Measuring the geometry and topology of large-scale structure using SURFGEN: methodology and preliminary results

Jatush V. Sheth,1⋆ Varun Sahni,1⋆ Sergei F. Shandarin2⋆ and B. S. Sathyaprakash3⋆

1Inter University Centre for Astronomy and Astrophysics, Pune, India
2Department of Physics and Astronomy, University of Kansas, KS 66045, USA
3Department of Physics and Astronomy, University of Cardiff, Cardiff CF2 3YB

Accepted 2003 March 20. Received 2003 March 6; in original form 2002 October 15

ABSTRACT
Observations of the Universe reveal that matter within it clusters on a variety of scales. On scales between 10 and 100 Mpc, the Universe is spanned by a percolating network of superclusters interspersed with large and almost empty regions – voids. This paper, the first in a series, presents a new Ansatz that can successfully be used to determine the morphological properties of the supercluster–void network. The Ansatz is based on a surface modelling scheme (SURFGEN), developed explicitly for the purpose, which generates a triangulated surface from a discrete data set representing (say) the distribution of galaxies in real (or redshift) space. The triangulated surface describes, at progressively lower density thresholds, clusters of galaxies, superclusters of galaxies and voids. Four Minkowski functionals (MFs) – surface area, volume, extrinsic curvature and genus – describe the geometry and topology of the supercluster–void network. On a discretized and closed triangulated surface, the MFs are determined using SURFGEN. Ratios of the MFs provide us with an excellent diagnostic of three-dimensional shapes of clusters, superclusters and voids. MFs can be studied at different levels of the density contrast and therefore probe the morphology of large-scale structure on a variety of length-scales. Our method for determining the MFs of a triangulated isodensity surface is tested against both simply and multiply connected eikonal surfaces such as triaxial ellipsoids and tori. The performance of our code is thereby evaluated using density distributions that are pancake-like, filamentary, ribbon-like and spherical. Remarkably, the first three MFs are computed to better than 1 per cent accuracy while the fourth (genus) is known exactly. SURFGEN also gives very accurate results when applied to Gaussian random fields. We apply SURFGEN to study morphology in three cosmological models, ΛCDM, τCDM and SCDM, at the present epoch. Geometrical properties of the supercluster–void network are found to be sensitive to the underlying cosmological parameter set. For instance, the percolating supercluster in ΛCDM turns out to be more filamentary but topologically simpler than superclusters in τCDM and SCDM. It occupies just 0.6 per cent of the total simulation-box volume yet contains about 4 per cent of the total mass. Our results indicate that the surface modelling scheme to calculate MFs is accurate and robust and can successfully be used to quantify the topology and morphology of the supercluster–void network in the universe.

Key words: methods: numerical – galaxies: statistics – cosmology: theory – large-scale structure of Universe.

1 INTRODUCTION
The existence of the supercluster–void network is one of the most intriguing observational features of our Universe. Redshift surveys of galaxies confirm that on very large scales the bulk of matter in the Universe is concentrated in clusters and superclusters of galaxies which are separated by large almost empty regions, appropriately called voids. At moderate density thresholds δ ~ 1, the supercluster network percolates even though it occupies a tiny fraction of the total volume, and so has a small filling fraction.

The morphology of the supercluster–void network is quite complex and has inspired evocative descriptions such as being

⋆E-mail: jvs@iucaa.ernet.in (JVS); varun@iucaa.ernet.in (VS); sergei@ukans.edu, sergei@ku.edu (SFS); B.Sathyaprakash@astro.cf.ac.uk (BSS)

© 2003 RAS
Quantifying large-scale structure using SURFGEN

2 MINKOWSKI FUNCTIONALS AND TRIANGULATED SURFACES

The large-scale structure of the Universe can be studied on various scales by considering the geometry and topology of excursion sets of the density contrast \( \delta(x) \), defined as \( E_{MF} = \{ x \mid \delta(x) > \delta_{TH} \} \) for overdense regions (clusters, superclusters), and as \( E_{MF} = \{ x \mid \delta(x) < \delta_{TH} \} \) for underdense regions (voids). By specifying a given density

'‘honeycomb-like’, ‘bubble-like’, ‘a filamentary network’, ‘Swiss cheese’, ‘cosmic web’, etc. Indeed the intricate weave of large-scale structure is becoming increasingly evident, as results from progressively larger galaxy redshift surveys (CFA, IRAS, PSCz, LCRS, 2dFGRS, SDSS) demonstrate. Quantifying the properties of the supercluster–void network is clearly one of the cardinal tasks facing cosmology today. Since the network evolved from an almost featureless Gaussian random field (as suggested by observations of the cosmic microwave background), it is important to understand how such a rich and complex cosmic tapestry – characterized by prominent non-Gaussian features (clusters, superclusters, voids) – could have arisen from small and statistically random initial conditions. Furthermore, since the network is highly evolved it is unlikely that its principal features can be described within a strictly perturbative framework which breaks down when \( \delta \geq 1 \). It is therefore both interesting and revealing that approximations that successfully describe non-linear dynamics (such as the Zel’dovich approximation and the adhesion model) generically predict the formation of filaments and pancakes, which interweave and percolate to form the large-scale structure of the Universe (Gurbatov, Saichev & Shandarin 1985; 1989; Shandarin & Zeldovich 1989). Furthermore, this interweaving network of filaments, sheets and voids – predicted by both the Zel’dovich approximation and the adhesion model – is readily seen in \( N \)-body simulations of gravitational clustering for a wide variety of initial power spectra (Klypin & Shandarin 1993; Sathyaprakash, Sahni & Shandarin 1996). It therefore appears that a filamentary distribution of large-scale structure is an almost generic outcome of pressureless (dark) matter clustering from Gaussian initial conditions. [A theoretical model for understanding the morphology of large-scale structure can be found in the ‘cosmic web’ hypothesis of Bond, Kohman & Pogosyan (1996). Insightful observations of the statistical and dynamical techniques used in studies of large-scale structure can be found in Sahni & Coles (1995), Martínez & Saar (2002) and van de Weygaert (2002).]

A number of statistical measures have been advanced to quantify the pattern made by galaxies as they cluster in our Universe. Prominent among them are percolation analysis (Zeldovich, Einasto & Shandarin 1982; Shandarin & Zeldovich 1983), counts in cells (Janes & Demarque 1983; de Lapparent, Geller & Huchra 1991), minimal spanning trees (Barrow, Sonoda & Bhavsar 1985), the genus measure (Gott, Melott & Dickinson 1986), etc. [To some extent these methods complement the traditional approach to quantify clustering using the hierarchy of correlation functions (Peebles 1980), which become cumbersome to evaluate beyond the three-point function.] Fairly recently Mecke, Buchert & Wagner (1994) introduced Minkowski functionals (MFs) to cosmology. MFs contain information pertaining to geometry, connectivity (percolation) and topology (genus). In addition, the ratios of MFs quantify the morphology of large-scale structure. This has led to the construction of a set of measures called shapefinders which tell us whether the distribution of matter in superclusters/voids is spherical, planar, filamentary, etc. (Sahni, Sathyaprakash & Shandarin 1998; Sathyaprakash, Sahni & Shandarin 1998). [Applications of shapefinders to galaxy catalogues and \( N \)-body simulations has been discussed in Basilakos, Plionis & Rowan-Robinson (2001) and Kolokotronis, Basilakos & Plionis (2002). Moment-based studies of morphology pre-dating the shapefinders can be found in Babul & Starkman (1992) and Luo & Vishniac (1995.) Thus, between them, the four MFs contain valuable information regarding both the geometrical as well as the topological distribution of matter in the Universe. Since MFs are additive in nature, one can both glean information regarding individual objects (galaxies, clusters, voids) as well as describe the supercluster–void network in its totality.

There have been three major attempts made to study the morphology of the large-scale structure using MFs. These efforts differ in their approach of evaluating the MFs: (1) Boolean grain models study the MFs of surfaces which result due to intersecting spheres decorating the input point set (Mecke et al. 1994). (2) Krofton’s formulæ make it possible to calculate MFs on a density field defined on a grid (Schmalzing & Buchert 1997). In this case the MFs are calculated by using the information of the number of cells in 1D (vertices), 2D (faces) and 3D (cuboids). (3) An alternative, resolution-dependent approach consists in employing the Koenderink invariants (Schmalzing & Buchert 1997; Schmalzing et al. 1999).

This paper introduces a radically new and powerful approach for computing the MFs. This approach consists of constructing isodensity surfaces using an elaborate surface modelling scheme defined in terms of excursion sets of a density field. These surfaces are triangulated and the MFs are evaluated for the resulting closed polyhedral surface. A similar algorithm has been used in studies of 2D cosmic microwave background maps (Novikov, Feldman & Shandarin 1999; Shandarin et al. 2002). However, developing a 3D algorithm requires a new component that builds a surface of approximately constant (to linear order) density. It has been implemented and reported in this paper.

The MFs so constructed are tested against known results for simply connected surfaces (triaxial ellipsoids) as well as multiply connected surfaces (triaxial tori). In all cases we find that our Ansatz for determining MFs on polyhedra gives results that are in excellent agreement with exact continuum values for these quantities. We have further tested the performance of SURFGEN against Gaussian random fields for which the analytic prediction for MFs is available and find that our calculations reproduce the analytic results to a remarkable precision. This gives us confidence that our Ansatz is well suited for application to large \( N \)-body simulations and 3D galaxy redshift surveys. Encouraged by the success of these preliminary tests, we apply our method to simulations of three cosmological models due to the Virgo Consortium.

The rest of this paper is arranged as follows. In Section 2, we provide a brief overview of MFs and present an algorithm that computes MFs for a triangulated polyhedral surface. In Section 3, we describe the SURFGEN algorithm, which triangulates a general class of surfaces including isodensity surfaces describing superclusters or voids. In Section 4 we test our method against standard eikonal surfaces – spheres, ellipsoids, tori – and demonstrate the accuracy of our algorithm. Special boundary conditions are needed to reproduce the analytic expressions for MFs of a Gaussian random field. The related discussion along with the results for Gaussian random fields is presented at the end of the paper in the Appendix. Section 5 is devoted to a detailed morphological study of cosmological \( N \)-body simulations. The models investigated are \( \Lambda \)CDM, rCDM and SCDM at the present cosmological epoch. Our conclusions are presented in Section 6, which also discusses possible further applications of SURFGEN and the MFs.
threshold one effectively defines an isodensity surface over which the MFs should be evaluated. Depending upon whether the surface encloses an overdense region of space or an underdense region, the surface refers to a cluster/supercluster or a void. Note that the term ‘cluster’ is used for any region in the excursion set \( \delta > \delta_{TH} \) which in general is different from the astronomical definition. The terminology reflects the cluster analysis jargon. The following four MFs describe the morphological properties of an isodensity surface in three dimensions:\footnote{Genus measures the number of handles in excess of the number of isolated underdense regions (voids) that the surface of a cluster exhibits.}

1. area \( S \) of the surface;
2. volume \( V \) enclosed by the surface;
3. integrated mean curvature \( C \) of the surface (or integrated extrinsic curvature)

\[
C = \frac{1}{2 \pi} \oint \left( \frac{1}{R_1} + \frac{1}{R_2} \right) \, dS,
\]

where \( R_1 \) and \( R_2 \) are the principal radii of curvature at a given point on the surface;

4. integrated intrinsic (or Gaussian) curvature \( \chi \) of the surface – also called the Euler characteristic

\[
\chi = \frac{1}{2\pi} \oint \left( \frac{1}{R_1 R_2} \right) \, dS.
\]

A related measure of topology is the genus\footnote{There are \((n+1)\) MFs in \(n\) dimensions each defined for an \((n-1)\)-dimensional hypersurface. Two-dimensional MFs have been used to analyse the LCRS slices by Bharadwaj et al. (2000) and the anisotropy of the cosmic microwave background by Schmalzing & Gorski (1998) and by Novikov et al. (1999).} \( G = 1 - \chi/2 \). Multiply connected surfaces have \( G > 0 \) while \( G = 0 \) for a simply connected surface such as a sphere. (The topological properties of all orientable surfaces are equivalent to those of a sphere with one or more ‘handles’. Thus a torus is homeomorphic to a sphere with one handle, while a pretzel is homeomorphic to a sphere with two handles, etc.) While the genus provides information about the connectivity of a surface, the remaining three MFs are sensitive to local surface deformations and hence characterize the geometry and shape of large-scale structure at varying thresholds of the density (Sahni et al. 1998).

In nature one seldom comes across surfaces that are perfectly smooth and differentiable. As a result the expressions in [1]–[4], which make perfect sense for manifolds \( C^0 \), \( n \geq 2 \), are woefully inadequate when it comes to determining the MFs for real data sets, which are grainy and quite often sparse. Below we describe how one can determine the MFs (and derived quantities) for surfaces derived from real data by interpolation. The interpolation scheme, which describes isodensity contours as triangulated surfaces, will be presented in detail in the next section.

We construct a polyhedral surface using an assembly of triangles in which every triangle shares its sides with each of its three neighbouring triangles.

(i) The area of such a triangulated surface is

\[
S = \sum_{i=1}^{N_T} S_i,
\]

where \( S_i \) is the area of the \( i \)-th triangle and \( N_T \) is the total number of triangles that compose a given surface.

(ii) The volume enclosed by this polyhedral surface is the summed contribution from \( N_T \) tetrahedra

\[
V = \sum_{i=1}^{N_T} V_i,
\]

where \( V_i \) is the volume of an individual tetrahedron whose base is a triangle on the surface. \( (n_i \vec{P}^i) \) is the scalar product between the outward-pointing normal \( \vec{n} \) to this triangle and the mean position vector of the three triangle vertices, for which the \( j \)-th component is given by

\[
\vec{P}^i = \frac{1}{2} \left( P^i_1 + P^i_2 + P^i_3 \right).
\]

The subscript \( i \) in (4) refers to the \( i \)-th tetrahedron, while the vectors \( P_1, P_2 \) and \( P_3 \) in (5) define the location of each of three triangle vertices defining the base of the tetrahedron relative to an (arbitrarily chosen) origin (Fig. 1). The vertices are ordered anticlockwise. This ensures that the contribution to the volume from tetrahedra whose base triangles lie on opposite sides of the origin subtract out. Thus for both possibilities we get the correct value for the enclosed volume (Fig. 1).

(iii) The extrinsic curvature of a triangulated surface is localized in the triangle edges. As a result the integrated mean curvature \( C \) is determined by the formula

\[
C = \frac{1}{2} \sum_{i,j} \ell_{ij} \cdot \phi_{ij} \epsilon,
\]

where \( \ell_{ij} \) is the edge common to adjacent triangles \( i \) and \( j \), and \( \phi_{ij} \) is the angle between the normals to these triangles \( \vec{n}_i \) and \( \vec{n}_j \),

\[
\cos \phi_{ij} = \hat{n}_i \cdot \hat{n}_j.
\]

The summation in (6) is carried out over all pairs of adjacent triangles. It should be noted that, for a completely general surface, the extrinsic curvature can be positive at some (convex) points and negative at some other (concave) points on the surface. To accommodate this fact one associates a number \( \epsilon = \pm 1 \) with every triangle edge in (6): \( \epsilon = 1 \) if the normals on adjacent triangles diverge away from the surface, indicating a locally convex surface, while \( \epsilon = -1 \) if the normals converge towards each other outside the surface, which is indicative of a concave surface. In the former case the centre of curvature of the surface lies within the surface body, whereas in the latter case the centre of curvature lies outside of the surface body (Fig. 2).

(iv) The genus of a triangulated closed polyhedral surface is given by the convenient expression

\[
G = 1 - \chi/2,
\]

\[
\chi = N_T - N_E + N_V,
\]

where \( \chi \) is the Euler characteristic of the triangulated surface, \( N_T \), \( N_E \) and \( N_V \) are, respectively, the total number of triangles, triangle edges and triangle vertices defining the surface.

(v) As demonstrated in Sahni et al. (1998) and Sathyaprakash et al. (1998) the ratios of MFs called ‘shapefinders’ provide us with an excellent measure of morphology. Therefore, in addition to determining MFs we shall also determine the shapefinders, \( T \) (thickness), \( B \) (breadth) and \( L \) (length), which are defined as follows:

\[
T = \frac{3V}{S}, \quad B = \frac{S}{C}, \quad L = \frac{C}{4\pi(G+1)},
\]

© 2003 RAS, MNRAS 343, 22–46
Figure 1. The volume enclosed by the surface is estimated by vectorially summing over the volumes of individual tetrahedra, each having its base at one of the triangles composing the surface and its apex at an arbitrarily chosen point that we call the origin. Note that the origin can lie both within as well as outside the surface. The left panel shows how SURFGEN estimates the volume when the origin lies outside the surface. In this case, the contribution (to the enclosed volume) from triangles falling in the same solid angle carries opposite sign for tetrahedra OABC and OFED. Clearly the volume of OFED must be subtracted from the volume OABC in order to give the true volume enclosed by the surface. Anticlockwise orientation of the vertices used in determining the normals of the base triangles helps us bring this about. The right panel shows the second possibility for which the two triangles ACB and DEF form tetrahedra whose fourth (common) vertex lies within the surface. In this case, the contributions from ODEF and OACB add to give the total volume. This figure can be seen in colour in the on-line version of the journal on Synergy.

Figure 2. Illustrated here is the algorithm to compute the local contribution to the integrated mean curvature: (a) \( \Delta ABC \) and \( \Delta CBD \) are two triangles sharing a side BC of length \( \ell \). The normals \( \hat{n}_1 \) and \( \hat{n}_2 \) to the two triangles lie in the planes orthogonal to BC. Their projection in the plane of the paper diverges from a point O inside the surface. Thus the triangles correspond to a convex neighbourhood and \( \epsilon = 1 \) in equation (6). The angle between the two triangles is denoted as \( \phi \). (b) Shown here are a similar pair of triangles, but with inverted sense of the normals, so that the sense of the surface is opposite to that shown in (a). We note that, in the present case, the two normals converge outside the surface, so that their projections in the plane of the paper would meet at \( O' \), which lies outside the surface. Thus this represents a concave neighbourhood, for which \( \epsilon = -1 \) in equation (6). The angle \( \phi \) represents the internal angle between \( \hat{n}_1 \) and \( \hat{n}_2 \).

where \( G = 0 \) for simply connected surfaces and \( G > 0 \) for multiply connected regions. The three shapefinders (describing an object) have dimensions of length and provide us with an estimate of the ‘extension’ of the object along each of the three spatial directions: \( T \) is the shortest and thus describes the characteristic thickness of the object; \( L \) is the longest and thus characterizes the length of the object; \( B \) is intermediate and can be associated with the breadth of the object. This simple interpretation obviously is relevant only for fairly simple shapes. For example, a triaxial ellipsoid has the values of \( T, B \) and \( L \) closely related to the lengths of the three principal axes: shortest, intermediate and longest respectively. It is worth stressing that \( L \) defined by equation (9) quantifies the characteristic length between holes which should be taken into account when interpreting the results (e.g. Table 3 below).

(vi) An excellent indicator of ‘shape’ is provided by the dimensionless shapefinder statistics

\[
P = \frac{B - T}{B + T}, \quad F = \frac{L - B}{L + B}
\]

where \( P \) and \( F \) are measures of planarity and filamentarity respectively (\( P, F \leq 1 \)). A sphere has \( P = F = 0 \), an ideal filament has \( P = 0, F = 1 \) while \( P = 1, F = 0 \) for an ideal pancake. Other
The third category are triangulated to model the surfaces.

interesting shapes include ‘ribbons’ for which $P \sim F \sim 1$. When combined with the genus measure, the triplet $\{P, F, G\}$ provides an example of shape-space, which incorporates information about the topology as well as the morphology of superclusters and voids.\(^3\)

Having presented an Ansatz to calculate MFs and shapefinders on a triangulated surface, we now discuss the surface generating algorithm (SURFGEN) which creates triangulated surfaces corresponding to isodensity contours evaluated at any desired threshold of the density field.

3 SURFACE GENERATING ALGORITHM (SURFGEN)

SURFGEN is a versatile and powerful prescription which allows us to generate and study surfaces whose physical origin can be quite varied and different. Intensity surfaces (isophotes), isotherms and isodensity surfaces constructed from 3D data provide examples of 2D that can be generated and studied using SURFGEN and the MFs discussed in the previous section.\(^4\) The present paper is the first of a series in which SURFGEN and MFs will be used in conjunction to make a detailed morphological study of the supercluster–void network in the Universe. SURFGEN will be applied to a density field $\rho(r)$ defined on a rectangular cubic lattice.

The SURFGEN algorithm is constructed as follows:

(i) A point particle distribution (from $N$-body simulations, galaxy catalogues) is used to reconstruct the density field on the vertices of a cubic lattice via a cloud-in-cell (CIC) approach. The lattice itself consists of a large number of closely packed elementary cubes. Any given cube is characterized by the value of the density at its eight vertices.

(ii) An appropriate density threshold $\rho_{TH}$ is chosen and lattice points at this threshold are found using linear interpolation. The lattice cells that involve such interpolated points along their sides are marked for triangulation, which is to be done in the next step.

(iii) SURFGEN determines a closed polyhedral surface at the threshold $\rho_{TH}$ by triangulating elementary cubes while simultaneously moving across the lattice. Triangulated elementary cubes are then put together to make up the full 2D triangulated surface.

The process of triangulation is based on the following observation. For any arbitrary threshold of the density, all elementary cubes of the density field fall into one of the following three classes:

(i) those which have all eight vertices below the density threshold (underdense cubes) – Fig. 3(I);

(ii) those which have all eight vertices above the density threshold (overdense cubes) – Fig. 3(II);

(iii) cubes having both overdense and underdense vertices (surface cubes) – Fig. 3(III).

When modelling clusters and superclusters, overdense cubes will be enclosed by our surface while underdense cubes will be excluded by it. (For voids, the situation will be reversed.) Contour surfaces at a prespecified density threshold will lie entirely within surface cubes. Thus the properties of surface cubes are vitally important for this surface-reconstruction exercise.

We now describe how SURFGEN constructs a surface at the desired threshold $\rho_{TH}$ by triangulating surface cubes. We work under the assumption that the underlying density field $\rho(r)$ is continuous, so that in moving from an overdense site to an underdense site (along a cube edge) we invariably encounter a point at which $\rho = \rho_{TH}$. SURFGEN classifies a given lattice cube in terms of the number of points (on the edges) where $\rho = \rho_{TH}$. The precise location of these points (found by interpolation) tells us exactly where our isodensity surface will intersect the cube. Since there are eight vertices and each vertex can be either overdense (1) or underdense (0), the number of possible configurations of the cube are $2^8 = 256$. Clearly the triangulation must be invariant upon interchanging ‘1’s with ‘0’s, and this reduces the number of independent configurations to 128. Upon invoking rotational symmetry this number further reduces to only 14. (For instance, although there are eight cubes that have a single overdense vertex, any two members of this family are related via the 3D rotation group. Therefore a single scheme of triangulation suffices to describe all members of this group.) These 14 configurations of triangulated cubes are shown in Fig. 4 (also see Lorenson & Cline 1987).

The surface intersecting a given surface cube forms quite clearly only a portion of the full isodensity surface which we are interested in constructing. By explicitly demanding that the density field be continuous across the isodensity surface, one joins triangles across the faces of neighbouring cubes, thus constructing the full continuous polyhedral surface. This then is the basic prescription that may be used to construct a closed triangulated surface at any desired value of the density threshold. However, a few comments about triangulation are in order: (i) The triangulation of the cubes having four contrasting vertices is not unique. To illustrate this, let us consider Fig. 5, where the last cube belonging to the middle row of Fig. 4

\[^3\]Non-geometrical shape statistics based on mass moments, etc., can give misleading results when applied to large-scale structure, as demonstrated in Sathyaprakash et al. (1998).

\[^4\]SURFGEN is a modified and extended version of the marching cubes algorithm (MCA) which is used in the field of medical imaging to render high-resolution images of internal organs by processing the output generated by X-ray tomography (Lorenson & Cline 1987).

---

**Figure 3.** Given a density field defined on a lattice, the elementary lattice cubes fall into one of the above three classes: (I) cubes that are completely underdense, (II) cubes that are completely overdense, and finally (III) cubes that have a few vertices overdense and the remaining vertices underdense. Cubes referring to the third category are triangulated to model the surfaces.
is reproduced in the left panel. We note that this cube is symmetric with respect to an interchange of underdense sites with overdense sites. As a result, we could just as well triangulate it according to the prescription shown in the right panel of Fig. 5. This degeneracy in the way a given cube can be triangulated is inherent also in the last two cubes of the last row of Fig. 4. Noticeably, at all such places where we encounter one of these cubes, one scheme of triangulation has to be preferred over the other, otherwise the surface could become discontinuous. Thus to create a unique triangulated surface one must triangulate such cubes in tandem with their neighbours, ensuring that the continuity of the surface is maintained. (ii) There are instances when two neighbouring cubes have a density configuration that does not lead to the complete closure of the given surface. The resulting surface has a hole (Fig. 6a) which must be filled by constructing two additional triangles which close the surface (Fig. 6b). The probability of hole formation is finite every time the common face shared by two cubes shows two vertices of type 1 along one diagonal and of type 0 along the other diagonal.
Thus, in order to generate a closed polyhedral surface at a given value of the density threshold $\rho = \rho_{TH}$, SURFGEN uses the 14 independent triangulations of an elementary lattice cube in conjunction with an algorithm which ensures that the surface is continuous. We should emphasize that surfaces generated in this manner need not be simply connected (see Fig. 7). Indeed, isodensity surfaces constructed at moderate density thresholds in $N$-body simulations as well as 3D galaxy surveys tend to show a large positive value for the genus (equation 8). This issue will be discussed in detail in forthcoming paper(s).

To enable an online calculation of the MFs, our surface modelling scheme also adheres to the following requirements:

(i) The vertices of all the triangles are stored in an anticlockwise order. This enforces a uniform prescription on normals which always point outward from the surface being modelled. Information regarding the directionality of normals is of great importance since it is used for calculating both the volume as well as the mean curvature.

(ii) For online calculation of the mean curvature, information concerning a given triangle must be supplemented with information concerning the triplet of triangles which are its neighbours. This allows us to determine unambiguously the contribution to the mean curvature from a local neighbourhood.

(iii) The total number of triangles, triangle edges and triangle vertices of a triangulated surface are counted, which enables us to determine its genus when surface construction is complete.

We might add that, although in this paper the density field is reconstructed from $N$-body particle positions by means of a CIC prescription, this is not a necessary prerequisite for SURFGEN, which is versatile enough to be used to triangulate isodensity contours of fields reconstructed using more elaborate techniques such as Delauney tesselations (Schaap & van de Weygaert 2000; van de Weygaert 2002), Wiener reconstruction or adaptive smoothing. As it turns out, these schemes tend to maintain the structural complexity in the system without diluting the patterns in an isotropic manner. However, since SURFGEN constructs a triangulated surface by ‘marching across neighbouring lattice cubes’ (Lorenson & Cline 1987), the density field must be specified on a cubic lattice for this approach to give meaningful results.5

To summarize, our surface modelling code incorporates the triangulation of the entire set of 256 possible configurations of a cube within a single scheme and makes possible online computation of the area, the volume, the integrated mean curvature and the genus of a triangulated surface. The code has been tested on a variety of standard density distributions and eikonal morphologies as well as on Gaussian random fields. In the next section, we present results based on this analysis.

4 RESULTS FOR STANDARD EIKONAL SURFACES

In order to test the accuracy of our Ansatz we generate triangulated surfaces whose counterpart continuum surfaces have known (analytically calculable) Minkowski functionals.

4.1 Spherical structures

In the first exercise, we consider spherically symmetric density distributions and generate surfaces of constant density for a variety of

---

5 We should emphasize, however, that the Ansatz developed in Section 2 for determining the MFs for a triangulated surface can easily be adapted to other triangulation schemes such as those described in O’Rourke (1998) and van de Weygaert (2002).
density thresholds starting from a chosen maximum radius down to grid size. A convenient density distribution for this exercise is
\[
\rho(i, j, k) = \begin{cases} 
\rho_0 / R, & (i, j, k) \neq (i_0, j_0, k_0), \\
\rho_0, & (i, j, k) = (i_0, j_0, k_0),
\end{cases}
\]
where \( R \) is the distance between \((i, j, k) \) and the centre,
\[
R = \sqrt{(i - i_0)^2 + (j - j_0)^2 + (k - k_0)^2}.
\]
Since the density field falls off as the inverse of the distance from the centre at \((i_0, j_0, k_0)\), any threshold \( \rho_{\text{TH}} \) is associated with a sphere of radius \( R = \rho_0 / \rho_{\text{TH}} \). Thus larger spheres correspond to surfaces of lower constant density in this model. We assume \( i_0 = j_0 = k_0 = 15 \) for the centre of the sphere in our numerical calculations.

Applying the shapefinders (9) and (10) to a sphere of radius \( R \) one finds the simple results \( T = B = L = R \) and \( F = P = 0 \). Fig. 8 compares the area, volume, mean curvature and shapefinders measured for a triangulated sphere against known analytical values for these quantities. This figure clearly demonstrates that exact values, and values computed using triangulation, match exceedingly well down to the very lowest scale.

We should mention that the genus evaluated using equation (8) is identically zero for spheres of all radii and for all possible deformations of an ellipsoid (to be discussed next). This provides an excellent independent endorsement of our methodology by demonstrating that our triangulations of a sphere and ellipsoid indeed result in closed, continuous surfaces.

4.2 Triaxial ellipsoid
A triaxial ellipsoid is an excellent shape with which to test a morphological statistic. This is because, depending upon the relative scales of the three axes, a triaxial ellipsoid can be oblate, prolate or spherical. We saw previously that MFs and shapefinders give extremely accurate results for spherical surfaces. We now demonstrate that this remains true even for surfaces that are highly planar or filamentary.

The parametric form of an ellipsoid having axes \( a, b, c \) and volume \( V = \frac{4}{3} \pi abc \) is
\[
r = a(\sin \theta \cos \phi) \hat{x} + b(\sin \theta \sin \phi) \hat{y} + c(\cos \theta) \hat{z},
\]
where \( 0 < \phi < 2\pi, 0 < \theta < \pi \). For the purposes of our study, we systematically deform a triaxial ellipsoid and study how its shape evolves in the process. When shrunk along a single axis a triaxial ellipsoid becomes planar. Simultaneous shrinking along a second axis makes it cigar-like or filamentary.

4.2.1 Oblate spheroids
We first study the accuracy of our triangulation scheme for planar configurations by considering oblate deformations of an ellipsoid. In this case two axes \( a \) and \( b \) are held fixed \( (a \sim b) \), while the third axis \( c \) is slowly shrunk leading to an increasingly planar surface. Our results for this case are compiled in two figures.

Fig. 9 shows MFs as they evolve with the dimensionless variable \( \epsilon \). [The exact results for the MFs that we quote are based on the analytical expressions for MFs given in Sahni et al. (1998). We refer the reader to that paper for more details.] Fig. 9

Figure 8. Minkowski functionals and dimensional shapefinders for a sphere. Note that the results for the MFs evaluated using our surface modelling scheme (SURFGEN) virtually coincide with the exact values.

© 2003 RAS, MNRAS 343, 22–46
clearly demonstrates that the values of MFs obtained using triangulation match the exact values to a remarkable degree of accuracy. Indeed, for a wide range in \(c/a\) the two distinctly different calculational algorithms give virtually indistinguishable results, thereby indicating that the triangulation Ansatz is, for all practical purposes, exact!

Fig. 10 shows the evolution of all the shapefinders, i.e. \(T, B, L, P\) and \(F\), together with the percentage errors in the estimation of dimensional shapefinders. We find that \(L\) is estimated to greatest accuracy with maximum error of \(\sim 0.4\) per cent, while \(T\) and \(B\) can be determined to an accuracy of \(\sim 0.8\) per cent. We further note (right panels of Fig. 10) that planarity grows from an initially low value \(\sim 0.0\) to a large final value \(\sim 0.4\) as the ellipsoid becomes increasingly oblate. Filamentarity, on the other hand, remains small at \(\sim 0.09\). Both filamentarity and planarity are determined to great accuracy by SURFGEN.

### 4.2.2 Prolate spheroids

Next we study prolate deformations of our oblate ellipsoid. We start with \(b \simeq a\), \(c \ll a\), and shrink the second axis \(b\) while keeping \(a\) and \(c\) fixed, so that finally \(c \sim b \ll a\) and our initially oblate ellipsoid becomes prolate. Our results are again summarized in two figures.
Fig. 11 reveals very good agreement between measured and true values of MFs, with the former tending to be slightly smaller than the latter.

Turning to the shapefinders, we find that, with the possible exception of extremely prolate figures, the shapefinders are remarkably well determined. Indeed, even the extremely prolate ellipsoid with axis ratio $b/a < 0.2$ has a largest error in $T$ of only $\sim 1.7$ per cent, while errors in $B$ and $L$ never exceed $\sim 1$ per cent (Fig. 12). This figure also shows the evolution of planarity ($P$) and filamentarity ($F$) as our ellipsoid becomes increasingly more prolate. In keeping with our expectation, $F$ steadily increases from a small initial value $\sim 0.1$ to $\sim 0.44$. Planarity drops from its large initial value to a small final value $P \sim 0.1$ and is slightly underestimated for $b/a \gtrsim 0.3$. We therefore conclude that both $P$ and $F$ are determined by SURFGEN to a sufficient accuracy for prolate ellipsoidal figures.

4.3 The torus and its deformations

Next we extend our analysis to manifolds that are multiply connected by considering the deformations of a torus with an elliptical cross-section which we shall refer to as an elliptical torus. A torus is an important surface on which to test SURFGEN for two reasons: it is multiply connected and, unlike an ellipsoid, it contains regions that are convex (on its outside) as well as concave (on its inside). The elliptical torus can be described by three parameters $a$, $b$, and $c$.
Table 1. Minkowski functionals for some extreme deformations of an elliptical torus. $\Delta C/C$, $\Delta S/S$ and $\Delta V/V$ give the percentage error in the determination of the MFs using SURFGEN. An accuracy of better than 1 per cent is achieved for the three MFs, curvature ($C$), surface area ($S$) and volume ($V$), while the genus is determined exactly.

<table>
<thead>
<tr>
<th>Surface</th>
<th>$b$, $a$, $c$</th>
<th>$C$ (per cent)</th>
<th>$\Delta C/C$</th>
<th>$S$ (per cent)</th>
<th>$\Delta S/S$</th>
<th>$V$ (per cent)</th>
<th>$\Delta V/V$</th>
<th>Genus</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sphere-with-hole</td>
<td>(40.0, 37.9, 37.9)</td>
<td>$7.93 \times 10^2$</td>
<td>+0.47</td>
<td>$5.98 \times 10^4$</td>
<td>−0.01</td>
<td>$1.13 \times 10^6$</td>
<td>−0.001</td>
<td>1</td>
</tr>
<tr>
<td>Ribbon</td>
<td>(140.0, 19.9, 1.99)</td>
<td>$2.8 \times 10^3$</td>
<td>+0.03</td>
<td>$7.1 \times 10^4$</td>
<td>−0.42</td>
<td>$1.1 \times 10^5$</td>
<td>−0.830</td>
<td>1</td>
</tr>
<tr>
<td>Pancake</td>
<td>(60.0, 57.9, 3.86)</td>
<td>$1.2 \times 10^3$</td>
<td>+0.40</td>
<td>$8.8 \times 10^4$</td>
<td>−0.18</td>
<td>$2.7 \times 10^5$</td>
<td>+0.590</td>
<td>1</td>
</tr>
<tr>
<td>Filament</td>
<td>(50.02, 3.52, 3.52)</td>
<td>$9.92 \times 10^2$</td>
<td>+0.52</td>
<td>$6.95 \times 10^3$</td>
<td>+0.18</td>
<td>$1.21 \times 10^4$</td>
<td>−0.580</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 2. SURFGEN-determined values of the shapefinders $T$ (thickness), $B$ (breadth), $L$ (length), $P$ (planarity) and $F$ (filamentarity) describing extreme deformations of an elliptical torus. Also given alongside are the percentage errors in their estimation.

<table>
<thead>
<tr>
<th>Surface</th>
<th>$T$ (per cent)</th>
<th>$B$ (per cent)</th>
<th>$L$ (per cent)</th>
<th>$P$ (per cent)</th>
<th>$F$ (per cent)</th>
<th>Genus</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sphere-with-hole</td>
<td>56.86</td>
<td>0.02</td>
<td>75.44</td>
<td>−0.48</td>
<td>31.57</td>
<td>0.48</td>
</tr>
<tr>
<td>Ribbon</td>
<td>4.67</td>
<td>1.08</td>
<td>25.63</td>
<td>−0.47</td>
<td>109.96</td>
<td>0.03</td>
</tr>
<tr>
<td>Pancake</td>
<td>9.09</td>
<td>0.74</td>
<td>74.19</td>
<td>−0.59</td>
<td>47.11</td>
<td>0.41</td>
</tr>
<tr>
<td>Filament</td>
<td>5.23</td>
<td>−0.95</td>
<td>7.01</td>
<td>−0.43</td>
<td>39.49</td>
<td>0.46</td>
</tr>
</tbody>
</table>

are its radii of curvature in two mutually orthogonal directions. The elliptical torus reduces to the more familiar circular torus when $a = c$. We choose to work with the elliptical torus because changing the values of $a$, $b$, $c$ can give rise to a large variety of surfaces all of which (by definition) are multiply connected but which have very different shapes. Thus our surface modelling scheme SURFGEN can be put to a rigorous test. The parametric form for the elliptical torus is

$$r = (b + c \sin \phi) \cos \theta + (b + c \sin \phi) \sin \theta \gamma + a \cos \phi \hat{z},$$

(14)

where $a$, $c < b$, $0 \leq \phi < 2\pi$ and $\theta < 2\pi$. We shall compare results due to SURFGEN with the exact analytical results for four kinds of tori -- a nearly spheroidal torus, a ribbon, a pancake, and a filament. Tables 1 and 2 refer to these four surfaces; relevant figures can be found in Sahni et al. (1998). Table 1 shows the estimated values of the MFs and the percentage error in their estimation. Table 2 shows all the shapefinders for these surfaces.\(^6\) This table also shows the percentage errors in the estimation of the shapefinders.

We should emphasize that the genus estimated by SURFGEN for all these deformations has the correct value of unity. This gives an independent check on the self-consistency of the triangulation scheme since the absence of even a single triangle out of several thousand would create an artificial ‘hole’ in our surface and give the wrong value for the genus -- a situation that has never occurred for any of the several dozen deformations of either the ellipsoid or the torus.

Our results clearly demonstrate that the surface modelling scheme SURFGEN provides values for the MFs which are in excellent agreement with exact analytical formulae. We have further tested SURFGEN against a set of Gaussian random field samples. Owing to a subtle issue of the boundary conditions involved in the corresponding test, we present the corresponding results in the Appendix. For the time being, it suffices to make a note that our Ansatz for studying isodensity contours and their morphology works exceedingly well for both simply connected as well as multiply connected surfaces and on Gaussian random fields.

\(^6\) It should be noted that the definition of the third shapefinder $L$ is modified from that reported in Sahni et al. (1998). The new definition of $L$ is as given in equation (9). The newly adopted definition of $L$ incorporates treatment of simply connected and multiply connected surfaces within a single scheme.

In the next section, we apply SURFGEN to cosmological $N$-body simulations and give an example of the study of the morphology in three models of structure formation: $\Lambda$CDM, $\tau$CDM and SCDM.

5 APPLICATIONS: MORPHOLOGY OF THE VIRGO SIMULATIONS

As a first application of the SURFGEN code, we report here a morphological study of cosmological simulations. We use dark matter distributions of three cosmological models simulated by the Virgo Consortium. The models are (1) a flat model with $\Omega_m = 0.3$ (LCDM), and two models with $\Omega_m = 1$, namely (2) one with the standard CDM power spectrum (SCDM) and (3) a flat model with the same power spectrum as the $\Omega_m = 0.3$ LCDM model (also referred to as $\tau$CDM). The shape parameter $\Gamma (= \Omega_m)h$ for SCDM is 0.5, whereas for the other two models $\Gamma = 0.21$. The amplitude of the power spectrum in all models is set so as to reproduce the observed abundance of rich galaxy clusters at the present epoch. A detailed discussion of the cosmological parameters and simulations can be found in Jenkins et al. (1998).

The data consist of $256^3$ particles in a box of size $239.5h^{-1}$ Mpc. In order to carry out a detailed morphological study, we first need to smooth the data and recover the underlying density field. The issue of density field reconstruction has been well explored by several groups studying the topology of the large-scale structure. Here we follow the smoothing technique used by Springel et al. (1998), which they adopted for their preliminary topological analysis of the Virgo simulations. We fit a $128^3$ grid on to the box. Thus, the size of each cell is $1.875 h^{-1}$ Mpc. Since there are on average eight particles per cell, the smoothing is not too sensitive to shot noise. The samples are fairly rich and we can smooth the fields in two steps. In the first, we apply a cloud-in-cell (CIC) smoothing technique to construct a density field on the grid. Next we smooth this field with a Gaussian kernel which offers us an extra smoothing length-scale.

There are various criteria to fix this smoothing scale which depend on the relative sparseness/richness of the samples. To avoid any discreteness effect (so that there are no regions in the survey where density is undefined), the simplest criterion for sparse samples has been to set $\lambda = 2.5\ell_g$, where $\ell_g$ is the grid size. There are other sophisticated criteria to deal with sparse samples that are tested
and reported in the literature, which need not be reiterated here on account of our samples being rich. We are left with two choices: either to smooth the fields at a scale comparable to \( r_0 \), the correlation scale, or to probe the morphology at smaller length-scales with an interest to study the regularity in the behaviour of the MFs; of these, the latter is more promising. The average interparticle distance in the present case is \( \sim 0.94 h^{-1} \) Mpc, which is far outweighed by the grid size and the correlation length. Hence, we adopt a modest scale of smoothing, \( \lambda = 2 h^{-1} \) Mpc, which is small compared to usual standards as well as enough to provide the necessary smoothing. The Gaussian kernel for smoothing that we adopt here is

\[
W(r) = \frac{1}{(2\pi)^{3/2} \lambda^3} \exp \left( -\frac{r^2}{2\lambda^2} \right).
\]  

In this section we first study the global MFs for all three models. Next we investigate the morphology of the percolating supercluster network. This is followed by a statistical study of the morphology of smaller structures. We summarize our results in the next section.

5.1 Global Minkowski functionals

We scan the density fields at 100 values of the density threshold \( \rho_{\text{TH}} \), all equispaced in the filling factor

\[
FF_V = \int \Theta(\rho - \rho_{\text{TH}}) d^3 x,
\]

where \( \Theta(x) \) is the Heaviside Theta function. \( FF_V \) measures the volume fraction in regions that satisfy the ‘cluster’ criterion \( \rho_{\text{cluster}} \geq \rho_{\text{TH}} \) at a given density threshold \( \rho_{\text{TH}} \). In the following, we use \( FF_V \) as a parameter to label the density contours.

At each level of the density field (labelled by \( FF_V \)), we construct a catalogue of clusters (overdense regions) based on a friends-of-friends (FOF) algorithm. Next we (i) run the SURFGEN code on each of these clusters to model surfaces for each of them and (ii) determine the MFs for all clusters at the given threshold (these are referred to in the literature as partial MFs). Global MFs are partial MFs summed over all clusters. Thus, at each level of the density, we compute the partial MFs first and then the global MFs. Our plots of global MFs as functions of \( FF_V \) are shown in Fig. 13. This figure shows some interesting features. The volume curve is the same for all models and grows linearly with the volume fraction; this is simply a restatement of the definition of ‘filling factor’.

The clusters/superclusters discussed in the present paper are defined as connected overdense regions lying above a prescribed density threshold. Owing to the large smoothing scale adopted, the overdensity in clusters ranges from \( \delta \sim 1 \) to \( \delta \sim 10 \), which makes them more extended (\( \gtrsim \) few Mpc) and less dense than the galaxy clusters in (for instance) the Abell catalogue.

---

Figure 13. Global MFs for \( \Lambda \)CDM, \( \tau \)CDM and SCDM are shown as functions of the volume fraction (\( \equiv \) volume filling factor, \( FF_V \)). The three cosmological models have appreciably different morphology and hence can be distinguished from one another on the basis of morphological measures. This figure can be seen in colour in the on-line version of the journal on Synergy.
Notice that the amplitude of the remaining three MFs (area, curvature and genus) is substantially greater in SCDM than in \(\Lambda\)CDM, with \(\tau\)CDM falling between the two. This could mean that large-scale structure is much more ‘spongy’ in SCDM with percolating structures in this model showing many more ‘holes’ (or tunnels) and resulting in a large value for the genus. A relative shift to the left of the position of the peak of the genus curve in \(\Lambda\)CDM from FFV = 0.5 is indicative of the bubble shift, implying more clumpiness in this model compared to the other two models. It should be noted that, for a Gaussian random field, the peak occurs at FFV = 0.5. Gravitational clustering modulates the genus curve by lowering its peak and producing a shift to the left (the ‘bubble shift’) or to right (the ‘meatball shift’) depending upon the other model parameters. The relatively large value of the surface area and the curvature may indicate that moderately overdense superclusters have many more ‘twists and turns’ in SCDM than in either \(\tau\)CDM or \(\Lambda\)CDM, at identical values of the filling factor. In all models the genus curve has a large negative value at large values of the volume fraction. The reason for this is that the percolating supercluster, at extremely low density thresholds (high FFV), occupies most of the volume. Voids exist as small isolated bubbles in this vast supercluster and lead to a large negative genus whose value is of the same order as the total number of voids. In the opposite case, at very high density thresholds associated with small filling fractions, we probe the morphology of isolated and simply connected clusters. Since the genus for these clusters vanishes, the global topology of large-scale structure at high density thresholds generically approaches zero for all the models. The vertical lines in the lower right panel show the percolation threshold in SCDM (dashed), \(\tau\)CDM (dot-dashed) and \(\Lambda\)CDM (solid) which occurs at intermediate density thresholds corresponding to density contrast \(\delta \sim 1\). It is clear from this figure that the percolating supercluster has a relatively simple topology at the onset of percolation. The rapid increase in genus value as one moves to lower density thresholds (larger FFV) reflects the progressive increase in the ‘sponge-like’ topology of the percolating supercluster, which is more marked in SCDM than in \(\Lambda\)CDM. As in the case of the percolation statistic (Dominik & Shandarin 1992), the difference between models is brought out much better if, instead of plotting the MFs against the volume fraction (equivalently ‘volume filling factor’), we choose the ‘mass fraction’ defined as FFM = \(M_{\text{total}}^{-1} \int \rho(\delta \geq \rho_{TH}) \, dV\) (see Fig. 14). [The ‘mass fraction’ can be thought of as the ‘volume or mass filling factor’ in (initial) Lagrangian space, while the ‘volume fraction’ is the filling factor in (final) Eulerian space.] By employing this parameter for studying the global MFs, we essentially probe the morphology of the isodensity contours (which may refer to different thresholds of density or density contrast, but) which enclose the same fraction of the total mass. Since clustering tends to pack up the mass into progressively smaller regions of space, such a study connects to aspects of gravitational clustering in a direct way. Thus, we see that at all the values

\[ \frac{\delta}{\Lambda} \]

\[ \frac{\delta}{\Lambda} \]

\[ \frac{\delta}{\Lambda} \]

\[ \frac{\delta}{\Lambda} \]

\[ \frac{\delta}{\Lambda} \]

\[ \frac{\delta}{\Lambda} \]

\[ \frac{\delta}{\Lambda} \]

\[ \frac{\delta}{\Lambda} \]

\[ \frac{\delta}{\Lambda} \]

\[ \frac{\delta}{\Lambda} \]

\[ \frac{\delta}{\Lambda} \]

\[ \frac{\delta}{\Lambda} \]

\[ \frac{\delta}{\Lambda} \]

\[ \frac{\delta}{\Lambda} \]

\[ \frac{\delta}{\Lambda} \]

\[ \frac{\delta}{\Lambda} \]

\[ \frac{\delta}{\Lambda} \]

\[ \frac{\delta}{\Lambda} \]

\[ \frac{\delta}{\Lambda} \]

\[ \frac{\delta}{\Lambda} \]

\[ \frac{\delta}{\Lambda} \]

\[ \frac{\delta}{\Lambda} \]

\[ \frac{\delta}{\Lambda} \]

\[ \frac{\delta}{\Lambda} \]

\[ \frac{\delta}{\Lambda} \]

\[ \frac{\delta}{\Lambda} \]

\[ \frac{\delta}{\Lambda} \]

\[ \frac{\delta}{\Lambda} \]

\[ \frac{\delta}{\Lambda} \]

\[ \frac{\delta}{\Lambda} \]

\[ \frac{\delta}{\Lambda} \]

\[ \frac{\delta}{\Lambda} \]

\[ \frac{\delta}{\Lambda} \]

\[ \frac{\delta}{\Lambda} \]

\[ \frac{\delta}{\Lambda} \]

\[ \frac{\delta}{\Lambda} \]

\[ \frac{\delta}{\Lambda} \]

\[ \frac{\delta}{\Lambda} \]

\[ \frac{\delta}{\Lambda} \]

\[ \frac{\delta}{\Lambda} \]

\[ \frac{\delta}{\Lambda} \]

\[ \frac{\delta}{\Lambda} \]

\[ \frac{\delta}{\Lambda} \]

\[ \frac{\delta}{\Lambda} \]

\[ \frac{\delta}{\Lambda} \]

\[ \frac{\delta}{\Lambda} \]

\[ \frac{\delta}{\Lambda} \]

\[ \frac{\delta}{\Lambda} \]

\[ \frac{\delta}{\Lambda} \]

\[ \frac{\delta}{\Lambda} \]

\[ \frac{\delta}{\Lambda} \]

\[ \frac{\delta}{\Lambda} \]

\[ \frac{\delta}{\Lambda} \]

\[ \frac{\delta}{\Lambda} \]

\[ \frac{\delta}{\Lambda} \]

\[ \frac{\delta}{\Lambda} \]

\[ \frac{\delta}{\Lambda} \]

\[ \frac{\delta}{\Lambda} \]

\[ \frac{\delta}{\Lambda} \]

\[ \frac{\delta}{\Lambda} \]

\[ \frac{\delta}{\Lambda} \]

\[ \frac{\delta}{\Lambda} \]

\[ \frac{\delta}{\Lambda} \]

\[ \frac{\delta}{\Lambda} \]

\[ \frac{\delta}{\Lambda} \]

\[ \frac{\delta}{\Lambda} \]

\[ \frac{\delta}{\Lambda} \]

\[ \frac{\delta}{\Lambda} \]

\[ \frac{\delta}{\Lambda} \]

\[ \frac{\delta}{\Lambda} \]

\[ \frac{\delta}{\Lambda} \]

\[ \frac{\delta}{\Lambda} \]

\[ \frac{\delta}{\Lambda} \]

\[ \frac{\delta}{\Lambda} \]

\[ \frac{\delta}{\Lambda} \]

\[ \frac{\delta}{\Lambda} \]

\[ \frac{\delta}{\Lambda} \]

\[ \frac{\delta}{\Lambda} \]

\[ \frac{\delta}{\Lambda} \]

\[ \frac{\delta}{\Lambda} \]

\[ \frac{\delta}{\Lambda} \]

\[ \frac{\delta}{\Lambda} \]

\[ \frac{\delta}{\Lambda} \]

\[ \frac{\delta}{\Lambda} \]

\[ \frac{\delta}{\Lambda} \]

\[ \frac{\delta}{\Lambda} \]

\[ \frac{\delta}{\Lambda} \]

\[ \frac{\delta}{\Lambda} \]

\[ \frac{\delta}{\Lambda} \]

\[ \frac{\delta}{\Lambda} \]

\[ \frac{\delta}{\Lambda} \]

\[ \frac{\delta}{\Lambda} \]

\[ \frac{\delta}{\Lambda} \]

\[ \frac{\delta}{\Lambda} \]

\[ \frac{\delta}{\Lambda} \]

\[ \frac{\delta}{\Lambda} \]
Figure 15. Here we study the evolution of the volume fraction and the mass fraction with respect to the density contrast for all three models. ΛCDM shows maximum density contrast. SCDM and τCDM form another class of models which have similar density contrasts but smaller than ΛCDM. The vertical line across both panels refers to the mean density, which can be used as a marker to study the mass fraction that is contained in the overdense volume (the volume occupied by regions above the mean density). We note that the overdense volume in ΛCDM is ∼25 per cent, which is about 10 per cent smaller than that occupied by the other two models. At the same time, the mass that this volume encloses is ∼67 per cent and is about 5 per cent larger than that enclosed by the overdense volume in the other two models. This figure can be seen in colour in the online version of the journal on Synergy.

5.2 Cluster abundance and percolation

In this subsection we first discuss the percolation properties of the three density fields being studied. In this context we study how the total number of clusters and the fractional volume in the largest cluster vary as we scan through a set of density levels corresponding to equispaced fractions of volume and total mass. Fig. 15 shows the volume fraction and mass fraction as a function of the density contrast. It is interesting to note that SCDM and τCDM show the same pattern of behaviour, which differs from ΛCDM. It is further noted that at high density thresholds (δ ≳ 1) the filling factor (at a constant density threshold) is greater in ΛCDM than in SCDM or τCDM, while exactly the reverse is true for underdense regions. This figure, which relates FF₆₄,₉ to δ, serves as a reference point to all our subsequent percolation studies. Fig. 15 also shows that, relative to other models, more mass occupies less space in ΛCDM. Thus almost 67 per cent of the total mass in the ΛCDM universe resides in just 25 per cent of the volume. (In the case of τCDM and SCDM, ∼62 per cent of the mass occupies ∼34 per cent of the volume.)

A statistical pair that helps to quantify the geometry of large-scale structure and its morphology are the ‘number of clusters statistic’ (NCS) and the ‘largest cluster statistic’ (LCS); both are shown in Fig. 16. The NCS shows cluster abundance as a function of the density contrast (lower left), mass fraction (lower middle) and volume fraction (lower right). The LCS measures the fractional cluster volume occupied by the largest cluster:

\[ LCS = \frac{V_{LC}}{\sum_i V_i}, \]

where \( V_i \) is the volume of the \( i \)th cluster in the sample and \( V_{LC} \) is the volume occupied by the largest cluster. Summation is over all clusters evaluated at a given threshold of the density and includes the largest cluster. In Fig. 16 LCS is shown as a function of the density contrast.
density contrast (upper left), mass fraction (upper middle) and volume fraction (upper right).

The square, triangle and circle in NCS denote critical values of the parameters \( \delta = \delta_{\text{cluster max}} \), \( F_{\text{FV}} = F_{\text{FF cluster max}} \), and \( F_{\text{FM}} \) at which the cluster abundance peaks in a given model. Similar symbols on the LCS curves denote values of \( \delta = \delta_{\text{perc}} \), \( F_{\text{FV}} = F_{\text{FF perc}} \) and \( F_{\text{FM}} \) at which the largest cluster first spans across the box in at least one direction. This is commonly referred to as the ‘percolation threshold’. It is important to note that the number of clusters in SCDM is considerably greater than the number of clusters in \( \Lambda \)CDM at most density thresholds. It is also interesting that percolation takes place at higher values of the density contrast (and correspondingly lower values of the volume filling factor) in the case of \( \Lambda \)CDM, \( \delta_{\text{perc}} \approx 2.3 \) for \( \Lambda \)CDM, \( \delta_{\text{perc}} \approx 1 \) for SCDM and \( \tau \)CDM. Furthermore, while in the case of \( \Lambda \)CDM, \( \delta_{\text{cluster max}} \approx \delta_{\text{perc}} \), in the case of SCDM and \( \tau \)CDM, \( \delta_{\text{cluster max}} \lesssim \delta_{\text{perc}} \). Thus in the latter two models, as the density threshold is lowered, the cluster abundance initially peaks, then, as the threshold is lowered further, neighbouring clusters merge to form the percolating supercluster. In \( \Lambda \)CDM, on the other hand, the threshold at which the cluster abundance peaks also signals the formation of the percolating supercluster. Having said this we would like to add a word of caution: the analysis of one realization alone does not allow us to assess reliably the statistical fluctuations in the threshold of percolation. Therefore the final conclusion can be drawn only after a study of many realizations of each model.

Clearly both the LCS threshold \( \delta_{\text{perc}} \) and the NCS threshold \( \delta_{\text{cluster max}} \) contain important information and most of our subsequent description of supercluster morphology will be carried out at one of these two thresholds. Table 3 shows the MFs and associated shapefinders for the 10 largest (most voluminous) superclusters compiled at the percolation threshold for the three cosmological models \( \Lambda \)CDM, \( \tau \)CDM and SCDM. Table 3 also lists values for the planarity and filamentarity for these superclusters along with the mass that these enclose and their genus value.

Our \((239.5 h^{-1} \text{Mpc})^3 \) \( \Lambda \)CDM universe contains 1334 clusters and superclusters at the percolation threshold. Of these the smallest is quasi-spherical with a radius of a few megaparsecs while the largest is extremely filamentary and percolates through the entire simulation box (see Fig. 17). From Figs 16 and 17 we find that the percolating \( \Lambda \)CDM supercluster is a slim but massive object.\(^{10}\) It contains 4.5 per cent of the total mass in the universe yet occupies only 0.6 per cent of the total volume. In SCDM the percolating supercluster contains 4.4 per cent of the total mass and occupies 1.2 per cent of the total volume, i.e. a similar mass occupies twice as much volume than in the \( \Lambda \)CDM. The \(
\sim 10^7 \) \( \Lambda \)CDM clusters with \( \delta \gtrsim \delta_{\text{perc}} \) occupy 4.4 per cent of the total volume and contain 33 per cent of the total mass in their central half.

\(^{10}\)Surface visualization is a difficult task especially if we wish to follow tunnels through superclusters to check whether our visual impression of the genus agrees with that calculated using the Euler formula (8). MATLAB has been used for surface plotting and the ‘reality’ of tunnels is ascertained by rotating surfaces and by viewing the supercluster at various angles.
the universe. Gravitational clustering thus ensures that most of the mass in the $\Lambda$CDM universe is distributed in coherent filamentary regions which occupy very little volume but contain much of the mass. Of the 1334 $\Lambda$CDM clusters and superclusters identified at the percolation threshold, the 10 most voluminous are extremely filamentary and contain close to 40 per cent of the total cluster mass. (Of the 1334 regions which occupy very little volume but contain much of the universe, gravitational clustering thus ensures that most of the clusters, or SCDM; see Fig. 16.) Another interesting feature of gravitational clustering is that, although the volume occupied by clusters and superclusters identified at the percolation threshold is 4.5 × 10$^{4}$ M$_{\odot}$, it should be noted that the interpretation of $L$ as the ‘linear length’ of a supercluster can be misleading for the case of superclusters having a large genus. In this case $L \times (G + 1)$ provides a more realistic estimate of supercluster length since it allows for its numerous twists and turns. The morphology of the objects is conveyed through their planarity $P$ and filamentarity $F$. We note that the most voluminous and massive structures are highly filamentary in all the models. The largest supercluster is most filamentary in the case of $\Lambda$CDM and least so in the case of $r$CDM.

<table>
<thead>
<tr>
<th>Model</th>
<th>Mass ($M_{\odot}$)</th>
<th>Volume ($h^{-1}$ Mpc)$^{3}$</th>
<th>Area ($h^{-1}$ Mpc)$^{2}$</th>
<th>Curvature ($h^{-1}$ Mpc)</th>
<th>Genus</th>
<th>Shapefinders $T$ ($h^{-1}$ Mpc)</th>
<th>Shapefinders $B$ ($h^{-1}$ Mpc)</th>
<th>Shapefinders $L$ ($h^{-1}$ Mpc)</th>
<th>$P$</th>
<th>$F$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Lambda$CDM</td>
<td>$\delta = 2.31$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$1.19 \times 10^{17}$</td>
<td>$1.37 \times 10^{5}$</td>
<td>$7.57 \times 10^{4}$</td>
<td>$1.03 \times 10^{4}$</td>
<td>19</td>
<td>$5.42$</td>
<td>$7.36$</td>
<td>40.94</td>
<td>0.15</td>
<td>0.70</td>
<td></td>
</tr>
<tr>
<td>$r$CDM</td>
<td>$\delta = 1.37$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$1.66 \times 10^{17}$</td>
<td>$1.70 \times 10^{4}$</td>
<td>$9.89 \times 10^{4}$</td>
<td>$1.44 \times 10^{4}$</td>
<td>20</td>
<td>$5.14$</td>
<td>$6.88$</td>
<td>54.57</td>
<td>0.14</td>
<td>0.78</td>
<td></td>
</tr>
<tr>
<td>$\delta = 1.19$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$4.15 \times 10^{16}$</td>
<td>$4.24 \times 10^{4}$</td>
<td>$2.60 \times 10^{4}$</td>
<td>$4.07 \times 10^{4}$</td>
<td>2</td>
<td>$4.90$</td>
<td>$6.38$</td>
<td>107.94</td>
<td>0.13</td>
<td>0.89</td>
<td></td>
</tr>
<tr>
<td>$3.43 \times 10^{16}$</td>
<td>$3.52 \times 10^{4}$</td>
<td>$2.04 \times 10^{4}$</td>
<td>$2.93 \times 10^{4}$</td>
<td>3</td>
<td>$5.19$</td>
<td>$6.96$</td>
<td>58.27</td>
<td>0.15</td>
<td>0.79</td>
<td></td>
</tr>
<tr>
<td>$2.02 \times 10^{16}$</td>
<td>$1.95 \times 10^{4}$</td>
<td>$1.07 \times 10^{4}$</td>
<td>$1.58 \times 10^{4}$</td>
<td>3</td>
<td>$5.47$</td>
<td>$7.21$</td>
<td>29.47</td>
<td>0.14</td>
<td>0.61</td>
<td></td>
</tr>
<tr>
<td>$1.82 \times 10^{16}$</td>
<td>$1.94 \times 10^{4}$</td>
<td>$1.12 \times 10^{4}$</td>
<td>$1.56 \times 10^{4}$</td>
<td>2</td>
<td>$5.19$</td>
<td>$7.08$</td>
<td>42.05</td>
<td>0.15</td>
<td>0.71</td>
<td></td>
</tr>
<tr>
<td>$1.61 \times 10^{16}$</td>
<td>$1.60 \times 10^{4}$</td>
<td>$9.64 \times 10^{3}$</td>
<td>$1.39 \times 10^{4}$</td>
<td>3</td>
<td>$4.97$</td>
<td>$6.95$</td>
<td>27.58</td>
<td>0.17</td>
<td>0.60</td>
<td></td>
</tr>
<tr>
<td>$1.34 \times 10^{16}$</td>
<td>$1.36 \times 10^{4}$</td>
<td>$7.59 \times 10^{3}$</td>
<td>$1.08 \times 10^{4}$</td>
<td>2</td>
<td>$5.37$</td>
<td>$7.05$</td>
<td>28.59</td>
<td>0.14</td>
<td>0.60</td>
<td></td>
</tr>
<tr>
<td>$1.30 \times 10^{16}$</td>
<td>$1.34 \times 10^{4}$</td>
<td>$8.07 \times 10^{3}$</td>
<td>$1.25 \times 10^{4}$</td>
<td>2</td>
<td>$4.97$</td>
<td>$6.47$</td>
<td>33.07</td>
<td>0.13</td>
<td>0.67</td>
<td></td>
</tr>
<tr>
<td>$1.29 \times 10^{16}$</td>
<td>$1.25 \times 10^{4}$</td>
<td>$7.11 \times 10^{3}$</td>
<td>$9.98 \times 10^{3}$</td>
<td>2</td>
<td>$5.27$</td>
<td>$7.12$</td>
<td>26.50</td>
<td>0.15</td>
<td>0.58</td>
<td></td>
</tr>
<tr>
<td>$1.30 \times 10^{16}$</td>
<td>$1.21 \times 10^{4}$</td>
<td>$6.87 \times 10^{3}$</td>
<td>$9.88 \times 10^{3}$</td>
<td>2</td>
<td>$5.27$</td>
<td>$6.95$</td>
<td>26.22</td>
<td>0.14</td>
<td>0.58</td>
<td></td>
</tr>
</tbody>
</table>

© 2003 RAS, MNRAS 343, 22–46

5.3 Cluster morphology

Cosmic density fields contain a wealth of information. As demonstrated in Fig. 16, at density thresholds significantly lower than $\delta_{\text{perc}}$, most clusters merge to form a single percolating supercluster. On the other hand, the slightly larger NCS density contrast, $\delta_{\text{cluster max}}$, provides an excellent threshold at which to study the morphology of individual objects since it is precisely at $\delta_{\text{cluster max}}$ that the cluster abundance peaks. We study the morphology and topology of large-scale structure in a two-fold manner: (i) properties of all

Table 3. The 10 most voluminous superclusters (determined at the percolation threshold) are listed with the mass they enclose, their associated MFs (volume, area, curvature and genus) and shapefinders $T$, $B$, $L$, $P$ and $F$. The first row in each cosmological model describes the percolating supercluster and appears in bold. It is interesting that in all three cosmological models the top 10 superclusters contain roughly 40 per cent of the total mass in overdense regions with $\delta \geq \delta_{\text{perc}}$. The precise numbers are 40 per cent for $\Lambda$CDM, 37.8 per cent for SCDM and 45.6 per cent for $r$CDM. It should also be noted that the mass of a typical supercluster in $\Lambda$CDM is somewhat smaller than that in the other two models mainly on account of the fact that the adopted particle mass in $\Lambda$CDM simulations is smaller than in simulations of SCDM and $r$CDM. (The particle mass in $\Lambda$CDM is 6.80 × 10$^{10}$ h$^{-1}$ M$_{\odot}$ and that in SCDM and $r$CDM is 2.27 × 10$^{11}$ h$^{-1}$ M$_{\odot}$; Jenkins et al. 1998). It should be noted that the interpretation of $L$ as the ‘linear length’ of a supercluster can be misleading for the case of superclusters having a large genus. In this case $L \times (G + 1)$ provides a more realistic estimate of supercluster length since it allows for its numerous twists and turns. The morphology of the objects is conveyed through their planarity $P$ and filamentarity $F$. We note that the most voluminous and massive structures are highly filamentary in all the models. The largest supercluster is most filamentary in the case of $\Lambda$CDM and least so in the case of $r$CDM.
superclusters are analysed at one of the two thresholds, \( \delta_{\text{perc}} \) and \( \delta_{\text{cluster max}} \); (ii) the morphology of the largest supercluster is extensively probed as a function of the density contrast.

Fig. 18 shows the evolution of the morphology and topology of the largest cluster as we scan through different threshold levels of the \( \Lambda \)CDM density field. The density contrast, volume fraction and genus for a few of these levels have been labelled and are given in the top-right corner of the figure. We note that at high density thresholds the largest cluster occupies a very small fraction of the total volume and its shape is characterized by significant filamentarity \((\sim 0.67)\) and negligible planarity \((\sim 0)\). We note a sharp increase in the filamentarity of the largest cluster as we approach the percolation threshold (shown in the figure as a solid triangle). At smaller values of the density contrast, \( \delta \lesssim 3 \), supercluster filamentarity rapidly drops while its planarity considerably increases. The drop in filamentarity of the supercluster is accompanied by a growth in its complexity, with the result that, at moderately low thresholds \( (\delta \sim 0.5) \), the percolating supercluster can contain several hundred tunnels. At very low thresholds \( (\delta \lesssim 0.5) \) the percolating supercluster is an isotropic object possessing negligible values of both planarity \((P)\) and filamentarity \((F)\).

Since \( \delta_{\text{perc}} \) and \( \delta_{\text{cluster max}} \) contain information pertaining to morphology and connectivity, we compile partial MFs for individual clusters at both these thresholds for the three cosmological models \( \Lambda \)CDM, SCDM and SCDM. We find it convenient to work with the cumulative probability function (CPF), which we define as the normalized count of clusters with the value of a quantity \( Q \) to be greater than \( q \) at a given value \( q \). The quantity \( Q \) could be one of the MFs or one of the shapefinders. We shall study the dependence of CPF with \( Q \) on a log–log scale.

Figure 17. The largest (percolating) supercluster in \( \Lambda \)CDM. This cluster is selected at the density threshold that marks the onset of percolation \((\delta_{\text{perc}} = 2.3)\). As demonstrated in the figure, the cluster at this threshold percolates through the entire length of the simulation box. It is important to note that the percolating supercluster occupies only a small fraction of the total volume and its volume fraction (filling factor) is only 0.6 per cent. Our percolating supercluster is a multiply connected and highly filamentary object. Its visual appearance is accurately reflected in the value of the shapefinder diagnostic assigned to this supercluster: \((T, B, L) = (5.63, 7.30, 70.03) h^{-1} \text{Mpc}\) and \((P, F, G) = (0.13, 0.81, 6)\). This figure can be seen in colour in the on-line version of the journal on Synergy.

Our results are presented in Figs 19 and 20 for clusters compiled at the NCS threshold \( \delta_{\text{cluster max}} \). For large values of the MFs \((V, A, C)\) the CPF declines rapidly. Although the curves may look similar, some of them are statistically very different from their peers. The first four columns of Table 4 show the results of the Kolmogorov–Smirnov (KS) test applied to the distribution of the MFs. (We do not carry out a similar exercise for the genus since the vast majority of clusters at the NCS threshold are simply connected.) In particular the CPF of the masses clearly distinguish the \( \Lambda \)CDM model from SCDM and SCDM, in agreement with Fig. 19.

Fig. 20 shows the CPF for planarity and filamentarity in our cluster sample. The last two columns of Table 4 show the results of the KS test for these statistics. For all three pairs they are among the three best discriminators of the models. The value of KS statistic \( \Delta d \) suggests that the signal comes from relatively small values of \( P(t < 0.1) \) and \( F(t < 0.25) \). The more conspicuous differences in the tail of the distribution are not statistically significant owing to the poor statistics. Clusters in \( \Lambda \)CDM are the least anisotropic, a fact that corresponds to their relatively early formation and therefore longer evolution. From Fig. 20 we also find that clusters at the NCS threshold are significantly more filamentary than they are planar. This appears to be a generic prediction of gravitational clustering as demonstrated by Arnol’d et al. (1982), Klypin & Shandarin (1983), Sathyaprakash et al. (1996) and Bond et al. (1996).

Fig. 21 is a scatter plot of shapefinders \( T, B \) and \( L \) for clusters in \( \Lambda \)CDM defined at the percolation threshold. The strong correlation between \( T \) and \( B \) in the left panel indicates that two (of three) dimensions defining any given cluster assume similar
values and are of the same order as the correlation length. (Note that $T \simeq B \simeq 5\, h^{-1}\, \text{Mpc}$ for the largest superclusters in Table 3.) The clustering of objects near $T \simeq B$ ($P \simeq 0$) in this panel suggests that our clusters/superclusters are either quasi-spherical or filamentary. The scatter plot between $B$ and $L$ in the right panel of Fig. 21 breaks the degeneracy between spheres and filaments. The mass dependence of morphology is highlighted in this panel in which larger dots denote more massive objects. This figure clearly reveals that more massive clusters/superclusters are, as a rule, also more filamentary, while smaller, less massive objects are more nearly spherical. The concentration of points near the ‘edge’ of the scatter plot in the right panel arises because of an abundance of low-mass compact quasi-spherical objects with $L \simeq B$ ($F \simeq 0$). The large number of massive superclusters with $F > 0.5$ serves to highlight the important fact that the larger and more massive elements of the supercluster chain consist of highly elongated filaments as much as $\sim 100\, h^{-1}\, \text{Mpc}$ in length, with mean diameter $\sim 5\, h^{-1}\, \text{Mpc}$ (see Table 3). It is important to reiterate that almost 40 per cent of the total overdense mass resides in the 10 largest objects listed in Table 3, while the remaining 60 per cent is distributed among 1324 clusters.

We also find short filaments to be generally thinner than longer ones, in agreement with predictions made by the adhesion model with regard to the formation of hierarchical filamentary structure during gravitational clustering (Kofman et al. 1992).

To probe the morphology of clusters and superclusters further, we define the notion of shape-space in Fig. 22. Shape-space is 2D with the filamentarity (of a cluster) plotted along the $x$-axis and its filamentarity along the $y$-axis. (One can also incorporate a third dimension showing the genus.) The first panel in Fig. 22 is a scatter plot of $P$ and $F$ for clusters in $\Lambda$CDM. The sizes of dots in the middle panel are proportional to cluster mass. We note that the most massive structures are also very filamentary. In the right panel we try to relate the shape of the structures with their topology by scaling the size of the dots with the genus value of clusters having a given morphology $(P, F)$. As shown here, clusters that are multiply connected (larger dots are indicative of more complicated topologies) are also more filamentary. Together, the three panels show us that more massive superclusters are frequently very filamentary and often also topologically quite complex. We also see that a large number of less massive superclusters are simply connected and prolate. These structures are a few megaparsecs across along their two

![Shape-Space trajectory of the Largest Supercluster](image)

Figure 18. The morphological evolution of the largest supercluster in $\Lambda$CDM is shown as a series of open squares in shape-space $(F, P)$. Each square corresponds to a different value of the density threshold, which is progressively lowered from a large initial value ($\delta = 6.9$; leftmost square) until the mean density level ($\delta = 0$; lowermost square). The legend in the top-right corner lists the density contrast, the associated volume fraction and the genus of the largest supercluster at six monotonically decreasing values of the density contrast (1 $\rightarrow$ 6). At the highest ($\delta \approx 6.9$), the largest cluster appears to have a large filamentarity and a small planarity. The solid triangle (labelled 3) refers to the percolation threshold at which the largest cluster breaks the degeneracy between spheres and filaments. The middle panel is a scatter plot of $B$ versus $L$ (in the right panel of Fig. 21). The scatter plot in the right panel arises because of an abundance of low-mass filamentary structures. The concentration of points near the ‘edge’ of the scatter plot in the right panel arises because of an abundance of low-mass compact quasi-spherical objects with $L \simeq B$ ($F \simeq 0$). The large number of massive superclusters with $F > 0.5$ serves to highlight the important fact that the larger and more massive elements of the supercluster chain consist of highly elongated filaments as much as $\sim 100\, h^{-1}\, \text{Mpc}$ in length, with mean diameter $\sim 5\, h^{-1}\, \text{Mpc}$ (see Table 3). It is important to reiterate that almost 40 per cent of the total overdense mass resides in the 10 largest objects listed in Table 3, while the remaining 60 per cent is distributed among 1324 clusters.

We also find short filaments to be generally thinner than longer ones, in agreement with predictions made by the adhesion model with regard to the formation of hierarchical filamentary structure during gravitational clustering (Kofman et al. 1992).

To probe the morphology of clusters and superclusters further, we define the notion of shape-space in Fig. 22. Shape-space is 2D with the filamentarity (of a cluster) plotted along the $x$-axis and its filamentarity along the $y$-axis. (One can also incorporate a third dimension showing the genus.) The first panel in Fig. 22 is a scatter plot of $P$ and $F$ for clusters in $\Lambda$CDM. The sizes of dots in the middle panel are proportional to cluster mass. We note that the most massive structures are also very filamentary. In the right panel we try to relate the shape of the structures with their topology by scaling the size of the dots with the genus value of clusters having a given morphology $(P, F)$. As shown here, clusters that are multiply connected (larger dots are indicative of more complicated topologies) are also more filamentary. Together, the three panels show us that more massive superclusters are frequently very filamentary and often also topologically quite complex. We also see that a large number of less massive superclusters are simply connected and prolate. These structures are a few megaparsecs across along their two

© 2003 RAS, MNRAS 343, 22–46
shorter dimensions and $\sim 20$ Mpc along the third, and therefore have appreciable filamentarity ($F \sim 0.3$).

6 DISCUSSION AND CONCLUSIONS

This paper presents a new technique for studying the geometrical and topological properties of large-scale structure using Minkowski functionals (MFs). Given a density field reconstructed by appropriately smoothing a point data set consisting (say) of a distribution of galaxies, (i) our Ansatz constructs closed polyhedral surfaces of constant density corresponding to excursion sets of a density field, using for this purpose a surface generating triangulation scheme (SURFGEN), and (ii) SURFGEN evaluates the MFs (volume, surface area, extrinsic curvature and genus), thus providing full morphological and topological information of 3D isodensity contour surfaces corresponding to a given data set. Evaluated in this manner, the MFs can be used to study the properties of individual objects lying either above (clusters, superclusters) or below (voids) a given density threshold. They can also be used to study the morphological properties of the full supercluster–void network at (say) the percolation threshold. The ratios of MFs (shapefinders) are used to probe the shape of isodensity surfaces which sample the distribution of large-scale structure at different thresholds of the density. (The highest density thresholds correspond to galaxies and clusters of galaxies, moderate thresholds correspond to superclusters, while the lowest density thresholds characterize voids.) The performance of our Ansatz has been tested against both simply and multiply connected surfaces such as the sphere, triaxial ellipsoid and triaxial torus. These three eikonal bodies can be smoothly deformed to give surfaces that are spheroidal, pancake-like (oblate) and filament-like (prolate), etc. Analytically known values of the MFs for these surfaces allow us to test both our surface modelling scheme and our evaluation of MFs from triangulation. Remarkably, we find that volume, area and integrated extrinsic curvature are determined to an accuracy of better than 1 per cent on all length-scales for both simply and multiply connected surfaces. The integrated intrinsic curvature (genus) is evaluated exactly. We also find that SURFGEN is remarkably accurate at calculating the MFs for Gaussian random fields (see the Appendix). Having validated the performance of the code against eikonal surfaces and Gaussian random fields, we apply our method to cosmological simulations of large-scale structure performed by the Virgo Consortium. We study the geometry and topology of large-scale structure in three cosmological scenarios – $\Lambda$CDM, $\tau$CDM and SCDM. All three cosmologies are analysed at the present epoch ($z = 0$) using global MFs, partial MFs and shapefinders. (The redshift evolution of geometry and topology will be discussed in a companion paper.) Our main conclusions are summarized below:

(i) Using the MFs we show that, like other diagnostics of clustering, supercluster morphology too is sensitive to the underlying cosmological parameter set characterizing our Universe. Although the three cosmological models considered by us, $\Lambda$CDM, $\tau$CDM and SCDM, display features that are qualitatively similar, the
Figure 20. Cumulative probability function (CPF) for the filamentarity ($F$) and the planarity ($P$) of superclusters selected at the density threshold $\delta_{\text{cluster max}}$. The curves have the same meaning for both panels. This figure can be seen in colour in the on-line version of the journal on Synergy.

Table 4. The Kolmogorov–Smirnov statistic $d$ (first row for each pair of models) and the probability (second row) that the two data sets are drawn from the same distribution. The models $\Lambda$CDM, $\tau$CDM and SCDM are compared on the basis of three Minkowski functionals: volume ($V$), area ($A$) and integrated mean curvature ($C$), as well as mass ($M$) and the shapefinders, thickness ($T$), breadth ($B$), length ($L$), planarity ($P$) and filamentarity ($F$). Boldface highlights the three smallest probabilities for each pair of models.

<table>
<thead>
<tr>
<th>Models</th>
<th>$M$</th>
<th>$V$</th>
<th>$A$</th>
<th>$C$</th>
<th>$T$</th>
<th>$B$</th>
<th>$L$</th>
<th>$P$</th>
<th>$F$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Lambda$CDM–$\tau$CDM</td>
<td>$9.4 \times 10^{-4}$</td>
<td>0.036</td>
<td>0.094</td>
<td>0.56</td>
<td>$5.0 \times 10^{-3}$</td>
<td>0.026</td>
<td>0.56</td>
<td>$6.5 \times 10^{-9}$</td>
<td>$2.2 \times 10^{-4}$</td>
</tr>
<tr>
<td>$\Lambda$CDM–SCDM</td>
<td>$4.6 \times 10^{-3}$</td>
<td>0.048</td>
<td>0.069</td>
<td>0.17</td>
<td>$2.9 \times 10^{-3}$</td>
<td>0.018</td>
<td>0.16</td>
<td>$3.2 \times 10^{-7}$</td>
<td>$6.3 \times 10^{-3}$</td>
</tr>
<tr>
<td>$\tau$CDM–SCDM</td>
<td>0.034</td>
<td>0.024</td>
<td>0.032</td>
<td>0.041</td>
<td>0.021</td>
<td>0.023</td>
<td>0.042</td>
<td>0.052</td>
<td>0.046</td>
</tr>
<tr>
<td></td>
<td>0.54</td>
<td>0.90</td>
<td>0.61</td>
<td>0.32</td>
<td>0.97</td>
<td>0.92</td>
<td>$0.29$</td>
<td>$0.096$</td>
<td>$0.19$</td>
</tr>
</tbody>
</table>

Percolation is reached at moderate values of the density contrast ranging from $\delta_{\text{perc}} \simeq 2.3$ for $\Lambda$CDM to $\delta_{\text{perc}} \simeq 1.2$ for SCDM. The abundance of clusters reaches a maximum value at (or very near) the percolation threshold and the percolating supercluster occupies a rather small amount of space in all three cosmological models. Thus the fraction of total simulation-box volume contained in the percolating supercluster is least in $\Lambda$CDM (0.6 per cent) and greatest in SCDM (1.2 per cent). When taken together, all overdense objects at the percolation threshold occupy 4.4 per cent of the total volume in the $\Lambda$CDM model. For comparison, the volume fraction in overdense regions at the percolation threshold is $\sim 16$ per cent in an idealized, continuous Gaussian random field (Shandarin & Zeldovich 1989). This fraction can increase up to $\sim 30$ per cent for Gaussian fields generated on a grid (Yess & Shandarin 1996; Sahni, Sathyaprakash & Shandarin 1997). The
fact that clusters and superclusters occupy a very small fraction of the total volume appears to be a hallmark of the gravitational clustering process which succeeds in placing a large amount of mass (~30 per cent of the total, in the case of ΛCDM) in a small region of space (~4 per cent). The low filling fraction of the percolating supercluster in ΛCDM (0.006) strongly suggests that this object is either planar or filamentary (Sahni et al. 1997) and a definitive answer to this issue is provided by the shapefinder statistic.

(iv) Shapefinders were introduced to quantify the visual impression one has of the supercluster–void network of being a cosmic web of filaments interspersed with large voids (Sahni et al. 1998). By applying the shapefinder statistics via SURFGEN to realistic N-body simulations, we have demonstrated the following: (1) Most of the mass in the Universe is contained in large superclusters, which are also extremely filamentary; the vast abundance of smaller clusters and superclusters tends to be prolate or quasi-spherical. (2) Of the three cosmological models, the percolating supercluster in ΛCDM is the most filamentary (F ≃ 0.81) and the supercluster in τCDM the least (F ≃ 0.7). (3) The percolating supercluster in ΛCDM is topologically a much simpler object than its counterpart in τCDM, with the former having only 26 tunnels compared to 19 in the latter. Other differences between the models are quantified in Figs 13–22. We also show that, among various morphological parameters, the planarity and filamentarity of clusters and superclusters are two of the most powerful statistics to discriminate between the models (see Table 4).

To summarize, this paper has demonstrated that MFs and shapefinders evaluated using SURFGEN provide sensitive probes of the geometry, topology and shape of large-scale structure and help in distinguishing between rival cosmological models. Having established the strength and versatility of SURFGEN it would clearly be interesting to apply it to other important issues, related to those addressed in this paper, including: (i) geometry and morphology of underdense regions (voids); (ii) geometry and morphology of strongly overdense regions (clusters); (iii) redshift-space distortions of supercluster/void morphology; (iv) the time evolution of the supercluster–void network; and (v) analysis of superclusters and voids in fully 3D surveys such as 2dFGRS and SDSS. SURFGEN is also likely to provide useful insights into other physical and astrophysical situations in which matter is distributed anisotropically such as in the interstellar medium (Lazarian & Pogosyan 1996; Boumis et al. 2002). We hope to return to some of these subjects in the near future.

ACKNOWLEDGMENTS

We acknowledge useful discussions with Andrew Benson, Shaun Cole, Suman Datta, Carlos Frenk, Adrian Jenkins, Madan Rao, C. D. Ravikumar, Tarun Deep Saini, Yuri Shtanov, Volker Springel and David Weinberg. Our insights into programming techniques of computational geometry were honed by the excellent monograph on the subject by Joseph O’Rourke. Part of this work was done during a visit by JVS, VS and SFS to the Department of Physics and Astronomy, University of Cardiff, under a PPARC Grant (Ref. PPA/V/S/1998/00041). JVS is supported by a Research Fellowship of the Council of Scientific and Industrial Research (CSIR), India. VS and SFS acknowledge support from the Indo–US collaborative project DST/NSF/US (NSF-RP087)/2001. The simulations studied in this paper were carried out by the Virgo Supercomputing Consortium using computers based at the Computing Centre of the Max-Planck Society in Garching and at the Edinburgh Parallel Computing Centre. The data are publicly available at www.mpa-garching.mpg.de/NumCos.
Figure 22. Shape-space for $\Lambda$CDM. Left panel (multiplicity): the dots in this panel have area proportional to the number of clusters with a given (binned) value of filamentarity and planarity. Centre panel (mass): the dots have area proportional to the total mass contained in clusters having a given (binned) value of filamentarity and planarity. Note that more massive superclusters are also more filamentary. Right panel (genus): the dots have area proportional to the total genus value of superclusters with a given (binned) value of filamentarity and planarity. This figure demonstrates the correlation between the mass of a supercluster, its shape and its genus. More massive superclusters are, as a rule, very filamentary and also topologically multiply connected. As in the previous figure, all objects are determined at the percolation threshold.

REFERENCES
The samples are coincident with the parameter $\lambda$, $\nu$. This leaves them with zero mean and unit variance. The MFs are $f_i$ grid units. All the $n$ analytical results against which the performance of SURFGEN is to $y$ describe the samples that we use and summarize the an-

APPENDIX A: GAUSSIAN RANDOM FIELDS 

AND THE ROLE OF BOUNDARY CONDITIONS

This appendix demonstrates the great accuracy with which SURFGEN determines Minkowski functionals (MFs) for Gaussian random fields (hereafter GRFs). Before embarking on our discussion, we briefly describe the samples that we use and summarize the analytical results against which the performance of SURFGEN is to be tested.

We work with three realizations of a Gaussian random field with a power-law power spectrum ($n = -1$) on a $128^3$ grid. Each realization of the field is smoothed with a Gaussian kernel of length $\lambda = 2.5$ grid units. All the fields are normalized by the standard deviation. This leaves them with zero mean and unit variance. The MFs are evaluated at a set of equispaced levels of the density field which coincide with the parameter $\nu$ on account of $\sigma$ being unity ($\rho_{TH} = \nu \sigma = \nu$). $\nu$ is used to label the levels and is related to the volume filling fraction through the following equation:

$$
\text{FF}_{\nu}(v) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp(-r^2/2) \, dr.
$$

The samples are finally used for ensemble averaging.

The MFs of a GRF are fully specified in terms of a length-scale $\lambda_c$:

$$
\lambda_c = \sqrt{\frac{2\pi\xi(0)}{|\xi''(0)|}}; \quad \sigma^2 = \xi(0).
$$

$\lambda_c$ can be analytically derived from knowledge of the power spectrum. It can also be estimated numerically by evaluating the variance $\xi(0)$ of the field and the variance $\xi''(0)$ of any of its first spatial derivatives (for more details see Matsubara 2003).

For a GRF in three dimensions the four MFs (per unit volume) are (Tomita 1990; Matsubara 2003):

$$
V(v) = \frac{1}{2} - \frac{1}{2} \Phi \left( \frac{v}{\sqrt{2}} \right),
$$

$$
\Phi(x) = \frac{2}{\sqrt{\pi}} \int_{0}^{x} \exp \left( -\frac{t^2}{2} \right) \, dt, \tag{A3}
$$

where $\Phi(x)$ is the error function;

$$
S(v) = \frac{2}{\lambda_c} \sqrt{\frac{2}{\pi}} \exp \left( -\frac{v^2}{2} \right), \tag{A4}
$$

$$
C(v) = \sqrt{\frac{2\pi}{\lambda_c^2}} v \exp \left( -\frac{v^2}{2} \right), \tag{A5}
$$

$$
G(v) = \frac{1}{\lambda_c \sqrt{2\pi}} (1 - v^2) \exp \left( -\frac{v^2}{2} \right). \tag{A6}
$$

Let us now describe the boundary conditions which we adopt and which incorporate the periodic nature of the GRFs generated using a fast Fourier transform (FFT) routine. GRFs defined on a grid are, by construction, periodic in nature. As a result, two overdense excursions sets visibly separated from each other inside the box and touching two opposite sides of the box would constitute a single cluster. To make this more clear, let us turn to Fig. A1, which shows four clusters (labelled 1, 2, 3 and 4 respectively) in the left panel. The boundaries of these clusters – as drawn by an analogue of SURFGEN in 2D – are shown in bold, and we note that the sides of the box (bold dashed lines) are incorporated when we define the clusters. However, owing to the periodicity in the density field, these four clusters are in fact one single cluster (right panel). We note that the portion of the contours due to the boundaries of the box occurs in the interior of the actual cluster (shown with a bold dashed line).
It is easy to extrapolate this situation to 3D. Note that the interpolation method employed by SURFGEN will correctly estimate the global volume of the excursion set even if we worked with the field as it is. In 2D this is the same as estimating the area enclosed by the full cluster without having to piecethe four ‘clusters’ (labelled by 1, 2, 3 and 4) together. Our estimation of global genus will also turn out to be exact and, as can be shown, this will hold true even if either of these four constituent clusters exhibits a non-trivial topology. On the other hand, if we work with the same set of contours for estimating global area and mean curvature, we will wrongly be adding an excess contribution due to the sides of the box. (In 2D, this translates to adding the contribution of the dashed line segments to the perimeter of the contour in the right panel of Fig. A1.) The actual area and mean curvature contribution come from the interface between overdense and underdense regions (such as the 2D contours drawn in solid bold line in both the panels) as opposed to the artificial boundaries of the clusters (shown dashed).

We note that incorporating periodic boundary conditions in order to construct a total contour as shown in the right panel of Fig. A1 is non-trivial from the numerical point of view. When the excursion set touches all the boundaries (as, for example, at a very low threshold of density, below the percolation threshold), the corresponding algorithm could further run into an infinite loop, which makes it impracticable for implementation. In light of this, it would be more manageable if we restricted ourselves to using contours which are closed at the boundaries of the box (like the ones in the left panel of Fig. A1), but devise a way to exclude the excess contribution to surface area and mean curvature from the boundaries of the box. A uniform prescription of this type, when applied to all the appropriate clusters at all the thresholds of density, will lead to correct estimation of all the global MFs. This, then, is the choice of boundary conditions incorporated in our calculation of a GRF. A further cautionary note concerns the familiar W-shape of the genus curve. Recall that the genus curve peaks at $\nu = 0$ signifying a sponge-like topology of the medium at the mean density threshold. Note further that $G(\nu=\pm\sqrt{3}) = -2G(0)\exp(-3/2)$ are the two negative minima of the genus curve. In fact for $|\nu| > 1$, the genus is always negative, and approaches zero from below when $|\nu| \gg 1$. For $\nu < -1$, this negative genus refers to isolated underdense regions (bubbles) in the overdense excursion set. Similarly, the negative genus for $\nu > 1$ should be interpreted as being caused by isolated overdense regions (meatballs) in the underdense excursion set. To summarize, in order to determine the genus for a GRF we need to estimate the genus of the overdense excursion set for $\nu \leq 0$ and the genus of the underdense excursion set for $\nu > 0$, thereby making use of the symmetry property of the GRFs. (Note, however, that in our earlier determination of the genus curve for $N$-body simulations in this paper, we have dealt with the topology of overdense regions only, because of which the genus curve is asymmetric with respect to $\nu = 0$ and is always positive when $\nu > 1$.)

Having elaborated on the nature of boundary conditions used in our evaluation of MFs for a GRF, we now present our results. Fig. A2 shows the global MFs averaged over three realizations of the density field (solid lines with $1\sigma$ error bars) along with the exact analytical results (shown dotted). There are in all 40 levels for every realization. Fig. A2 clearly shows the remarkable agreement

Figure A2. Minkowski functionals for a GRF ($P(k) \sim k^{-1}$), smoothed with a Gaussian filter with $\lambda = 2.5$ grid units. Values of the four MFs determined using SURFGEN are shown as solid lines together with the $1\sigma$ scatter. Exact analytical results are shown as dotted lines.
between exact theoretical results (A3)–(A6) and numerical estimates obtained using SURFGEN. We therefore conclude that SURFGEN determines the MFs of a GRF to great precision.

The above discussion demonstrates the important role played by boundary conditions in reproducing the analytical predictions for GRFs. Having tested the performance of SURFGEN against GRFs, we are faced with the following choice of boundary conditions when dealing with N-body simulations and mock/real galaxy catalogues: either (i) we could correct for the boundaries when dealing with the overdense regions which encounter the faces of the survey volume, or (ii) we could avoid doing this and treat all measurements with a tacit understanding that area and mean curvature contain an excess contribution which arises because of boundary effects. It should be noted that the boundary effects become important once the system has percolated, i.e. when there are a large number of clusters touching the boundaries of the box. The corresponding value of the filling factor \( FF_V > FF_{\text{perc}} \) is useful only when one wishes to compare two samples through the trends in their global MFs. Since the structural elements of the cosmic web (superclusters, voids) are usually identified near the percolation threshold, and are therefore not very sensitive to boundary effects, we have chosen option (ii) in our analysis of N-body simulations in this paper. Option (ii) is also more suited for dealing with real galaxy catalogues which do not satisfy periodic boundary conditions. We should also point out that since we compare two samples under the same conditions – the same volume and dimensions of the box, same resolution, etc. – the excess contribution to the area and mean curvature caused by box boundaries enters in identical fashion for both the samples. We therefore prefer option (ii) to (i) and all our simulations in Section 5 are analysed without correcting for contributions from the boundaries. (It may also be noted that the contribution from the boundaries becomes less important as we deal with larger survey volumes.)

This paper has been typeset from a \TeX\ file prepared by the author.