

# Multiscale phenomenology of the cosmic web

Miguel A. Aragón-Calvo,<sup>1,2</sup> Rien van de Weygaert<sup>2\*</sup> and Bernard J. T. Jones<sup>2</sup>

<sup>1</sup>The Johns Hopkins University, 3701 San Martin Drive, Baltimore, MD 21218, USA

<sup>2</sup>Kapteyn Astronomical Institute, University of Groningen, PO Box 800, 9700 AV, Groningen, the Netherlands

Accepted 2010 June 28. Received 2010 June 23; in original form 2010 June 2

## ABSTRACT

We analyse the structure and connectivity of the distinct morphologies that define the cosmic web. With the help of our multiscale morphology filter (MMF), we dissect the matter distribution of a cosmological  $\Lambda$  cold dark matter  $N$ -body computer simulation into cluster, filaments and walls. The MMF is ideally suited to address both the anisotropic morphological character of filaments and sheets, and the multiscale nature of the hierarchically evolved cosmic matter distribution. The results of our study may be summarized as follows. (i) While all morphologies occupy a roughly well-defined range in density, this alone is not sufficient to differentiate between them given their overlap. Environment defined only in terms of density fails to incorporate the intrinsic dynamics of each morphology. This plays an important role in both linear and non-linear interactions between haloes. (ii) Most of the mass in the Universe is concentrated in filaments, narrowly followed by clusters. In terms of volume, clusters only represent a minute fraction and filaments not more than 9 per cent. Walls are relatively inconspicuous in terms of mass and volume. (iii) On average, massive clusters are connected to more filaments than low-mass clusters. Clusters with  $M \sim 10^{14} M_{\odot} h^{-1}$  have on average two connecting filaments, while clusters with  $M \geq 10^{15} M_{\odot} h^{-1}$  have on average five connecting filaments. (iv) Density profiles indicate that the typical width of filaments is  $2 h^{-1}$  Mpc. Walls have less well-defined boundaries with widths between 5 and 8 Mpc  $h^{-1}$ . In their interior, filaments have a power-law density profile with slope  $\gamma \approx -1$ , corresponding to an isothermal density profile.

**Key words:** methods: data analysis – methods: numerical – cosmology: theory – large-scale structure of Universe.

## 1 INTRODUCTION

The megaparsec matter distribution in the Universe represents a dynamical system of great structural and topological complexity, the cosmic web.

Early attempts to map the large-scale distribution of galaxies in the Universe (Gregory, Thompson & Tift 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). These preliminary studies suggested that the Universe on the large scales could be described as a cellular system (Joeveer & Einasto 1978) or a *cosmic web* (Bond, Kofman & Pogosyan 1996).

This has been confirmed in recent times by large galaxy surveys like the Two-degree Field 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 advent of these large 3D maps of the local Universe unveiled a cosmos of considerable richness and complexity, featuring intricate filamentary structures. These structures can be seen on scales from a few megaparsecs up to tens and even hundreds of megaparsecs. They include immense elongated and semiplanar patterns (Diaferio & Geller 1997; Tittley & Henriksen 2001; Bharadwaj & Pandey 2004; Ebeling, Barrett & Donovan 2004; Pimblet & Drinkwater 2004; Pimblet, Drinkwater & Hawkrigg 2004; Stevens et al. 2004; Pimblet 2005a) and include huge wall-like structures such as the Coma Great Wall (Geller & Huchra 1989) and the Sloan Great Wall (Gott et al. 2005), with its size being more than  $400 h^{-1}$  Mpc, the largest known structure in the nearby Universe. Similar web-like structures have also been discovered at high redshifts (Broadhurst et al. 1990; Cohen et al. 1996; Ouchi et al. 2004).

\*E-mail: weygaert@astro.rug.nl

### 1.1 Gravitational formation

The cosmic web can be seen as the most salient 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 profusely 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 2005a; Springel et al. 2005; Dolag et al. 2006). The filaments connect to the cluster nodes and act as the transport channels along which matter flows into the clusters.

Fundamental understanding of anisotropic collapse on cosmological scales came with the seminal study by Zel'dovich (1970), who recognized the key role of the large-scale tidal force field in shaping the cosmic web (also see Icke 1973). In addition to the anisotropic nature of gravitational collapse, the multiscale character of the cosmic mass distribution is also an important characteristic signature of the gravitational formation of structure (Zel'dovich 1970; Icke 1973; Eisenstein & Loeb 1995; Bond, Kofman & Pogosyan 1996; Eisenstein, Loeb & Turner 1997; van de Weygaert & Bond 2008; Shandarin & Sunyaev 2009). Cosmic structure formation is a hierarchical process as a result of the amplitude distribution of fluctuations over the different scales. Small-scale fluctuations in scenarios with a primordial power spectrum  $P(k) \propto k^n$ , where  $n > -3$ , have a larger amplitude than those on larger scales. As a result, collapsed clumps of matter will aggregate into larger systems and eventually merge to form even larger structures.

The description of the megaparsec matter distribution as an interconnected network or as a cosmic web is not a coincidence. While the Zel'dovich approximation describes the evolution of a cellular distribution made up of pancakes, it does not really offer an explicit dynamical explanation for the observed connectivity between morphologies. Early computer simulations already 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). Bond et al. (1996) introduced the ‘cosmic web’ theory, which provides a natural explanation for the emergence of the filamentary network as well as the relation between the morphological components of the cosmic web. The theory emphasizes the close relation between the peaks in the density field and the overall web-like network: *knowledge of the tidal field at a few relevant locations in a region provides all the information needed to predict the resulting large-scale matter configuration*. In the primordial density field, this can be traced back to the simple quadrupolar pattern in the density distribution implied by a local shear configuration (see van de Weygaert & Bertschinger 1996; van de Weygaert & Bond 2008). This distribution will naturally evolve into a cluster–filament–cluster configuration, the structural basis of the cosmic web.

### 1.2 Web analysis

Despite the multitude of elaborate qualitative descriptions, it has remained a major challenge to characterize the structure, geometry and topology of the cosmic web. Quantities as basic and general as the mass and volume content of clusters, filaments, walls and voids are still not well established or defined. Since there is not yet a common framework to objectively define filaments and walls,

the comparison of results of different studies concerning properties of the filamentary network – such as their internal structure and dynamics, evolution in time and connectivity properties – is usually rendered cumbersome and/or difficult.

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 that one finds in the cosmic matter distribution have prevented the use of a simple and straightforward tool box. Over the years, a variety of heuristic measures were proposed to analyse specific aspects of the spatial patterns in the large-scale Universe. It is only in recent years that these have led to a more solid and well-defined machinery for the description and quantitative analysis of the intricate and complex spatial patterns of 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 & Zel'dovich 1989; Shandarin, Sheth & Sahni 2004; Shandarin & Sunyaev 2009), while a related graph-theoretical construct, the minimum spanning tree of the galaxy distribution, was extensively probed and analysed by Bhavsar and collaborators (Barrow et al. 1985; Graham & Clowes 1995; Colberg 2007) in an attempt to develop an objective measure of filamentarity. Finding filaments joining neighbouring clusters has been tackled, using quite different techniques, by Colberg et al. (2005a) and Pimblet (2005b).

More general filament finders have been put forward by a number of authors. Following specific physical criteria, Gonzalez & Padilla (2010) recently proposed 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, has recently been 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; Forero-Romero et al. 2008; Sousbie et al. 2008a,b; Bond, Strauss & Cen 2010a,b). The 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 the 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 proposed by Hahn et al. (2007a) (also see Hahn et al. 2007b; Forero-Romero et al. 2008). Its great virtue is that it is based on the structure of the tidal field, which links it directly to our theoretical understanding of the formation and evolution of the cosmic web.

Instead of using the tidal field configuration, one may also try to link directly to the morphology of the density field itself. Usually, this allows a more detailed view of the intricacies of the multiscale matter distribution, although it is usually more sensitive to noise and more indirectly coupled to the underlying dynamics of structure formation than the tidal field morphology. A single scale dissection of the large-scale density field into its various morphological components based on the Hessian of the density field has been followed 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).

In this study, we follow the more elaborate multiscale formalism of the multiscale morphology filter (MMF), introduced by Aragón-Calvo et al. (2007b). The MMF assigns a morphology of the local density field in terms of its multiscale second-order variations in the local density field. Instead of restricting the analysis to one particular scale, the MMF explicitly addresses the multiscale nature of the cosmic density field 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. It represents a complete and self-consistent framework that allows us to identify and isolate specific morphologies in an objective way (see Fig. 1). A somewhat similar multiscale approach is the metric space technique described by Wu, Batuski & Khalil (2009), who applied it to a morphological analysis of the DR5 of the SDSS.

A more recent development is that of the Spineweb procedure (Aragón-Calvo et al. 2008, 2010), which traces the various features of the cosmic web on pure topological grounds by invoking the *watershed transform*. The watershed transform is a key instrument

for the segmentation of a density field, and as such is also ideally suited for tracing the boundaries between the identified segments. Spineweb identifies the filaments and sheets with the boundaries of watershed basins. The latter are the influence areas in and around cosmic voids.

### 1.3 Cosmic environment and galaxy formation

One of the main reasons for our interest in outlining the filamentary cosmic web concerns the question whether and to what extent the web-like environment influences the properties and evolution of galaxies. Most studies of environmental influences limit themselves to the density, but various indications argue for a more intricate connection.

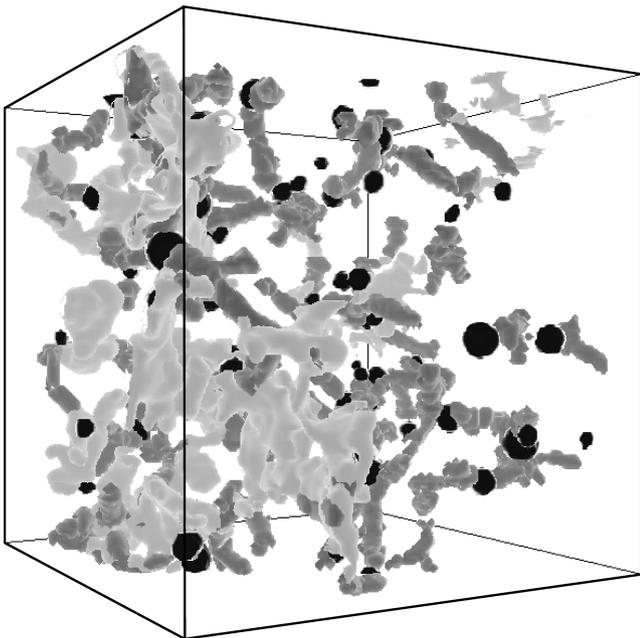
In at least one aspect, we may immediately suspect a significant relation between the tidally induced morphological nature of the cosmic environment and the galaxy. The tidally induced rotation of galaxies implies a link between the galaxy formation process and the surrounding external matter distribution. With the cosmic web as a direct manifestation of the large-scale tidal field, we would therefore expect a connection with the angular momentum of galaxies or galaxy haloes. The theoretical studies of Sugerman, Summers & Kamionkowski (2000) and Lee & Pen (2000) were important in pointing out that this connection should be visible in the orientation of galaxy spins with the surrounding large-scale structure.

Equipped with some of the filament and wall detection techniques described above, recent  $N$ -body simulations have been able to find, amongst others, 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; Paz, Stasyszyn & Padilla 2008; Hahn 2009; Hahn et al. 2009; Zhang et al. 2009). In the case of haloes located in large-scale walls, they seem to agree that both the spin vector and the major axis of inertia lie in the plane of the wall. In the case of the alignment of haloes with their embedding filaments, Aragón-Calvo et al. (2007a) and Hahn et al. (2007a) found evidence for a mass and redshift dependence, which has been confirmed by the studies of Paz et al. (2008) and Zhang et al. (2009).

While the MMF (Aragón-Calvo et al. 2007b) proved successful in elucidating halo shape and spin alignment characteristics in filaments and sheets in  $N$ -body simulations (Aragón-Calvo et al. 2007a; Zhang et al. 2009), in recent work Jones et al. (2010) also succeeded in identifying and tracing filaments in the SDSS survey which contained manifestly aligned galaxies. In this paper, we use the MMF to look in more detail at the intrinsic properties of the web-like structures themselves.

### 1.4 Intention and outline

We address the cosmic web in terms of its basic morphologies – clusters, filaments and walls – identified on the basis of the MMF. We will mainly focus on filaments given the fact that they are the most prominent components of the cosmic web and largely delineate its outline. Walls are far less prominent, more tenuous and highly complex. We will therefore pay less attention to them. In some cases, we will also include clusters and voids in our analysis but always in the context of the filament–wall network. Instead of seeking to provide a comprehensive list of properties of the morphologies in the cosmic web, this work presents a general view of the cosmic web from the point of view of their morphological components and introduces some tools for their characterization. Some of our results



**Figure 1.** MMF segmentation of the mass distribution in the cosmic web. The cosmic web is delineated by filaments (dark grey) and walls (light grey). Clusters (dark grey) are located at the intersection of filaments. Each of these elements is indicated by isodensity contours (on a Gaussian scale of  $R_f = 2 h^{-1}$  Mpc). Only the largest structures are shown for clarity.

confirm previous findings, while others have not been presented before.

This study is organized as follows. In Section 2 we describe the cosmological simulation on which this work is based, including the resampling of the discrete particle set into a regular grid of density values using the adaptive and morphology preserving Delaunay tessellation field estimator (DTFE) technique. Section 3 briefly describes the steps followed in the morphological characterization of the cosmic web by means of the MMF. Section 4 contains a qualitative presentation of the various morphological components of the cosmic web, while the corresponding quantitative inventory in terms of mass, volume and density is the subject of Section 5.

In Section 6, we proceed to describe the filamentary network and some of its global properties such as mass function, length distribution and density profiles. Finally, we summarize our findings in Section 7.

## 2 N-BODY SIMULATIONS AND HALO CATALOGUES

The work presented here is based on a cosmological  $N$ -body simulation containing only dark matter particles. The simulation follows the evolution of a set of particles ‘tracing’ the underlying density field from a given set of initial conditions until the present time. We adopted the concordance  $\Lambda$  cold dark matter ( $\Lambda$ CDM) cosmological model  $\Omega_m = 0.3$ ,  $\Omega_\Lambda = 0.7$ ,  $h = 0.7$  and  $\sigma_8 = 1.0$ . Its size ( $150 h^{-1}$  Mpc) makes it suitable to study large structures comparable to those seen in present galaxy surveys. The large number of particles ( $512^3$ ) allows us to achieve a mass resolution of  $1.34 \times 10^9 M_\odot h^{-1}$  per particle. The mass resolution and simulation box were chosen as a compromise between a box large enough to contain a significant amount of large structures and at the same time the ability to resolve haloes down to a few times  $10^{11} M_\odot$  (given the computational resources available). The simulation was performed using the public version of the parallel Tree-PM code GADGET 2 (Springel 2005), running on eight double processor nodes on the Linux cluster at the University of Groningen. Initial conditions at redshift  $z = 50$  with  $512^3$  dark matter particles were generated using the transfer function given by Bardeen et al. (1986).

### 2.1 The $N$ -body data

We stored 20 snapshots starting at redshift 9 in logarithmic intervals of the expansion factor until the present time. Additionally, we generated  $256^3$  and  $128^3$  versions following the ‘averaging’ prescription described in Klypin et al. (2001). These lower resolution simulations were used to compute the density fields and to get a preliminary impression of the structures present in the simulation box (see Table 1).

(i) The low-resolution version ( $n_{\text{part}} = 128^3$ ) is used to compute some properties of filaments such as linear density and for visualization purposes. This resolution per particle of this simulation

allows us to resolve the main features of the large-scale distribution and at the same time is sparse enough to allow a clear visualization of the particles (see Fig. 3, later). This is the simulation we use when (in the following sections) we refer to dark matter particles, unless we state something different.

(ii) The intermediate resolution ( $n_{\text{part}} = 256^3$ ) is used to compute density fields in order to take full advantage of the spatial information with the computing resources available.

(iii) The high-resolution version ( $n_{\text{part}} = 512^3$ ) is used to produce the HOP and FracHOP halo and subhalo catalogues (see Section 2.3).

### 2.2 The density field

The output of the  $N$ -body simulation consists of a discrete set of particles. This is translated into a continuum volume-filling density field sampled on a regular 3D grid. Crucial for the ability of the MMF to identify anisotropic features such as filaments and walls is the use of a morphologically unbiased and optimized continuous density field retaining all features visible in a discrete galaxy or particle distribution.

We therefore use the DTFE, introduced by Schaap & van de Weygaert (2000) [see Schaap (2007); van de Weygaert & Schaap (2009) for extensive descriptions], to reconstruct the underlying density field. It uses the Voronoi and Delaunay tessellations of the particle distribution to obtain an optimal local density estimate and subsequently interpolate these values linearly over the simulation volume. For our purpose of detecting web-like features, the DTFE method has several important characteristics as follows.

(i) The self-adaptive nature of the Delaunay tessellation, and resulting sensitivity to all levels of substructure present in the particle distribution, makes it highly suited for a multiresolution analysis of the hierarchically evolved large-scale matter distribution.

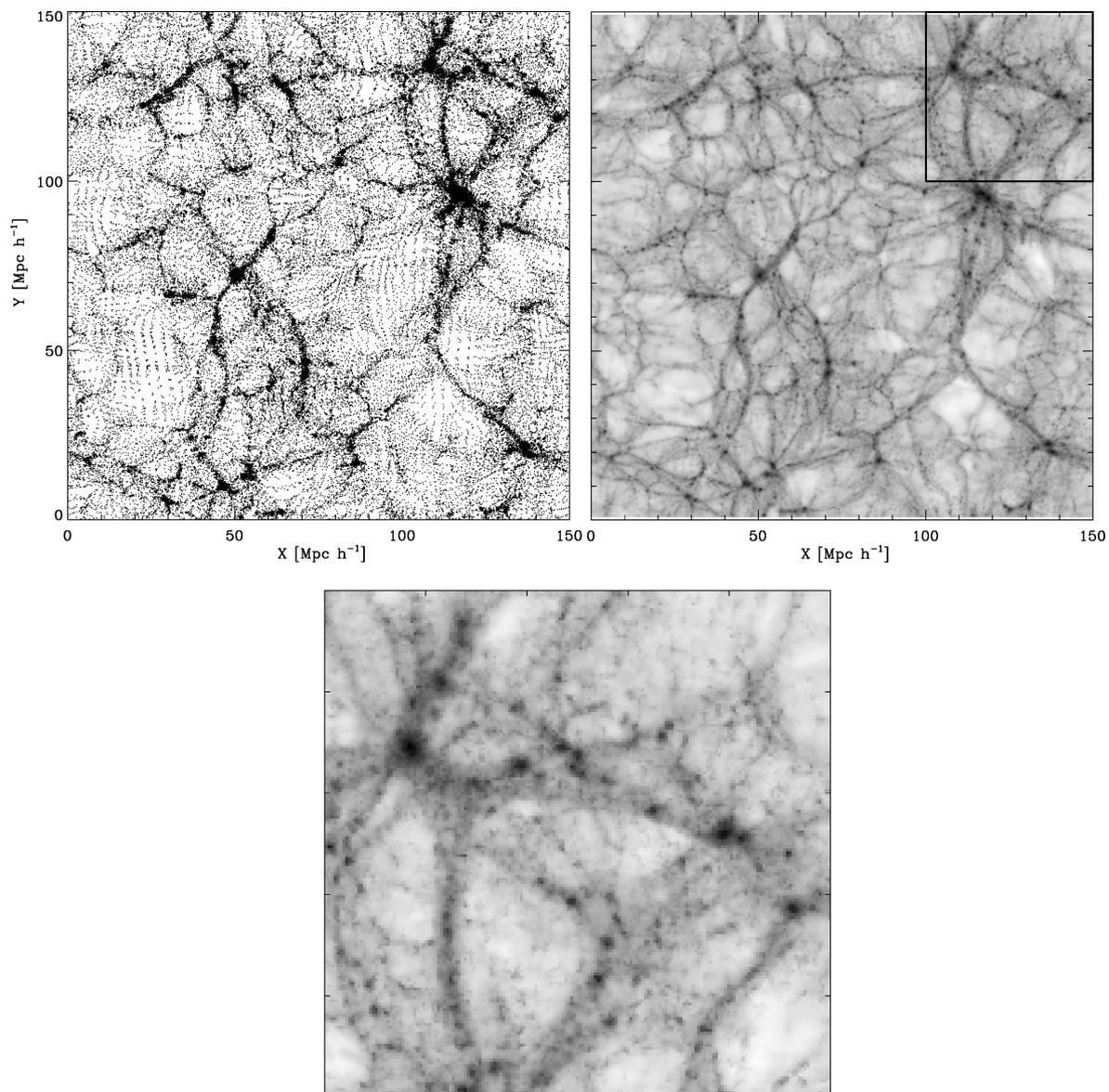
(ii) The Delaunay tessellation follows the intrinsic anisotropies of the spatial matter distribution, resulting in a density field reconstruction which accurately traces and outlines the intricate and complex spatial patterns in the cosmic web.

(iii) It does not introduce significant artificial features. The main artefacts concern diffuse low-density tetrahedral wings at the boundary between dense and underdense regions and a rather noisy reconstruction of the underdense regions.

Fig. 2 shows a slice of  $25 h^{-1}$  Mpc along the  $z$ -axis of a sub-box of  $50 \times 50 h^{-2}$  Mpc<sup>2</sup>. The top-left panel shows the particle distribution and the top right-panel shows the corresponding DTFE density field. A zoomed region is shown in the bottom panel. DTFE yields a density field reconstruction in which nearly all structures present in the particle distribution are represented. It finds highly dense clumps of matter as well as the tenuous voids. Also note the absence of artificial blobs in the inner regions of the voids. An additional virtue concerns the absence of a grid-like imprint in the DTFE density field, even though this is still visible in the particle distribution.

**Table 1.** Parameters of the  $N$ -body simulation used in this study.

Name	Box size ( $h^{-1}$ Mpc)	$\Omega_m$	$\Omega_\Lambda$	$h$	$\sigma_8$	$N_{\text{part}}$	$M_{\text{part}}$ ( $M_\odot$ )	Softening (kpc)
150 <sub>High</sub>	150.0	0.3	0.7	0.7	1	$512^3$	$2.09 \times 10^9$	18/6
150 <sub>Med</sub>	150.0	0.3	0.7	0.7	1	$256^3$	$1.67 \times 10^{10}$	36/12
150 <sub>Low</sub>	150.0	0.3	0.7	0.7	1	$128^3$	$1.34 \times 10^{11}$	72/24



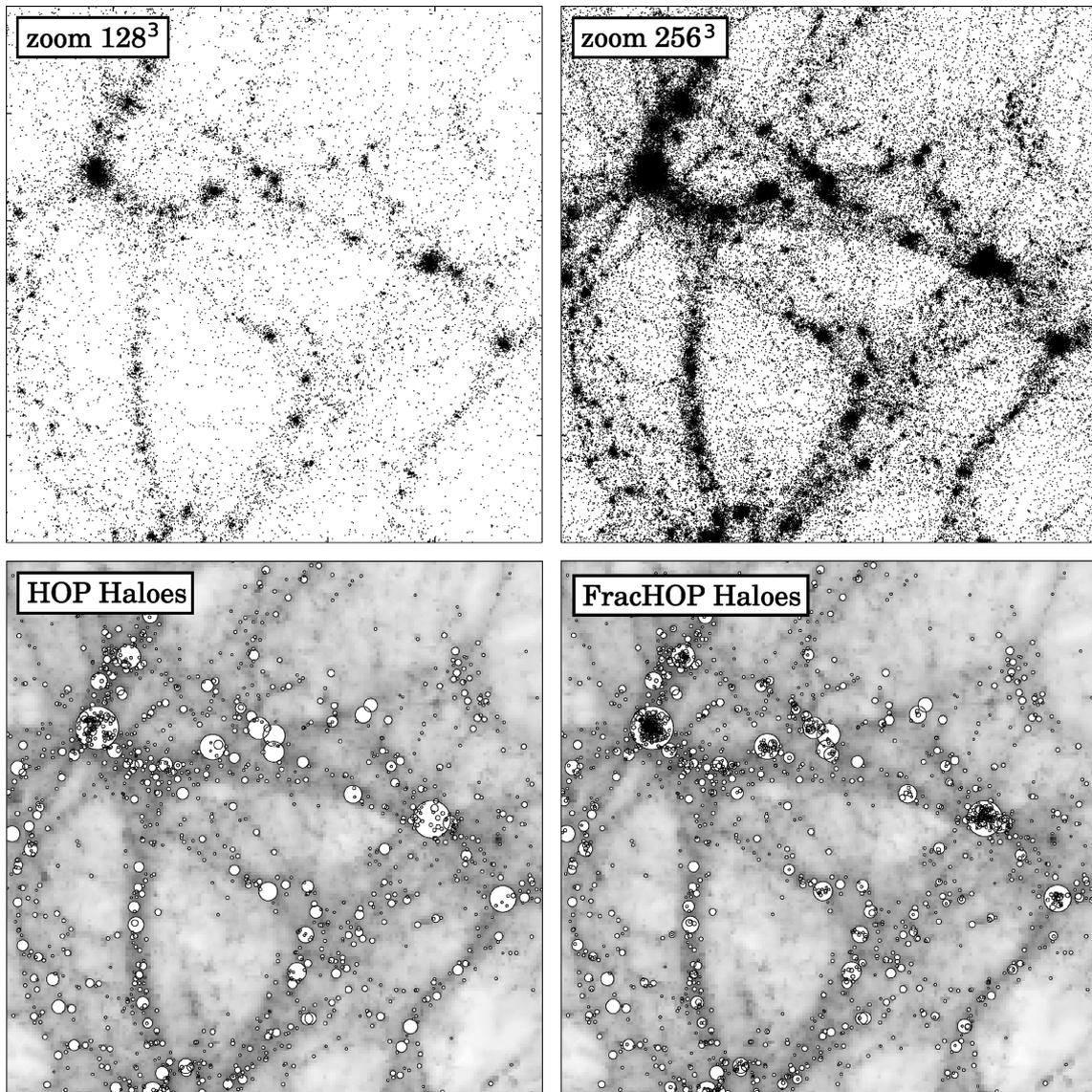
**Figure 2.**  $N$ -body simulation. The particle distribution and density fields of the  $\Lambda$ CDM simulation, which form the data sample for this study, are shown. Particles in a slice of  $25 h^{-1}$  Mpc along the  $z$ -axis (top left) and its corresponding DTFE density field (top right) are shown in a logarithmic scale. The bottom panel shows the zoomed region indicated in the top-right corner of the top-right panel. For details of the simulation, see the text.

Accordingly, DTFE is used to process the particle distribution into a continuous density field  $f_{\text{DTFE}}$  (top-right frame in Fig. 2).

### 2.3 Haloes and subhaloes

We used the HOP group finder (Eisenstein & Hut 1998) for the identification of self-bound virialized haloes. Each of these haloes is considered a parent candidate which may contain one or more subhaloes. HOP links particles by associating each particle with the densest of its  $n$ -closest neighbours, until it finally reaches the particle that is its own densest neighbour. For allocating particles to their haloes, we applied HOP with standard parameters  $\delta_{\text{out}} = 80$ ,  $\delta_{\text{saddle}} = 120$  and  $\delta_{\text{peak}} = 160$ . Although the groups identified with HOP are nearly identical to those found with FoF (Davis et al. 1985), they are less prone to involve artificial bridges between close groups.

Galactic haloes embedded in groups of galaxies and clusters of galaxies are identified with subhaloes in the computer simulation. They are the bound groups that are clearly defined against the diffuse background particles that form the halo in which they are embedded. In order to identify the bound subhaloes inside larger groups otherwise identified as single virialized objects, we use the FracHOP algorithm developed by Aragón-Calvo (2007). It is an elaboration of the HOP halo finder (Eisenstein & Hut 1998) and exploits the topological properties of nested local maxima smoothed on a fixed scale. It starts by rerunning HOP, exclusively for the particles identified with parent haloes in the first HOP halo identification step. To this end, it uses a Gaussian-smoothed density field with a kernel size of  $R_f = 35 h^{-1}$  kpc, so that substructure on smaller scale is suppressed. The subhalo identification is performed without running REGROUP, so that all particles are assigned to their local maximum in the smoothed density field. The centre of mass of each of the resulting candidate subhaloes is determined iteratively, while



**Figure 3.** Simulation particle and halo distribution in a slice of  $10 h^{-1}$  Mpc along the  $z$ -axis. The top panels show particles from the  $128^3$  and  $256^3$  simulations (left and right respectively) in the zoomed region indicated in the top-right panel of Fig. 2. The bottom panels show the haloes identified with HOP (left) and FracHOP (right). The circles are located at the centre of mass of the haloes. The radius of the circles are scaled with the mass of the halo as  $r \propto M^{1/3}$ .

unbound particles are removed. The end product is a listing of the subhaloes within the simulation box.

Fig. 3 shows the distribution of particles in the  $128^3$  and  $256^3$  simulations as well as the distribution of HOP and FracHop haloes plotted on top of the density field. Both HOP and FracHop haloes closely follow the patterns of the cosmic web, revealing that haloes are fair tracers of the large-scale matter distribution. The only difference between the two is that a given HOP halo can be formed by several FracHOP haloes. The distribution of haloes delineates the cosmic web in a more sparse and smooth way compared to the particles.

### 3 MORPHOLOGICAL SEGMENTATION: THE MMF FORMALISM

The MMF is used for identifying and characterizing the different morphological elements of the large-scale matter distribution in

the cosmic web (Aragón-Calvo et al. 2007b). The formalism has been developed on the basis of visualization and feature extraction techniques in computer vision and medical research (Florack et al. 1992). The technology, finding its origin in computer vision research, has been optimized within the context of feature detections in medical imaging. Frangi et al. (1998) and Sato et al. (1998) presented its operation for the specific situation of detecting the web of blood vessels in a medical image.

The MMF morphological segmentation takes account of the multiscale nature of the matter distribution by means of a scale-space analysis, looking for morphological structures of a mathematically specified type in a multiscale, scale-independent, manner. The scale-space analysis presumes that the specific structural characteristic is quantified by some appropriate parameter. Examples are density, eccentricity, orientation and curvature. The MMF filters these data to produce a hierarchy of maps having different resolutions and subsequently selects at each point the dominant parameter

value from the hierarchy in order to construct a scale-independent map.

The MMF-based dissection and visualization of the cosmic web in its three basic components allows us to concentrate on the significant features of the cosmic matter distribution and reach a level of abstraction by avoiding spurious details. For the visualization of the intricate filament–cluster network, this is particularly useful. In this section, we briefly summarize the steps involved in the morphological segmentation of the cosmic web obtained from the  $N$ -body cosmological simulation. A detailed step-by-step description of the MMF algorithm can be found in Aragón-Calvo et al. (2007b).

### 3.1 Scale space

The DTFE density field  $f_{\text{DTFE}}$  is the starting point of the morphological segmentation. The density field is smoothed over a range of scales by means of a hierarchy of spherically symmetric Gaussian filters  $W_G$  having different widths  $R_n$ . The  $n$ th-level smoothed version of the DTFE-reconstructed field  $f_{\text{DTFE}}$  is assigned  $f_n$ :

$$f_n(\mathbf{x}) = \int d\mathbf{y} f_{\text{DTFE}}(\mathbf{y}) W_G(\mathbf{y}, \mathbf{x}),$$

where  $W_G$  denotes a Gaussian filter of width  $R_n$ :

$$W_G(\mathbf{y}, \mathbf{x}) = \frac{1}{(2\pi R^2)^{3/2}} \exp\left(-\frac{|\mathbf{y} - \mathbf{x}|^2}{2R_n^2}\right). \quad (1)$$

Scale space itself is constructed by stacking these variously smoothed data sets, yielding the family  $\Phi$  of smoothed density maps  $f_n$ :

$$\Phi = \bigcup_{\text{levels } n} f_n. \quad (2)$$

A data point can be viewed at any of the scales where scaled data have been generated. The crux of the concept of scale space is that the neighbourhood of a given point will look different at each scale. There are potentially many ways of making a comparison of the scale dependence of local environment. We address the local ‘shape’ of the density field.

### 3.2 Local shape

The local shape of the density field at any of the scales  $R_n$  in the scale-space representation of the density field can be quantified on the basis of the Hessian matrix,  $\tilde{\mathcal{H}}_{ij} = \nabla_{ij} f_n(\mathbf{x})$ :

$$\begin{aligned} \frac{\partial^2}{\partial x_i \partial x_j} f_n(\mathbf{x}) &= f_{\text{DTFE}} \otimes \frac{\partial^2}{\partial x_i \partial x_j} W_G(R_n) \\ &= \int d\mathbf{y} f(\mathbf{y}) \frac{(x_i - y_i)(x_j - y_j) - \delta_{ij} R_S^2}{R_S^4} W_G(\mathbf{y}, \mathbf{x}), \end{aligned} \quad (3)$$

where  $x_1, x_2, x_3 = x, y, z$  and  $\delta_{ij}$  is the Kronecker delta. In other words, at each level  $n$  of the scale-space representation the Hessian matrix is evaluated by means of a convolution with the second derivatives of the Gaussian filter, also known as the Marr (or, less appropriately, ‘Mexican Hat’) wavelet. In order to properly compare the values of the Hessian arising from the differently scaled variants of the data that make up the scale space, the Hessian is renormalized,  $\tilde{\mathcal{H}} = R_S^2 \mathcal{H}$ , where  $R_S$  is the filter width that has been used.

The eigenvalues  $\lambda_i$  of the Hessian matrix determine the local morphological signal, dictated by the local shape of the density distribution. A small eigenvalue indicates a low rate of change of the field values in the corresponding eigendirection and vice versa.

**Table 2.** Morphology and eigenvalue configuration. The eigenvalue conditions specify clusters, filaments and walls, each having a density higher than the background. A negative eigenvalue indicates that the feature reaches a maximum along the corresponding direction (and vice versa), while a small eigenvalue indicates a low rate of change of the field values in the corresponding eigendirection (and vice versa). This leads to the morphological relationships listed in this table.

Structure	$\lambda$ ratios	$\lambda$ constraints
Cluster node (blob)	$\lambda_1 \simeq \lambda_2 \simeq \lambda_3$	$\lambda_3 < 0; \lambda_2 < 0; \lambda_1 < 0$
Filament	$\lambda_1 \simeq \lambda_2 \gg \lambda_3$	$\lambda_3 < 0; \lambda_2 < 0$
Sheet	$\lambda_1 \gg \lambda_2 \simeq \lambda_3$	$\lambda_3 < 0$

We denote these eigenvalues by  $\lambda_a(\mathbf{x})$  and arrange them so that  $\lambda_1 \geq \lambda_2 \geq \lambda_3$ :

$$\left| \frac{\partial^2 f_n(\mathbf{x})}{\partial x_i \partial x_j} - \lambda_a(\mathbf{x}) \delta_{ij} \right| = 0, \quad a = 1, 2, 3 \quad (4)$$

with  $\lambda_1 > \lambda_2 > \lambda_3$ .

The  $\lambda_i(\mathbf{x})$  are coordinate-independent descriptors of the behaviour of the density field in the locality of the point  $\mathbf{x}$  and can be combined to create a variety of morphological indicators. The criteria we used for identifying a local blob-like cluster, filamentary or sheet-like morphology are listed in Table 2. Evidently, the eigenvalues corresponding to a blob (cluster) morphology are a subset of the eigenvalue subset related to filamentary morphologies. In turn, the eigenvalue set of the latter is a subset of the wall eigenvalues.

### 3.3 Multiscale structure identification

In practice, we are interested in the local morphology as a function of scale. In order to establish how it changes with scale, we evaluate the eigenvalues and eigenvectors of the renormalized Hessian  $\tilde{\mathcal{H}}$  of each data set in the scale space  $\Phi$ .

Since we are looking for three distinct structural morphologies – cluster blobs, walls and filaments – the practical implementation of the segmentation consists of a sequence of three stages. Because curvature components are used as structural indicators, the blobs need to be eliminated before looking for filaments, after which the filaments have to be eliminated before looking for walls. This results in the MMF procedure following the sequence ‘clusters  $\rightarrow$  filaments  $\rightarrow$  walls’. At each of these three steps, the regions and scales are identified at which the local matter distribution follows the corresponding eigenvalue signature.

In practice, the MMF defines a set of morphology masks, morphology response filters and morphology filters for each of the three different morphological components: clusters, filaments and walls. Their form is dictated by the particular morphological feature they seek to extract, via the eigenvalues at each level in scale space and the criteria for each of the corresponding morphologies (Table 2). The local value of the morphology response depends on the local shape and spatial coherence of the density field. The morphology signal  $\Psi(\mathbf{x})$  at each location is then defined to be the one with the maximum response across the full range of smoothing scales. Formally, we denote  $\Psi$  by the name of *scale-space map stack*.

### 3.4 Morphology thresholds

The final step in the MMF feature identification concerns the removal of noisy structures. To this end, the MMF invokes global morphology thresholds in order to separate the *texture noise* from

genuine structures (see Aragón-Calvo et al. 2007b). Regions with a morphology response  $\Psi(\mathbf{x})$  lower than the global threshold  $\tau$  are omitted from the list of identified structures.

The value of the thresholds  $\tau_B$ ,  $\tau_f$  and  $\tau_w$  for clusters, filaments and walls is determined on the basis of the measured dependence of the structure detection rate as a function of the value of the morphology signal  $\Psi$ . All clusters with a morphology value less than the threshold  $\tau_B$  are considered to be small insignificant blobs. The threshold is chosen such that these are erased, but not the large gravitationally bound clumps. In the case of filaments and walls, the threshold value is determined on the basis of the percolation properties of the network of filaments and walls. The threshold values  $\tau_f$  and  $\tau_w$  are defined as the morphology signal value  $\Psi$  for which the population of filaments and walls reaches its maximum number: at lower values, the filaments and walls start to percolate.

In a distribution where all filaments or walls have similar properties such as contrast and physical extent, this is the perfect choice. In the real Universe, however, there is a large variation in the contrast and size of filaments and walls. The same global criteria are therefore applied to faint as well as prominent structures. As a result, there is a systematic inclusion of low-density regions forming the boundary of faint structures. While most mass is concentrated in high-density regions, most of the volume of space concerns low-density regions. As a result, the MMF has some bias towards low-density structures. This might be alleviated by the use of more restrictive threshold values. However, this would imply the loss of very faint structures. A more preferable but as yet not practical approach would be the use of a local threshold value which would

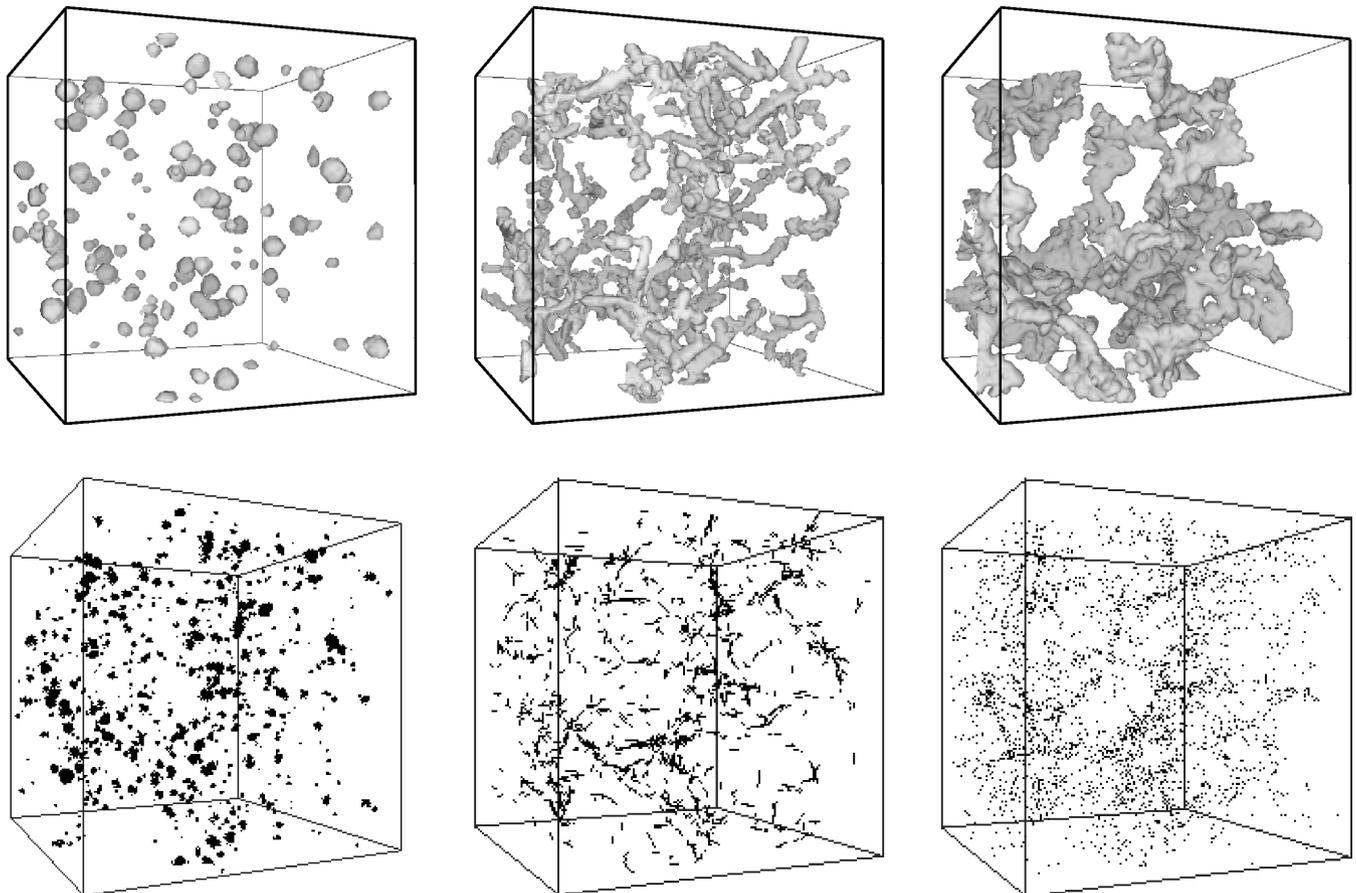
account for the significance of features within the environmental context.

### 3.5 MMF product

The end product of the MMF procedure is a map segmented in clusters, filaments and walls (Fig. 4). These have been identified as the most outstanding structures and vary in scale over the full range of scales represented in scale space (equation 2). Following the thresholding of the scale-space map stack  $\Psi$  on the basis of cosmological and astrophysical considerations, we are left with the *object map*  $\mathcal{O}$ . For each of the different morphologies – clusters, filaments and walls – these consist of the physically recognizable objects in the cosmic web.

## 4 MORPHOLOGICAL SEGMENTATION: COSMIC WEB COMPONENTS

Fig. 1 shows the morphological segmentation of the  $150 h^{-1}$  Mpc simulation ( $150_{\text{Low}}$  in Table 1) obtained with the MMF. This figure illustrates the large-scale distribution of matter as an interconnected network of filaments (dark grey) defining the boundaries of walls (light grey), and the clusters (black) located at the intersections of the network. The spatial distribution for each of the individual morphological components is shown in Fig. 4, by means of isodensity surfaces (top column) and by means of the particles enclosed by these surfaces (bottom column).



**Figure 4.** Top: surfaces enclosing regions of space identified as clusters, filaments and walls in the simulation box (left-hand, middle and right-hand panels, respectively). Bottom: the particles enclosed by the surfaces in the top panels. Only the largest objects are shown for clarity.

For clarity, Figs 1 and 4 only show the largest structures. Including all the objects would quickly have produced an image saturated with walls and filaments. By restricting the number of structures included, it is easier to identify the individual components of the cosmic web. Each morphological component is well differentiated and occupies, by construction, mutually exclusive regions. The variety of sizes of the clusters is a nice illustration of the ability of the MMF to identify structures at different scales (also see Section 4.1).

Careful inspection of Figs 1 and 4 reveals the close physical affiliation of the different morphological components. Cluster blobs are located at the tips of filaments, and filaments tend to be found at the boundaries of walls. This confirms theoretical expectations (Zel'dovich 1970; Bond et al. 1996).

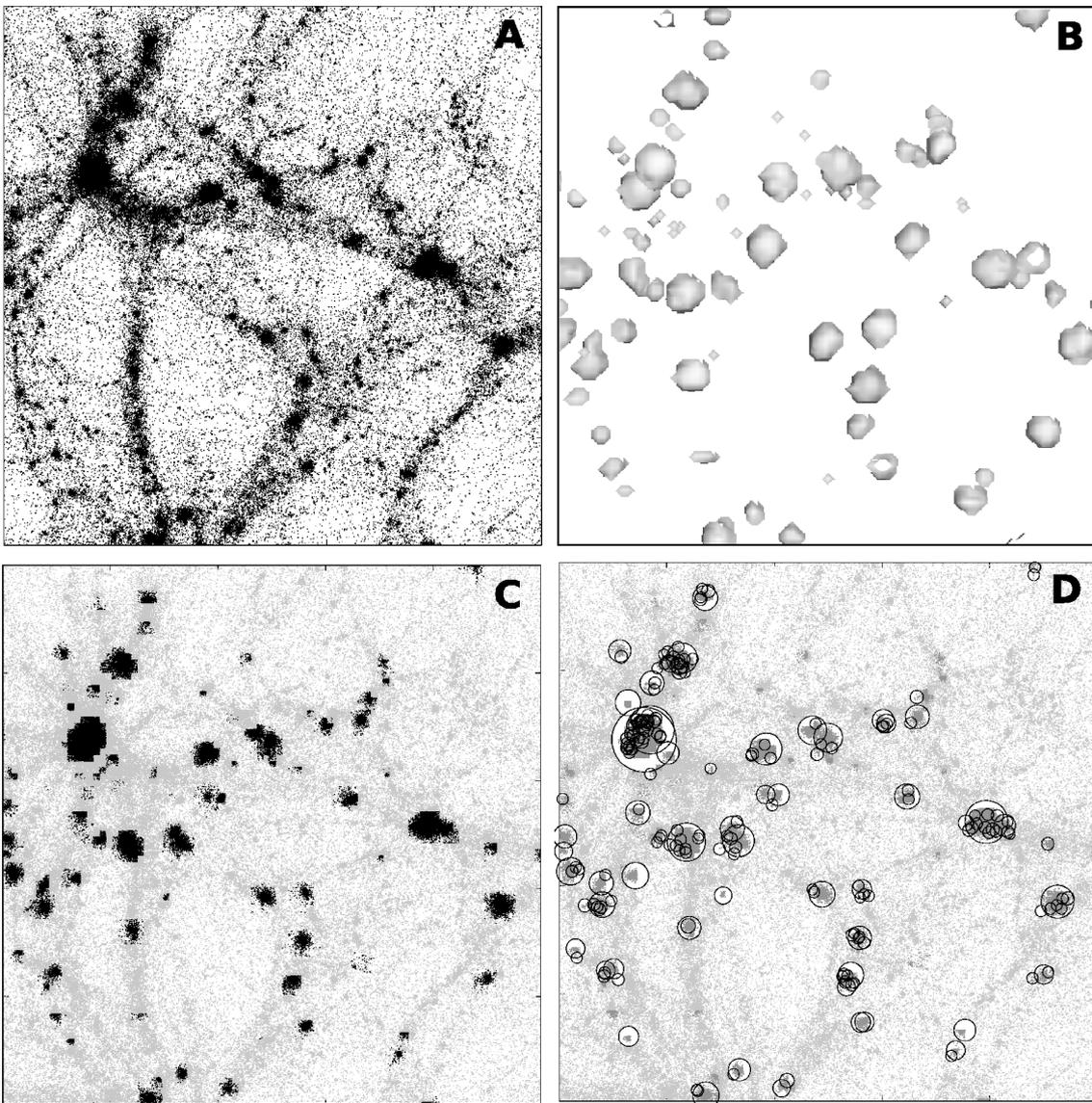
Following the MMF segmentation of the matter distribution according to their intrinsic morphology and scale, it is straightforward

to compute the global properties of each morphological component, such as mass and volume content. For other properties, such as the length and density profiles of filaments, additional post-processing steps are necessary. These analysis procedures will be described in the next sections. First, we present a qualitative and illustrative impression of each of the main morphologies, starting with clusters and followed by filaments and walls.

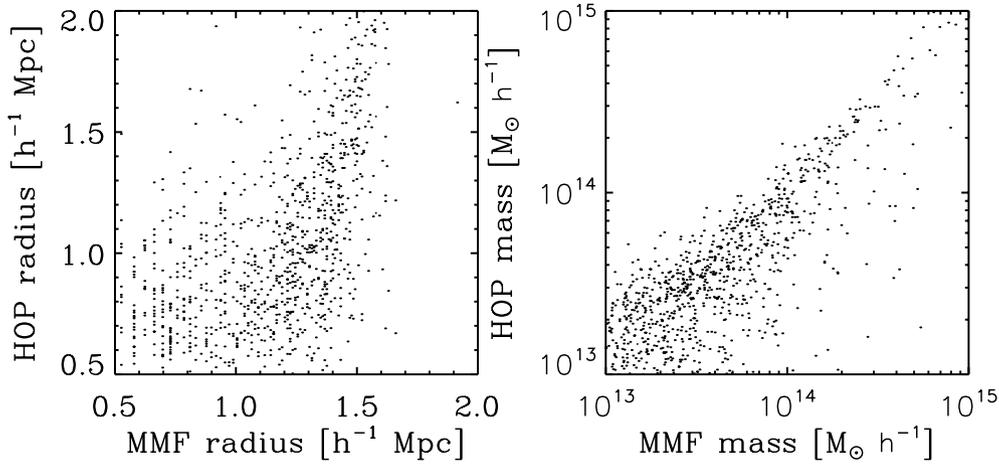
#### 4.1 Clusters

A more detailed view of the cluster distribution is presented in Fig. 5. We illustrate the cluster distribution by means of four panels as follows.

- (i) The particle distribution.
- (ii) Surfaces defining clusters identified with the MMF.



**Figure 5.** MMF cluster and halo identifications. (a) All particles inside a sub-box of the simulation. (b) Surfaces enclosing regions identified by the MMF as clusters. Note that projection effects may distort the real size of the surfaces enclosing clusters. (c) Particles located inside MMF clusters (dark grey). The rest of the particles are shown in light grey colour. (d) Haloes identified with HOP that have their centre of mass inside regions identified as clusters by the MMF. The MMF manages to identify the dense clusters at their characteristic scale. This may be seen from panels (c) and (d). The match is reasonably good (although not perfect due to the intrinsic differences between HOP and the MMF methods).



**Figure 6.** Halo identification comparison. Scatter plot of the radius (left) and mass (right) of haloes identified with HOP and the MMF. The radius of the HOP haloes corresponds to the virial radius

- (iii) Particles identified inside the MMF clusters.
- (iv) HOP haloes corresponding to the MMF clusters.

The size of the objects seen in the particle distribution as well as the HOP haloes is related to the size of the clusters identified with the MMF. The match is not perfect, as one may expect due to the intrinsic differences between HOP and the MMF. This is illustrated in Fig. 6, where we compare the radius and mass of clusters identified with HOP and the MMF.

The radius of HOP clusters is defined as the distance from the centre of mass to the outermost particle. For the MMF clusters, we computed the radius from

$$R = \left( \frac{3}{4\pi} V_{\text{blob}} \right)^{1/3}, \quad (5)$$

where  $V_{\text{blob}}$  is the volume of all pixels defining an individual blob (cluster). The masses of HOP and MMF clusters are well correlated. This is not surprising, since most of the mass of the cluster is located in the dense inner regions of the cluster. On the other hand, the scatter between the radius of HOP and MMF clusters is large. This is a result of the way in which the radius is estimated. The distance of the most distant particle from the centre of mass is sensitive to small fluctuations in the periphery of the clusters. Resolution effects also influence the estimated radius of MMF clusters, since the density field grid size is large compared to the radius of the smallest clusters.

The third major effect which influences the radius estimate of massive clusters is the often substantial intrinsic elongation of clusters. By virtue of the MMF formalism, clusters identified with the MMF tend to be spherically symmetric. However, in general the shape of virialized clusters tends to depart from sphericity, which can be most clearly observed in computer simulations (see e.g. Araya et al. 2009). Massive clusters are often highly elongated in the direction of the filaments connected to them and along which most merging clumps are moving (van Haarlem & van de Weygaert 1993). In fact, the infall of matter along the filamentary transport channels amplifies the elongation and alignments of the clusters with respect to the filaments and neighbouring clusters (van Haarlem & van de Weygaert 1993). The strongest contribution to this effect is that by the merging of two or more clusters, which shares the highly anisotropic nature of the more gradual accretion of most of the matter.

Once all clusters are identified, the corresponding cluster particles are removed from the particle sample. The cluster-free particle distribution is subsequently analysed for its filament population.

## 4.2 Filaments

On the basis of the ‘cluster-free’ particle distribution, we compute the DTFE density field. This density field only contains structures associated with filaments and walls and is used as an input for the identification of filaments in the MMF pipeline.

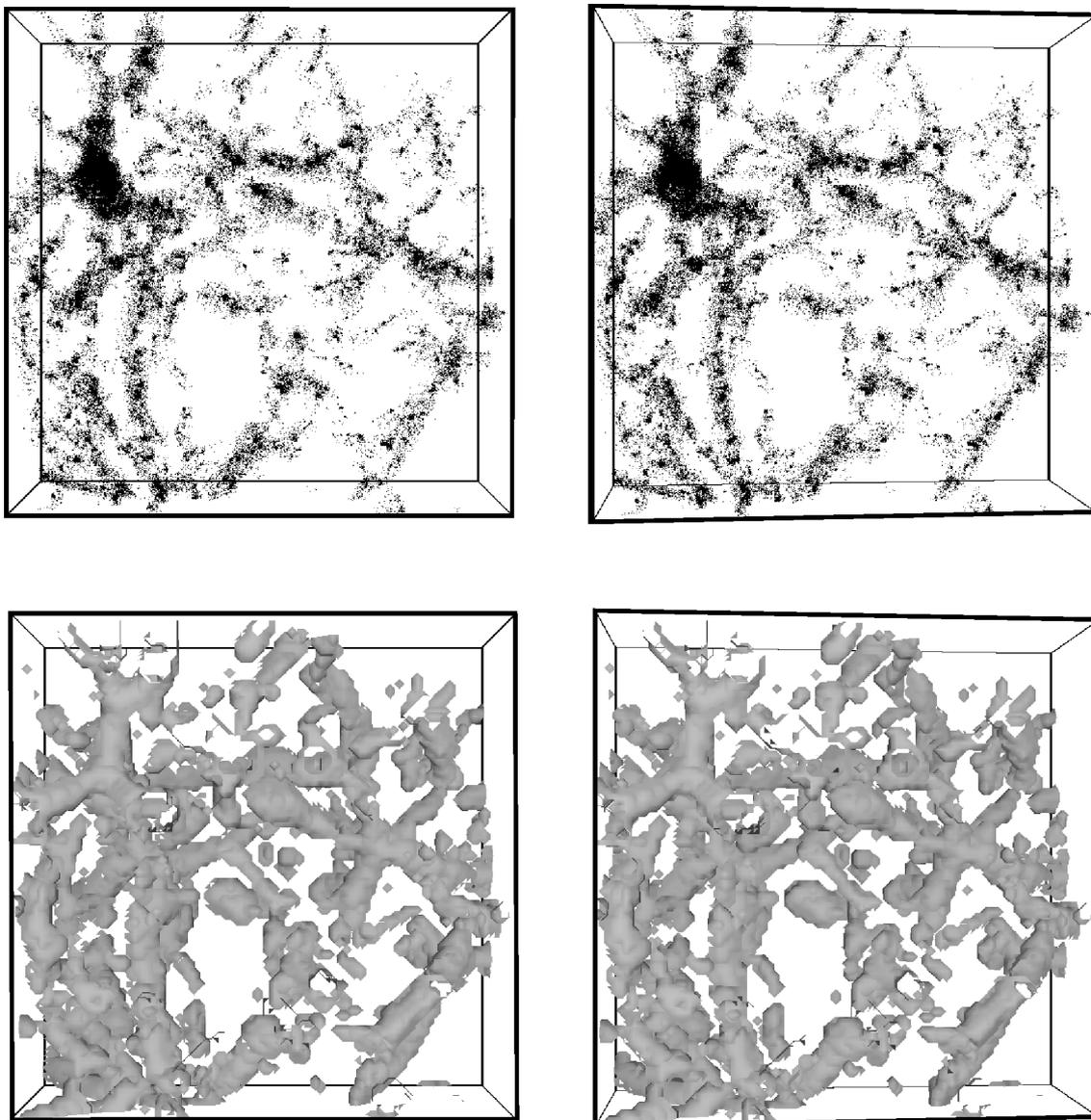
The 3D visualization allows us to appreciate the 3D nature of the filaments, in particular also highlighting their connectivity. Fig. 7 shows a 3D view of filaments inside a sub-box with a side of  $50 h^{-1}$  Mpc. The isosurfaces enclosing regions of space identified as filaments are shown in the bottom panels. The particles they contain are shown in the top panels. Note that the regions where clusters are located (e.g. the large cluster at the top-left corner) are also covered by the filament mask but that we explicitly exclude these regions from the filament mask.

Note that there are some small particles and clumps, which appear to be mostly isolated and bear no relation to the surrounding structures. These correspond to regions that have a spurious filamentary nature and are near the resolution and filament detection limit of the MMF. We discard this spurious identification from the filament mask.

Filaments identified with the MMF are complex objects that pervade throughout the sample volume, connecting each of its regions. The MMF follows the intrinsic scale of the real filaments, which feature a large range of lengths and widths. The corresponding filament particle distribution is far from homogeneous: along the filaments we find dense haloes as well as a pervasive medium of diffuse particles. It is clear that a description of such complex systems requires advanced methods of analysis which are sensitive to the anisotropic and multiscale nature of the matter distribution. The MMF is an example of such a specifically designed instrument.

## 4.3 Walls

After removing the particles located inside clusters and filaments we proceed to the last step in the morphological segmentation, the identification of walls in the density field. Walls are the most tenuous coherent structures in the large-scale Universe. Their identification



**Figure 7.** MMF filaments. 3D stereoscopic view (cross-eyed) of the isosurfaces enclosing filaments (bottom panels) and the enclosed particle distribution (top panels). The box corresponds to the zoomed region shown in Fig. 3.

poses a major challenge (Shandarin et al. 2004), and their tenuous nature and complex topology and shapes make them the most difficult morphology to characterize.

Fig. 8 shows a 3D view of the walls inside a volume of  $50 h^{-1}$  Mpc size. The top panel shows the particles located within the walls, with the bottom panels depicting the corresponding walls by means of isosurfaces. The DTFE density field on the basis of the cluster- and filament-free particle sample contains only features of a planar nature, although it also contains various noisy features that cannot be clearly identified with any of the three basic morphologies.

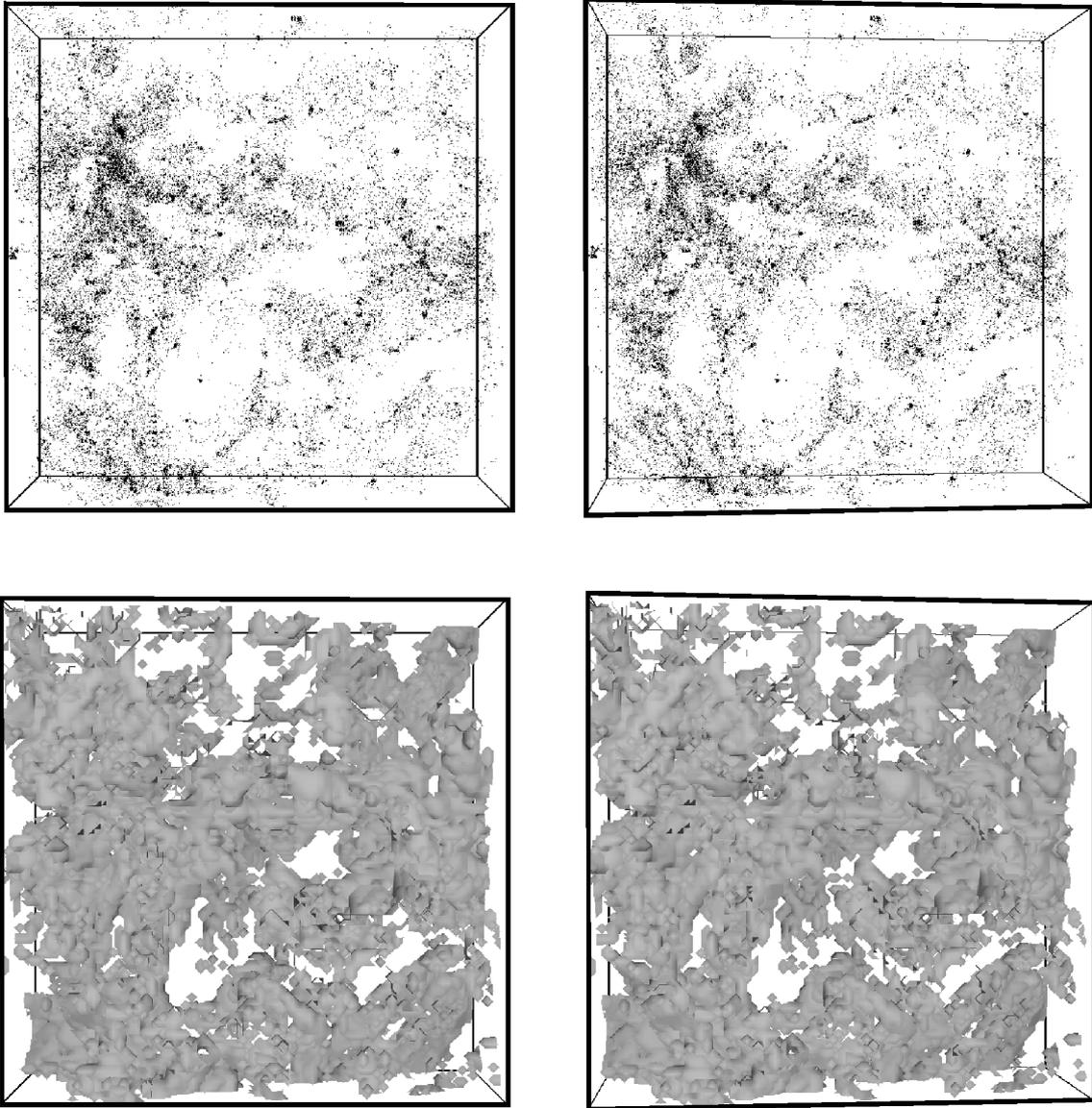
In general, walls identified with the MMF are far from smooth planar objects. They tend to have crumpled shapes with no obvious topology. Because they are multiply connected objects, it is virtually impossible to isolate individual walls from the complex web of walls. Evidently, the MMF identifies the walls successfully. However, as a result of their complex nature, we restrict ourselves to assessing their basic properties.

## 5 COSMIC WEB INVENTORY

### 5.1 Mass and volume content

In order to understand the role of clusters, filaments and walls in the shaping of the cosmic web, it is crucial to determine their relative abundances in terms of volume and mass. Such quantities may provide rough estimates of the dominance of one morphology with respect to the other in shaping the cosmic web and driving its overall dynamics.

Table 3 lists some basic characteristics for each of the morphological elements. These include an inventory of the cosmic web in terms of volume and mass content. The mass content of a particular morphology is measured by adding the total mass of the particles enclosed within the boundaries of that morphology. The fraction of the occupied volume is determined by adding the volume of all voxels enclosed within the morphological boundaries.



**Figure 8.** MMF walls. 3D stereoscopic view (cross-eyed) of the isosurfaces enclosing walls (bottom panels) and the enclosed particle distribution (top panels). The box corresponds to the zoomed region shown in Fig. 3.

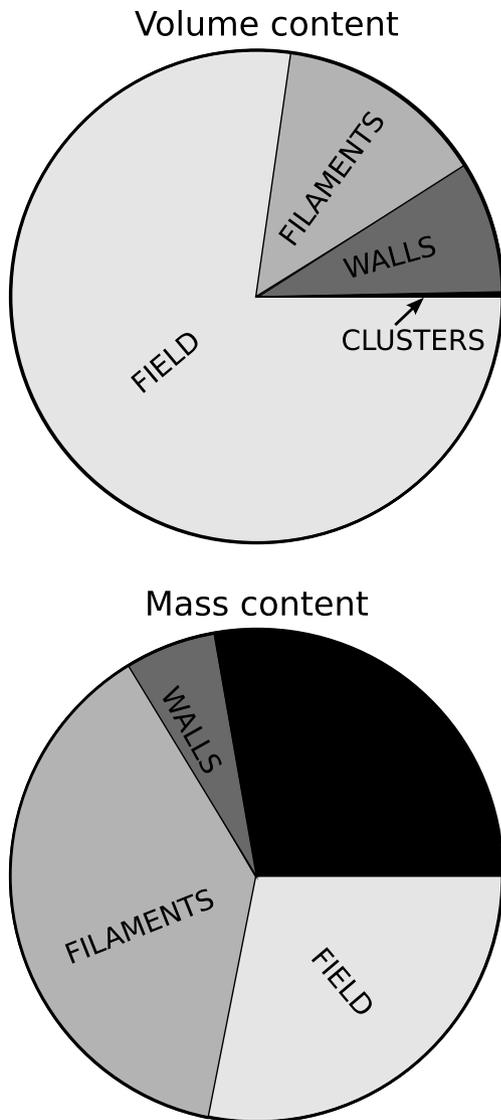
**Table 3.** Inventory of the cosmic web. The volume, mass content and a few statistical characteristics of the density distribution of the individual structural morphologies are listed. The mean, median, standard deviation and kurtosis are computed from the distribution of the overdensity  $1 + \delta = \rho/\bar{\rho}$ .

	Clusters	Filaments	Walls	Voids
Volume filling (per cent)	0.4	8.8	4.9	85.9
Mass content (per cent)	28.1	39.2	5.5	27.2
Mean overdensity	73.0	4.5	1.1	0.3
Median overdensity	11.5	1.7	0.9	0.3
Standard deviation	58.8	11.4	2.61	0.52
Kurtosis	58.7	44.8	160.5	142.1

The resulting cosmic inventory is summarized in the pie diagrams of Fig. 9. The stark contrast between the volume and the mass share of the clusters, filaments and sheets is a direct manifestation of the large density differences between the different morphologies and a

direct indication of the dynamical importance of these elements. The density contrast differences are also an indication for the different evolutionary stages in which they reside, as gravitational collapse proceeds faster as we go from walls  $\rightarrow$  filaments  $\rightarrow$  clusters.

Not surprisingly, clusters occupy the smallest volume fraction in the cosmic web, occupying only 0.38 per cent. Despite this, they also represent a major share of the cosmic mass: 28 per cent of the total mass resides within cluster regions. This not only makes them by far the densest objects of the megaparsec Universe, but also makes them the dynamically dominant component of the cosmic web (see e.g. Bond et al. 1996). The largest fraction of the mass in the Universe,  $\approx 39$  per cent, resides in filaments, which occupy around 10 per cent of the total volume. Although their density is lower than that of clusters, they represent the most salient component of the cosmic web via their function as the all-pervasive bridges between all structural features in the megaparsec Universe. Walls contain a substantially smaller fraction of the mass,  $\approx 5.5$  per cent. They also occupy a relatively small volume, at  $\approx 4.9$  per cent, even less than



**Figure 9.** Pie diagram showing an inventory of the cosmic web in terms of volume (top) and mass (bottom). We distinguish clusters, filaments, walls and void regions (or ‘field’).

that occupied by filaments. It certainly means that walls have been relatively unimportant in the recent formation history of the cosmic web.

It is instructive to compare the present-day morphological inventory with that in the primordial density field. On the basis of the (tidal) deformation tensor distribution in the primordial Gaussian field Doroshkevich (1970) showed that in the *linear regime* 92 per cent of the mass will collapse into walls, filaments or clusters. Filaments and walls would each take 42 per cent of the share, clusters 8 per cent, while the remaining 8 per cent would correspond to underdense voids. While we may already expect that this direct link between the primordial deformation tensor and morphology is too simplistic, it is the subsequent quasi-linear and non-linear evolution which changes these numbers substantially.

The MMF is a density-field-based criterion and performs better as the density field becomes more prominent and non-linear: it selects only those regions with a clear morphology and contrast. Following the same deformation tensor criteria with respect to the primordial density field, Pogosyan et al. (1998) showed that fila-

ments will be much more prominent in the high-density regions, which tend to develop faster in the subsequent non-linear evolution. Walls are more biased to lower density regions and at all times will therefore occur less prominent than filaments. The MMF sensitivities will therefore be naturally biased towards the filaments and clusters in the mass distribution. Moreover, the MMF is less likely to properly identify the outer infall regions of clusters, filaments and walls and instead tends to relegate part of these to the field as they have not yet emerged as fully developed structures. Here, for simplicity, we identify these field regions with the larger low-density voids.

## 5.2 Density segregation of the cosmic web

The differences in mass and volume content derived for each morphology correspond to different density ranges. It is often assumed that these elements mark out a unique density regime, with no overlapping values. This assumption is the basis for the use of overdensity as one of the most widely used criteria to identify clusters (Lacey & Cole 1994; Eke et al. 1996) and filaments (Shandarin et al. 2004; Dolag et al. 2006).

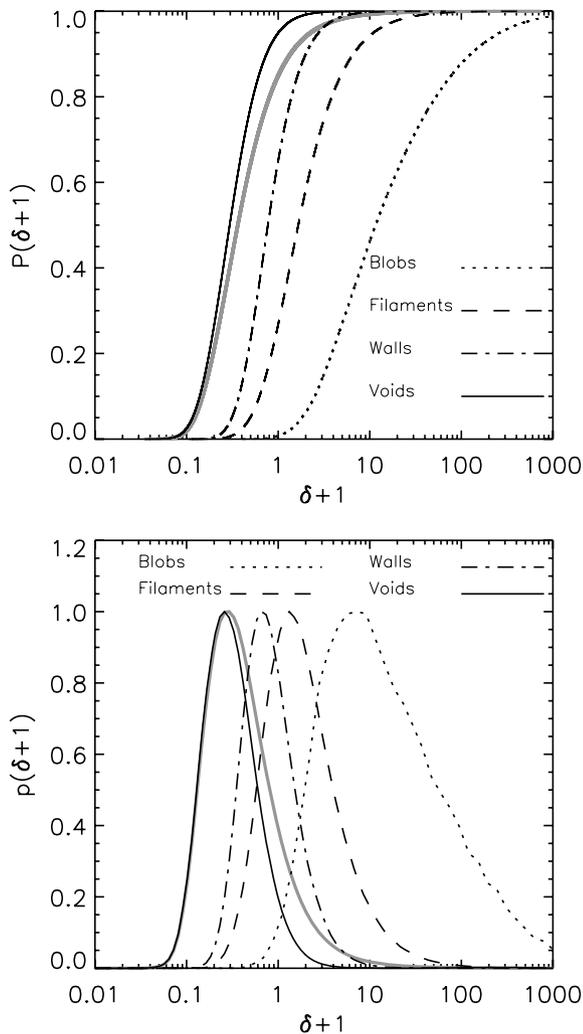
Fig. 10 shows the cumulative and probability distributions of the overdensity  $\delta$  within the regions identified as clusters, filaments, walls and field. The bottom panel shows that each morphology occupies a characteristic range in density.

- (i) Clusters are the densest objects, with a median overdensity of  $\sim 11.5$  and a mean overdensity of  $\sim 73$ . The range of overdensities extends to more than  $\delta \sim 100$  and even  $\delta \sim 1000$  within the large virialized clusters.
- (ii) Filaments and walls have medium overdensities, with a mean overdensity of  $\langle 1 + \delta \rangle \sim 4.5$  for filaments and  $\langle 1 + \delta \rangle \sim 1.1$  for walls and a median overdensity of  $\sim 1.7$  and  $\sim 0.9$ , respectively.
- (iii) The field should be mostly identified with the most underdense void regions. On average, the void regions correspond to underdensities of  $\delta \sim -0.7$ .

The density values for filaments and walls partially coincide with the density range expected for collapsed objects, since they concern values  $\delta \geq 6$  at which spherical objects turn around into collapse. However, in particular for walls a major fraction of the enclosed space has a substantially lower density. To a large extent this concerns the lower density in the outer realms which surround the dense inner regions of clusters, filaments and walls. There is also a bias towards low densities in structures identified with the MMF as a result of the morphology threshold criteria used by the MMF to separate real and noisy structures (see Section 3.4).

The considerable level of overlap in density between the various morphologies also means that a pure density criterion for structural identification does not provide an accurate description of reality. A morphological segmentation in terms of density alone would require at least a non-overlapping low-density tail. However, the fact that this does not seem to be the case implies that there is a substantial contamination with other morphological elements if one resorts to a pure density-based criterion: a (global) density threshold would be a poor discriminator of morphology. Even when each morphology can be associated with a specific density range, in general an additional, more sophisticated characterization is required.

The use of the MMF method to disentangle the cosmic web into its basic morphologies, independent of their density contrast, is clearly justified by the results presented in Fig. 10. However, we do have to take care of the fact that the MMF density estimates of



**Figure 10.** Cumulative (top) and density (bottom) probability distribution of the density contrast in clusters (blobs), filaments, walls and the void regions (dotted, dashed, dot–dashed and solid, respectively). The thick grey lines indicate the distribution corresponding to all the volume.

filaments and walls tend to be systematically lower than the actual values (see Section 3.4).

## 6 FILAMENTS IN THE COSMIC WEB

Without doubt, the most salient features of the cosmic web are the large filamentary networks, which are interconnected across tens and even hundreds of megaparsecs. In this section, we focus specifically on the filaments identified with the MMF and study their general properties such as length, density profile, connectivity, etc.

### 6.1 Filament and wall compression

Filaments and walls have a complex topology. Their general shape may be far from idealized lines and planes. Properties such as direction, density profiles, extent and other measures derived from these quantities are difficult to interpret or meaningless without a proper reference point. We address this problem by defining their ‘heartline’ in a similar way as the centre of mass in spherical clusters is used as reference point. We define the 1D and 2D counterparts for

filaments and walls. They will be referred as the ‘spine’ of filaments and the ‘plane’ of walls, respectively.

In order to infer the idealized lines and planes from the complex filaments and walls, we performed an iterative algorithm that compresses structures along their perpendicular direction (normal to the spine of the filament or the plane of the wall) by moving each particle (or halo) to the centre of mass inside a spherical window centred on the particle until its position converges. The movement of the particles is restricted along the perpendicular direction to the spine of the filament or the plane of the wall (see Appendix A for details). This procedure *enhances* filaments and walls compressing them closer to idealized structures: filaments become 1D lines while walls are compressed to nearly planar 2D planes (see Fig. A1).

In the determination of spines and planes, we based ourselves on dark matter haloes instead of particles. Spines and planes derived from the raw dark matter particles tend to cross the centres of large haloes since they contain most of the matter in the neighbourhood. This gives the same result as derived with the use of haloes. However, in computing the density profile the difference between particles and haloes can become important. The density profile is dominated at small scales by large haloes, giving the false impression of highly dense cuspy cores or even worse, producing several ‘cores’ of a single filament (Colberg et al. 2005a).

### 6.2 The filamentary network

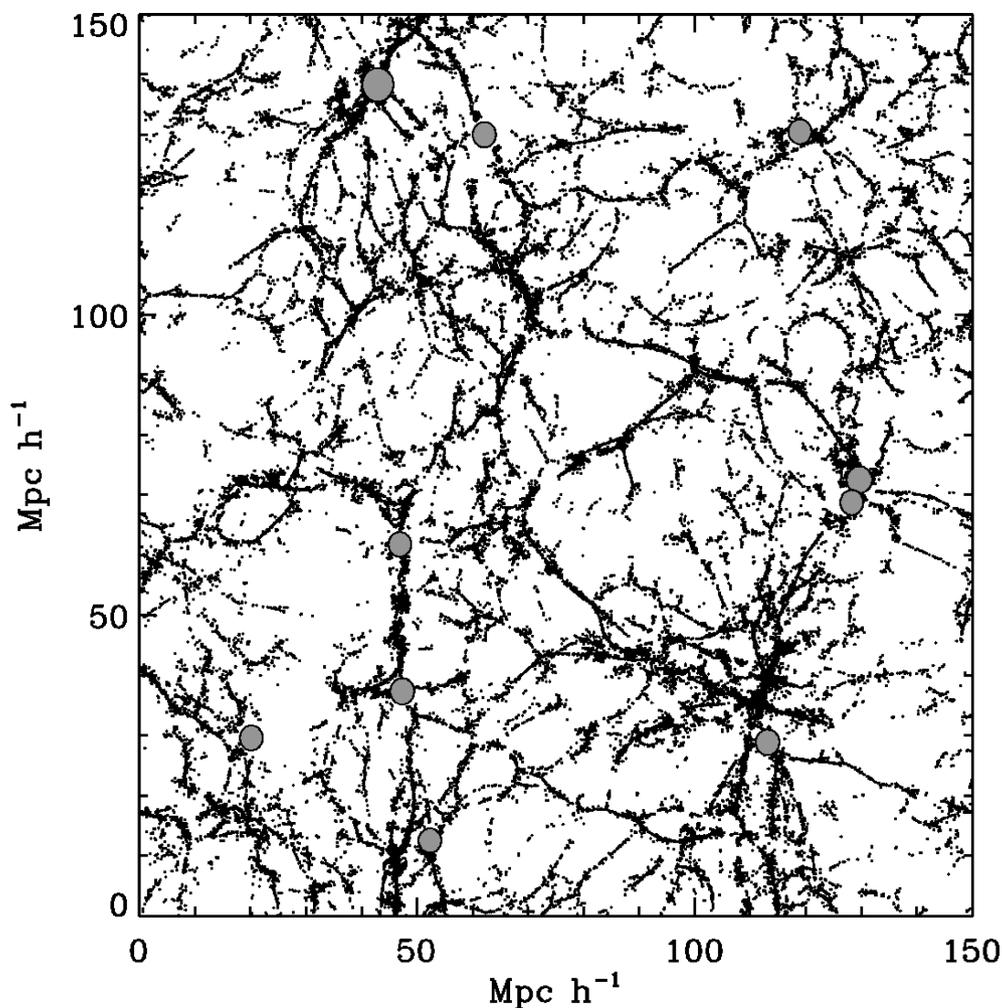
Fig. 11 shows a slice of the simulation box in which filaments have been compressed to delineate their spines. Grey circles indicate the location of clusters with masses above  $10^{14} M_{\odot} h^{-1}$ . This figure presents the cosmic web as a network of interconnected filaments spread all over the simulation box. The clusters sit at the intersections or ‘nodes’ of the network. The filamentary network permeates all regions of space, even the very underdense voids. One important aspect of the filamentary network is its cellular nature (Joeveer & Einasto 1978; Zeldovich et al. 1982) which defines a multiscale system marked by structures over a range of scales (Sheth 2004; Sheth & van de Weygaert 2004; Shen et al. 2006). Large voids are delineated by thick large filaments. Each of these voids contains subsystems of smaller filaments delineating smaller mini voids which in turn form the basis for even smaller systems (Dubinski et al. 1993; Schmidt, Ryden & Melott 2001; Sheth & van de Weygaert 2004; Shen et al. 2006; Aragón-Calvo et al. 2010). Also, large empty voids contain extremely tenuous but rich filamentary systems only seen in high-resolution simulations (van de Weygaert & van Kampen 1993; Gottlöber et al. 2003; Colberg et al. 2005b).

### 6.3 Filamentary network: percolation and connectivity

The raw output of the MMF is an object Map  $\mathcal{O}$ , which defines which pixels belong to a given morphology (see Aragón-Calvo et al. 2007b). Given the coherence of the filamentary (and sheet-like) patterns in the cosmic web, such pixels connect with each other into large volume pervading complexes. This network connects filaments of a large variety of sizes, in turn branching into smaller filaments which are often multiply connected.

Such connectivity information, as well as the individual properties of filaments, is not explicit in the object map. To this end, we assess the percolation behaviour of the *density field contained in the filamentary network*. This avoids the ambiguity in the segmentation of morphologies based on density alone (see Section 5.2).

The percolation analysis studies the change in number, properties and/or connectivity of objects defined as regions of space



**Figure 11.** Filamentary network in a slice of  $20 h^{-1}$  Mpc. Back dots indicate dark matter particles in filaments after the compression algorithm. Grey circles indicate the location of clusters with  $M \geq 10^{14} M_{\odot} h^{-1}$ . The size is scaled proportional to their mass.

above a given density contrast threshold  $\delta_{\text{th}} \equiv \rho_{\text{th}}/\bar{\rho} - 1$ . By varying the threshold across the complete range of density values in the matter distribution, we obtain a systematically evolving population of structures, each characteristic for the value  $\delta_{\text{th}}$  (Zeldovich et al. 1982; Shandarin 1983; Shandarin & Zeldovich 1983; Klypin 1988).

We assess the change in filamentary complexes as a function of the density contrast threshold:

$$1 + \delta_{\text{th}} \equiv \left( \frac{4}{3} \pi l_{\text{link}}^3 \right)^{-1}, \quad (6)$$

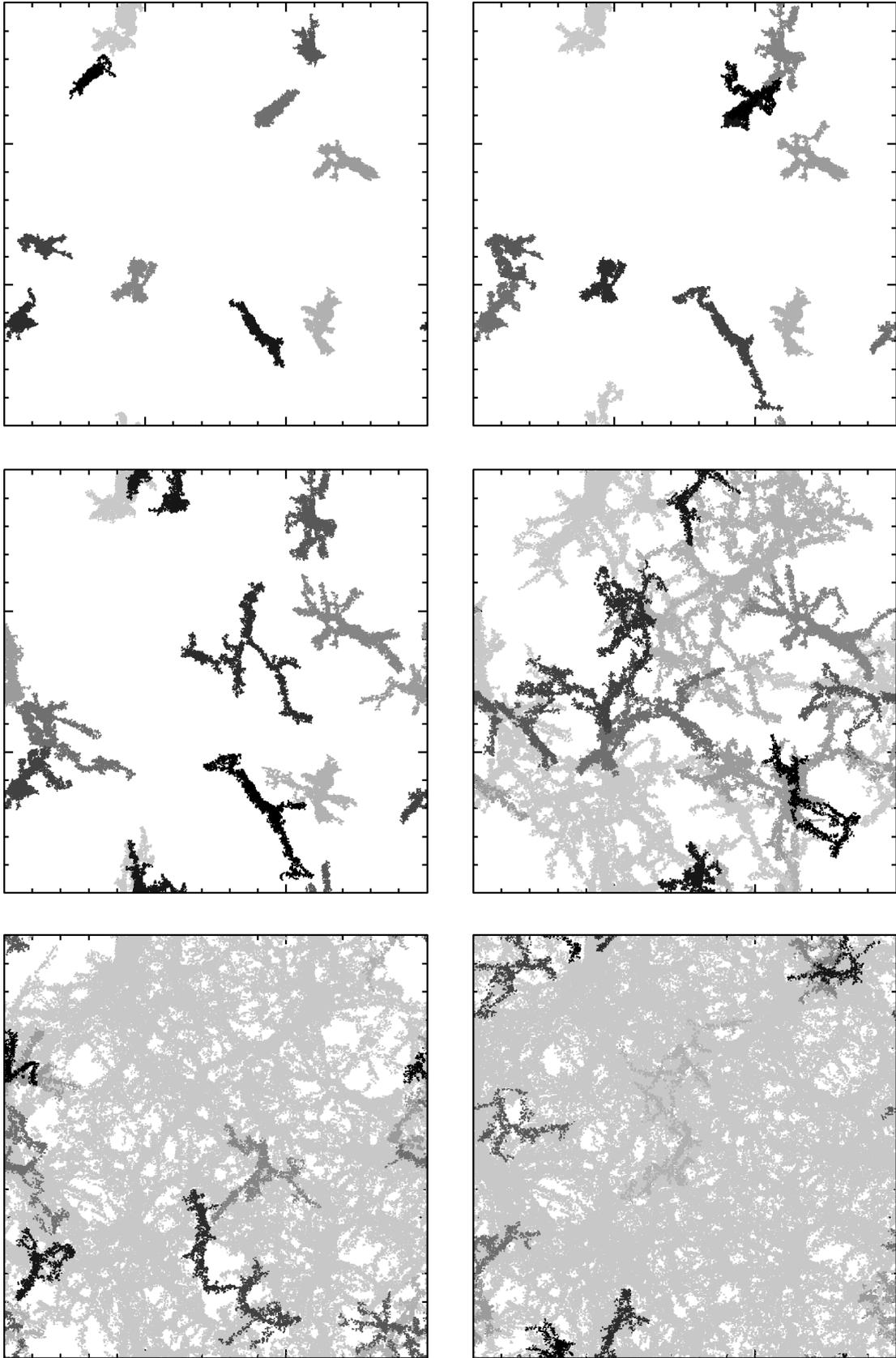
where  $l_{\text{link}}$  is a linking length between two dark matter particles. By iteratively associating particles with separations  $d \leq l_{\text{link}}$ , we produce a catalogue of filamentary complexes. This procedure is rather similar to the identification of clusters using the FoF algorithm. Following their identification, we rank the filamentary configurations by their mass, i.e. the number of particles they contain. This results in a mass ordered list consisting of the most massive filament, second most massive filament, third most massive, etc.

At high densities, filaments are isolated objects with a simple shape and topology. As the value of  $\delta_{\text{th}}$  decreases, the filaments grow steadily while more mass from their surroundings is added

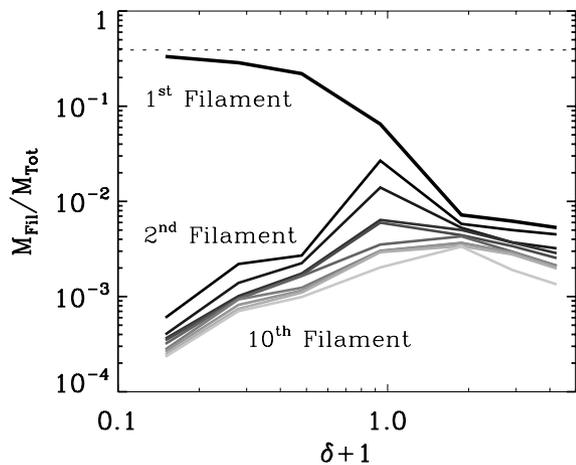
to them.<sup>1</sup> While they grow, they branch into increasingly complex structures. At some point, at the merging threshold  $\delta_{\text{th}} \sim 2$ , there is a rather sudden transition in the way the filaments grow. The steady inclusion of mass from adjacent lower density regions no longer constitutes the main growth process. Instead, the merging of existing filamentary complexes into super-filaments becomes the main mode of structure growth. Descending to even lower densities, the filaments continue to merge until, rather abruptly, at one particular density value a single superstructure emerges which spans the entire volume: this marks the *percolation* transition. As a result, opposite faces of the simulation box are connected.

The growth process is illustrated in Fig. 12. It follows the development of the 10 most massive filaments along a range of decreasing density thresholds  $\delta > \delta_{\text{th}}$  (going from top left to bottom right). The figure shows the filaments at thresholds  $1 + \delta_{\text{th}} = 0.2, 0.4, 0.9, 1.8, 2.9$  and  $4.2$  (from top to bottom and left to right). In order to distinguish them, each of the filaments is plotted with a different grey tone, with the lighter shades corresponding to more massive structures. The panels highlight the non-linear nature of the percolation

<sup>1</sup> By restricting ourselves to the filament pixels, the structures remain confined to filaments and do not flood into walls or voids.



**Figure 12.** Filament percolation and connectivity. The 10 most massive filamentary structures at different density contrast thresholds,  $\delta_{\text{th}} = 0.2, 0.4, 0.9, 1.8, 2.9, 4.2$  (from top to bottom and left to right). In order to differentiate between them, we plot each structure with a different grey tone. The lighter shades correspond to more massive structures. See the text for the description.



**Figure 13.** Mass of the first up to the 10th largest filaments, in units of the total mass in the simulation box. The horizontal dotted line indicates the total mass content in filaments (see Table 3).

process, with the initial gradual growth suddenly transiting into the merging of filaments and the emergence of super-complexes.

The corresponding growth in mass of the 10 most massive filaments, as a function of threshold density  $\delta_{\text{th}}$ , is plotted in Fig. 13. It depicts the mass fraction of each of these filaments. At high values of  $\delta_{\text{th}}$ , the 10 filaments have similar masses, confirming the impression obtained from Fig. 12. At low, percolating, density values there is a distinct difference between the largest filament and all other filaments. Towards the lowest density values,  $\delta \approx -0.8$ , the most massive filament has absorbed the major share of filamentary objects. It asymptotically attains a mass of  $\sim 40$  per cent of the total mass, which is the entire mass enclosed by the filamentary network (see Table 3). Meanwhile, the mass of the remaining filaments decreases continuously. As their more massive peers get absorbed into the percolating principal filament, the remaining isolated objects represent ever smaller specimens of the filament population.

The largest structure in the percolation process, referred to as the *principal percolating filament*,<sup>2</sup> carries important information on the topology of the density field (Shandarin et al. 2004). The principal percolating filament has significantly different properties than the rest of the (much smaller) filaments. It is a space-covering network connecting all regions of space and hardly changes significantly once it has reached the percolation threshold.

#### 6.4 Isolating individual filaments

The ability to recognize individual features such as filaments is natural to the human brain. However, the analogue for computational recognition still represents a major challenge. To *identify and isolate* the individual elements forming the interconnected network, we need to invoke post-processing procedures. This involves the definition and introduction of user-specified measures.

As we have seen in previous sections, strictly speaking the filamentary network is a system that connects *all* filamentary features. In this sense it does not constitute a sample of individual isolated structures, so that any attempt towards the identification of individual filaments necessarily involves a level of subjectivity. And even though the MMF formalism provides us with an objective measure

<sup>2</sup> We use the term filament since it is contained inside the filament object map (see Aragón-Calvo et al. 2007b)

of filamentariness at each position of space, it remains far from trivial for the MMF to dissect the filamentary network into objectively defined individual filaments.

One strategy to dissect the filamentary network into individual objects is by exploiting its percolation properties. Following the argument of Shandarin et al. (2004) that individual objects identified by means of density thresholds should be studied *before percolation occurs*, we use the same principle to select the density threshold for defining individual filaments.

From Fig. 13 we see that the merging density threshold, below which the identified features start to merge with each other, is in the order  $\delta_{\text{th}} \sim 1$ . In practice, we use a somewhat larger value of the threshold in order to eliminate the possibility of filament mergers via thin tenuous bridges whose significance might be questioned (this is a known problem of the FoF algorithm for halo detection; Eisenstein & Hut 1998). Visual inspection indicates that the differences are small, a consequence of the restriction to regions identified by the MMF as filaments. After some experimentation, we use the value  $\delta_{\text{th}} = 3$  as the density threshold for the identification of individual filaments.

#### 6.5 Filament classification

Fig. 14 show four examples of typical filaments. For visualization purposes, we plot dark matter particles taken from the  $128^3$  simulation in grey colour. Superimposed on these, we plot the spine of the filament by means of the black particles. Haloes delineate the same structures, be it more diluted. In general, we find that the level of complexity of filaments is related to the surrounding large-scale matter configuration. Filaments in the vicinity of massive clusters form more complex systems than those connected to less massive clusters.

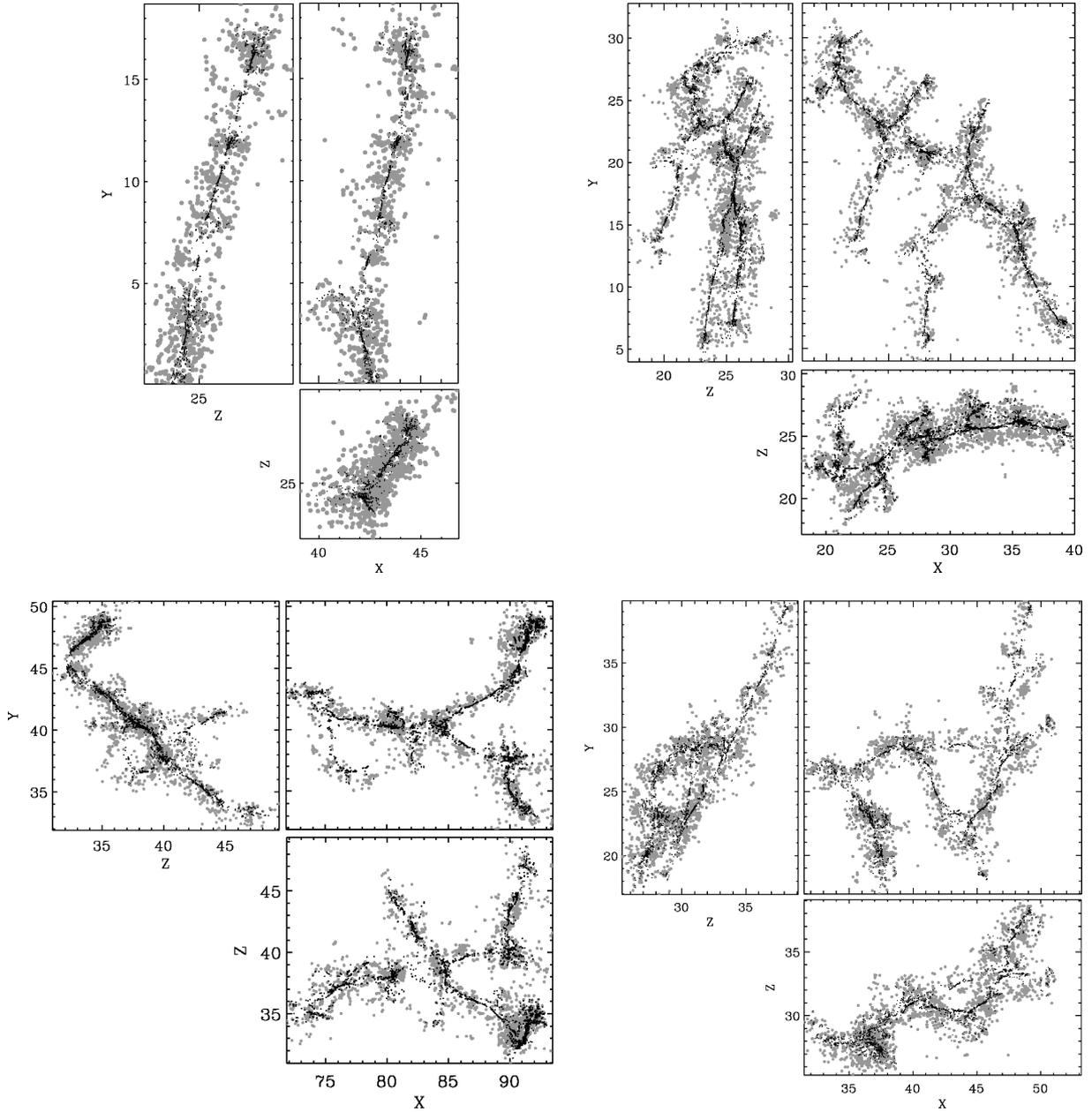
The fractal nature of the filamentary network makes it difficult to classify individual filaments since in principle they form a percolating network that includes all filaments. Also, the branching properties of our filaments ultimately depend on the resolution limit of our simulation. This is a natural consequence of the hierarchical development of the cosmic web (Sheth 2004; Sheth & van de Weygaert 2004; Shen et al. 2006). Ultra-high-resolution  $N$ -body simulations show that even in the most underdense regions, one can find systems of tenuous filaments extending along the whole physical extent of the voids (van de Weygaert & van Kampen 1993; Gottlöber et al. 2003; Platen, van de Weygaert & Jones 2007).

On the basis of a rough phenomenological inventory of the shape and morphology of the filaments in our simulations, we distinguish four basic types of filaments.

(i) *Line filaments* do not have branches (or very few) and are mostly straight with lengths in the order of  $5\text{--}20 h^{-1} \text{Mpc}$ . They are often found as ‘bridges’ between massive clusters. Shorter filaments are also straighter than large ones. They may be identified with the *intracluster filaments* studied by Colberg et al. (2005a) and filament Types 0, I and II in the classification of Pimblet & Drinkwater (2004).

(ii) *Grid filaments* are often found crossing vast regions with no massive clusters crossing them. They form the surrounding ‘net’ enclosing large voids and are almost invariably 2D, suggesting that walls are in fact delineated by these kind of filaments. Even when they consist of several branches, one can often identify a main ‘path’ with smaller filaments running almost perpendicular to it.

(iii) *Star filaments* have a well-defined ‘centre’, usually a cluster or large group from which several ‘arms’ stretch out. Star filaments



**Figure 14.** Filament diversity. Top: three orthogonal projections of typical line (top) and star (bottom) filaments. Dark matter particles are indicated by filled grey circles. The spine of the filament is also shown, by black dots, in order to better delineate the shape of the filaments. Bottom: three orthogonal projections of typical grid (top) and complex (bottom) filaments. Dark matter particles are indicated by filled grey circles.

can be considered a smaller version of grid filaments. They are also 2D structures, suggesting that grid and star filaments represent the same kind of structures.

(iv) *Complex filaments* do not have a clear shape; they are often multiply connected and it is difficult to define a main path or direction. These filaments can be found in almost any environment.

## 6.6 Length of filaments

While describing filaments in terms of their mass is straightforward (see Section 6.3), the determination of their length and related physical properties involves several complications because of the branching nature of the filamentary network. The length of com-

plex systems composed of several interconnected branches is not straightforwardly or uniquely defined. One may even argue that it is a rather meaningless concept in the case of filaments with strong branching such as grid and star filaments, for which it is almost impossible to define a main spine.

We may consider various options for the definition of the length of a filament as follows.

(i) The total length of all the branches of filament, related to the fractal nature of the filamentary network (Martinez et al. 1990). Its definition presents several practical complications such as the identification of the branching points in the main path.

(ii) The length of the main path of the filament (Pimblet & Drinkwater 2004; Colberg et al. 2005a). This definition presumes

the existence of such a main path, whose definition introduces an arbitrary choice of defining criteria.

Given the high level of complexity of the algorithms used to identify the total length of filamentary systems, we postpone their study for future work. Instead, here we concentrate on the length of the main path of the filament. To first order, this can be identified with the thickest or longest branch.

We are particularly interested in the length of filaments connecting to clusters. These do not only appear as the nodes of the cosmic web, but they also define the formation sites of filaments (Bond et al. 1996). They therefore provide a natural way of dissecting the filamentary network. Accordingly, we proceed different than in the percolation-based dissection described in Section 6.4. For the construction of the filament catalogue, we use the haloes instead of the dark matter particles to trace the filamentary network. This has the advantage of being faster when computing the length of the filaments and of ignoring the irrelevant small details.

We start by collecting all clusters with a mass  $M \geq 10^{14} M_{\odot} h^{-1}$ . Around the locations of these clusters, we cut spheres with a radius of  $2 h^{-1}$  Mpc. The filaments contained in these spheres involve smaller individual objects, which are easier to handle. From this set of isolated filaments, we produce a catalogue using an FoF algorithm with a linking length corresponding to a density threshold of  $\delta_{\text{th}} + 1 = 4$ . Note that in this way we exclude large superstructures, as the filaments are broken up at cluster nodes.

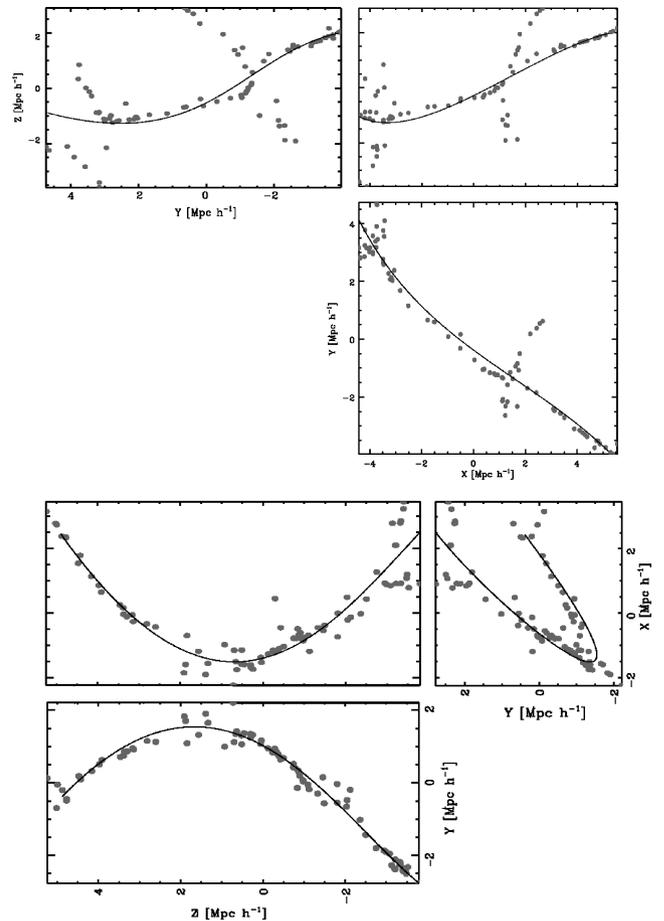
To analyse the length of the identified filaments, we use a third-order polynomial fit to describe and quantify their shape (see Appendix B for details). Two representative filaments are shown in Fig. 15, along three mutually orthogonal directions. The best polynomial fit is superimposed on the related halo distribution. The fit manages to closely follow even the most intricate filaments and also ignores small branches. This is particularly visible in the case of the *grid* filament in the top panel, whose main branch is crossed by several smaller filaments.

The length and mass distribution of the resulting filament sample is shown in Fig. 16. Small filaments are clearly more abundant than the large ones, as we may expect for a hierarchically evolved distribution. Filaments with lengths in excess of a few tens of megaparsecs are extremely rare. In terms of their mass, we also see that there are hardly any filaments with masses larger than  $\sim 10^{14} M_{\odot} h^{-1}$ . In other words, while the largest and most massive filaments are the most prominent structures in the cosmic web, they represent only a minor fraction of the entire filament population.

## 6.7 The density profiles of filaments

Filaments are far from being smooth uniform structures. In most cases it is possible to identify a highly dense spine surrounded by more diffuse matter. Filaments are also populated by compact dense haloes. This yields a resemblance of filaments to a pearl necklace, with haloes suspended along the bridging spine between massive clusters. Despite the prominence of the inner realms of filaments, there are only a handful of studies addressing their density profile with respect to the filament's spine (e.g. Colberg et al. 2005a; Dolag et al. 2006). This is even more true for the density distribution of walls, which has only been addressed in a few theoretical studies (see e.g. Zel'dovich 1970; Shandarin & Zel'dovich 1989).

Visual inspection of the observed cosmic web reveals that there are indeed filaments spanning across several tens of megaparsecs. Examples of these are the 'spine' of the Pisces-Perseus supercluster (Gregory, Thompson & Tifft 1981; Giovanelli & Haynes 1985)

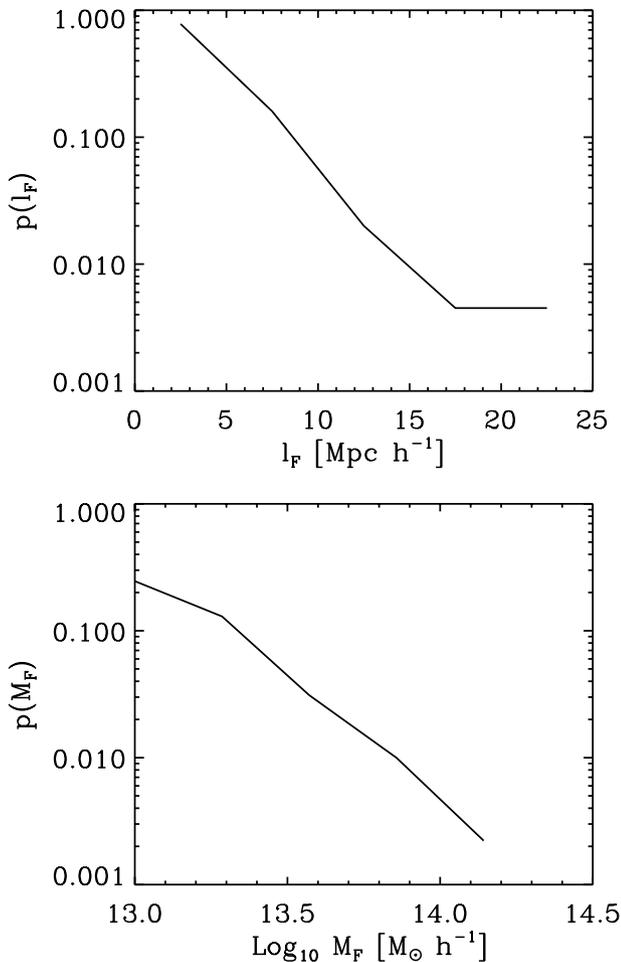


**Figure 15.** Polynomial fit (solid line) of two of the largest filaments in the simulation. Dark matter haloes are represented by grey circles.

and the planar filamentary system known as the Sloan Great Wall (Gott et al. 2005; Platen 2009). However, such massive systems are marked by a substantial degree of substructure, containing numerous clusters and a filigree of small-scale filaments.

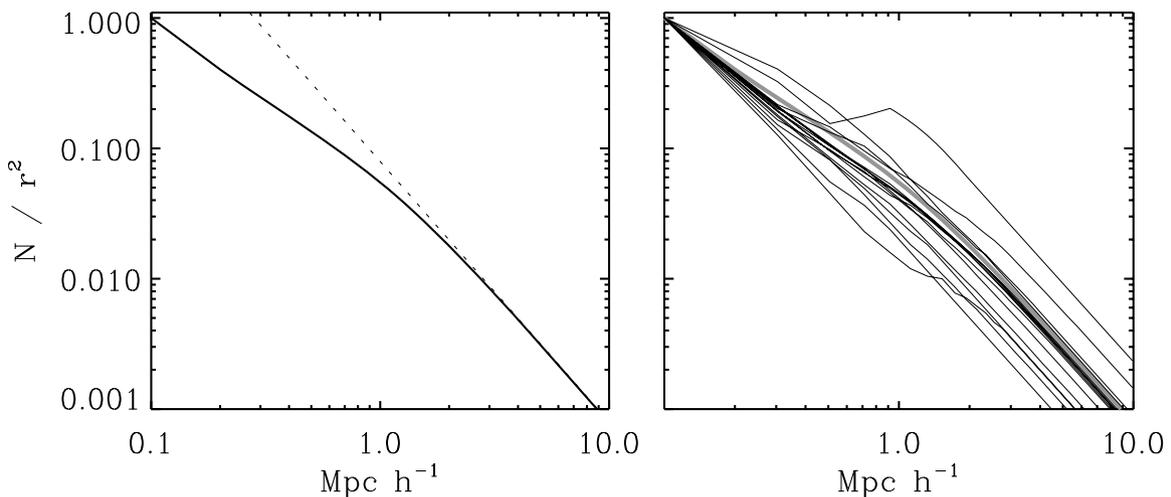
In general, large haloes dominate the density profile. This manifests itself in a cuspy profile centred at the spine of the host filaments and even occasionally involving several peaks near the centre. Following this observation, we use haloes instead of dark matter particles, thereby avoiding contamination of the radial profile at small scales. The density profiles of the filaments are determined by counting the *enclosed number of haloes*, weighted by their mass, in bins of increasing radial distance from the spine of the filament. The radial distance of a halo with respect to the spine of its host filament is defined as the displacement of a halo *before* and *after* we apply the compression algorithm described in Section 6.1 (see Appendix A for details). To enable the mutual comparison of the density profiles of all filaments, they are all scaled to a value  $N(r=0) = 1$  at the centre of each filament. The resulting density profile, averaged over all filaments, is shown in the left-hand panel of Fig. 17. A similar analysis for walls reveals that they have less well-defined boundaries with widths ranging between 5 and  $8 h^{-1}$  Mpc.

The average filament profile has a power-law shape,  $N(r) \propto r^{-2}$ , beyond a radius of  $r \sim 2 h^{-1}$  Mpc. For a 1D entity like a filament, this implies that mass is no longer attached to the filament at larger radii. In other words, the radius of  $r \sim 2 h^{-1}$  Mpc marks the average maximum extent of a filament. Within this range, the profile turns to



**Figure 16.** Probability density distribution of length (top panel) and mass (bottom panel) of filaments (see the text for details). Only filaments with more than 10 haloes were considered.

a power-law shape with a slope  $\gamma \approx -1$ , which corresponds roughly to an isothermal profile for a filamentary entity. The profile slope transition at around  $r \sim 2 h^{-1} \text{Mpc}$  provides a simple criterion for defining the width of filaments. The fact that there appears to be



**Figure 17.** Filament density profile. Left: average enclosed density profile of haloes inside a filament. The dotted line corresponds to a power-law fit  $N(r) \propto r^{-2}$ . Right: density profiles of 15 individual filaments. Superimposed (thick grey line) is the average enclosed density profile. To enable their mutual comparison, all density profiles are scaled to a value  $N(r=0) = 1$  at the centre of the filament.

only a small variation in this width (see Fig. 17) means that we may have some confidence in using this one particular value.

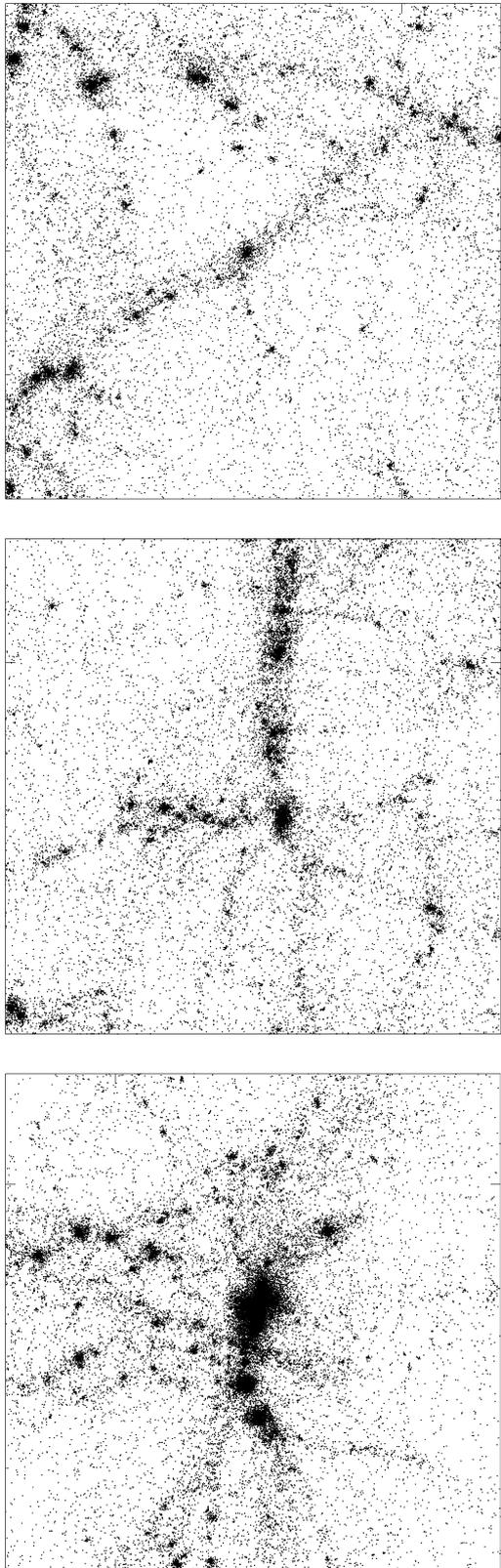
To provide an impression of the variation in the density distribution around filaments, we show the density profile for 15 individual filaments in the right-hand panel of the same figure. While the individual profiles differ, their variation is restricted to a rather small range. We also find some variation in the maximum extent of filaments, which confirms the impression obtained from the observed cosmic web and had already been noted in previous studies (Colberg et al. 2005a; Dolag et al. 2006).

### 6.8 Cluster bridges

Filaments are closely affiliated to clusters. According to the cosmic web theory (Bond et al. 1996), filaments are the transport channels along which clusters accrete mass (see van Haarlem & van de Weygaert 1993; Diaferio & Geller 1997; Colberg et al. 1999). Massive clusters are formed at the sites of rare high-density peaks in the primordial matter distribution (Bardeen et al. 1986) and dominate the dynamical evolution of the cosmic matter distribution. Particularly in the high-density areas, the overwhelming coherent gravitational tidal forces between two cluster peaks are responsible for anisotropic collapse of the surrounding matter towards elongated filaments (Bond & Myers 1996; Bond et al. 1996). The strength of the filamentary bridges is expected to depend on the mass of the generating cluster, their mutual distance and their mutual orientation (Bond et al. 1996). This has been confirmed by numerous numerical simulations (see e.g. Colberg, Krughoff & Connolly 2005a; Sousbie, Colombi & Pichon 2009; Gonzalez & Padilla 2010). A visual assessment of observations as well as simulations also suggests that the more prominent clusters are associated with filamentary systems of a higher complexity (Colberg et al. 1999, 2005a; Pimblett & Drinkwater 2004).

Fig. 18 shows three examples of groups and clusters with masses in the range of  $10^{13}$ – $10^{14} M_\odot h^{-1}$ . A first impression is that massive clusters appear to be embedded within a more complex filamentary environment. On the basis of this figure, we may make a few direct observations:

- (i) The top panel shows a cluster connected to two filaments. Its mass,  $M \sim 10^{13} M_\odot h^{-1}$ , is characteristic for a group of galaxies



**Figure 18.** Three examples of clusters and the filaments to which they are connected. The three clusters have masses in the range of  $10^{13}$ – $10^{14} M_{\odot} h^{-1}$ . Top: a group-like cluster, with  $M \sim 10^{13} M_{\odot} h^{-1}$ , connected to two filaments. Middle: a medium-sized cluster, with  $M \sim 2\text{--}3 \times 10^{13} M_{\odot} h^{-1}$ , connected to several filaments. Bottom: a massive cluster, with  $M \sim \times 10^{14} M_{\odot} h^{-1}$ , at the centre of a highly complex environment of filamentary branches.

consisting of a few tens of galaxies. Such filaments may be the result of a gravitationally induced fragmentation of a longer filament. The infall pattern of matter into these clusters is highly anisotropic and is mainly restricted to the direction of the connecting filaments.

(ii) The cluster in the middle panel has a mass of  $M \approx 2\text{--}3 \times 10^{13} M_{\odot} h^{-1}$ . It is connected to several filaments. It is also indicative of the fact that we find several other clusters in its vicinity.

(iii) The cluster in the bottom panel is embedded in a highly complex environment. Several filaments can be seen branching in a range of different directions. It is not possible to identify a main filament to which the cluster is connected.

Following these observations along with others obtained from the simulation, we find that the number of filamentary extensions of a cluster is closely related to the mass of the cluster.

### 6.8.1 Cluster mass and filament connections

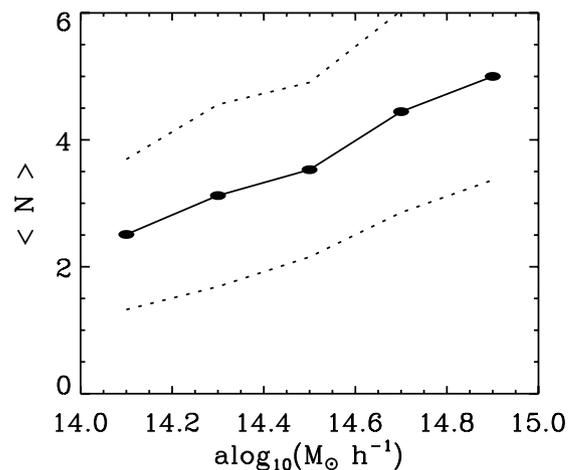
In order to quantify the relation between the mass of a cluster and the number of connected filaments, we used the filament catalogue described in Section 6.6. To this end, we applied the following criteria.

(i) A filament is connected to a cluster if it has at least one halo within a sphere of a radius of  $3 h^{-1} \text{ Mpc}$  from the centre of the cluster.

(ii) Only clusters with  $M \geq 10^{14} M_{\odot} h^{-1}$  were considered in our analysis.

Fig. 19 suggests an almost linear relation between the mass of a cluster and the number of filaments that are connected to the cluster, although the dispersion around the relation is rather large. The most massive clusters, the ones with a mass in excess of  $M \sim 10^{15} M_{\odot} h^{-1}$ , may form a node connecting to five or even six different filaments. Low-mass clusters with a mass of  $\leq 10^{14} M_{\odot} h^{-1}$ , on the other hand, tend to have rather simple connections of two and not more than three filaments.

The figure suggests that some clusters may even have only one filament connected to them. However, this is not observed in the cosmic web and would be difficult to justify from a physical point of view. It is mainly the result of the method used to assign filaments



**Figure 19.** Cluster mass and filament connections. The figure plots the mean number of filaments,  $\langle N \rangle$ , connected to a cluster as a function of the mass  $M_{\text{cl}}$  of the cluster (solid line). The dotted line indicates the  $1\sigma$  dispersion of the data around the mean (see the text).

**Table 4.** Inventory of the cosmic web in terms of volume and mass content.

	Clusters (per cent)	Filaments (per cent)	Walls (per cent)	Voids (per cent)
Volume filling	0.38	8.79	4.89	85.94
Mass content	28.1	39.2	5.45	27.25
Relative density	73	4.45	1.11	0.31

to clusters, which may in some situations miss a few faint filaments which would be connected to these low-mass clusters.

The large dispersion in Fig. 19 reflects the complexity of the cosmic web. The final matter configuration in the neighbourhood of a cluster depends not only on its mass but also on the geometrical configuration of the surrounding clusters (Bond et al. 1996). Other studies have found a similar relation based on intracluster filaments found in  $N$ -body simulations (Colberg et al. 2005a) and visually identified filament–cluster connections from the 2dF galaxy survey (Pimblet & Drinkwater 2004).

## 7 CONCLUSIONS

We provide a qualitative and quantitative description of the cosmic web in terms of its morphological constituents. We focused on filaments, to a lesser degree on walls. The basis for this work is a large  $N$ -body simulation of a  $\Lambda$ CDM cosmology with dark matter particles. The morphological segmentation was done with the MMF.

(i) The mass content, volume content and mean density of the cosmic web are quantitatively summarized in Table 4.

(ii) Each morphology of the cosmic web has a characteristic density. The distribution of densities however, has a large overlap. Density alone can give an indication of the morphology, but it cannot be used to unambiguously segment the cosmic web into its morphological constituents.

(iii) We offer a qualitative classification of filaments based on their visual properties in four types: *line*, *star*, *grid* and *complex*.

(iv) The density profile of filaments indicates that their typical radial extent is of the order of  $2 h^{-1}$  Mpc, although there are significant variations between filaments. In their interior, filaments have a power-law density profile with slope  $\gamma \approx -1$ , corresponding to an isothermal density profile.

(v) We find a relation between the mass of a cluster and the number of filaments it has connected. More massive clusters have more filaments in average. Clusters with a mass of  $\sim 10^{14} M_{\odot} h^{-1}$  have on average two connecting filaments while clusters with a mass of  $\sim 10^{15} M_{\odot} h^{-1}$  have on average five connecting filaments.

Having analysed and described the structure of the filamentary network, in the subsequent paper we will address the velocity flows and the dynamics of the network and of individual clusters. Also, we plan to experiment with the MMF detection technique, and instead of addressing the multiscale character of the density field develop a version based on the dynamically more relevant gravitational potential field.

In a related study (Jones et al. 2010), we have applied the MMF to detecting filaments in the SDSS galaxy redshift survey and identified edge-on galaxies within their realm. This allowed us to address and answer the question whether there are significant alignments of galaxy spins along the spine of the SDSS filaments. This is expected following the tidal torque theory for galaxy angular momentum generation. The MMF indeed allowed us to identify a subset of galaxies and filaments conforming to a significant alignment.

The application of the MMF to real galaxy surveys introduces a few important additional challenges. One aspect concerns the survey volume selection. While a volume-limited survey would guarantee an ideal homogeneous coverage, in practice it involves a severe reduction of spatial resolution and hence the feasibility of identifying crucial aspects of the anisotropic features in the cosmic web. For magnitude-limited galaxy redshift surveys, the DTFE density field reconstruction enables a correction for the diminishing sampling density at higher density. None the less, an appropriate MMF web analysis will still be restricted to the more densely sampled regions out to the peak of the radial survey selection function. A recent meticulous analysis by Platen et al. (2010) has assured us that our web analysis tool box can be successfully tuned towards controlling this issue.

An additional artefact that may severely affect the identification of filaments is the imprint of redshift distortions. The large peculiar motions in and around the virialized cluster cores generate artificial radially directed elongated features, better known as the *Fingers of God*. In practice, we remove all filaments within a limiting angle  $\eta_c$  around the line of sight. Jones et al. (2010) show that this successfully recovers a statistically proper filament distribution. Large-scale cosmic flows are known to enhance the contrast of anisotropic filamentary features (see e.g. Shandarin 2009). However, a proper correction for this would at least demand a densely sampled environment so that one may model the corresponding force field. This work is in progress and has not yet been implemented in our tool box.

## ACKNOWLEDGMENT

We are grateful to the (anonymous) referee for the careful and constructive appraisal of our manuscript, leading to substantial improvements.

## REFERENCES

- Aragón-Calvo M. A., 2007, PhD thesis, Univ. Groningen  
Aragón-Calvo M. A., van de Weygaert R., Jones B. J. T., 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., 2008, preprint (arXiv:0809.5104)  
Aragón-Calvo M. A., van de Weygaert R., Araya-Melo P. A., Platen E., Szalay A. S., 2010, *MNRAS*, 404, 89  
Araya-Melo P. A., Reisenegger A., Meza A., van de Weygaert R., Dünner R., Quintana H., 2009, *MNRAS*, 399, 97  
Bardeen J. M., Bond J. R., Kaiser N., Szalay A. S., 1986, *ApJ*, 304, 15  
Barrow J. D., Bhavsar S. P., Sonoda D. H., 1985, *MNRAS*, 216, 17  
Bharadwaj S., Pandey B., 2004, *ApJ*, 615, 1  
Bond J. R., Myers S. T., 1996, *ApJS*, 103, 1  
Bond J. R., Kofman L., Pogosyan D., 1996, *Nat*, 380, 603  
Bond N., Strauss M., Cen R., 2010a, *MNRAS*, 406, 1609

- Bond N. A., Strauss M., Cen R., 2010b, MNRAS, submitted (arXiv:1003.3273)
- Broadhurst T. J., Ellis R. S., Koo D. C., Szalay A. S., 1990, Nat, 343, 726
- Chazal F., Cohen-Steiner D., M erigot Q., 2009, INRIA Rapport de Recherche 6930
- Choi E., Bond N. A., Strauss M. A., Coil A. L., Davis M., Willmer C. N. A., 2010, MNRAS, 406, 320
- Cohen J. G., Cowie L. L., Hogg D. W., Songaila A., Blandford R., Hu E. M., Shopbell P., 1996, ApJ, 471, L5
- Colberg J. M., 2007, MNRAS, 375, 337
- Colberg J. M., White S. D. M., Jenkins A., Pearce F. R., 1999, MNRAS, 308, 593
- Colberg J. M., Krughoff K. S., Connolly A. J., 2005a, MNRAS, 359, 272
- Colberg J. M., Sheth R. K., Diaferio A., Gao L., Yoshida N., 2005b, MNRAS, 360, 216
- Colless M. et al., 2003, preprint (astro-ph/0306581)
- Colombi S., Pogosyan D., Souradeep T., 2000, Phys. Rev. Lett., 85, 5515
- Davis M., Efstathiou G., Frenk C. S., White S. D. M., 1985, ApJ, 292, 371
- de Lapparent V., Geller M. J., Huchra J. P., 1986, ApJ, 302, L1
- Diaferio A., Geller M. J., 1997, ApJ, 481, 633
- Dolag K., Meneghetti M., Moscardini L., Rasia E., Bonaldi A., 2006, MNRAS, 370, 656
- Doroshkevich A. G., 1970, Astrophysics, 6, 320
- Doroshkevich A. G., Kotok E. V., Poliudov A. N., Shandarin S. F., Sigov I. S., Novikov I. D., 1980, MNRAS, 192, 321
- Dubinski J., da Costa L. N., Goldwirth D. S., Lecar M., Piran T., 1993, ApJ, 410, 458
- Ebeling H., Barrett E., Donovan D., 2004, ApJ, 609, L49
- Eisenstein D. J., Hut P., 1998, ApJ, 498, 137
- Eisenstein D. J., Loeb A., 1995, ApJ, 439, 520
- Eisenstein D. J., Loeb A., Turner E. L., 1997, ApJ, 475, 421
- Eke V. R., Cole S., Frenk C. S., Navarro J. F., 1996, MNRAS, 281, 703
- Florack L. M. J., ter Haar Romeny B. M., Koenderink J. J., Viergever M. A., 1992, Image Vision Comput., 10, 376
- Forero-Romero J. E., Hoffman Y., Gottl ober S., Klypin A., Yepes G., 2008, MNRAS, 396, 1815
- Frangi A. F., Niessen W. J., Vincken K. L., Viergever M. A., 1998, Lecture Notes in Computer Science Vol. 1496, Medical Image Computing and Computer-Assisted Intervention-MICCAI'98. Springer-Verlag, Berlin
- Geller M. J., Huchra J. P., 1989, Sci, 246, 897
- Genovese C. R., Pero-Pacifco M., Verdinelli I., Wasserman L., 2010, preprint (arXiv:1003.5536)
- Giovanelli R., Haynes M. P., 1985, AJ, 90, 2445
- Gonzalez R. E., Padilla N. E., 2010, MNRAS, in press (arXiv:0912.0006)
- Gott J. R. I., Juri c M., Schlegel D., Hoyle F., Vogeley M., Tegmark M., Bahcall N., Brinkmann J., 2005, ApJ, 624, 463
- Gottl ober S., Łokas E. L., Klypin A., Hoffman Y., 2003, MNRAS, 344, 715
- Graham M. J., Clowes R. G., 1995, MNRAS, 275, 790
- Gregory S. A., Thompson L. A., Tift W. G., 1978, BAAS, 10, 622
- Gregory S. A., Thompson L. A., Tift W. G., 1981, ApJ, 243, 411
- Hahn O., 2009, PhD thesis, ETH
- Hahn O., Porciani C., Carollo C. M., Dekel A., 2007a, MNRAS, 375, 489
- Hahn O., Porciani C., Carollo C. M., 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
- Icke V., 1973, A&A, 27, 1
- 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. Kluwer, Dordrecht, p. 241
- Jones B. J. T., van de Weygaert R., Arag on-Calvo M. A., 2010, MNRAS, in press (arXiv:1001.4479)
- Kirshner R. P., Oemler A. Jr, Schechter P. L., Shectman S. A., 1981, ApJ, 248, L57
- Klypin A. A., 1988, Acta Cosmologica, 15, 101
- Klypin A., Kravtsov A. V., Bullock J. S., Primack J. R., 2001, ApJ, 554, 903
- Lacey C., Cole S., 1994, MNRAS, 271, 676
- Lee J., Pen U.-L., 2000, ApJ, 532, L5
- Martinez V. J., Jones B. J. T., Dominguez-Tenreiro R., van de Weygaert R., 1990, ApJ, 357, 50
- Melott A. L., 1983, MNRAS, 204, 7P
- Novikov D., Colombi S., Dor e O., 2006, MNRAS, 366, 1201
- Ouchi M. et al., 2004, ApJ, 611, 685
- 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
- Pimblet K. A., 2005a, Publ. Astron. Soc. Aust., 22, 136
- Pimblet K. A., 2005b, MNRAS, 358, 256
- Pimblet K. A., Drinkwater M. J., 2004, MNRAS, 347, 137
- Pimblet K. A., Drinkwater M. J., Hawkrigg M. C., 2004, MNRAS, 354, L61
- Platen E., 2009, PhD thesis, Univ. Groningen
- Platen E., van de Weygaert R., Jones B. J. T., 2007, MNRAS, 380, 551
- Platen E., van de Weygaert R., Arag on-Calvo M. A., Jones B. J. T., Vegter G., 2010, MNRAS, submitted
- Pogosyan D. Yu., Bond J. R., Kofman L., Wadsley J., 1998, in Colombi S., Mellier Y., eds, IAP Meeting Vol. 14, Wide Field Surveys in Cosmology. Editions Fronti eres, Paris, p. 61
- Sathyaprakash B. S., Sahni V., Shandarin S. F., 1996, ApJ, 462, L5
- Sato Y., Nakajima D., Atsumi H., Koller T., Gerig G., Yoshida S., Kikinis R., 1998, IEEE Medical Image Analysis, 2, 143
- Schaap W. E., 2007, PhD thesis, Univ. Groningen
- Schaap W. E., van de Weygaert R., 2000, A&A, 363, L29
- Schmidt J. D., Ryden B. S., Melott A. L., 2001, ApJ, 546, 609
- Shandarin S. F., 1983, Soviet Astron. Lett., 9, 104
- Shandarin S. F., 2009, J. Cosmology Astropart. Phys., 2, 31
- Shandarin S. F., Sunyaev R., 2009, A&A, 500, 19
- Shandarin S. F., Zeldovich I. B., 1983, Comments Astrophys., 10, 33
- Shandarin S. F., Zel'dovich Ya. B., 1989, Rev. Modern Phys., 61, 185
- Shandarin S. F., Sheth J. V., Sahni V., 2004, MNRAS, 353, 162
- 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
- Shen J., Abel T., Mo H. J., Sheth R. K., 2006, ApJ, 645, 783
- Sheth J. V., 2004, MNRAS, 354, 332
- Sheth R. K., van de Weygaert R., 2004, MNRAS, 350, 517
- 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, 1
- Sousbie T., Colombi S., Pichon C., 2009, MNRAS, 393, 457
- Springel V., 2005, MNRAS, 364, 1105
- Springel V. et al., 2005, Nat, 435, 629
- Stevens J. A., Page M. J., Ivison R. J., Smail I., Carrera F. J., 2004, ApJ, 604, L17
- Stoica R. S., Mart nez V. J., Mateu J., Saar E., 2005, A&A, 434, 423
- Stoica R. S., Mart nez V. J., Saar E., 2007, Appl. Statistics, 56, 459
- Stoica R. S., Mart nez V. J., Saar E., 2010, A&A, 510, 38
- Sugerman B., Summers F. J., Kamionkowski M., 2000, MNRAS, 311, 762
- Tegmark M., SDSS Collaboration, 2004, ApJ, 606, 702
- Tittley E. R., Henriksen M., 2001, ApJ, 563, 673
- van de Weygaert R., Bertschinger E., 1996, MNRAS, 281, 84
- van de Weygaert R., Bond J. R., 2008, in Plionis M., Lopez-Cruz O., Hughes D., eds, Lecture Notes in Physics Vol. 740, A Pan-Chromatic View of Clusters of Galaxies and the Large-Scale Structure. Springer, Berlin, p. 335
- van de Weygaert R., Schaap W. E., 2009, in Mart nez V., Saar E., Mart nez-Gonzalez E., Pons M., eds, Lecture Notes in Physics Vol. 665, Data Analysis in Cosmology. Springer-Verlag, Berlin, p. 291
- van de Weygaert R., van Kampen E., 1993, MNRAS, 263, 481
- van den Bosch F. C., 2002, MNRAS, 331, 98

van Haarlem M., van de Weygaert R., 1993, *ApJ*, 418, 544  
 Wu Y., Batuski D. L., Khalil A., 2009, *ApJ*, 707, 1160  
 Zel'dovich Ya. 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: FILAMENT AND WALL COMPRESSION

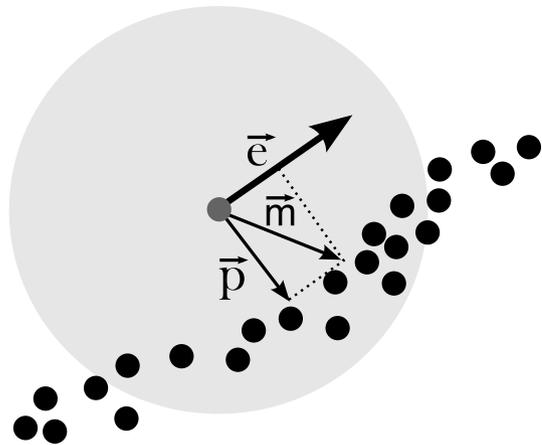
In order to enhance filaments and walls and to morph them closer to idealized lines and planes, we developed a *compression* algorithm. The algorithm displaces particles in the direction of increasing density, towards the spine of the filaments or the planes that define the walls. The algorithm can be applied to particles as well as to haloes, after weighing them by their mass.

The compression procedure exploits the information produced by the MMF. Not only does it identify the filaments and walls in the cosmic web, but also their local direction represented by the eigenvectors of the Hessian matrix (see Arag3n-Calvo et al. 2007b; Section 3). The smallest eigenvector of the Hessian matrix traces the local direction of a filament, while the largest eigenvector locally defines the normal to the wall. This information is exploited to iteratively displace particles to the local centre of mass within a given radius. It is a widely used method for computing the centre of mass in spherical or semispherical haloes (van den Bosch 2002).

We start by defining the *heartline* of filaments and walls in a way similar to determining the centre of mass in spherical haloes. This heartline is used as a reference point. The 1D and 2D heartlines for filaments and walls are referred to as the *spine* of the filaments and the *plane* of the walls, respectively.

Our compression algorithm involves the crucial constraint that the displacement of particles to the centre of mass follows the direction *perpendicular to the filament or wall*. As a result, this process transforms thick structures into thin lines or planes without affecting their length.

In summary, the algorithm has the following steps (for simplicity, we only refer to particles; haloes will be treated equivalently).



**Figure A2.** Illustration of the compression steps for an individual galaxy. It shows how the galaxy is displaced towards a position on the heartline of the filament, along a direction perpendicular to the filament in which it is embedded.

(i) For each particle  $i$  we find all neighbours inside a top-hat window of a given radius  $R_{\text{top}}$ , centred at the particle's position  $x_i$ . The top-hat radius should be large enough to enclose the filament or wall in order to minimize the number of iterations needed.

(ii) The centre of mass  $m_i$  of the particles inside the top-hat window is computed, along with the vector defined by the particle's position and the centre of mass  $m_i$ .

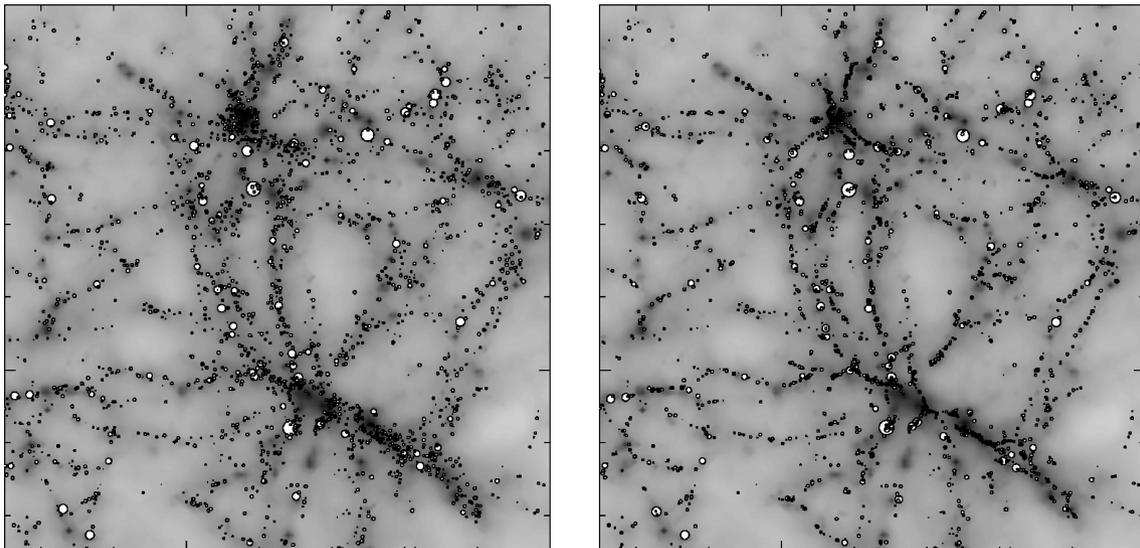
(iii) The particle is displaced to the centre of mass *along the perpendicular direction of the filament/wall*:

$$p = (e * m) * e \sin(\theta), \quad (\text{A1})$$

where  $e$  is the vector indicating the direction of the spine of the filament or the normal to the plane of the wall. The angle  $\theta$  is the angle between the vector  $e$  and the centre of mass  $m$  (see Fig. A2).

The eigenvectors  $e$  are computed from the Hessian matrix smoothed at the characteristic scale of filaments,  $\sim 2-3 h^{-1} \text{ Mpc}$  (see Section 6.7).

(iv) After having performed the process for all particles, we compute the dispersion between the 'original' and 'compressed'



**Figure A1.** Filament compression illustrated. Haloes in filaments before (left) and after (right) the application of the compression algorithm. Haloes are indicated as white circles whose sizes are scaled according to their mass. The grey background delineates the density field plotted in a logarithmic scale.

positions. We repeat the complete process until the dispersion between consecutive iterations changes by a given factor or until the total dispersion is less than a prescribed value. This value specifies the convergence of the method.

The compression algorithm is rather insensitive to the size of the top-hat window, which may be in the range of  $R_{\text{th}} \sim 1\text{--}5 h^{-1} \text{ Mpc}$  for filaments and  $R_{\text{th}} \sim 2\text{--}8 h^{-1} \text{ Mpc}$  for walls. The lower limit depends on the mean interparticle separation, because there must be at least two particles inside the search window in order to displace the particle. The maximum value of  $R_{\text{th}}$  depends on how close the filaments and walls are to each other. If we would chose for a larger radius, particles from adjacent structures would be included which would invalidate the compression procedure.

We continue the compression algorithm recursively until the dispersion between consecutive compressed positions is less than  $0.25 h^{-1} \text{ Mpc}$ . It is important to note that in the compression algorithm, we only account for particles contained in the morphological population under consideration.

## APPENDIX B: LENGTH OF FILAMENTS

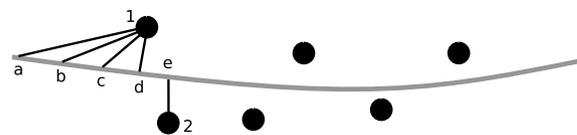
The determination of the length of filaments involves two steps. The first is the compression of the filaments, following the procedure outlined in Appendix A. The second step consists of fitting a polynomial to the particle or halo distribution along the filament.

By fitting to a polynomial, we smooth the small-scale variations that may remain following the compression procedure. If we would not include this step, and instead chose to add the segments of the minimum spanning tree defined by the particle distribution, we would end up with a filament whose size would be larger than is strictly representative.

As a compromise between simplicity and the ability to follow intricate and complex filament shapes, we chose to use polynomials of the third order. A visual inspection of several filaments assured us that the third-order polynomials are indeed sufficient for modelling even the most intricate filaments. They manage to follow each change in direction.

The positions  $\mathbf{r}_i = (x_i, y_i, z_i)$  of each of the particles/haloes  $i$  belonging to a filament are fitted to a polynomial of the form

$$\begin{aligned} x &= a_1 + b_1 t + c_1 t^2 + d_1 t^3, \\ y &= a_1 + b_1 t + c_1 t^2 + d_1 t^3, \\ z &= a_1 + b_1 t + c_1 t^2 + d_1 t^3, \end{aligned} \quad (\text{B1})$$



**Figure B1.** Identification of the extreme of a filament. The fitting curve (grey line) is closely sampled at points  $a, b, c, d, e$ . Haloes are represented as large black dots. The closest halo to points  $a, b, c$  and  $d$  is halo 1. At point  $e$ , the closest halo changes to halo 2 indicating that the fitting curve is ‘inside’ the filament.

where the parameter  $t$  is defined as the distance from an arbitrary location  $(x_0, y_0, z_0)$ :

$$t = \sqrt{(x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2}. \quad (\text{B2})$$

In practice we chose a set of values for  $(x_0, y_0, z_0)$ , usually the corners of the simulation box, and select the one that gives the best fit according to the criterion of having the smallest mean square difference:

$$\epsilon = \frac{1}{N} \chi^2. \quad (\text{B3})$$

If  $\epsilon$  turns out to be larger than a given threshold, whose value is determined by means of experimentation, we reject the fit.

One technical difficulty remaining is the determination of the extreme points of the filament for the fitting curve. This is a non-trivial task, and if not considered properly may lead to wrong length determinations. We follow a simple but effective method to identify the extremes, as follows.

- (i) The polynomial curve is closely sampled and distances to all the particles in the filament are computed starting from one extreme of the fitted curve.
- (ii) For each point in the fitting curve, the closest halo is identified.
- (iii) We identify the point in the fitting curve at which the identity of the closest halo changes. This indicates that the fitting curve is no longer ‘outside’ the filament but ‘inside’ it (see Fig. B1).
- (iv) We repeat the procedure for the second extreme of the fitting curve.

Following the previous steps, the fitted polynomial is used to compute the length of the filament.

This paper has been typeset from a  $\text{\LaTeX}$  file prepared by the author.