The 3D skeleton: tracing the filamentary structure of the Universe

T. Sousbie,[1,2]★ C. Pichon,[1,2]★ S. Colombi,[1]★ D. Novikov[3]★ and D. Pogosyan[4]★

[1]*Institut d'Astrophysique de Paris & UPMC, 98 bis boulevard Arago, 75014 Paris, France*
[2]*Centre de Recherche Astrophysique de Lyon, 9 avenue Charles André, 69561 Saint Genis Laval, France*
[3]*Astrophysics, Blackett Laboratory, Imperial College London, London SW7 2AZ*
[4]*Department of physics, University of Alberta, 11322-89 Avenue, Edmonton, Alberta, Canada T6G 2J1*

ABSTRACT

The skeleton formalism, which aims at extracting and quantifying the filamentary structure of our Universe, is generalized to 3D density fields. A numerical method for computing a local approximation of the skeleton is presented and validated here on Gaussian random fields. It involves solving equation $(\mathcal{H}\nabla\rho \times \nabla\rho) = 0$, where $\nabla\rho$ and $\mathcal{H}$ are the gradient and Hessian matrix of the field. This method traces well the filamentary structure in 3D fields such as those produced by numerical simulations of the dark matter distribution on large scales, and is insensitive to monotonic biasing.

Two of its characteristics, namely its length and differential length, are analysed for Gaussian random fields. Its differential length per unit normalized density contrast scales like the probability distribution function of the underlying density contrast times the total length times a quadratic Edgeworth correction involving the square of the spectral parameter. The total length-scales like the inverse square smoothing length, with a scaling factor given by $0.21\,(5.28 + n)$ where n is the power index of the underlying field. This dependency implies that the total length can be used to constrain the shape of the underlying power spectrum, hence the cosmology.

Possible applications of the skeleton to galaxy formation and cosmology are discussed. As an illustration, the orientation of the spin of dark haloes and the orientation of the flow near the skeleton is computed for cosmological dark matter simulations. The flow is laminar along the filaments, while spins of dark haloes within 500 kpc of the skeleton are preferentially orthogonal to the direction of the flow at a level of 25 per cent.

Key words: cosmology: theory – dark matter – large-scale structure of Universe.

1 INTRODUCTION

Recent galaxy surveys like 2dF (Colless et al. 2003) or Sloan Digital Sky Survey (SDSS) (Gott et al. 2005) emphasized the complexity of the matter distribution in the Universe which presents large-scale structures such as filaments, clusters or walls on the boundaries of low-density bubbles (voids). On the theoretical side, the currently favoured scenario suggests that the Universe evolved from Gaussian initial conditions to form the structures that are observed nowadays. Numerical simulations have successfully captured the main features of the observed filamentary distribution, both statistically and visually. The skeleton formalism in 2D was introduced in (Novikov, Colombi & Doré 2006) (NCD) and aims at making possible the extraction and analysis of these filamentary structures. This paper extends it to three dimensions in order to describe the Universe's large-scale matter distribution and its dynamical environment.

In the literature, various steps towards a quantitative description of the large structures have been suggested. Statistical tools such as correlation functions (e.g. Peebles 1980) and power spectra (e.g. Peacock 1998) have been widely used and have been successful in describing matter distribution and constraining cosmological parameter. Recently, fast algorithms have been designed for first and second order (Szapudi et al. 2005), as well as higher order statistics (counts in cells etc.) as in Croton et al. (2004) or Kulkarni et al. (2007). The Minkowski functionals have also been very popular since their first applications to matter density field topology (see e.g. Gott, Melott & Dickinson 1986). By studying the average properties of excursion sets, they allow the extraction of characteristic numbers that reflect the topology of the field such as the genus, computed from the mean curvature of isodensity surfaces (see Hamilton, Gott & Weinberg 1986). This approach is in fact very powerful and has been used to test various properties of matter distribution such as its Gaussianity in Doroshkevich (1970), Gott et al. (1986), Winitzki

★E-mail: sousbie@iap.fr (TS); pichon@iap.fr (CP); colombi@iap.fr (SC); novikov@astro.ox.ac.uk (DN); pogosyan@phys.ualberta.ca (DP)

& Kosowsky (1998) or more recently in Hikage, Komatsu & Matsubara (2006). They can also be used as so-called 'shape finders' (Sahni, Sathyaprakash & Shandarin 1998) and have been successfully applied to observed data sets (see e.g. Hikage et al. 2002; Sheth & Sahni 2005; James, Lewis & Colless 2007 for application to the SDSS and other galaxy distribution surveys). A large number of good reviews on the subject can be found (e.g. Melott 1990 or Kerscher 2000 and references therein).

These topological and statistical estimators analyse the distribution of observed galaxies globally and uniformly, and make little attempt at recovering the precise geometry of the matter distribution, i.e. they do not focus on specific regions (such as clumps, voids and filaments). Focusing on the identifiable regions of the Universe, the peak patches theory (Bond & Myers 1996a) attempts to describe cosmic structures formation through the identification of the collapse of the dense regions near the density peak and surrounding patches. In this framework, the evolution of patches hierarchy can be understood

from the measurement of only a few characteristics of the patches, while assuming that their flow does not depend on their internal non-linear dynamics.

This line of thought has been extended in the Cosmic Web paradigm (Bond, Kofman & Pogosyan 1996), which has emphasized that the large-scale spatial distribution of galaxy clusters and the filaments between them can be understood as mildly non-linear enhancements of the high-density peaks and filamentary ridges already present in the initial Gaussian density field. Recently, Hanami (2001) presented the so-called skeleton tree formalism: it analyses the process of hierarchical merging and extends the language of the peak patch through the analysis of the ridges of the density field in an abstract space corresponding to the usual three dimensions augmented by the smoothing length. The structure of voids in the large-scale dark matter distribution also has an extended history of theoretical modelling – see e.g. Hoffman & Shaham (1982), Icke (1984) or Bertschinger (1985) – while various void identifiers have been designed (see e.g. Platen, van de Weygaert & Jones 2007 and references therein).

One of the first attempts to develop an algorithm to detect and trace the filaments in the particle distribution has been the minimal spanning tree technique proposed by Doroshkevich in Doroshkevich (1970) and Barrow, Bhavsar & Sonoda (1985). Starting from a point distribution (a galaxy survey or a dark matter simulation), this method constructs the graph that connects all the dots with the property of never forming closed paths and being of minimal total lengths. Interesting statistical features can be extracted from it like the shape of the clusters or the length of the trunk (the longest path) and branches which are characteristic of the filamentarity of the distribution (Pearson & Coles 1995). More recently, other techniques with the same aim have been developed, such as Stoica et al. (2005), which uses a marked point process in order to recover the filament locations or Aragon-Calvo et al. (2007a) which provides an automatic segmentation of the different galactic distribution components using multiscale morphology filtering.

The 3D skeleton described in this paper focuses on the critical lines of a distribution, i.e. the set of lines joining the critical points in order to be able to compute the characteristic features of the underlying field (such as the total length of the filaments in a cosmological dark matter distribution). The skeleton provides a simple mathematical definition of the filaments of a density field based on Morse theory – see e.g. Milnor (1963), Colombi, Pogosyan & Souradeep (2000), Jost (2002) or Novikov, Colombi & Doré (2006) – and thus allows their extraction as well as their characterization.

Section 2 defines the local skeleton of large-scale structures. Section 3 introduces the numerical algorithm for constructing the local skeleton, and discusses its properties near the critical points (Appendix A gives a more detailed description of the algorithm). Section 4 investigates the evolution of its differential and total length as a function of the properties of the underlying field. Appendix C sketches the derivation of this differential length. Possible applications to cosmology and galaxy formation are discussed in Section 5, where two illustrations regarding the nature of the dark matter flow near the skeleton are given.

2 THE LOCAL SKELETON: THEORY

A comprehensive definition of the skeleton and how its local approximation in two dimensions is derived can be found in Novikov et al. (2006). To sum up, the so-called 'real' skeleton is by definition the subset of critical lines joining the saddle points of a field to its maxima while following the gradient's direction (while critical lines link *all* kinds of critical points together). It is easy to picture that applying this definition to a 2D field (an altitude map in a mountainous region for instance) allows the extraction of the ridges of that distribution. Although simple in appearance, this definition presents the drawback that it is in essence non-local: the presence of the skeleton in a given subregion may depend on the presence of a saddle point in a different subregion. In order to enforce locality, an approximation can in fact be derived using Taylor expansion in the vicinity of the critical points (i.e. local maxima and saddle points), leading to a second-order approximation of the skeleton: the local skeleton.

2.1 The 2D local skeleton

Defining the *local critical lines* as the set of points where the gradient of the field is an extremum along an isodensity contour, it can be shown (Novikov et al. 2006) that this set of points obeys the equation

$$
\mathcal{S} \equiv \frac{\partial \rho}{\partial r_1} \frac{\partial \rho}{\partial r_2} \left(\frac{\partial^2 \rho}{\partial r_1^2} - \frac{\partial^2 \rho}{\partial r_2^2} \right)
$$
$$
+ \frac{\partial^2 \rho}{\partial r_1 \partial r_2} \left(\left[\frac{\partial \rho}{\partial r_2} \right]^2 - \left[\frac{\partial \rho}{\partial r_1} \right]^2 \right) = 0, \tag{1}
$$

where r_1 and r_2 denote space coordinates and $\rho(r_1, r_2)$ is the density field. Equation (1) can be rewritten as

$$
\mathcal{S} = \det(\mathcal{H} \nabla \rho, \nabla \rho) = 0, \tag{2}
$$

where $\mathcal{H} \equiv \partial^2 \rho / \partial r_1 \partial r_2$ is the Hessian (second derivatives matrix) of the field. The solution to equation (2) can be interpreted mathematically as the set of points where the gradient of the field is an eigenvector of the Hessian (i.e. gradient and main curvature axis are aligned), which is clearly a local property of the field.

The local skeleton is defined as the subset of the local critical lines that is an approximation of the skeleton. Selecting this subset can be achieved by enforcing an additional condition: the gradient should be *minimal* [i.e. every point of the local skeleton of coordinates *r* should also be a local minimum of the isodensity contour at density $\rho(\boldsymbol{r})$]. That is, the second eigenvalue of the Hessian should be negative:

$$
\lambda_2 < 0, \quad \text{and} \quad \mathcal{H} \nabla \rho = \lambda_1 \nabla \rho, \tag{3}
$$

where λ_i are the eigenvalues of the Hessian and $\lambda_2 < \lambda_1$.

2.2 The 3D local skeleton

Let us now derive the generalization of the notion of the local skeleton to a 3D space. The philosophy is essentially the same but minor differences arise.

Starting from the same definition as in 2D, the local skeleton should be the set of points where the density is an extremum along an isodensity contour. Let (u, v) be a coordinate system along an isocontour $(r_1(u, v), r_2(u, v), r_3(u, v))$ where $r_i, i \in \{1..3\}$ are the three space coordinates. The definition of an isocontour implies that

$$\begin{cases} \dfrac{\partial \rho}{\partial r_1} \dfrac{\mathrm{d}r_1}{\mathrm{d}u} + \dfrac{\partial \rho}{\partial r_2} \dfrac{\mathrm{d}r_2}{\mathrm{d}u} + \dfrac{\partial \rho}{\partial r_3} \dfrac{\mathrm{d}r_3}{\mathrm{d}u} = 0 \\[2mm] \dfrac{\partial \rho}{\partial r_1} \dfrac{\mathrm{d}r_1}{\mathrm{d}v} + \dfrac{\partial \rho}{\partial r_2} \dfrac{\mathrm{d}r_2}{\mathrm{d}v} + \dfrac{\partial \rho}{\partial r_3} \dfrac{\mathrm{d}r_3}{\mathrm{d}v} = 0. \end{cases} \tag{4}$$

Moreover, as the gradient of the field ρ has to be an extremum:

$$\frac{\mathrm{d}}{\mathrm{d}u}(|\nabla \rho|^2) = 0 \quad \text{and} \quad \frac{\mathrm{d}}{\mathrm{d}v}(|\nabla \rho|^2) = 0. \tag{5}$$

Using equations (4) and (5), let us derive the equation of the local critical lines, which should only depend on the field and its first- and second-order spatial derivatives, similarly to equation (1). To do so, a coordinate system along the isocontour is needed but, as opposed to the 2D case, any coordinate system defined on an isocontour will be singular in some place as an isocontour is a closed surface. In order to avoid this problem, we choose to define three coordinates systems and swap from one to another when it becomes singular.

Defining three 1D coordinate systems s_i (see Fig. 1) so that for different values of s_i, one remains in the plane (r_j, r_k) where $i \neq j \neq k$ and $i, j, k \in \{1..3\}$. The coordinate system s_i is singular wherever $\nabla \rho$ is proportional to r_i. The 3D local critical lines satisfy equations (4) and (5) for $u \equiv s_i$ and $v \equiv s_j$ with $i \neq j$. For any s_i, these read

$$\frac{\mathrm{d}}{\mathrm{d}s_i}(|\nabla \rho|^2) = 0, \quad \text{and} \quad \frac{\partial \rho}{\partial r_1} \frac{\mathrm{d}r_1}{\mathrm{d}s_i} + \frac{\partial \rho}{\partial r_2} \frac{\mathrm{d}r_2}{\mathrm{d}s_i} + \frac{\partial \rho}{\partial r_3} \frac{\mathrm{d}r_3}{\mathrm{d}s_i} = 0. \tag{6}$$

Choosing $i \neq j \neq k \in \{1..3\}$, this system becomes after doing some algebra

$$\begin{aligned} \mathcal{S}_i &\equiv \frac{\partial^2 \rho}{\partial r_j \partial r_k} \left(\frac{\partial \rho}{\partial r_j}^2 - \frac{\partial \rho}{\partial r_k}^2 \right) \\ &+ \frac{\partial \rho}{\partial r_j} \frac{\partial \rho}{\partial r_k} \left(\frac{\partial^2 \rho}{\partial r_k^2} - \frac{\partial^2 \rho}{\partial r_j^2} \right) \\ &- \frac{\partial \rho}{\partial r_i} \left(\frac{\partial \rho}{\partial r_k} \frac{\partial^2 \rho}{\partial r_i \partial r_j} - \frac{\partial \rho}{\partial r_j} \frac{\partial^2 \rho}{\partial r_i \partial r_k} \right) = 0. \end{aligned} \tag{7}$$

One can check that equation (7) reduces to equation (1) in the 2D case, assuming that the field is constant in the third dimension [the first two terms of equation (7) are the same as in equation (1)]. The local critical lines are thus the set of points that satisfies

$$\boldsymbol{S} \equiv \begin{pmatrix} \mathcal{S}_i \\ \mathcal{S}_j \end{pmatrix} = \boldsymbol{0}, \ i \neq j \in \{1, 2, 3\}. \tag{8}$$

It is interesting to note that, as in the 2D case, equation (8) defines the local critical line as the set of points where the gradient of the density is an eigenvector of its Hessian matrix (the gradient and the principal curvature axis are collinear):

$$\mathcal{S} = (\mathcal{H} \cdot \nabla \rho \times \nabla \rho) = \boldsymbol{0}. \tag{9}$$

Once again, in order to require that the skeleton traces only the ridges of the distribution (i.e. the filaments in 3D), retrieving the subset of local critical lines that define the local skeleton can be achieved by
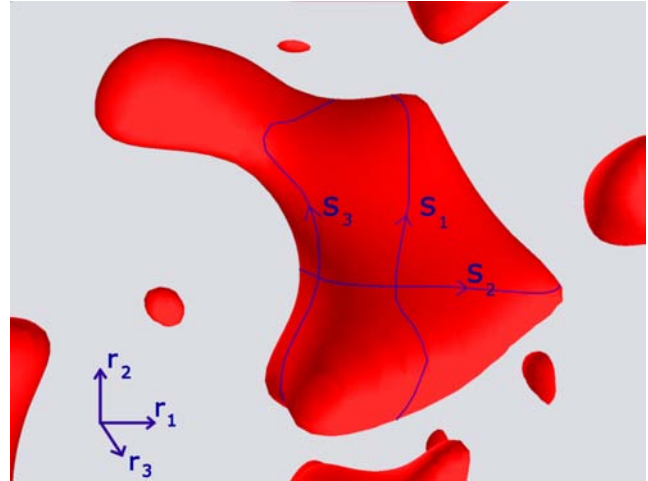

Figure 1. Definition of the coordinate system on an isocontour.

enforcing a negativity condition on the weakest eigenvalues of the Hessian:

$$\lambda_2 < 0, \quad \lambda_3 < 0, \quad \mathcal{H} \nabla \rho = \lambda_1 \nabla \rho. \tag{10}$$

That is, the local skeleton is the subset of the local critical where the norm of the 3D gradient is minimal along the 2D isodensity contours (as opposed to simply extremal). Note that from equation (8) it is straightforward to show that any monotonic function of the field will have exactly the same skeleton as the field itself.

3 IMPLEMENTATION AND FEATURES

3.1 Implementation

Equation (8) is at the basis of the numerical implementation of the local skeleton determination developed here. The details of the algorithm are described in Appendix A, while the optimal choice of resolution and smoothing is presented in Appendix B. All the computations were performed using a specially developed C package: SKELEX[1] (Skeleton Extractor). This package also includes a flexible OpenGL visualization tool that was used for making the figures in this paper.

Fig. 2 presents the skeleton obtained for a density field sampled from a numerical simulation of dark matter distribution on a $50 \, h^{-1}$ Mpc box with 512^3 particles using GADGET-2 (Springel 2005). The lighter colours represent denser regions and the blue skeleton appears to match quite well what one could identify as the filaments by eye. Note that the skeleton is both a tracer of the topology (it links a subset of the critical points) and the geometry of its underlying density. Hence it can be used to compare the geometrical and topological properties of various fields, e.g. the temperature and the dark matter distribution in hydrodynamical simulations. See also Fig. A3 for a graphical description of how the local skeleton is drawn.

3.2 The local skeleton branching properties

Let us now describe some global branching properties of the critical lines and the local skeleton. Important ingredients of the skeleton

---

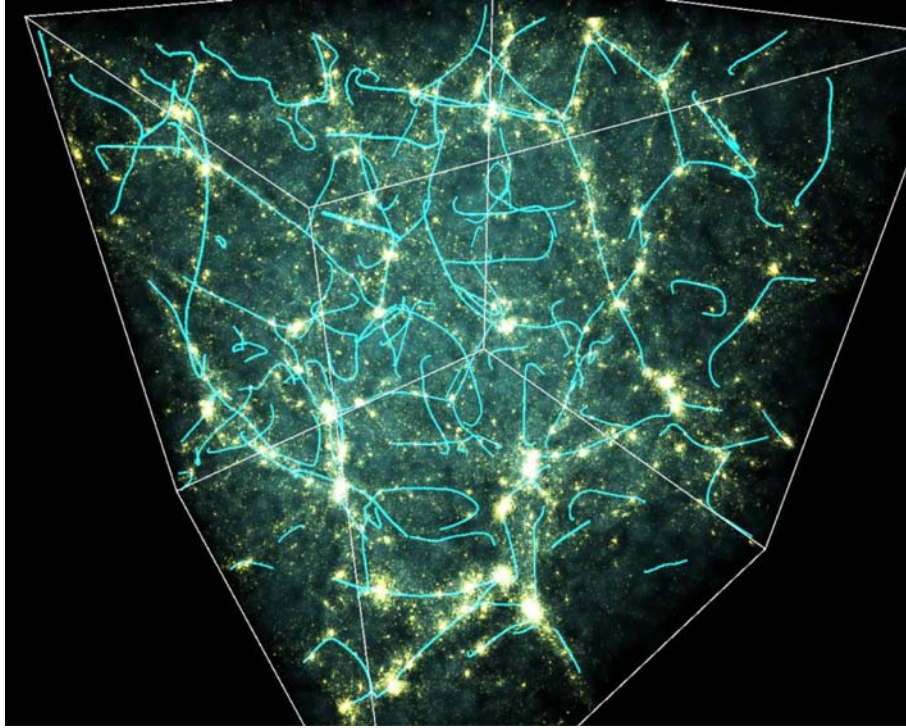[1] Available on request from the authors.

Figure 2. The final 3D skeleton derived from a 50-Mpc standard ΛCDM simulation run with sc gadget-2 using 512^3 particles. This result is obtained after post treating the skeleton using the method described in Appendix A.

are the extrema of the field. Indeed, the 'real skeleton' is defined as a set of critical lines that connect maxima to saddle points. Much of the topological behaviour of the skeleton is related to the distribution of such extremal points. For the *local* skeleton described this paper, the role of the extrema is similar but the whole set of critical lines encompass additional branches linking all kind of field extrema together.

Since the local skeleton is based on a local second-order approximation of the density field, ρ, its properties can be understood through the properties of the gradient $\nabla \rho$ and Hessian matrix $\mathcal{H}(\rho)$ only. The eigenvalues of $\mathcal{H}$ define the local curvature at any point, thus separating space into distinct regions depending on the sign of these eigenvalues λ_i. Within a 3D space, as by definition $\lambda_j < \lambda_i$ if $j > i$, there exist four of these regions. Let I be the number of negative eigenvalues, then the regions where I is equal to 0, 1, 2 and 3. This classification applies to critical points of the field in particular, where $\nabla \rho = 0$, the maxima ($I = 3$) and minima ($I = 0$) existing within local clumps and voids, respectively, while two types of saddle points can be distinguished: the filaments type saddle points (for $I = 2$) and the pancake type ones (for $I = 1$).

Fig. 3 illustrates a second-order approximation of the density field in the vicinity of the field extrema. The total set of critical lines form a fully connected path linking all the critical points together and exactly six branches pass through each of them in the direction of the three eigenvectors of the Hessian. Empirically, it is possible to picture the typical behaviour of the whole set of critical lines. Defining $E = \{0, 1, 2, 3\}$ and considering a given critical point where $I = n$, if $i < j < k \in E - \{n\}$, this critical point C_n is usually linked to three other pairs of critical points C_i, C_j and C_k (where $I = i, j$ and k, respectively) by critical lines aligned with eigenvectors associated with eigenvalues λ_1, λ_2 and λ_3, respectively, at point C_n. Most of the time, each of these branches connect to critical points C_i, C_j and C_k along the eigenvectors associated with eigenvalues λ_1, λ_2

(a) $0 > \lambda_1 \geq \lambda_2 \geq \lambda_3$ ($I = 0$)   (b) $\lambda_1 > 0 > \lambda_2 \geq \lambda_3$ ($I = 1$)

(c) $\lambda_1 \geq \lambda_2 > 0 > \lambda_3$ ($I = 2$)   (d) $\lambda_1 \geq \lambda_2 \geq \lambda_3 > 0$ ($I = 3$)
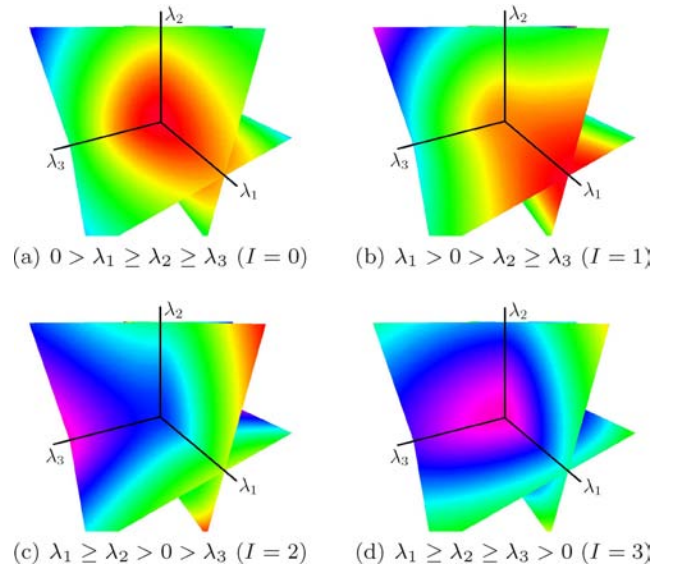
Figure 3. Illustration of a second-order approximation of the density field around a maximum ($I = 0$), filament ($I = 1$) and pancake ($I = 2$) saddle point and a minimum ($I = 3$). The colour stands for the density, ranging from purple in low-density regions to red in high-density regions. The axes are the eigenvectors of the Hessian, and give the direction of the six branches of the local critical lines going through these critical points (i.e. where the gradient of the field and the eigenvectors of H are aligned). The skeleton is the subset of these critical lines linking maxima (Fig. 3a) and filament saddle points (Fig. 3b), in the direction of the eigenvector associated with λ_1.

and λ_3, respectively, evaluated at points C_i, C_j and C_k, respectively. In this picture, the critical lines can be seen as a fully connected path linking all the different regions defined by the sign of the eigenvalues of $\mathcal{H}$.

The overdense filamentary structure correspond to the subset of the critical lines that constitute an approximation of the 'real' skeleton (i.e. the 'ridges' of the distribution). This part is the one which links maxima ($I = 3$) and filamentary saddle points ($I = 2$). The typical behaviour of such lines is the following: in the immediate vicinity of a non-degenerate maximum, two branches of the skeleton exist, stretching in the eigendirection that corresponds to λ_1. Following one of the branches, denoting as $\lambda_{||}$ an eigenvalue whose eigenvector is parallel to the skeleton and $\lambda_{\perp,1,2}$ as two eigenvalues associated to eigenvectors in the perpendicular directions. Near the maximum, $0 > \lambda_{||} = \lambda_1 > \lambda_{\perp,1} > \lambda_{\perp,2}$. As one follows a branch one probable outcome is the change of sign of $\lambda_{||}$, in which case the branch will typically end in a saddle point of a filamentary type along its λ_1 direction. There is always another branch that starts from this saddle point on the other side, thus this type of branches have a fully connected structure. However, another possible outcome is that one of the orthogonal eigenvalues changes faster than $\lambda_{||}$ as one moves away from the maximum and becomes positive before the saddle point is reached. In this case the branch of the local skeleton formally terminates, which however in reality often means that the skeleton splits at this point in two new branches.

Such branching of the skeleton is especially frequent near the maxima of the field, where it accounts for how multiple filamentary sections can end up in a single dark matter halo. Studying how skeleton segments merge is relevant for questions such as the multipole structure of matter inflow on to dark haloes (Aubert, Pichon & Colombi 2004; Pichon & Aubert 2006). This property of skeleton segments to end outside of the critical points is specific to the local definition of the skeleton, in contrast to the 'real' skeleton whose segments are always connected on both ends.

4 THE SKELETON LENGTH FOR SCALE-FREE GAUSSIAN RANDOM FIELDS

Before considering general cosmological density fields, the local skeleton of scale-free Gaussian random fields ρ with null average value $\langle \rho \rangle = 0$ will be investigated. For convenience, it is useful to define some spectral parameters that depend on the spectral index n and on the smoothing length. In the statistical description of the skeleton of a random density field (Appendix C), the following spectral parameters appear to play a role:

$$\sigma_0^2 = \langle \rho^2 \rangle, \tag{11}$$

$$\sigma_1^2 = \langle (\nabla \rho)^2 \rangle, \tag{12}$$

$$\sigma_2^2 = \langle (\Delta \rho)^2 \rangle, \tag{13}$$

$$\sigma_3^2 = \langle (\nabla \Delta \rho)^2 \rangle. \tag{14}$$

This introduces three linear scales into the skeleton theory

$$R_0 = \frac{\sigma_0}{\sigma_1}, \quad R_* = \frac{\sigma_1}{\sigma_2}, \quad \tilde{R} = \frac{\sigma_2}{\sigma_3}, \tag{15}$$

where the first two have a well-known meaning of typical separation between zero crossing of the field R_0 and mean distance between extrema, R_* (Bardeen et al. 1986), while the third one, $\tilde{R}$ is, by analogy, the typical distance between the inflection points.

Out of three scales two dimensionless ratios may be constructed that are intrinsic parameters of the theory

$$\gamma \equiv \frac{R_*}{R_0} = \frac{\sigma_1^2}{\sigma_0 \sigma_2}, \quad \tilde{\gamma} \equiv \frac{\tilde{R}}{R_*} = \frac{\sigma_2^2}{\sigma_3 \sigma_1}, \tag{16}$$
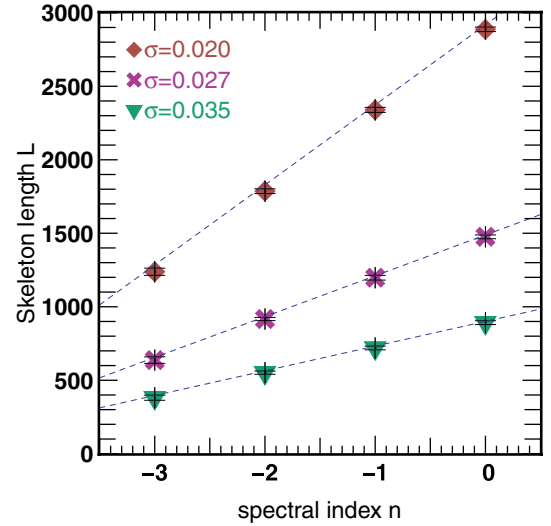
Figure 4. Total length L of the skeleton per unit box size for different smoothing lengths $\sigma = 0.020, 0.027, 0.035$; measured over 25 realizations of Gaussian random fields as a function of the spectral index n. While L depends linearly on the spectral index n, it grows as a power of σ. The dotted lines represent the fits obtained using the function: $L \approx 0.21(n + 5.28)\sigma^{-2}$.

where γ says how frequent encountering a maximum between two zero crossings of the field is, while $\tilde{\gamma}$ describes, on average, how many inflection points are between two extrema. For Gaussian fields, these parameters can be easily calculated from the power spectrum. Both γ and $\tilde{\gamma}$ range from 0 up to 1. For reference, for the power-law spectra with index $n > -3$, smoothed at small scales with a Gaussian window,

$$\gamma = \sqrt{\frac{n+3}{n+5}}, \quad \tilde{\gamma} = \sqrt{\frac{n+5}{n+7}}. \tag{17}$$

Note that cosmologically relevant density power spectra have $n > -3$ and thus, while γ can attain low values, $\tilde{\gamma}$ are always close to unity.[2]

Appendix C introduces a statistical description of the skeleton for the Gaussian and non-Gaussian random field. This section presents the numerical measurements of the properties of the skeleton for scale-free Gaussian fields.

The first quantity of interest is the total length of the skeleton, L_{tot}. In the context of cosmology, L_{tot} can be linked to the total length of the filaments linking clusters together and in that sense reflects the history of matter accretion as well as the initial distribution of matter (which is supposed to be similar to a Gaussian random field with a scale-dependent effective spectral index similar to the ones considered here). Fig. 4 presents the result of the measurement of the total length L_{tot} of the skeleton per unit box size as a function of the spectral index and for different smoothing lengths σ (within the range of validity of the algorithm as described in Appendix B). These measurements are carried over 25 realizations of scale-free 256^3 Gaussian random fields as a function of the spectral index n. The sensitivity of the skeleton to the value of the spectral index is clear on this plot and, if L_{tot} appears to be a linear function of the

---

[2] Cosmological density fields, therefore, have of order one inflection point per extremum, unlike, e.g. a mountain range, where one encounters many inflection points on a way from a mountain top to the bottom.
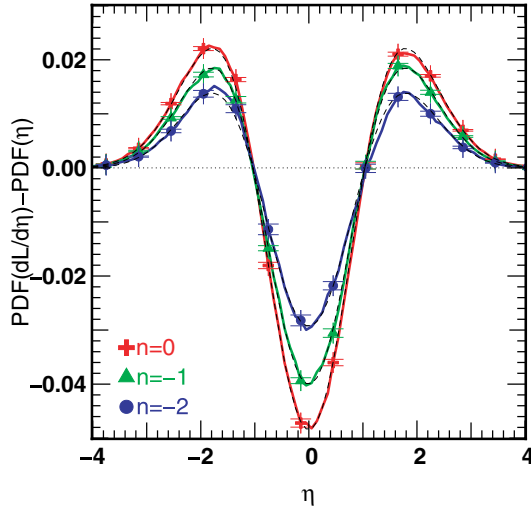
Figure 5. Difference between the PDF of the density field and the normalized differential length of the skeleton $\mathrm{d}L/\mathrm{d}\eta$ as a function of the density contrast $\eta = \rho/\sigma_0$. Each curve represents the average value and variance of the measured value of $\mathrm{d}L/\mathrm{d}\eta$ over 25 different realizations of scale-free Gaussian fields, for different values of the spectral index $n = 0, -1, -2$. The dotted curves represent the estimation obtained by fitting data using equation (19) (see Table 1 for values of the parameters).

spectral index, it is also clear that it grows as a power law of the smoothing length. The dotted lines on Fig. 4 shows the result of such a fit of the data and seems to work very well. A very good approximation of L_{tot} per unit box size is thus given by the function:

$$L_{\mathrm{tot}} = 0.21(n + 5.28)\sigma^{-2.00}. \tag{18}$$

As expected, the exponent of σ is measured to be exactly 2. It can be proved with a simple argument that this should be the case for scale-free Gaussian fields. In fact, for such fields, computing the skeleton over a grid of volume l^3 and smoothed on a scale σ is equivalent to computing the skeleton on a grid of volume $(\alpha\, l)^3$ while smoothing on a scale $(\alpha\sigma)$ and rescaling the result by a factor of $1/\alpha$. Because of the scale invariance, we also have $L(\sigma) = \alpha^{-3} L(\alpha\sigma)$ and so $L(\sigma) \propto \sigma/\sigma^3 = \sigma^{-2}$.

Interestingly, the dependence on the spectral index n is close to $n + 5$ which argues for filaments being relatively straight between extrema, see Appendix C. A visual examination of the filaments confirms this picture.

Now consider the differential length of the skeleton, $\mathrm{d}L/\mathrm{d}\eta(\eta)$ where $\eta \equiv \rho/\sigma_0$ is the normalized density contrast. This quantity represents the expected length of skeleton that can be measured in a given distribution between density contrasts η and $\eta + \mathrm{d}\eta$. Fig. 5 shows the normalized function $\mathrm{d}L/\mathrm{d}\eta(\eta)$ as a function of the normalized density contrast η from which was subtracted the probability distribution function (PDF) of the field (which, within the range of sampling and finite volume effects approximations, is a Gaussian function). These values were also averaged over 25 realizations of Gaussian fields with spectral index $n = 0, -1, -2$ sampled on 256^3 pixel grids and for a smoothing length $\sigma = 0.027$. This value was chosen as a compromise between finite volume effect and differentiability of the field on a grid discussed in Appendix B. Considering the error bars, it is clear that the value of $\mathrm{d}L/\mathrm{d}\eta(\eta)$ is directly linked to the spectral index n.

It is shown in Appendix C that $\mathrm{d}L/\mathrm{d}\eta(\eta)$ can be written using an Edgeworth expansion (see also Novikov et al. 2005 for the cor-

Table 1. Measured values of the first three non-null terms in the Edgeworth expansion, equation (19), for three different values of the spectral index $n = 0, -1, -2$. These results are obtained by fitting equation (19) on the data presented in Fig. 5 on which the dotted lines represent the fitted function. The measurements show very good agreement, whatever the value of n.

	C_2	C_4	C_6
$n = 0$	0.219	0.006	-0.001
$n = -1$	0.212	0.002	-0.002
$n = -2$	0.206	-0.005	-0.008
	0.21 ± 0.005	0.001 ± 0.005	-0.004 ± 0.003

responding proof and fit in 2D):

$$\frac{\mathrm{d}L}{\mathrm{d}\eta}(\eta) = \frac{L_{\mathrm{tot}}}{\sqrt{2\pi}} \exp(-\eta^2/2) \left(\sum_{n \geqslant 0} C_{2n} \gamma^{2n} H_{2n}\left(\eta/\sqrt{2}\right) \right), \tag{19}$$

where L_{tot} is the total length of the skeleton, $C_0 = 1$ and H_{2n} are Hermite polynomials (using the Probabilist's convention). Fig. 4 demonstrates that this expansion also works very well in the 3D case. Remarkably, equation (19) does not depend on $\tilde{\gamma}$ which again argues for the picture of a stiff behaviour of the skeleton for cosmological scale-invariant density fields (see Appendix C). Table 1 presents the values of the first three coefficients C_{2n} obtained by fitting the measurements presented in Fig. 5 (the dotted line of Fig. 5 are the result of these fits). Not only does equation (19) allows a very good fit of the measured data, but it also appears that only the first-order term is non-null and the differential length of the skeleton of a Gaussian random field with spectral parameter γ is thus given by

$$\frac{\mathrm{d}L}{\mathrm{d}\eta}(\eta) = \frac{L_{\mathrm{tot}}}{\sqrt{2\pi}} \exp(-\eta^2/2) \left[1 + 0.21\gamma^2(\eta^2 - 1) \right]. \tag{20}$$

The only non-null coefficients in the expansion are thus $C_0 = 1$ and $C_2 = 0.21$, to be contrasted to $C_2 = 0.17$ in the 2D case. Equation (20) can be used as a test of non-Gaussianity like any other topological estimator, such as the genus, the PDF, etc. as discussed in Novikov et al. (2006), since departure from the shape of equation (20) must appear when the skeleton's differential length is computed while the underlying field is not Gaussian.[3]

For the matter distribution in the Universe, the filaments are overdense regions along which matter flows. In that sense, they are less subject to numerical or observational noise and contain most of the information about the underlying matter distribution. The skeleton length can thus be seen as a method for measuring the power spectrum which naturally weights information in different regions according to their importance.

5 ILLUSTRATION: DYNAMICAL ENVIRONMENT OF FILAMENTS

Drawing the skeleton allows us to pin down the nature of the flow around the filaments. Indeed one may roughly define three dynamically distinct regions in large-scale structures: voids, clusters and filaments. The first two have been investigated in some detail. The filaments represent a fairly unexplored venue. Beyond the kinematics (velocity distribution, spin, etc.), the photometric and spectroscopic properties of galaxies (colour, age, metallicity, etc.), their morphology (ellipticals versus spirals, Gini number, asymmetry) or

---

[3] Of course, given the properties of the skeleton, this will not apply if the non-Gaussianity involves only a (monotonic) bias.
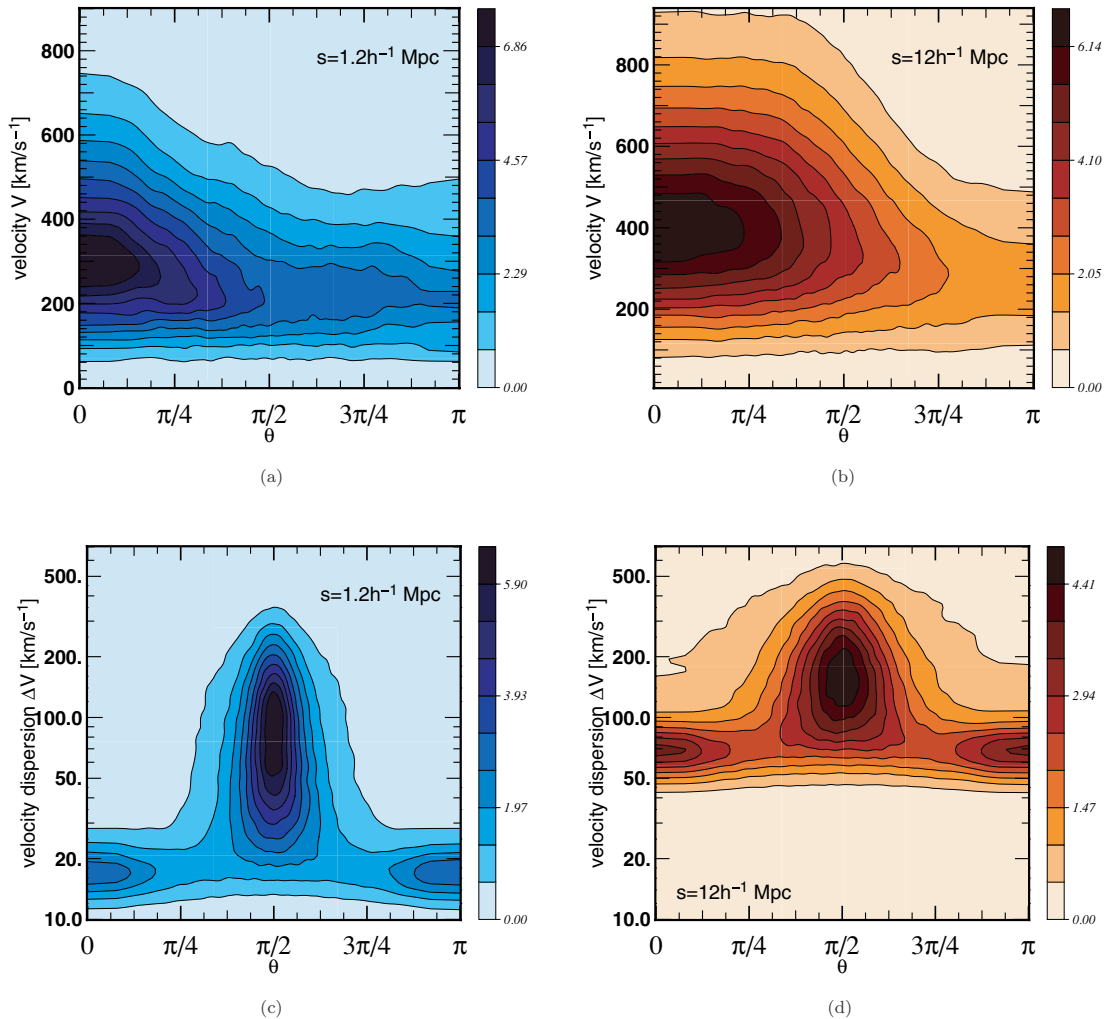
Figure 6. Top panels: PDF of the velocity field V of the dark matter along the skeleton as a function of its angle θ with the skeleton and its norm. The measurements were achieved on a $100\,h^{-1}$ Mpc and $1000\,h^{-1}$ Mpc dark matter simulation featuring 512^3 particles and a standard ΛCDM model, smoothed over a scale $s = 1.2$ and $12\,h^{-1}$ Mpc (left- and right-hand panels, respectively). The skeleton is oriented in the direction of increasing density. Dark matter appears to be flowing along the filaments in the direction of higher density regions (i.e. haloes). Bottom panels: PDF of main eigenvector of the velocity dispersion tensor ΔV_{ij} as a function of its angle θ with the skeleton and its eigenvalue amplitude. The peak of the PDF corresponds to high velocity dispersion orthogonal to the filaments, which is coherent with the picture of dark matter being accreted orthogonally by the filaments before flowing along them. Note the increase in velocity dispersion with scale (left- and right-hand panels) as well as the larger angular dispersion in the dark matter flow. This trend is also found while considering the same simulation at higher z.

the IGM (gas temperature, WHIM detection, fraction of gas/metals in the filaments, etc.), could also be investigated as a function of the distance to, and along the filaments.

In this section, two examples simply illustrate how the skeleton can be used to explore the environment of filaments in cosmological simulations.

5.1 Dark matter flow near the skeleton

Fig. 6 displays PDF of different characteristics of the dark matter flow along the skeleton. In order to understand the correlations between the filaments and the velocity field, we computed the PDF of its angle relative to the skeleton as a function of its intensity (top panels), and the PDF of the angle between its largest eigenvector and the skeleton as a function of the norm of the corresponding eigenvalue (bottom panels). These measurements were achieved by first sampling the field characteristics on a grid, averaging particles velocities $V \equiv \langle v \rangle$ and dispersion tensor, $\Delta V_{ij}^2 \equiv \langle (v_i - \langle v_i \rangle)(v_j$

$- \langle v_j \rangle) \rangle$ over each cell, and then computing for each segment the distance-weighted average of their PDF. Left- and right-hand panels yield the resulting PDF computed in a $100\,h^{-1}$ and $1000\,h^{-1}$ Mpc dark matter standard Lambda cold dark matter (ΛCDM) model simulation, respectively, at redshift $z = 0$ and using 512^3 particles. In both cases, the density and velocity fields where sampled on a 512^3 pixels grid and smoothed over $\sigma_p = 6$ pixels (i.e. $s = 1.2\,h^{-1}$ Mpc and $s = 12\,h^{-1}$ Mpc, respectively). The skeleton segments being oriented in the direction of increasing density, an angle of $\theta = 0$ means that dark matter is flowing along the filament in the direction of higher density regions.

The flows appears to be laminar and its amplitude increases with scale: this is expected since on larger scales the clusters are more massive, the potential difference is larger, hence the flow towards them is faster. Most dark matter particles have a mean velocity of about 300 (respectively 400) km s^{-1} along the filament and a dispersion of about 100 (respectively 150) km s^{-1} orthogonal to the filaments for the two scales considered here. The angular spread
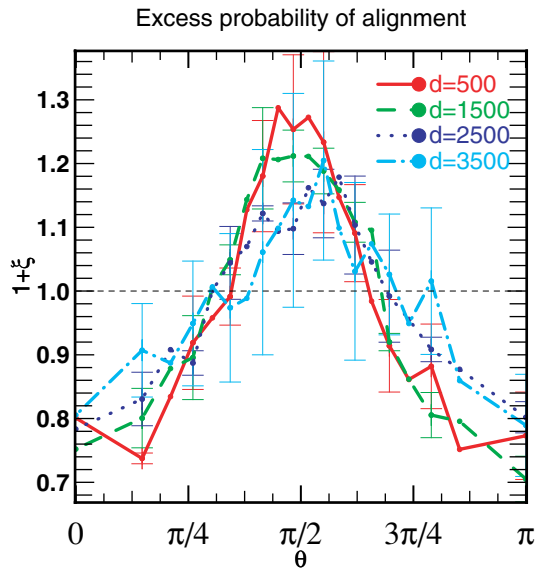
Excess probability of alignment

Figure 7. Excess probability of spin alignment with the local skeleton computed from the average of three 512^3 $50 h^{-1}$ Mpc ΛCDM simulations at different distances: $d \in [0, 500]$, $[500, 1500]$, $[1500, 2500]$ and $[2500, 3500]$ h^{-1} kpc from the closest skeleton segments. This figure demonstrates that on average the spin of dark matter haloes tends to be orthogonal to the local filaments at a level of 25 per cent for distances shorter than 500 kpc. The simulation is analysed at redshift zero.

(panels 6a and b) also increases with scale, from about 30° to about 45°, reflecting the larger internal heat of the filament, also seen in Figs 6(c) and (d).

The qualitative shape of this PDF may be explained by the advection of new haloes on to the 'highways' corresponding to the mean flow. The first eigenvector of the dispersion tensor is on average clearly orthogonal to the filament, reflecting the velocity of dark matter falling on to the filaments. Note that the distribution is decreasing monotonously with θ in Fig. 6(a): some dark matter particles statistically even move downhill, and their relative fraction decreases with scale. The filaments are collecting matter away from the underdense regions. Smaller filaments empty smaller voids, which tend to get depleted earlier than larger ones; hence this may explain why the flow becomes more orderly at smaller scale as accretion diminishes.

Note that the redshift evolution (not shown here) of this distribution follows closely its scale evolution, the $z = 15$ PDF over $100 h^{-1}$ Mpc resembling the $z = 0$ PDF over $1000 h^{-1}$ Mpc (Sousbie 2006).

The detailed nature of the flow should eventually be investigated in a smoothing scale-independent manner, in order to derive universal features which would only depend on the cosmology and the initial power spectrum. Its evolution with redshift or with the cosmology should also be systematically analysed.

5.2 Dark matter spin-skeleton connection

The geometric orientation of the spin of dark matter haloes corresponds to another feature of the large-scale structure which can be characterized using the skeleton. The spin of dark haloes was computed using the classical friend-of-friend algorithm with 0.2 times the interparticle distance as linking length and retaining only haloes containing more than 100 particles. Fig. 7 displays the excess prob-

ability of alignment of the haloes' spins with the closest skeleton segment for different distances $[0, 0.5]$, $[0.5, 1.5]$, $[1.5, 2.5]$ and $[2.5, 3.5]$ h^{-1} Mpc. This probability reaches 25 per cent for an angle $\theta = \pi/2$ between the spin and the skeleton: the spin of dark matter haloes is preferentially orthogonal to the filament they belong to. This trend accounts for the fact that the filaments are the locus of laminar flow where haloes coalesce along the direction of the filaments parallel to the mean flow, hence acquiring momentum orthogonal to the flow, as observed in Aubert et al. (2004) and Aragon-Calvo et al. (2007b).

6 CONCLUSION AND PERSPECTIVES

The 3D skeleton formalism is a well-defined framework for studying the filamentary structure of a distribution. The 'real' skeleton is defined as the set of critical lines joining saddle points to maxima of the field along the gradient. A local approximation of it was introduced in Section 2 along with a numerical method allowing a fast retrieval of the locus of the filaments from a sampled field (see also Appendix A). This method involves computing the null isodensity surfaces of each component of a function $\mathcal{S} = (\mathcal{H} \cdot \nabla\rho \times \nabla\rho)$ of the gradient, $\nabla\rho$, and Hessian matrix, $\mathcal{H}$ of this field.

The ability to localize and characterize the filamentary structure of matter distribution in the Universe opens the prospect of many applications for the skeleton as discussed in Sections 4 and 5. It has been shown in Section 4 that for a Gaussian random field, the total length of the skeleton per unit volume depended only of the chosen smoothing length σ and spectral index n, with a specific functional from which was both fitted from simulations and motivated in Appendix C. In this sense, the local skeleton provides a direct measurement of the local shape of the power spectrum, $P(k)$, on various scales depending on the smoothing applied to the underlying field. Though there exist other ways to measure the power spectrum of a given distribution, the skeleton length is promising as it relies only on the filamentary structure of the distribution. A forthcoming paper will investigate in more details the expected scalings on the shape of the power spectrum. The analysis of the length of the skeleton of the galaxy distribution in the SDSS as a measurement of cosmological parameter Ω_m can be found in Sousbie et al. (2007). This paper addresses the issue of implementing the present algorithm on observational data sets and mock catalogues. In particular, it is shown there that the effect of redshift distortion is well accounted for by comparing data from large-scale surveys to such catalogues.

The skeleton may also be used as an isotropy probe. It corresponds in fact to a good candidate for the Alcock–Paczynski (Alcock & Paczynski 1979) test, since the apparent longitudinal to transverse length of skeleton segments should directly constrain the curvature of space in a manner which is bias independent. This test will be presented in a forthcoming paper.

It was demonstrated in Section 4 that the dark matter flow in the vicinity of filaments was dominantly laminar along the filaments and shows signs of orthogonal accretion corresponding to the infall of dark matter collected from the voids. It was also shown that the spin of dark matter haloes were preferentially orthogonal to the filament's direction, a feature which can be understood as a consequence of merger events taking place along these filaments. A clear virtue of the local skeleton is that since it relies on a local expansion of the field, it can deal with truncated/masked fields, segmented or vanishing ridges or isolated structures. Note finally

that the fit, equation (20), opens the prospect of using the local skeleton to estimate the bias in observed surveys. The idea is to compute the PDF of galaxies on the one hand, which depends on the mass to light ratio of the sample, and the differential length (equation 20) on the other hand. Since the former depends on the bias, whereas the later does not, comparing the two should give an estimate of the bias. On the other hand, the local formulation of the skeleton presents some limitations. Mainly, it is not fully connected: it has by construction (since it is drawn from a second-order Taylor expansion of the field) only two segments per maxima whereas full connection would require three or more. A consequence is that it cannot represent merging filaments.

One could also use the curvature and torsion of filaments as cosmological probes, since the acceleration of the Universe induced by the cosmological constant is likely to straighten the filaments, though the fact that the local skeleton has only two segments near its maxima (the other segments must branch out) is likely to introduce some artefacts. The topology and geometry of the skeleton near the density peaks and the redshift evolution of the skeleton of the large-scale structures may prove of interest, for instance to study the frequency of reconnection, though again the local skeleton is not ideal in this respect. It would also be interesting to construct the skeleton in higher dimensions, for instance in space–time, to trace the events lines, but again connection is critical. In a forthcoming paper, an alternative algorithm for the identification of the skeleton, loosely based on a least action formulation, will be presented. It is complementary to the solution presented in this paper and will allow us to tackle those points for which the local skeleton is less efficient. Finally, the 3D skeleton algorithm could possibly be applied to other fields of research, such as neurology, in order to trace the neural network.

ACKNOWLEDGMENTS

REFERENCES

Alcock C., Paczynski B., 1979, Nat, 281, 358
Aragon-Calvo M. A., Jones B. J. T., van de Weygaert R. M. J., der Hulst V., 2007a, preprint (arXiv:0705.2072)
Aragon-Calvo M. A., van de Weygaert R., Jones B. J. T., van der Hulst J. M., 2007b, ApJ, 655, L5
Aubert D., Pichon C., 2007, MNRAS, 374, 877
Aubert D., Pichon C., Colombi S., 2004, MNRAS, 352, 376A
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
Bertschinger E., 1985, ApJS, 58, 1
Bond J. R., Kofman L. A., Pogosyan D., 1996, Nat, 380, 63
Colless M. et al., 2003, preprint (astro-ph/0306581)
Colombi S., Pogosyan D., Souradeep T., 2000, Phys. Rev. Lett., 85, 5515
Croton D. J. et al., 2004, MNRAS, 352, 1232
Doroshkevich A. G., 1970, Astrofizica, 6, 581 (Astrophys., 6, 320)
Doroshkevich A. G., Tucker D. L., Lin H., Turchaninov V., Fong R., 2001, MNRAS, 322, 369
Gott J. R., III, Melott A. L., Dickinson M., 1986, ApJ, 306, 341
Gott J. R., III, Jurić M., Schlegel D., Hoyle F., Vogeley M., Tegmark M., Bahcall N., Brinkmann J., 2005, ApJ, 624, 463
Hamilton A. J. S., Gott J. R., III, Weinberg D., 1986, ApJ, 309, 1
Hanami H., 2001, MNRAS, 327, 721
Hikage C. et al., 2002, PASJ, 54, 707
Hikage C., Komatsu E., Matsubara T., 2006, ApJ, 653, 11
Hoffman Y., Shaham J., 1982, ApJ, 262, L23
Icke V., 1984, MNRAS, 206, 1
James J. B., Lewis, G. F., Colless M., 2007, MNRAS, 375, 128
Jost J., 2002, Riemannian Geometry and Geometric Analysis, 3rd edn. Springer-Verlag, Berlin
Kerscher M., 2000, Lecture Notes in Physics, 554, 36
Kulkarni G. V., Nichol R. C., Sheth R. K., Seo H.-J., Eisenstein D. J., Gray A., 2007, MNRAS, 378, 1196
Lorensen W., Harvey E., 1987, Cline. Marching Cubes: A High Resolution 3D Surface Construction Algorithm. Computer Graphics (SIGGRAPH 87 Proceedings), p. 163
Melott A. L., 1990, Phys. Rep., 193, 1
Milnor J., 1963, Morse Theory. Princeton Univ. Press, Princeton, NJ
Novikov D., Colombi S., Doré O., 2006, MNRAS, 366, 1201
Peacock J. A., 1998, Cosmological Physics. Cambridge Univ. Press, Cambridge
Pearson R. C., Coles P., 1995, MNRAS, 272, 231
Peebles P. J. E., 1980, The Large-Scale Structure of the Universe. Princeton Univ. Press, Princeton, NJ
Platen E., van de Weygaert R., Jones B. J. T., 2007, MNRAS, 380, 551
Sahni V., Sathyaprakash B. S., Shandarin S. F., 1998, ApJ, 495, L5
Sheth J. V., Sahni V., 2005, preprint (astro-ph/0502105)
Springel V., 2005, MNRAS, 364, 1105
Sousbie T., 2006, PhD thesis, http://hal-insu.archives-ouvertes.fr/
Sousbie T., Pichon C., Courtois H., Colombi S., Novikov D., 2007, ApJL, in press, preprint (astro-ph/0602628)
Stoica R. S., Martínez V. J., Mateu J., Saar E., 2005, A&A, 434, 423
Szapudi I., Pan, J., Prunet S., Budavári T., 2005, ApJ, 631, L1
Winitzki S., Kosowsky A., 1998, New Astron., 3, 75

APPENDIX A: NUMERICAL IMPLEMENTATION

All the computations were performed using a specially developed C package: SKELEX[4] (Skeleton Extractor). This package also includes a flexible OpenGL visualization tool that was used for making the figures in this paper.

The first step before computing the skeleton requires obtaining a density field from a discrete point-like distribution. This is achieved by smoothing appropriately the density field on a grid so that it is not singular (i.e. is sufficiently differentiable) but still contains all the topological information. The density field is computed using cloud-in-cell interpolation and convolving the result with Gaussian windows of different widths. As was shown in section B, the grid size and smoothing length are decisive parameters. It is then necessary to compute first and second derivatives of the field on the grid, which can be done using finite difference or Fourier transform method. Choosing one method or the other does not seem to have any influence on the resulting skeleton if the field is smooth enough (which is anyway a necessary condition).

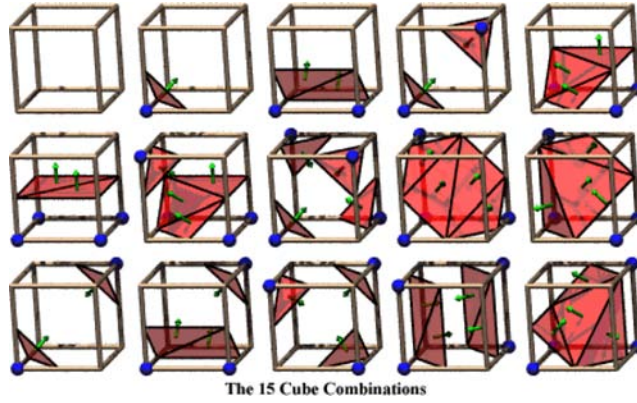[4] Available on request from the authors.

Figure A1. Illustration of the different possible configurations of a grid cell used for marching cube algorithm. Given a field f and isocontour $f = 0$, a blue ball represents a vertex where $f > 0$. It is then easy to build the isocontour by linearly interpolating the value of f along the edges. This picture was borrowed from James Sharman's web site, http://www.exaflop.org/docs/marchcubes/ind.html.
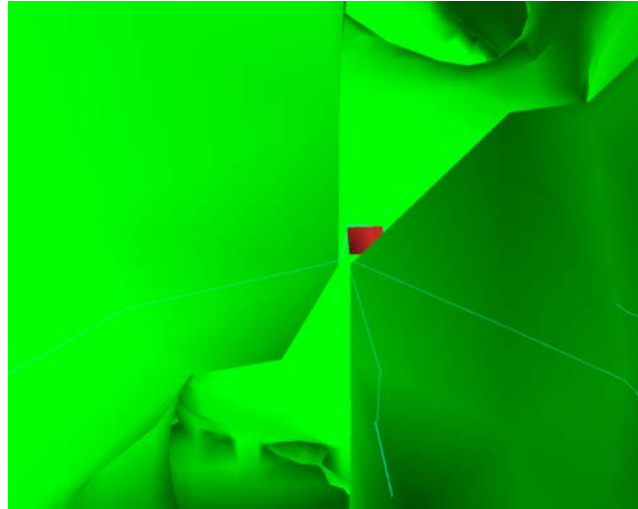

Figure A2. Illustration of a drawback of marching cubes algorithm. The green surface is an isosurface solution of equation (7) and the light blue line is the resulting skeleton. The red diamond represent a field maximum. It is clear on this picture that the algorithm misses the part close to that maximum, thus creating a spurious hole in the skeleton.

The next step involves solving the system of equations (8); the solution of this system corresponding to the intersection of two of the three solutions of equations (7). This is done by computing the 3D meshes of the 2D surfaces that are solution to these equations: the skeleton is at the intersection of two of them, depending on the value of the gradient at the point considered. Solving equation (7) is equivalent to finding the null isocontour of field $\mathcal{S}_i$, which can be done using the marching cube algorithm (Lorensen & Harvey 1987). The basic idea is to consider every cell of the grid as an individual cube. One can then compute the value of every $\mathcal{S}_i$ for the eight vertices and it is easy to check whether the isosurface intersects the cube or not. In fact, every vertex is above or below a threshold value (in this case 0), which gives a total of $2^8 = 256$ types of intersections (only 15 of them being intrinsically different) that can be pre-computed as illustrated in Fig. A1. The exact positions of the intersections are computed using quadratic interpolation. This yields the position of the intersections of the grid and the isocontour, and defines triangles that smartly link those intersection vertices: one can then reconstruct a very good approximation of what the isocontour is.

Which surfaces should be used for each cell is decided by computing $d_k = \det(\boldsymbol{r}_i, \boldsymbol{r}_j, \nabla \rho), i \neq j \neq k \in \{1, 2, 3\}$ and selecting only the two $\mathcal{S}_k$ for which d_k is maximal. This gives two surfaces defined by triangles whose intersection can be efficiently computed: it amounts to computing the intersection of triangle pairs only. It is then straightforward to compute the eigenvalue of the Hessian for every segment and keep or reject them depending on the previously defined criteria (equation 10) in order to draw the *local* skeleton. The exact same method was used for efficiently and consistently finding the extrema and saddle points of the field. Indeed, if one defines three fields $f_i = \partial \rho / \partial r_i$, those critical points are the intersections of the three isocontour surfaces $f_i = 0$. One can then decide if a critical point is a maximum, minimum or saddle point by checking the value of eigenvalues of the Hessian (i.e. the curvature). Although marching cubes algorithms are very efficient for computing isodensity contour, they present some drawbacks for ambiguous configurations. Indeed, as illustrated on Fig. A2, some configurations are degenerate and one cannot decide where the isosurface should pass. This problem happens most of the time around critical points where the value of the field can go above and below the threshold within one cell. It induces the loss of small skeleton segments.

In order to obtain a smooth skeleton that does not present holes and to retrieve the connectivity information (i.e. to be able to follow the skeleton from one point to another), a three steps post-processing is applied. Here the algorithm is based on a weighted marking system to
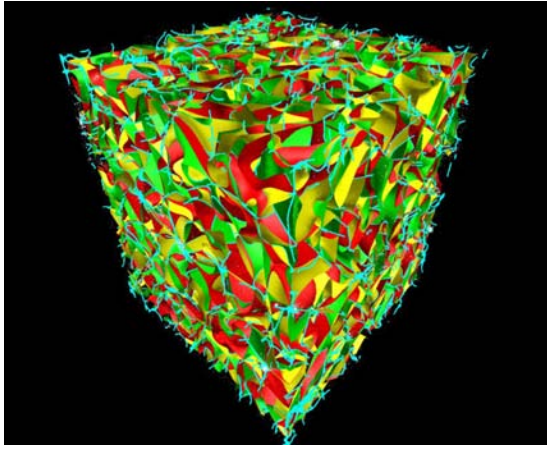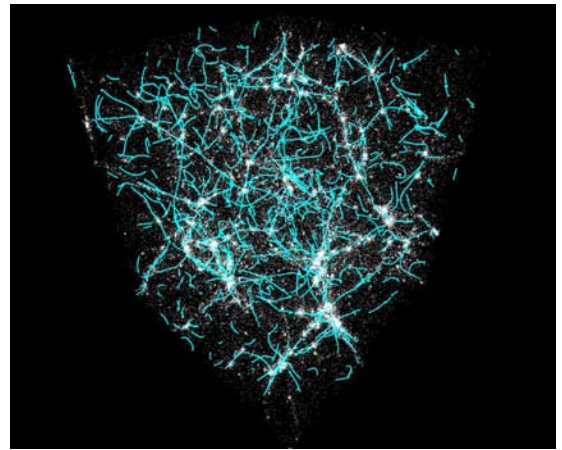
(a) The three fields isosurfaces the intersection of which constitutes the critical lines
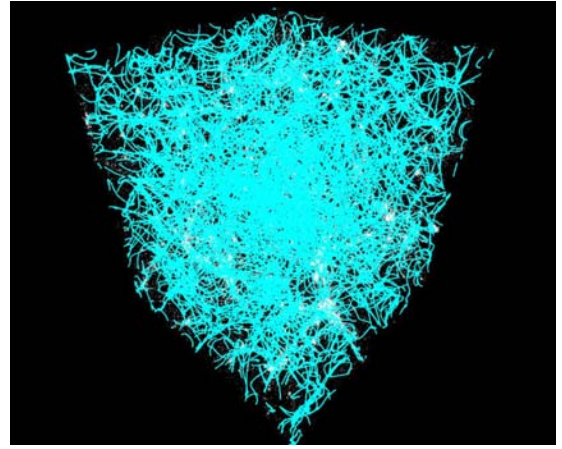
(b) The resulting critical lines made of all the intersection of any two of the three isosurfaces shown in 3(a).
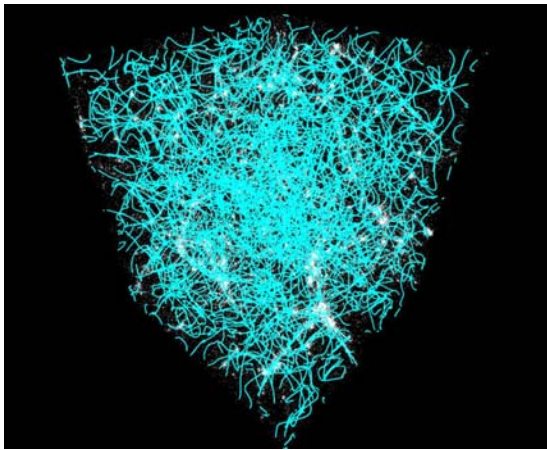
(c) The local critical lines obtained by selecting only the two least degenerate fields depending on the value of the gradient.

(d) The skeleton obtained after enforcing condition 10 on the local critical lines: $\lambda_1 > 0$ and $\lambda_2, \lambda_3 < 0$.

Figure A3. Illustration of the process of the skeleton computation. White points are dark matter particles extracted from a standard ΛCDM simulation run using sc gadget-2. The skeleton is defined as the intersection of two (among three) isosurfaces (Fig. 3a). Defining the curvature as λ_i with $H\nabla\rho = \lambda_i \nabla\rho$ and $\forall j > i, \lambda_j < \lambda_i$ (ρ being the density and H its Hessian), it is possible to select only some parts of the skeleton depending on the value of λ_i and retrieve only the filaments (Fig. 3d). Using a simple post treatment, it is then possible to remove insignificant pieces and obtain the precise locus of the filaments (Fig. 2).

achieve this result (where the weights are assigned depending on the relative importance of the selection criteria). (i) The branches that were missed around the extrema are regenerated using the fact that the skeleton around an extremum is along the main curvature axis (i.e. along the first eigenvector of the Hessian). So for each extremum, marks are given to all skeleton segment, favouring those at small distances and with similar orientation as the main eigenvector of $\mathcal{H}$. Each extremum is eventually connected to the segment with the highest mark. (ii) The gaps between segments in the sequence of skeleton branches are filled. Starting from segments connected to extrema, all segments are visited iteratively: for the running segment, a mark is now assigned to all other unprocessed segments, based upon their relative distance, their relative angle and relative orientation. Note that the corresponding cost functions are non-linear: for instance segments with too large a relative angle are given an exponentially negative mark. (iii) Finally, all segments which have not been considered during step (ii) are dropped. The process is illustrated on Fig. A3, and the resulting skeleton is shown on Fig. 2. A detailed accounting of all stages of the skeleton extraction, including the post treatment is given in Sousbie (2006) (which gives the exact marking scheme described above), while the code is available upon request from the authors.

From a performance point of view, this method presents the advantage of being both fast and robust. The computational cost in fact mainly scales as the number of pixels in grid N_g^3; the cost of computing the isosurfaces intersections is negligible given the possibility of computing only the intersections of faces belonging to the same pixel. It is moreover memory efficient and can be trivially parallelized: the computation can be done on subgrid regions and then merged. On a modern computer, the memory requirement corresponds to the requirement to store one subgrid and its three isosurface, which can be arbitrarily small, and the computational time for a 128^3 pixel grid is of the order of a few seconds on a modern desktop computer while only a few tens of minutes is necessary for a grid of 1024^3 pixel.

APPENDIX B: SMOOTHING LENGTH AND RESOLUTION

One aspect of the numerical implementation that deserves special attention is the issue of smoothing. In the main text, we consider the total skeleton of Gaussian random fields, focusing mainly on two of its properties: its length L and differential length $dL/d\eta$. The algorithm
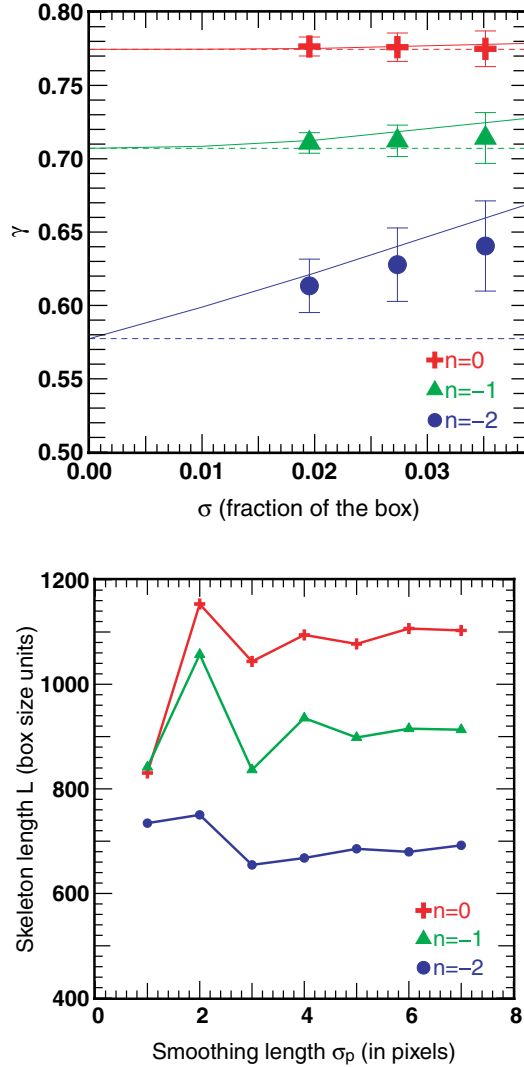
Figure B1. Top: evolution of the measured spectral parameter γ (see equation 16) for 25 realizations of Gaussian random fields with spectral index $n = 0$ (red crosses), $n = -1$ (green triangles) and $n = -1$ (blue discs) as a function of the smoothing scale σ expressed in box size units. The three continuous lines represent the expected theoretical values, measured by integrating the power spectrum truncated to grid limit frequencies. The dotted lines are the theoretical expectations (equation 17) without accounting for finite volume effects. For higher values of σ, the finite box size effects have more influence and the measured value of γ tends to differ from the correct one, thus limiting the maximal smoothing scale Bottom: evolution of the measured length of the total skeleton in box size units as a function of the smoothing length in pixels σ_p, for different values of the spectral index n and while keeping the smoothing scale to a constant fraction of the box size $\sigma \approx 0.031$. The measurements are obtained by resampling one initial realization of a Gaussian random field (generated over a 256^3 pixel grid) on smaller resolution grids and smoothing the resulting fields over the appropriate number of pixels. The measured length of the total skeleton appears to become stable for values of σ_p above a limit of 4 to 5 pixels at least, which corresponds to $\sigma > 0.19$ for a field sampled on a 256^3 pixels grid.

presented in this paper deals with the numerical computation of the skeleton of a discretized realization of a given field. It is thus important in the first instance to be able to deal with the influence of this discretization on the measured skeleton properties (see e.g. Colombi et al. 2000).

The statistical properties of a scale-free Gaussian random field can be described using only two numbers: its spectral index n and the amplitude A of its power spectrum $P(k) = Ak^n$, where k is the wavenumber. The skeleton formalism is totally independent of the amplitude of the field, so only the value of n is of interest to us. Consider a realization of a 3D scale-free Gaussian random field with spectral index n on a N_g^3 pixel grid. In order to ensure sufficient differentiability, this field is convolved to a Gaussian kernel whose scale σ is expressed per unit box size. The value of σ limits the size of the smallest scale that can be considered, while the finite size of the grid imposes an upper limit. Fig. B1 presents the measured value of the spectral parameter $\gamma^2 = (n + 3)/(n + 5)$ as a function of σ, for 25 realizations of Gaussian random fields with spectral index $n \in \{0, -1, -2\}$, together with the theoretical value, measured by integrating the power spectrum truncated to grid limit frequencies. As expected, a departure from theory is observed for higher values of σ, especially for fields with lower spectral index where most of the power is concentrated on small values of k (i.e. on large scales). This sets an upper limit on the value of the smoothing scale and so we will only be considering fields smoothed on scales $\sigma \leqslant 0.035$.

The other constraint on the value of σ arises from the fact that the skeleton computation algorithm requires a field that is continuously differentiable twice in the finite difference scheme sense. This means that the smoothing length should be large enough for the computational

errors on field derivatives to be negligible. These considerations imply a lower limit on the smoothing length value expressed in number of pixels $\sigma_{\rm p} = \sigma \, N_{\rm g}$. In order to estimate this limit, we again generated Gaussian random fields with different spectral indices over a 256^3 ($N_{\rm g} = 256$) pixels grid and downsampled them on grids with eight different values of $N_{\rm g}$ ranging from $N_{\rm g} = 64$ up to $N_{\rm g} = 224$. Fig. B1 presents the evolution of the measured skeleton length for these realizations, each of them being computed for a smoothing scale corresponding to a constant fraction of the box size $\sigma \approx 0.031$ but to different values of $\sigma_{\rm p}$ ranging from $\sigma_{\rm p} = 1$ up to $\sigma_{\rm p} = 7$. One would clearly expect the length of the skeleton to depend only on the value of σ as long as the numerical approximations are negligible, which seems to be the case only for values of $\sigma_{\rm p}$ at least of order 5 pixels. For a given sampling $N_{\rm g}$, this limits the possible smoothing scale to $\sigma > 5/N_{\rm g}$. As was noted previously, this exact value depends on the considered spectral index, so we chose to consider the worst case, $n = 1$, where the fluctuations of the field do not dampen on small scales thus making the field naturally not smooth on any scale.

In this paper, all fields considered are sampled over $N_{\rm g} = 256$ cubic grids, so in order to respect the constraints described above, the fields are smoothed on scales in the range $0.02 < \sigma < 0.035$.

APPENDIX C: THE THEORETICAL DIFFERENTIAL LENGTH OF THE SKELETON

C1 Average length of the skeleton per unit volume

To find the average length per unit volume, $\mathcal{L}(\rho_{\rm th})$, of the critical lines[5] that are above the threshold $\rho_{\rm th}$ consider the vicinity of the points through which the local critical line passes, $S_i = 0$, $S_j = 0$. (Since the sets of conditions $(S_i, S_j) = (0, 0)$, $i \neq j$ is degenerate, without loss of generality one can assume a particular choice for i and j). Define the set of points, $\mathcal{E}$, in the excursion $\rho > \rho_{\rm th}$ near the critical line solutions that satisfy $-\Delta S_i/2 \leqslant S_i \leqslant \Delta S_i/2$ and $-\Delta S_j/2 \leqslant S_j \leqslant \Delta S_j/2$ where ΔS_i and ΔS_j are sufficiently small so that the linear expansion $\Delta S_i \approx \nabla S_i \cdot {\rm d}\,{\boldsymbol r}$, $\Delta S_j \approx \nabla S_j \cdot {\rm d}\,{\boldsymbol r}$ holds.[6] The fraction of the total volume the set $\mathcal{E}$ occupies (the filling factor) is

$$\mathcal{V}(\rho_{\rm th}, \Delta S_i, \Delta S_j) = \int_{\rho > \rho_{\rm th}} {\rm d}\rho \int_{-\Delta S_i/2}^{\Delta S_i/2} {\rm d}S_i \int_{-\Delta S_j/2}^{\Delta S_j/2} {\rm d}S_j \int {\rm d}^3(\nabla S_i) \, {\rm d}^3(\nabla S_j) \mathcal{P}(\rho, S_i, S_j, \nabla S_i, \nabla S_j), \tag{C1}$$

where $\mathcal{P}(\rho, S_i, S_j, \nabla S_i, \nabla S_j)$ is the joint PDF of the quantities $(\rho, S_i, S_j, \nabla S_i, \nabla S_j)$. Here the seemingly redundant distribution of the gradients ∇S_i and ∇S_j was introduced to have the expression for the fraction of the total volume occupied by a differential subset of $\mathcal{E}$ that has specific values of the gradients $\nabla S_i, \nabla S_j$ [within ${\rm d}^3(\nabla S_i)$ and ${\rm d}^3(\nabla S_j)$]:

$$\mathrm{d}\mathcal{V}(\rho_{\rm th}, \Delta S_i, \Delta S_j, \nabla S_i, \nabla S_j) = {\rm d}^3(\nabla S_i) \, {\rm d}^3(\nabla S_j) \int_{\rho > \rho_{\rm th}} {\rm d}\rho \int_{-\Delta S_i/2}^{\Delta S_i/2} {\rm d}S_i \int_{-\Delta S_j/2}^{\Delta S_j/2} {\rm d}S_j \mathcal{P}(\rho, S_i, S_j, \nabla S_i, \nabla S_j). \tag{C2}$$

Since the area, Σ, of a section locally orthogonal to the such subset, is simply (modulo some trigonometry) given by

$$\Sigma(\Delta S_i, \Delta S_j, \nabla S_i, \nabla S_j) = \Delta S_i \Delta S_j / |\nabla S_i \times \nabla S_j|,$$

dividing ${\rm d}\mathcal{V}$ by Σ, integrating over all possible gradients $\nabla S_i, \nabla S_j$ and then taking the limit $(\Delta S_i, \Delta S_j) \to (0, 0)$ yields the length per unit volume of the skeleton that is above the threshold $\rho_{\rm th}$:

$$\begin{aligned} \mathcal{L}(\rho_{\rm th}) &= \lim_{(\Delta S_i, \Delta S_j) \to (0,0)} \int \frac{{\rm d}\mathcal{V}(\rho_{\rm th}, \Delta S_i, \Delta S_j, \nabla S_i, \nabla S_j)}{\Sigma(\Delta S_i, \Delta S_j, \nabla S_i, \nabla S_j)} \\ &= \int_{\rho > \rho_{\rm th}} {\rm d}\rho \int {\rm d}^3(\nabla S_i) \, {\rm d}^3(\nabla S_j) \, |\nabla S_i \times \nabla S_j| \mathcal{P}(\rho, S_i = 0, S_j = 0, \nabla S_i, \nabla S_j). \end{aligned} \tag{C3}$$

This generalizes the calculation of NCD to three dimensions: the length of the local skeleton is defined by the properties of the density field and its partial derivatives up to third order, as expected.

In order to understand the scalings involved in the computation of $\mathcal{L}$, let us rewrite this equation in terms of dimensionless quantities

$$\sigma_0 \eta \equiv \rho_{th}, \quad \sigma_0 x \equiv \rho, \quad \sigma_1 x_i \equiv \frac{\partial \rho}{\partial r_i}, \quad \sigma_2 x_{ij} \equiv \frac{\partial^2 \rho}{\partial r_i \partial r_j}, \quad \sigma_3 x_{ijk} \equiv \frac{\partial^3 \rho}{\partial r_i \partial r_j \partial r_k}, \quad \sigma_2 \sigma_1^2 s_i \equiv S_i, \quad \sigma_3 \sigma_1^2 \nabla s_i \equiv \nabla S_i, \tag{C4}$$

with, following Bardeen et al. (1986),

$$\sigma_n^2 \equiv \int \frac{k^2 {\rm d}k}{2\pi^2} P(k) k^{2n}, \tag{C5}$$

where $P(k)$ is the power spectrum of ρ. Equation (9) and its gradient can be written more conveniently using the totally antisymmetric tensor, ϵ^{ijk}, as

$$s_i = \sum_{jkl} \epsilon^{ijk} x_{jl} x_l x_k, \quad \text{and} \quad \nabla_m s_i \equiv \nabla \hat{s}_i(x_k, x_{kl}, x_{klm}) = \sum_{jkl} \epsilon^{ijk} \left(x_{jlm} x_l x_k + \tilde{\gamma} [x_{jl} x_{lm} x_k + x_{jl} x_{km} x_l] \right). \tag{C6}$$

---

[5] the distinction is made here between the theoretical expectation, $\mathcal{L}(\rho_{\rm th})$, in this section and the estimator, L, in the main text.

[6] In such small neighbourhood of a critical line there are no other critical lines since the solution for linearized skeleton equations is unique. Note that the linear expansion of S_i breaks near the extrema of the field, where $\nabla S_i = 0$, which allows several critical lines to intersect at such points. However, extremal points are of measure zero as far as the computation of the length of the skeleton is concerned.

The variances of all the random quantities defined in equation (C4) do not depend on spectral parameters (they are pure numbers) except for ∇s_i. Keeping that in mind, in this new notation, equation (C3) becomes

$$\mathcal{L}(\eta) = \left(\frac{\sigma_3}{\sigma_2}\right)^2 \int_{x > \eta} \mathrm{d}x \int \mathrm{d}^3(\nabla s_i)\, \mathrm{d}^3(\nabla s_j)\, |\nabla s_i \times \nabla s_j| \mathcal{P}(x, s_i = 0, s_j = 0, \nabla s_i, \nabla s_j). \tag{C7}$$

Equation (C7) is the formal expression for the length per unit volume of the total set of critical lines above the threshold η.

Let us express the joint distribution function $\mathcal{P}(\eta, s_i, s_j, \nabla s_i, \nabla s_j)$ in terms of the joint distribution function of the underlying field and its derivatives $P(x, x_k, x_{km}, x_{klm})$. Expressions (C6) for s_i and ∇s_i involve up to third derivatives of the field x, thus, accounting for the symmetries in the derivative tensors of the second and third order is, one deals with a set of 20 independent variables $(x, x_k, x_{kl}, x_{klm})$. $\mathcal{P}$ is obtained as a marginalization over the field distribution

$$\mathcal{P}(\eta, s_i, s_j, \nabla s_i, \nabla s_j) = \int \mathrm{d}x\, \mathrm{d}^3 x_k \mathrm{d}^6 x_{kl} \mathrm{d}^{10} x_{klm}\, P(x, x_k, x_{kl}, x_{klm}) \delta_D(x - \eta) \delta_D(\hat{s}_i(x_k, x_{kl}) - s_i) \delta_D(\hat{s}_j(x_k, x_{kl}) - s_j)$$

$$\times \delta_D(\nabla \hat{s}_i(x_k, x_{kl}, x_{klm}) - \nabla s_i) \delta_D(\nabla \hat{s}_j(x_k, x_{kl}, x_{klm}) - \nabla s_j), \tag{C8}$$

which yields the appropriate 9D probability density. Substituting the expression (C8) into equation (C7), differentiating with respect to η, and accounting for two delta functions in ∇s_i and ∇s_j yields

$$\frac{\partial \mathcal{L}}{\partial \eta} = \left(\frac{1}{\tilde{R}}\right)^2 \int \mathrm{d}^3 x_k \mathrm{d}^6 x_{kl} \mathrm{d}^{10} x_{klm}\, |\nabla \hat{s}_i \times \nabla \hat{s}_j| P(\eta, x_k, x_{kl}, x_{klm})\, \delta_D(\hat{s}_i(x_k, x_{kl})) \delta_D(\hat{s}_j(x_k, x_{kl})), \tag{C9}$$

where σ_3/σ_2 is rewritten in terms $\tilde{R}$ with the help of equation (16).

Expression (C9) gives the differential length per unit volume of the total set of critical lines. Note that $\nabla \hat{s}_i$ and $\nabla \hat{s}_j$ are now functions of (x_k, x_{kl}, x_{klm}) and $\hat{s}_i$ and $\hat{s}_j$ are function of (x_k, x_{kl}) given by equation (C6). The two delta functions couple the different x_k, x_{kl}, accounting for the fact that the integral should be restricted to the intersection of the two isosurfaces, i.e. along the critical lines. The modulus in $|\nabla \hat{s}_i \times \nabla \hat{s}_j|$ makes the summation of skeleton segments non-algebraic, which complicates the further reduction of equation (C9). For the set of local critical lines, there are no restriction to the region of integration. If one is interested in the local skeleton, the integration should be restricted to regions where the condition given by equation (10) holds.

The total length of the critical lines per unit volume is

$$\mathcal{L}_{tot} = \int_{-\infty}^{\infty} \mathrm{d}\eta \frac{\partial \mathcal{L}}{\partial \eta} = \left(\frac{1}{\tilde{R}}\right)^2 \int \mathrm{d}^3 x_k \mathrm{d}^6 x_{kl} \mathrm{d}^{10} x_{klm}\, |\nabla \hat{s}_i \times \nabla \hat{s}_j| P(x_k, x_{kl}, x_{klm}) \delta_D(\hat{s}_i(x_k, x_{kl})) \delta_D(\hat{s}_j(x_k, x_{kl})). \tag{C10}$$

C2 $\partial \mathcal{L}/\partial \eta$ for a Gaussian random field

Since at a point the value of a Gaussian field does not correlate with its derivatives of odd orders (this is easy to understand using symmetries in Fourier space), the joint distribution function $P(x, x_k, x_{kl}, x_{klm})$ can be factorized as

$$P(x, x_k, x_{kl}, x_{klm}) = P_0(x, x_{kl}) P_1(x_k, x_{klm}). \tag{C11}$$

In P_0, the only dependence on the power spectrum of the underlying field is in the parameter γ (cf. equation 16) that describes the correlation between the field and its second derivatives. Similarly, $P_1(x_i, x_{ijk})$ only involves $\tilde{\gamma}$ which describes the correlation between the gradient of the field and its third derivatives. Therefore $\partial \mathcal{L}/\partial \eta$ depends only on η, $\tilde{R}$, γ and $\tilde{\gamma}$, as argued in the main text. Note that, by symmetry, $\partial \mathcal{L}/\partial \eta$ for the total set of critical lines should be an even function of η. The total length of the skeleton, $\mathcal{L}_{\text{tot}}$, which follows from marginalization of the equation (C9) over η may depend only on $\tilde{\gamma}$ and $\tilde{R}$ since the integration of $P_0(\eta, x_{kl})$ over all η cancels the dependency on γ.

C2.1 The 'stiff' filament approximation

The $1/\tilde{R}^2$ scaling in equation (C9) reflects the basic fact that, by definition, the local skeleton lines are almost straight within a volume, $\sim \tilde{R}^3$, that contains one inflection point. A straight segment through such volume has the length $\sim \tilde{R}$, thus, the expected length per unit volume is $\sim 1/\tilde{R}^2$. The dependence on the spectral index is then $1/\tilde{R}^2 \propto (n+7)/\sigma^2$, where σ is the smoothing length. Is this the scaling with n that one should expect in the simulations? Let us write formally

$$\nabla s_i \times \nabla s_j = \boldsymbol{A}(x_k, x_{kl}, x_{klm}) + \tilde{\gamma} \boldsymbol{B}(x_k, x_{kl}, x_{klm}) + \tilde{\gamma}^2 \boldsymbol{C}(x_k, x_{kl}). \tag{C12}$$

Suppose the last term dominates statistically.[7] Then, since $\tilde{\gamma}/\tilde{R} = 1/R_*$, and given that $\boldsymbol{C}(x_k, x_{lm})$ does not depend on the third derivative of the field (which can then be integrated out), equation (C12) becomes

$$\frac{\partial \mathcal{L}}{\partial \eta} \approx \left(\frac{1}{R_*}\right)^2 \int \mathrm{d}^3 x_k \mathrm{d}^6 x_{kl}\, |\boldsymbol{C}(x_k, x_{lm})|\, P_0(\eta, x_{kl})\, P_1(x_k) \delta_D(\hat{s}_i(x_k, x_{kl})) \delta_D(\hat{s}_j(x_k, x_{kl})). \tag{C13}$$

It is easy to foresee when this regime is valid. The same argument as before implies that the $1/R_*^2$ scaling arises when the skeleton is almost straight within a volume that contains one extremum, $\sim R_*^3$, rather than one inflection point. This is supported by the fact that the integral

---

[7] Or equivalently assume that the magnitude of derivative of the Hessian is negligible relative to the magnitude of the Hessian.

does not depend on the third derivatives, thus inflection points play no role, and no dependence on $\tilde{\gamma}$ remains. This picture corresponds to a skeleton connecting extrema with relatively straight segments. The scaling is then $1/R_*^2 \propto (n+5)/\sigma^2$. We call this regime and the expression (C13) that describes it 'the stiff approximation'.

For the total length of the critical lines, integration over η gives

$$\mathcal{L}_{tot} \approx \left(\frac{1}{R_*}\right)^2 \int \mathrm{d}^3 x_k \mathrm{d}^6 x_{kl} \; |\boldsymbol{C}(x_k, x_{lm})| \; P_0(x_{kl}) P_1(x_k) \delta_D(\hat{s}_i(x_k, x_{kl})) \delta_D(\hat{s}_j(x_k, x_{kl})) \propto (n+5)\sigma^{-2} \tag{C14}$$

strictly, since the integral is just a pure number. This is very close to the scaling with n that was found in the numerical fit, equation (18). The differential length in the stiff regime is then only the function of γ times $\mathcal{L}_{tot}$. This demonstrates theoretical consistency between the scaling $\sim (n+5)\,\sigma^{-2}$ of $\mathcal{L}_{tot}$ and insensitivity of $\partial \mathcal{L}/\partial \eta$ to $\tilde{\gamma}$ for scale-free Gaussian random fields that was observed in the simulations.

C2.2 Joint distribution of the field and its derivatives for a Gaussian random field

The full expression $P_0(x, x_{kl})$ is given in Bardeen et al. (1986). Introducing variables

$$u \equiv -\Delta x = -(x_{11} + x_{22} + x_{33}), \quad v \equiv \frac{1}{2}(x_{33} - x_{11}), \quad w \equiv \sqrt{\frac{1}{12}}(2x_{22} - x_{11} - x_{33}), \tag{C15}$$

in place of diagonal elements of the Hessian (x_{11}, x_{22}, x_{33}) one finds that $u, v, w, x_{12}, x_{13}, x_{23}$ are uncorrelated. Importantly, the field, x is only correlated with $u = \Delta x$ and

$$\langle xu \rangle = \gamma, \quad \langle xv \rangle = 0, \quad \langle xw \rangle = 0, \quad \langle x x_{kl} \rangle = 0, \; k \neq l, \tag{C16}$$

where γ is the same quantity as in equation (16). The full expression of $P_0(x, x_{kl})$ is then

$$P_0(x, x_{kl}) \, \mathrm{d}x \, \mathrm{d}^6 x_{kl} = \frac{(15)^{5/2}}{(2\pi)^{7/2}(1-\gamma^2)^{1/2}} \exp\left[-\frac{(x-\gamma u)^2}{2(1-\gamma^2)} - \frac{u^2}{2}\right] \exp\left[-\frac{15}{2}(v^2 + w^2 + x_{12}^2 + x_{13}^2 + x_{23}^2)\right] \, \mathrm{d}x \, \mathrm{d}u \, \mathrm{d}v \, \mathrm{d}w \, \mathrm{d}x_{12} \, \mathrm{d}x_{13} \, \mathrm{d}x_{23},$$

and is described by only one correlation parameter γ.

A similar procedure can be performed for the joint probability of the first and third derivatives of the fields, $P_1(x_i, x_{ijk})$, by defining the following nine parameters (see also Hanami 99):

$$u_i \equiv \nabla_i u, \quad v_i \equiv \frac{1}{2}\epsilon^{ijk}\nabla_i(\nabla_j\nabla_j - \nabla_k\nabla_k)x, \; \text{with} \; j < k, \quad \text{and} \quad w_i \equiv \sqrt{\frac{5}{12}}\nabla_i\left(\nabla_i\nabla_i - \frac{3}{5}\Delta\right)x, \tag{C17}$$

and replacing the variables ($x_{i11}, x_{i22}, x_{i33}$) with ($u_i, v_i, w_i$). In that case, the only cross-correlations in the vector ($x_1, x_2, x_3, u_1, v_1, w_1, u_2, v_2, w_2, u_3, v_3, w_3, x_{123}$) which do not vanish are between the same components of the gradient and the gradient of the Laplacian of the field:

$$\langle x_i u_i \rangle = \tilde{\gamma}/3, \quad i = 1, 2, 3, \tag{C18}$$

where $\tilde{\gamma}$ is the same quantity as in equation (16).

This allows us to write

$$P_1(x_i, x_{ijk}) \mathrm{d}^3 x_i \, \mathrm{d}^{10} x_{ijk} = \frac{105^{7/2} 3^3 \mathrm{d}^3 w_i \, \mathrm{d}^3 v_i \, \mathrm{d}x_{123}}{(2\pi)^{13/2}(1-\tilde{\gamma}^2)^{3/2}} \exp\left[-\frac{105}{2}\left(x_{123}^2 + \sum_{i=1}^{3}(v_i^2 + w_i^2)\right)\right] \prod_{i=1}^{3} \mathrm{d}u_i \, \mathrm{d}x_i \exp\left[-\frac{3(u_i - \tilde{\gamma}x_i)^2}{2(1-\tilde{\gamma}^2)} - \frac{3x_i^2}{2}\right].$$

C2.3 Dependence of the differential length on threshold and spectral parameters

What is the dependence of the skeleton differential length on the parameter γ and the threshold η? Let us look at the structure of the integrals involved with respect to the variable u. Importantly, the arguments of the delta functions $S_i \neq S_i(u)$ and $\nabla S_i \times \nabla S_j$, given by equation (C12), is $\sim \sqrt{Q_4(u)}$ where $Q_4(u)$ is a positive quartic in u. Inserting the expressions for P_1 and P_0 into equation (C9) one sees that the integral over u in $\partial \mathcal{L}/\partial \eta$ has the form

$$\mathcal{I}(\gamma, \eta) = \int_{-\infty}^{\infty} \frac{\sqrt{Q_4(u)} \exp\left(-u^2/2\right)}{(2\pi)^{1/2}(1-\gamma^2)^{1/2}} \exp\left[-\frac{(\eta-\gamma u)^2}{2(1-\gamma^2)}\right] \, \mathrm{d}u, \tag{C19}$$

where $Q_4(u)$, of course, also depends on $v, w, u_k, v_k, w_k, x_{123}, x_{k<l}$ and possibly $\tilde{\gamma}$, but not on γ.

In the trivial limit $\gamma \to 0$ the coupling between u and the field value η vanishes and the differential length is reduced to the PDF of η:

$$\mathrm{d}\mathcal{L}/\mathrm{d}\eta \propto \exp\left[-\frac{\eta^2}{2}\right] = \frac{\mathcal{L}_{tot}}{(2\pi)^{1/2}} \exp\left[-\frac{\eta^2}{2}\right]. \tag{C20}$$

For non-vanishing γ, following NCD, the differentiation of equation (C19) shows that $\mathcal{I}(\gamma, \eta)$ obeys the equation

$$\gamma \frac{\partial \mathcal{I}}{\partial \gamma} = -\frac{\partial}{\partial \eta}\left[\eta \mathcal{I}(\gamma, \eta) + \frac{\partial \mathcal{I}}{\partial \eta}\right],$$

whose solution involve even Hermite polynomials (retaining only the convergent solution at large η):

$$\mathcal{I}(\gamma, \eta) = \sum_{n=0}^{\infty} c_{2n} \gamma^{2n} H_{2n}(\eta/\sqrt{2}) \exp\left[-\frac{\eta^2}{2}\right]. \tag{C21}$$

Due to the orthogonality property of Hermite polynomials, c_{2n} is given by

$$c_{2n} = \lim_{\gamma \to 1} \int \mathrm{d}x \, H_{2n}(x/\sqrt{2}) \exp(-x^2/2) \mathcal{I}(x, \gamma) = \int \mathrm{d}x \, H_{2n}(x/\sqrt{2}) \exp(-x^2/2) \sqrt{Q_4(x)} = c_{2n}(v, w, u_i, u_i, w_i, x_{123}, x_{ij}). \tag{C22}$$

The integration of equation (C21) over $v, w, u_k, u_k, w_k, x_{123}, x_{k<l}$ (while accounting for the rest of the integrant corresponding to P_0 and P_1 together with the two delta functions) yields the functional form of $\mathrm{d}\mathcal{L}/\mathrm{d}\eta$, equation (19), where C_{2n} is a pure number in the stiff approximation, but may depend on $\tilde{\gamma}$ in general.

In the stiff approximation, equation (C13) can be investigated semi-analytically. This analysis is the subject of a subsequent paper.

This paper has been typeset from a TEX/LATEX file prepared by the author.