



NEXUS: tracing the cosmic web connection

Marius Cautun,[★] Rien van de Weygaert and Bernard J. T. Jones

Kapteyn Astronomical Institute, University of Groningen, PO Box 800, NL-9747 AV Groningen, the Netherlands

Accepted 2012 November 14. Received 2012 November 13; in original form 2012 September 10

ABSTRACT

We introduce the NEXUS algorithm for the identification of cosmic web environments: clusters, filaments, walls and voids. This is a multiscale and automatic morphological analysis tool that identifies all the cosmic structures in a scale free way, without preference for a certain size or shape. We develop the NEXUS method to incorporate the density, tidal field, velocity divergence and velocity shear as tracers of the cosmic web. We also present the NEXUS+ procedure which, taking advantage of a novel filtering of the density in logarithmic space, is very successful at identifying the filament and wall environments in a robust and natural way.

To assess the algorithms we apply them to an N -body simulation. We find that all methods correctly identify the most prominent filaments and walls, while there are differences in the detection of the more tenuous structures. In general, the structures traced by the density and tidal fields are clumpier and more rugged than those present in the velocity divergence and velocity shear fields. We find that the NEXUS+ method captures much better the filamentary and wall networks and is successful in detecting even the fainter structures. We also confirm the efficiency of our methods by examining the dark matter particle and halo distributions.

Key words: methods: data analysis – techniques: image processing – cosmology: theory – large-scale structure of Universe.

1 INTRODUCTION

Early attempts to map the large-scale distribution of galaxies in the universe (Gregory & Thompson 1978; de Lapparent Geller & Huchra 1986; Geller & Huchra 1989; Shectman et al. 1996) revealed that galaxies are far from being evenly distributed across the nearby Universe. On the contrary, the mass distribution delineated by galaxies seems to form an intricate network of compact and dense associations interconnected by tenuous ‘bridges’ or ‘filaments’ surrounded by surprisingly vast empty regions (Kirshner et al. 1981). Preliminary studies suggested that the universe on large scales could be described as a cellular system (Jøeuvre & Einasto 1978) or a *cosmic web* (Bond, Kofman & Pogosyan 1996). This has been confirmed in recent times by large galaxy surveys such as the 2dF Galaxy Redshift Survey (2dFGRS; Colless et al. 2003), the Sloan Digital Sky Survey (SDSS; e.g. Tegmark et al. 2004) and the Two Micron All Sky Survey (2MASS) redshift survey (Huchra et al. 2005).

The cosmic web can be seen as the most prominent manifestation of the anisotropic nature of gravitational collapse, the motor behind the formation of structure in the cosmos (Peebles 1980). N -body computer simulations have illustrated how a primordial field of tiny Gaussian density perturbations transforms into a pronounced and

intricate filigree of filamentary features, dented by dense compact clumps at the nodes of the network (Jenkins et al. 1998; Colberg, Krughoff & Connolly 2005; Springel et al. 2005; Dolag et al. 2006). The description of the megaparsec matter distribution as an interconnected network or a cosmic web is not a coincidence. Even early computer simulations indicated the close connection between each morphological component, namely that clusters sit at the intersection of filaments and filaments are formed at the intersection of walls (Doroshkevich et al. 1980; Melott 1983; Shapiro, Struck-Marcell & Melott 1983; Pauls & Melott 1995; Sathyaprakash, Sahni & Shandarin 1996).

One of the main reasons for our interest in outlining the cosmic web concerns the question whether and how far the web-like environment influences the properties and evolution of galaxies. Recent N -body simulations have found that the filamentary or sheet-like nature of the environment has a distinct influence on the shape and spin orientation of dark matter haloes (Aragón-Calvo et al. 2007a; Hahn et al. 2007a,b, 2009; Paz, Stasyszyn & Padilla 2008; Hahn 2009; Zhang et al. 2009). Other recent works (Jones, van de Weygaert & Aragón-Calvo 2010; Tempel, Stoica & Saar 2012) have shown that indeed there is an alignment, even though weak, of galaxies and the filaments they lie within. In this paper we propose new robust and flexible methods that allow a better identification of the cosmic web environments and hence help us to better understand how environments influence the formation and evolution of dark matter haloes and galaxies.

[★]E-mail: cautun@astro.rug.nl

1.1 Cosmic web detection

Identifying the components of the cosmic web is a major challenge due to the overwhelming complexity of the individual structures as well as their connectivity, the lack of structural symmetries, its intrinsic multiscale nature and the wide range of densities found in the cosmic matter distribution. Over the years, a variety of heuristic measures were forwarded to analyse specific aspects of the spatial patterns in the large-scale Universe, but only recently these have lead to a more solid and well-defined machinery for identifying the cosmic web. Nearly without exception, these methods borrow extensively from other branches of science such as image processing, mathematical morphology, computational geometry and medical imaging.

The connectedness of elongated supercluster structures in the cosmic matter distribution was first probed by means of percolation analysis, introduced and emphasized by Zel'dovich and coworkers (Zeldovich, Einasto & Shandarin 1982; Shandarin & Zeldovich 1989; Shandarin, Sheth & Sahni 2004; Shandarin, Habib & Heitmann 2010), while a related graph-theoretical construct, the minimum spanning tree of the galaxy distribution, was extensively analysed by Bhavsar and collaborators (Barrow, Bhavsar & Sonoda 1985; Graham & Clowes 1995; Colberg 2007) in an attempt to develop an objective measure of filamentarity. Both Colberg et al. (2005) and Pimbblet (2005) set out to identify filaments and their adjoining clusters, using quite different techniques.

More general filament finders have been put forward by a number of authors. Following specific physical criteria, González & Padilla (2010) recently forwarded an interesting and promising combination of a tessellation-based density estimator and a dynamical binding energy criterion. A thorough mathematical non-parametric formalism involving the medial axis of a point cloud, as yet for 2D point distributions, was proposed by Genovese et al. (2010). It is based on a geometric representation of filaments as the medial axis of the data distribution. Also solidly rooted within a geometric and mathematical context is the more generic geometric inference formalism developed by Chazal, Cohen-Steiner & Mérigot (2009). It allows the recovery of geometric and topological features of the supposedly underlying density field from a sampled point cloud on the basis of distance functions. Stoica et al. (2005) and Stoica, Martínez & Saar (2007, 2010) use a generalization of the classical Candy model to locate and catalogue filaments in galaxy surveys. This approach has the advantage that it works directly with the original point process and does not require the creation of a continuous density field. However, computationally it is very demanding.

The more recent formalisms that are intent on characterizing the full range of web-like formalisms usually exploit the morphological information in the gradient and Hessian of the density field or potential field, i.e. the tidal field (see e.g. Aragón-Calvo et al. 2007a,b; Hahn et al. 2007a,b; Sousbie et al. 2008a; Forero-Romero et al. 2009; Bond, Strauss & Cen 2010a,b). Morse theory (see Colombi, Pogosyan & Souradeep 2000) forms the basis of the *skeleton analysis* by Novikov, Colombi & Doré (2006) (2D) and Sousbie et al. (2008a) (3D). It identifies morphological features with the maxima and saddle points in the density field and results in an elegant and mathematically rigorous tool for filament identification. However, it is computationally intensive, focuses mostly on filaments and is strongly dependent on the smoothing scale of the density field. A more elaborate classification scheme on the basis of the manifolds in the tidal field – involving all morphological

features in the cosmic matter distribution – has been forwarded by Hahn et al. (2007a) (also see Hahn et al. 2007b; Forero-Romero et al. 2009).

Instead of using the tidal field configuration, one may also try to link directly to the morphology of the density field itself. Though this allows a more detailed view of the intricacies of the multiscale matter distribution, it is usually more sensitive to noise and less directly coupled to the underlying dynamics of structure formation than the tidal field morphology. A single-scale dissection of the density field into its various morphological components has been done by Bond et al. (2010a), and applied to N -body simulations and galaxy redshift samples (also see Bond et al. 2010b; Choi et al. 2010). A more elaborate formalism is the multiscale morphology filter (MMF), introduced by Aragón-Calvo et al. (2007b). It looks at structure from a scale-space point of view, treating the spatial structure in D dimensions in the context of an explicit $D + 1$ dimensional space (Florack et al. 1992; Lindeberg 1998). The $D + 1$ dimensional scale-space consists of the D -dimensional spatial structure at a range of spatial resolution scales. The MMF formalism subsequently assigns a local morphology based on an evaluation of the multiscale second-order variations in the local density field. Instead of restricting the analysis to one particular scale, by evaluating the density field Hessian over a range of spatial scales and determining at which scales and locations the various morphological signatures are most prominent, the MMF explicitly addresses the multiscale nature of the cosmic structures. A somewhat similar multiscale approach was followed by the metric space technique described by Wu, Batuski & Khalil (2009), who applied it to a morphological analysis of the Data Release 5 (DR5) of the SDSS.

A more recent development is that of the SpineWeb procedure (Aragón-Calvo et al. 2010a), which traces the various features of the cosmic web on pure topological grounds by invoking the Watershed Transform (WT). The WT was introduced by Platen, van de Weygaert & Jones (2007) as the basis of the Watershed Void Finder (WVF) technique (Platen et al. 2007) which identifies cosmic voids with the watershed basins. The SpineWeb procedure elaborates on this, by identifying the central axis of filaments and the inner plane of walls with the boundaries between the watershed segments of the density field. While the basic SpineWeb procedure involved one scale, the full procedure allows a multiscale topological characterization of the cosmic web (Aragón-Calvo et al. 2010b). However, to do so it must invoke some implicit assumptions on the connectivity of the various topological features.

1.2 Intention and outline

The goal of this paper is to present two new algorithms (NEXUS and NEXUS+) for the detection of cosmic web environments and to assess their effectiveness. Elaborating on the multiscale scale-space context of the rudimentary density field MMF (Aragón-Calvo 2007), the NEXUS and NEXUS+ formalism represent a complete and versatile instrument for the structural and physical study of the cosmic web. Simultaneously taking into account the multiscale nature of the cosmic mass distribution, NEXUS and NEXUS+ explicitly operate on a diversity of physical fields that are relevant to the formation and evolution of the cosmic web. The extension beyond the density field, towards the use of information contained in the tidal field, velocity divergence and velocity shear to trace the large-scale structure, is a key aspect of NEXUS and

NEXUS+. The new formalism allows us to compare the environments traced by both the positional and velocity part of the phase space. We focus most of the analysis on the detection of filaments and walls, since these are the most challenging environments to identify. We find that our new methods are very efficient at tracing the cosmic web, resulting in very high quality filaments and walls.

This study is organized as follows. In Sections 2 and 3 we describe the NEXUS and NEXUS+ methods which we use for the identification of the cosmic web, including a comparison of the two algorithms on a toy model. This is followed in Section 4 with an extension of the tools to use a multitude of cosmological fields (density and tidal field versus the velocity divergence and velocity shear) as tracers of the cosmic environments. This way we take full advantage of the full 6D information contained in phase space.

The second part of the paper is focused on assessing how these methods cope with the complex and hierarchical structures present in the universe. To do so we use the Delaunay Tessellation Field Estimator (DTFE) density and velocity divergence from N -body simulations as inputs to our algorithms – see Section 5. Section 6 presents the cluster, filament and wall environments identified in the simulation and compares the results of the different methods. To assess the quality of the detections we look at the dark matter particle and halo distributions in each environment and also study the effects of a multiscale versus single-scale approach in tracing the cosmic web. Finally, we summarize our findings in Section 8.

2 NEXUS : GENERAL FORMALISM FOR MULTISCALE MORPHOLOGICAL ANALYSIS

The NEXUS algorithm is a scale-space method for morphologically segmenting the cosmic web into its three distinct features: clusters, filaments and walls. The environment identification is performed in a scale and user independent way to account for the multiscale nature of the large-scale structure, which is the result of the hierarchical evolution of the cosmic mass distribution. The method is derived from the field of medical imaging (Frangi et al. 1998; Sato et al. 1998; Li, Sone & Doi 2003) where it is used to identify nodules, vessels and walls in two- and three-dimensional images. An earlier and simpler version of the method was introduced in cosmology under the name MMF in Aragón-Calvo et al. (2007b). The MMF involved a rudimentary treatment of scale-space analysis and restricted itself to the use of the density field as tracer of the cosmic web environment.

In general, the scale-space formalism can be applied to any input field to detect point-, line- and sheet-like structures in the field values, and thus lends itself to applications involving a range of quantities dynamically relevant for the formation and evolution of the cosmic web. Following this observation, we have embedded the scale-space formalism in the physical framework of cosmic web formation.

For this purpose we have defined two classes of the algorithm, the NEXUS and the NEXUS+ formalism. The main difference between NEXUS and NEXUS+ concerns the filter used for constructing the representation of the field at different resolutions in scale-space. The NEXUS technique uses a Gaussian filter for smoothing while NEXUS+ uses a log-Gaussian filter (more on that in Section 3). We will demonstrate in this study that they yield a substantially more realistic and robust representation of filaments, walls and their mutual connectivity, over the range of scales covered by the scale-space representation.

2.1 NEXUS: general algorithm description

The NEXUS algorithm detects the point-, line- and sheet-like structures¹ for a generic input field f . For large-scale structure, these features correspond to clusters, filaments and walls. To keep the notations clear, we limit our discussion to the cosmic web environments, but there is no loss of generality. The NEXUS algorithm consists of the following six steps.

- (i) Applying a Gaussian filter of width R_n to the input field.
- (ii) Computing the Hessian matrix eigenvalues for the filtered field.
- (iii) Assigning to each point a cluster, filament and wall signature using the Hessian eigenvalues.
- (iv) Repeating steps (I) to (III) over a range of smoothing scales (R_0, R_1, \dots, R_N) to construct the scale-space representation of the field.
- (v) Combining the results of all scales to obtain a scale-independent cluster, filament and wall signature.
- (vi) Using physical criteria to determine the detection threshold corresponding to valid environments.

In the following we elaborate on each step of the algorithm and give the details necessary for the implementation of the method.

2.1.1 Step I: applying Gaussian smoothing

A Gaussian filter of width R_n is applied to the input field $f(\mathbf{x})$. This gives rise to a smoothed field $f_{R_n}(\mathbf{x})$ given by

$$f_{R_n}(\mathbf{x}) = \int \frac{d^3k}{(2\pi)^3} e^{-k^2 R_n^2 / 2} \hat{f}(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{x}}, \quad (1)$$

where $\hat{f}(\mathbf{k})$ is the Fourier transform of the input field $f(\mathbf{x})$.

2.1.2 Step II: computing Hessian eigenvalues

The Hessian of the filtered field is computed as

$$\mathbf{H}_{ij,R_n}(\mathbf{x}) = R_n^2 \frac{\partial^2 f_{R_n}(\mathbf{x})}{\partial x_i \partial x_j}, \quad (2)$$

where \mathbf{H}_{ij,R_n} represents the i,j entry of the \mathbf{H}_{R_n} Hessian matrix. The R_n^2 term is a renormalization factor that has to do with the multiscale nature of the NEXUS algorithm. It makes sure that the same weight is assigned when comparing the Hessian value at different scales. Using equation (1), the Fourier transform of the Hessian reads

$$\hat{\mathbf{H}}_{ij,R_n}(\mathbf{k}) = \mathbf{H}_{ij,R_n}(\mathbf{k}) \hat{f}(\mathbf{k}), \quad (3)$$

with \mathbf{H} the Hessian kernel function given by

$$\mathbf{H}_{ij,R_n}(\mathbf{k}) = -k_i k_j R_n^2 e^{-k^2 R_n^2 / 2}. \quad (4)$$

The kernel function characterizes which Fourier components of the input field give contributions to the Hessian matrix. The dependence of the Hessian kernel on k is shown in Fig. 1. At a given smoothing scale R_n , the \mathbf{H} kernel has a peak at $k_{\text{peak}} = \sqrt{2}/R_n$, with a sharp drop-off for higher k and a linear fall for smaller k . Therefore, for a given scale R_n , only the Fourier components of f

¹ The algorithm given here applies to detecting the point-, line- and sheet-like structures corresponding to maxima in the field values. If we are interested in the same structures but for the minima of the field values, than we need to apply the same algorithm to $-f$. For example the cosmic web environments correspond to maxima in density but to minima in velocity divergence.

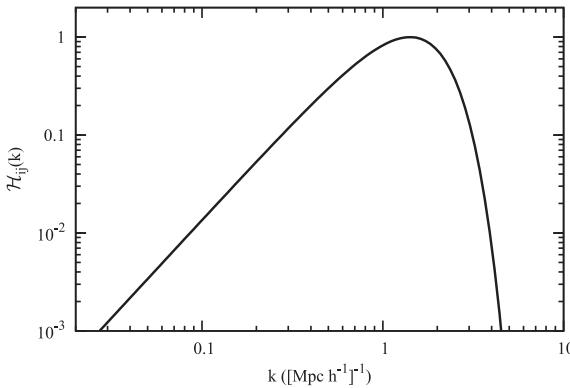


Figure 1. The amplitude of the kernel function H_{ij} for a smoothing radius of $1 h^{-1} \text{ Mpc}$. The values were normalized such that H_{ij} has a maximum value of 1. Note the logarithmic axes.

around the peak k_{peak} give an important contribution to the Hessian matrix.

The NEXUS formalism depends only on the eigenvalues of the Hessian matrix, eigenvalues given by

$$\det(\mathbf{H}_{ij, R_n}(\mathbf{x}) - \lambda_{a, R_n}(\mathbf{x})\delta_{ij}) = 0, \quad \text{with } \lambda_1 \leq \lambda_2 \leq \lambda_3. \quad (5)$$

2.1.3 Step III: computing environment signature

The eigenvalues of the Hessian matrix can be used to assign a cluster, filament and wall characteristic to every point \mathbf{x} using the expected behaviour given in Table 1. This is the environment signature and is denoted with $\mathcal{S}(\mathbf{x})$. The first step in computing the signature is to define the shape strength \mathcal{I} . This gives a quantitative description of the approximate relations given in the middle column of Table 1. The shape strength is defined as

$$\mathcal{I} = \begin{cases} \left| \frac{\lambda_3}{\lambda_1} \right| & \text{cluster}, \\ \left| \frac{\lambda_2}{\lambda_1} \right| \Theta \left(1 - \left| \frac{\lambda_3}{\lambda_1} \right| \right) & \text{filament}, \\ \Theta \left(1 - \left| \frac{\lambda_2}{\lambda_1} \right| \right) \Theta \left(1 - \left| \frac{\lambda_3}{\lambda_1} \right| \right) & \text{wall}, \end{cases} \quad (6)$$

where we use the notation $\Theta(x) = x\theta(x)$ for clarity, with $\theta(x)$ the step function ($\theta(x) = 1$ if $x \geq 0$, 0 otherwise). The strength \mathcal{I} is large when the eigenvalues at \mathbf{x} correspond to a prominent structure and small otherwise. The cluster/filament/wall signature is defined as

$$\mathcal{S} = \mathcal{I} \times \begin{cases} |\lambda_3| \theta(-\lambda_1)\theta(-\lambda_2)\theta(-\lambda_3) & \text{cluster}, \\ |\lambda_2| \theta(-\lambda_1)\theta(-\lambda_2) & \text{filament}, \\ |\lambda_1| \theta(-\lambda_1) & \text{wall}, \end{cases} \quad (7)$$

Table 1. Hessian eigenvalue relationships for the different environments of the cosmic web. The second column gives the qualitative relationships between the eigenvalues (conditions that are implemented analytically in equation 6) while the third column gives strict eigenvalues constraints implemented in equation (7).

Structure	Soft constraints	Strict constraints
Cluster	$ \lambda_1 \simeq \lambda_2 \simeq \lambda_3 $	$\lambda_1 < 0; \lambda_2 < 0; \lambda_3 < 0$
Filament	$ \lambda_1 \simeq \lambda_2 \gg \lambda_3 $	$\lambda_1 < 0; \lambda_2 < 0$
Wall	$ \lambda_1 \gg \lambda_2 ; \lambda_1 \gg \lambda_3 $	$\lambda_1 < 0$

where the $\theta(-\lambda_a)$ factors (with $a = 1, 2, 3$) incorporate the right-most column of Table 1. The $|\lambda_a|$ term gives the intensity of the morphological feature and can be used to discriminate between real signals (large $|\lambda_a|$) and noise (small $|\lambda_a|$).

2.1.4 Step IV: computing the environmental signature over a range of smoothing scales

The previous three steps are repeated over a range of smoothing scales (R_0, R_1, \dots, R_N). The hierarchy of smoothing scales is taken as $R_n = (\sqrt{2})^n R_0$ with R_0 the smallest scale at which one expects to find structures (Sato et al. 1998). Taking an even smaller step between any two successive smoothing scales makes only minor differences. In practice we choose R_0 equal to the grid spacing of the input field. We found that for the detection of the most prominent features of the cosmic web it is sufficient to consider smoothing scales in the range 0.5 – $4 h^{-1} \text{ Mpc}$. In this respect it is good to note that outstanding features of the cosmic web are visible within a particular range of scales centred around the transition scale between linear and non-linear structures.

The result of this step is a signature function for each scale $\mathcal{S}_{R_n}(\mathbf{x})$ which characterizes the environmental response of point \mathbf{x} at the R_n smoothing scale.

2.1.5 Step V: scale-space stacking

The signature of the given set of scales is combined to obtain the overall signature. This is a scale-independent map characterizing the degree to which the point \mathbf{x} is part of a cluster, filament or wall. A structure of a given size will give the largest signature for a smoothing scale of the same size. Therefore, the overall signature at a point is the maximum signature over all the scales:

$$\mathcal{S}(\mathbf{x}) = \max_{\text{levels } n} \mathcal{S}_{R_n}(\mathbf{x}). \quad (8)$$

2.1.6 Step VI: computing the detection threshold

The signature has a wide range of values, with the large ones corresponding to strong structures and the small ones coming from noise and null detections. This can be appreciated in Fig. 2, which shows the cluster, filament and wall signature. Therefore, the last step in the algorithm involves the use of physical criteria to find the threshold signature that discriminates between valid and invalid detections. Signature values larger than the threshold correspond to real structures while the rest are spurious detections. The threshold signature for clusters is found by requiring that the identified objects are virialized, whereas for filaments and walls the threshold is given by the dependence of the filament/wall mass with environmental signature.

The procedure to determine the signature threshold for cluster detection is illustrated in the upper panel of Fig. 3. Clusters are the largest and most recently formed fully virialized objects (Voit 2005). We use this definition to determine the signature threshold for cluster identification. We test for virialization by requiring that the average density of the cluster is larger than $\Delta = 370$, which is the value given by the spherical collapse model at $z = 0$ (Gunn & Gott 1972). From Fig. 3 it can be seen that the fraction of objects with an average density larger than the virialization threshold changes very fast from 0 to 1 as we increase the cluster signature \mathcal{S}_c . We then take the signature threshold as the value where half of the objects have a density larger than Δ .

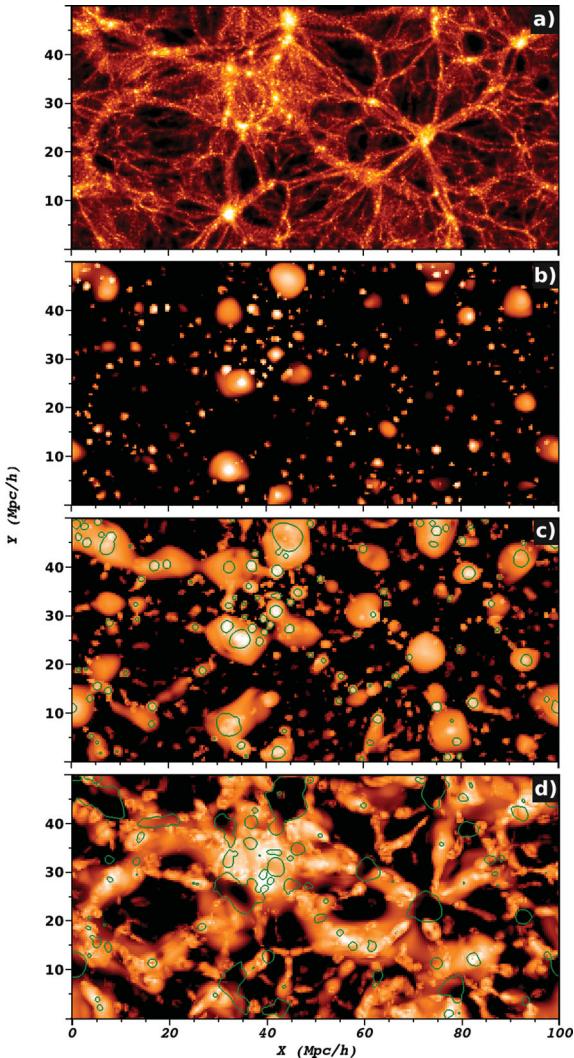


Figure 2. A $1 h^{-1}$ Mpc slice through (a) density field, (b) node signature, (c) filament signature and (d) wall signature obtained from an N -body simulation. The white, orange and black show the high, medium and low values of density and environment signature. The green line contours show the regions with high node signature (panel c) and with high filament signature (panel d).

The filament and wall identification is performed by limiting our detections to only the most prominent filamentary and wall regions. We find that the same method can be successfully used for the recognition of both filaments and walls. Let us denote with $M_f(S_f)$ the mass in filaments with a signature value larger or equal to S_f . As S_f decreases, more and more regions are included and hence $M_f(S_f)$ increases. Most of the change in this function is restricted to a small range in S_f values and it gives a natural way of discriminating between real and spurious detections. In Appendix A we show that the mass change with signature,

$$\Delta M_f^2 = \left| \frac{dM_f}{d \log S_f} \right|, \quad (9)$$

gives a natural and robust method of defining the most prominent filamentary components of the cosmic web. Similarly, in the case of walls, we can define ΔM_w^2 using the above equation with M_f and S_f replaced by their corresponding quantities for wall environments, M_w and S_w . The quantity $M_w(S_w)$ is the mass in regions that have a

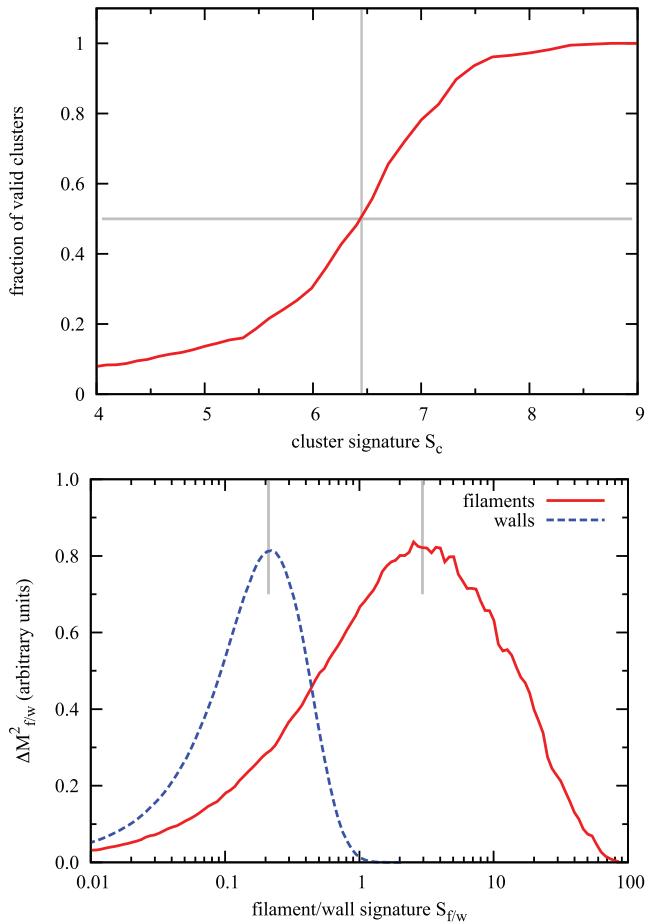


Figure 3. Upper panel: the dependence of the fraction of clusters with density larger than the virial density versus the cluster signature S_c . The intersection of the grey lines shows the cluster detection threshold. Lower panel: determination of the detection threshold for filaments and walls. The peak of ΔM^2 (shown by the grey vertical lines) corresponds to the signature threshold for filament and wall identification (see text for details).

wall signature value larger or equal to S_w . The ΔM^2 dependence for both filaments and walls is shown in Fig. 3. We use the pronounced ΔM^2 peak to delineate the valid environments, which are the points with signatures larger than the position of the ΔM^2 peak. All other points with smaller signature are considered null detections. This threshold method reproduces very well the filamentary and wall network visible in both the cosmic density and velocity divergence fields.

The algorithm performs the environment detection by applying the above steps first to clusters, then to filaments and finally to walls. This sequence (first clusters, then filaments and finally walls) has to be followed due to presence of anisotropic clusters and filaments that give mixed environmental signatures. This can be appreciated from panel (c) of Fig. 2 where on top of the filament signature we show the contours corresponding to large cluster signature. We see that there are several regions that have both a large cluster and filamentary characteristic. This is due to non-spherical clusters that have a large filamentary signature. Similar anisotropic cluster/filaments may give a strong wall signature (see panel d of Fig. 2). To overcome this cross-contamination, a point is part of a filament only if it was not previously identified as in a cluster. Similarly a point is in a wall if it was not previously identified as part of a cluster or filament. This procedure makes sure that each point is assigned a

single classification: cluster, filament, wall or field (everything else that is not a cluster, filament or wall).

3 NEXUS+: LOGARITHMIC FORMALISM FOR MULTISCALE MORPHOLOGICAL ANALYSIS

The NEXUS algorithm is very efficient in detecting the environments in a field f where all the structures correspond to the same order of magnitude values of f . However, the method faces some challenges when the structures in f are present over orders of magnitude in field values. To better understand this, we present a test example in Fig. 4. It shows three filaments characterized by different intensities: 1, 10 and 100 (from left to right). For the NEXUS method to identify all the three filaments, the threshold needs to be so low that the stronger filaments are detected as extending much beyond their input data boundaries. The higher intensity peaks give a significant signal even at large distances, due to the combination of the Gaussian filter not dropping off fast enough and the high field value of the peak.

One way to remedy this problem is to replace the Gaussian filter with a new smoothing method that takes into account the large range in values. For that we introduce the log-Gaussian filter, which is a Gaussian filter in logarithm space. By replacing the Gaussian filter in NEXUS with the log-Gaussian filter we obtain the NEXUS+ formalism. The results of the new method are presented in frame (c) of Fig. 4. It clearly shows that the new filtering procedure recovers much better the three filaments.

3.1 NEXUS+: general algorithm description

The main difference between NEXUS and NEXUS+ is the use of the log-Gaussian filter instead of the Gaussian one. The steps of the NEXUS+ algorithm are the same as the steps of NEXUS with the exception of steps (I) and (II).

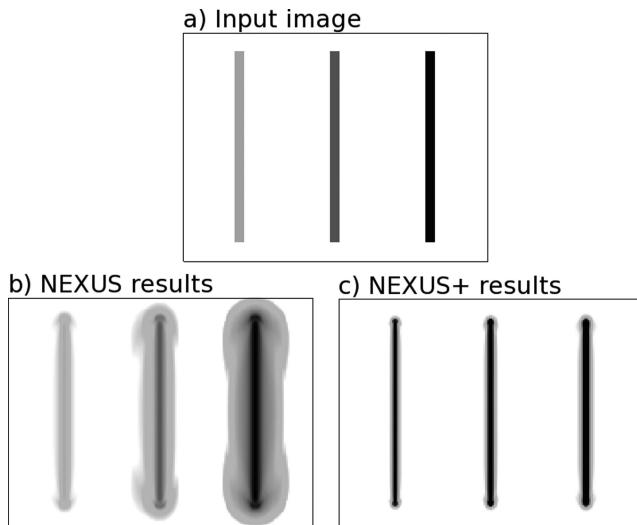


Figure 4. The filaments detected using the NEXUS and NEXUS+ methods when applied to a test image. The input test image (upper panel) contains three filaments with the same width but of different intensities: 1, 10 and 100 (from left to right). Frame (b) shows the NEXUS filament signature with a threshold low enough such that also the weakest filament (leftmost one) is visible. Panel (c) depicts the same as frame (b) but for the NEXUS+ method.

3.1.1 Step I: applying log-Gaussian smoothing

A log-Gaussian filter of width R_n is applied to the input field f . To this end, we introduce the field g , the logarithm of field f ,

$$g = \log_{10} f, \quad (10)$$

and the field g_{R_n} , the smoothed logarithm at scale R_n ,

$$g_{R_n}(\mathbf{x}) = \int d^3y g(\mathbf{y}) W_{G,R_n}(\mathbf{x}, \mathbf{y}), \quad (11)$$

with W_{G,R_n} the Gaussian filter of width R_n .

Following the introduction of these quantities, the application of the log-Gaussian filter consists of three main steps.

(i) Computing the logarithm of the field f , $g = \log_{10} f$.

(ii) Applying the Gaussian filter of width R_n to g to obtain the smoothed logarithm g_{R_n} .

(iii) Computing the smoothed field f_{R_n} by taking the exponential of the smoothed logarithm g_{R_n} .

In practice, we perform the convolution of the field g with the Gaussian filter in Fourier space,

$$g_{R_n}(\mathbf{x}) = \int \frac{d^3k}{(2\pi)^3} e^{-k^2 R_n^2 / 2} \hat{g}(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{x}}, \quad (12)$$

involving the simple multiplication of the Fourier field component $\hat{g}(\mathbf{k})$ with the Gaussian exponential,

$$\hat{g}_{R_n}(\mathbf{k}) = e^{-k^2 R_n^2 / 2} \hat{g}(\mathbf{k}). \quad (13)$$

Subsequently, the resulting NEXUS+ smoothed field f_{R_n} is obtained by evaluating

$$f_{R_n}(\mathbf{x}) = C_{R_n} 10^{g_{R_n}}. \quad (14)$$

The variable C_{R_n} is a multiplication constant that assures the mean of the input field is the same before and after filtering.

3.1.2 Step II: computing Hessian eigenvalues

The second step is the same as for the NEXUS algorithm, but since the smoothing filter is different some of the equations will also change. Now the Hessian of the smoothed field:

$$\mathbf{H}_{ij,R_n}(\mathbf{x}) = R_n^2 \frac{\partial^2 f_{R_n}(\mathbf{x})}{\partial x_i \partial x_j}, \quad (15)$$

can be written in Fourier space using

$$\mathbf{H}_{ij,R_n}(\mathbf{k}) = -k_i k_j R_n^2 \hat{f}_{R_n}(\mathbf{k}). \quad (16)$$

Please note that now one cannot formulate $\hat{f}_{R_n}(\mathbf{k})$ as a simple analytical expression as in the case of the NEXUS algorithm's equation (3). For the NEXUS+ algorithm one needs to perform steps (I) and (II) separately.

The rest of the steps are the same as NEXUS steps (III)–(VI) described in Section 2.1. It is important to note that because NEXUS+ uses the logarithm of f it can only be applied to input fields that have positive values at every point.

3.2 NEXUS+ on the density field

The major challenge of structure detection lies in the fact that the non-linear density field, following its evolution, ranges over many orders of magnitude between the underdense and overdense regions. Structures are present over the whole range of values in density. To deal with this challenge we can use two approaches: either take

the density logarithm (see Section 4.3) or use a different algorithm that takes into account the approximative lognormal shape of the density distribution. Here we take the former approach and apply the NEXUS+ algorithm to the density field.

The strength of the NEXUS+ algorithm can be easily appreciated if one compares the density field using Gaussian versus log-Gaussian smoothing. While there is a one to one mapping between density and density logarithm, this relation does not hold when one compares the smoothed density with the smoothed logarithm of the density. The results of the two methods are shown in Fig. 5. We immediately observe that the Gaussian filtered density is dominated by several peaks with a typical spherical shape. On the other hand, the log-Gaussian results seem to trace much better the large-scale structure.

Most of the differences between the two results come from the very high density peaks. When applying a Gaussian smoothing, these higher density peaks gets smoothed up to large distances and dominate the signal coming from other less dense neighbouring regions. This leads to a loss of information about the large-scale structure around these peaks. In the case of log-Gaussian smoothing, by taking the logarithm of the input field the contrast between these very high density peaks and their neighbourhoods is greatly reduced. Therefore the contribution of the peaks will not be dominant, even though the log-Gaussian filter has the same spatial extension as the Gaussian one.

4 TRACER FIELDS OF THE COSMIC WEB: EXTENDING THE NEXUS ALGORITHM BEYOND DENSITY

There are various methods that attempt to identify the components of the cosmic web. These not only implement different detection techniques, but in many cases differ in the nature of the field used to trace the underlying cosmic structure. In other words, the variation in the results of different methods should not only be ascribed to the algorithms used, but also to the differences in the tracer fields. In this section we extend the NEXUS method to a multitude of cosmic web tracers: density, tidal field, velocity divergence and shear as well as to the density logarithm (see Table 2). By doing so we not only find the field with the best footprint of the cosmic web, but also gain better understanding of the evolution and structure of the cosmic web.

The most widely used tracers of the cosmic web are the density (Aragón-Calvo et al. 2007b; Sousbie et al. 2008b; Soubie, Pichon & Kawahara 2011) and the tidal field (Hahn et al. 2007a; Forero-Romero et al. 2009). On the other hand, the use of the density logarithm for the classification of the cosmic web is a novel method that we introduce here. Later on we will argue why this is a natural structure tracer field that one should consider. While there have been a fair share of methods using the positional information of the phase space, there are very few works that use the velocity field as cosmic web tracer. Most interesting is the work by Shandarin (2011), which emphasized the importance of the velocity field for understanding the emerging patterns in the matter distribution. Based on this, he used the variance of the velocity field as a measure of the local environment. In a sequence of studies that follows up on this idea (Abel, Hahn & Kaehler 2011; Neyrinck 2012; Shandarin, Habib & Heitmann 2012), the full phase-space structure of the mass distribution is used for an impressively accurate dynamical characterization of morphological structure of the cosmic web. Following this promising avenue, we are also working

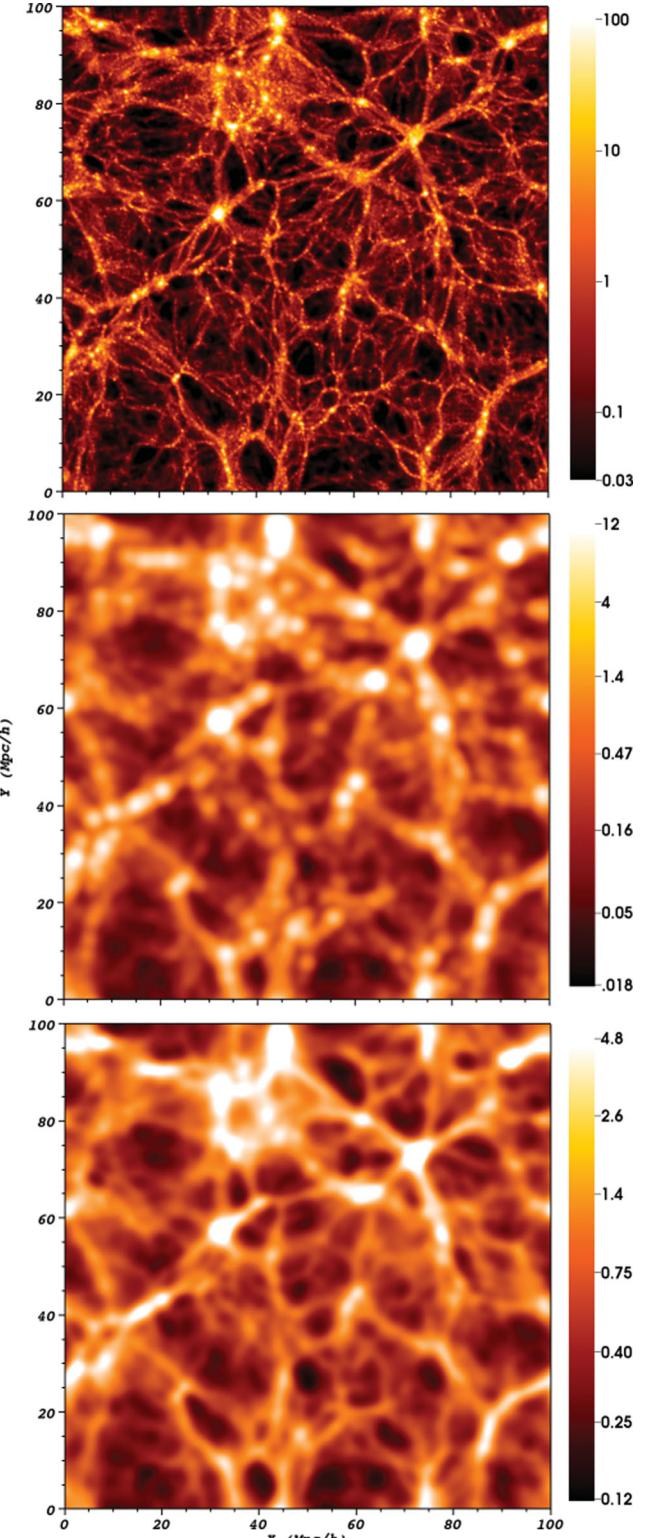


Figure 5. Comparing Gaussian and log-Gaussian density smoothing (see text for details). The upper panel shows the initial density field, while the central and lower panels give the Gaussian and log-Gaussian smoothed density, respectively. Both cases were obtained using a $1 h^{-1} \text{ Mpc}$ smoothing. The scale in the lower two panels was selected by fitting the density histogram with a lognormal distribution and plotting the values in the $\text{peak} - 3\sigma$ to $\text{peak} + 3\sigma$ range (with peak and σ the peak and standard deviation of the lognormal distribution).

Table 2. The methods resulting from the extension of the NEXUS algorithm to several tracer fields of the cosmic web.

Method name	Tracer field
NEXUS_den	Density field δ
NEXUS_tidal	Tidal field T
NEXUS_denlog	Density logarithm field $\log_{10}(1 + \delta)$
NEXUS_veldiv	Velocity divergence field θ
NEXUS_velshear	Velocity shear field σ

on relating the NEXUS and NEXUS+ formalism to the structure found by these phase-space-based methods.

4.1 NEXUS_den: tracing the cosmic web using the density field

The density is one of the obvious fields used for environmental detection due to the sharp contrast of the clusters and filaments compared to most of remaining void-dominated volume. To apply the NEXUS algorithm on the density one has to just insert the density δ in equation (3). For simplicity, we denoted this method as NEXUS_den. To better understand the behaviour of the NEXUS_den method we need to rewrite the Hessian matrix (see equations 3 and 4) to

$$\hat{\mathbf{H}}_{ij,R_n}(\mathbf{k}) = -\frac{k_i k_j}{k^2} \hat{\delta}(\mathbf{k}) k^2 R_n^2 e^{-k^2 R_n^2/2}, \quad (17)$$

where the initial equation was multiplied by the unit factor $\frac{1}{k^2} k^2$. It is immediately obvious that the Hessian is given by two distinct parts: the tidal field² multiplied by a bandpass filter. The bandpass filter is made of two distinct components: the $k^2 R_n^2$ high pass filter and the Gaussian $e^{-k^2 R_n^2/2}$ low pass filter. Simple calculations show that the maximum of the band pass filter is at $k = \frac{\sqrt{2}}{R_n}$, while the shape of the filter is very similar to the one in Fig. 1. Therefore detecting the cosmic web structures in the density field is equivalent to identifying those structures in a bandpass filtered tidal field.

When applying the NEXUS_den formalism, an additional step has to be taken on top of those described in Section 2.1 and mask the density field when detecting filaments and walls. For filaments identification we need to set the density to 0 in the cluster regions. In the absence of this mask, the cluster regions will give a large, unrealistic, filamentary signature.³ Similarly when identifying walls, we need to set the density to 0 in both the cluster and filament regions.

² The first part of equation (17) is the same as the Fourier transform of the tidal field given by equation (21) in Section 4.2 up to the multiplication factor $4\pi G \bar{\rho}$. This factor has no effect on the final results since it only rescales the Hessian eigenvalues.

³ The filamentary signature of cluster regions will be large, even though the filamentary shape strength $\mathcal{I}_{\text{filament}}$ given by equation (6) is small in those regions. This is since the filamentary signature (see equation 7) depends on $|\lambda_2|$ which has very large values in the cluster regions and will compensate for the small values of $\mathcal{I}_{\text{filament}}$. So this additional $|\lambda_2|$ factor that discriminates between signal and noise also introduces false detections. More generally, this false detection problem is important when the typical values of the tracer field for the different environments are orders of magnitude apart. It can be easily corrected by using the masking procedure described in the text.

4.2 NEXUS_tidal: tracing the cosmic web using the tidal field

The tidal field is the driver of anisotropic gravitational collapse and it is an essential ingredient for the formation and evolution of the cosmic web (Zel'dovich 1970; Gurbatov, Saichev & Shandarin 1989; Bond et al. 1996). It is only natural to use it for the detection and understanding of the cosmic structures (Hahn et al. 2007a; Forero-Romero et al. 2009). This mode of the NEXUS method is indicated as NEXUS_tidal.

The tidal field is given by

$$\mathbf{T}_{ij}(\mathbf{x}) = \frac{\partial^2 \phi_{\text{grav}}(\mathbf{x})}{\partial x_i \partial x_j}, \quad (18)$$

with ϕ_{grav} the gravitational potential. The latter is related to the density via the Poisson equation:

$$\nabla^2 \phi_{\text{grav}}(\mathbf{x}) = 4\pi G \bar{\rho} \delta(\mathbf{x}), \quad (19)$$

where G is the gravitational constant, $\bar{\rho}$ is the average matter density and $1 + \delta(x) = \rho(x)/\bar{\rho}$ is the overdensity. The Poisson equation is easily solved in Fourier space to obtain

$$\hat{\phi}_{\text{grav}}(\mathbf{k}) = -4\pi G \bar{\rho} \frac{1}{k^2} \hat{\delta}(\mathbf{k}), \quad (20)$$

leading to the following expression for the Fourier components of the tidal field:

$$\hat{\mathbf{T}}_{ij}(\mathbf{k}) = 4\pi G \bar{\rho} \frac{k_i k_j}{k^2} \hat{\delta}(\mathbf{k}). \quad (21)$$

Identifying the cosmic environments traced by the tidal field is done by applying the NEXUS algorithm on the gravitational potential. The Hessian matrix of the potential ϕ_{grav} given by equation (20) reduces to

$$\hat{\mathbf{H}}_{ij,R_n}(\mathbf{k}) = 4\pi G \bar{\rho} \frac{k_i k_j}{k^2} \hat{\delta}(\mathbf{k}) e^{-k^2 R_n^2/2}. \quad (22)$$

This is exactly the tidal field smoothed over with a Gaussian filter.

The difference between the Hessian matrix of NEXUS_tidal and NEXUS_den consists in the additional $k^2 R_n^2$ high pass filter present in the case of the second method – compare equations (17) and (22). Therefore variations in the result of the two methods come from excluding the low-frequency modes in the NEXUS_den case and not in the NEXUS_tidal. The same conclusion can be reached by looking at the Fourier transform amplitude of the input fields: the density versus the gravitational potential. These are shown in Fig. 6. The first is more flat, while for the second the low frequencies have much larger amplitudes. This means that the large-scale modes give a much larger contribution for NEXUS_tidal than for NEXUS_den. These effects are illustrated in Fig. 7. The NEXUS_den environments have a very clumpy appearance and are very sensitive to small-scale structures. On the other hand, NEXUS_tidal is only responsive to the large-scale modes and cannot trace the smaller details of the matter distribution.

It is important to note that the environment characteristics in the gravitational potential are different from the ones in the density field. According to the cosmic web theory, the clusters, filaments and walls are given by the strength and sign of the first, second and third eigenvalues of the tidal field (Bond et al. 1996). This can be easily implemented within the NEXUS framework by changing the environmental signature from equation (7) to

$$\mathcal{S} = \theta(\lambda_a) \lambda_a, \quad (23)$$

with $a = 1$ for clusters, $a = 2$ for filaments and $a = 3$ for walls.

As in the case of the NEXUS_den method, we need to apply a cluster mask when identifying filaments and a combined cluster and

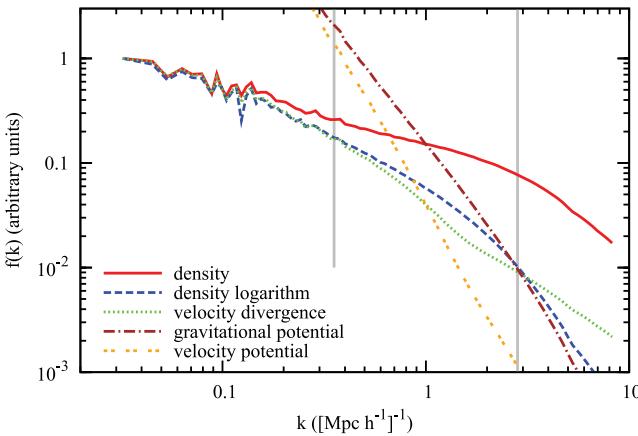


Figure 6. The Fourier transform amplitude for the density, density logarithm, gravitational potential, velocity divergence and velocity potential fields. The spectra were obtained directly from the DTFE interpolated fields used as input for the NEXUS method. The curves were shifted vertically to better emphasize the differences between the two grey lines which mark the peak of the \mathcal{H}_R function with $R = 4$ (left line) and $0.5 h^{-1} \text{ Mpc}$ (right line). The two smoothing radii represent the upper and lower limits of the smoothing scales set used in the NEXUS algorithm.

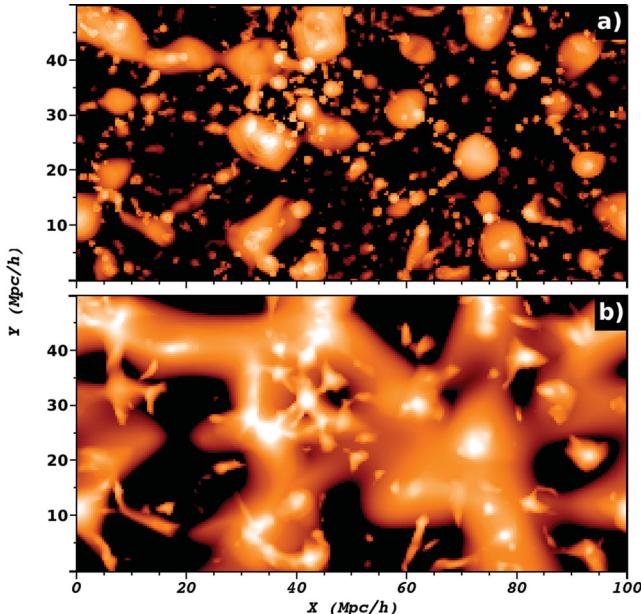


Figure 7. The filament signature computed using (a) NEXUS_{den} and (b) NEXUS_{tidal} methods. The graph shows a thin $1 h^{-1} \text{ Mpc}$ slice. The white, orange and black correspond to high, medium and low signature values.

filament mask when identifying walls. This procedure sets the density to 0 in the mask regions, after which the gravitational potential is computed using equation (19).

4.3 NEXUS_{denlog}: tracing the cosmic web using the density logarithm

Using the density logarithm for cosmic web detection is an approach that has not been explored until now. We were motivated to apply the NEXUS algorithm to the density logarithm and not the density itself because of a multitude of reasons.

(i) The NEXUS method works best when all structures correspond to similar values in the input field, while the density ranges from 0.01 in underdense versus 10^4 and higher in overdense regions. We expect to find structures over orders of magnitude in density values and simply using density biases the results towards high-density structures. By taking the density logarithm the orders of magnitude difference is reduced to values of -2 in voids to around 4 in cluster regions.⁴

(ii) The non-linear density field is close to a lognormal distribution, when smoothed on scales of a few Mpc (Coles & Jones 1991).

(iii) The large-scale structure is best made visible when rendering the density logarithm and not the density itself.

The NEXUS_{denlog} method consists in replacing the input field f in the NEXUS algorithm with $\log_{10}(1 + \delta)$. The main difference between this method and NEXUS_{den} consists in the reduced contrast between underdense and overdense regions as well as a much steeper spectrum towards large scales for the density logarithm (see Fig. 6). Because of the reduced contrast between underdense and overdense regions when looking at the density logarithm, there is no need to apply the mask described for the NEXUS_{den} method.

While we were motivated by the same reasons as above to develop the NEXUS+ algorithm, there is a large difference between the NEXUS+ and NEXUS_{denlog} methods as will be clearly visible in the results of Section 6. In NEXUS_{denlog} we identify the cosmic web using the logarithm of the density $\log_{10}(1 + \delta)$, while for the NEXUS+ method we trace the environments using the density field smoothed with the log-Gaussian filter.

4.4 NEXUS_{veldiv}: tracing the cosmic web using the velocity divergence

The NEXUS_{den}, NEXUS_{tidal} and NEXUS_{denlog} methods use only half of the phase space, the positional information, for identifying the elements of the cosmic web. It is interesting from both a theoretical and practical point of view to see how the remaining phase space can also be used to trace the large-scale structure. The natural candidates for this are the velocity divergence and the velocity shear, due to the one-to-one connection between these quantities and the density and tidal field in the linear regime.

The velocity divergence is defined as

$$\theta(\mathbf{x}) = \frac{1}{H} \nabla \cdot \mathbf{v}(\mathbf{x}), \quad (24)$$

where we divide by the Hubble factor H such that θ is a dimensionless quantity. The velocity divergence is easily computed as an output of the DTFE method (Bernardeau & van de Weygaert 1996; Romano-Díaz & van de Weygaert 2007). According to linear theory, the velocity divergence is related to the density field via

$$\theta(\mathbf{x}) = -f\delta(\mathbf{x}), \quad (25)$$

with f the linear velocity growth factor (see Peebles 1980). So in the linear regime, any structure in the density field should also be present in velocity divergence. A similar relation between θ and δ holds true also for the non-linear regime, but in a more complex way (for details see Nusser et al. 1991; Chodorowski & Lokas 1997; Bernardeau et al. 1999). The differences between the structures

⁴ Note that the DTFE density field will always have a density different from zero even in the emptiest voids. The typical DTFE density contrast in voids at redshift $z = 0$ is $0.01\text{--}0.1$.

detected using density versus velocity divergence probe the effects of the non-linear evolution on the cosmic web components.

The NEXUS_veldiv method is the application of the NEXUS algorithm on the negative of the velocity divergence $-\theta$. We choose the minus sign because of equation (25). The NEXUS_veldiv Hessian matrix is given by

$$\hat{\mathbf{H}}_{ij,R_n}(\mathbf{k}) = \frac{k_i k_j}{k^2} \hat{\theta}(\mathbf{k}) k^2 R_n^2 e^{-k^2 R_n^2/2}. \quad (26)$$

This is the product of the velocity shear given by equation (32) multiplied by a bandpass filter. This is exactly the same as for the NEXUS_den method, but with the tidal field replaced by the velocity shear. The main difference between NEXUS_veldiv and NEXUS_den can be easily seen in Fig. 6: the velocity divergence and density have the same Fourier components at large scales, but the density has a more flattened drop at smaller scales.

In contrast to the NEXUS_den method, we choose not to apply a mask for the velocity-related methods. While there are still orders of magnitude variation in the velocity divergence between overdense and underdense regions, this difference is not as large as for the density field. The major challenge in applying a mask arises because the velocity divergence can take both positive and negative values, so there is no a priori well motivated value that we can use in the mask regions.

4.5 NEXUS_velshear: tracing the cosmic web using the velocity shear

The velocity shear is the symmetric part of the velocity gradient, with the ij component defined as

$$\sigma_{ij}(\mathbf{x}) = \frac{1}{2H} \left(\frac{\partial v_j}{\partial x_i} + \frac{\partial v_i}{\partial x_j} \right), \quad (27)$$

where v_i is the i component of the velocity. We normalize the velocity shear by the Hubble constant to keep the same notations as in the case of the velocity divergence. To obtain the velocity shear, the velocity is rewritten as the sum of the potential and rotational flows:

$$\mathbf{v} = \nabla \phi_{\text{vel}} + \nabla \times \mathbf{A}_{\text{vel}}, \quad (28)$$

where ϕ_{vel} is the scalar velocity potential and \mathbf{A}_{vel} is the vector potential. Inserting this last equation into the velocity shear gives

$$\sigma_{ij}(\mathbf{x}) = \frac{1}{H} \frac{\partial^2 \phi_{\text{vel}}(\mathbf{x})}{\partial x_i \partial x_j}. \quad (29)$$

Therefore, the velocity shear depends only on the velocity potential. It is interesting to notice that the velocity divergence is also given only by the velocity potential via

$$\theta = \frac{1}{H} \nabla^2 \phi_{\text{vel}}. \quad (30)$$

This last equation can be inverted and used to solve for the potential ϕ_{vel} to obtain its Fourier components as

$$\hat{\phi}_{\text{vel}}(\mathbf{k}) = -H \frac{1}{k^2} \hat{\theta}(\mathbf{k}). \quad (31)$$

On the basis of this equation, one can infer the velocity potential starting from the velocity divergence output of the DTFE method. Combining these, we obtain that the components of the velocity shear can be expressed in terms of θ as

$$\hat{\sigma}_{ij}(\mathbf{k}) = \frac{k_i k_j}{k^2} \hat{\theta}(\mathbf{k}). \quad (32)$$

Thus the relation between the velocity shear and the tidal tensor in the linear regime is given by

$$\sigma_{ij} = -f \mathbf{T}_{ij}. \quad (33)$$

The NEXUS_velshear method involves using the negative of the velocity potential $-\phi_{\text{vel}}$ as input field to the NEXUS algorithm. The negative sign comes, as in the case of NEXUS_veldiv, from the minus in relation equation (33). Inserting the expression for ϕ_{vel} in the Hessian matrix gives

$$\hat{\mathbf{H}}_{ij,R_n}(\mathbf{k}) = -\frac{k_i k_j}{k^2} \hat{\theta}(\mathbf{k}) e^{-k^2 R_n^2/2}. \quad (34)$$

As in the case of the NEXUS_tidal method, the environment characteristics are different in velocity shear compared to velocity divergence. So the environment signature has to be changed to the expression given by equation (23).

5 N-BODY SIMULATIONS AND HALO CATALOGUES

To test our structure finding algorithms, we apply them to cosmological N -body simulations containing only dark matter particles. We adopted the Λ cold dark matter (Λ CDM) cosmological model with $\Omega_m = 0.26$, $\Omega_\Lambda = 0.74$, $h = 0.71$, $\sigma_8 = 0.8$ and $n_s = 1$. We performed two 512^3 particle simulations in a 100 and $200 h^{-1}$ Mpc periodic boxes. The force resolution was fixed in comoving coordinates up to $z = 4$, to 15 and $21 h^{-1}$ kpc. Afterwards it was fixed in physical coordinates to 5 and $7 h^{-1}$ kpc, respectively.

The $100 h^{-1}$ Mpc simulation (mass resolution of $5.4 \times 10^8 h^{-1} M_\odot$) was chosen in order to resolve haloes to a few times $10^{10} h^{-1} M_\odot$ and at the same time to have a reasonable cosmological volume whose smallest mode is still evolving linearly. We used this small volume simulation for visualization and resolution studies. The $200 h^{-1}$ Mpc simulation (mass resolution of $4.3 \times 10^9 h^{-1} M_\odot$) is used for computing the quantitative results since a larger volume gives better statistics.

The simulations were performed using the public version of the parallel TREE-PM code GADGET2 (Springel 2005) on a Linux cluster at the University of Groningen, the Netherlands. The initial conditions for both simulations were generated at the $z = 50$ redshift using the transfer function given by Bardeen et al. (1986).

5.1 Density and velocity divergence fields

The output of the N -body simulation consists of a discrete set of particles. This needs to be interpolated to a continuum volume-filling density and velocity divergence fields that will be used as input for the cosmic web environment detection algorithm. It is crucial for the environment detection procedure, especially for anisotropic features such as filaments and walls, that the interpolation method used to obtain the continuous fields retains all the scale and geometry information of the discrete galaxy or particle distribution.

For these reasons we use the DTFE, introduced by Schaan & van de Weygaert (2000) (for additional details see van de Weygaert & Schaan 2009; Cautun & van de Weygaert 2011), to reconstruct the underlying density and velocity divergence fields. For the environment detection algorithm, the DTFE method has the following important advantages.

- (i) Preserves the multiscale character of the discrete distribution.
- (ii) Preserves the local geometry of the discrete distribution.

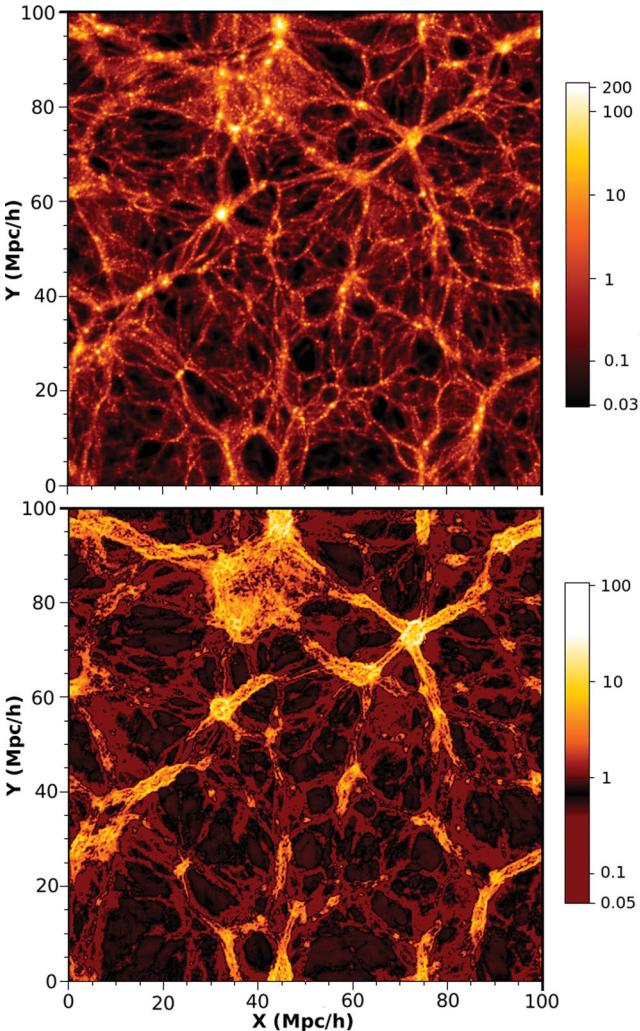


Figure 8. A $1 h^{-1}$ Mpc slice from the N -body simulation illustrating the DTDE density $1 + \delta$ (upper panel) and absolute value of the velocity divergence θ (lower panel).

(iii) Does not depend on user defined parameters or choices.

The continuous DTDE density and velocity divergence fields are sampled on a 256^3 and 512^3 grid for the 100 and $200 h^{-1}$ Mpc simulations, respectively, such that there is a $0.4 h^{-1}$ Mpc grid spacing in both cases. Fig. 8 shows a thin slice of the grid sampled density and velocity divergence fields. Note the level of detail in the structures, even inside voids, and the one-to-one correspondence between the DTDE density and velocity divergence features. Fig. 9 gives a 3D rendering of the density and velocity divergence fields in a larger volume – the same volume that later on will be used to visualize the cosmic web environments.

5.2 Halo and subhalo catalogues

We use the AMIGAs Halo Finder (AHF) by Knollmann & Knebe (2009) to identify the dark matter haloes. The AHF halo finder is the successor of the MLAPM Halo Finder (MHF) halo finder by Gill, Knebe & Gibson (2004). AHF uses adaptive mesh refinement to identify the density peaks which it classifies as the halo and subhalo centres. Afterwards it grows the objects around their centres until the

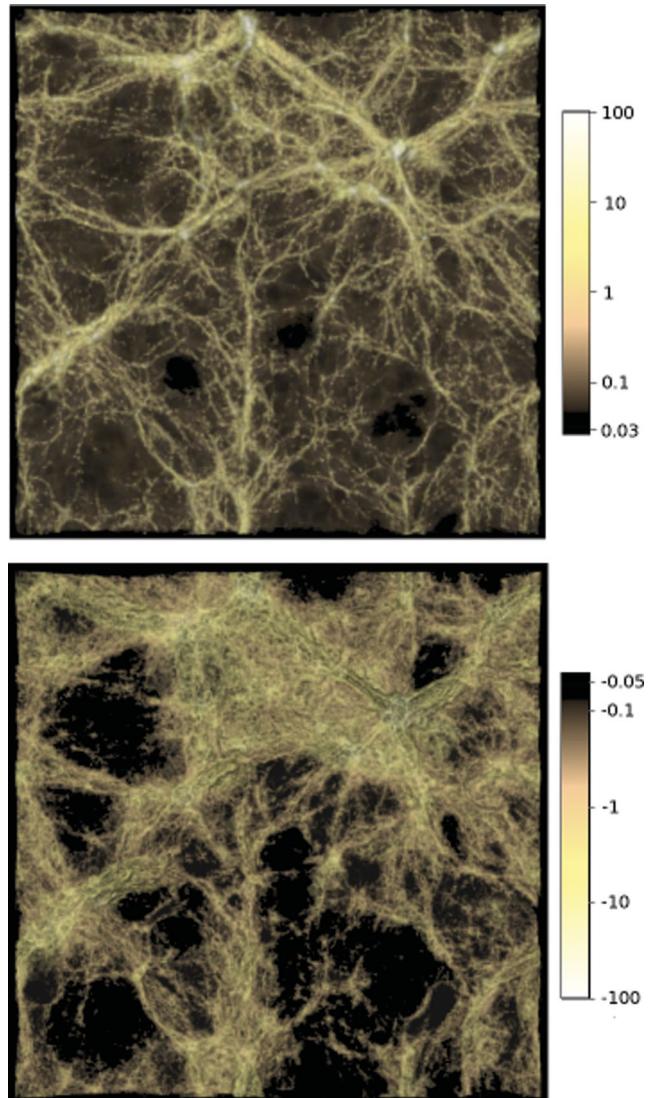


Figure 9. A 3D volume rendering of the density $1 + \delta$ (upper panel) and velocity divergence θ (lower panel). Note that we only show the negative values of the velocity divergence. The picture represents a $100 \times 100 \times 100 (h^{-1} \text{ Mpc})^3$ volume in an N -body simulation. We used the same volume to illustrate the cosmic web environments in Figs 12–16.

spherically averaged density contrast reaches the virial density.⁵ The last step consists in removing the gravitationally unbound particles. The AHF halo and subhalo catalogues are complete up to haloes with 50 or more particles (for a complete description see Knollmann & Knebe 2009).

6 THE COSMIC WEB ENVIRONMENTS

6.1 Clusters

The point-like objects detected by the NEXUS and NEXUS+ methods correspond to large overdensities in the density field. They range from very massive to very small mass objects. We will see later in this section that these objects correspond to dark matter haloes.

⁵ The virial density is automatically computed by AHF and depends on both cosmological parameters and redshift.

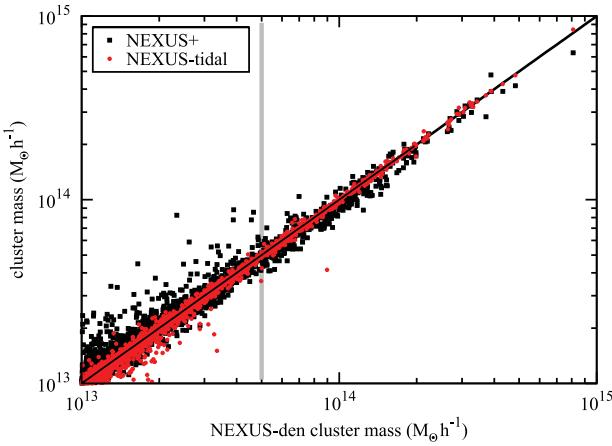


Figure 10. Comparison between the mass of cluster-like objects detected using the NEXUS_{den}, NEXUS_{tidal} and NEXUS₊ methods. The grey vertical line delineates the objects with mass larger than $5 \times 10^{13} h^{-1} M_{\odot}$ while the diagonal black line shows a one-to-one relationship.

On the other hand, the cosmic clusters are the largest and most recent to form virialized objects (Voit 2005). To be able to identify our point-like objects with actual clusters, we need to limit our detections to only the most massive objects. In this study we consider as clusters the objects with mass larger than $5 \times 10^{13} h^{-1} M_{\odot}$. This is a compromise between studying the most massive objects and having a large sample of such objects in our simulation.

The cluster detection method performs very well for the NEXUS_{den}, NEXUS_{tidal} and NEXUS₊ methods. In contrast, it does not give reliable detections for the NEXUS_{veldiv} and NEXUS_{velshear} methods. We suspect this deficiency is due to high vorticity in the cluster regions, vorticity that is not captured in the velocity divergence or shear fields and therefore is not taken into account in our methods.

The NEXUS_{den}, NEXUS_{tidal} and NEXUS₊ methods detect the same clusters, with similar mass and volume associated with each object. This can clearly be seen in the mass comparison plot shown in Fig. 10. The NEXUS_{den} and NEXUS_{tidal} methods give very similar results, while we find a larger scatter when comparing these results with NEXUS₊.

There is a very close connection between the cluster-like objects of our methods and the most massive dark matter haloes. This is shown in Fig. 11 where we use AHF to identify the dark matter haloes. The top panel shows that there is a very good correlation between the AHF halo mass and its corresponding object identified using our methods. We show that this relation holds down to masses of $10^{13} h^{-1} M_{\odot}$. This lower limit is a limitation of the grid spacing size and not of the method. We see that NEXUS_{den}, and also NEXUS_{tidal} (not shown), gives a much better correlation of the cluster mass with the corresponding AHF halo mass, while in the case of NEXUS₊ there is a larger scatter. A similar comparison is done between the cluster volumes and the AHF halo volume – see centre frame of Fig. 11. Again there is a good match between the methods, but with a much wider scatter than in the case of the mass comparison. This is understandable since the volume is much more sensitive to the lower density regions around clusters, while the mass is dominated by the inner high-density regions. In this case the scatter is especially large in the NEXUS₊ results than for the NEXUS_{den} and NEXUS_{tidal} methods.

Finally, we compare the NEXUS_{den} results and AHF haloes only for the most massive objects. In the bottom panel of Fig. 11 it

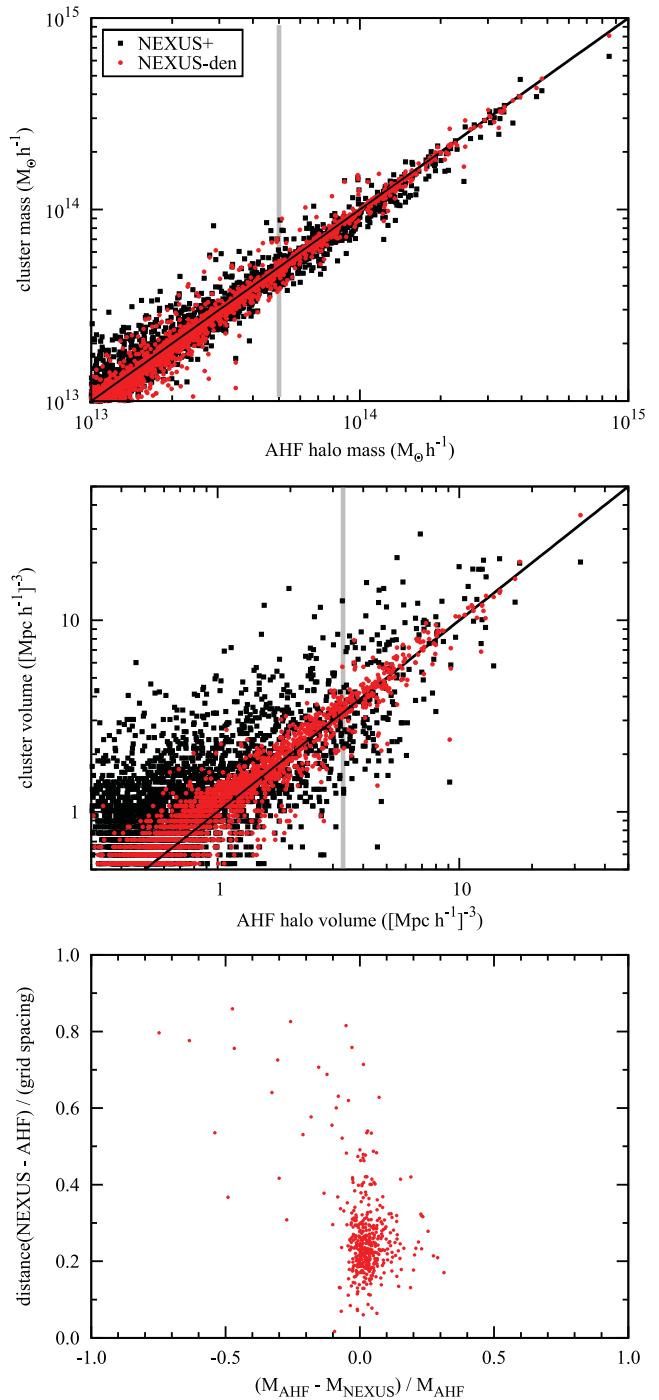


Figure 11. Comparison of cluster-like objects with their corresponding AHF haloes. The top panel shows the mass comparison while the middle panel shows the volume comparison for the NEXUS₊ and NEXUS_{den} objects. The grey vertical line delineates the objects with mass larger than $5 \times 10^{13} h^{-1} M_{\odot}$ while the diagonal black line shows a one-to-one relationship. The lower panel gives the mismatch in centre position and mass difference between the AHF haloes and NEXUS_{den} clusters (only for objects with mass larger than $5 \times 10^{13} h^{-1} M_{\odot}$).

can be seen that the clusters agree within 10 % to their corresponding AHF halo mass and that their centres are at most 0.4 grid spacing distance from the AHF halo centre. The few outliers are objects that are merged in one of the methods and detected as distinct objects in the second one.

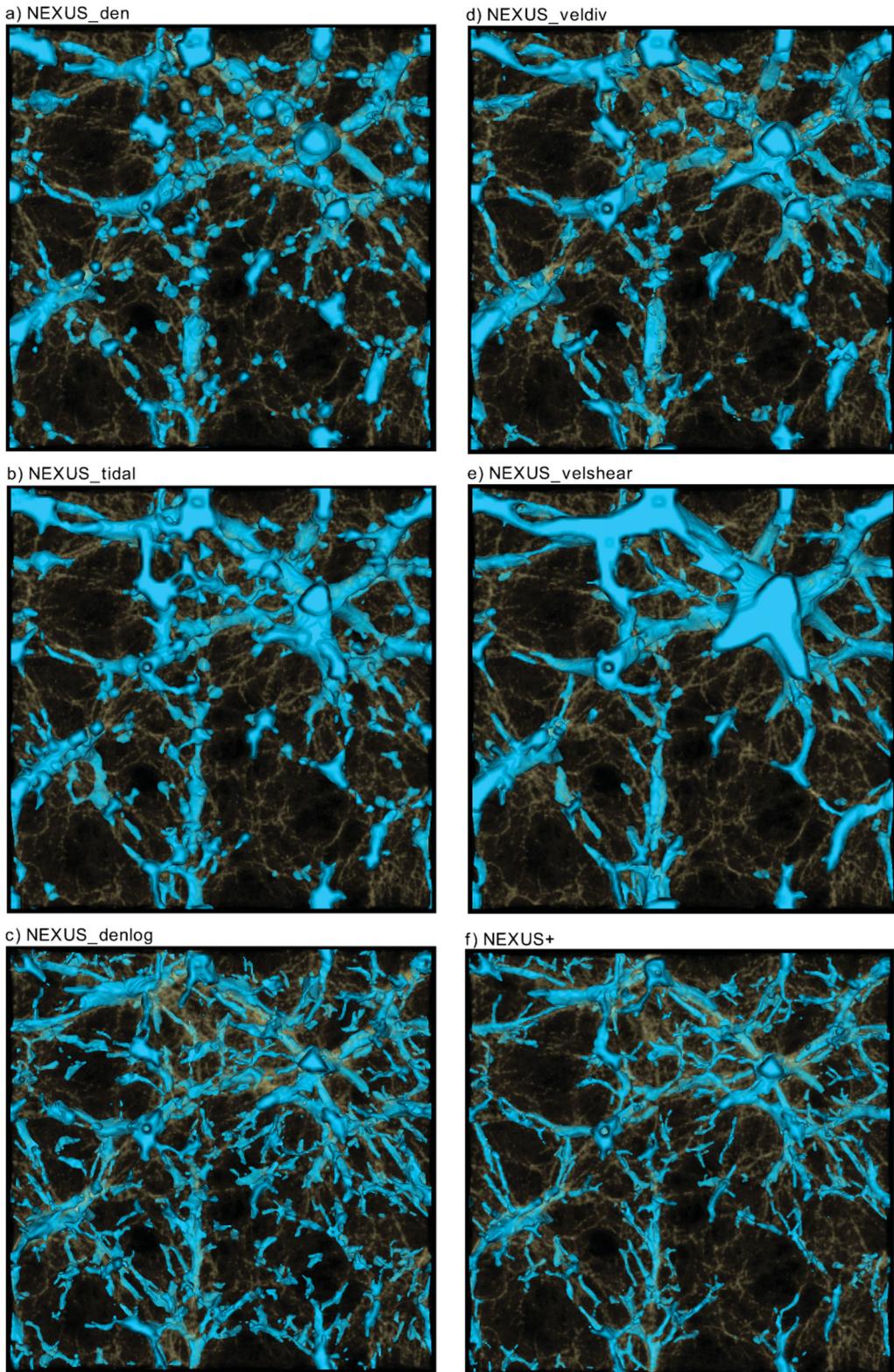


Figure 12. A 3D rendering of the filaments in a $100 \times 100 \times 10 (h^{-1} \text{Mpc})^3$ volume of the *N*-body simulation. The faint background shows the density field. The filaments were obtained using (a) NEXUS_den, (b) NEXUS_tidal, (c) NEXUS_denlog, (d) NEXUS_veldiv, (e) NEXUS_velshear and (f) NEXUS+. This is the same region of the simulation volume as the density field shown in Fig. 9.

We see that NEXUS+ is less reliable in the detection of cosmic web clusters than the NEXUS_den or NEXUS_tidal methods. The very localized nature of the log-Gaussian filter means that the contribution of the highly dense centre does not have a large effect on the periphery of the cluster. Because of this the outer boundaries of the NEXUS+ clusters are more dependent on the substructure at the periphery. One can overcome this ‘weakness’ in the NEXUS+ method by identifying the clusters using NEXUS_den or NEXUS_tidal.

6.2 Filaments

After clusters, filaments are the most noticeable feature of the cosmic web. These structures are shown in Fig. 12, as identified by the six methods proposed in this work. To obtain these results we have restricted our discussion to the most significant filaments, which we define as any continuous region with a volume larger than $10 (h^{-1} \text{ Mpc})^3$. By doing so we discard small regions, typically around isolated haloes which even though show a local filament signature, are not embedded in the larger network.

Comparing the filamentary maps with the density field rendering of the same volume from Fig. 9, we see that all the methods succeed in identifying the strongest filaments, but there are differences when it comes to the smaller, less pronounced filaments. We immediately observe that the cosmic web filaments form an interconnected network, with the most massive filaments acting as the backbone. These very pronounced filaments branch into thinner ones that slowly disappear into lower density regions. The backbone filaments are clearly visible in the central panel while the branching into fainter structures is most pronounced in the lower right-hand panel of Fig. 12. A visual inspection of clusters (not shown here) shows that these reside at the intersection of the most prominent filaments, which serve as the highways along which matter is transported to the clusters.

A visual comparison of the filamentary maps obtained using the six methods leads to the following conclusions.

(i) The filamentary structures in the NEXUS_den and NEXUS_veldiv are more clumpy and have less large-scale cohesion when compared to the NEXUS_tidal and NEXUS_velshear results. This is in line with our expectations, since the large-scale modes bring a larger contribution for the latter methods (see Fig. 6).

(ii) The methods using the tidal and velocity shear fields are biased towards the most significant structures and miss most of the filaments present in the less dense regions. This less pronounced filament, marginally seen by the NEXUS_den and NEXUS_veldiv methods, are very well reproduced in the NEXUS_denlog and NEXUS+ methods. This supports the view that the cosmic web has structures present over a large range in density values.

(iii) While the most pronounced filaments are detected by all the methods, their thickness varies between the different methods. The NEXUS_denlog and NEXUS+ filaments have typical diameters around $2 h^{-1} \text{ Mpc}$, while for the rest the typical diameter is around $4 h^{-1} \text{ Mpc}$. The thinner filaments mostly constitute the inner regions of the much thicker filaments detected using the other methods.

(iv) While NEXUS_denlog finds the same structures as the other methods in the high-density regions, it finds a much richer filamentary network in the underdense regions. This is because any small changes in the density field in these regions can lead to a large contrast in the density logarithm. The Poissonian sampling noise is especially important for density determination in the underdense regions due to sparse sampling. This makes the NEXUS_denlog

method especially sensitive to Poissonian noise in the void-like regions.

To summarize, the variation in the six methods manifests itself as mostly differences in the detection of smaller filaments and thickness differences in the very prominent structures. Since the large-scale modes contribute much more to NEXUS_tidal and NEXUS_velshear, these methods identify only the largest filaments that correspond to the peaks of the large-scale modes. On the other hand, the NEXUS_denlog and NEXUS+ methods are much more sensitive to the less pronounced structures, finding an important filamentary network also in the underdense regions. While NEXUS_den and NEXUS_veldiv are in between the two classes of results, they are much closer in character to the NEXUS_tidal and NEXUS_velshear methods.

6.3 Walls

The wall environments detected using the six methods are presented in Fig. 14 where, on top of the walls in orange, we superimposed the filaments in light blue. In the case of filament identification one can use the density and velocity divergence maps to judge the success of the detection method, but this is much more difficult for walls. This is a consequence of both the smaller contrast and the planar nature of these structures. The presence of sheet-like structures in the distribution of matter on cosmological scales can be easily inferred from Fig. 13, where we show the dark matter particles after removing the particles located in clusters and filaments. The most striking structures are the line-like arrangements of particles visible especially in the upper part of the figure. These are sheets that are perpendicular on the projection plane and not filaments missed by our detection algorithm. There are additional walls along the projection plane (e.g. centre upper part, between the three line-like structures) but these are less easily detected visually.

By comparing the particle distribution from Fig. 13 to the NEXUS walls in Fig. 14 it is clear that the algorithm is very successful in

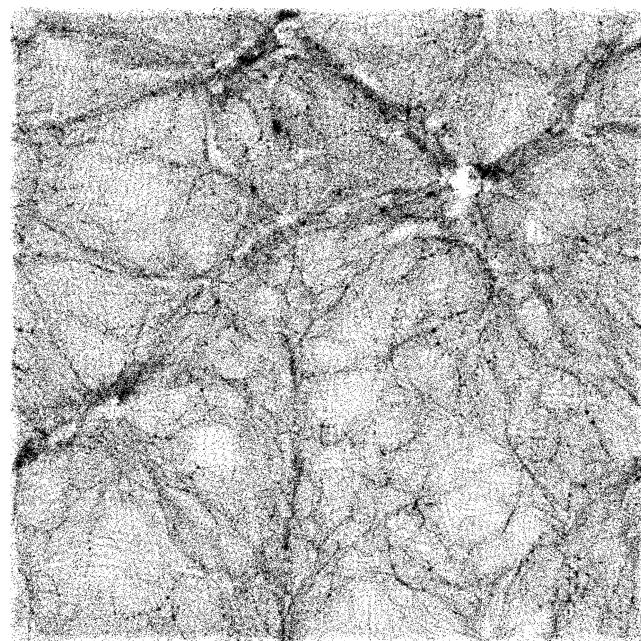


Figure 13. The dark matter particles left after taking out the particles located in cluster and filament environments. This is a projection of the volume of the N -body simulation as in Figs 9, 12 and 14.

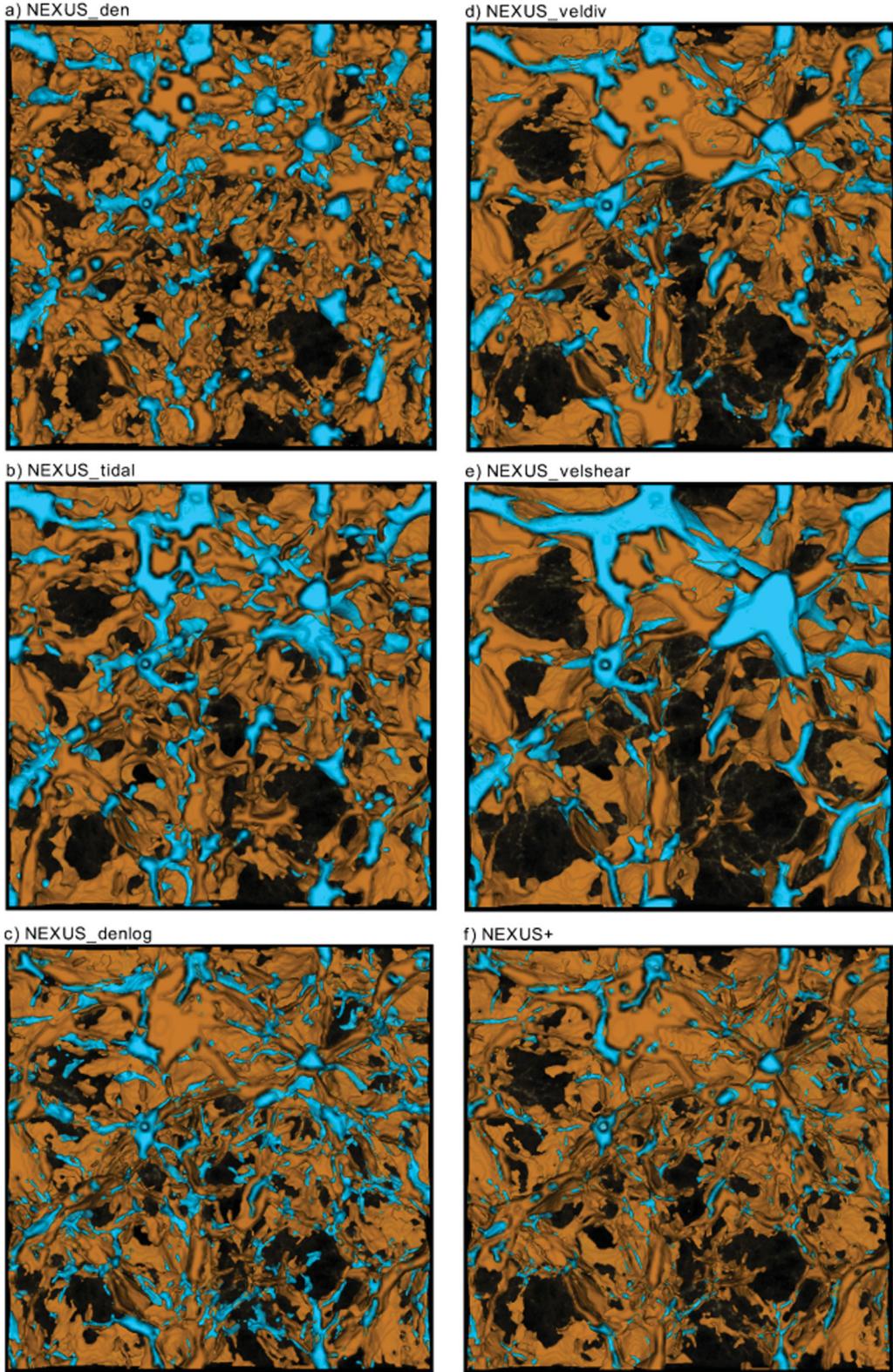


Figure 14. A 3D rendering of the walls (orange) in a $100 \times 100 \times 10 (h^{-1} \text{Mpc})^3$ volume of the *N*-body simulation. The light blue depicts the filaments while the faint background shows the density field. The walls were obtained using (a) NEXUS_{den}, (b) NEXUS_{tidal}, (c) NEXUS_{denlog}, (d) NEXUS_{veldiv}, (e) NEXUS_{velshear} and (f) NEXUS₊. This is the same region of the simulation volume as the density field shown in Fig. 9 and the filaments in Fig. 12.

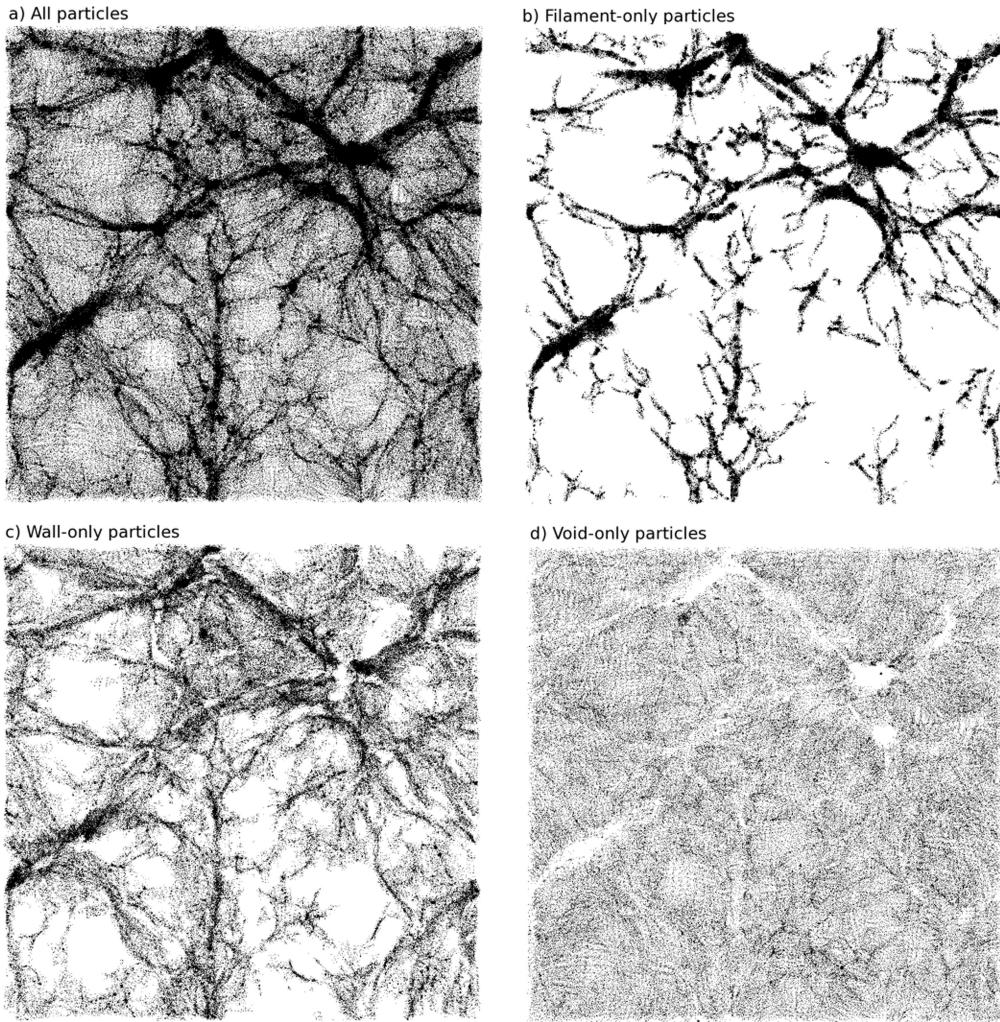


Figure 15. The dark matter particles in the different environments of the cosmic web. The panels give (a) all, (b) filament-only, (c) wall-only and (d) void-only particles. The environments were identified using the NEXUS+ method.

identifying both the prominent as well as the tenuous cosmic web walls. The resulting objects form large continuous planar structures which delineate the different voids. The aspect of the sheets resembles very much that of walls of biological cells that have been stacked on top of each other, with the cells having a wide range of sizes. Even though the cosmic web walls are continuous for large portions of their surface, they are still punctured by a large number of holes that allow for void percolation. While these holes are dependent on the wall identification method, they are a sign of the diffuse and sparse nature of the cosmic walls.

As in the case of the filamentary network, the prominent walls are detected by all the methods, but there are differences when it comes to the sheets in the more underdense regions. Some of the most important features and differences between the six results can be summarized as the following.

(i) The NEXUS_den walls have a very clumpy appearance and this is also true to a lesser extent for the NEXUS_tidal results. This is in contrast with the other methods where the walls have a much more planar and sheet-like look. The clumpy appearance is due to the composition of walls, which are tenuous structures with sporadic haloes from place to place. The concentration of mass in the haloes compared to their neighbourhood regions gives the clumpy structure of the NEXUS_den and NEXUS_tidal walls.

(ii) The NEXUS_veldiv and NEXUS_velshear walls have a smooth planar appearance since the velocity field, compared to the density, is less affected by the sparse Poissonian sampling of the mildly underdense regions. These are the regions that make most of the volume in walls. We find the same smooth planar look also for the NEXUS_denlog and NEXUS+ walls.

(iii) We find that for all the methods the larger gaps in the walls are present at the same locations, indicating that, at least occasionally, there is no clear boundary between adjacent voids.

(iv) The NEXUS_velshear method seems to be the most conservative tracer of walls. It detects only the most prominent wall regions.

A visual inspection favours the NEXUS+, NEXUS_denlog and NEXUS_veldiv as best tracers of the cosmic sheets with the remaining methods either giving clumpy detections or missing some of the more tenuous structures.

By comparing the filaments and walls in Fig. 13 we find that most if not all the filament volume elements are embedded in walls. Though not shown, we find that also the clusters are fully embedded in filaments. This suggests that the strict classification where each volume element is assigned to a single environment can be extended by realizing that clusters are embedded in filaments and that in turn filaments are embedded in walls.

Table 3. The mass and volume filling fractions for the environments identified using NEXUS+. The results presented here are for the larger $200 h^{-1}$ Mpc simulation. The mass and volume filling fractions for the other five methods are given in Appendix A.

Environment	Mass fraction (per cent)	Volume fraction (per cent)
Clusters	8.0	0.027
Filaments	51.3	4.35
Walls	24.0	16.8
Voids	16.7	78.8

6.4 Cosmic web dark matter particle and halo population

An interesting and important issue for the study of structure formation is our understanding on how far the cosmic structure in the dark matter distribution is reflected in the halo distribution. In Fig. 15 we show the particles in our numerical simulation split according to the environment in which they reside. The filament environment correctly traces the largest linear particle concentrations. This can clearly be seen in the prominent filaments. While the thinner structures are less populated with dark matter particles, they are identified as filaments due to their higher local contrast. NEXUS+ finds that on average filaments have an overdensity $1 + \delta$ around 11, as can be seen by comparing the mass and volume filling fractions given in Table 3. On the other hand, walls are much more tenuous structures

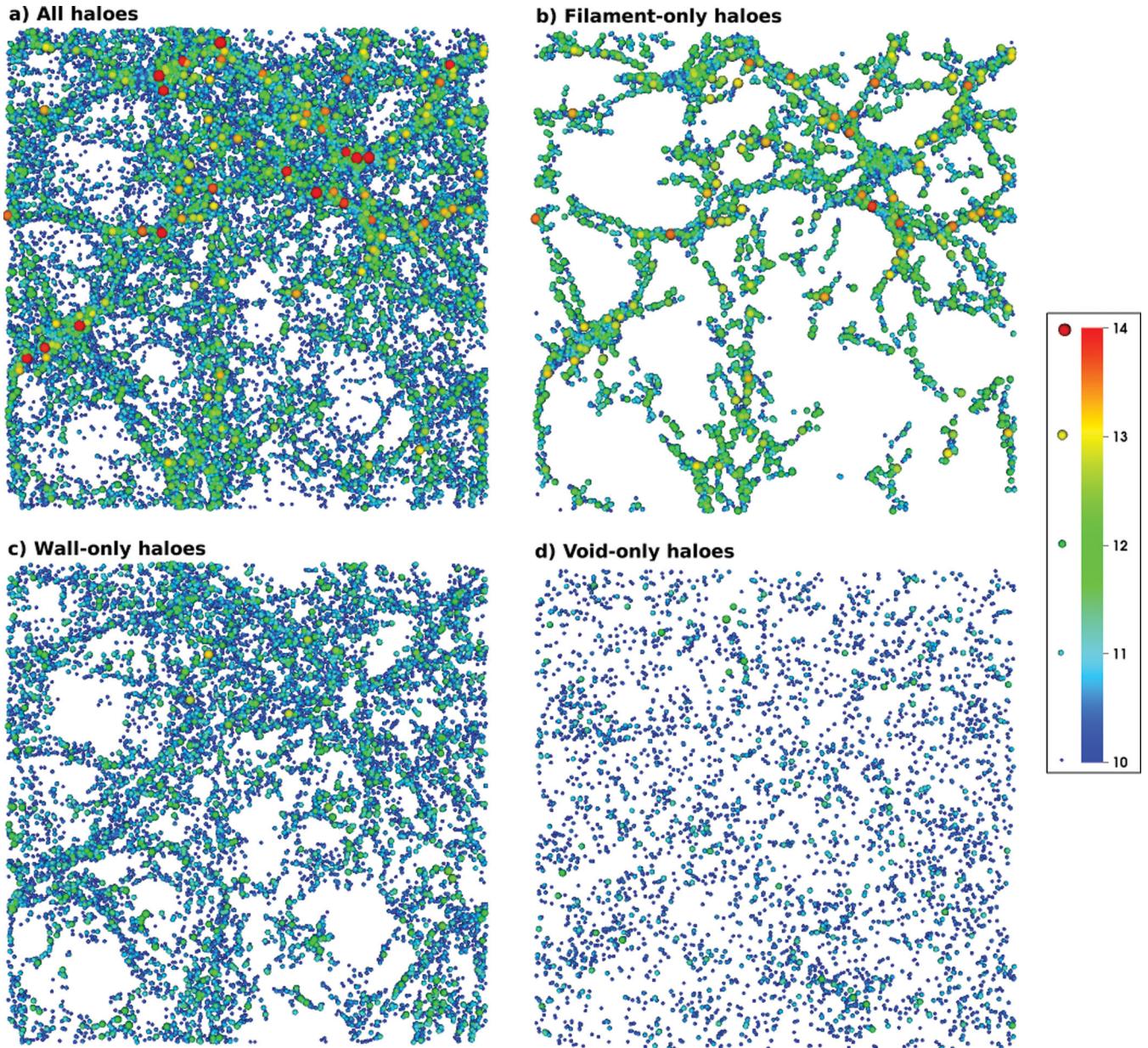


Figure 16. The dark matter haloes as a function of the environment they reside in. The panels give (a) all, (b) filament-only, (c) wall-only and (d) void-only haloes. The environments were identified using the NEXUS+ method. The colour and size of the points are proportional to the mass of the halo they represent. The labels in the legend correspond to $\log_{10}(M/h^{-1} M_\odot)$. Note that most of the smaller mass haloes around massive ones are not visible due to the larger size with which we show the massive haloes.

and this is visible in the particle distribution. When seen edge on, the sheets appear as prominent structures in the particle distribution. But when looked at face on, they are sparsely populated by particles and hence difficult to identify. When comparing the particle distributions in walls and voids there seems to be little difference in the particle densities. This is just a projection effect, with walls having on average an overdensity around 1.4 versus an overdensity of 0.2 for voids (see Table 3 for details). The success of the structure finding method can be easily assessed by observing that the dark matter particles in void regions do not have any significant structure present in their distribution.

A more interesting picture is found when looking at the variation of the dark matter halo populations with the cosmic web environment. This is shown in Fig. 16 where the haloes are coloured and scaled according to their mass. Fig. 16, in conjunction with Fig. 17 which gives the cumulative halo mass function split according to environment, offers a very suggestive picture. We find that all of the massive haloes with $M \geq 5 \times 10^{13} h^{-1} M_\odot$ are located in cluster environments.⁶ Moreover, clusters are the most crowded regions when it comes to haloes of all masses, with halo overdensities about 10 times larger than in filaments and about 100 times larger than in the full simulation box (see lower panel in Fig. 17).

When looking at filaments we see that these environments are also crowded when it comes to haloes. Most of the $10^{12} h^{-1} M_\odot$ and higher mass haloes are located in filaments, while the lower mass objects are a factor of 10 more common in filaments than on average in the universe. From Fig. 16 we see that even the more tenuous filaments have a large number of haloes which shows the power of our method to correctly identify the filamentary environments.

The walls are dominated by low-mass objects, with the sheets containing only a significant share of the $10^{12} h^{-1} M_\odot$ and lower mass haloes. For a few times $10^{11} h^{-1} M_\odot$ and lower mass haloes, the walls have a similar halo density as the average universe. Compared to voids, the sheets clearly have a much higher halo density and are populated by more massive haloes. In contrast, the voids are very sparsely populated, with extremely few $10^{11} h^{-1} M_\odot$ and higher mass haloes. As in the case of the particle distribution, there does not seem to be any significant structures present in the void halo distribution.

6.5 Single-scale versus multiscale analysis

One of the frequent questions that surface when dealing with cosmological structure identification is the optimal value of the smoothing scale (see Hahn et al. 2007a; Forero-Romero et al. 2009). Often this constitutes a major limitation of single-scale approaches. Moreover, until now, the choice of one smoothing length versus another was mostly heuristic. While our multiscale approach does not suffer from these problems, it is certainly a very interesting question to find which scales are the most important and how they compare with results from other works.

We exemplify the effects of a single-scale versus a multiscale approach in Fig. 18. To obtain the single-scale results we restrict the NEXUS+ method to a single filter radius. We then show this result for several values of the filter radius. For small smoothing scales, the method detects most of the filaments. While it reproduces correctly

⁶ Remember that the $5 \times 10^{13} h^{-1} M_\odot$ mass threshold was introduced as a lower mass cut-off in Section 6.1. While this lower mass threshold depends on the cluster definition one chooses, the method is very successful in identifying as clusters all the haloes above the cut-off mass.

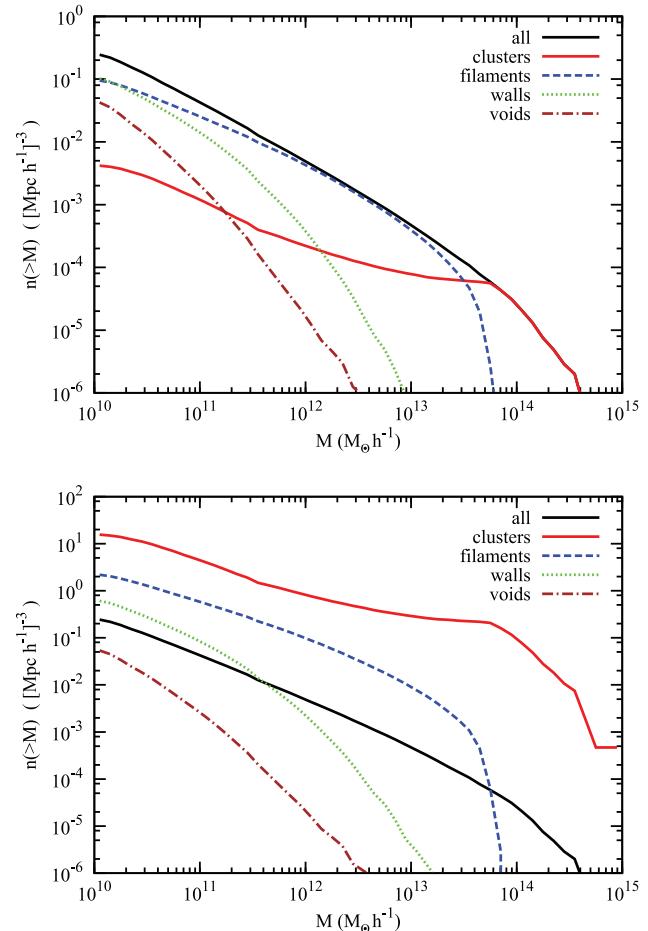


Figure 17. The cumulative halo mass function segmented according to the components of the cosmic web as identified by NEXUS+. The upper panel gives the cumulative mass function normalized according to the volume of the whole simulation box. The lower panel gives the cumulative mass function normalized according to the volume of each environment (see Table 3).

the small filaments, it greatly underestimates the thickness of the more prominent objects. As the filter radius is increased, most of the thinner filaments are missed while the larger ones are detected as being thicker and thicker. In conclusion, applying a filter of a given radius makes the environment detection method sensitive only to objects similar or larger than that smoothing scale. Moreover, at a given filter radius, all filaments have very similar diameters which are given by the value of the smoothing radius.

By comparing the single-scale results versus the full NEXUS+ results (see Fig. 18, panel d) we conclude that the $2 h^{-1}$ Mpc smoothing scale gives the closest match to the multiscale filaments. However, there are a lot of important differences, with many thin filaments missing in the single-scale picture. Another striking difference is the smaller diameters for the larger filaments in the single-scale versus the full NEXUS+ results. These results strongly show the need of a multiscale approach to be able to fully trace all the features of the cosmic web.

When restricted to a single scale, the NEXUS_den and NEXUS_veldiv methods show an even stronger difference between results at different filter radii. This trend is so strong that it is very difficult to find a single smoothing scale that matches even remotely the multiscale results. Even the single-scale NEXUS_tidal and

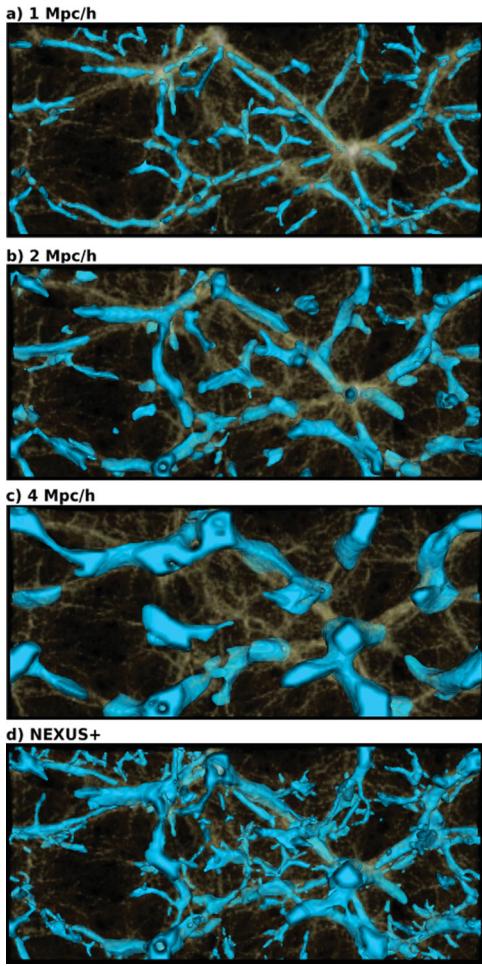


Figure 18. A comparison of the single-scale versus the multiscale approach. Panels (a)–(c) show the single-scale filaments obtained by restricting the NEXUS+ method to the single filter radius of 1, 2 and $4 h^{-1}$ Mpc, respectively. Panel (d) shows the full NEXUS+ results.

NEXUS_velshear methods show significant differences between scales, though less prominent than the results in Fig. 18. The weaker dependence on scale is due to the steep drop at small wavelengths of the gravitational and velocity potentials spectra (see Fig. 6) which means that large-scale modes contribute more than for the other methods. For the single-scale NEXUS_tidal and NEXUS_velshear methods we find that $1 h^{-1}$ Mpc smoothing radius results are the closest match to the full multiscale results. This is in contrast with Hahn et al. (2007a), who argue for a $2.1 h^{-1}$ Mpc filter radius.

Another very interesting question is finding the smoothing scale at which a region has the largest environmental signature. This is illustrated in Fig. 19. We show in different colours the filter radius which gives the strongest filamentary characteristic for a given region. The small filter radii give the strongest signature in the thinner filaments while the larger smoothing scales give a stronger signal for the major filaments. A closer inspection of the prominent filaments shows that their central axes give the strongest filamentary response for small filter radii. Increasing the size of the filter adds larger and larger filamentary regions around this inner central axis. This is visible in panels (b)–(d) of Fig. 19. This is due to the inner structure of the filaments, with their central axis having a larger density than their periphery.

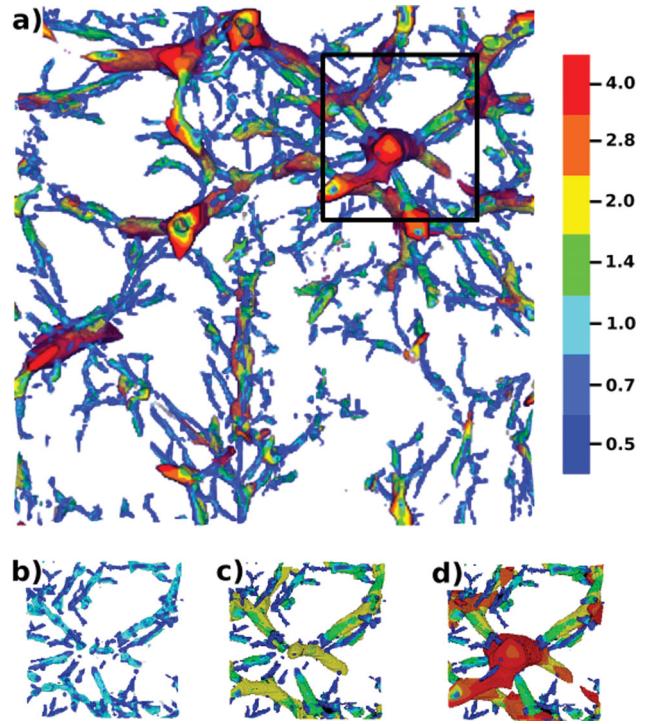


Figure 19. The NEXUS+ filaments coloured according to the smoothing scale that gives the largest filamentary signature for that voxel. The corresponding colour bar indicates the filter scale in units of h^{-1} Mpc. The lower panels (b)–(d) show a smaller region from the larger volume presented in panel (a). These panels show the filaments detected using a maximum of (b) 1, (c) 2 and (c) $4 h^{-1}$ Mpc smoothing scales.

7 COMPARISON TO OTHER STRUCTURE FINDING ALGORITHMS

In this section we discuss the abilities and virtues of the NEXUS and NEXUS+ algorithms with respect to those of other structure finding algorithms. It will underline the advantages and disadvantages of the instruments described and introduced in this paper. In our discussion we limit ourselves to Hessian and topological-based methods. There are a few additional methods that find filaments using directly the particle/galaxy distribution (Stoica et al. 2005, 2007, 2010; Chazal et al. 2009). While these approaches have the advantage that one does not need to compute the density field, they depend on many free parameters that make their use cumbersome.

A first group of cosmic web identification methods consists of algorithms that use the Hessian of the density or gravitational potential. Both the NEXUS and NEXUS+ algorithms are part of this class.

Hahn et al. (2007a) (hereafter referred to as HPCD) proposed the use of the tidal field eigenvalues for environment classification.⁷ The method uses the criterion that all positive eigenvalues identify clusters, one negative eigenvalue corresponds to filaments while two negative eigenvalues trace sheets. The result of this method applied to our simulation is shown in Fig. 20. It is immediately clear that while the inner region of the HPCD results correspond to NEXUS_tidal (compare to panel b in Fig. 12), the HPCD filaments extend to much larger diameters, encompassing substantial parts of wall and void regions. This leads to a large cross-contamination of

⁷ This method can be easily implemented as a special case of our NEXUS_tidal algorithm.

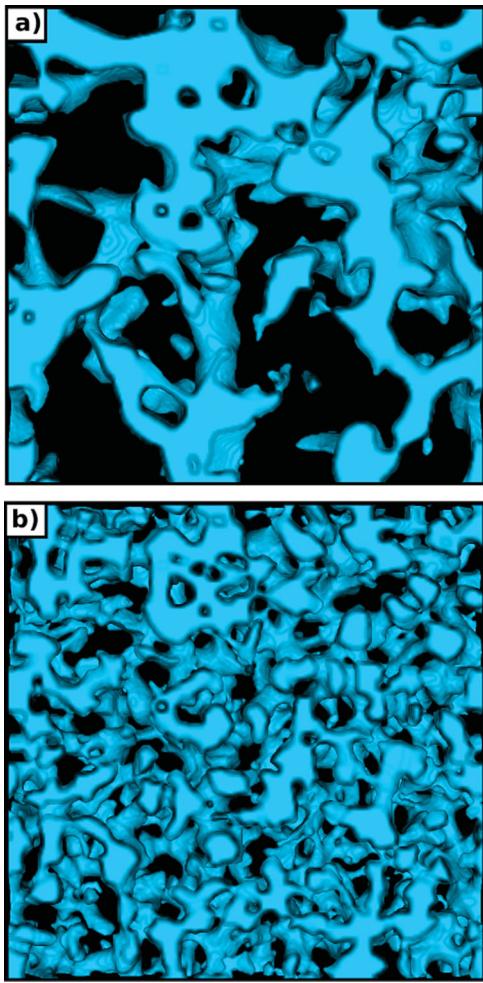


Figure 20. The filamentary network as identified by the algorithms presented in (a) Hahn et al. (2007a) and (b) Zhang et al. (2009) (see text for details). The results are obtained for a $2 h^{-1}$ Mpc filter width. It shows the same volume of the N -body simulation as in Figs 9 and 12.

the cosmic web components. A similar method to HPCD, but using the density Hessian instead of the tidal field, was used by Zhang et al. (2009). From Fig. 20, panel (b), it is obvious that this last method is not very successful. It leads to even more misclassified regions than the HPCD method. There are two main shortcomings of the two methods: use of a single-scale approach and the absence of a threshold to distinguish between significant and spurious detections.

The absence of a detection threshold in the HPCD method was pointed out in Forero-Romero et al. (2009). They suggested that an eigenvalue threshold in the range 0.2–0.4 gives results in agreement with the visual impression of the cosmic web. Using NEXUS_tidal restricted to the $2 h^{-1}$ Mpc filter radius we do obtain that the filament and wall threshold is 0.42 and 0.2, respectively. While we do confirm the results of Forero-Romero et al. (2009), we stress that NEXUS_tidal has the advantage of a multiscale approach and does not employ user-dependent arguments for specifying the detection threshold.

It is the multiscale character of NEXUS and NEXUS+ which are crucial for their successful analysis of emerging structures and patterns in the hierarchically evolving cosmic mass distribution. The first version of the multiscale formalism that we have developed into NEXUS and NEXUS+ is the MMF, described and introduced in Aragón-Calvo et al. (2007b). NEXUS is an extension of MMF to

a more physical and versatile algorithm. While MMF is restricted to the density field, the NEXUS method incorporates, among others, the tidal and velocity fields. Both MMF and NEXUS_den find the same filamentary structures, while the second method gives a much better identification of walls (for details see Appendix A). More importantly, NEXUS+ seems optimal at capturing the structural intricacies of the cosmic web, and thus represents an important advancement within the context of the multiscale scale-space formalism that we have been developing.

An additional important class of structure identification procedures is based on topological considerations, in principle following an analysis of the Morse–Smale complex of the density field (Sousbie et al. 2008a; Aragón-Calvo et al. 2010a; Sousbie 2011). Given the fact that in general, cosmic density fields behave like a proper Morse function, the assumption is that filaments and walls in the mass distribution can be identified with the manifolds in the density field connecting maxima via saddle points of the field (Sousbie 2011). Filaments are identified with the line connecting two maxima via a saddle point, and walls with the sheet separating the regions around two minima and centred around a saddle point. Aragón-Calvo et al. (2010a) developed a similar strategy by first delineating the watershed basins around the minima in the field, following the WVF procedure introduced by Platen et al. (2007) in which the watershed basins are identified with the voids in the cosmic mass distribution. The topological character of the boundary is determined locally via the number of touching watershed basins. Walls are the 2D manifolds separating two watershed basins, the filaments are their boundaries and identified as the locations where three basins touch each other.

NEXUS and NEXUS+ do find that the largest filaments are between massive density peaks, which is in agreement with the implicit assumptions of the topological methods. But on top of the prominent structures, we also find, especially in the NEXUS+ results, an important network of thinner objects which branches into voids. These tenuous objects contradict the hypothesis that filaments are always located between density maxima. We also find that walls are not fully continuous sheets and that they sometimes stop as they branch into lower density regions. Thus, the voids fully percolate, with large regions without a clear boundary between adjacent voids. Also, we should note that while the topological based methods can detect the central axis of filaments and inner plane of walls, they cannot assign a natural thickness to the structures. As our results show, both the filament and wall environments have a wide range of sizes. This makes it rather challenging for the topological methods to detect and outline in a natural fashion the regions belonging to a filament around the identified central filament axis.

8 CONCLUSIONS AND PROSPECTS

This work presents the NEXUS and NEXUS+ methods, which are multiscale and automatic algorithms used for the segmentation of the cosmic web into its distinct components: clusters, filaments, walls and voids. We have shown that the environments identified with the two methods correspond very well to the structures visible in the density and velocity divergence fields as well as in the dark matter particle and halo distributions. The success of the method lies in two important ingredients: the use of a scale-free approach that makes sure that the algorithm detects structures of all sizes, and the use of a physically motivated threshold to distinguish valid environments from spurious detections. Another strength of the two algorithms is that they do not depend on user set parameters and

therefore can be easily applied in a consistent way to multiple data sets.

We have extended the NEXUS method to detect the cosmic web as traced by the density (NEXUS_den), tidal field (NEXUS_tidal), velocity divergence (NEXUS_veldiv) and velocity shear fields (NEXUS_velshear). We find that NEXUS_den and NEXUS_tidal are very efficient in identifying the cluster regions when compared with the most massive haloes present in the simulation. The methods perform similarly in the detection of filament and wall regions, with all the prominent structures detected by all methods. The only differences arise in the identification of the more tenuous structures. We find that NEXUS+ performs better in tracing the weaker filaments and walls.

From all the methods presented in this paper, we find that NEXUS+ is the most successful one in tracing the cosmic web. Its main advantage comes from the use of the log-Gaussian filter which is designed to better deal with the orders of magnitude difference in the density field between low- and high-density regions. The filamentary and wall-like environments detected with NEXUS+ contain complex networks of prominent structures that branch out into more tenuous ones until they finally disappear out into underdense regions.

We showed in Section 7 that NEXUS and NEXUS+ have several advantages compared to other Hessian-based methods: the use of a multiscale approach and a physically motivated threshold for identifying the significant environments. Compared to topological methods, our tools are able to detect the filamentary/wall regions and not only their central axis/plane. Moreover, we do not make the assumptions that filaments extend between density maxima and walls separate density minima basins. Our results show that these assumptions do not always hold.

Equipped with the NEXUS and NEXUS+ methods, we plan to address a range of cosmological issues. The ability to identify filaments and walls over a range of scales, in both numerical and observational data sets, allows us to study not only environmental factors affecting the formation and evolution of dark matter haloes and galaxies, but also the hierarchical build-up of the cosmic web itself. Our first priority lies with a systematic study of environmental factors affecting the evolution of galaxies. The fact that we can identify galaxies and haloes within finely outlined filaments and sheets will allow us to determine which physical characteristics are most sensitive to the environment, which galaxies and haloes are most sensitive to large-scale influences, and to study the causes that give rise to this dependency.

A major point of our interest is performing a systematic comparison between the structures traced by the density, velocity and tidal fields in order to understand which of these physical influences are most decisive in determining the global outline of the cosmic web. In addition to our current focus on the large-scale structure of the dark matter distribution, we will also direct our study to the structure of the gaseous cosmic web. Comparison of the intergalactic medium (IGM) with the dark matter structures in numerical simulations will be instrumental in understanding how the cosmic web can be traced both in the galaxy and cosmic gas distribution. This will be essential for relating the distribution of H I in the local Universe to the overall large-scale structure found in the galaxy distribution (Popping & Braun 2011), and will help understand recent findings such as a small-scale H I filament in a void (Beygu et al., in preparation).

Finally, the NEXUS and NEXUS+ procedures are perfectly suited for a systematic appraisal of the structures found in maps produced by galaxy redshift surveys such as SDSS and 2MASS Redshift Survey (2MRS).

ACKNOWLEDGMENTS

We are grateful to Miguel Aragón-Calvo for useful discussions about the MMF method and for his early IDL implementation. We also thank Wojciech Hellwing and Carlos Frenk for inspiring discussions and suggestions. We are grateful to the anonymous referee whose comments and questions improved the presentation of this paper. All simulations and analysis were performed on the Gemini machines at the Kapteyn Astronomical Institute, the Netherlands.

REFERENCES

- Abel T., Hahn O., Kaehler R., 2012, MNRAS, 427, 61
 Aragón-Calvo M. A., 2007, PhD thesis, Univ. Groningen, the Netherlands
 Aragón-Calvo M. A., Jones B. J. T., van de Weygaert R., van der Hulst J. M., 2007a, ApJ, 655, L5
 Aragón-Calvo M. A., Jones B. J. T., van de Weygaert R., van der Hulst J. M., 2007b, A&A, 474, 315
 Aragón-Calvo M. A., Platen E., van de Weygaert R., Szalay A. S., 2010a, ApJ, 723, 364
 Aragón-Calvo M. A., van de Weygaert R., Araya-Melo P. A., Platen E., Szalay A. S., 2010b, MNRAS, 404, L89
 Bardeen J. M., Bond J. R., Kaiser N., Szalay A. S., 1986, ApJ, 304, 15
 Barrow J., Bhavsar S., Sonoda D., 1985, MNRAS, 216, 17
 Bernardeau F., van de Weygaert R., 1996, MNRAS, 279, 693
 Bernardeau F., Chodorowski M. J., Łokas E. L., Stompor R., Kudlicki A., 1999, MNRAS, 309, 543
 Bond J. R., Kofman L., Pogosyan D., 1996, Nat, 380, 603
 Bond N. A., Strauss M. A., Cen R., 2010a, MNRAS, 406, 1609
 Bond N. A., Strauss M. A., Cen R., 2010b, MNRAS, 409, 156
 Cautun M. C., van de Weygaert R., 2011, preprint (arXiv:1105.0370)
 Chazal F., Cohen-Steiner D., Mérigot Q., 2009, INRIA Rapp. Récherche, 6930
 Chodorowski M. J., Łokas E. L., 1997, MNRAS, 287, 591
 Choi E., Bond N. A., Strauss M. A., Coil A. L., Davis M., Willmer C. N. A., 2010, MNRAS, 406, 320
 Colberg J., 2007, MNRAS, 375, 337
 Colberg J. M., Krughoff K. S., Connolly A. J., 2005, MNRAS, 359, 272
 Coles P., Jones B., 1991, MNRAS, 248, 1
 Colless M. et al., 2003, preprint (astro-ph/0306581)
 Colombi S., Pogosyan D., Souradeep T., 2000, Phys. Rev. Lett., 85, 5515
 de Lapparent V., Geller M., Huchra J., 1986, ApJ, 302, L1
 Dolag K., Meneghetti M., Moscardini L. E. R., Bonaldi A., 2006, MNRAS, 370, 656
 Doroshkevich A. G., Kotok E. V., Poliudov A. N., Shandarin S. F., Sigov I. S., Novikov I. D., 1980, MNRAS, 192, 321
 Florack L., ter Haar Romeny B., Koenderink J., Viergever M., 1992, Image Vision Comput., 10, 376
 Forero-Romero J. E., Hoffman Y., Gottlöber S., Klypin A., Yepes G., 2009, MNRAS, 396, 1815
 Frangi A. F., Niessen W. J., Vincken K. L., Viergever M. A., 1998, in Wells W., Colchester A., Delp S., eds, Multiscale vessel enhancement filtering, Medical Image Computing and Computer-Assisted Intervention-MICCAI'98, Lecture Notes in Computer Science, Vol. 1496. Springer-Verlag, Berlin, p. 130
 Geller M. J., Huchra J. P., 1989, Sci, 246, 897
 Genovese C. R., Perone-Pacifico M., Verdinelli I., Wasserman L., 2010, preprint (arXiv:1003.5536)
 Gill S. P. D., Knebe A., Gibson B. K., 2004, MNRAS, 351, 399
 González R. E., Padilla N. D., 2010, MNRAS, 407, 1449
 Graham M., Clowes R., 1995, MNRAS, 275, 790
 Gregory S. A., Thompson L. A., 1978, ApJ, 222, 784
 Gunn J. E., Gott J. R. III, 1972, ApJ, 176, 1
 Gurbatov S. N., Saichev A. I., Shandarin S. F., 1989, MNRAS, 236, 385
 Hahn O., 2009, PhD thesis, ETH, Zürich
 Hahn O., Carollo C. M., Porciani C., Dekel A., 2007a, MNRAS, 375, 489 (HPCD)

- Hahn O., Carollo C. M., Porciani C., Dekel A., 2007b, MNRAS, 381, 41
 Hahn O., Porciani C., Carollo C. M., Dekel A., 2009, MNRAS, 398, 1742
 Huchra J. et al., 2005, in Fairall K. P., Woudt P. A., eds, ASP Conf. Ser. Vol. 239, Nearby Large-Scale Structures and the Zone of Avoidance. Astron. Soc. Pac., San Francisco, p. 135
 Jenkins A. et al., 1998, ApJ, 499, 20
 Joeveer M., Einasto J., 1978, in Longair M. S., Einasto J., eds, Proc. IAU Symp. 79, Large Scale Structures in the Universe. p. 241
 Jones B. J. T., van de Weygaert R., Aragón-Calvo M. A., 2010, MNRAS, 408, 897
 Kirshner R. P., Oemler A. J., Schechter P. L., Shectman S. A., 1981, ApJ, 248, L57
 Knollmann S. R., Knebe A., 2009, ApJS, 182, 608
 Li Q., Sone S., Doi K., 2003, Medical Phys., 30, 2040
 Lindeberg T., 1998, Int. J. Comput. Vision, 30, 77
 Melott A. L., 1983, MNRAS, 204, 7p
 Neyrinck M. C., 2012, MNRAS, 427, 494
 Novikov D., Colombi S., Doré O., 2006, MNRAS, 366, 1201
 Nusser A., Dekel A., Bertschinger E., Blumenthal G. R., 1991, ApJ, 379, 6
 Pauls J. L., Melott A. L., 1995, MNRAS, 274, 99
 Paz D. J., Stasyszyn F., Padilla N. D., 2008, MNRAS, 389, 1127
 Peebles P. J. E., 1980, The Large Scale Structure of the Universe. Princeton Univ. Press, Princeton, NJ
 Pimbblet K., 2005, MNRAS, 358, 256
 Platen E., van de Weygaert R., Jones B. J. T., 2007, MNRAS, 380, 551
 Popping A., Braun R., 2011, A&A, 527, A90
 Romano-Díaz E., van de Weygaert R., 2007, MNRAS, 382, 2
 Sathyaprakash B. S., Sahni V., Shandarin S. F., 1996, ApJ, 462, L5
 Sato Y., Nakajima S., Shiraga N., Atsumi H., Yoshida S., Koller T., Gerig G., Kikinis R., 1998, Medical Image Analysis, 2, 143
 Schaap W. E., van de Weygaert R., 2000, A&A, 363, L29
 Shandarin S. F., 2011, J. Cosmol. Astropart. Phys., 5, 15
 Shandarin S. F., Zeldovich Y. B., 1989, Rev. Modern Phys., 61, 185
 Shandarin S. F., Sheth J. V., Sahni V., 2004, MNRAS, 353, 162
 Shandarin S., Habib S., Heitmann K., 2010, Phys. Rev. D, 81, 103006
 Shandarin S., Habib S., Heitmann K., 2012, Phys. Rev. D, 85, 083005
 Shapiro P. R., Struck-Marcell C., Melott A. L., 1983, ApJ, 275, 413
 Shectman S. A., Landy S. D., Oemler A., Tucker D. L., Lin H., Kirshner R. P., Schechter P. L., 1996, ApJ, 470, 172
 Sousbie T., 2011, MNRAS, 414, 350
 Sousbie T., Pichon C., Colombi S., Novikov D., Pogosyan D., 2008a, MNRAS, 383, 1655
 Sousbie T., Pichon C., Courtois H., Colombi S., Novikov D., 2008b, ApJ, 672, L1
 Sousbie T., Pichon C., Kawahara H., 2011, MNRAS, 414, 384
 Springel V., 2005, MNRAS, 364, 1105
 Springel V. et al., 2005, Nat, 435, 629
 Stoica R. S., Martínez V. J., Mateu J., Saar E., 2005, A&A, 434, 423
 Stoica R. S., Martínez V. J., Saar E., 2007, J. R. Stat. Soc. Ser. C Appl. Stat., 56, 459
 Stoica R. S., Martínez V. J., Saar E., 2010, A&A, 510, A38
 Tegmark M. et al., 2004, ApJ, 606, 702
 Tempel E., Stoica R. S., Saar E., 2012, preprint (arXiv:1207.0068)
 van de Weygaert R., Schaap W., 2009, in Martínez V. J., Saar E., Martínez-González E., Pons-Bordería M.-J., eds, Lecture Notes in Physics, Vol. 665, Data Analysis in Cosmology. Springer-Verlag, Berlin, p. 291
 Voit G. M., 2005, Rev. Modern Phys., 77, 207
 Wu Y., Batuski D., Khalil A., 2009, ApJ, 707, 1160
 Zel'dovich Y. B., 1970, A&A, 5, 84
 Zeldovich I. B., Einasto J., Shandarin S. F., 1982, Nat, 300, 407
 Zhang Y., Yang X., Faltenbacher A., Springel V., Lin W., Wang H., 2009, ApJ, 706, 747

APPENDIX A: OPTIMAL FILAMENT AND WALL DETECTION

This section deals with identifying the signature threshold used for the detection of the cosmic web filaments and walls. When ap-

plying the environment detection algorithm every region of space is assigned an environment signature \mathcal{S} (for definition see equation 7). A large signature corresponds to very prominent structures while a zero or small one corresponds to null detections. Since many regions of space will have a signature value between the two extremes, we need to identify a signature threshold that differentiates between valid structures and spurious detections. All regions with signatures larger than the threshold will correspond to valid environments.

The simplest way to define the cosmic web is using the tidal field, since this is what drives the anisotropic collapse. Using the eigenvalues of the tidal tensor one defines filaments and walls as regions with one and two negative eigenvalues (see HPCD). The major problem with this approach is that it gives unrealistic looking environments, with only a small fraction of the volume (~ 10 per cent) occupied by voids – this is in stark contrast with the observational data.

A second approach is to detect only the most significant filaments and walls. This is the procedure we follow in this paper. Aragón-Calvo et al. (2007b) have argued that the percolation threshold of filaments/walls offers a natural way of identifying the prominent structures. We find that indeed the percolation threshold gives a good identification of the filamentary network, but fails in detecting the sheets. The walls obtained via this method are made of many small patches that do not show the large-scale cohesion expected for void boundaries (see Fig. A1). Another downside is the dependence of the percolation threshold on the grid resolution used to analyse the data. But more importantly for this work, many large differences between the results of the six methods described here can be attributed to different percolation properties of the filamentary/wall networks in each method. We analyse this in more details later on.

Most of the mass in filaments and walls is given by regions with a narrow range in environment signatures. This is illustrated by the red curve of Fig. A2 which gives the mass fraction M in filament/wall regions as a function of the signature threshold S . The rapid increase in the filament/wall mass can be appreciated when computing the mass change:

$$\Delta M^2 = \left| \frac{dM^2}{d \log \mathcal{S}} \right|. \quad (\text{A1})$$

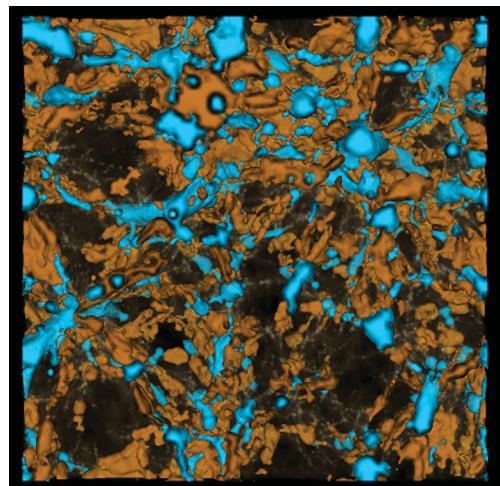


Figure A1. A 3D volume rendering of the NEXUS_{den} filaments (blue) and walls (orange) when the detection threshold is taken as the percolation point.

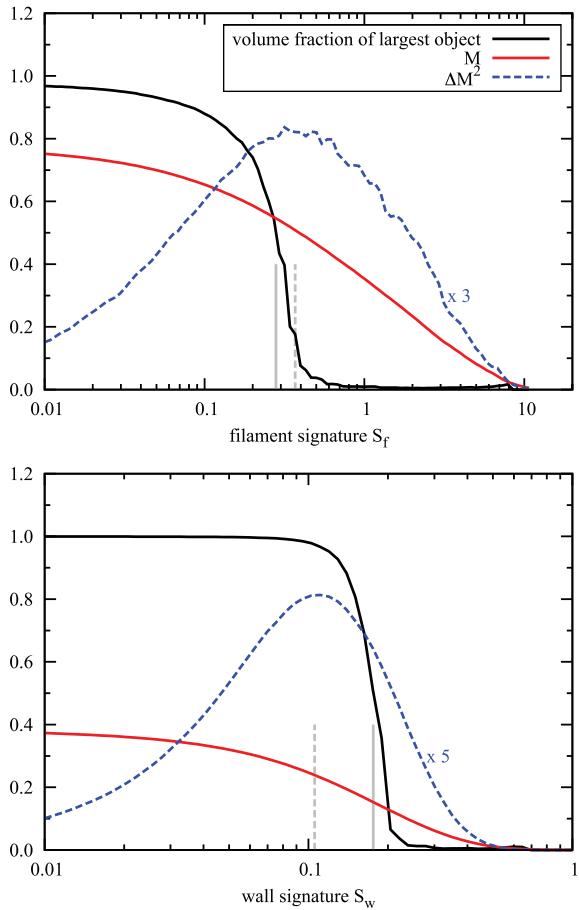


Figure A2. The dependence of the mass fraction in the components of the cosmic web as a function of environment signature. The two panels show the dependence for filaments and walls obtained using the NEXUS_{den} method, respectively. The continuous red curve gives the mass fraction M in filaments/walls while the dashed blue curve gives ΔM^2 (see text for definition). The black curve gives the volume fraction of the largest filament/wall. The sharp transition of this curve from 0 to 1 marks the percolation threshold. The vertical grey lines mark the percolation point (filled line) and the peak of ΔM^2 (dashed line).

This is represented by the dashed blue curve in Fig. A2. We found that the peak position in ΔM^2 gives a robust and natural way of identifying the most significant filaments and walls. The environments detected using the ΔM^2 peak threshold are shown in Figs 12 and 14. This new threshold captures very well the filamentary and wall features seen in the density and velocity divergence fields from Fig. 9. Moreover the peak of ΔM^2 for the NEXUS_{den} filaments is very close to the percolation threshold and hence reproduces the success of using percolation as a good tracer of the filamentary network.⁸

⁸ The percolation threshold and the ΔM^2 peak are similar only for the NEXUS_{den} filaments, in general there are large offsets between the two values. This was expected since Aragón-Calvo et al. (2007b) have shown the success of the percolation threshold for filament detection only when using the density field as tracer of cosmic web environments.

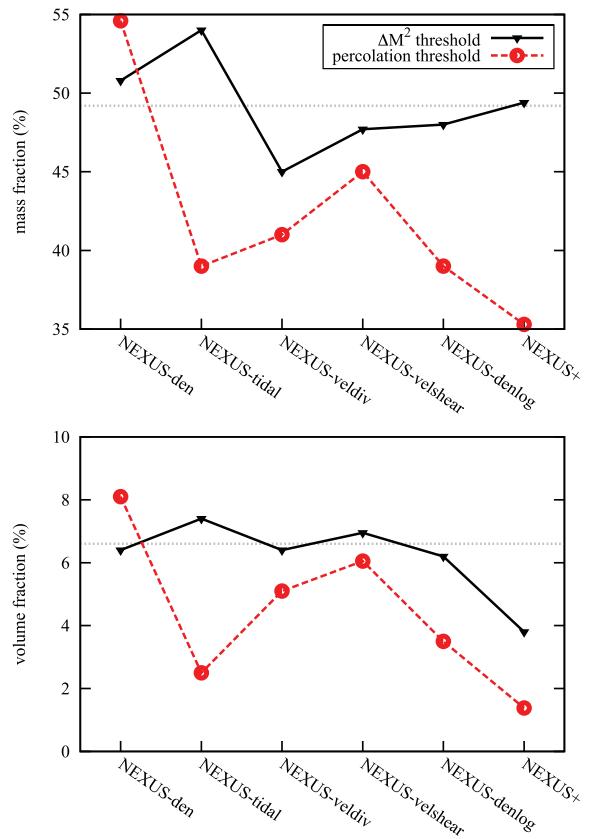


Figure A3. Comparison of the mass fraction (upper panel) and volume fraction (lower panel) in filaments detected using the six methods described in this paper. The continuous black line describes the filaments detected using the ΔM^2 threshold method (described in this section) while the dashed red line shows the filaments detected using the percolation threshold (Aragón-Calvo et al. 2007b). The horizontal dotted line gives the average mass and volume fraction for the ΔM^2 results.

There are two major improvements when using the ΔM^2 peak versus the percolation point as the detection threshold for the most significant environments. The walls detected via the ΔM^2 method have a bigger large-scale cohesion than the results of the percolation method. This can clearly be seen by comparing the NEXUS_{den} walls from Fig. 14 and the percolation walls in Fig. A1. The second enhancement comes from a more robust detection threshold. This can be seen in Fig. A3 where we compare the mass and volume fraction in filaments detected using the six methods introduced in this paper. Notice the large variation in mass fraction between the NEXUS_{den} and NEXUS+ results for the percolation method – with the former having almost twice as much mass than the latter one. In the case of the ΔM^2 threshold all methods give a similar mass fraction, with a much smaller scatter around the mean. The same behaviour can be seen in the volume fraction plot, where the ΔM^2 threshold gives more consistent values. The only exception is the NEXUS+ volume fraction whose lower value is due to the method itself and not the detection threshold used (see Section 6.2 for more details).