclosed by the higher density structures in the web-like network. All three curves are marked by a steady shift toward higher values

of α. While we find a minor shift of the maximum of β_1, the shift is clear for β_2. It is interesting that a similar shift toward larger α values has has been predicted by Sheth & van de Weygaert [52] in the context of their dynamical model of the evolving void hierarchy. In their two-barrier excursion set formalism, this shift is a manifestation of the merging of smaller voids into ever larger ones, accompanied by the destruction of a large number of small voids inside gravitationally collapsing over-dense regions.

C. Alpha Tracks and Multiscale Homology

When assessing the corresponding evolution of alpha tracks in Figure 9, we find a systematic behavior that is similar to what we have seen for the Voronoi clustering models. As the evolution proceeds, we find fewer tunnels and voids at small scale, while their numbers are still substantial at values of α at which we have only very few remaining components.

However, as the mass distribution evolves, the shifts between the alpha tracks are relatively small compared to those seen in the Voronoi clustering models. The mass distribution retains its multiscale character. The hierarchical evolution of the multiscale mass distribution leads to a more or less self-similar mapping of the LCDM alpha tracks to higher α values. At each phase, we find dominant clusters and conspicuous filamentary features, although their scale shifts as the mass distribution advances from the mildly linear to the quasi-linear phase. This contrasts the evolution of the Voronoi clustering models, which is characterized by clear transitions between distinct topological patterns outlined by the mass distribution.

VII. PROBING DARK ENERGY

To assess the utility of our topological methods, we look at the possibility to use the homology of the Cosmic Web towards determining the nature of dark energy in the Universe. The parameters for what might now be called "the standard model for cosmology" have been established with remarkable precision.

However, there remains the great mystery of the nature of the so-called *dark energy*, which appears to make up 73% of the total cosmic energy. Dark energy, or the cosmological constant, has been dominating the dynamics of the Universe since the last 7 Gigayears and has pushed it into an accelerated expansion. The simplest model for the dark energy is Einstein's cosmological constant, Λ: it makes a time-independent contribution to the total energy density in the Friedman-Lemaitre equations. Such models are referred to as *LCDM models*. However, there are numerous, possibly more plausible models in which the dark energy evolves as a function of time. These models are generally described in terms of a time- or redshift-dependent function $w(z)$ that
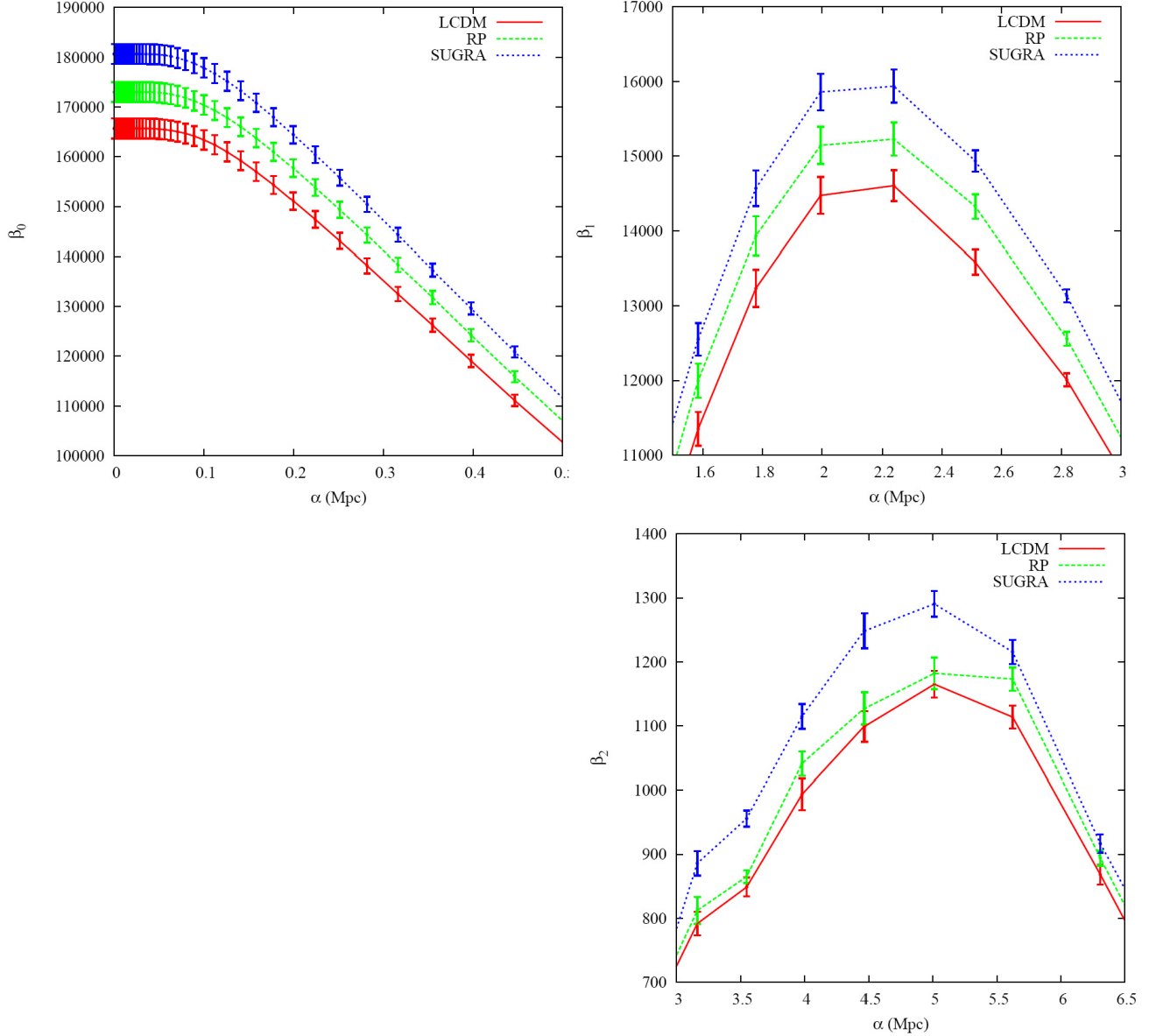
Figure 11. Homology and dark energy. The Betti number curves for three cosmological models with different dark energy models. The curves are determined from the distribution of collapsed dark matter halos in the corresponding scenarios. Red: the standard LCDM scenario. Green: the Ratra-Peebles quintessence model. Blue: the supergravity SUGRA model. From left to right: $\beta_0(\alpha)$, $\beta_1(\alpha)$, and $\beta_2(\alpha)$. The curves demonstrate that homology is capable of using the web-like matter distribution to discriminate between cosmological scenarios with different dark energy content.

describes the history of equation of state of the dark energy component, i.e. the relation between pressure and density:

$$p \quad = \quad w(z)\rho c^2. \qquad (18)$$

Different models for dark energy produce different functions $w(z)$, and a number of simple parametrizations can be found in the literature. The classical Einstein cosmological constant corresponds to $w(z) = -1$.

In principle, the Cosmic Web should be one of the strongest differentiators between the various dark energy models. As a result of the different growth rates at comparable cosmic epochs, we will see different structures in the matter distribution. To illustrate this idea, we run three N-body simulations with identical initial conditions but different dark energy equation of state: the *standard LCDM model* and two *quintessence models*. The latter assume that the Universe contains an evolving quintessence scalar field, whose energy content manifests itself as dark energy. The two quintessence models are the *Ratra-Peebles (RP) model* and the *SUGRA model*; see [53], [54], [55] for a detailed

description.

While the general appearance of the emerging Cosmic Web is structurally similar, there are visible differences in the development. On small to medium scales, the quintessence models are less clustered; see [25] for an illustration. As a result, the alpha shapes contain a larger number of small objects. In Figure 11, we see that this manifests itself in values of β_0 that are systematically higher for the quintessence RP and SUGRA models than for the LCDM model. The two right frames in Figure 11 concentrate on the medium scale range, $1.5h^{-1}\mathrm{Mpc} < \alpha < 3.0h^{-1}\mathrm{Mpc}$ for β_1 (top righthand panel) and $3.0h^{-1}\mathrm{Mpc} < \alpha < 6.5h^{-1}\mathrm{Mpc}$ for β_2 (bottom righthand panel). The error bars on the obtained curves are determined on the basis of the obtained values for independent samples. Particularly encouraging is the fact that homology is sensitive to the subtle differences we see in the pattern of the Cosmic Web. The size of the error bars show that it is feasible to find significant and measurable differences between the outcome of the different cosmologies.

VIII. BETTI NUMBERS OF GAUSSIAN RANDOM FIELDS

To interpret our results for the web-like cosmic matter distribution, we compare them with those for the initial condition out of which our Universe arose. The behavior of the Betti numbers in a Gaussian random field is a reference point for any further assessment of their behavior in the more complex environment of the Cosmic Web.

According to the current paradigm, structure in the Universe grew by gravitational instability out of tiny primordial density perturbations. The evidence provided by the temperature fluctuations in the cosmic microwave background [56], [57], [58], [59] suggests that the character of the perturbation field is that of a homogeneous and isotropic spatial Gaussian process. Such primordial Gaussian perturbations in the gravitational potential are a natural product of an early inflationary phase of our Universe.

Before proceeding to the Betti numbers of a Gaussian random field, we present the necessary nomenclature, focusing on the three-dimensional situation. For fundamentals on Gaussian random fields, we refer to the standard work by Adler and Taylor [60], and for their cosmological application to the seminal study by Bardeen, Bond, Kaiser, and Szalay [61]. Here we follow the notation from van de Weygaert and Bertschinger [62]. The first papers on the homology and persistence of Gaussian random fields are Adler et al. [63] and Park et al. [64]; see also [65]. Here, we discuss some of the main findings, and we refer to the mentioned papers for more detailed treatments. We alert the reader to the fact that the analysis in this section concerns itself with density fields, which is different from earlier sections in which we studied distance fields defined by particle distributions.

A. Gaussian Random Fields

A *random field*, f, on a spatial volume assigns a value, $f(\mathbf{x})$, to each location, $\mathbf{x}$, of that volume. The fields of interest, such as the primordial density or velocity field, are smooth and continuous. The stochastic properties of a random field are defined by its N-*point joint probabilities*, where N can be any arbitrary positive integer. To denote them, we write $\mathbf{X} = (\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_N)$ for a vector of N points and $\mathbf{f} = (f_1, f_2, \ldots, f_N)$ for a vector of N field values. The joint probability is

$$\mathrm{Prob}[f(\mathbf{x}_1) = f_1, \ldots, f(\mathbf{x}_N) = f_N] \;=\; \mathcal{P}_\mathbf{X}(\mathbf{f})\,\mathrm{d}\mathbf{f}\,, \tag{19}$$

which is the probability that the field f at the locations $\mathbf{x}_i$ has values in the range f_i to $f_i + \mathrm{d}f_i$, for each $1 \le i \le N$. Here, $\mathcal{P}_\mathbf{X}(\mathbf{f})$ is the probability density for the field realization vector $\mathbf{f}$ at the location vector $\mathbf{X} = (\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_N)$. For a *Gaussian random field*, the joint probabilities for $N = 1$ and $N = 2$ determine all others. Specifically, the probability density functions take the simple form

$$\mathcal{P}_\mathbf{X}(\mathbf{f}) \;=\; C \cdot \exp\left[-\mathbf{f}\mathsf{M}^{-1}\mathbf{f}^T/2\right]\,, \tag{20}$$

where $C = 1/[(2\pi)^N (\det \mathsf{M})]^{1/2}$ normalizes the expression, making sure that the integral of $\mathcal{P}_\mathbf{X}(\mathbf{f})$, over all $\mathbf{f} \in \mathbb{R}^N$, is equal to 1. Here, we assume that each 1-point distribution is Gaussian with zero mean. The matrix M^{-1} is the inverse of the $N \times N$ covariance matrix with entries

$$M_{ij} \;=\; \langle f(\mathbf{x}_i)f(\mathbf{x}_j)\rangle\,, \tag{21}$$

in which the angle bracket denotes the ensemble average of the product, over the 2-point probability density function. In effect, M is the generalization of the variance of a 1-point normal distribution, and we indeed have $\mathsf{M} = [\sigma_0^2]$ for the case $N = 1$.

Equation (20) shows that a Gaussian random is fully specified by the autocorrelation function, $\xi(r)$, which expresses the correlation between the density values at two points separated by a distance $r = |\mathbf{r}|$,

$$\xi(r) = \xi(|\mathbf{r}|) \equiv \langle \mathbf{f}(\mathbf{x})\mathbf{f}(\mathbf{x}+\mathbf{r})\rangle\,. \tag{22}$$

Here we use the *statistical cosmological principle*, which states that statistical properties of e.g. the cosmic density distribution in the Universe are uniform throughout the Universe. It means that the distribution functions and moments of fields are the same in each direction and at each location. The latter implies that ensemble averages depend only on one parameter, namely the distance between the points. In other words, the entries in the matrix are the values of the *autocorrelation function* for the distance between the points: $M_{ij} = \xi(r_{ij})$, with $r_{ij} = \|\mathbf{x}_i - \mathbf{x}_j\|$.

An impression of a typical Gaussian random field may be obtained from the three-dimensional realization shown in the left panel of Figure 12. The field is chosen using
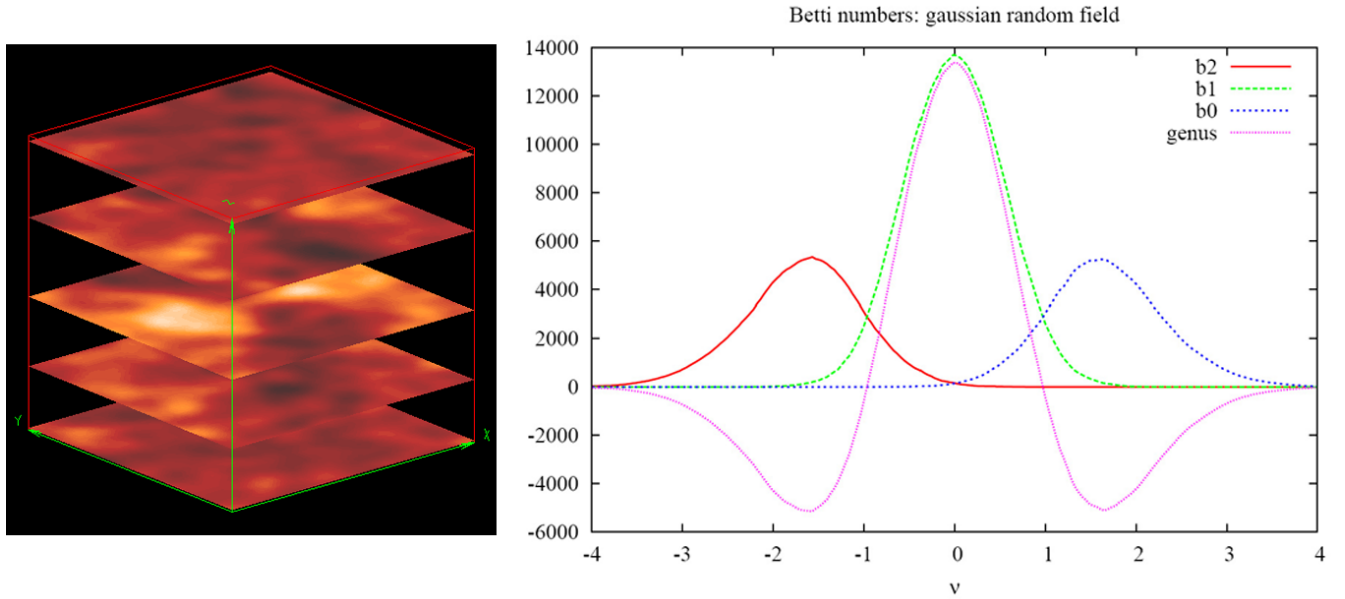
Figure 12. Left: five slices through a realization of a Gaussian random field, with the color reflecting the density, running from bright yellow (high) to dark red (low). It is an LCDM density field in a box of $100h^{-1}\mathrm{Mpc}$, Gaussian filtered on a scale of $5h^{-1}\mathrm{Mpc}$. Right: expected Betti numbers and expected *reduced genus* of the superlevel sets in a Gaussian random field, as functions of the density threshold, ν.

the power spectrum for the standard LCDM cosmology with some Gaussian filtering; see e.g. [61], [66]. The image illustrates that Gaussian random fields are symmetric, i.e. negative values are as likely as positive values.

B. Power Spectrum

A stochastic random density field is composed of a spectrum of density fluctuations, each of a different scale. The relative amplitudes of small-scale and large-scale fluctuations is of decisive influence on the outcome of the subsequent gravitational evolution of the density field and on the emerging patterns in the spatial density distribution.

To describe the multiscale composition of a density field, we write it as a sum of individual harmonic waves, i.e. in terms of its Fourier sum. Each of the waves is specified by its *wave vector* $\mathbf{k} = (k_x, k_y, k_z) \in \mathbb{R}^3$, describing the direction and spatial frequency of the wave. The latter is determined by the *magnitude* of the wave vector,

$$k = |\mathbf{k}| = \sqrt{k_x^2 + k_y^2 + k_z^2}, \qquad (23)$$

which is the inverse of the wavelength, $\lambda = 2\pi/k$. Subsequently, we write the field $f(\mathbf{x})$ as the Fourier integral,

$$f(\mathbf{x}) = \int \frac{d\mathbf{k}}{(2\pi)^3} \, \hat{f}(\mathbf{k}) \, e^{-i\mathbf{k}\cdot\mathbf{x}}, \qquad (24)$$

where $e^{i\varphi} = \cos\varphi + i\sin\varphi$, as usual, and $\mathbf{k} \cdot \mathbf{x}$ the inner product between the wave vector, $\mathbf{k}$, and position vector, $\mathbf{x}$. The Fourier components, $\hat{f}(\mathbf{k}) \in \mathbb{C}$, represent the contributions by the harmonic wave $\exp(i\mathbf{k}\cdot\mathbf{x})$ to the field $f(\mathbf{x})$. They are given by the inverse Fourier transform,

$$\hat{f}(\mathbf{k}) = \int d\mathbf{x} \, f(\mathbf{x}) \, e^{i\mathbf{k}\cdot\mathbf{x}}. \qquad (25)$$

Because the density field is always real, $f(\mathbf{x}) \in \mathbb{R}$, the Fourier components $\hat{f}(\mathbf{k}) \in \mathbb{C}$ obey the symmetry constraint

$$\hat{f}(\mathbf{k}) = \hat{f}^*(-\mathbf{k}). \qquad (26)$$

It identifies $\hat{f}(\mathbf{k})$ with the complex conjugate of the Fourier component of the wave vector $-\mathbf{k}$, i.e. of the wave with the same spatial frequency oriented in the opposite direction.

The *power spectrum* is formally defined as the mean square of the Fourier components, $\hat{f}(\mathbf{k})$, of the field. It can be computed from the Fourier components using the Dirac delta function, $\delta_D : \mathbb{R} \to \mathbb{R}$, which is the limit of the normal distribution, whose variance goes to zero. Most importantly, the integral of the Dirac delta function is assumed to be 1. Now, we have

$$\langle \hat{f}(\mathbf{k})\hat{f}(\mathbf{k}') \rangle = (2\pi)^{3/2} \, P(k) \, \delta_D(\mathbf{k} - \mathbf{k}'), \qquad (27)$$

and we can compute the power spectrum accordingly. We note that because of the statistical isotropy of the field, we have $P(\mathbf{k}) = P(\mathbf{k}')$, whenever $|\mathbf{k}| = |\mathbf{k}'|$. We can therefore introduce the 1-dimensional power spectrum, defined by $P(k) = P(\mathbf{k})$ for every $\mathbf{k}$, which we refer to by the same name, for convenience.

The *power spectrum* is the Fourier transform of the autocorrelation function $\xi(\mathbf{x})$, which is straightforward to infer from equations (22) and (27),

$$\xi(x) = \int \frac{d\mathbf{k}}{(2\pi)^3} \, P(k) e^{-i\mathbf{k}\cdot\mathbf{x}}, \qquad (28)$$

where $\mathbf{x}$ is a point in the volume with $|\mathbf{x}| = x$.

Because of the (statistical) isotropy of the field $f(\mathbf{x})$, it is straightforward to perform the angular part of the integral

over the 2-sphere, to yield the following integral expression for the autocorrelation function:

$$\xi(x) \;=\; \int \frac{k^2\,\mathrm{d}k}{2\pi^2}\,P(k)\,\frac{\sin kx}{kx}\,, \qquad (29)$$

with $k = |\mathbf{k}|$ and $x = \mathbf{x}.$.

From the above we find that the power spectrum is a complete characterization of a homogeneous and isotropic Gaussian random field. Within its cosmological context, the power spectrum encapsulates a wealth of information on the parameters and content of the Universe. Its measurement is therefore considered to be a central key for the understanding of the origin of the cosmos.

C. Fourier Decomposition

The unique nature of Gaussian fluctuations is particularly apparent when considering the stochastic distribution of the field in terms of its Fourier decomposition. The assumption of Gaussianity means that the stochastic distribution of the Fourier components

$$\begin{aligned} \hat{f}(\mathbf{k}) &= \hat{f}_{\mathrm{re}}(\mathbf{k}) + \mathrm{i}\hat{f}_{\mathrm{im}}(\mathbf{k}) \\ &= |\hat{f}(\mathbf{k})|\,e^{\mathrm{i}\,\theta(\mathbf{k})} \end{aligned} \qquad (30)$$

involves a random phase, $\theta(\mathbf{k})$, in the complex plane. This follows from the mutual independence of its real and imaginary parts, each of which is a Gaussian variable. While the phase, $\theta(\mathbf{k})$, has a uniform distribution over the interval $[0, 2\pi]$, the amplitude, $r = |\hat{f}(\mathbf{k})|$, has a Rayleigh distribution:

$$\mathcal{P}(r) \;=\; \frac{r}{s}\exp\left[-\frac{r^2}{2s}\right], \qquad (31)$$

where $s = P(k)$ is the power spectrum value at spatial frequency k.

Subsequently determining the probability of the field realization $f(\mathbf{x})$ in terms of the probability density function of its Fourier decomposition, one finds the interesting result that it is the product of the individual probability distributions of each of the individual Fourier components $\hat{f}(\mathbf{k})$, each Gaussian distributed with zero mean and variance $\sigma^2(k) = P(k)$. It is most straightforward to appreciate this by assessing the probability of the entire field $f(\mathbf{x})$.

To infer the probability $\mathcal{P}[f]$ of $f(\mathbf{x})$, we take the N-point joint probabilities for the limit of $N \to \infty$ with uniform spatial sampling, the summations appearing in equation (20) may be turned into integrals. The resulting expression for the infinitesimal probability $\mathcal{P}[f]$ with measure $\mathcal{D}[f]$,

$$\mathcal{P}[f] \;=\; \mathrm{e}^{-S[f]}\,\mathcal{D}[f]\,, \qquad (32)$$

involves the probability density of the field, $\exp\left(-S[f]\right)$. The square brackets in $\mathcal{P}[f]$ and $S[f]$ indicate that these are functionals, i.e., they map the complete function $f(\mathbf{x})$ to one number. The probabilit density is similar to the quantum-mechanical partition function in path integral form, where

S is the action functional (see [62]). For a Gaussian random field the expression for the action S may be inferred from equation (20),

$$S[f] \;=\; \frac{1}{2}\int \mathrm{d}\mathbf{x}_1\int \mathrm{d}\mathbf{x}_2\, f(\mathbf{x}_1)K(\mathbf{x}_1 - \mathbf{x}_2)f(\mathbf{x}_2)\,, \qquad (33)$$

where K is the functional inverse of the correlation function ξ,

$$\int \mathrm{d}\mathbf{x}\, K(\mathbf{x}_1 - \mathbf{x})\xi(\mathbf{x} - \mathbf{x}_2) \;=\; \delta_D(\mathbf{x}_1 - \mathbf{x}_2)\,, \qquad (34)$$

and δ_{D} the Dirac delta function. By transforming this expression for the action $S[f]$ to Fourier space, one finds the integral expression (see [62], appendix B),

$$S[f] \;=\; \int \frac{\mathrm{d}\mathbf{k}}{(2\pi)^3}\,\frac{|\hat{f}(\mathbf{k})|^2}{2P(k)}\,. \qquad (35)$$

This immediately demonstrates that a Gaussian field has the unique property of its Fourier components being mutually independent. It has the practical virtue of considerably simplifying the construction of Gaussian fields by sampling its Fourier components.

D. Betti Numbers

To determine the Betti numbers of a Gaussian field, we consider the density values sampled on a regular grid and assess the topology of the superlevel sets. We adopt a 64^3 grid, and sample the field inside a box of size $100h^{-1}\mathrm{Mpc}$. The density values are smoothed on a filter scale $2.0h^{-1}\mathrm{Mpc}$. Subsequently, we determine the Betti numbers of superlevel sets, the agglomerate of regions consisting of the voxels whose density is in excess of $\nu\sigma(R_f)$, where $\sigma^2(R_f)$ is the variance of the density field on the filter scale R_f.

Having outlined the superlevel sets, parameterized by the threshold value ν, we subsequently determine their Betti numbers. To this end, we use the code of Kerber [67], which determines the cycles for the superlevel sets sampled on a grid, and subsequently produces the corresponding homology groups and their ranks, i.e. the Betti numbers $\beta_p(\nu)$. By following this procedure over a threshold level range $\nu \in [-5.0, 5.0]$, we can evaluate the systematic behavior of the Betti numbers as function of ν. Note that a different algorithm was used in [64]. The fact that the results are similar is reassuring.

The right panel of Figure 12 depicts the curves for $\beta_0(\nu)$, $\beta_1(\nu)$, and $\beta_2(\nu)$. The first major impression is the dominant peak of $\beta_1(\nu)$ at $\nu = 0$ (green dashed curve). This relates to the fact that tunnels prevail the topological structure of the outlined regions, which in a Gaussian field are known to possess an intricate sponge-like topology [9]. The region of points with density value at most $\nu = 0$ consists of very few massive structures that percolate through the volume and, by symmetry, the same is true for the region of points with density value at least $\nu = 0$. It is not hard to imagine that this goes along with a large number of tunnels.

Going away from $\nu = 0$ in either direction, we see that β_1 quickly drops to a negligible level. This marks the transition towards a topology marked by individual objects, either clumps for positive ν or voids for negative ν, and this goes along with the disappearance of tunnels. Proceeding to higher levels, we find a quick rise of β_0 (blue dotted curve), the number of over-dense regions in the volume. This is a manifestation of the breaking up into individual high-density regions of the percolating volume at $\nu = 0$. The number of components reaches its maximum for ν approximately equal to $\sqrt{3}$. Beyond that value, the number of individual clumps decreases rapidly, as a result of the decreasing probability of regions having a density that high. It is interesting that the behavior of β_2 (red solid curve) is an almost perfect reflection of the behavior of β_0 through the zero density level. It documents the growing prominence of voids for negative and shrinking density threshold. Here we find a maximum value for ν approximately equal to $-\sqrt{3}$.

We note that we would get perfect symmetry if we replaced the number of components by the number of gaps between them, which is $\beta_0 - 1$. As mentioned, β_2 goes to zero rapidly as ν increases beyond 0, and $\beta_0 - 1$ goes to zero rapidly as ν decreases below 0. However, at $\nu = 0$, both are small but clearly positive. The expected number of components at $\nu = 0$ is thus small but larger than 1, with an expected number of voids that is precisely one less than for the components.

E. Gaussian Fields Versus the Cosmic Web

It is interesting to compare these results with the homology we find in the web-like configuration of the evolved Universe. Besides the disappearance of symmetry between high-density and low-density regions, reflected in a substantially different behavior of β_0 and β_2, there are a few additional differences as well as similarities.

In evolving LCDM density fields, we find that β_1 is almost always in excess of β_2: the number of tunnels in the Cosmic Web tends to be several factors higher than the number of enclosed voids; see Figure 10. However, the advanced nonlinear mass distribution is marked by a substantially higher number of individual components than found in the Gaussian field. This partially reflects the difference of an alpha shape based analysis and one based on level sets defined for a fixed Gaussian filter radius. In the alpha shape analysis, the substructure of an intrinsically multiscale mass distribution emerging through an hierarchical is not lost. At small values of α one finds the small clumps that in a Gaussian filtered field have been removed.

F. Genus and Homology

It is interesting to relate the results on the Betti numbers depending on a threshold value with the analytically known

expression for the expected *reduced genus*[11] in a Gaussian random field [9], [10]. This expression is

$$g(\nu) = -\frac{1}{8\pi^2}\left(\frac{\langle k^2 \rangle}{3}\right)^{3/2}(1 - \nu^2)e^{-\nu^2/2}, \quad (36)$$

where $\langle k^2 \rangle = \langle |\nabla f|^2 \rangle / \langle f^2 \rangle$. Other than its amplitude, the shape does not depend on the power spectrum but only on whether or not the field is Gaussian. Recall from (10) that the genus of an isodensity surface is directly related to the Betti numbers of the enclosed superlevel sets, via the alternating sum

$$g(\nu) = -\beta_0(\nu) + \beta_1(\nu) - \beta_2(\nu). \quad (37)$$

We may therefore compare the outcome of our superlevel set study to the expected distribution for the genus in (36). In Figure 12, the magenta dotted dashed line is the genus $g(\nu)$ computed from the alternating sum (37). It closely matches the predicted relation (36); see [64]. It is most reassuring that the peaks of β_0 and β_2 are reached at the threshold value where, according to (36), $g(\nu)$ has its two minima, at $\nu = -\sqrt{3}$ and $\nu = \sqrt{3}$.

Having established the cosmologically crucial Gaussian basis of the Betti number analysis, a few more interesting findings follow from our analysis in [64]. While the shape of the genus function is independent of the power spectrum, it turns out that the Betti numbers do reflect the power spectrum. In particular, for small absolute values of ν, there are substantial differences between the Betti numbers in Gaussian fields with different power spectra. This makes them potentially strong discriminators between cosmological scenarios. On the other hand, the differences are perfectly symmetric between $\beta_0 - 1$ and β_2, and β_1 is symmetric itself. The Betti numbers may therefore be useful for tracing non-Gaussianities in the primordial Universe, a major point of interest in current cosmological studies.

G. Gaussian Betti Correlations

Finally, [64] also assess the correlations in Gaussian fields between the various Betti numbers. They find that the Betti numbers are not independent of one another. For example, β_0 near its maximum ($\nu = 1.7$) is positively correlated with β_1 at low threshold levels, while there are other levels where β_0 and β_1 are anti-correlated.

The important implication is that we need to take into account the mutual dependence of the Betti numbers. This will be largely dependent on the nature of the density field. On the other hand, in general the correlation is not perfect. In other words, the Betti numbers contains complementary

---

[11]In section II-A we remarked that the genus g in cosmological studies is slightly differently defined than the usual definition of the genus G (eqn. 1). The genus, g, in these studies has been defined as the number of holes minus the number of connected regions: $g = G - c$. In this review we distinguish between these definitions by referring to g as the *reduced genus*.
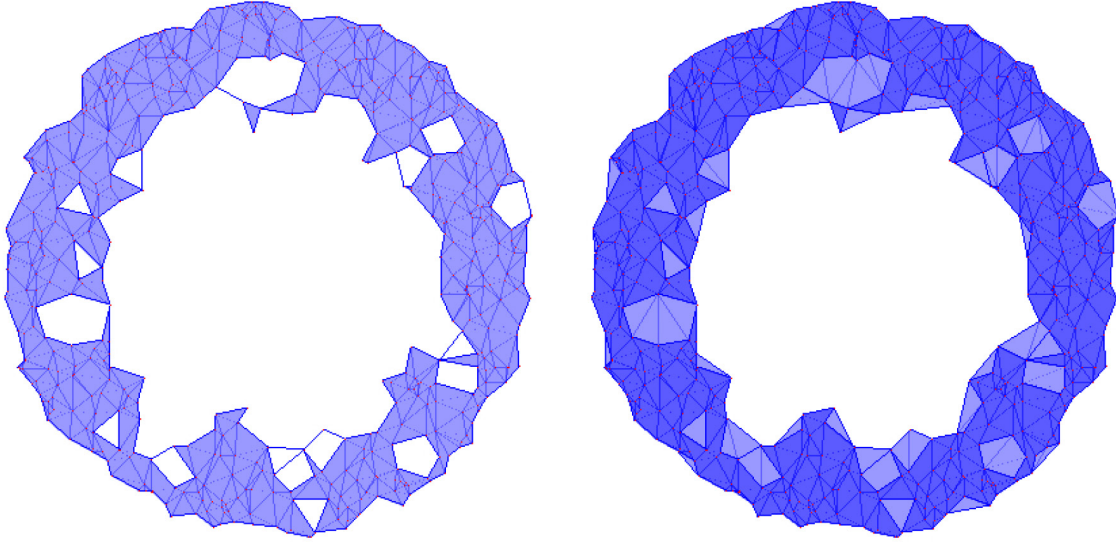
Figure 13. Illustrating the idea of persistence. The alpha shapes of a set of points randomly chosen in a 2-dimensional annulus. Left: for a small value of α, the alpha shape has 18 holes. Right: for a somewhat larger value of α, the alpha shape has only two holes. One of these holes is new, while the other already existed on the left, albeit in somewhat different form. We say the latter hole *persists* over the entire interval delimited by the two alpha shapes.

information on the topological structure of e.g. the cosmic mass distribution.

IX. PERSISTENCE

The one outstanding issue we have not yet addressed systematically is the hierarchical substructure of the Cosmic Web data. In standard practice, the multiscale nature of the mass distribution tends to be investigated by means of user-imposed filtering. Persistence rationalizes this approach by considering the range of filters at once, and by making meaningful comparisons. At the same time, it deepens the approach by combining it with topological measurements: the homology groups and their ranks.

Persistence entails the conceptual framework and language for separating scales of a spatial structure or density distribution. It was introduced by Edelsbrunner, Letscher and Zomorodian [68]; see also the recent text on computational topology [43] of which persistent homology forms the core. By separating the scales in a mass distribution, one may analyze and map the topological hierarchy of the cosmos. Within this context, substructures can be separated by determining the range of scales over which they exist. Here, we find a close link to Morse theory, in which the crucial new idea contributed by persistence is the pairing of all critical points depending on a global topological criterion. A recent discussions of persistence in the cosmological context can be found in the work of Sousbie et al. [29], [69].

A. Persistence of Alpha Shapes

Alpha shapes provide the perfect context to illustrate the essential idea of persistence, which is to follow components,

tunnels, and voids over the entire range of the parameter, α. The 2-dimensional illustration in Figure 13 shows how the small scale holes in the left frame disappear when we increase the parameter in the right frame. This is not to say that the value of α chosen on the right is perfect for the data set; indeed, we see a new, larger hole appear that did not yet exist on the left. The key point is that each feature is *born* at some value of α, and *dies* at another value of α. The interval between birth and death sets the position of the feature within the structural hierarchy. In particular, the length of the interval, that is, the absolute difference between the values of α at the birth and at the death, is called the *persistence* of the feature.

The full implementation of persistence in our topological study will be addressed in future publications; e.g. [65]. Here, we briefly address its effect on the Betti number analysis. We assess the matter distribution for three cosmological scenarios: LCDM, and RP and SUGRA quintessence. Each is used in a many-body simulation, generating a discrete point distribution in 3-dimensional space that reflects the intricate multi-scale organization of matter. For each dataset, we compute the entire range of alpha shapes, encoded for efficiency reasons in the Delaunay triangulation of the points, and we follow each cycle through the sequence of alpha shapes, recording when it is born and when it dies. Separating the results for different homological dimensions ($p = 0$ for components, $p = 1$ for tunnels, and $p = 2$ for voids), we show the statistics for $p = 1$ in Figure 14. On the left, we see the 1-dimensional persistence diagrams of the three datasets. Each dot represents a 1-cycle in the alpha
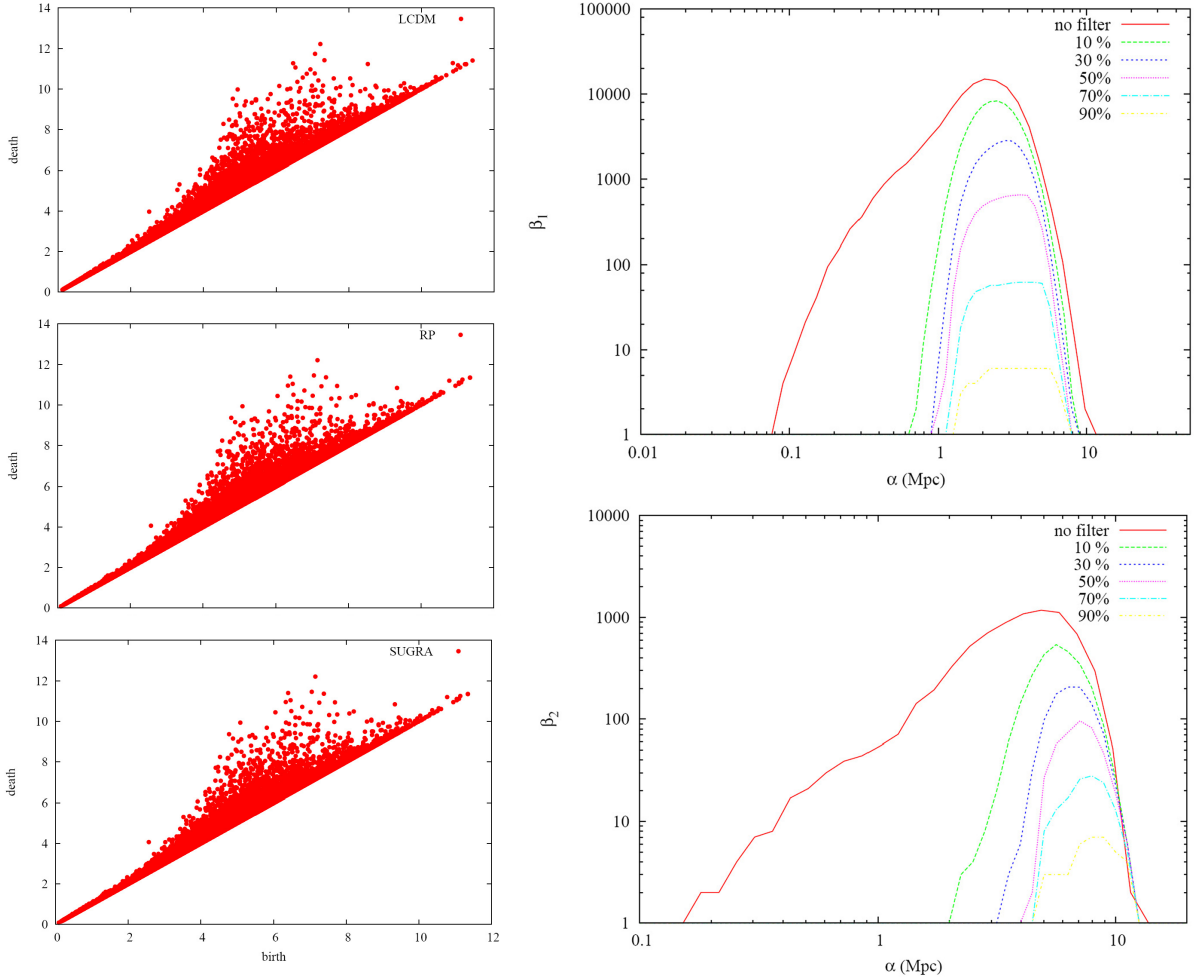
Figure 14. Persistent homology for the Cosmic Web simulations. Left: the 1-dimensional persistence diagrams of the LCDM simulation and of two quintessence simulations representing the RP and SUGRA dark energy models. Right: the curves describing the evolution of the first and second Betti numbers for growing value of α. In addition to giving the Betti numbers, at every value of α, we show how many of the counted features belong to the top 90%, 70%, 50%, 30%, 10% most persistent features in the entire population.

shape filtration. Its horizontal coordinate gives the value of α at which the 1-cycle is born, which happens when its last edge is added to the alpha shape. The vertical coordinate of the dot gives the value of α at which the 1-cycle dies, which happens when the last triangle completes a membrane filling the tunnel formed by the cycle.

In all three diagrams, we see a substantial number of 1-cycles that die shortly after they are born. They correspond to dots close to and right above the diagonal line in the diagram. This is typical for natural data, in which we see features appear and disappear in rapid progression at all times. This "topological noise" creates a deceivingly complicated picture, obscuring the cleaner picture of more persistent features. In our datasets, there are also a good number of 1-cycles that remain for a while after they are born. They correspond to dots that are further up in the persistence diagram, with the vertical distance to the

diagonal representing the measured persistence.

B. Persistence and Scale

The promise of persistence is that it gives access to data on different scale levels. This is particularly important for the Cosmic Web, where we observe structure on almost every scale. The goal would be to separate scales, or quantify them in such a way that relations between scales become explicit.

To probe this aspect of persistence, we use a *persistence parameter*, λ, that restricts our attention to features of persistence at least λ. In our experiment, we adjusted the parameter so that a specified percentage of the features (dots) have persistence at least λ. All dots with persistence less than λ are removed. The effect on the Betti numbers of the LCDM simulation data is illustrated in the two right frames of Figure 14. The top frame shows six evolutions of β_1 for increasing values of α: for 100%, 90%, 70%, 50%, 30%,

and 10% of the most persistent 1-cycles. The bottom frame shows the same information for β_2.

The features with small persistence can be considered noise. They may be reflections of numerical inaccuracies in computations, artefacts of the data representation, measuring errors, or genuine features that are too small for us to distinguish them from noise. In practical astrophysical circumstances, datasets are indeed beset by a variety of artefacts. As a result, the alpha shapes of discrete point distributions, such as the galaxies in our local cosmic environment, will reflect the noise. The induced irregularities in the number of components, tunnels, and voids do not represent any real topological structure but nonetheless influence the Betti numbers we collect. Persistence diagrams offer the real opportunity to remove the noise without introducing side-effects, and they may help in defining the usually ill-specified noise distribution.

X. CONCLUSIONS AND PROSPECTS

In this overview we have discussed and presented an analysis of the homology of cosmic mass distributions. We have argued that Betti numbers of density level sets and of the corresponding isodensity surfaces provide us with a far more complete description of the topology of web-like patterns in the megaparsec Universe than the well-known genus analysis or the topological and geometric instruments of Minkowski functionals. The genus of isodensity surfaces and the Euler characteristic, the main topological Minkowski functional, are directly related to Betti numbers and may be considered as lower-dimensional projections of the homological information content in terms of Betti numbers.

We have established the promise of alpha shapes for measuring the homology of the megaparsec galaxy distribution in the Cosmic Web. Alpha shapes are well-defined subsets of Delaunay tessellations, selected following a strictly defined scale parameter α. These simplicial complexes constitute a filtration of the Delaunay tessellations, and are homotopy equivalent to the distance function field defined by the point distribution. Alpha shape analysis has the great advantage of being self-consistent and natural, involving shapes and surfaces that are entirely determined by the point distribution, independent of any artificial filtering.

By studying the Betti numbers and several Minkowski functionals of a set of heuristic Voronoi clustering models, as well as dark energy cosmological scenarios (LCDM, and RP and SUGRA quintessence), we have illustrated the potential for exploiting the cosmological information contained in the Cosmic Web. We have analyzed the evolution of Betti numbers in an LCDM simulation, and compared the discriminative power of Betti curves for distinguishing between different dark energy models. In addition, we have addressed the significance of Betti numbers in the case of Gaussian random fields, as these primordial density fields represent the reference point for any further analysis of the subsequent cosmic structure evolution.

Related Work
The mathematical fundamentals of our project, including a full persistence analysis, will be extensively outlined and defined in the upcoming study of Pranav, Edelsbrunner et al. [65]. In addition to an evaluation on the basis of the simplicial alpha complexes, it also discusses and compares the homology analysis on the basis of isodensity surfaces, superlevel and sublevel sets, and their mutual relationship. This includes an evaluation of the results obtained on the basis of a piecewise linear density field reconstructions on Delaunay tessellations, following the DTFE formalism [26], [27], [28].

In related work, [29] and [69] recently published an impressive study on the topological analysis of the Cosmic Web within the context of the skeleton formalism [70], [71], [72]. Their study used the DTFE density field reconstructions and incorporated the complexities of a full persistence analysis to trace filamentary features in the Cosmic Web.

In two accompanying letters we will address cosmological applications of homology analysis. In [25] the ability of Betti numbers to discriminate between different dark energy cosmological scenarios will be demonstrated. Crucial for furthering our understanding and appreciation of the significance of measured Betti numbers in a cosmological context, is to know their values in Gaussian random density fields. With cosmic structure emerging from a primordial Gaussian field, the homology of these fields forms a crucial reference point. In the study by Park et al. [64] we have analyzed this aspect in more detail.

REFERENCES

[1] J. Bond, L. Kofman, and D. Pogosyan, "How filaments are woven into the cosmic web," *Nature*, vol. 380, pp. 603–606, 1996.

[2] R. van de Weygaert and J. R. Bond, "Observations and Morphology 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, & D. Hughes, Ed., vol. 740, 2008, pp. 409–+.

[3] M. Colless and 2dF consortium, "The 2df galaxy redshift survey: Final data release," pp. 1–32, 2003, astroph/0306581.

[4] M. Tegmark and et al., "The three-dimensional power spectrum of galaxies from the sloan digital sky survey," *Astrophys. J.*, vol. 606, pp. 702–740, 2004.

[5] J. Huchra and et al., "The 2mass redshift survey and low galactic latitude large-scale structure," in *Nearby Large-Scale Structures and the Zone of Avoidance*, ser. ASP Conf. Ser. Vol., A. Fairall and P. Woudt, Eds., vol. 239. Astron. Soc. Pacific, San Francisco, 2005, pp. 135–146.

[6] P. Peebles, *The Large Scale Structure of the Universe*. Princeton Univ. Press, 1980.

[7] 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, & D. Hughes, Ed., vol. 740, 2008, pp. 335–+.

[8] V. Springel and et al., "Simulations of the formation, evolution and clustering of galaxies and quasars," *Nature*, vol. 435, pp. 629–636, 2005.

[9] J. Gott, M. Dickinson, and A. Melott, "The sponge-like topology of large-scale structure in the universe," *Astrophys. J.*, vol. 306, pp. 341–357, 1986.

[10] A. J. S. Hamilton, J. R. Gott, III, and D. Weinberg, "The topology of the large-scale structure of the universe," *Astrophys.J.*, vol. 309, pp. 1–12, Oct. 1986.

[11] C. Park and Y.-R. Kim, "Large-scale Structure of the Universe as a Cosmic Standard Ruler," *Astrophys. J. Lett.*, vol. 715, pp. L185–L188, Jun. 2010.

[12] J. R. Gott, III, J. Miller, T. X. Thuan, S. E. Schneider, D. H. Weinberg, C. Gammie, K. Polk, M. Vogeley, S. Jeffrey, S. P. Bhavsar, A. L. Melott, R. Giovanelli, M. P. Hayes, R. B. Tully, and A. J. S. Hamilton, "The topology of large-scale structure. III - Analysis of observations," *Astrophys. J.*, vol. 340, pp. 625–646, May 1989.

[13] C. Park, J. R. Gott, III, A. L. Melott, and I. D. Karachentsev, "The topology of large-scale structure. VI - Slices of the universe," *Astrophys. J.*, vol. 387, pp. 1–8, Mar. 1992.

[14] B. Moore, C. S. Frenk, D. H. Weinberg, W. Saunders, A. Lawrence, R. S. Ellis, N. Kaiser, G. Efstathiou, and M. Rowan-Robinson, "The topology of the QDOT IRAS redshift survey," *Mon. Not. R. Astron. Soc.*, vol. 256, pp. 477–499, Jun. 1992.

[15] M. S. Vogeley, C. Park, M. J. Geller, J. P. Huchra, and J. R. Gott, III, "Topological analysis of the CfA redshift survey," *Astrophys. J.*, vol. 420, pp. 525–544, Jan. 1994.

[16] C. Park, J. Kim, and J. R. Gott, III, "Effects of Gravitational Evolution, Biasing, and Redshift Space Distortion on Topology," *Astrophys. J.*, vol. 633, pp. 1–10, Nov. 2005.

[17] J. R. Gott, Y.-Y. Choi, C. Park, and J. Kim, "Three-Dimensional Genus Topology of Luminous Red Galaxies," *Astrophys. J. Lett.*, vol. 695, pp. L45–L48, Apr. 2009.

[18] Y. Zhang, V. Springel, and X. Yang, "Genus Statistics Using the Delaunay Tessellation Field Estimation Method. I. Tests with the Millennium Simulation and the SDSS DR7," *Astrophys. J.*, vol. 722, pp. 812–824, Oct. 2010.

[19] Y.-Y. Choi, C. Park, J. Kim, J. R. Gott, D. H. Weinberg, M. S. Vogeley, S. S. Kim, and for the SDSS Collaboration, "Galaxy Clustering Topology in the Sloan Digital Sky Survey Main Galaxy Sample: A Test for Galaxy Formation Models," *Astrophys. J. Suppl.*, vol. 190, pp. 181–202, Sep. 2010.

[20] K. Mecke, T. Buchert, and H. Wagner, "Robust morphological measures for large-scale structure in the universe," *Astron. Astrophys.*, vol. 288, pp. 697–704, 1994.

[21] J. Schmalzing and T. Buchert, "Beyond Genus Statistics: A Unifying Approach to the Morphology of Cosmic Structure," *Astrophys. J. Lett.*, vol. 482, pp. L1+, Jun. 1997.

[22] J. Schmalzing, T. Buchert, A. Melott, V. Sahni, B. Sathyaprakash, and S. Shandarin, "Disentangling the cosmic web. i. morphology of isodensity contours," *Astrophys. J.*, vol. 526, pp. 568–578, 1999.

[23] V. Sahni, B. Sathyprakash, and S. Shandarin, "Shapefinders: A new shape diagnostic for large-scale structure," *Astrophys. J.*, vol. 507, pp. L109–L112, 1998.

[24] M. A. A. Calvo, S. F. Shandarin, and A. Szalay., "Geometry of the cosmic web: Minkowski functionals from the delaunay tessellation," *International Symposium on Voronoi Diagrams in Science and Engineering*, vol. 0, pp. 235–243, 2010.

[25] R. van de Weygaert, P. Pranav, B. Jones, G. Vegter, P. Bos, C. Park, and W. Hellwing, "Probing dark energy with betti-analysis of simulations," *Astrophys. J. Lett.*, 2011, to be subm.

[26] W. E. Schaap and R. van de Weygaert, "Continuous fields and discrete samples: reconstruction through Delaunay tessellations," *Astron. Astrophys.*, vol. 363, pp. L29–L32, Nov. 2000.

[27] 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, & M.-J. Pons-Bordería, Ed., vol. 665, 2009, pp. 291–413.

[28] M. C. Cautun and R. van de Weygaert, "The DTFE public software - The Delaunay Tessellation Field Estimator code," *ArXiv e-prints*, May 2011.

[29] T. Sousbie, "The persistent cosmic web and its filamentary structure - I. Theory and implementation," *Mon. Not. R. Astron. Soc.*, pp. 511–+, Apr. 2011.

[30] H. Edelsbrunner, D. Kirkpatrick, and R. Seidel, "On the shape of a set of points in the plane," *IEEE Trans. Inform. Theory*, vol. 29, pp. 551–559, 1983.

[31] T. Dey, H. Edelsbrunner, and S. Guha, "Computational topology," in *Advances in Discrete and Computational Geometry*, B. Chazelle, J. Goodman, and R. Pollack, Eds. American Mathematical Society, 1999, pp. 109–143.

[32] G. Vegter, "Computational topology," in *Handbook of Discrete and Computational Geometry, 2nd edition*, J. E. Goodman and J. O'Rourke, Eds. Boca Raton, FL: CRC Press LLC, 2004, ch. 32, pp. 719–742.

[33] A. Zomorodian, *Topology for Computing*, ser. Cambr. Mon. Appl. Comp. Math. Cambr. Univ. Press, 2005.

[34] R. van de Weygaert, E. Platen, G. Vegter, B. Eldering, and N. Kruithof, "Alpha shape topology of the cosmic web," *International Symposium on Voronoi Diagrams in Science and Engineering*, vol. 0, pp. 224–234, 2010.

[35] F. Hoyle, M. S. Vogeley, J. R. Gott, III, M. Blanton, M. Tegmark, D. H. Weinberg, N. Bahcall, J. Brinkmann, and D. York, "Two-dimensional Topology of the Sloan Digital Sky Survey," *Astrophys. J.*, vol. 580, pp. 663–671, Dec. 2002.

[36] H. Edelsbrunner, "Alpha shapes - a survey," in *Tessellations in the Sciences; Virtues, Techniques and Applications of Geometric Tilings*, J. R. . V. I. R. van de Weygaert, G. Vegter, Ed. Springer Verlag, 2010.

[37] E. Muecke, *Shapes and Implementations in three-dimensional geometry*, 1993, phD thesis, Univ. Illinois Urbana-Champaign.

[38] H. Edelsbrunner and E. Muecke, "Three-dimensional alpha shapes," *ACM Trans. Graphics*, vol. 13, pp. 43–72, 1994.

[39] H. Edelsbrunner, M. Facello, and J. Liang, "On the definition and the construction of pockets in macromolecules," *Discrete Appl. Math.*, vol. 88, pp. 83–102, 1998.

[40] J. Liang, H. Edelsbrunner, P. Fu, P. Sudhakar, and S. Subramaniam, "Analytical shape computation of macromolecules: I. molecular area and volume through alpha shape," *Proteins: Structure, Function, and Genetics*, vol. 33, pp. 1–17, 1998.

[41] ——, "Analytical shape computation of macromolecules: Ii. inaccessible cavities in proteins," *Proteins: Structure, Function, and Genetics*, vol. 33, pp. 18–29, 1998.

[42] J. Liang, C. Woodward, and H. Edelsbrunner, "Anatomy of protein pockets and cavities: measurement of binding site geometry and implications for ligand design," *Protein Science*, vol. 7, pp. 1884–1897, 1998.

[43] H. Edelsbrunner and J. Harer, *Computational Topology, An Introduction*. American Mathematical Society, 2010.

[44] C. J. A. Delfinado and H. Edelsbrunner, "An incremental algorithm for betti numbers of simplicial complexes," in *Symposium on Computational Geometry*, 1993, pp. 232–239.

[45] B. Eldering, *Topology of Galaxy Models*, 2006, mSc thesis, Univ. Groningen.

[46] G. Kauffmann, J. M. Colberg, A. Diaferio, and S. D. M. White, "Clustering of galaxies in a hierarchical universe - I. Methods and results at z=0," *Mon. Not. R. Astron. Soc.*, vol. 303, pp. 188–206, Feb. 1999.

[47] R. van de Weygaert and V. Icke, "Fragmenting the universe. II - Voronoi vertices as Abell clusters," *Astron. Astrophys.*, vol. 213, pp. 1–9, Apr. 1989.

[48] R. van de Weygaert, *Voids and the geometry of large scale structure*, 1991, phD thesis, University of Leiden.

[49] R. v. d. Weygaert, "Voronoi tessellations and the cosmic web: Spatial patterns and clustering across the universe," in *ISVD '07: Proceedings of the 4th International Symposium on Voronoi Diagrams in Science and Engineering*. Washington, DC, USA: IEEE Computer Society, 2007, pp. 230–239.

[50] V. Icke, "Voids and filaments," *Mon. Not. R. Astron. Soc.*, vol. 206, pp. 1P–3P, Jan. 1984.

[51] G. Vegter, R. van de Weygaert, E. Platen, N. Kruithof, and B. Eldering, "Alpha shapes and the topology of cosmic large scale structure," *Mon. Not. R. Astron. Soc.*, 2010, in prep.

[52] R. K. Sheth and R. van de Weygaert, "A hierarchy of voids: much ado about nothing," *Mon. Not. R. Astron. Soc.*, vol. 350, pp. 517–538, May 2004.

[53] B. Ratra and P. J. E. Peebles, "Cosmological consequences of a rolling homogeneous scalar field," *Phys. Rev. D.*, vol. 37, pp. 3406–3427, Jun. 1988.

[54] C. de Boni, K. Dolag, S. Ettori, L. Moscardini, V. Pettorino, and C. Baccigalupi, "Hydrodynamical simulations of galaxy clusters in dark energy cosmologies," *Mon. Not. R. Astron. Soc.*, vol. 414, p. 780, 2011.

[55] P. B. C., R. van de Weygaert, K. Dolag, and V. Pettorino, "Void shapes as probes of the nature of dark energy," *Mon. Not. R. Astron. Soc.*, 2011.

[56] G. F. Smoot, C. L. Bennett, A. Kogut, E. L. Wright, J. Aymon, N. W. Boggess, E. S. Cheng, G. de Amici, S. Gulkis, M. G. Hauser, G. Hinshaw, P. D. Jackson, M. Janssen, E. Kaita, T. Kelsall, P. Keegstra, C. Lineweaver, K. Loewenstein, P. Lubin, J. Mather, S. S. Meyer, S. H. Moseley, T. Murdock, L. Rokke, R. F. Silverberg, L. Tenorio, R. Weiss, and D. T. Wilkinson, "Structure in the COBE differential microwave radiometer first-year maps," *Astrophys.J.Lett.*, vol. 396, pp. L1–L5, Sep. 1992.

[57] C. L. Bennett, M. Halpern, G. Hinshaw, N. Jarosik, A. Kogut, M. Limon, S. S. Meyer, L. Page, D. N. Spergel, G. S. Tucker, E. Wollack, E. L. Wright, C. Barnes, M. R. Greason, R. S. Hill, E. Komatsu, M. R. Nolta, N. Odegard, H. V. Peiris, L. Verde, and J. L. Weiland, "First-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Preliminary Maps and Basic Results," *Astrophys. J. Suppl.*, vol. 148, pp. 1–27, Sep. 2003.

[58] D. N. Spergel, R. Bean, O. Doré, M. R. Nolta, C. L. Bennett, J. Dunkley, G. Hinshaw, N. Jarosik, E. Komatsu, L. Page, H. V. Peiris, L. Verde, M. Halpern, R. S. Hill, A. Kogut, M. Limon, S. S. Meyer, N. Odegard, G. S. Tucker, J. L. Weiland, E. Wollack, and E. L. Wright, "Three-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Implications for Cosmology," *Astrophys.J.Suppl.*, vol. 170, pp. 377–408, Jun. 2007.

[59] E. Komatsu, K. M. Smith, J. Dunkley *et al.*, "Seven-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Interpretation," *eprint arXiv:1001.4538*, January 2010.

[60] R. J. Adler and J. E. Taylor, "Random fields and geometry," 2007.

[61] J. M. Bardeen, J. R. Bond, N. Kaiser, and A. S. Szalay, "The statistics of peaks of Gaussian random fields," *Astrophys.J.*, vol. 304, pp. 15–61, May 1986.

[62] R. van de Weygaert and E. Bertschinger, "Peak and gravity constraints in Gaussian primordial density fields: An application of the Hoffman-Ribak method," *Mon. Not. R. Astron. Soc.*, vol. 281, pp. 84–+, Jul. 1996.

[63] R. J. Adler, O. Bobrowski, M. S. Borman, E. Subag, and S. Weinberger, "Persistent homology for random fields and complexes," *Collections*, vol. 0, no. 0000, pp. 1–6, 2010. [Online]. Available: http://arxiv.org/abs/1003.1001

[64] C. Park, P. Chingangbam, R. van de Weygaert, G. Vegter, I. Kim, J. Hidding, W. Hellwing, and P. Pranav, "Betti numbers of gaussian random fields," *Astrophys. J.*, 2011, to be subm.

[65] P. Pranav, H. Edelsbrunner, R. van de Weygaert, and G. Vegter, "on the alpha and betti of the universe: Multiscale persistence of the cosmic web," *Mon. Not. R. Astron. Soc.*, 2011, to be subm.

[66] D. J. Eisenstein and W. Hu, "Baryonic Features in the Matter Transfer Function," *Astrophys.J.*, vol. 496, pp. 605–+, Mar. 1998.

[67] P. Bendich, H. Edelsbrunner, and M. Kerber, "Computing robustness and persistence for images," *IEEE Trans. Vis. Comput. Graph.*, vol. 16, no. 6, pp. 1251–1260, 2010.

[68] H. Edelsbrunner, D. Letscher, and A. Zomorodian, "Topological persistence and simplification," *Discrete and Computational Geometry*, vol. 28, pp. 511–533, 2002.

[69] T. Sousbie, C. Pichon, and H. Kawahara, "The persistent cosmic web and its filamentary structure - II. Illustrations," *Mon. Not. R. Astron. Soc.*, pp. 530–+, Apr. 2011.

[70] T. Sousbie, C. Pichon, H. Courtois, S. Colombi, and D. Novikov, "The Three-dimensional Skeleton of the SDSS," *Astrophys. J. Lett.*, vol. 672, pp. L1–L4, Jan. 2008.

[71] T. Sousbie, C. Pichon, S. Colombi, D. Novikov, and D. Pogosyan, "The 3D skeleton: tracing the filamentary structure of the Universe," *Mon. Not. R. Astron. Soc.*, vol. 383, pp. 1655–1670, Feb. 2008.

[72] T. Sousbie, S. Colombi, and C. Pichon, "The fully connected N-dimensional skeleton: probing the evolution of the cosmic web," *Mon. Not. R. Astron. Soc.*, vol. 393, pp. 457–477, Feb. 2009.