Tecnociencia, Vol. 28, N°2: 123-156
Julio-Diciembre 2026
ISSN L 2415-0940

Mapping the Cosmic Web: A Statistical Comparison of Filamentary Structure in SDSS and Simulated Galaxy Distributions

Mapeo de la red cósmica: comparación estadística de la estructura filamentaria en SDSS y distribuciones de galaxias simuladas

Franklin Simón Vásquez Guardia
Universidad de Panamá, Facultad de Ciencias Naturales y Exactas, Departamento de Física, Panamá.
franklin-s.vasqywez-g@up.ac.pa https://orcid.org/0000-0002-1413-6617

Fecha de recepción: 2 de febrero de 2026
Fecha de aceptación: 4 de mayo de 2026

DOI: https://doi.org/10.48204/j.tecno.v8n2.a10863

ABSTRACT

We present a statistical analysis of large-scale filamentary structure using Sloan Digital Sky Survey (SDSS) data and synthetic cosmological simulations. A 112 630-galaxy catalog within the redshift range $0.06 < z < 0.09$ is used to construct a two-dimensional density field via angular binning and Gaussian smoothing, after which a morphology-based filament detection algorithm is applied.

To quantify filament strength, we define a percentile-based observable given by the 98th percentile of the filament intensity field. This procedure is applied consistently to both observational data and an ensemble of 10 000 simulated realizations, yielding a statistical distribution of threshold values. The observed threshold is compared to the simulation distribution using cumulative distribution functions and dispersion measures. We find that the SDSS value lies within the expected range, indicating statistical consistency with ΛCDM (Lambda Cold Dark Matter) cosmological models.

As a complementary analysis, we fix the filament intensity threshold to the observed value and measure the fraction of the field exceeding this threshold. The observed fraction lies toward the lower tail of the simulated distribution, suggesting a mild tendency toward less spatially extended high-intensity filamentary structures.

Finally, we discuss limitations related to survey geometry and the projection of the three-dimensional galaxy distribution onto a two-dimensional field, which can reduce the apparent contrast and spatial coherence of filamentary structures.

KEYWORDS

Large-scale structure, cosmic filaments, SDSS, density field, filament detection, percentile threshold, cosmological simulations, ΛCDM.

RESUMEN

Se presenta un análisis estadístico de la estructura filamentaria a gran escala del universo utilizando datos del Sloan Digital Sky Survey (SDSS) y simulaciones cosmológicas sintéticas. A partir de un catálogo de 112 630 galaxias en el intervalo $0.06 < z < 0.09$, se construye un campo de densidad bidimensional mediante binning angular y suavizado gaussiano, sobre el cual se aplica un algoritmo de detección de filamentos.

Para cuantificar la prominencia de las estructuras, se introduce un observable basado en percentiles, definido como el umbral correspondiente al percentil 98 del campo de intensidad filamentaria. Este procedimiento se replica en un conjunto de 10 000 realizaciones simuladas, generando una distribución estadística de umbrales comparables. El valor observado se contrasta con la distribución simulada mediante funciones de distribución acumulada y medidas de dispersión. Los resultados muestran que el umbral observado es consistente con el rango esperado de las simulaciones, indicando consistencia estadística con modelos cosmológicos tipo ΛCDM (materia oscura fría con constante cosmológica).

De manera complementaria, se analiza la fracción del campo que excede un umbral fijo definido por el valor observado. En este caso, la fracción medida se ubica hacia la cola inferior de la distribución simulada, sugiriendo una leve tendencia hacia estructuras filamentarias menos extendidas en los datos.

Finalmente, se discuten las limitaciones asociadas a la geometría observacional y a la proyección bidimensional del campo, las cuales pueden afectar la reconstrucción de la estructura tridimensional y reducir el contraste y la coherencia espacial de las estructuras filamentarias.

PALABRAS CLAVE

Estructura a gran escala, filamentos cósmicos, SDSS, campo de densidad, detección de filamentos, percentiles, simulaciones cosmológicas, ΛCDM

INTRODUCTION

The large-scale structure of the Universe emerges from the gravitational amplification of primordial density perturbations within an expanding spacetime (Peebles, 1980; Dodelson, 2003). In the standard cosmological framework, this evolution is described by the ΛCDM (Lambda Cold Dark Matter) model, in which the present-day energy density of the Universe is dominated by dark energy, followed by dark matter and baryonic matter, with observations indicating a spatially flat geometry (Planck Collaboration, 2018).

The origin of structure is traced back to the inflationary paradigm, which posits a phase of rapid exponential expansion in the early Universe. During this period, quantum fluctuations were stretched to cosmological scales, generating nearly Gaussian and approximately scale-invariant perturbations (Mukhanov, 2005; Planck Collaboration, 2018). These fluctuations act as the seeds for structure formation and are commonly described in terms of a primordial power spectrum of the form

$$P(k) \propto k^{n_s}$$

where n_s is the scalar spectral index.

Although these initial conditions are well described by a Gaussian random field, the subsequent gravitational evolution of the Universe is inherently nonlinear. As density perturbations grow, gravitational collapse induces mode coupling and drives the system away from Gaussianity, leading to the formation of anisotropic and highly structured features such as filaments, sheets, and clusters (Bond, Kofman, & Pogosyan, 1996). As a result, the present-day matter distribution cannot be fully characterized by linear statistical descriptors alone. Nevertheless, the large-scale morphology of the cosmic web retains imprints of the initial conditions, providing a complementary means of probing the underlying cosmological model beyond the linear regime.

As structure evolves, matter organizes into the cosmic web, a network of voids, sheets, filaments, and clusters. Filaments, in particular, are elongated overdense structures that connect massive nodes such as galaxy clusters and serve as channels for matter accretion (Bond et al., 1996; Springel et al., 2005). Their geometry and spatial distribution encode information about both the primordial fluctuations and the nonlinear processes governing structure formation, making them a valuable probe of cosmology.

Observationally, the large-scale distribution of galaxies traces the underlying matter density field. Wide-field surveys such as the Sloan Digital Sky Survey (SDSS) have mapped millions of galaxies across large cosmological volumes, enabling detailed studies of the cosmic web (York et al., 2000). However, galaxies are not a direct representation of the total matter distribution, but rather biased tracers of it. The efficiency of galaxy formation depends on local environmental conditions, leading to a nontrivial relationship between galaxy number density and total matter density (Peebles, 1980). Despite this, galaxies remain a powerful observational probe, as their spatial distribution traces the gravitational potential wells dominated by dark matter.

An example of the large-scale galaxy distribution observed by SDSS is shown in Figure 1, where galaxies are plotted in right ascension–redshift space. The characteristic wedge-like geometry reflects both the survey footprint and the radial nature of redshift measurements, revealing filamentary patterns and large-scale clustering directly in observational data.

In this work, we focus on the statistical characterization of filamentary structure in a low-redshift slice of SDSS data ($0.06 < z < 0.09$). Building on the framework described above, we construct a two-dimensional density field from the observed galaxy distribution through angular binning and Gaussian smoothing and apply a morphology-based filament detection method to obtain a filament intensity map. Rather than identifying individual filaments as discrete objects, we adopt a field-based approach and define a scalar observable based on the upper percentile of the filament intensity distribution.

This observable is then compared with results obtained from synthetic universes generated using simplified ΛCDM-inspired models. By applying the same analysis pipeline to both observational data and simulated realizations, we assess whether the strength of filamentary structure in the observed Universe is statistically consistent with the model. This approach emphasizes real-space morphology and provides a complementary perspective to traditional analyses based on correlation functions or power spectra (Peacock, 1999).

THEORETICAL BACKGROUND AND METHODOLOGICAL FRAMEWORK

Cosmological Framework

The large-scale dynamics of the Universe are well described by the Friedmann–Lemaître–Robertson–Walker (FLRW) metric, which assumes spatial homogeneity and isotropy. In comoving coordinates, the metric takes the form

$$ds^2 = -c^2 dt^2 + a^2(t)\left[\frac{dr^2}{1 - kr^2} + r^2 d\Omega^2\right],$$

where $a(t)$is the scale factor, k characterizes the spatial curvature, and $d\Omega^2$ represents the angular component (Dodelson, 2003; Hogg, 1999).

Comoving coordinates are used to factor out the expansion of the Universe, allowing distances between objects to remain constant in the absence of peculiar motions and providing a convenient framework for describing structure formation.

The evolution of the scale factor is governed by the Friedmann equation,

$$H^2(a) = H_0^2[\Omega_r a^{-4} + \Omega_m a^{-3} + \Omega_k a^{-2} + \Omega_\Lambda],$$

where $H(a)$ is the Hubble parameter accounting for expansion, H_0 is its present-day value, and Ω_r, Ω_m, Ω_k, and $\Omega_\Lambda$denote the normalized density parameters corresponding to radiation, matter, spatial curvature, and dark energy, respectively (Peebles, 1980; Dodelson, 2003).
The curvature density parameter is defined as

$$\Omega_k = -\frac{kc^2}{a_0^2 H_0^2},$$

such that a flat Universe corresponds to Ω_k=0, while positive or negative values indicate open or closed spatial geometries depending on the sign of k. Observational constraints from the cosmic microwave background are consistent with a nearly flat Universe (Planck Collaboration, 2018).

Within the ΛCDM framework, the present-day energy budget of the Universe can be approximated as

$$\Omega_\Lambda \approx 0.68,\ \ \Omega_{\rm cdm} \approx 0.27,\ \ \Omega_b \approx 0.05,\ \ \Omega_k \approx 0,$$

with the total matter density given by Ω_m in:

$$\Omega_m = \Omega_{\rm cdm} + \Omega_b.$$

where $\Omega_{\rm cdm}$ and $\Omega_{\rm b}$ correspond to cold dark matter and baryonic matter, respectively.

Power Spectrum and Initial Conditions

Within the ΛCDM framework, baryonic matter constitutes the visible component traced by galaxies, while cold dark matter dominates the gravitational potential wells responsible for structure formation (Planck Collaboration, 2018).

The growth of structure is driven by gravitational instability acting on primordial density perturbations. These perturbations are conveniently described in terms of the density contrast

$$\delta(\mathbf{x}) = \frac{\rho(\mathbf{x}) - \bar{\rho}}{\bar{\rho}},$$

where $\rho(\mathrm{x})$ is the local matter density and $\bar{\rho}$ is the mean density of the Universe (Peebles, 1980).

In the linear regime, the statistical properties of the density field are fully characterized by the matter power spectrum $P(k)$, defined as the Fourier transform of the two-point correlation function. For a statistically homogeneous and isotropic field,

$$\langle \delta(\mathbf{k})\delta^*(\mathbf{k}') \rangle = (2\pi)^3 \delta_D(\mathbf{k} - \mathbf{k}')P(k),$$

where $\delta(\mathrm{k})$ is the Fourier transform of the density contrast and δ_D is the Dirac delta function (Peacock, 1999).

The primordial power spectrum is commonly parameterized as a power law,

$$P(k) \propto k^{n_s},$$

where n_s is the scalar spectral index. Observations indicate that n_s is close to unity, corresponding to nearly scale-invariant fluctuations (Planck Collaboration, 2018). These initial perturbations grow under gravitational instability, eventually evolving into the nonlinear structures observed in the present-day Universe (Peebles, 1980; Dodelson, 2003).

Density Field Reconstruction and Smoothing

While the theoretical density field is continuous, observations provide only a discrete sampling of the underlying matter distribution through galaxies. Moreover, galaxies are biased tracers of the total matter field, reflecting the complex physics of galaxy formation (Peebles, 1980; Kaiser, 1984).

To approximate a continuous field, the observed galaxy distribution is converted into a density field through spatial binning. This process yields a discrete estimate of the density contrast, which serves as a proxy for the underlying matter distribution.

However, the resulting field is affected by shot noise due to the finite number of galaxies. To suppress these fluctuations and enhance coherent structures, a Gaussian smoothing filter is applied,

$$\delta_{\text{smooth}}(\mathbf{x}) = G_\sigma * \delta(\mathbf{x}),$$

where G_σ is a Gaussian kernel with characteristic scale σ. This operation reduces small-scale noise while preserving large-scale morphological features relevant to the cosmic web (Martínez & Saar, 2002)

The Cosmic Web and Filamentary Structure

As density perturbations evolve into the nonlinear regime, gravitational collapse proceeds anisotropically, leading to the formation of the cosmic web. This structure is characterized by a hierarchy of voids, sheets, filaments, and nodes, which emerge from collapse along one, two, or three spatial directions, respectively (Bond et al., 1996).

Filaments correspond to elongated overdense regions that connect massive nodes such as galaxy clusters. They act as channels for matter transport and play a central role in the assembly of large-scale structure. Because their geometry reflects both the initial conditions and the subsequent nonlinear evolution, filamentary structures provide a sensitive probe of cosmological models.

Filament Detection

To characterize filamentary structure, morphology-based detection methods operate on the smoothed density field to identify anisotropic features associated with elongated structures. These methods typically rely on local properties of the field, such as gradients, curvature, or directional coherence, to distinguish filaments from isotropic overdensities and underdense regions.

Several well-established techniques have been developed to extract filamentary structures from cosmological data. For example, the Discrete Persistent Structures Extractor (DisPerSE) identifies filaments using topological persistence, tracing the skeleton of the density field through critical points and their connectivity (Sousbie, 2011). Alternatively, multiscale approaches such as the NEXUS and Multiscale Morphology Filter (MMF) methods classify structures based on the eigenvalues of the Hessian of the density field, enabling a scale-dependent identification of filaments, sheets, and clusters (Aragón-Calvo et al., 2007; Cautun et al., 2013).

Rather than explicitly implementing one of these discrete structure extraction algorithms, the approach adopted in this work follows a field-based methodology. Instead of identifying individual filaments as isolated objects, the method produces a continuous scalar field that encodes the degree of filamentarity at each point in the density field.

This representation allows for a statistical characterization of the filament network without requiring explicit segmentation or topological reconstruction. As a result, it facilitates direct comparison with theoretical models and simulated realizations through global observables derived from the filament intensity distribution.

OBSERVATIONAL DATA AND INSTRUMENTATION

The Sloan Digital Sky Survey

The observational data used in this work are drawn from the Sloan Digital Sky Survey (SDSS), one of the most extensive spectroscopic and photometric surveys of the sky. SDSS has mapped millions of galaxies over large cosmological volumes using a dedicated 2.5 meter telescope, providing precise measurements of angular positions and redshifts (York et al., 2000).

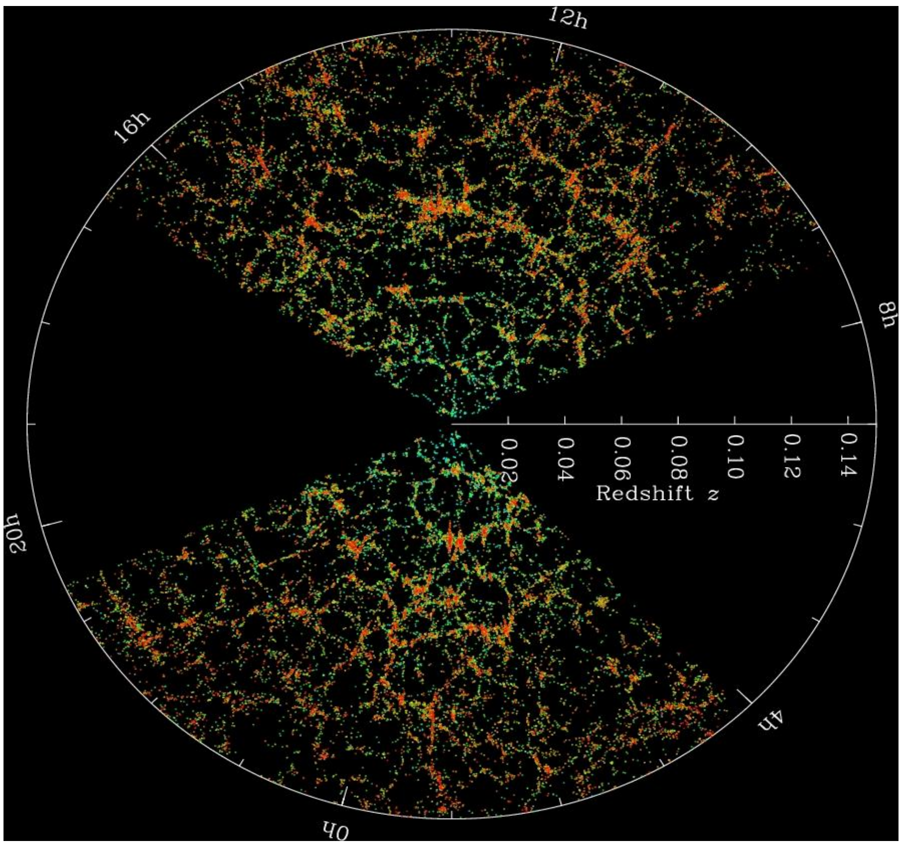
The survey operates in a drift-scan mode, covering large contiguous regions of the sky in a series of stripes. Spectroscopic observations provide redshift measurements, which allow the reconstruction of the three-dimensional distribution of galaxies. The resulting dataset forms a powerful probe of large-scale structure and the cosmic web.
An example of the SDSS galaxy distribution is shown in Figure 1, where galaxies are plotted in right ascension–redshift space. The characteristic wedge-shaped geometry reflects both the angular footprint of the survey and the radial nature of redshift measurements, with the two lobes corresponding to distinct survey regions.

The survey provides photometric measurements in five optical bands (*u*,*g*,*r*,*i*,*z*) and spectroscopic data that yield redshift estimates for millions of galaxies. The combination of angular positions (right ascension and declination) with redshift information enables the reconstruction of the three-dimensional distribution of matter in the nearby Universe.

Figure 1

SDSS Survey Geometry in RA–Redshift Space

Visualization of the Sloan Digital Sky Survey (SDSS) galaxy distribution in right ascension and redshift, illustrating the characteristic wedge-shaped geometry of the survey footprint. The radial axis corresponds to redshift, which serves as a proxy for distance, while the angular coordinate traces sky coverage. Image credit: Sloan Digital Sky Survey (SDSS), retrieved from https://classic.sdss.org/includes/sideimages/sdss_pie2.php



Data Products and Selection

The data was obtained through the SDSS CasJobs interface, which provides SQL-based access to the survey database (Sloan Digital Sky Survey Collaboration, 2023). Each entry

provides the right ascension (RA), declination (Dec), and spectroscopic redshift z of individual galaxies.
In this analysis, we restrict the sample to a narrow redshift interval, $0.06 < z < 0.09$, corresponding to a relatively local volume of the Universe and a lookback time of approximately 0.8 –1.2 Gyr in a standard ΛCDM cosmology. This selection balances the need for sufficient galaxy counts with the requirement of minimizing projection effects, ensuring that large-scale structures remain coherent within the slice.

The final dataset consists of 112,630 galaxies, providing a statistically robust sampling of the large-scale structure within this redshift range.

The spatial distribution of the selected galaxies is shown in Figure 2, where the SDSS data are displayed in right ascension–redshift coordinates. The characteristic wedge-shaped structure reflects both the survey geometry and the radial selection imposed by the redshift cut.

The combination of angular coordinates and redshift information allows the reconstruction of the three-dimensional spatial distribution of galaxies. By converting redshift into comoving distance under the assumed cosmological model, the galaxy catalog can be embedded in a three-dimensional coordinate system.

While the SDSS catalog provides angular positions and redshifts, physical interpretation of large-scale structure requires reconstructing the galaxy distribution in three-dimensional space.

To achieve this, redshift is converted into comoving distance, which represents a coordinate system that factors out the expansion of the Universe. In comoving coordinates, distances between objects remain fixed in the absence of peculiar velocities, allowing direct comparison of structures across cosmic time (Hogg, 1999).

The comoving radial distance $\chi(z)$ is computed as

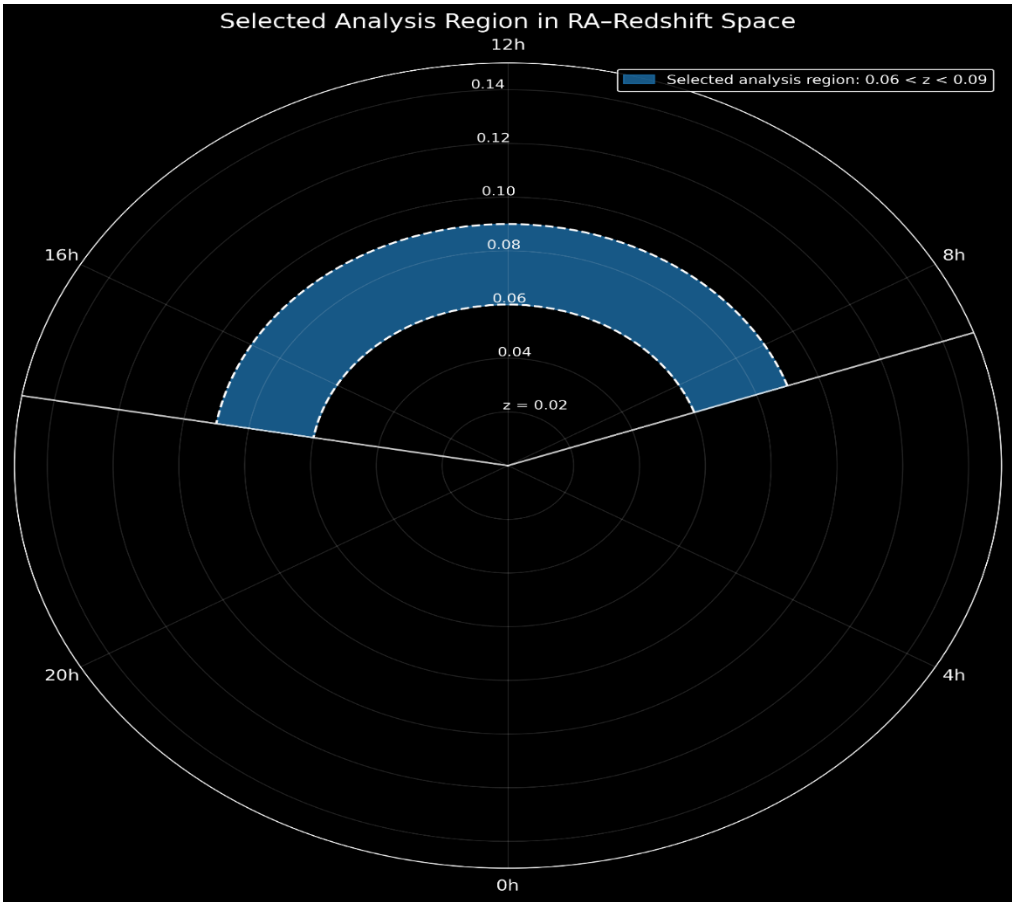
$$\chi(z) = c\int_0^z \frac{dz'}{H(z')},$$

$$x = \chi(z)\cos(\delta)\cos(\alpha), y = \chi(z)\cos(\delta)\sin(\alpha), z = \chi(z)\sin(\delta),$$

where α and δ are the right ascension and declination, respectively.

Figure 2
SDSS Galaxy Distribution in RA–Redshift Space

Distribution of galaxies in right ascension and redshift for the selected SDSS sample. The characteristic wedge-like geometry arises from the survey footprint combined with the applied redshift cut, defining the angular and radial boundaries of the analyzed volume.

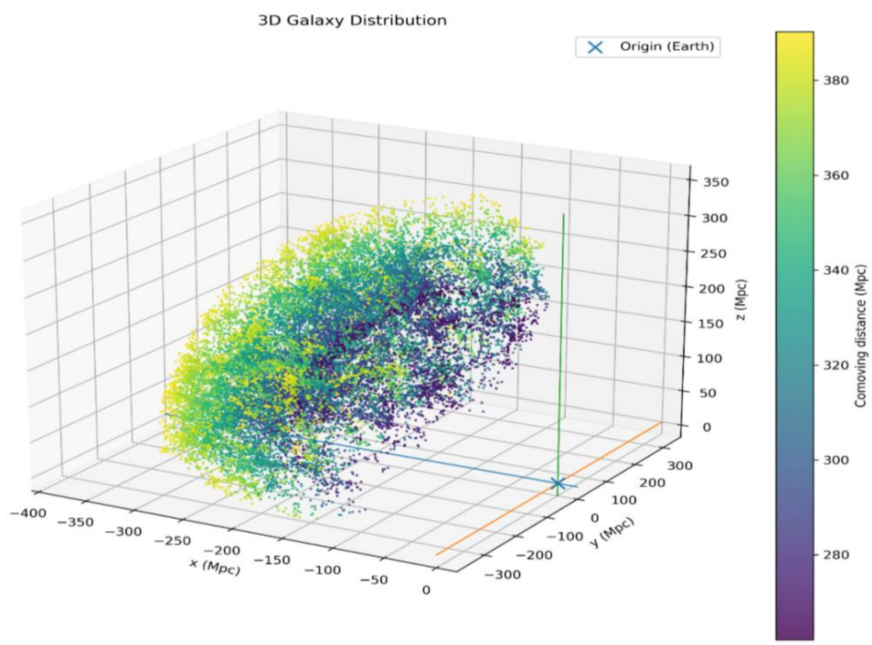


The resulting three-dimensional distribution is shown in Figure 3, where the observer is located at the origin and the galaxy sample occupies a wedge-shaped volume extending outward in comoving space. The geometry reflects both the angular footprint of the survey and the imposed redshift limits.

The reconstruction is performed in comoving coordinates, which factor out the expansion of the Universe and provide a physically meaningful spatial framework. It should be noted, however, that the distances are inferred from redshift measurements and therefore include distortions due to peculiar velocities. Furthermore, while this representation provides useful visualization of the galaxy distribution, the filament analysis itself is performed on a projected two-dimensional density field.

This representation highlights the filamentary nature of the galaxy distribution in three dimensions and provides a useful framework for interpreting large-scale structure.

Figure 3
Three-Dimensional Galaxy Distribution
Three-dimensional reconstruction of the galaxy distribution in comoving coordinates. The observer is located at the origin, and the data occupy a wedge-shaped volume determined by the survey footprint and redshift limits.



Survey Footprint and Observational Geometry
Due to the drift-scan observing strategy of SDSS, the survey footprint is not rectangular but instead follows curved stripes on the celestial sphere. When projected onto right ascension and declination coordinates, this results in a non-uniform, irregularly shaped observational region.

In the context of this analysis, regions of the sky outside the survey coverage contain no data and must be explicitly masked. This is achieved by constructing a footprint mask derived from the observed galaxy distribution, ensuring that subsequent analysis steps operate only within the valid observational domain.

This masking procedure is essential to avoid artificial features arising from empty regions and to ensure that both observational data and simulated fields are treated consistently when applying the analysis pipeline.

Computational Tools and Data Handling
The analysis pipeline is implemented in Python (Python Software Foundation, 2023) using standard scientific computing libraries. Numerical operations and array manipulation are performed using NumPy (Harris et al., 2020), while visualization is handled with Matplotlib (Hunter, 2007). The construction and smoothing of the density field rely on functions from SciPy, particularly for Gaussian filtering (Virtanen et al., 2020).

Custom scripts are used to process the SDSS catalog, construct density fields, and apply filament detection methods. These tools enable efficient handling of large datasets and ensure reproducibility of the analysis across both observational and simulated data.

METHODOLOGY

Overview of the Analysis Pipeline
The methodology developed in this work aims to quantify filamentary structure in galaxy surveys through a reproducible and statistically interpretable pipeline. The approach consists of four main stages: (i) construction of a two-dimensional density field from the galaxy catalog, (ii) smoothing of the field to suppress shot noise, (iii) extraction of a filament intensity field, and (iv) definition of a percentile-based observable used for comparison with simulations.

Rather than identifying individual filaments as discrete objects, the analysis adopts a field-based approach, in which filamentary structure is encoded in a continuous scalar field. This enables a statistical characterization of the cosmic web and facilitates direct comparison between observational data and synthetic realizations.

Construction of the Density Field
The starting point of the analysis is the discrete distribution of galaxies in angular coordinates. To obtain a discretized approximation to the continuous matter distribution, the galaxy positions are projected onto a two-dimensional grid in right ascension and declination.

This is achieved by binning the galaxy sample into a regular grid of $N \times N$ cells, where each cell counts the number of galaxies falling within its angular boundaries. The resulting map can be expressed as

$$H(\alpha,\delta) = \sum_{i=1}^{N_{\mathrm{gal}}} \mathcal{W}_i\,(\alpha,\delta),$$

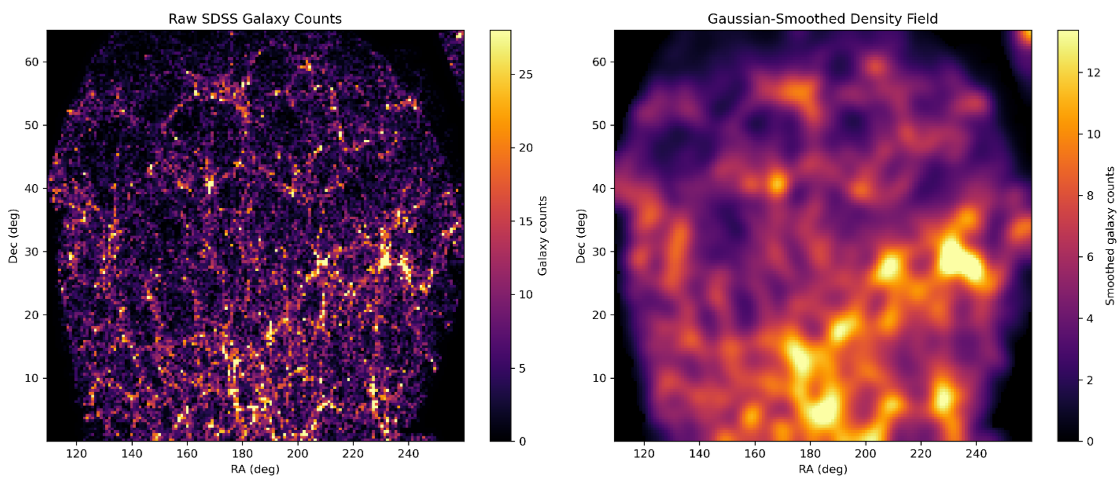
where $W_i(\alpha,\delta)$ is a window function that assigns each galaxy to the corresponding grid cell. In practice, this corresponds to constructing a two-dimensional histogram of galaxy positions. The resulting field represents the raw projected galaxy density, where $\alpha$and δdenote right ascension and declination, respectively.

This procedure converts the discrete galaxy catalog into a pixelized density field, enabling subsequent analysis using image-based and statistical techniques. The resulting raw density map, together with its Gaussian-smoothed counterpart, is shown in Figure 4.

Figure 4

Raw and Smoothed Galaxy Density Fields

Left: raw binned galaxy counts in RA–Dec space. Right: Gaussian-smoothed density field highlighting large-scale structure. The smoothing suppresses shot noise and reveals coherent filamentary features.



Density Contrast and Gaussian Smoothing

The raw density field is affected by shot noise due to the discrete nature of galaxy sampling. To suppress small-scale fluctuations and enhance coherent large-scale structures, the density field is smoothed using a Gaussian kernel.

The smoothed density field is defined as

$$H_{\mathrm{smooth}}(\alpha,\delta)=G_\sigma * H(\alpha,\delta),$$

where G_σ is a Gaussian kernel with standard deviation σ, and $*$denotes convolution (Martínez & Saar, 2002).

The choice of σ sets the characteristic scale of the structures being probed. Larger values of σ emphasize extended structures such as filaments and clusters, while smaller values retain finer details but are more sensitive to noise.

This smoothing step is essential for revealing the continuous morphology of the cosmic web and provides the input for filament detection. In particular, it ensures that the analysis focuses on physically meaningful, extended structures rather than fluctuations induced by discrete sampling.

Survey Mask and Valid Analysis Region

The SDSS observational footprint does not cover a contiguous rectangular region of the sky, but instead follows a complex geometry determined by the survey scanning strategy (York et al., 2000). As a result, the projected density field contains regions with no data corresponding to areas outside the survey coverage.

To ensure that the analysis is restricted to physically meaningful regions, a survey mask is constructed directly from the observed galaxy distribution. This is achieved by identifying pixels in the raw density field that contain galaxy counts and expanding this footprint slightly using a low-level smoothing operation. The resulting mask defines a valid analysis region, excluding empty areas where no observational data are available.

Formally, the mask is represented as a binary field,

$$M(\alpha,\delta)=\begin{cases}1, & \text{if the pixel belongs to the observed footprint}\\ 0, & \text{otherwise}\end{cases}$$

All subsequent quantities, including the smoothed density field and filament intensity map, are evaluated only within this valid region. Pixels outside the mask are excluded from statistical calculations and are treated as undefined in visual representations.

This masking procedure is essential to prevent artificial structures from appearing at the boundaries of the survey and to ensure a consistent comparison between observational data and simulated realizations. The same masking procedure is applied to both datasets to maintain statistical consistency.

Filament Intensity Field

The smoothed density field provides a continuous representation of the large-scale matter distribution, but does not directly isolate filamentary structures. To extract these features, a morphology-based filament detection algorithm is applied to the smoothed field.

This procedure produces a scalar field,

$$f_{\text{map}}(\alpha,\delta),$$

referred to as the *filament intensity map*, which assigns to each pixel a value representing the degree to which the local environment resembles a filamentary structure. High values of f_{map} correspond to elongated, ridge-like features connecting overdense regions, while low values indicate voids or more isotropic environments (Bond, Kofman, & Pogosyan, 1996). Unlike methods that identify individual filaments as discrete objects, this approach treats filamentarity as a continuous field. This allows for a statistical characterization of the cosmic web without requiring explicit segmentation or topological classification of structures.

The filament intensity field therefore encodes the morphological information of the density field in a form suitable for defining global statistical observables, which are introduced in the following section.

This field-based representation differs from discrete skeleton-based methods such as DisPerSE, which explicitly reconstruct the filamentary network through topological analysis of critical points and their connectivity (Sousbie, 2011), as well as multiscale approaches such as NEXUS/MMF, which classify structures using the eigenvalues of the density field Hessian across multiple spatial scales (Cautun et al., 2013). Nevertheless, the continuous formulation adopted here retains sensitivity to filamentary morphology while enabling a straightforward statistical treatment of the resulting field.

Percentile-Based Filament Observable

To quantify the filamentary structure of the cosmic web in a manner that is both robust and comparable across datasets, we define a scalar observable based on the distribution of filament intensities.

Given the filament intensity field $f_{\text{map}}(\alpha,\delta)$, we consider its statistical distribution within the valid survey region defined by the mask $M(\alpha,\delta)$. Rather than selecting structures using an absolute threshold, which would depend sensitively on normalization and smoothing scale, we adopt a percentile-based approach.

Specifically, we define the threshold value

$$T_{98}=P_{98}\left(f_{\text{map}}\right),$$

where P_{98} denotes the 98th percentile of the filament intensity distribution evaluated over all pixels satisfying M=1. This definition ensures that the threshold is determined relative to the intrinsic distribution of filament strengths in the field.
A corresponding binary filament map is then constructed as

$$f_{\text{binary}}(\alpha,\delta)=\begin{cases}1, & f_{\text{map}}(\alpha,\delta)>T_{98}\\ 0, & \text{otherwise}\end{cases}$$

which isolates the regions associated with the highest filament intensities.

The percentile thresholding procedure is illustrated in Figure 5, where the filament intensity field is shown together with the contour corresponding to the 98th percentile. This contour delineates the regions selected as significant filamentary structures and defines the subset of the field used in the construction of the binary filament map.

The quantity T_{98} itself is treated as the primary observable of this work. It encodes the characteristic strength of the most prominent filamentary structures in the field and provides a scalar measure that can be directly compared between observational data and simulated realizations.

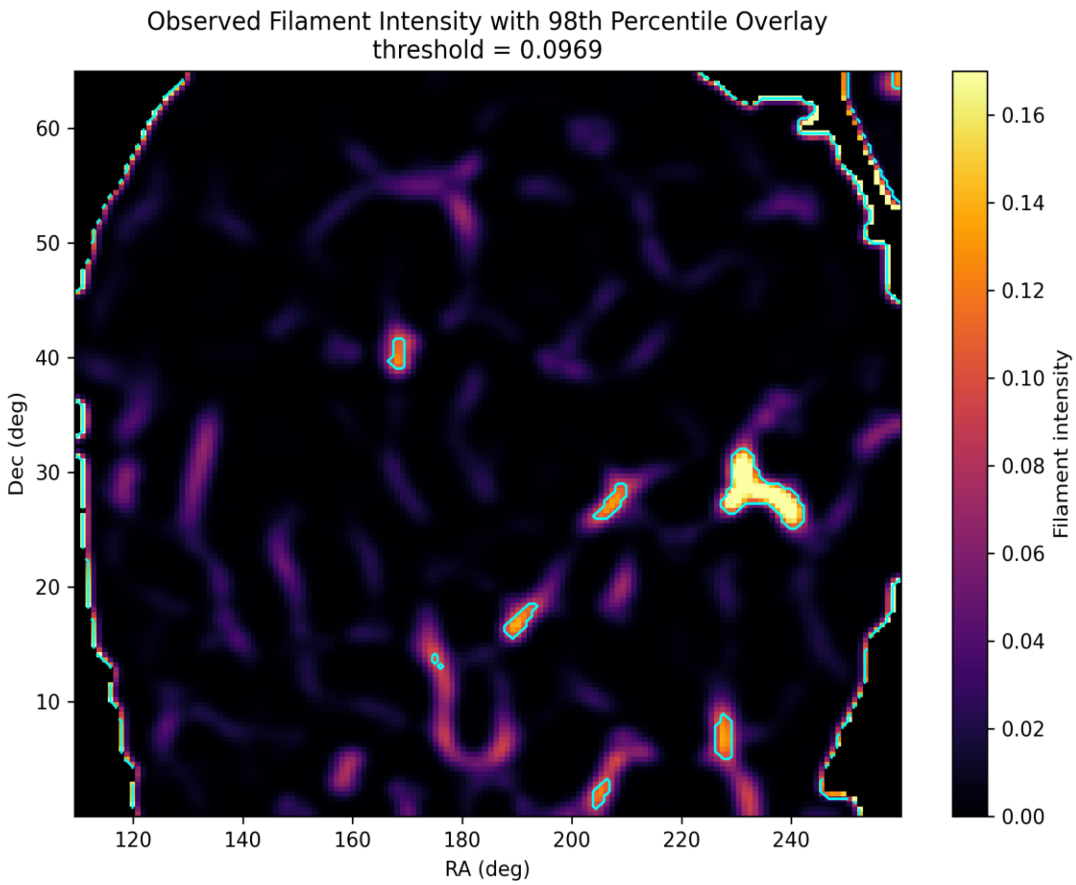
By construction, this approach avoids arbitrary threshold choices and is insensitive to overall normalization differences, making it particularly well suited for statistical comparisons across datasets with different sampling properties.

It is important to emphasize that, by construction, the fraction of pixels identified as filamentary is fixed by the percentile definition. As a result, the filament fraction does not constitute an independent observable in this framework. Instead, the physically meaningful quantity is the threshold value T_{98} itself, which encodes the relative intensity required to isolate the most prominent filamentary structures within the field.

It is useful to emphasize that the percentile-based threshold and the fixed-threshold fraction probe complementary aspects of filamentary structure. The percentile threshold T_{98} is primarily sensitive to the upper tail of the filament intensity distribution and therefore reflects the characteristic strength of the most prominent filamentary features. In contrast, the fixed-threshold filament fraction probes the spatial extent and connectivity of these structures, providing a measure of how widespread high-intensity filamentary regions are within the field.

Figure 5

Filament Intensity Field with 98th Percentile Threshold Overlay
Filament intensity map derived from the smoothed density field, with contours indicating the 98th percentile threshold. The contour delineates the regions classified as high-significance filamentary structures and defines the corresponding binary filament selection. The filament intensity is a dimensionless quantity derived from the morphology of the smoothed density field, encoding the degree of filamentary structure at each location.



Simulation Pipeline and Statistical Comparison

To interpret the observational measurement of filamentary structure, we construct an ensemble of synthetic density fields designed to approximate the statistical properties of large-scale structure within a simplified ΛCDM-inspired framework.

The simulations are based on stochastic realizations of the density field, generated from an underlying power spectrum characterized by an effective spectral parameter $n_s$and a characteristic scale parameter k_0. In this work, these parameters are treated as phenomenological inputs controlling the large-scale clustering behavior of the field, rather than direct measurements of the physical primordial power spectrum.

In the simulations presented in this work, the effective spectral parameter and characteristic scale were set to $n_s = 0.15$ and $k_0 = 0.05$, respectively. These values were selected to reproduce the large-scale clustering behavior observed in the SDSS data at a qualitative level, ensuring that the simulated density fields exhibit comparable visual and statistical properties.

While these parameters are not intended to represent precise cosmological measurements, their values define the statistical baseline against which the observational data are compared and are therefore reported here to ensure reproducibility of the analysis.

The initial density field is modeled using Gaussian or lognormal statistics. Gaussian random fields provide a natural description of the linear regime of structure formation, where density fluctuations are small and statistically homogeneous (Peacock, 1999).

However, the evolved matter distribution exhibits significant non-Gaussian features due to nonlinear gravitational collapse. To approximate this behavior, lognormal transformations are often employed as a simple extension that captures the skewness of the density field while preserving its large-scale correlations (Coles & Jones, 1991).

To produce realistic galaxy distributions, additional observational effects are incorporated into the simulations. First, a linear galaxy bias parameter is introduced to account for the fact that galaxies trace the underlying matter density field in a nontrivial way (Peebles, 1980). Second, Poisson sampling is applied to convert the continuous density field into a discrete galaxy distribution, mimicking the stochastic nature of galaxy counts.

Further observational realism is introduced through a completeness factor, which models the finite detection efficiency of the survey, and a redshift-space distortion (RSD) proxy, implemented as a Gaussian perturbation along the line of sight to approximate the impact of peculiar velocities on observed redshifts (Hamilton, 1998).

Each simulated realization is constructed to match the number of galaxies in the observational dataset, ensuring a consistent level of sampling. The resulting simulated catalogs are then projected onto the same angular grid used for the SDSS data, and the full analysis pipeline, including density field construction, smoothing, masking, filament detection, and percentile thresholding, is applied identically to both observational and simulated datasets.

It is important to emphasize that the parameter n_s used in this context should be interpreted as an effective spectral-tilt parameter within a simplified two-dimensional model, rather than a direct representation of the physical scalar spectral index inferred from cosmological observations (Planck Collaboration, 2018). Accordingly, the simulated fields should not be interpreted as full physical realizations of ΛCDM cosmology, but rather as controlled stochastic models designed to reproduce key statistical properties of large-scale structure.

The goal of this approach is to provide a consistent statistical baseline for comparison with observational data, rather than a detailed dynamical model of structure formation.

To illustrate the outcome of this procedure, a representative simulated realization is shown in Figure 6, where the filament intensity field is displayed together with the contour corresponding to the 98th percentile threshold. This visualization demonstrates how the same analysis pipeline applied to the observational data produces comparable filamentary structures in the simulated field, allowing for a direct statistical comparison between the two.

This representative realization provides a visual reference for interpreting the statistical distribution of filament strength discussed in the following section.

It is important to emphasize that the goal of this simulation framework is not to reproduce the full nonlinear dynamics of structure formation, as achieved in high-resolution N-body simulations, but rather to provide a statistically controlled baseline for testing morphology-based observables.

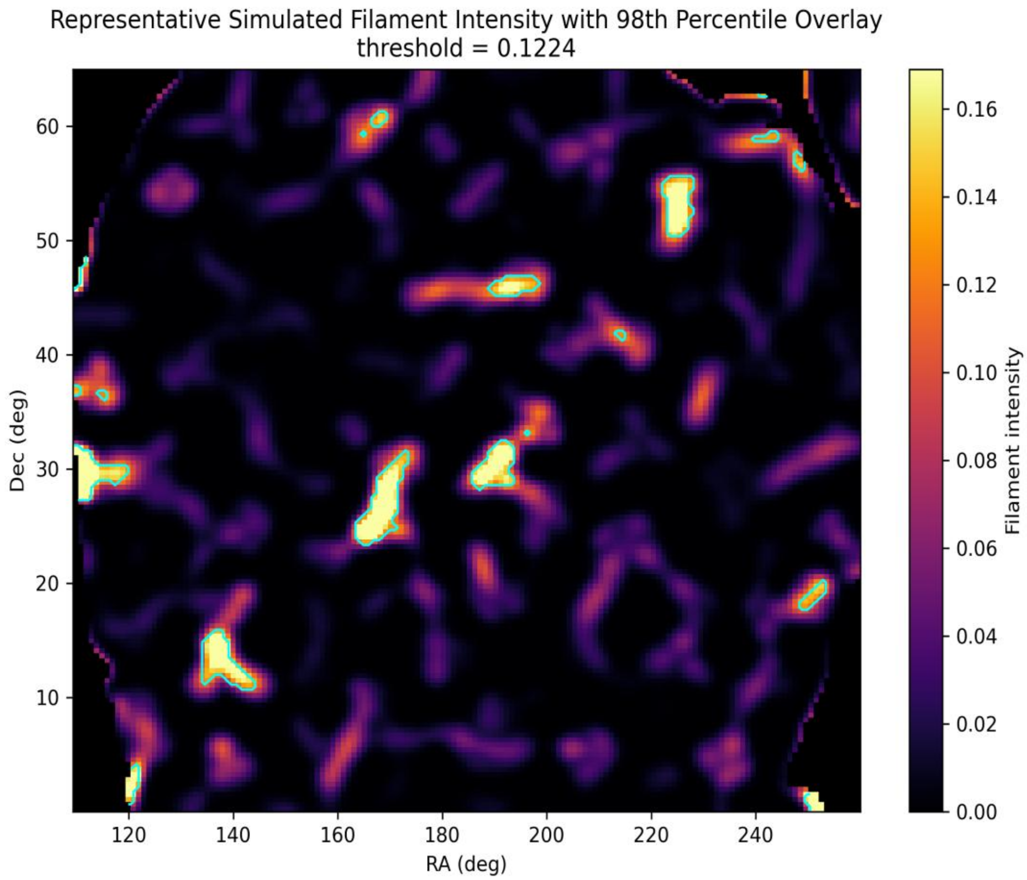
In this context, the use of Gaussian and lognormal fields is justified by their ability to reproduce key statistical properties of large-scale structure, including clustering behavior and density contrast distributions, at the scales probed in this work. For the relatively narrow redshift interval considered ($0.06 < z < 0.09$), where the analysis focuses on large-scale projected morphology rather than detailed halo dynamics, this approach provides a computationally efficient and physically motivated proxy.

Accordingly, the comparison performed here should be interpreted as a validation of the statistical methodology used to quantify filamentary structure, rather than a direct test of ΛCDM predictions at the level of full dynamical simulations.

Different choices of these parameters are not expected to qualitatively alter the main conclusions of the statistical comparison, as the analysis focuses on relative statistical properties rather than the absolute calibration of the model.

Figure 6
Representative Simulated Filament Intensity Field with 98th Percentile Threshold
Filament intensity map derived from a simulated realization, shown together with the contour corresponding to the 98th percentile threshold. Regions enclosed by the contour indicate the most significant filamentary structures selected by the percentile-based criterion. The same

analysis pipeline applied to the observational data is used, ensuring consistency in the definition of the filament observable.



Fixed-Threshold Filament Fraction Analysis

In addition to the percentile-based observable defined in the previous sections, a complementary analysis is performed by fixing the filament intensity threshold to the value measured in the observational data. In this formulation, the observable is defined as the fraction of the field exceeding a fixed intensity threshold,

$$F(T_{\text{obs}})=\frac{N(f_{\text{map}}>T_{\text{obs}})}{N_{\text{total}}},$$

where T_{obs} corresponds to the 98th percentile of the SDSS filament intensity field.

Unlike the percentile-based approach, where the filament fraction is fixed by construction, the fixed-threshold observable allows the spatial extent of filamentary regions to vary across realizations. It therefore provides an independent measure of how widespread high-intensity filamentary structures are under a fixed physical criterion. However, in contrast to the percentile-based observable, it is also sensitive to the normalization and smoothing scale of the filament intensity field.

The regions selected by this threshold in the SDSS data are shown in Figure 7, where pixels exceeding T_{obs}=0.0969 are highlighted. These regions define the observed filament fraction used in the subsequent statistical comparison.

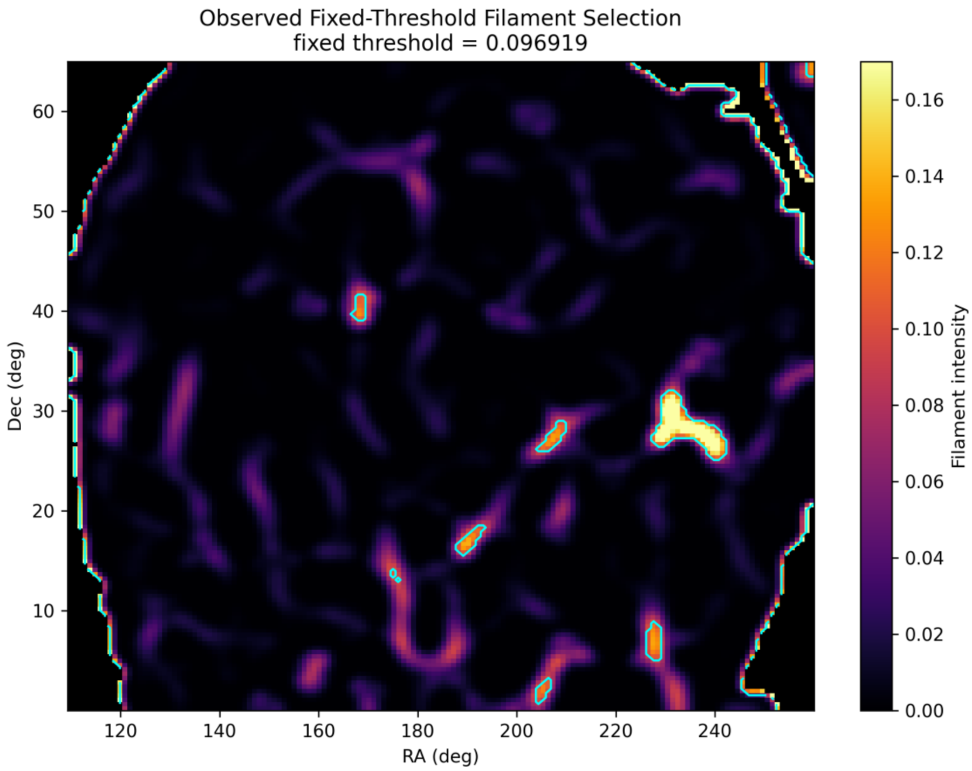
The same fixed threshold is then applied to each simulated realization, and the corresponding filament fraction is computed using an identical analysis pipeline. This ensures a consistent comparison between observational data and simulations.

Taken together, the percentile-based and fixed-threshold approaches provide two complementary characterizations of filamentary structure. While the former defines filaments through a fixed fraction of the highest-intensity regions, the latter fixes the intensity scale and allows the spatial extent of filaments to vary. This dual formulation enables a direct comparison between the strength and spatial coverage of filamentary structures in both observational data and simulations.

Figure 7

Fixed-Threshold Filament Selection in SDSS Data

Filament intensity map derived from the SDSS galaxy sample within the redshift range $0.06 < z < 0.09$. *Regions exceeding the fixed intensity threshold* $T_{\text{obs}} = 0.0969$*are highlighted, corresponding to the value obtained from the 98th percentile of the observed filament intensity field. These regions define the filament fraction* $F(T_{\text{obs}})$, *representing the spatial extent of high-intensity filamentary structures under a fixed threshold criterion.*



RESULTS

The statistical properties of the filament intensity threshold T_{98} are characterized using an ensemble of simulated realizations processed through the same analysis pipeline as the observational data. This enables a direct comparison between the observed SDSS value and the distribution expected under the adopted model. Histograms are constructed using a fixed binning scheme to enable consistent visual comparison across different observables.

The distribution of T_{98} values obtained from the simulations is shown in Figure 8. Each value corresponds to the 98th percentile of the filament intensity field in an individual realization. The resulting distribution spans a broad range, with a mean value of

$$\langle T_{98}^{\mathrm{sim}} \rangle \approx 0.132$$

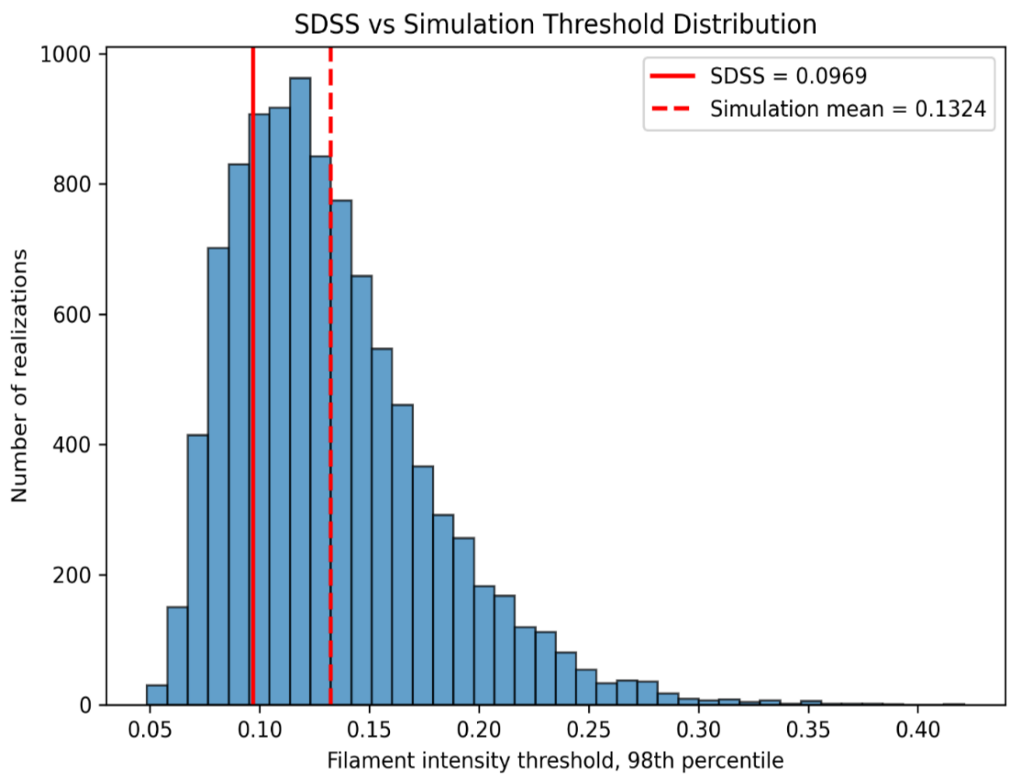
and a standard deviation of

$$\sigma_T \approx 0.045,$$

reflecting the intrinsic variability of filament strength across realizations.

Figure 8

Distribution of Simulated Filament Thresholds Histogram of filament intensity thresholds



T_{98} obtained from simulated realizations. The solid red line indicates the observed SDSS value, while the dashed red line marks the simulation mean. The spread reflects realization-to-realization variability in filament strength.

The observed SDSS value,

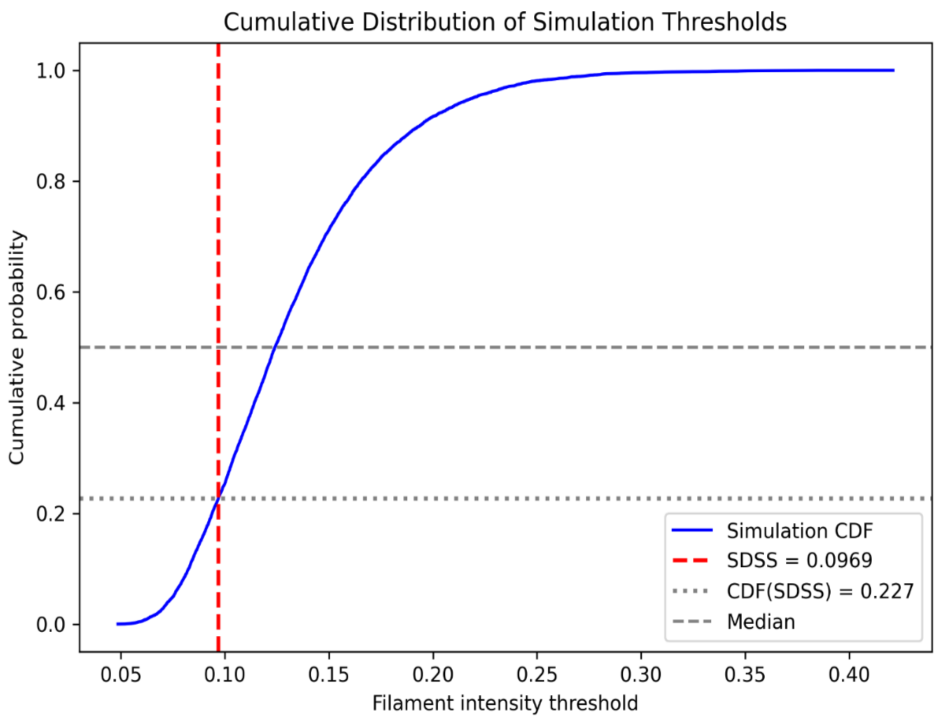
$$T_{98}^{\mathrm{obs}}=0.0969,$$

is indicated as a vertical line in Figure 8; this value lies below the simulation mean and lies approximately 0.78 standard deviations below the simulation mean, remaining well within the expected statistical dispersion.
To quantify the statistical significance of this comparison, we compute the cumulative distribution function (CDF) of the simulated threshold values, shown in Figure 9. Evaluating the CDF at the observed value yields

$$P(T_{98}^{\mathrm{sim}}\leq T_{98}^{\mathrm{obs}})\approx 0.227$$

indicating that approximately 23% of the simulated realizations produce threshold values equal to or lower than the observed one.

Figure 9
Cumulative Distribution of Simulated Thresholds
Cumulative distribution function of simulated filament intensity thresholds. The solid red line indicates the observed SDSS value, and the corresponding cumulative probability is shown.



A quantitative summary of these results is provided in Table 1. The corresponding Z-score, $Z \approx -0.78$, and empirical p-value, $p \approx 0.42$, confirm that the observed threshold is statistically consistent with the distribution predicted by the model.

By construction, the filament fraction associated with percentile thresholding is fixed and therefore does not provide independent statistical information. The primary observable in this analysis is the threshold value T_{98}, which encodes the relative strength of filamentary structures.

Table 1

Summary of observational and simulated filament statistics.
The table compares the filament fraction and percentile-based intensity threshold measured from SDSS data with the distribution obtained from 10 000 simulated realizations. The Z-score and empirical p-value quantify the statistical consistency between observation and model.

Quantity	Observed (SDSS)	Simulations (Mean ± Std)
Number of galaxies	112 630	112 630
Redshift range	0.06 < z < 0.09	0.06 < z < 0.09
Percentile threshold	98%	98%
Filament fraction	0.020017	0.020017 ± 0.000000
Filament intensity threshold T_{98}	0.096919	0.132385 ± 0.045288
Z-score (threshold)	—	-0.7831
Empirical p-value	—	0.4157
CDF(SDSS)	—	0.2271
Number of simulations	—	10 000

By construction, the filament fraction is fixed by the percentile definition and therefore does not provide independent statistical information. The primary observable in this analysis is the threshold value T_{98}, which varies across realizations and encodes the relative strength of filamentary structures.

To complement the percentile-based analysis, we fix the filament intensity threshold to the observed SDSS value and measure the fraction of the field exceeding this threshold. In this formulation, the observable corresponds to the spatial extent of high-intensity filamentary regions under a fixed physical criterion.

The distribution of fixed-threshold filament fractions obtained from the simulations is shown in Figure 10. Each value represents the fraction of pixels with filament intensity greater than the fixed threshold T_{obs}=0.0969in an individual realization. The simulated distribution has a mean value of

$$\langle F_{\mathrm{sim}} \rangle \approx 0.0299$$

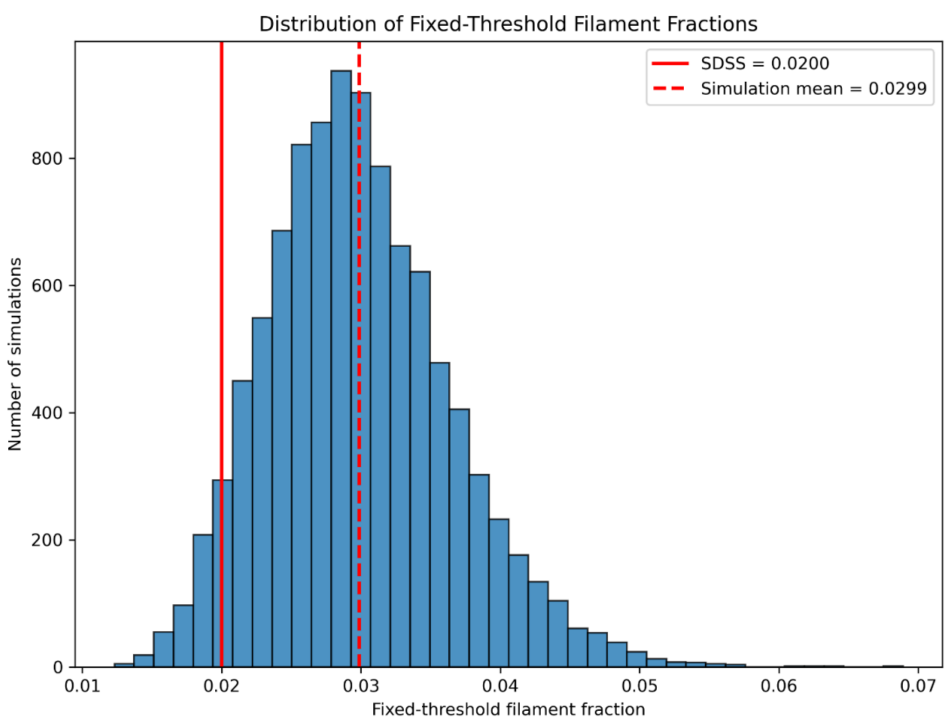
and a standard deviation of

$$\sigma_F \approx 0.0066.$$

Figure 10

Distribution of Fixed-Threshold Filament Fractions

Histogram of filament fractions obtained from simulated realizations using a fixed threshold equal to the SDSS value. The solid red line indicates the observed SDSS fraction, while the dashed red line marks the simulation mean. The displacement of the observed value toward the lower tail indicates reduced spatial coverage of high-intensity filamentary regions.



The observed SDSS value,

$$F_{\text{obs}} \approx 0.0200,$$

is indicated as a vertical line in Figure 10. This value lies below the mean of the simulated distribution and is located toward its lower tail.
To quantify this deviation, we compute the cumulative distribution function of the fixed-threshold filament fractions, shown in Figure 11.

Evaluating the CDF at the observed value yields

$$P(F_{\text{sim}} \leq F_{\text{obs}}) \approx 0.049,$$

indicating that approximately 5% of simulated realizations produce filament fractions equal to or lower than the observed value.

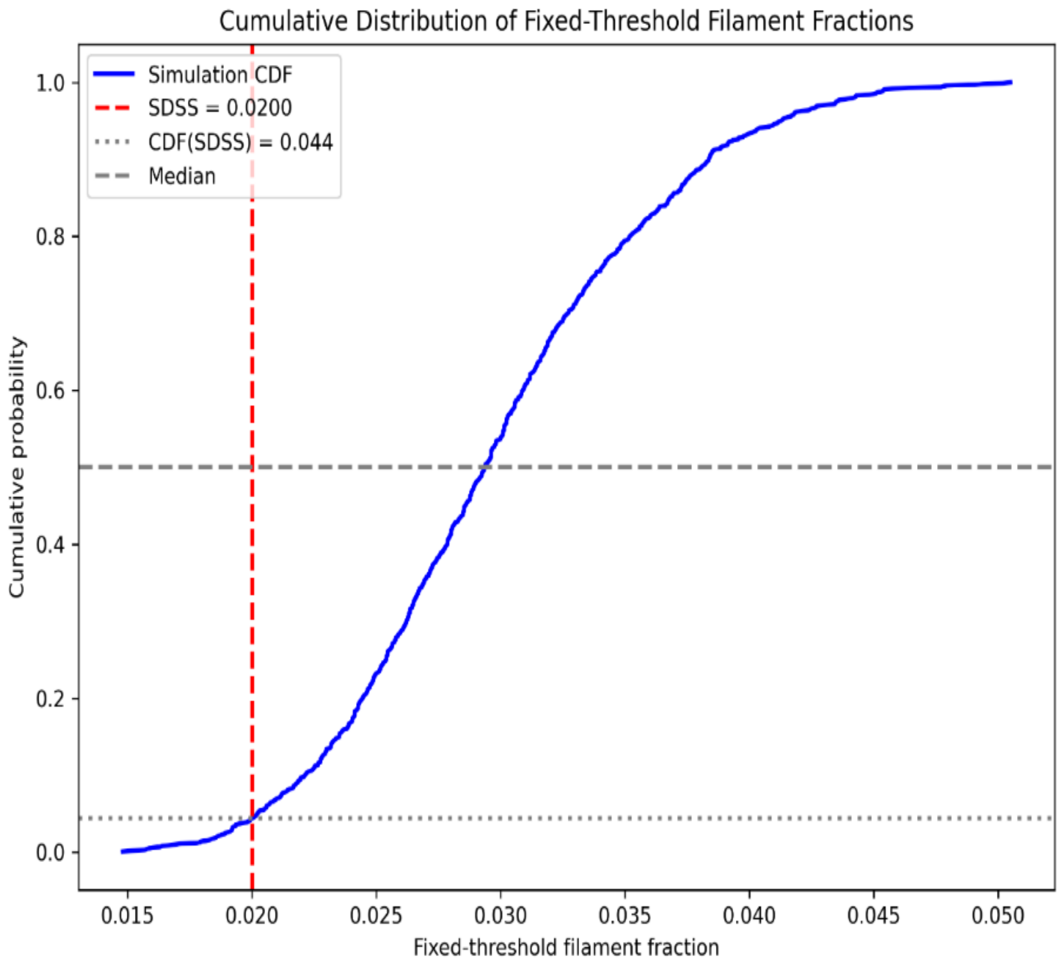
Interpreted as a one-sided probability, this places the observation near the lower tail of the distribution. The corresponding two-sided p-value is approximately $p \approx 0.098$, indicating that the deviation is not statistically significant at conventional confidence levels. A quantitative summary of these results is provided in Table 2.

Table 2
Summary of fixed-threshold filament fraction statistics

Quantity	Observed (SDSS)	Simulations (Mean ± Std)
Number of galaxies	112 630	112 630
Redshift range	0.06 < z < 0.09	0.06 < z < 0.09
Fixed threshold T_{obs}	0.096919	0.096919
Filament fraction $F(T_{\text{obs}})$	0.020017	0.029875 ± 0.006584
Z-score	—	-1.4973
Empirical p-value	—	0.0982
CDF(SDSS)	—	0.0491
Number of simulations	—	10 000

Figure 11

Cumulative Distribution of Fixed-Threshold Filament Fractions
Cumulative distribution function of simulated filament fractions under the fixed-threshold definition. The solid red line indicates the observed SDSS value, and the corresponding cumulative probability is shown.



As a final summary of the main results, a compact comparison of both observables is provided in Table 3

Table 3
Summary of filament observables.

Observable	SDSS (Observed)	Simulations (Mean ± Std)	Z-score	CDF(SDSS)	p-value
Threshold T_{98}	0.0969	0.1324 ± 0.0453	-0.78	0.227	0.416
F(T_{obs})	0.0200	0.0299 ± 0.0066	-1.50	0.049	0.098

DISCUSSION

The results presented in Section 5 indicate that the observed filament intensity threshold T_{98}^{obs} is statistically consistent with the distribution obtained from the ensemble of simulated realizations. The observed value lies within one standard deviation of the simulation mean and corresponds to a cumulative probability of approximately 23%, indicating that similar or lower filament strengths occur in a substantial fraction of the simulated universes.

This agreement suggests that the filamentary structure observed in the SDSS dataset is compatible with expectations from a stochastic model constructed to approximate large-scale structure within a simplified ΛCDM framework. From a physical perspective, this reflects the fact that filamentary structures arise naturally from the anisotropic gravitational collapse of initial density perturbations, forming the interconnected cosmic web of filaments and nodes (Bond, Kofman, & Pogosyan, 1996; Colberg et al., 2005)

Although the observed value lies within the expected dispersion, it is positioned in the lower half of the simulated distribution, suggesting slightly weaker filamentary structure in the SDSS sample compared to the model mean. While not statistically significant, this trend may reflect observational effects such as survey geometry, selection biases, and incomplete sampling, all of which can reduce the apparent contrast of filamentary structures in projected data.

A complementary perspective is obtained by fixing the filament intensity threshold to the observed value and measuring the fraction of the field exceeding this threshold. In this case, the observed filament fraction lies close to the lower tail of the simulated distribution, with a cumulative probability of approximately 5%, corresponding to a two-sided p-value of approximately 0.10. This indicates that only a small fraction of simulated realizations produces filamentary regions that are as spatially limited as those observed in the SDSS data.

Although this deviation does not reach conventional levels of statistical significance, it reveals a systematic trend that is not captured by the percentile-based observable alone. While the intensity-based measure is fully consistent with the model, the extent-based observable indicates that high-intensity filamentary regions in the SDSS data are less spatially extended than in typical simulated realizations.

This asymmetry suggests that, although the amplitude of filamentary structures is well reproduced, their spatial connectivity may differ slightly between observations and the simplified model. This discrepancy likely reflects both limitations of the stochastic modeling approach, which appears to reproduce filament amplitudes more accurately than their topology, and the effects of projecting the three-dimensional galaxy distribution onto a two-dimensional field.

The contrast between these two analyses highlights the importance of observable definition in the statistical characterization of filamentary structure. The percentile-based threshold probes the intensity scale of filamentary features, while the fixed-threshold fraction probes their spatial extent. The combination of these observables provides a more complete description of the filamentary network and reveals differences that are not apparent when considering either measure alone.

It is important to recognize the limitations of the simulation framework adopted in this work. The synthetic realizations are based on simplified statistical models and do not fully capture the nonlinear dynamics of structure formation or the detailed physics of galaxy formation. In addition, the analysis is performed on a projected two-dimensional density field, which compresses three-dimensional information along the line of sight. The projection of the three-dimensional galaxy distribution onto a two-dimensional density field introduces line-of-sight mixing, whereby physically distinct structures overlap in the projected map. This effect acts as a smoothing mechanism that reduces contrast and preferentially suppresses thin, elongated filamentary features.

This projection-induced smoothing affects the two observables differently. The percentile-based threshold T_{98}, which depends on the upper tail of the intensity distribution, remains relatively robust to projection effects, as it is dominated by the highest-density regions. In contrast, the fixed-threshold filament fraction $F_{(Tobs)}$, which measures the spatial extent of high-intensity regions, is more sensitive to this smoothing. As a result, projection effects are expected to reduce the measured filament fraction, biasing it toward lower values.

This provides a natural explanation for the tendency of the SDSS filament fraction to lie near the lower tail of the simulated distribution, without requiring a fundamental discrepancy between observations and the underlying cosmological model. The choice of the redshift interval $0.06<z<0.09$ partially mitigates this effect by limiting projection depth, although some loss of structural information remains unavoidable.

Finally, the dependence of the results on the smoothing scale σ was not explored and represents an additional source of systematic uncertainty. Despite these limitations, the methodology developed here provides a computationally efficient and statistically robust framework for quantifying filamentary structure in galaxy surveys, enabling direct comparison between observations and simulations in real space (Peacock, 1999).

These results therefore emphasize that simplified statistical models can capture the amplitude of filamentary structures, while more detailed simulations may be required to fully reproduce their spatial connectivity and topology.

CONCLUSIONS

This work presented a statistical analysis of filamentary structure in the large-scale distribution of galaxies using observational data from the Sloan Digital Sky Survey and a suite of synthetic cosmological simulations. By constructing a smoothed density field and applying a morphology-based filament detection method, we defined a continuous filament intensity field and introduced scalar observables to quantify filamentary structure.

Two complementary observables were considered. The first, based on percentile thresholding, measures the intensity scale required to select the most prominent filamentary regions. The second, based on a fixed threshold, measures the spatial extent of high-intensity filamentary structures. Together, these observables provide a more complete characterization of the filamentary network.

The percentile-based analysis shows that the observed threshold T_{98} lies within the expected dispersion of the simulations, with a Z-score of approximately -0.78 and an empirical p-value of ≈ 0.42. This indicates that the overall strength of filamentary structures in the SDSS data is statistically consistent with predictions from a simplified ΛCDM-inspired model.

The fixed-threshold analysis reveals a complementary trend. When evaluated at a constant intensity threshold, the observed filament fraction lies near the lower tail of the simulated distribution, with a cumulative probability of approximately 5%. This suggests that the spatial extent of high-intensity filamentary regions in the observed data is somewhat reduced relative to the model predictions, although not at a level of strong statistical significance.

Taken together, these results support the interpretation that large-scale filamentary structure arises naturally from gravitational structure formation within the ΛCDM paradigm, while also highlighting the sensitivity of filament observables to the choice of statistical definition. The combination of intensity-based and extent-based measures provides a more nuanced view of the cosmic web than either observable alone.

The methodology developed in this work provides a computationally efficient and flexible framework for analyzing filamentary structure in galaxy surveys. While the current implementation relies on simplified simulations and two-dimensional projections, it can be naturally extended to more realistic three-dimensional analyses and higher-fidelity cosmological models. These results highlight the importance of combining multiple statistical observables when characterizing the cosmic web, as different definitions probe distinct physical aspects of large-scale structure.
Future work could explore the dependence of filament observables on redshift, smoothing scale, and survey completeness, as well as extend the analysis to larger datasets and more realistic simulations. Such developments would enable more stringent tests of structure formation and further clarify the role of filamentary morphology as a probe of cosmology.

BIBLIOGRAPHIC REFERENCES

Aragón-Calvo, M. A., Jones, B. J. T., van de Weygaert, R., & van der Hulst, J. M. (2007). The multiscale morphology filter: Identifying and extracting spatial patterns in the galaxy distribution. *Astronomy & Astrophysics, 474*(1), 315–338. https://doi.org/10.1051/0004-6361:20077644

Bond, J. R., Kofman, L., & Pogosyan, D. (1996). How filaments of galaxies are woven into the cosmic web. *Nature, 380*(6575), 603–606. https://doi.org/10.1038/380603a0

Cautun, M., van de Weygaert, R., Jones, B. J. T., & Frenk, C. S. (2013). Nexus: Tracing the cosmic web connection. *Monthly Notices of the Royal Astronomical Society, 429*(2), 1286–1308. https://doi.org/10.1093/mnras/sts441

Colberg, J. M., Krughoff, K. S., & Connolly, A. J. (2005). Intercluster filaments in a ΛCDM universe. Monthly Notices of the Royal Astronomical Society, 359(1), 272–282. https://doi.org/10.1111/j.1365-2966.2005.08897.x

Coles, P., & Jones, B. (1991). A lognormal model for the cosmological mass distribution. *Monthly Notices of the Royal Astronomical Society, 248*(1), 1–13. https://doi.org/10.1093/mnras/248.1.1

Dodelson, S. (2003). *Modern cosmology.* Academic Press.

Hamilton, A. J. S. (1998). Linear redshift distortions: A review. In D. Hamilton (Ed.), *The evolving universe* (pp. 185–275). Springer. https://doi.org/10.1007/978-94-011-4960-0_17

Harris, C. R., Millman, K. J., van der Walt, S. J., et al. (2020). Array programming with NumPy. *Nature, 585*, 357–362. https://doi.org/10.1038/s41586-020-2649-2

Hogg, D. W. (1999). Distance measures in cosmology. *arXiv:astro-ph/9905116*. https://arxiv.org/abs/astro-ph/9905116

Hunter, J. D. (2007). Matplotlib: A 2D graphics environment. *Computing in Science & Engineering, 9*(3), 90–95. https://doi.org/10.1109/MCSE.2007.55

Martínez, V. J., & Saar, E. (2002). *Statistics of the galaxy distribution.* Chapman & Hall/CRC.

Mukhanov, V. (2005). *Physical foundations of cosmology.* Cambridge University Press.

Peacock, J. A. (1999). *Cosmological physics.* Cambridge University Press.

Peebles, P. J. E. (1980). *The large-scale structure of the universe.* Princeton University Press.

Planck Collaboration. (2018). Planck 2018 results. VI. Cosmological parameters. *Astronomy & Astrophysics, 641*, A6. https://doi.org/10.1051/0004-6361/201833910

Python Software Foundation. (2023). Python language reference, version 3.x. https://www.python.org

Sousbie, T. (2011). The persistent cosmic web and its filamentary structure: Theory and implementations. *Monthly Notices of the Royal Astronomical Society, 414*(1), 350–383. https://doi.org/10.1111/j.1365-2966.2011.18394.x

Sloan Digital Sky Survey Collaboration. (2023). SDSS SkyServer and CasJobs. https://skyserver.sdss.org/CasJobs/

Springel, V., White, S. D. M., Jenkins, A., et al. (2005). Simulations of the formation, evolution and clustering of galaxies and quasars. *Nature, 435*(7042), 629–636. https://doi.org/10.1038/nature03597

Virtanen, P., Gommers, R., Oliphant, T. E., et al. (2020). SciPy 1.0: Fundamental algorithms for scientific computing in Python. *Nature Methods, 17*, 261–272. https://doi.org/10.1038/s41592-019-0686-2

Wall, J. V., & Jenkins, C. R. (2012). *Practical statistics for astronomers* (2nd ed.). Cambridge University Press.

York, D. G., Adelman, J., Anderson, J. E., et al. (2000). The Sloan Digital Sky Survey: Technical summary. *The Astronomical Journal, 120*(3), 1579–1587. https://doi.org/10.1086/301513