On Statistics and Dynamics
of Cosmic Structure

Jens Schmalzing

München 1999

On Statistics and Dynamics
of Cosmic Structure

Dissertation

der

Fakultät für Physik

der

Ludwig–Maximilians–Universität

vorgelegt von

Jens Schmalzing

aus

Nürnberg

München

den

24. September 1999

Contents

Contents

4

1 Zusammenfassung

In the beginning the Universe was created.
This has made a lot of people very angry
and been widely regarded as a bad move.
(Adams 1986)

Bereits kurze Zeit nach diesem denkwürdigen Ereignis, bei einer Rotverschiebung von etwa 1100 und einem Weltalter von 300,000 Jahren, entsendet die letzte Streufläche eine Strahlung, die wir heute im Mikrowellenbereich als Kosmische Hintergrundstrahlung wahrnehmen. Ihre sehr kleinen Schwankungen werden über den Sachs–Wolfe–Effekt (Sachs & Wolfe 1967) als Abbild der Schwankungen der damaligen Materieverteilung interpretiert. Diese winzigen Dichtefluktuationen sind in der jetzigen Epoche offensichtlich zu ausgeprägten Strukturen angewachsen. Unser Ziel ist es, diese Strukturen quantitativ zu beschreiben, um so Einsichten in ihre Entstehung zu gewinnen. Da das Universum letztendlich durch die zufälligen primordialen Quantenfluktuationen[1] seine Struktur erhält (Börner 1993), fassen wir es als eine Realisierung eines Zufallsprozesses auf, und machen es dadurch der Bewertung mittels statistischer Methoden zugänglich.

1.1 Integralgeometrie

Jenseits der seit langem (Totsuji & Kihara 1969; Peebles 1974) in der Kosmologie verwendeten Zweipunktmaße werden in letzter Zeit Charakteristiken populär, die auch die höheren Ordnungen einer Verteilung erfassen. Zur mathematischen Beschreibung der Galaxienverteilung bedienen wir uns der Theorie der Punktprozesse (Daley & Vere-Jones 1988), während wir eine kontinuierliche Verteilung etwa der Dunklen Materie als Zufallsfeld (Adler 1981) auffassen wollen. Was die Größen angeht, die wir zur Untersuchung berechnen werden, konzentrieren wir uns auf morphologische Maße, also solche, die Topologie und Geometrie eines Körpers quantifizieren, und hier speziell auf die Minkowskifunktionale.

Deren mathematische Heimat stellt die Integralgeometrie dar. Für einen profunden Einstieg in das Gebiet verweisen wir auf die Standardliteratur (Schneider 1993; Stoyan et al. 1995). Wir

---

[1]Alternative Szenarien zum Ursprung von Struktur, nämlich die Defekttheorien, die kosmische Strukturen auf Phasenübergänge im frühen Universum zurückführen, enthalten ebenfalls eine Zufallskomponente.

Tabelle 1.1 Geometrische Interpretationen der Minkowskifunktionale in einer, zwei und drei Dimensionen. χ bezeichnet die Eulercharakteristik, die als rein topologische Größe den Zusammenhang einer Menge mißt.

d	1	2	3
V_0	Länge	Fläche	Volumen
V_1	χ	Umfang	Oberfläche
V_2	–	χ	Integrale mittlere Krümmung
V_3	–	–	χ

wollen festhalten, daß der Satz von Hadwiger (1957) unter sehr allgemeinen Voraussetzungen die Zahl der unabhängigen morphologischen Maße in d Dimensionen auf $d+1$ beschränkt. Dabei handelt es sich gerade um die Minkowskifunktionale V_μ, $\mu = 0, \ldots, d$. Darüber hinaus sind alle Minkowskifunktionale im zwei– sowie im dreidimensionalen Raum als wohlbekannte geometrische Größen interpretierbar, die in Tabelle 1.1 kurz zusammengefaßt sind.

Zur Untersuchung von Punktverteilungen im Raum wurde von Mecke et al. (1994) das verallgemeinerte Boolesche Kornmodell in die Kosmologie eingeführt. Zu einer Menge $\{\mathbf{x}_i\}$, $i = 1 \ldots N$ von Punkten, die einen Punktprozeß realisiert, konstruieren wir einen Körper, indem wir jeden Punkt mit einer Kugel vom Radius r verzieren und die Vereinigung

$$A_r = \bigcup_{i=1}^{N} B_{\mathbf{x}_i} \tag{1.1}$$

betrachten. Der Radius dient dabei als diagnostischer Parameter. Die Berechnung der Minkowskifunktionale des Kornmodells erfolgt mittels der partiellen Funktionale und ist bei Mecke et al. (1994) knapp und bei Schmalzing (1996) detailliert dargestellt.

Für Dichtefelder $u(\mathbf{x})$ im d–dimensionalen Raum erweist es sich als zweckmäßig, die Morphologie der Exkursionsmenge

$$Q_\nu = \{\, \mathbf{x} \,|\, u(\mathbf{x}) \geq \nu \,\} \tag{1.2}$$

über eine Schwelle ν zu untersuchen. Hier dient die Schwelle ν als diagnostischer Parameter. Schmalzing & Buchert (1997) geben zur Berechnung der Minkowskifunktionale einer Exkursionsmenge zwei Alternativen an, die auf der Croftonschen Schnittformel (Crofton 1868), beziehungsweise auf der räumlichen Mittelung von Koenderinkinvarianten (Koenderink 1984; ter Haar Romeny et al. 1991) beruhen. Letztere Methode läßt sich für beliebige Räume konstanter Krümmung verallgemeinern, was sie besonders für die Anwendung auf Karten der Temperaturschwankungen des Mikrowellenhintergrundes interessant macht (siehe hierzu Kapitel 4 beziehungsweise Abschnitt 1.3).

In unseren Analysen betrachten wir nicht nur die Minkowskifunktionale selbst, wir versuchen außerdem, aus den Minkowskifunktionalen quantitative Maße für den Gehalt an filamentartige und

planaren Strukturen in einer Punkt– oder Dichteverteilung zu gewinnen. Für zusammenhängende Einzelteile einer Isodichtekontur verwenden wir dabei in Kapitel 5 (siehe auch Abschnitt 1.4) die von Sahni et al. (1998) vorgeschlagenen isoperimetrischen Verhältnisse. Im Falle des Booleschen Kornmodells leiten wir dagegen Filamentarität und Planarität aus den globalen Funktionalen durch Anpassung an ein stark vereinfachtes Spielzeugmodell her (Schmalzing & Diaferio 1999). Letztere Methode wird in Kapitel 6 (siehe auch Abschnitt 1.5) erklärt und angewandt.

1.2 Die Entwicklung großräumiger Strukturen

Auf den extragalaktischen Längenskalen, die uns in dieser Arbeit interessieren werden, ist das Wachstum von Struktur wesentlich von der gravitativen Wechselwirkung dominiert. Nach dem Standardmodell der Kosmologie (Goenner 1994) kann die Raumzeit gut durch eine Friedmann–Robertson–Walker–Metrik

$$\mathrm{d}s^2 = \mathrm{d}t^2 - a(t)^2 \left(\frac{\mathrm{d}r^2}{1 - kr^2} + r^2 \mathrm{d}\Omega^2 \right) \tag{1.3}$$

genähert werden, die ein räumlich homogenes und mit dem Skalenfaktor $a(t)$ expandierendes Universum beschreibt. Vor diesem Hintergrund vollzieht sich die Entwicklung von Strukturen aus den anfänglich sehr kleinen Dichteschwankungen im wesentlichen durch die Selbstgravitation der Dunklen Materie, die keiner anderen Wechselwirkung unterliegt und daher gerne durch ein einkomponentiges, druckfreies Fluidum beschrieben wird. In der Newtonschen Näherung betrachten wir die Dichte ρ, die Geschwindigkeit $\mathbf{v}$ sowie die Beschleunigung $\mathbf{g}$ als von den nichtrotierenden Eulerschen Koordinaten $(\mathbf{x}, t)$ abhängige Felder, und erhalten ein System partieller Differentialgleichungen, das Euler–Newton System

$$\begin{aligned}
\partial_t \rho + \nabla (\rho \mathbf{v}) &= 0 \\
\partial_t \mathbf{v} + (\mathbf{v}\nabla) \, \mathbf{v} &= \mathbf{g} \\
\nabla \mathbf{g} &= \Lambda - 4\pi G \rho \\
\nabla \times \mathbf{g} &= 0.
\end{aligned} \tag{1.4}$$

Alternativ können wir Lagrangekoordinaten $\mathbf{X}$ sowie ihre Abbildung $\mathbf{x} = \mathbf{f}(\mathbf{X}, t)$ auf die Eulerschen Positionen einführen. Es zeigt sich, daß das aus dieser Variablentransformation resultierende Lagrange–Newton–System nur noch die Abbildung $\mathbf{f}$ enthält, und daß die Eulerschen Felder auf einfache Weise daraus berechnet werden können. Trotz dieser Reduktion auf ein einziges Vektorfeld bleiben die nichtlinearen partiellen Differentialgleichungen das Hauptproblem.

Für die näherungsweise Lösung dieses Systems sind eine Vielzahl von Ansätzen vorgeschlagen worden (siehe Sahni & Coles 1995 und Referenzen). Wir verwenden eine spezielle Entwicklung der Abbildung $\mathbf{f}$ in zweiter Ordnung (Buchert 1996), um eine Separation in orts– und zeitabhängige Anteile zu erreichen. In Verallgemeinerung der Arbeiten von Weiß et al. (1996) auf beliebige Hintergrundkosmologien zeigen wir eine Methode zur numerischen Abschätzung der Materieverteilung auf großen Skalen auf. Durch ihre hohe Effizienz ist sie der Simulation

der vollen Newtonschen Dynamik auf großen Skalen überlegen. Dies ist insbesondere bei Untersuchungen zur kosmischen Varianz der Fall, wofür wir in Kapitel 6 (siehe auch Abschnitt 1.5) ein Beispiel geben.

1.3   Anisotropien der Kosmischen Hintergrundstrahlung

Die Anisotropien der Hintergrundstrahlung entwickeln sich in der Epoche der Rekombination aus den primordialen Fluktuationen. Dabei sind die Schwankungen in der Temperatur der Strahlung und die Schwankungen in der Materiedichte gekoppelt, und geben daher Aufschluß über die sehr kleinen Dichtefluktuationen, aus denen im Laufe der Zeit die ausgeprägten Strukturen entstanden, die im heutigen Universum erscheinen. Nach der kurzen Klärung der integralgeometrischen Grundlagen in Kapitel 2 geben wir in diesem Kapitel an, wie aus Karten, die aus Satellitendaten gewonnen werden, die Minkowskifunktionale des kosmischen Signals trotz Einschränkungen bei der Beobachtung berechnet werden können.

Von entscheidender Bedeutung für die Entstehung der primordialen Fluktuationen überhaupt ist die Frage nach der Art der Verteilung der Temperaturschwankungen. Handelte es sich dabei um ein Gaußsches Zufallsfeld, würden inflationäre Szenarien gestützt, während nicht–Gaußsche Signaturen auf Phasenübergänge oder exotische Mechanismen der Inflation hinweisen würden. Auch die Minkowskifunktionale sind zur Beantwortung dieser Frage prinzipiell geeignet. Es zeigt sich, daß offensichtlich nicht–Gaußsche Felder auch unter widrigen Umständen als solche erkannt werden können. Die Daten des COBE–Satelliten sind jedoch mit der Annahme eines Gaußschen Zufallsfeldes vereinbar, jedenfalls was ihre Minkowskifunktionale betrifft. Daß an anderer Stelle (Ferreira et al. 1998; Banday et al. 1999) durchaus signifikante Abweichungen vom Gaußschen Verhalten festgestellt wurden, zeigt die Komplementarität der verwendeten Methoden.

1.4   Die Verteilung der Dunklen Materie

Da die Dunkle Materie, wie schon der Name andeutet, nicht direkt beobachtbar ist, ist man beim Studium ihrer Verteilung auf approximative Lösungen oder auf numerische Experimente anhand von N–Körper–Simulationen angewiesen. Wir verwenden eine Reihe hochaufgelöster Simulationen (Melott & Shandarin 1993; Jenkins et al. 1997) zur Demonstration der in Kapitel 2 dargestellten Methoden. Zunächst klären wir jedoch die Frage, wie aus den Teilchen einer Simulation am zweckmäßigsten das zugrundeliegende Dichtefeld rekonstruiert werden kann. Dazu werden im allgemeinen Schätzer mit Glättungskernen konstanter Breite verwendet. Es zeigt sich, daß eine deutliche Verbesserung insbesondere für entwickelte Verteilungen dunkler Materie durch adaptive Glättung (Silverman 1986), also eine Anpassung der Glättungslänge an die lokale Dichte erreicht werden kann.

Die Anwendungen der Minkowskifunktionale auf Simulationen fördern reichhaltige Informa-

tionen zutage, wenn sie auf das globale Dichtefeld sowie auf das größte zusammenhängende Objekt darin angewandt werden. Dagegen erweist sich die Auswertung der Morphologie kleinerer Objekte als Fehlschlag, was die Unterscheidung unterschiedlicher Anfangsspektren betrifft. Allerdings können die geringen Unterschiede auch dadurch bedingt sein, daß das Anfangsdichtefeld durchwegs als Gaußsches Zufallsfeld modelliert wird, so daß die geringen Unterschiede auch als Hinweis auf universelles Verhalten interpretiert werden können. Anfangsbedingungen, die einer anderen Statistik als der Gaußschen folgen, sollten dann auch andere Verteilungen der Strukturen erkennen lassen.

1.5   Strukturen in Galaxienkatalogen

In Kapitel 6 gehen wir kurz auf die praktische Anwendung des Booleschen Kornmodells aus Kapitel 2 (siehe auch Abschnitt 1.1) auf Galaxienkataloge ein.

Deutliche Unterschiede treten beim Vergleich eines Galaxienkataloges, des Updated Zwicky Catalogue (UZC) von Falco et al. (1999), mit aus Simulationen gewonnenen Modellkatalogen (Diaferio et al. 1999a) zutage. Daraus könnte auf Unzulänglichkeiten der verwendeten Anfangsszenarien, zweier Varianten der Kalten Dunklen Materie, oder der zur Erzeugung von Galaxien eingesetzten semianalytischen Modellierung geschlossen werden. Ein anderer Ansatz, verweist auf die mit $70h^{-1}\mathrm{Mpc}$ relativ geringe Ausdehnung des Kataloges und der Simulationen. Ob dies tatsächlich ein Problem darstellt, muß eine momentan in Vorbereitung befindliche Arbeit anhand größerer Simulationen und vor allem tieferer Kataloge zeigen.

1.6   Ausblick

Die vorliegende Arbeit soll einen Einblick in aktuelle Entwicklungen der Statistik und Dynamik großräumiger Strukturen im Universum geben. Breiter Raum wird dabei der statistischen Auswertung mit Hilfe von Minkowskifunktionalen eingeräumt. Dieses übergeordnete Thema illustrieren wir exemplarisch durch die detaillierte Darlegung dreier Anwendungsbereiche, die ganz unterschiedliche Arten kosmologisch relevanter Daten verwerten. Wir hoffen, mit diesem Überblick einen Beitrag zu der wachsenden Akzeptanz zu leisten, die die Minkowskifunktionale in letzter Zeit als Alternative zu den weit verbreiteten Zweipunktstatistiken gewinnen.

1 Zusammenfassung

2 Integral geometry

Every chautauqua should have a list somewhere of valuable things to remember that can be kept in some safe place for times of future need and inspiration. [...] A standard toolkit comes with the cycle and is stored under the seat. This is supplemented with the following:
(Pirsig 1974)

Among the sciences, cosmology finds itself in the peculiar situation that it studies a unique object, namely the entire Universe. Nevertheless, it is possible to do statistics over different spatial regions of the Universe, and according to the Copernican Principle (Kanitscheider 1991), we can expect the same results regardless of position and direction – barring cosmic variance. The very first galaxy catalogues were looked upon with simple density estimation in cells Shane & Wirtanen (1954), and already revealed traces of clustering beyond galaxy cluster scales. Nowadays two–point characteristics introduced by Peebles (1974), most notably the traditional Fourier transform pair of power spectrum and correlation function, have become the prime methods of analysis. However, these quantities only describe a Gaussian density field in a unique and complete way, and are blind to the strongly non–Gaussian features seen in the today's distribution of light and matter.

One possible approach to capturing higher–order features of a pattern in space involves geo-metrical descriptors. A number of these statistics have been suggested to date. Among others, recent morphological analyses of the nearby Universe have been performed through measuring the void probability function of the CfA2 catalogue (Vogeley et al. 1994a), applying percolation techniques to the IRAS 1.2Jy catalogue (Yess et al. 1997) and the Las Campanas Redshift Survey (LCRS) (Shandarin & Yess 1998), or through various types of shape statistics, as the moment–of–inertia method suggested by Babul & Starkman (1992) and applied by Davé et al. (1996) to the CfA1 redshift survey and by Sathyaprakash et al. (1998b) to the IRAS 1.2Jy cat-alogue. The most wide–spread technique is probably the genus statistics introduced by Gott III et al. (1986), which has been used on various IRAS catalogues (Moore et al. 1992; Protogeros & Weinberg 1997; Canavezes et al. 1998; Springel et al. 1998), on the CfA2 redshift survey (Vogeley et al. 1994b) and, in a modified way, on the nearly two–dimensional LCRS (Colley 1997). With galaxy clusters, scales of several hundred $h^{-1}\mathrm{Mpc}$ have been probed using differ-ent morphological statistics (Coles et al. 1998; Plionis et al. 1992), including the Minkowski functionals (Kerscher et al. 1997).

Figure 2.1 Examples for simple structures with positive, vanishing, and negative Euler characteristic.
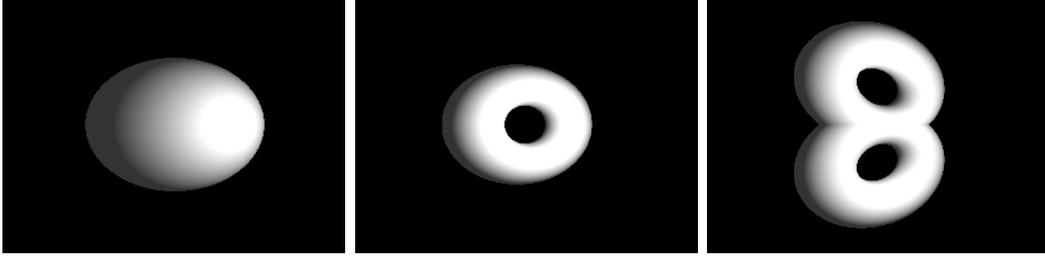
Among this plethora of methods, Minkowski functionals currently represent the most comprehensive description of the topological and geometric properties of a three-dimensional structure. Minkowski functionals include both the percolation analysis and the genus curve. Minkowski functionals can be applied both to density fields (Schmalzing & Buchert 1997) and to discrete point sets (Mecke et al. 1994). Because Minkowski functionals are additive they are robust measures even for sparse samples. Minkowski functionals also discriminate different cosmological models efficiently (Kerscher et al. 1997). Finally, Minkowski functionals can be used to define quantities which can be interpreted as a planarity $\mathcal{P}$ and a filamentary $\mathcal{F}$ and thus quantify our intuition about the shape of a structure (Sahni et al. 1998; Schmalzing et al. 1999a; Schmalzing & Diaferio 1999).

The present chapter contains a short summary of the mathematical foundations used in our work. We introduce the relevant results of integral geometry and related fields, and motivate their application to astrophysical datasets.

Table 2.1 Some of the $d + 1$ Minkowski functionals in d–dimensional Euclidean space may be interpreted as familiar geometric quantities (apart from numerical factors). This table summarizes the geometric interpretations of all Minkowski functionals for one, two and three dimensions. The symbol χ denotes the Euler characteristic, a purely topological quantity. It measures the connectivity of a set, being equal to the number of parts minus the number of tunnels.

d	1	2	3
V_0	length	area	volume
V_1	χ	circumference	surface area
V_2	–	χ	integrated mean curvature
V_3	–	–	χ

Figure 2.2 This plot shows the surface area and the integrated mean curvature, that is the Minkowski functionals V_1 and V_2, for spheroids of varying axis ratio. Prolate spheroids ("cigars") are characterised by large surface area and comparatively small mean curvature, while the situation is reversed for oblate spheroids ("pancakes").
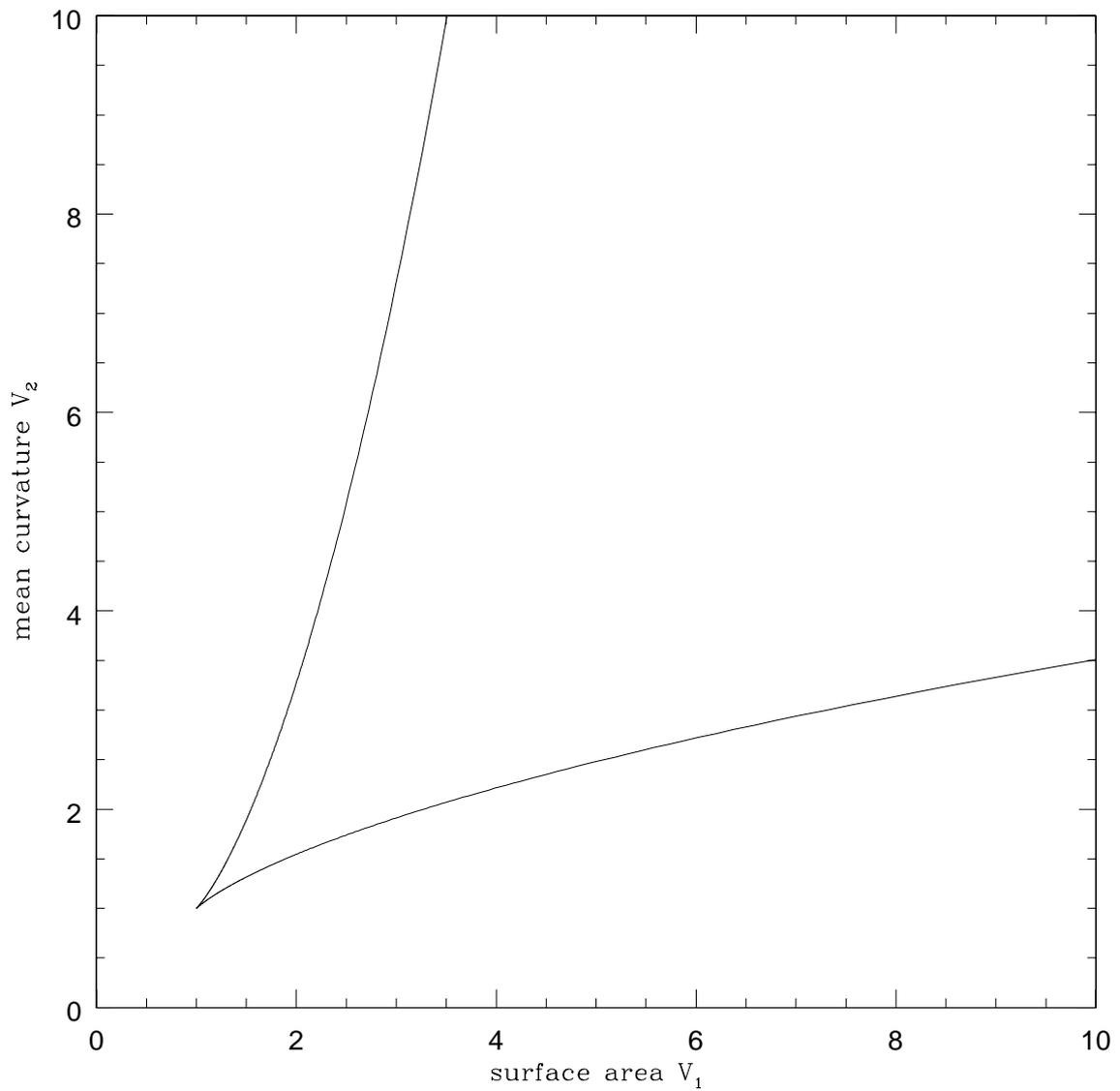
2.1 Overview

Stochastic problems that involved geometric objects provided a challenge for mathematicians since Buffon introduced his famous needle problem in the 18th century. Ambiguities such as Bertrand's paradox plagued the field until Poincaré found that it was possible to resolve them by associating probability measures with transformation groups. The work of Blaschke (1936) initiated the systematic study of probabilistic ideas in application to geometric problems, and started the evolution of what is now known as integral geometry. Most notably, Blaschke introduced the Quermaß integrals or Minkowski functionals. Later studies focussing on these quantities (Hadwiger 1957; Santaló 1976) provided further insight into the interplay of geometry and probability. Nowadays, the scope of the field has enlarged towards such diverse fields as differential geometry and stochastic processes, which is also reflected in the broad range of recent overviews (Weil 1983; Stoyan et al. 1995; Schneider 1993).

In our work, we are going to draw results from integral geometry and related fields to measure morphological properties of cosmological datasets. The central role in these studies is played by the Minkowski functionals. They were first introduced into large–scale structure analysis by Mecke et al. (1994), and later applied to catalogues of galaxies (Kerscher et al. 1998) and clusters of galaxies (Kerscher et al. 1997). A basic tutorial can be found in Schmalzing et al. (1996). Meanwhile, Minkowski functionals are enjoying increasing popularity as an alternative to the more standard low–order statistics.

One appealing property of the Minkowski functionals is their uniqueness and completeness as morphological measures, which follows by the theorem of Hadwiger (1957) on very general requirements. In d dimensions there exist $d + 1$ Minkowski functionals V_μ, $\mu = 0, \ldots, d$ on the convex ring, the set of all finite unions of convex bodies. Hadwiger's theorem states that any measure M on the convex ring that satisfies motion invariance, additivity, and continuity can be expressed as a linear combination of the Minkowski functionals. Furthermore, all the Minkowski functionals in up to three dimensions can be interpreted as well–known geometric quantities, as summarised in Table 2.1.

Figure 2.1 shows examples for bodies with different Euler characteristic. The Euler characteristic is a purely topological quantity and measures the connectivity of a set. It is equal to the number of parts minus the number of tunnels in the structure. Hence the egg has $V_3 = 1$, while the single and double loop have $V_3 = 0$ and $V_3 = -1$, respectively. The behaviour of the surface area V_1 and mean curvature V_2 is illustrated in Figure 2.2 on a simple family of spheroids, whose Minkowski functionals are known analytically (Hadwiger 1955). By varying the axis ratio of the spheroids while keeping the volume constant, we obtain a curve in the V_1–V_2–plane. Values are given in units of the corresponding functionals of a ball of the same size, so we are indeed showing the effects of shape. A ball would lie at the cusp at $(1, 1)$. Prolate spheroids ("cigars") are characterised by large surface area and comparatively small mean curvature, while the situation is reversed for oblate spheroids ("pancakes").

2.2 Point processes

Many phenomena, including the distribution of galaxies in the Universe, can be modelled by point processes. The underlying theory is explained in various standard works of mathematics (Diggle 1983; Daley & Vere-Jones 1988; Reiss 1993) and astrophysics (Peebles 1980). For our purpose, a point process distributes points on a d–dimensional support $\mathcal{D} \subset \mathbb{E}^d$, generating a random set of points $\{\mathbf{x}_i\}$, a so–called realization of the point process.

The complete statistical properties of a point process can be quantified using the density functions $\rho_n(\mathbf{x}_1 \ldots \mathbf{x}_n)$, where $n \in \mathbb{N}$. With probability $\rho_n(\mathbf{x}_1 \ldots \mathbf{x}_n)\mathrm{d}^d x_1 \ldots \mathrm{d}^d x_n$, a realization of the point process will contain points at the n positions $\mathbf{x}_1$ to $\mathbf{x}_n$ in infinitesimal volumes $\mathrm{d}^d x_1$ to $\mathrm{d}^d x_n$. All density functions can be summarised in a generating functional

$$\mathbf{P}[\mu(\mathbf{x})] = \sum_{n=1}^{\infty} \frac{1}{n!} \int \mathrm{d}\tau_n \rho_n(\mathbf{x}_1 \ldots \mathbf{x}_n)\mu(\mathbf{x}_1) \ldots \mu(\mathbf{x}_n), \tag{2.1}$$

with the integration measure defined by

$$\int \mathrm{d}\tau_n = \int_{\mathcal{D}} \mathrm{d}^d x_1 \ldots \mathrm{d}^d x_n. \tag{2.2}$$

An equivalent, but usually more meaningful description uses the hierarchy of correlation functions $\xi_n(\mathbf{x}_1 \ldots \mathbf{x}_n)$ (White 1979; Balian & Schaeffer 1989a; Balian & Schaeffer 1989b). The correlation function of order n gives the deviation of the density function of order n from the density function expected under the assumption that all partitions of the n points are independent. Furthermore, the correlation functions are normalised with the one–point particle density $\rho_1(\mathbf{x})$. While this definition cannot be written down in closed form for the individual functions, the connection between the generating functionals is very simple. Defining

$$\Xi[\mu(\mathbf{x})] = \sum_{n=1}^{\infty} \frac{1}{n!} \int \mathrm{d}\tau_n \xi_n(\mathbf{x}_1 \ldots \mathbf{x}_n)\mu(\mathbf{x}_1) \ldots \mu(\mathbf{x}_n), \tag{2.3}$$

in analogy to the generating functionals in Equation 2.1, we have

$$\Xi[\mu(\mathbf{x})] = \ln \mathbf{P}[\mu(\mathbf{x})/\rho_1(\mathbf{x})]. \tag{2.4}$$

Usually, we will deal with stationary point processes, where all characteristics are invariant under translations and rotations. In particular, this means that the one–point density $\rho_1(\mathbf{x})$ is independent of position, so we may write $\rho_1(\mathbf{x}) = \rho$.

Obviously, the morphology of a set of points is rather trivial. We have to put "flesh" on the skeleton before we can think of measuring topology and geometry. One possible approach (Mecke et al. 1994) uses a Boolean grain model (Wicksell 1925). From the set of points $\{\mathbf{x}_i\}$, we construct the union of balls B of a certain radius r centered around the points, that is

$$A_r = \bigcup_{i=1}^{N} B_{\mathbf{x}_i}. \tag{2.5}$$

We can then measure the Minkowski functionals $V_\mu(A_r)$ of this union set as functions of the radius. Throughout our applications, we divide these quantities by the sample volume, and thus obtain volume densities v_μ of Minkowski functionals.

We shall now establish a relation between the hierarchy of correlation functions of a point process and the average Minkowski functionals of a Boolean grain model constructed from the realizations. In order to cut down notation we always use a linear combination of Minkowski functionals, the so–called Minkowski polynomial $M(t)$ (Hadwiger 1975; Kellerer 1984). We define

$$M(t; K) = \sum_{\mu=0}^{d} \frac{\alpha_\mu t^\mu}{\mu!} V_\mu(K) \tag{2.6}$$

with coefficients[1]

$$\alpha_\mu = \frac{\omega_{d-\mu}}{\omega_d}. \tag{2.8}$$

Analogously, we construct a polynomial $m(t)$ from the densities v_μ of the Minkowski functionals per unit volume. From now on, we drop the variable t from the polynomial, and only reintroduce it when extracting the individual Minkowski functionals via

$$V_\mu(K) = \frac{1}{\alpha_\mu} \frac{\partial^\mu}{\partial t^\mu} M(t; K). \tag{2.9}$$

By taking the ensemble average and applying the additivity relation repeatedly, we obtain the formula given by Mecke et al. (1994),

$$m = 1 - \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} \int \mathrm{d}\tau_n \rho_n(\mathbf{x}_1, \dots, \mathbf{x}_n) M(B_{\mathbf{x}_1} \cap \dots \cap B_{\mathbf{x}_n}), \tag{2.10}$$

If the density functions ρ_n were independent of position, the integrals could be performed using the principal kinematic formula by Blaschke (1936) (see also Chern & Yien 1940 for the more general version employed here). Fixing one body A in space, and intersecting with a moving body B, we obtain after integration over all possible positions of B a simple factorization.

$$\int_{\mathcal{D}} \mathrm{d}^d x \, M(A \cap B_{\mathbf{x}}) = M(A)M(B), \tag{2.11}$$

---

[1]ω_μ denotes the volume of the μ–dimensional unit ball. Some important special values are $\omega_0 = 1$, $\omega_1 = 2$, $\omega_2 = \pi$, and $\omega_3 = \frac{4}{3}\pi$, while in general.

$$\omega_j = \frac{2\pi^{(j+1)/2}}{\Gamma((j+1)/2)}. \tag{2.7}$$

where equality holds to order t^d. In the case of the homogeneous Poisson process, where all higher–order correlation functions vanish, straightforward application of the principal kinematic formula to Equation (2.10) leads to the results given by Mecke (1994). Unfortunately, this approach is not viable in most situations where the density functions are position dependent. However, it is possible to using the factorization of the density functions into the correlation functions.

Equation (2.10) motivates the introduction of a hierarchy of "moments" P_n by weighting the n–point density function with the Minkowski polynomial of n intersecting balls,

$$P_n = \int \mathrm{d}\tau_n \rho_n \left(\mathbf{x}_1, \dots , \mathbf{x}_n \right) M \left(B_{\mathbf{x}_1} \cap \dots \cap B_{\mathbf{x}_n} \right). \tag{2.12}$$

Writing the density functions ρ_n as sums over all possible partitions into correlation functions ξ_n, and using the principal kinematic formula, it can be shown that the "cumulants" corresponding to the moments P_n are the Ξ_n given by

$$\Xi_n = \int \mathrm{d}\tau_n \xi_n \left(\mathbf{x}_1, \dots , \mathbf{x}_n \right) M \left(B_{\mathbf{x}_1} \cap \dots \cap B_{\mathbf{x}_n} \right). \tag{2.13}$$

With a generating functional $P[j]$ for the P_n given by

$$P[j] = \sum_{n=0}^{\infty} \frac{(-j)^n}{n!} P_n, \tag{2.14}$$

an analogous generating functional $\Xi[j]$ for the Ξ_n is obtained through

$$P[j] = \exp \left(\Xi[\rho j] \right). \tag{2.15}$$

Note the analogy with Equation (2.1). Setting $j = -1$, we directly arrive at the result

$$m = 1 - \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} P_n = 1 - \exp \left(\sum_{n=0}^{\infty} \frac{(-\rho)^n}{n!} \Xi_n \right). \tag{2.16}$$

It is interesting to note that the computational method used here can be regarded as a special case from the family of linked cluster theorems, which are often encountered in statistical physics (Negele & Orland 1988; Binney et al. 1992).

Recovering the Minkowski functionals from the polynomial via Equation (2.9), we obtain in $d = 3$ dimensions

$$\begin{aligned}
v_0 &= 1 - \exp \left(-\rho \overline{V}_0 \right) \\
v_1 &= \exp \left(-\rho \overline{V}_0 \right) \rho \overline{V}_1 \\
v_2 &= \exp \left(-\rho \overline{V}_0 \right) \left(\rho \overline{V}_2 - \frac{3\pi}{8} \rho^2 \overline{V}_1^2 \right) \\
v_3 &= \exp \left(-\rho \overline{V}_0 \right) \left(\rho \overline{V}_3 - \frac{9}{2} \rho^2 \overline{V}_1 \overline{V}_2 + \frac{9\pi}{16} \rho^3 \overline{V}_1^3 \right)
\end{aligned} \tag{2.17}$$

which looks remarkably similar to the result for the Poisson process, except that the bare Minkowski functionals of balls $V_\mu(B)$ are replaced by the quantities

$$\overline{V}_\mu = V_\mu(B) + \sum_{n=1}^{\infty} \frac{(-\rho)^n}{(n+1)!} \int_{\mathcal{D}} \mathrm{d}^d x_1 \ldots \mathrm{d}^d x_n$$

$$\xi_{n+1}(0, \mathbf{x}_1, \ldots, \mathbf{x}_n) V_\mu(B \cap B_{\mathbf{x}_1} \cap \ldots \cap B_{\mathbf{x}_n}). \quad (2.18)$$

Note that Equation (2.18) is a generalization of the result by White (1979) for the void probability function. This shows that all Minkowski functionals depend on the complete hierarchy of correlation functions through an asymptotic series, hence each order n can be expected to contribute significantly to the overall result. Comparisons on a high–resolution simulation of the dark–matter and halo distribution in the nearby Universe (Schmalzing et al. 1999b) have shown that Minkowski functionals do indeed reveal information beyond clustering statistics of low order, and are therefore a useful complement to the more standard correlation function and power spectrum analysis.

However, the Boolean grain model is not the only way of constructing non–trivial bodies from point sets. One can also smear the individual particles into a continuous field, and the first section of Chapter 5 discusses a few techniques for doing this. Moreover, not all data comes in points; therefore, we will leave point processes for a while and consider random fields.

2.3   Random fields

In principle, one can think of a random field $u(\mathbf{x})$ as a multivariate random variable, with a spatial domain $\mathcal{D}$ as index set. The statistical properties of a random field are fully specified by the statistics of all finite subsets $\{\mathbf{x}_1 \ldots \mathbf{x}_N\}$ from this domain, hence the problem is reduced to one of ordinary vector–valued random variables. The geometric properties of the index domain, however, allow for the definition quantities such as derivatives or integrals of random fields, which are random fields themselves. A full–blown mathematical course on the theory of random fields can be found in the classical book by Adler (1981). In the following, we will consider ways of evaluating Minkowski functionals given a realisation of a random field $u(\mathbf{x})$.

Consider a random field $u(\mathbf{x})$ on a d–dimensional support $\mathcal{D} \subseteq \mathbb{E}^d$. For a given threshold ν the excursion set Q_ν is the set of all points $\mathbf{x}$ with $u(\mathbf{x}) \geq \nu$. We wish to calculate its Minkowski functionals $v_k^{(d)}(\nu)$ per unit volume. Using methods from integral geometry and scale–space theory, we can develop formulae for the Minkowski functionals that can be easily turned into numerical prescriptions (Schmalzing & Buchert 1997). One set of estimators uses Crofton's formula and grid cell counting, while the other employs Koenderink invariants to reduce the problem to spatial averaging over invariants formed from derivatives of the random field.

Careful analysis reveals that both families of estimators are biased by the finite lattice constant a encountered with any practical realisation of a random field. The deviations are of order a^2 and can be evaluated analytically for a Gaussian random field (see Winitzki & Kosowsky 1997 for

some results in two dimensions). It turns out that the two estimates using Crofton's formula and Koenderink invariants, respectively, deviate from the true value in opposite directions. Hence their difference gives at least a rough idea of the errors associated with binning the continuous field onto a finite lattice.

Yet another method of calculating Minkowski functionals of isocontours has been suggested by Novikov et al. (1999). However, their interpolation method has so far only been implemented on small two–dimensional fields.

2.3.1   Crofton's intersection formula

Crofton's formula (Crofton 1868) relates the Minkowski functionals from different dimensions. For a body K in d dimensions we consider an arbitrary k–dimensional hyperplane E and calculate the Euler characteristic χ of the intersection $K \cap E$ in k dimensions. Integrating this quantity over the space $\mathcal{E}_k^{(d)}$ of all conceivable hyperplanes[2] we obtain the kth Minkowski functional of the body K in d dimensions. Crofton's formula states that

$$v_k^{(d)}(K) = \frac{\omega_d}{\omega_{d-k}\omega_k} \int_{\mathcal{E}_k^{(d)}} \mathrm{d}\mu_k(E) \chi^{(k)}(K \cap E). \tag{2.19}$$

If the body K is the excursion set of a homogeneous and isotropic random field sampled at L points of a cubic lattice of spacing a, we can as well average over the set $\mathcal{L}_k^{(d)}$ of all dual lattice hyperplanes only. The integral is then replaced by a summation over these hyperplanes with an additional normalization factor and we obtain

$$v_k^{(d)}(K) = \frac{\omega_d}{\omega_{d-k}\omega_k} \sum_{E \in \mathcal{L}_k^{(d)}} \frac{1}{a^k L} \frac{k!(d-k)!}{d!} \chi^{(k)}(K \cap E). \tag{2.20}$$

However, the Euler characteristic of a body in k dimensions can be approximated by counting the numbers $N_j(K)$ of all elementary lattice cells of dimension j within the body (Adler 1981). In three dimensions, for example, $N_3(K)$ gives the number of cubes within the body, $N_2(K)$ and $N_1(K)$ are the numbers of faces and edges of the lattice cubes, respectively, while $N_0(K)$ gives the number of lattice points contained in K. In general we have

$$\chi^{(k)}(K \cap E) = \sum_{j=0}^{k} (-1)^j N_j(K \cap E). \tag{2.21}$$

When summing equation (2.21) over all lattice hyperplanes, we notice that j–dimensional lattice cells of the bodies $K \cap E$ result from intersections of a $(d-k+j)$–dimensional lattice cell belonging to the body K with k–dimensional hyperplanes E, so there are $\begin{pmatrix} d-k+j \\ j \end{pmatrix}$

---

[2]The integration measure $\mathrm{d}\mu_k(E)$ is normalized to give $\int_{\mathcal{E}_k^{(d)}} \mathrm{d}\mu_k(E) = 1$.

of them. Therefore, we can actually perform the summation

$$\sum_{E \in \mathcal{L}_k^{(d)}} N_j(K \cap E) = \binom{d-k+j}{j} N_{d-k+j}(K) \tag{2.22}$$

and end up with the compact formula

$$v_k^{(d)}(K) = \frac{\omega_d}{\omega_{d-k}\omega_k} \frac{1}{a^k L} \sum_{j=0}^{k} (-1)^j \frac{k!(d-k+j)!}{d!j!} N_{d-k+j}(K). \tag{2.23}$$

For a homogeneous and isotropic random field in three dimensions the Minkowski functionals are thus given by

$$
\begin{aligned}
v_0(\nu) &= n_3, \\
v_1(\nu) &= \frac{2}{9a}(n_2 - 3n_3), \\
v_2(\nu) &= \frac{2}{9a^2}(n_1 - 2n_2 + 3n_3), \\
v_3(\nu) &= \frac{1}{a^3}(n_0 - n_1 + n_2 - n_3),
\end{aligned}
\tag{2.24}
$$

where the quantity $n_j = N_j/L$ is the number density of j–dimensional elementary cells contained in the isodensity contour; to be specific, n_3 is the number of grid volumes, n_2 counts the grid faces, and n_1 and n_0 denote the numbers of grid edges and lattice points, respectively.

Note that the last relation for the Euler characteristic $v_3 = \chi/|\mathcal{D}|$ per unit volume is not obtained from Crofton's formula but is already given as a special case of equation (2.21) which results from the Morse theorem (Morse & Cairns 1969). Being conceptually simple in comparison to the other approximations, it was already proved rigorously by Adler (1981) and used by Coles et al. (1996) to calculate the genus of isodensity contours in three dimensions. Adding Crofton's formula gives all the Minkowski functionals without further computational effort, since the calculation requires nothing but the counting of all elementary lattice cells, which is already necessary for calculating the genus.

2.3.2 Koenderink invariants

Again we consider a realisation of a random field $u(\mathbf{x})$ on a d–dimensional support $\mathcal{D} \subseteq \mathbb{E}^d$ and the excursion set Q_ν over a given threshold ν. We wish to calculate its Minkowski functionals $v_k^{(d)}(\nu)$ per unit volume. The volume functional v_0 is simply given by volume integration over the excursion set Q_ν, which can be written with a Heaviside step function Θ as

$$v_0(\nu) = \frac{1}{|\mathcal{D}|} \int_{\mathcal{D}} \mathrm{d}^d x \, \Theta(\nu - u(\mathbf{x})). \tag{2.25}$$

If the excursion set Q_ν has a smooth boundary ∂Q_ν, its other Minkowski functionals are given by surface integrals (Schneider 1978). We have

$$v_j(\nu) = \frac{1}{|\mathcal{D}|} \frac{1}{\omega_{j-1} \binom{d}{j}} \int_{\partial Q_\nu} \mathrm{d}s \, \mathfrak{S}_j \left(\kappa_1 \ldots \kappa_{d-1} \right). \tag{2.26}$$

Here $\mathrm{d}s$ denotes the surface element on Q_ν, κ_1 to κ_{d-1} are the boundary's $d-1$ principal curvatures, and $\mathfrak{S}_j$ is the jth elementary symmetric function defined by the polynomial expansion

$$\prod_{i=1}^{d-1} (x + \kappa_i) = \sum_{j=1}^{d} x^{d-j} \mathfrak{S}_j \left(\kappa_1 \ldots \kappa_{d-1} \right); \tag{2.27}$$

hence $\mathfrak{S}_1 = 1$, $\mathfrak{S}_2 = \kappa_1 + \ldots + \kappa_{d-1}$, and so on up to $\mathfrak{S}_d = \kappa_1 \ldots \kappa_{d-1}$.

For practical applications, it is quite cumbersome to follow the isodensity contour and evaluate the local curvatures directly. We use methods developed by Koenderink (Koenderink 1984; ter Haar Romeny et al. 1991) in two–dimensional image analysis to express these local curvatures in terms of geometric invariants formed from first and second derivatives. The calculations in two and three dimension were already done by Schmalzing (1996), so let us just summarise the results. In three dimensions, the isodensity contour is a surface and has two principal curvatures κ_1 and κ_2, and we obtain[3]

$$\kappa_1 + \kappa_2 = \frac{\epsilon_{ijm} \epsilon_{klm} u_{,i} u_{,jk} u_{,l}}{(u_{,n} u_{,n})^{3/2}},$$
$$\kappa_1 \kappa_2 = \frac{\epsilon_{ijk} \epsilon_{lmn} u_{,i} u_{,l} u_{,jm} u_{,kn}}{2(u_{,p} u_{,p})^2}. \tag{2.28}$$

In two dimensions, the isodensity contour is a line and hence there is only one principal curvature at each point,

$$\kappa = \frac{2 u_{,1} u_{,2} u_{,12} - u_{,1}{}^2 u_{,22} - u_{,2}{}^2 u_{,11}}{(u_{,1}{}^2 + u_{,2}{}^2)^{3/2}}. \tag{2.29}$$

Moreover, the surface integration (2.26) of these invariants may be transformed to the spatial mean over the whole volume. This adds to the integral a delta function for selecting the isodensity contour, and the field's gradient as the Jacobian of the transformation of the surface element. The resulting formula is

$$v_j(\nu) = \frac{1}{|\mathcal{D}|} \frac{1}{\omega_{j-1} \binom{d}{j}} \int_{\mathcal{D}} \mathrm{d}^d x \, \delta(\nu - u(\mathbf{x})) \left| \nabla u(\mathbf{x}) \right| \mathfrak{S}_j \left(\kappa_1 \ldots \kappa_{d-1} \right). \tag{2.30}$$

---

[3]The notation is chosen to emphasise the coordinate invariance of the integrand. We use indices following a comma to indicate differentiation with respect to the corresponding coordinate. ϵ_{ijk} denotes the components of the totally antisymmetric third–rank tensor normalised to $\epsilon_{123} = 1$. Summation over pairwise indices is understood.

If the density field is sampled at the points of some lattice, we can therefore estimate the Minkowski functionals by calculating the derivatives at each lattice point and replacing first the δ-function with a bin of finite width and then the average over all space with a summation over lattice points only.

Spaces of constant curvature

A promising cosmological application of Minkowski functionals uses random fields on the sphere, namely all–sky maps of the CMB. The two–dimensional sphere $\mathbb{S}^2$ belongs to the wider class of spaces of constant curvature.

A space of constant curvature is characterised by its dimension d, and by its curvature kK. The sign k equals either $+1$, or 0, or -1, for the spherical space $\mathbb{S}^d$, the Euclidean space $\mathbb{E}^d$ and the hyperbolic space $\mathbb{H}^d$, respectively. K is a positive constant of dimension $[\text{Length}]^{-2}$, hence its inverse square root $K^{-1/2}$ can be interpreted as the radius of curvature. Santaló (1976) shows how to obtain an integral geometry on such spaces. Curvature integrals as in Equation (2.26) can still be defined, if care is taken to use the geodesic curvatures κ_i. In the following, we will call these quantities the Minkowski functionals in curved spaces.

It turns out that some of the geometric interpretations are altered with respect to the flat case. While for a body K in flat space the curvature integral $V_d(K)$ is equal to the Euler characteristic $\chi(K)$, curved spaces require a generalized Gauss–Bonnet Theorem proved for arbitrary Riemannian manifolds by Allendoerfer & Weil (1943) and Chern (1944). The theorem states that the Euler characteristic is a linear combination of all Minkowski functionals as defined by Equation (2.26),

$$\chi(K) = \sum_{j=0}^{d} c_j V_j(K), \tag{2.31}$$

with the coefficients c_j given by

$$c_j = \begin{cases} \dbinom{d}{j} \dfrac{2(kK)^{(d-j)/2}}{\omega_{d-j}} & \text{if } d-j \text{ even,} \\ 0 & \text{if } d-j \text{ odd.} \end{cases} \tag{2.32}$$

Note that from the point of view of Hadwiger's theorem, which is also valid on curved spaces, all linear combinations of Minkowski functionals are equally suitable as morphological descriptors, so one may both use the integrated curvature V_d and the Euler characteristic χ as the last Minkowski functional. In the following, we will consider both quantities, because the integrated geodesic curvature is easier to calculate, and the Euler characteristic is easier to interpret. Obviously, in the case of Euclidean space $\mathbb{E}^d$, $k=0$ and all coefficients apart from c_d vanish, so $\chi(K) = V_d(K)$ and the original Gauss–Bonnet theorem is recovered.

2.3.3 Tomita's Formulae for Statistics of Random Interfaces

Just like the genus (Doroshkevich 1970), all Minkowski functionals are analytically known for a Gaussian random field, due to a highly instructive article by Tomita (1990). Meanwhile, it should be clear that we consider a random field $u(\mathbf{x})$ in d dimensions, construct the excursion set over a certain threshold ν, and calculate its mean Minkowski functionals $v_k^{(d)}(\nu)$ per unit volume.

In the previous section, Equation (2.30) has shown that all Minkowski functionals may be written as spatial averages over Koenderink invariants, which can in turn be expressed in terms of first and second derivatives of the field under consideration. We may as well average over all possible realizations at a fixed location, and hence it is sufficient to know the joint probability distribution of the field value and its first and second derivatives at the same point. The situation is particularly simple for the Gaussian random field, where we obtain a multivariate Gaussian distribution (Adler 1981) with zero mean and covariances given in terms of the field's correlation function $\xi(r)$ as

$$
\begin{aligned}
\langle u^2 \rangle &= \xi(0) \\
\langle u u_{,i} \rangle &= 0 \\
\langle u u_{,ij} \rangle &= -\xi''(0)\delta_{ij} \\
\langle u_{,i} u_{,j} \rangle &= \xi''(0)\delta_{ij} \\
\langle u_{,i} u_{,jk} \rangle &= 0 \\
\langle u_{,ij} u_{,kl} \rangle &= \tfrac{1}{3}\xi^{(4)}(0)(\delta_{ij}\delta_{kl} + \delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk}).
\end{aligned}
\tag{2.33}
$$

The task of inserting the corresponding probability density into the average of Equation (2.30) is a relatively simple one. It turns out that the result depends only on the value of the correlation function $\xi(0)$ and its second derivative $\xi''(0)$ at zero through two parameters σ and λ, with

$$
\sigma = \xi(0), \qquad \lambda = \sqrt{|\xi''(0)|/(2\pi\xi(0))}.
\tag{2.34}
$$

Both are easily measured from a given realization of the random field, since $\xi(0) = \langle u^2 \rangle$ is the variance of the field itself and $|\xi''(0)| = \langle u_{,i}^2 \rangle$ is the variance of any of its first derivatives (Adler 1981).

With these two parameters all Minkowski functionals in arbitrary dimension d can be calculated for a Gaussian random field (Tomita 1990). For the volume functional we obtain[4]

$$
v_0^{(d)}(\nu) = \tfrac{1}{2} - \tfrac{1}{2}\Phi\left(\frac{\nu}{\sqrt{2}\sigma}\right),
\tag{2.35}
$$

while all other Minkowski functionals are given by

$$
v_k^{(d)}(\nu) = \lambda^k \frac{\omega_d}{\omega_{d-k}\omega_k} H_{k-1}\left(\frac{\nu}{\sqrt{\sigma}}\right).
\tag{2.36}
$$

---

[4] $\Phi(x) = \frac{2}{\sqrt{\pi}}\int_0^x \mathrm{d}t \exp(-t^2)$ denotes the Gaussian error integral, while $H_n(x) = \left(-\frac{\mathrm{d}}{\mathrm{d}x}\right)^n \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}x^2\right)$ defines a Hermite function of order n.

Putting $d = 3$ we obtain the Minkowski functionals in three dimensions[5]

$$
\begin{aligned}
v_0(\nu) &= \tfrac{1}{2} - \tfrac{1}{2}\Phi\left(\tfrac{1}{\sqrt{2}}u\right) \\
v_1(\nu) &= \tfrac{2}{3}\frac{\lambda}{\sqrt{2\pi}}\exp\left(-\tfrac{1}{2}u^2\right) \\
v_2(\nu) &= \tfrac{2}{3}\frac{\lambda^2}{\sqrt{2\pi}}u\exp\left(-\tfrac{1}{2}u^2\right) \\
v_3(\nu) &= \frac{\lambda^3}{\sqrt{2\pi}}\left(u^2 - 1\right)\exp\left(-\tfrac{1}{2}u^2\right).
\end{aligned}
\tag{2.37}
$$

For other random fields derived from the Gaussian case, such as the χ^2 field, the F field, or the t field, Worsley (1994) gives the joint probability distribution necessary for calculating the Minkowski functionals through Equation (2.30). Some additional results including figures can be found in Chapter 4.

2.4 Isoperimetric inequalities

Minkowski functionals are no exception to Berry's principle – the original article by (Minkowski 1903) does *not* introduce the Quermaß integrals or Minkowski functionals, but rather related quantities known as mixed volumes. The mixed volume $V(K_1, \dots, K_n)$ of n convex bodies K_1 to K_n is a generalization of the ordinary volume $V(K)$ of a single body K. It can be expressed in terms of ordinary volumes of Minkowski sums of all combinations of the n bodies, and we have

$$
V(K_1, \dots, K_n) := \frac{1}{n!}\sum_{k=1}^{n}(-1)^{n+k}\sum_{i_1 < \dots < i_k} V(K_{i_1} \oplus \dots \oplus K_{i_k}),
\tag{2.38}
$$

where the Minkowski sum is defined as

$$
K \oplus L := \{x + y \,|\, x \in K, y \in L\};
\tag{2.39}
$$

Obviously, the Minkowski sum of two convex bodies is itself convex. It is clear from this definition that in the case of a single body, the mixed volume reduces to its ordinary volume. Furthermore, these quantities inherit many interesting properties from the volume and hence play a central role in the Brunn–Minkowski theory (see Schneider 1993 for an introduction).

A most useful inequality proved by Alexandrov (1937) and Fenchel (1936) states that

$$
V^2(K_1, K_2, \dots, K_n) \geq V(K_1, K_1, K_3 \dots, K_n)\, V(K_2, K_2, K_3 \dots, K_n).
\tag{2.40}
$$

---

[5]To relieve the notation of the overall normalization $\sqrt{\sigma}$, we use the dimensionless argument u, with $\nu = u\sqrt{\sigma}$.

The Minkowski functionals of a body K in d dimensions can be related to mixed volumes of two bodies via

$$V_j(K) = \frac{1}{\omega_j} V(\underbrace{K,\dots,K}_{d-j}, \underbrace{B,\dots,B}_{j}), \tag{2.41}$$

where B denotes the d–dimensional unit ball. Consult the book by Burago & Zalgaller (1988) for a broad discussion of the Alexandrov–Fenchel inequality and related issues.

In three dimensions, the Alexandrov–Fenchel inequality (2.40) leads to two independent non–trivial inequalities for the Minkowski functionals, namely

$$\begin{aligned} V_1{}^2 &\geq \frac{\pi}{4} V_0 V_2, \\ V_2{}^2 &\geq \frac{8}{3\pi} V_1 V_3. \end{aligned} \tag{2.42}$$

These two inequalities motivate the introduction of the Blaschke diagram (Hadwiger 1955). A convex body with Minkowski functionals V_j is mapped to a point (x, y) with coordinates

$$x := \frac{\pi V_0 V_2}{4 V_1^2}, \qquad y := \frac{8 V_1 V_3}{3 \pi V_2^2}. \tag{2.43}$$

A ball has $x = y = 1$. For convex bodies, all Minkowski functionals are non–negative, and satisfy the inequalities (2.42), so these points are confined to the unit square. Nevertheless, the convex bodies do not fill the whole unit square – the limiting isoperimetric inequality has yet to be found (Schneider 1993 and references therein).

Recently, Sahni et al. (1998) proposed a set of shapefinders derived from Minkowski functionals. One starts from the three independent ratios of Minkowski functionals that have dimension of length. Requiring that they yield the radius R if applied to a ball, we define

$$\text{Thickness } T := \frac{V_0}{2 V_1}, \qquad \text{Width } W := \frac{2 V_1}{\pi V_2}, \qquad \text{Length } L := \frac{3 V_2}{4 V_3}. \tag{2.44}$$

By the isoperimetric inequalities (2.42), we have $L \geq W \geq T$ for any convex body. Going one step further, Sahni et al. (1998) also define dimensionless shapefinders by

$$\text{Planarity } \mathcal{P} := \frac{W - T}{W + T}, \qquad \text{Filamentarity } \mathcal{F} := \frac{L - W}{L + W}. \tag{2.45}$$

These dimensionless shapefinders are related to the isoperimetric ratios via

$$\mathcal{P} = \frac{1 - x}{1 + x}, \qquad \mathcal{F} = \frac{1 - y}{1 + y}. \tag{2.46}$$

Obviously, a scatter plot of shapefinders is almost equivalent to the ordinary Blaschke diagram reflected at the point $(0.5, 0.5)$, so either method of presentation should convey the same morphological information. In our applications, we refer to scatter plots of shapefinders $(\mathcal{P}, \mathcal{F})$ as Blaschke diagrams.

3 The evolution of large–scale structure

Dich leder ich doch mit der Horex ab, Macker!
Wir baun einfach zwei Motor'n hinteranaer, denn siehst Du kein Land mehr!
(Brösel 1986)

According to the standard model of cosmology, the Universe evolved into its present state from a Big Bang, that is a hot, dense, and highly homogeneous state. The remains of this primeval fireball reach us in the photons of the Cosmic Microwave Background (CMB); having originated at the epoch of recombination, when matter and radiation decoupled, they reflect the distribution of matter when the Universe was a mere 300,000 years old. Measurements by the DMR instrument on board the COBE satellite have revealed very small fluctuations, of the order of 10^{-5}, in the CMB (Smoot et al. 1992). At the time of its origin, radiation and matter were still coupled, so fluctuations in the matter density were equally small. Since then, gravitational interaction alone has dominated the evolution of structures in the Universe.

In the following, we attempt to shed some light on the mechanisms of large–scale structure formation from an almost homogeneous initial state. It has been suggested (Pogosyan et al. 1996; Bond et al. 1996) that a "cosmic web" of filaments is already visible in the seeds of large–scale structure, in the initial fluctuations. Systematic investigations, e.g. by Arnol'd et al. (1982) in two dimensions, were often pursued with special focus on the Zel'dovich approximation. We will follow a similar approach. The framework for our considerations, Newtonian cosmology, is laid out in Section 3.1, and formulated in terms of the Lagrange–Newton system. In Section 3.2 we present and investigate a perturbation ansatz for this system of non–linear partial differential equations. After discussing some practical issues (Section 3.3), we present examples of particle distributions evolved from Cold Dark Matter initial conditions (Section 3.4), and conclude with an outlook on the prospects of our approximation scheme.

3.1   Newtonian cosmology

Let us begin by a brief summary of Newtonian cosmology (Peebles 1980; Padmanabhan 1993 and references therein). We assume that spacetime is well approximated by a Friedmann–

Robertson–Walker metric (see e.g. Weinberg 1972)

$$ds^2 = dt^2 - a(t)^2 \left(\frac{dr^2}{1 - kr^2} + r^2 d\Omega^2 \right), \tag{3.1}$$

and that inhomogeneities act predominantly as a one–component fluid. In the Eulerian picture, the state of this fluid is fully described by the density field $\rho(\mathbf{x}, t)$ and the velocity field $\mathbf{v}(\mathbf{x}, t)$ at each point $(\mathbf{x}, t)$ in the non–rotating Eulerian frame. Furthermore, the dynamical evolution of the system is governed by the acceleration field $\mathbf{g}(\mathbf{x}, t)$ which equals the gravitational field strength according to Einstein's equivalence principle.

Usually, this fluid is taken to be the dark matter, and accordingly modelled as pressureless dust. The dynamics of the Eulerian fields are contained in a system of partial differential equations: This system consists of the continuity equation, the Euler equation for zero pressure and field equations for the acceleration field. Together, these equations form the Euler–Newton system

$$\begin{aligned} \partial_t \rho + \nabla (\rho \mathbf{v}) &= 0 \\ \partial_t \mathbf{v} + (\mathbf{v}\nabla) \mathbf{v} &= \mathbf{g} \\ \nabla \mathbf{g} &= \Lambda - 4\pi G \rho \\ \nabla \times \mathbf{g} &= 0. \end{aligned} \tag{3.2}$$

A completely different, but equally complete description is obtained through the Lagrangian picture. We introduce Lagrangian coordinates $\mathbf{X}$, and map them onto the Eulerian positions $\mathbf{x}$ by

$$\mathbf{x} = \mathbf{f}(\mathbf{X}, t), \tag{3.3}$$

where the Lagrangian and Eulerian positions are chosen to coincide at some early time t_i,

$$\mathbf{X} = \mathbf{f}(\mathbf{X}, t_i). \tag{3.4}$$

The mapping $\mathbf{f}$ is sufficient to describe the complete system: The Eulerian density field is obtained from the continuity equation via $\rho = \rho_i / J$, where $J := \det \frac{\partial f_i}{\partial X_j}$ is the Jacobian of the mapping, the Eulerian velocity field equals[1] $\mathbf{v} = \dot{\mathbf{f}}$ by definition, and the Euler equation leads to $\mathbf{g} = \ddot{\mathbf{f}}$ for the acceleration field.

Transforming the field equations for $\mathbf{g}$ to the Lagrangian frame and using the integrals of acceleration and density leads to the Lagrange–Newton system. Introducing the functional determinant

$$\mathcal{J}(A, B, C) := \frac{\partial (A, B, C)}{\partial (X_1, X_2, X_3)} = \det \begin{pmatrix} A_{|1} & B_{|1} & C_{|1} \\ A_{|2} & B_{|2} & C_{|2} \\ A_{|3} & B_{|3} & C_{|3} \end{pmatrix}, \tag{3.5}$$

---

[1] The dot $\dot{}$ denotes the Lagrangian or convective time derivative given in terms of the Eulerian fields by $\dot{} := \partial_t + \mathbf{v}\nabla$. Furthermore, indices following a comma denote derivatives with respect to Eulerian coordinates, while derivation in the Lagrangian frame is indicated by indices following a vertical slash. Thus $u_{,i} := \partial u / \partial x_i$ and $u_{|i} := \partial u / \partial X_i$.

the field equations from Equation (3.2) become

$$\mathcal{J}(\ddot{f}_1, f_1, f_3) + \mathcal{J}(\ddot{f}_2, f_2, f_3) = 0$$
$$\mathcal{J}(\ddot{f}_2, f_2, f_1) + \mathcal{J}(\ddot{f}_3, f_3, f_1) = 0 \tag{3.6}$$
$$\mathcal{J}(\ddot{f}_3, f_3, f_2) + \mathcal{J}(\ddot{f}_1, f_1, f_2) = 0$$
$$\mathcal{J}(\ddot{f}_1, f_2, f_3) + \mathcal{J}(\ddot{f}_2, f_3, f_1) + \mathcal{J}(\ddot{f}_3, f_1, f_2) = \Lambda \mathcal{J}(f_1, f_2, f_3) - 4\pi G \rho_{\mathrm{i}} \tag{3.7}$$

While the number of functions involved has been reduced compared to the Euler–Newton system, the Lagrange–Newton system still consists of non–linear partial differential equations and is not exactly solvable. The brute–force approach of integrating the equations numerically usually follows the Eulerian description, but analytical considerations are generally more fruitful if the Lagrangian point–of–view is adopted. In the following, we will outline a perturbative expansion of the Lagrange–Newton system around the known solution of a homogeneous, expanding Universe.

3.2   Perturbative approach to the Lagrange–Newton system

A variety of approximation schemes for gravitational clustering have been proposed over the years, and are excellently reviewed and compared by Sahni & Coles (1995). Focusing on the Lagrange–Newton system, Buchert (1989) developed a systematic perturbation theory for the system of Equations (3.6) and (3.7). Various in–depth studies (Buchert 1992; Bildhauer et al. 1992; Bouchet et al. 1992) and basic tutorials (Buchert 1996; Bouchet et al. 1995) can be found in the literature.

It has been demonstrated that higher than second order solutions (Buchert & Ehlers 1993; Buchert 1994; Catelan 1995; Lachièze-Rey 1993) are not feasible from the practical point of view, since little accuracy is gained while calculations become increasingly messy. After all, the first–order solution, a generalized case of the classical Zel'dovich approximation (Zel'dovich 1970a), already reproduces the formation of filaments. The second–order solution also accounts for tidal forces and features secondary shell crossings (Buchert et al. 1997), which break the filaments into smaller lumps. Moreover, second–order schemes lead to a substantial correction of the delayed collapse time at first order (Karakatsanis et al. 1997).

3.2.1   The perturbation ansatz

Here, we use a perturbation ansatz up to second order that reduces to a set of easily tractable ordinary differential equations. The perturbative expansion reads

$$\mathbf{f}(\mathbf{X}, t) = a(t)\mathbf{X} + \varepsilon a(t) q_1(t) \nabla \Psi^{(1)}(\mathbf{X}) + \varepsilon^2 a(t) q_2(t) \nabla \Psi^{(2)}(\mathbf{X}). \tag{3.8}$$

3 The evolution of large–scale structure

From now on, all spatial derivatives are taken with respect to Lagrangian coordinates unless explicitly stated otherwise. The zeroth order of Equation (3.8) describes a homogeneous expanding Universe with scale factor $a(t)$, and the first order is known as the Zel'dovich approximation (Zel'dovich 1970b). Note that this is *not* the full Lagrangian perturbation approach[2]. We already use the result of Ehlers & Buchert (1997) that each order of perturbation theory separates into a time–dependent factor $q_i(t)$ and a Lagrangian position dependent factor, which is in addition irrotational and hence expressible as the gradient of a scalar potential $\Psi^{(i)}(\mathbf{X})$. Since the zeroth order is the homogeneous background solution of a uniformly expanding Universe, we also "expand" the initial density ρ_i using the background density $\bar\rho_i$ and the initial density contrast δ_i; the expansion stops at first order by definition, and we have

$$\rho_i(\mathbf{X}) = \bar\rho_i + \varepsilon\bar\rho_i\delta_i(\mathbf{X}). \tag{3.9}$$

Evaluating the derivatives and determinants of the Lagrange–Newton system with the ansatz chosen in Equation (3.8), we find that the curl equations (3.6) are automatically satisfied to order $\mathcal{O}(\varepsilon^2)$. The divergence equation (3.7) leads to one differential equation for each order in the perturbation expansion[3], namely

$$3\ddot{a}a^2 - \Lambda a^3 = -4\pi G\bar\rho_i$$
$$\left(\ddot{q}_1 a^3 + 2\dot{q}_1\dot{a}a^2 + 3q_1\ddot{a}a^2 - \Lambda q_1 a^3\right)\Psi^{(1)}_{,ii} = -4\pi G\bar\rho_i\delta_i$$
$$\left(\ddot{q}_2 a^3 + 2\dot{q}_2\dot{a}a^2 + 3q_2\ddot{a}a^2 - \Lambda q_2 a^3\right)\Psi^{(2)}_{,ii}$$
$$+ \left(2\ddot{q}_1 q_1 a^3 + 4\dot{q}_1 q_1\dot{a}a^2 + 3q_1^2\ddot{a}a^2 - \Lambda q_1^2 a^3\right)\frac{1}{2}\left(\Psi^{(1)}_{,ii}\Psi^{(1)}_{,jj} - \Psi^{(1)}_{,ij}\Psi^{(1)}_{,ji}\right) = 0. \tag{3.10}$$

Obviously, these equations separate into time– and position–dependent factors. Hence we obtain three ordinary differential equations in time for the scale factors a, q_1 and q_2, and two Poisson equations for the potentials $\Psi^{(1)}$ and $\Psi^{(2)}$. Let us start with the Poisson equations, since they can easily be solved by iteration starting from the initial density contrast δ_i. The two equations are

$$\Delta\Psi^{(1)} = -\delta_i \tag{3.11}$$

and

$$\Delta\Psi^{(2)} = \frac{1}{2}\left(\Psi^{(1)}_{,ii}\Psi^{(1)}_{,jj} - \Psi^{(1)}_{,ij}\Psi^{(1)}_{,ji}\right). \tag{3.12}$$

The equations for the scale factors deserve some more attention.

The zeroth order equation is the well–known Friedmann equation

$$3\ddot{a}a^2 - \Lambda a^3 = -4\pi G\bar\rho_i. \tag{3.13}$$

---

[2]General solutions to second order cover pages; see (Buchert & Ehlers 1993) for the general irrotational part.
[3]Summation over paired indices is understood.

We now introduce the function $y = a(t)/a_0$ and define three cosmological parameters, namely the Hubble constant H_0, the matter density parameter Ω_{m}, and the vacuum density parameter Ω_Λ in terms of the parameters used so far:

$$H_0 = \frac{\dot{a}_0}{a_0}, \qquad \Omega_{\mathrm{m}} = \frac{8\pi G \bar{\rho}_0}{3 H_0^2}, \qquad \Omega_\Lambda = \frac{\Lambda}{3 H_0^2}, \tag{3.14}$$

where the current epoch is denoted by t_0, and initial and final background density are related via $\bar{\rho}_{\mathrm{i}} = \bar{\rho}_0 a_0^3$. This casts the Friedmann equation into the form

$$\frac{\ddot{y}}{H_0^2} = \Omega_\Lambda y - \frac{\Omega_{\mathrm{m}}}{2 y^2}. \tag{3.15}$$

Multiplication by $2\dot{y}$ yields a total time derivative; so we obtain the integral form of the Friedmann equation

$$\frac{\dot{y}^2}{H_0^2} = \Omega_\Lambda y^2 + \frac{\Omega_{\mathrm{m}}}{y} + (1 - \Omega_{\mathrm{m}} - \Omega_\Lambda), \tag{3.16}$$

where the constant of integration was expressed in terms of the parameters by using the boundary conditions $\dot{y}(t_0) = H_0$, $y(t_0) = 1$. In general relativity, this integration constant can be interpreted via $k = -a_0^2 H_0^2 (1 - \Omega_{\mathrm{m}} - \Omega_\Lambda)$ as the curvature parameter, where time is usually rescaled to restrict the possible values to $k \in \{0, \pm 1\}$. At this point, a major draw back of Newtonian cosmology shows up. Although the theory does predict this curvature parameter, it fails to provide a physical motivation for spacetime curvature (Goenner 1994).

Since we have already split off the expanding background solution in our perturbation ansatz, we are not interested in solving of the Friedmann equation anyway. Instead, we will express the time–dependent factors q_1 and q_2 as functions of y. Then $\dot{q} = q' \dot{y}$ and $\ddot{q} = q'' \dot{y}^2 + q' \ddot{y}$, where the prime $'$ denotes the derivative with respect to y. The time derivatives of y are obtained from the two versions (3.15) and (3.16) of the Friedmann equation.

In order to solve the second–order differential equations for q_1 and q_2, some boundary conditions need to be fixed. By definition in Equation (3.4), a particle's initial Eulerian position equals its Lagrangian position. Hence $q_1(a_{\mathrm{i}}) = q_2(a_{\mathrm{i}}) = 0$. Furthermore, we require that initially the peculiar velocity $\mathbf{u} = \dot{\mathbf{f}} - \frac{\dot{a}}{a}\mathbf{f}$ [4] and the peculiar acceleration $\mathbf{w} = \ddot{\mathbf{f}} - \frac{\ddot{a}}{a}\mathbf{f}$ are aligned (see Bildhauer & Buchert 1991 or Susperregi & Buchert 1997 for a thorough discussion of the motivation of this assumption). Also, we will eliminate the decaying mode solution from the first–order scale factor by extrapolating back in time and requiring $q_1(a = 0) + 1 = 0$; in fact this is the only choice which removes the singularity of $q_1(a)$ at $a = 0$. Note, however, that the full first order solution is dominated by the decaying mode when evolved backwards in time and develops a Kasner–type singularity instead of a Big Bang singularity (Zentsova & Chernin 1980).

The equation for the first time–dependent factor q_1 is

$$\ddot{q}_1 a^3 + 2 \dot{q}_1 \dot{a} a^2 + 3 q_1 \ddot{a} a^2 - \Lambda q_1 a^3 = 4 \pi G \bar{\rho}_{\mathrm{i}}, \tag{3.17}$$

---

[4]The difference between relative velocity $\dot{\mathbf{f}} - \dot{a}\mathbf{X}$ and peculiar velocity $\dot{\mathbf{f}} - \frac{\dot{a}}{a}\mathbf{f}$ is discussed by Buchert (1992). Luckily, both are equal at initial time. Analogously, peculiar acceleration is $\ddot{\mathbf{f}} - \frac{\ddot{a}}{a}\mathbf{f}$ rather than $\ddot{\mathbf{f}} - \ddot{a}\mathbf{X}$.

and with the cosmological parameters introduced in Equation (3.14) and the Friedmann equation (3.15) for $\ddot{y}$ it becomes formally equal to the well–known evolution equation for the density contrast in linear theory (Peebles 1993)

$$\ddot{q}_1 y^3 + 2\dot{q}_1 \dot{y} y^2 - \frac{3}{2}\Omega_{\mathrm{m}} H_0^2 (q_1 + 1) = 0. \tag{3.18}$$

Changing parameters from t to y as described above and inserting the Friedmann equation (3.16) for $\dot{y}^2$ we obtain

$$\left(\Omega_\Lambda y^5 + \Omega_{\mathrm{m}} y^2 + (1 - \Omega_\Lambda - \Omega_{\mathrm{m}})y^3\right) q_1''$$
$$+ \left(3\Omega_\Lambda y^4 + \frac{3}{2}\Omega_{\mathrm{m}} y + 2(1 - \Omega_\Lambda - \Omega_{\mathrm{m}})y^2\right) q_1' + \left(-\frac{3}{2}\Omega_{\mathrm{m}}\right)(q_1 + 1) = 0. \tag{3.19}$$

Let us postpone the discussion of the general solution to Section 3.2.2, and just consider the special case $\Omega_\Lambda = 0$, $\Omega_{\mathrm{m}} = 1$ where we obtain

$$q_1(a) = -1 + C_1 a + C_2 a^{-3/2}. \tag{3.20}$$

We need two boundary conditions to fix the constants. With $q_1|_{\mathrm{i}} = 0$ from the initial coincidence of Eulerian and Lagrangian position, and $\frac{\partial q_1}{\partial a}\big|_{\mathrm{i}} = 1$ – as described above this corresponds to dropping the decaying mode solution – we obtain the unique solution

$$q_1(a) = a - 1. \tag{3.21}$$

The equation for the second scale factor q_2 is almost equal to Equation (3.19) for the first scale factor, apart from a non–constant right hand side. We have

$$\ddot{q}_2 a^3 + 2\dot{q}_2 \dot{a} a^2 + 3q_2 \ddot{a} a^2 - \Lambda q_2 a^3 = -2\ddot{q}_1 q_1 a^3 - 4\dot{q}_1 q_1 \dot{a} a^2 - 3q_1^2 \ddot{a} a^2 + \Lambda q_1^2 a^3. \tag{3.22}$$

Applying the same procedure as before leads to the equation

$$\left(\Omega_\Lambda y^5 + \Omega_{\mathrm{m}} y^2 + (1 - \Omega_\Lambda - \Omega_{\mathrm{m}})y^3\right) q_2''$$
$$+ \left(3\Omega_\Lambda y^4 + \frac{3}{2}\Omega_{\mathrm{m}} y + 2(1 - \Omega_\Lambda - \Omega_{\mathrm{m}})y^2\right) q_2' + \left(-\frac{3}{2}\Omega_{\mathrm{m}}\right)(q_2 + 1) = -\frac{3}{2}\Omega_{\mathrm{m}}(q_1 + 1)^2. \tag{3.23}$$

Together with two boundary conditions, the solution for this scale factor becomes unique. As described above, we choose

$$q_2|_{\mathrm{i}} = 0, \qquad \frac{\partial q_2}{\partial a}\bigg|_{\mathrm{i}} = 0 \tag{3.24}$$

to satisfy the initial condition $\mathbf{X} = \mathbf{f}(\mathbf{X}, t_{\mathrm{i}})$ for the mapping, and to reproduce the relation $\mathbf{u} \propto \mathbf{w}$ between peculiar velocity and peculiar acceleration, respectively.

Because of these boundary conditions, the solution contains a decaying part. For a critical Universe with $\Omega_{\mathrm{m}} = 1$ and $\Omega_\Lambda = 0$ it is explicitly given by

$$q_2(a) = -\tfrac{3}{7}a^2 + \tfrac{6}{5}a - 1 + \tfrac{8}{35}a^{-3/2}, \tag{3.25}$$

while for arbitrary parameter combinations, solutions can only be obtained numerically.

3.2.2 Solutions of the scale factor equation

Let us take a closer look at the solution of Equation (3.19) for the first–order scale factor q_1. Although the general solution was already derived by Heath (1977) and is also given in Bildhauer et al. (1992), the issue of choosing appropriate boundary conditions suitable for our purposes has not been addressed. Since the solutions also appear as the general solution of the homogeneous equation corresponding to Equation (3.23) for the second–order scale factor, they are well worth some detailed study.

Let us begin with stating that Equation (3.19) divided by y^2 is exact and can be integrated to yield the first–order equation

$$\left(\Omega_\Lambda y^3 + \Omega_{\mathrm{m}} + (1 - \Omega_\Lambda - \Omega_{\mathrm{m}})y\right) q_1' + \left(\frac{3}{2y}\Omega_{\mathrm{m}} + (1 - \Omega_\Lambda - \Omega_{\mathrm{m}})\right) q_1 = C, \qquad (3.26)$$

where C is a constant of integration to be determined. As explained in the main section, we now extrapolate into the epoch before t_{i} and define the growing mode solution as the one that has

$$q_1(a = 0) + 1 = 0, \qquad (3.27)$$

whence brute–force integration leads to a solution containing only one integral

$$q_1(a) + 1 = a^{-3/2}\left(\frac{\Omega_\Lambda a^3 + \Omega_{\mathrm{m}}a_0^3 + (1 - \Omega_\Lambda - \Omega_{\mathrm{m}})aa_0^2}{\Omega_\Lambda + \Omega_{\mathrm{m}}a_0^3 + (1 - \Omega_\Lambda - \Omega_{\mathrm{m}})a_0^2}\right)^{1/2}$$

$$\times C' \int\limits_0^a \mathrm{d}\xi \left(\frac{\xi a_0^3}{\Omega_\Lambda \xi^3 + \Omega_{\mathrm{m}}a_0^3 + (1 - \Omega_\Lambda - \Omega_{\mathrm{m}})\xi a_0^2}\right)^{3/2}, \quad (3.28)$$

with some constants absorbed into the modified constant of integration C'. Using the boundary condition $q_1(a_{\mathrm{i}}) = 0$ now fixes the constant C' at

$$C' = \left[\int\limits_0^1 \mathrm{d}\xi \left(\frac{\xi a_0^3}{\Omega_\Lambda \xi^3 + \Omega_{\mathrm{m}}a_0^3 + (1 - \Omega_\Lambda - \Omega_{\mathrm{m}})\xi a_0^2}\right)^{3/2}\right]^{-1}. \qquad (3.29)$$

Note that for the Einstein–de Sitter case $\Omega_{\mathrm{m}} = 1$ and $\Omega_\Lambda = 0$ this solution reduces to the well–known growing mode solution $q_1 + 1 = a$ and completely discards the decaying mode solution $q_1 + 1 = a^{-3/2}$. The general second–order solution for $\Omega_\Lambda = 0$ was obtained by Catelan (1995) and independently by Vanselow (1995). At least parts of the solution can also be found in (Bouchet et al. 1995).

For the general Friedmann–Robertson–Walker background Universe with arbitrary Ω_{m} and Ω_Λ, we can expand the initial slope of the scale factor as a series in the small parameter a_0^{-1}. This is

of the order of $\approx 10^{-3}$, so the series yields accurate initial conditions for the numerical solution of the differential equation (3.19). To fourth order in a_0^{-1} we have

$$
\left. \frac{\partial q_1}{\partial a} \right|_{\mathrm{i}} \approx 1 - \tfrac{4}{7} \frac{1 - \Omega_\Lambda - \Omega_\mathrm{m}}{\Omega_\mathrm{m}} a_0^{-1} + \tfrac{64}{147} \left(\frac{1 - \Omega_\Lambda - \Omega_\mathrm{m}}{\Omega_\mathrm{m}} \right)^2 a_0^{-2}
$$
$$
- \left(\tfrac{1376}{3773} \left(\frac{1 - \Omega_\Lambda - \Omega_\mathrm{m}}{\Omega_\mathrm{m}} \right)^3 + \tfrac{6}{11} \frac{\Omega_\Lambda}{\Omega_\mathrm{m}} \right) a_0^{-3} + \mathcal{O}\left(a_0^{-4} \right). \quad (3.30)
$$

The result of series expansion of the general solution from Equation (3.28) is

$$
q_1(a) + 1 = a - \tfrac{4}{7} \frac{1 - \Omega_\Lambda - \Omega_\mathrm{m}}{\Omega_\mathrm{m}} a(a-1) a_0^{-1}
$$
$$
+ \tfrac{8}{147} \left(\frac{1 - \Omega_\Lambda - \Omega_\mathrm{m}}{\Omega_\mathrm{m}} \right)^2 a(a-1)(7a+1) a_0^{-2} + \mathcal{O}\left(\left(\frac{a}{a_0} \right)^3 \right). \quad (3.31)
$$

Note that Ω_Λ itself only enters the third–order corrections, which we haven't written down here, so low–density Universes closed by a cosmological constant with $\Omega_\Lambda + \Omega_\mathrm{m} = 1$ are close to the Einstein–de Sitter case $q_1 + 1 = a$.

3.3 Practical issues

3.3.1 Initial conditions

The parameters Ω_Λ, Ω_m, and H_0 enter into the time–dependent scale factors q_1 and q_2. For the initial density field δ_i one normally uses a realisation of a Gaussian random field with a power spectrum determined by the type of dark matter and the background cosmology. The crucial part is of course the evaluation of the initial power spectrum.

Compared to the primordial power spectrum, usually modelled as the scale–invariant form suggested by Harrison (1970) and Zel'dovich (1972), the initial power spectrum is processed by a transfer function which contains the physics of the epoch of recombination. Analytical approximation formulae for a wide range of cosmological models are summarised in the classical articles by Bardeen et al. (1987) and Bardeen et al. (1986); more recent tables that include Cold Dark Matter with a non–vanishing baryon contribution are provided by Holtzman (1989).

Together, the primordial power spectrum and the transfer function determine the shape of the initial power spectrum. Its normalisation is not fixed by the cosmology, but must be tied to measurements. At the infrared end of the spectrum, COBE measurements of the CMB quadrupole $Q_{\mathrm{rms-PS}}$ provide the so–called COBE–normalisation described by Bunn et al. (1995), White & Bunn (1995) (note White & Bunn 1996), and Bunn & White (1997). A different possibility uses the parameter σ_8, which gives the density fluctuations on a scale of

Figure 3.1 The initial power spectra for Standard Cold Dark Matter (SCDM), together with the dynamical range of the boxes shown in our examples below.
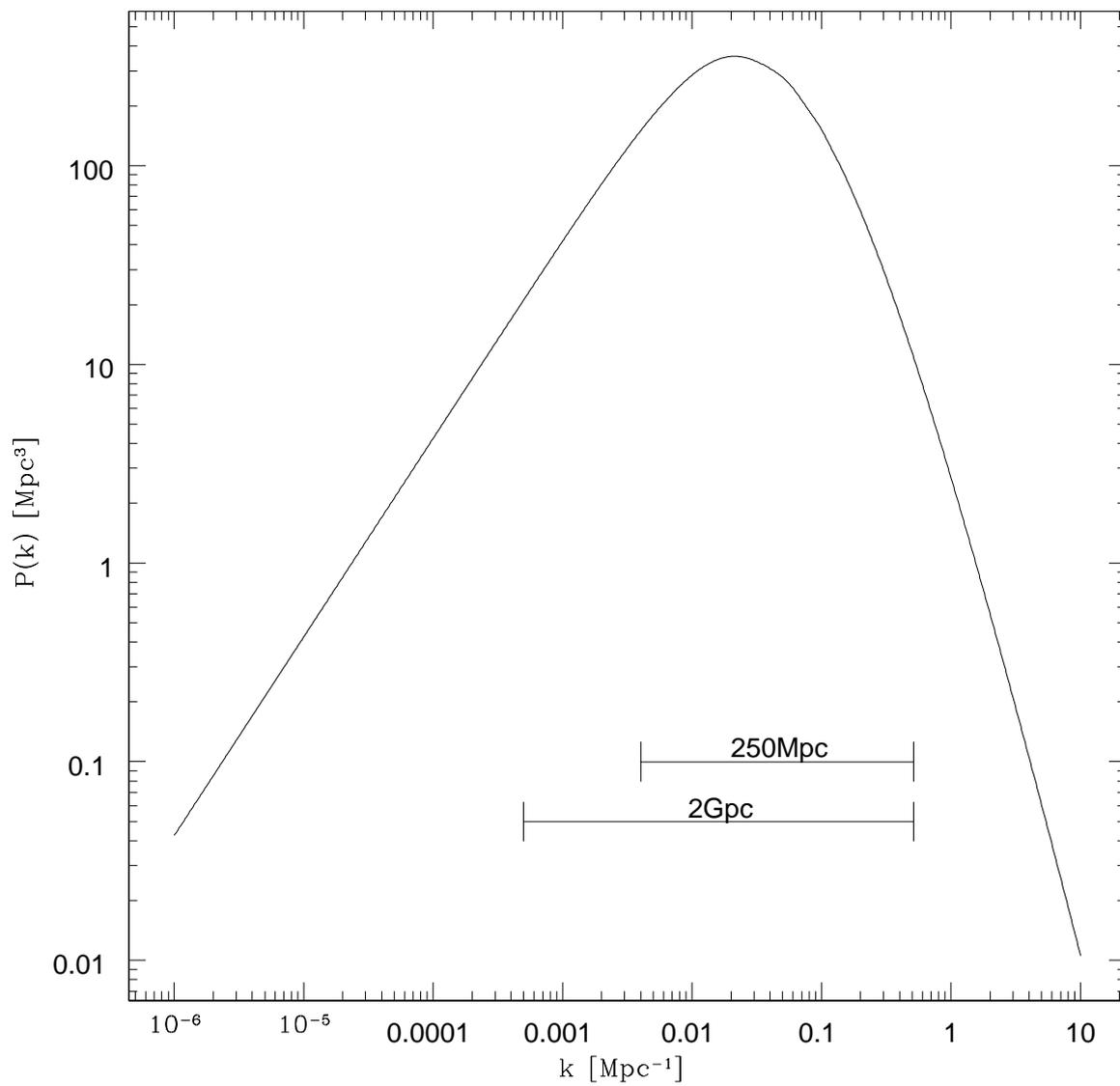
8Mpc and is measured through cluster abundances; hence the term "cluster normalisation". Borgani et al. (1994) have performed cluster normalised simulations of Abell/ACO clusters using the truncated Zel'dovich approximation.

From the practical point of view, the programs linger and graphics of the software package cosmics (Bertschinger 1995; Bode & Bertschinger 1995) are widely used for generating initial conditions for numerical simulations. The package implements linearised relativistic calculations reviewed for example by Bardeen (1980) or Mukhanov et al. (1992).

3.3.2   Optimization

When applied naïvely, Lagrangian perturbation schemes fail fairly miserably at reproducing the coherence of evolved large–scale structure. Coles et al. (1993a) have succeeded in enhancing the performance of the Zel'dovich approximation considerably by truncating the initial power spectrum at the scale of non–linearity. A detailed comparison to the related adhesion approximation was performed by Melott et al. (1994). We adopt a related approach by Weiß et al. (1996), that uses Gaussian smoothing of the initial field instead of a sharp Fourier space cutoff. Good values for the truncation length are given by multiples of the scale of non–linearity, where the factors for first and second order were determined with brute–force evaluation by Weiß et al. (1996) and Melott et al. (1995) in the Einstein–de Sitter case. Hamana (1998) confirms that the result is also valid in arbitrary background cosmologies. Other optimization schemes, using analytical calculations of the evolved correlation function, are discussed by Schneider & Bartelmann (1995).

Note that small scale power in the initial conditions is removed when truncating the power spectrum, so the mapping $\mathbf{f}$ becomes a smooth analytical function of Lagrangian position $\mathbf{X}$. This means that we can improve resolution by linearly interpolating in Lagrangian space. Tests, as shown in Figure 3.2, prove that accurate results are still obtainable when an interpolation factor of as much as ten is used.

3.3.3   Biasing

For the purpose of constructing large mock samples of galaxies there are two arguments why the erasure of small–scale features present in N–body simulations looses relevance. From a pragmatic point of view, the k–space cutoff is of the order of the Nyquist frequency of any presently conceivable N–body simulation. Also, small–scale filamentary structure hosts mainly low surface brightness galaxies which will be discarded when constructing magnitude–limited mock samples.

A distribution of galaxies may be obtained through Lagrangian biasing, as suggested by Buchert (1990) and implemented by Weiß & Buchert (1993) in the simulation of high–resolution pencil beam surveys, and by Blanchard et al. (1993) for the quasar density evolution. Also, Kates et al. (1991) implement a slightly different semianalytical scheme together with the truncated

Zel'dovich approximation. Alternatively, one can think of using ordinary Eulerian schemes (Dekel 1987; Rees 1987; Dekel & Rees 1987).

Obviously, catalogues of galaxies are observed in redshift space rather than real space. This issue is addressed by Hivon et al. (1995) with emphasis on Lagrangian perturbation schemes. Furthermore, on very large scales one should carefully take into account the light–cone effect in practice and theory (Colberg 1999; Yamamoto & Suto 1999).

3.4   Examples and Outlook

Figure 3.2 shows some slices of a 250Mpc or $125h^{-1}\mathrm{Mpc}$ SCDM box to demonstrate the power of Lagrangian interpolation. The example is of course artificial in the sense that the resolution is far too high given the physical limits of the perturbation schemes and of $\mathbf{f}$ itself (Buchert & Bartelmann 1991). However, the main point we would like to make is not the accurate reproduction of the full gravitational dynamics on small scales, but rather the emergence of finely resolved structures in the interpolated field, despite the relative coarseness of the original set of particles.

A more realistic situation is shown in Figure 3.3. This time, we reach a resolution of $1h^{-1}\mathrm{Mpc}$ by interpolating the original 128^3 particles in a $1h^{-1}\mathrm{Gpc}$ box to 1024^3. This resolution matches the method's limits. On the other hand, the box size is already comparable to the size of the observable Universe. Note that the whole run including interpolation and density reconstruction took less than half an hour of CPU–time on an ordinary work station, while running N–body simulations of comparable size (Colberg et al. 1998) consumes days on massive supercomputers.

The classical Zel'dovich approximation and its various refinements are often called the "poor man's N–body simulation". Weiß et al. (1996) showed that the method presented here becomes reasonably accurate on scales of $\approx 10h^{-1}\mathrm{Mpc}$, and Hamana (1998) has generalised their results for arbitrary background cosmologies. Generally speaking, the approximation scheme presented here can generate large numbers of fairly realistic distributions of dark matter and galaxies with minimal computational effort, whereas N–body simulations usually end up with a single realisation of the cosmological model under consideration.

More precisely, we are planning to use the method for systematic studies of cosmic variance in the near future. Another potentially interesting application is the evaluation of estimators for statistics of large–scale structure. Last but not least, the numerical implementation should be complemented with further analytical investigations, especially of post–Newtonian effects which should become relevant on the scales we reach.

Figure 3.2 Pretty pictures. To be more precise, this is a flat, critical Universe with Standard Cold Dark Matter spectrum in a 250Mpc or $125h^{-1}\mathrm{Mpc}$ box. Each panel shows a slice of the same initial conditions propagated to the present epoch for 128^3 particles. In the left column no interpolation is used, while right column interpolates by a factor of eight to 1024^3 particles. The density reconstruction was done by cloud–in–cell binning with 2^3 particles per cell on average. Both rows show the same model starting from the same initial conditions, but the top row gives the mere Zel'dovich approximation while for the bottom row second–order terms are also added. Note the richer fine structure of the caustics and multiple shell crossings. This Figure also shows that structure formation in the first–order model appears time–delayed compared to second order. This time–delay can be partly compensated by an optimized time shift investigated by Karakatsanis et al. (1997).
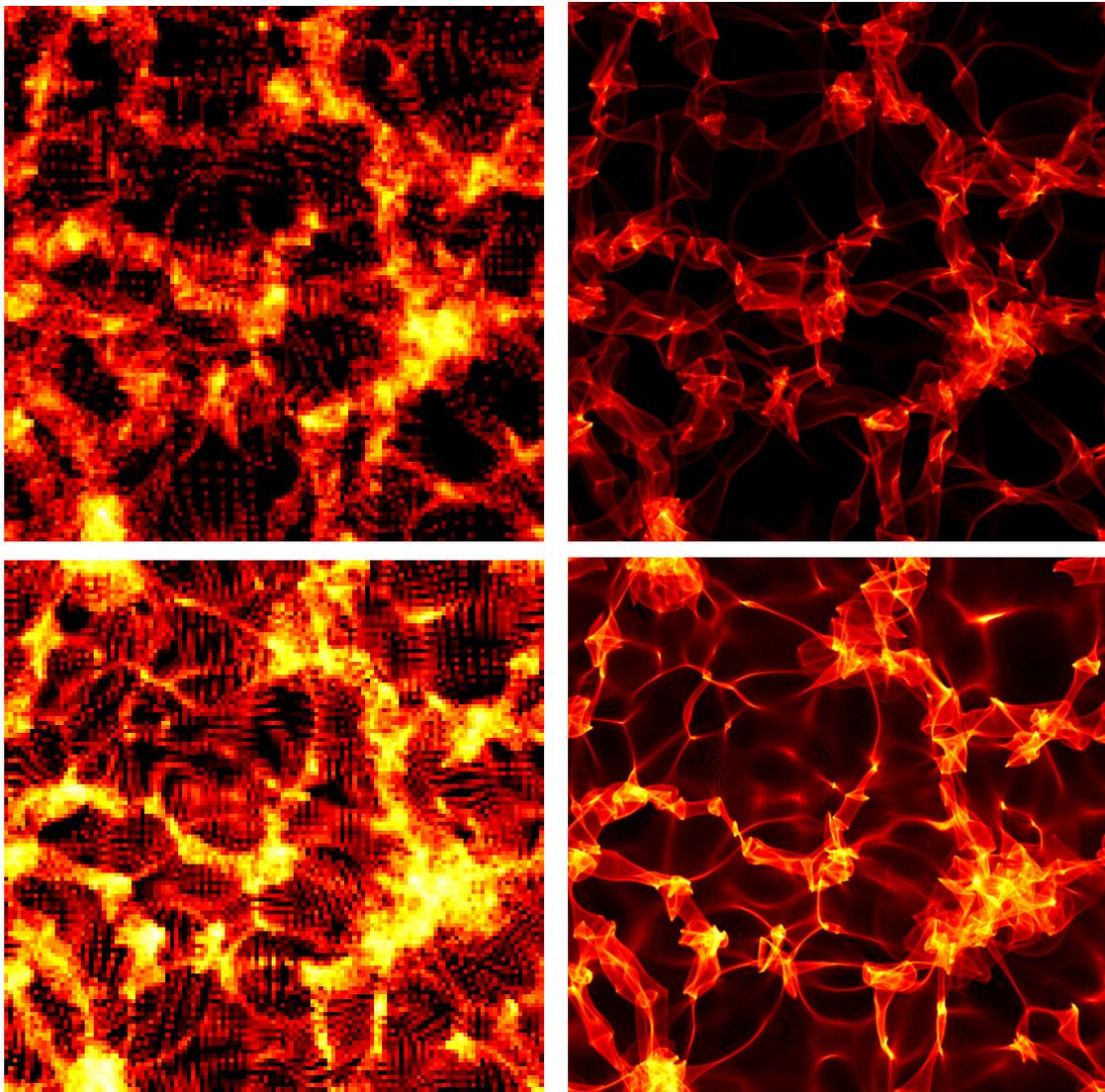
Figure 3.3 Another SCDM spectrum, this time in a box of 2Gpc or $1h^{-1}\mathrm{Gpc}$ sidelength. The top panel was obtained by propagating 128^3 particles from the initial conditions to today using optimized second order Lagrangian perturbation theory. After that, interpolation by a factor of eight increased resolution to 1024^3, as shown in the bottom panel.
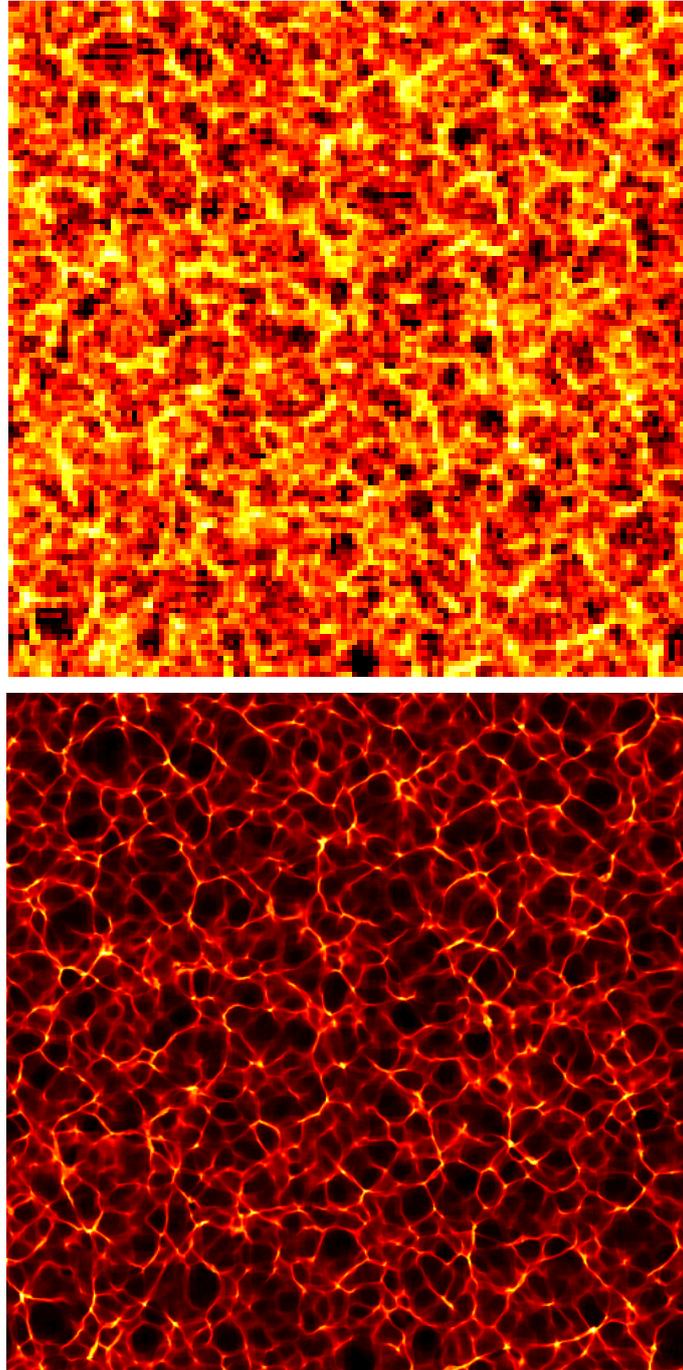
4 Anisotropies of the Cosmic Microwave Background

> *Don't stand there gawping like you've never seen the face of God before!*
> *(Gilliam & Jones 1983)*

The oldest signal accessible to mankind is the Cosmic Microwave Background discovered by Penzias & Wilson (1965). Consisting of photons that have been freely streaming since the Universe was only 300,000 years old, the Cosmic Microwave Background (CMB) provides valuable information on the early history of our Universe. Via the Sachs–Wolfe effect (Sachs & Wolfe 1967), its anisotropies mirror the matter fluctuations at the epoch of recombination (at redshift $z \approx 1,100$) and with them the seeds of the large–scale structure seen today.

Various methods of statistical analysis have been used on Cosmic Microwave Background maps. Among them are the two– and three–point correlation function (Hinshaw et al. 1996; Hinshaw et al. 1995), the power spectrum (Górski et al. 1996), skewness and kurtosis (Luo & Schramm 1993), multifractals (Pompilio et al. 1995), wavelet analysis (Hobson et al. 1998), and the extrema correlation function (Kogut et al. 1995). Another promising approach is the investigation of the morphology of hot and cold spots. Being complementary to the traditional approach via the hierarchy of correlation functions, it provides alternative methods for determining cosmological parameters (Torres et al. 1995). But above all, morphological statistics incorporate correlation functions of arbitrary order. Hence they are sensitive to signatures of non–Gaussianity in the temperature fluctuations, which would indicate the presence of topological defects such as strings or textures arising from phase transitions in the early Universe (Kaiser & Stebbins 1984; Stebbins 1988). Experiments to obtain high–resolution maps of large regions of the sky are still under development (Bersanelli et al. 1996; Bennett et al. 1995), and are widely expected to put an end to the ongoing debate (Albrecht et al. 1997; Spergel & Pen 1997).

In order to measure the morphology of CMB anisotropies, the Euler characteristic, or equivalently the genus, was suggested more than a decade ago (Coles & Barrow 1987; Coles 1988). Even to date, most applications are confined to genus statistics, although an early theoretical study by Gott III et al. (1990) also considers the boundary length, but failed to come up with a subsequent analysis of data. The analysis of the first–year COBE DMR data using genus statistics was done by Smoot et al. (1994); their work also contains a thorough discussion

of the performance of the method compared to other measures of non–Gaussianity. Further applications of topological methods on CMB anisotropies come from Torres (1994) and Torres et al. (1995), and genus calculations of the four–year COBE DMR data are due to Colley et al. (1996) and Kogut et al. (1996).

The genus can be placed in the wider framework of the Minkowski functionals (Minkowski 1903). Both the original Boolean grain model applicable to the analysis of point sets (Mecke et al. 1994), and the approach using excursion sets and isodensity contours of smoothed random fields by Schmalzing & Buchert (1997) will be explored in the following chapters. Here, we will apply Minkowski functionals to all–sky CMB maps. The application of all Minkowski functionals to CMB maps was suggested by Winitzki & Kosowsky (1997) and performed on the COBE DMR data by Schmalzing & Górski (1998). Subsequent works (Novikov et al. 1999; Naselsky & Novikov 1998; Hobson et al. 1998) have further explored the capabilities of this new method for various recent and upcoming CMB measurements.

As explained in Section 2.3.2, on a curved space the underlying concepts of integral geometry remain the same, and the relevant formulae are easily generalized (Santaló 1976). However, some formal obstacles will be encountered in the practical application to data. First of all, the temperature is not sampled continuously, but at discrete pixels (Section 4.1). Care must be taken to remove the effects of usually incomplete sky coverage, while retaining as much information as possible at the same time. Finally, the interesting cosmic signal is usually obscured by instrument noise. The influence of those issues in practice is tested in Section 4.2. That settled, Section 4.3 presents some examples, most notably an attempt to assess the presence or absence of non–Gaussian signatures in the COBE DMR maps.

4.1   Estimating Minkowski functionals from pixelised CMB sky maps

Throughout this section, we will illustrate the application of our method on a particular random field. In order to stick to a simple, analytically tractable model, we generate a Gaussian random field. Its angular power spectrum C_ℓ is chosen to reproduce the salient features of the DMR sky maps. Hence, we start from angular components given by the formula

$$C_{\ell,\mathrm{powerlaw}} = C_2 \frac{\Gamma\left(\ell + \frac{n-1}{2}\right)\Gamma\left(\frac{9-n}{2}\right)}{\Gamma\left(\ell + \frac{5-n}{2}\right)\Gamma\left(\frac{n+3}{2}\right)}, \tag{4.1}$$

derived from a power–law spectrum $P(k) \propto k^n$ by Bond & Efstathiou (1987), and smooth them with a Gaussian filter of $7°$ full width at half maximum (FWHM) to model the DMR beam. White noise with a fixed r.m.s. fluctuation level of $3\mu\mathrm{K}$ is then added to this "cosmic signal"; this is in practice done on the pixels in real space, but for comparison we may also evaluate the contribution to the angular power spectrum, which is

$$C_{\ell,\mathrm{noise}} = (3\mu\mathrm{K})^2 \tag{4.2}$$

Figure 4.1 The angular power spectrum shown in this plot is used for all tests of the estimators presented in this section. It was chosen as a rough model for the DMR maps while remaining analytically tractable. The complete power spectrum, indicated by a solid line, is the sum of two contributions, namely a Harrison–Zel'dovich spectrum normalized to $C_{10} = (7\mu\mathrm{K})^2$, convolved with a $7°$ FWHM beam to mimic a cosmic "signal" (dashed line), and random pixel noise (dotted line).
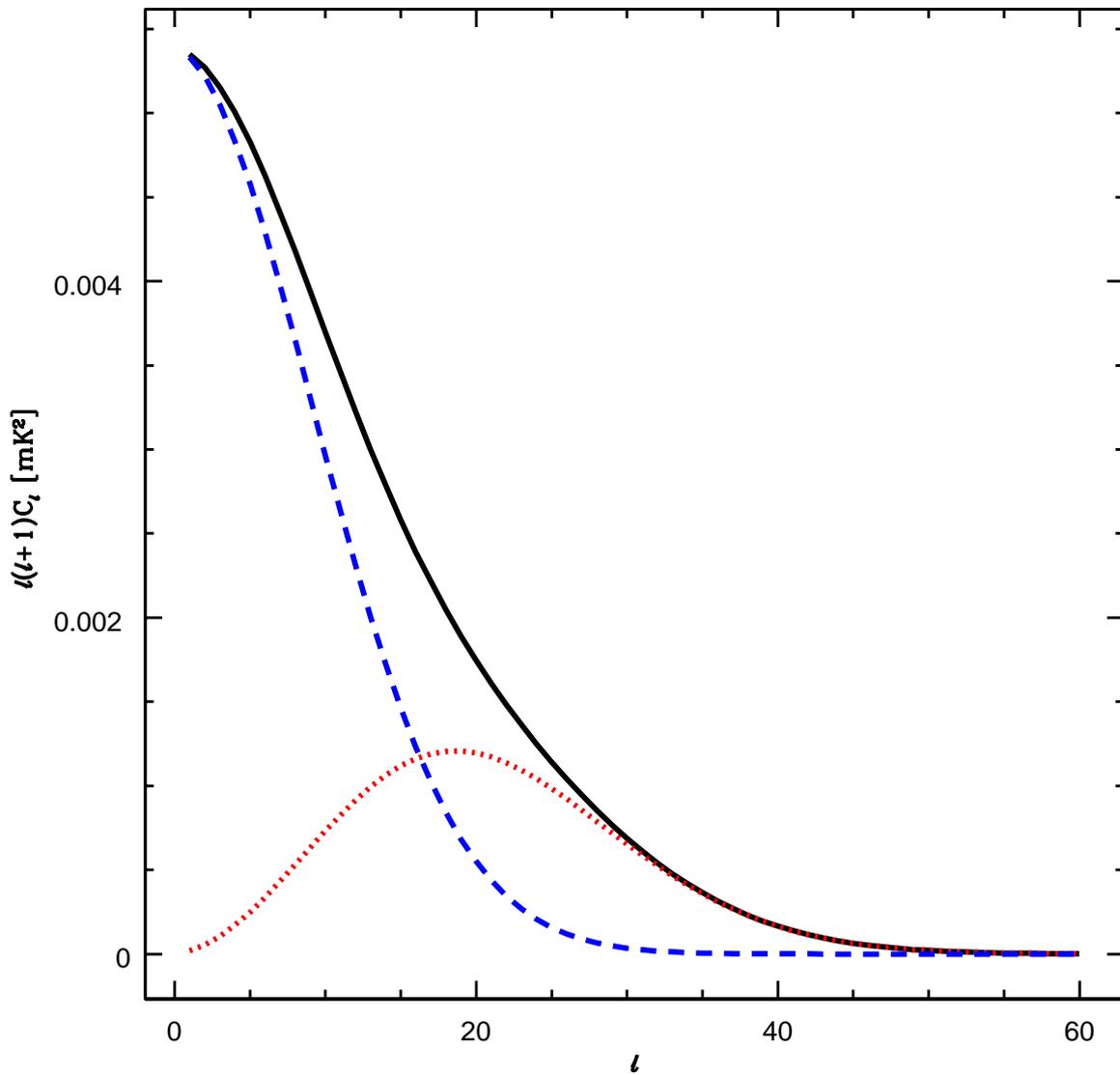
independent of ℓ. Finally, a Gaussian smoothing kernel of variance ω^2, given by

$$g_{\ell,\mathrm{Gauss}} = \exp\left(-\tfrac{1}{2}\omega^2\ell(\ell+1)\right) \tag{4.3}$$

is applied to reduce the noise level, and to obtain a regular field. Note that the normalization factor C_2 in Equation (4.1) is directly related to the CMB quadrupole. As pointed out by Górski et al. (1994) its particular value determined from the COBE DMR sky maps is highly dependent on the spectral index n; therefore it has become common usage to quote the multipole $C_{10} \approx (7\mu\mathrm{K})^2$ as a sufficiently spectrum independent normalization.

So the various contributions to the angular power spectrum C_ℓ for our example sum up to

$$C_\ell = g_{\ell,\mathrm{Gauss}}^2 \left(C_{\ell,\mathrm{powerlaw}}g_{\ell,\mathrm{beam}}^2 + C_{\ell,\mathrm{noise}}\right). \tag{4.4}$$

In Figure 4.1, the contributions of "signal" and noise are shown separately, and combined to the full spectrum. In order to obtain a regular field and suppress noise the combined spectrum is smoothed with a Gaussian filter; its width for this particular example is $3°$. Although the noise contribution has not been completely removed, it is considerably reduced without affecting the signal too strongly. See Section 4.2.4 for a broader discussion of this issue.

In order to estimate the Minkowski functionals from the discrete pixels of a map we attempt to follow the prescription outlined by Schmalzing & Buchert (1997) for cubic grids in three–dimensional Euclidean space (see also Chapter 2). Note that the first approach based on Crofton's formula and counting of elementary cells is not viable since a strictly regular pixelisation of the sphere does not exist. However, their second approach is based on averaging over invariants, and Equation (2.30) can easily be adapted to the sphere.

4.1.1 Covariant derivatives

Using the well–known parametrisation of the unit sphere through azimuth angle ϑ and polar angle φ we can express the covariant derivatives at a point $\mathbf{x} = (\vartheta, \varphi)$ in terms of the partial derivatives[1];

$$
\begin{aligned}
u_{;\vartheta} &= u_{,\vartheta}, \\
u_{;\varphi} &= \frac{1}{\sin\vartheta}u_{,\varphi}, \\
u_{;\vartheta\vartheta} &= u_{,\vartheta\vartheta}, \\
u_{;\vartheta\varphi} &= \frac{1}{\sin\vartheta}u_{,\vartheta\varphi} - \frac{\cos\vartheta}{\sin^2\vartheta}u_{,\varphi}, \\
u_{;\varphi\varphi} &= \frac{1}{\sin^2\vartheta}u_{,\varphi\varphi} + \frac{\cos\vartheta}{\sin\vartheta}u_{,\vartheta}.
\end{aligned}
\tag{4.5}
$$

---

[1]Note that we use indices following a semicolon, such as $u_{;i}$ to denote covariant differentiation of u with respect to the coordinate i, as opposed to partial derivatives where we write indices following a comma, e.g. $u_{,i}$.

The partial derivatives in turn are best calculated from the spherical harmonics expansion

$$u(\vartheta, \varphi) = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} a_{\ell m} Y_{\ell m}(\vartheta, \varphi). \tag{4.6}$$

This is simply done by replacing the harmonic function $Y_{\ell m}$ with its appropriate partial derivative. Since the functions $Y_{\ell m}$ depend on φ via sine and cosine functions only, the derivatives with respect to φ can be obtained analytically. Partial derivatives with respect to ϑ are calculated via recursion formulae constructed by differentiating the recursion for the associated Legendre functions P_ℓ^m, given for example by Abramowitz & Stegun (1970).

4.1.2 Integrals over invariants

On the two–dimensional sphere, the integrals summarized in Equation (2.30) become

$$v_j(\nu) = \frac{1}{4\pi R^2} \int_{\mathbb{S}^2} \mathrm{d}a \, \mathcal{I}_j, \tag{4.7}$$

with integrands $\mathcal{I}_j$ depending solely on the threshold ν, the field value u and its first– and second–order covariant derivatives. Explicitly,

$$\begin{aligned}
\mathcal{I}_0 &= \Theta(u - \nu), \\
\mathcal{I}_1 &= \frac{1}{4}\delta(u - \nu)\sqrt{u_{;\theta}^2 + u_{;\phi}^2}, \\
\mathcal{I}_2 &= \frac{1}{2\pi}\delta(u - \nu)\frac{2u_{;\theta}u_{;\phi}u_{;\theta\phi} - u_{;\theta}^2 u_{;\phi\phi} - u_{;\phi}^2 u_{;\theta\theta}}{u_{;\theta}^2 + u_{;\phi}^2}.
\end{aligned} \tag{4.8}$$

We still have to account for the finite number of sample points. If the random field is sampled at N pixels at locations $\mathbf{x}_i$ on the sphere, we only need to estimate the values $\mathcal{I}_j(\mathbf{x}_i)$ of the invariants from Equation (4.8) at each location. Furthermore, we need to replace the delta function with a bin of finite width Δ,

$$\delta(u - \nu) \approx \frac{1}{\Delta}\mathbf{1}_{[-\Delta/2, +\Delta/2]}(u - \nu), \tag{4.9}$$

where $\mathbf{1}_A$ is the indicator function of the set A, with $\mathbf{1}_A(x) = 1$ for $x \in A$, and $\mathbf{1}_A(x) = 0$ otherwise. The Minkowski functionals are then estimated by summation over all pixels[2]

$$v_j(\nu) \approx \frac{1}{N} \sum_{i=1}^{N} w_i \mathcal{I}_j(\mathbf{x}_i). \tag{4.10}$$

For incomplete sky coverage we must restrict the average to the unmasked pixels. This problem is addressed in detail in Section 4.2.

---

[2]We set the pixel weight factors w_i equal to 1, but this may be changed, if $\sum_{i=1}^{N} w_i = N$ is preserved.

4.1.3 Analytical expectation values

Putting $d = 2$ in Equation (2.35) we obtain the analytical expectation values for a Gaussian random field on the sphere at threshold ν.

$$
\begin{aligned}
v_0(\nu) &= \frac{1}{2} - \frac{1}{2}\Phi\left(\frac{\nu - \mu}{\sqrt{2\sigma}}\right), \\
v_1(\nu) &= \frac{\tau^{1/2}}{8\sigma^{1/2}}\exp\left(-\frac{(\nu - \mu)^2}{2\sigma}\right), \\
v_2(\nu) &= \frac{\tau}{2\pi^{3/2}\sigma}\frac{\nu - \mu}{\sqrt{2\sigma}}\exp\left(-\frac{(\nu - \mu)^2}{2\sigma}\right).
\end{aligned}
\tag{4.11}
$$

As stated before, the three parameters can be calculated from moments of the random field and its first derivatives, where care must be taken to use the covariant ones. We have

$$
\begin{aligned}
\mu &= \langle u \rangle \\
\sigma &= \langle u^2 \rangle - \mu^2 \\
\tau &= \frac{1}{2}\langle u_{;i}u_{;i} \rangle.
\end{aligned}
\tag{4.12}
$$

Alternatively, we may estimate the parameters directly from the power spectrum C_ℓ. Using the expansion from Equation (4.6) and the orthogonality properties of the spherical harmonics, we immediately obtain

$$
\begin{aligned}
\sigma &= \sum_{\ell=1}^{\infty}(2\ell + 1)C_\ell, \\
\tau &= \sum_{\ell=1}^{\infty}(2\ell + 1)C_\ell\frac{\ell(\ell + 1)}{2}.
\end{aligned}
\tag{4.13}
$$

4.2 Testing the estimators

4.2.1 Complete sky coverage

To begin with, let us look at the example without simulating the restrictions of incomplete sky coverage. Figure 4.2 shows the average Minkowski functionals of 1,000 realizations.

Looking at the general features of all curves, it can be seen that the area v_0 of hot spots decreases monotonically from the value of one at low threshold, when the whole sphere belongs to the excursion set, to a value of zero at high thresholds which are not passed by any of the pixels. The boundary length v_1 starts from a value of zero for a completely filled sphere. It reaches a maximum at intermediate thresholds, where the excursion set forms an interconnected pattern of patches and holes with a very long boundary. When the excursion set becomes emptier

Figure 4.2 Minkowski functionals for the example used throughout this Section. The areas indicate the averages and 1σ–fluctuations of an ensemble of 1,000 realizations computed on the whole DMR sky cube. The analytical expectation values from Equation (4.11) are almost exactly reproduced by the mean values (central lines). Fluctuations indicated by the shaded area are due to uncertainties in the parameter determination from a single map via Equation (4.12). They account for a large part of the overall fluctuations (empty area).
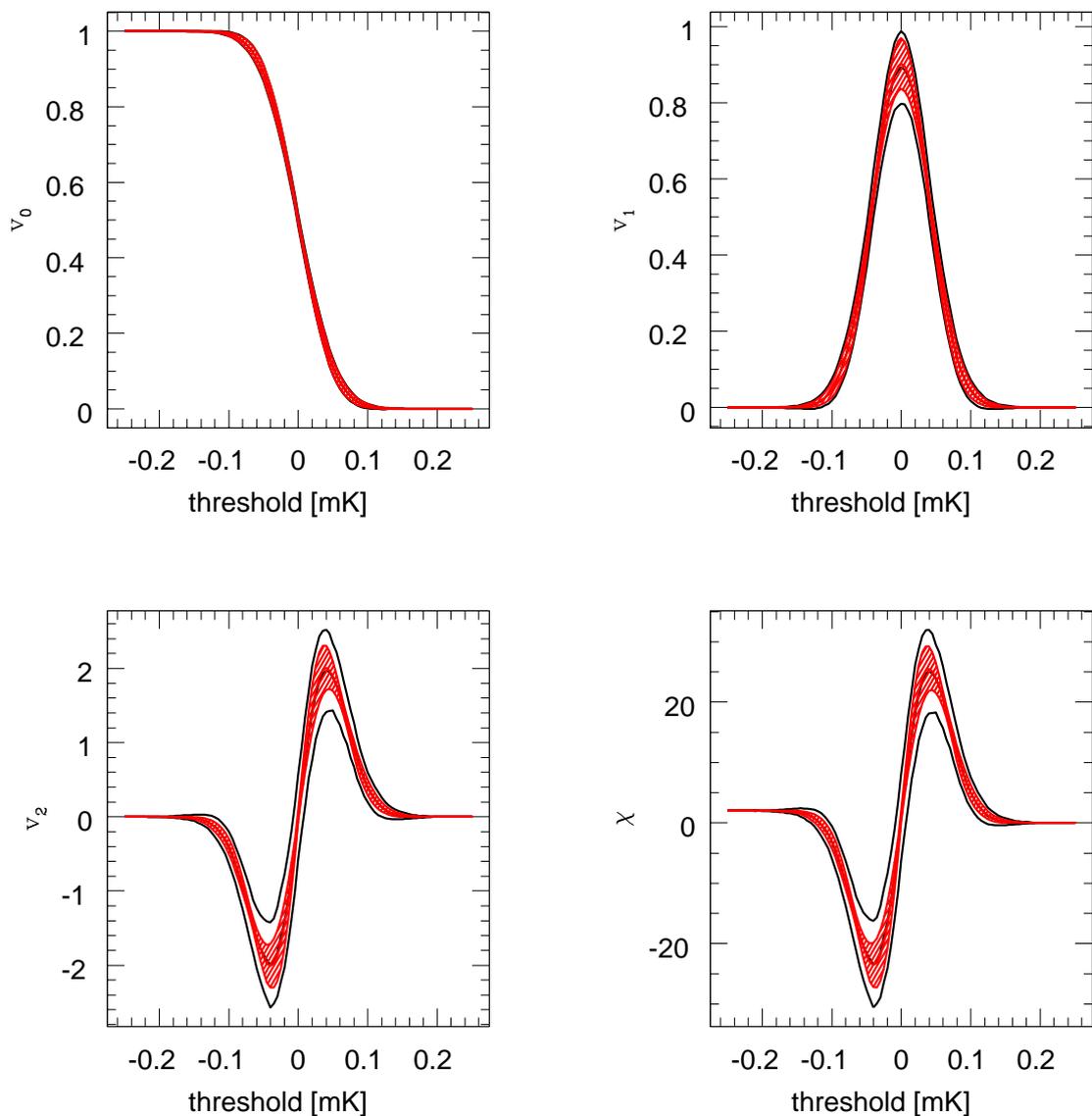
and emptier, the boundary length declines back to zero. For the random field shown in our example, the integrated geodesic curvature v_2 behaves largely similar to the Euler characteristic; the minor differences only become appreciable for fields with fewer features. Lastly, the Euler characteristic χ at low thresholds has a value of two for a closed sphere. With increasing threshold, the Euler characteristic declines to negative values as holes open in the excursion set and give a negative contribution. This downward trend gradually stops as individual hot spots emerge, so a minimum develops, and the Euler characteristic attains positive values. Finally, more and more hot spots fall below the growing threshold, so their number and hence the Euler characteristic decreases again, reaching a final value of zero.

A description of the individual curves can be found in the figure caption.

4.2.2   Uncertainties through incomplete sky coverage

In practice, a data set will suffer from incomplete sky coverage. In order to estimate the uncertainties introduced solely by the galactic cut, we first construct a single realization of the random field on the whole sky. The Minkowski functionals for this random field are calculated and roughly fit the analytical expectations, with fluctuations consistent with the areas shown in Figure 4.2. Then, we apply a series of straight galactic cuts with varying direction, but with constant width of $30°$; this value reduces the number of pixels to exactly half the original value. Figure 4.3 shows a comparison of the true values for one field and the fluctuations introduced by the sample variance of the rotating cuts. Note that the smaller number of pixels does increase the uncertainties, but the average is not affected – the estimator remains unbiased.

4.2.3   Boundary effects

The previous subsection dealt with a random field that was first realized on the whole sky, then smoothed with a Gaussian filter, and cut afterwards. In order to determine whether the galactic cut affects the estimators derived above, we use the COBE DMR pixels and the customized cut from the four–year data (Bennett et al. 1996). This time, we remove the pixels within the galactic cut before the smoothing kernel is applied.

It turns out that by using this procedure, which is actually the correct one for mimicking real data, the galactic cut severely affects the estimators, and leads to a systematic bias of as much as 1σ. Figure 4.4 shows the unbiased results from all–sky maps already displayed in Figure 4.2, in comparison the biased result obtained with the naïvely applied estimator.

A straightforward procedure to remove these biases from the estimators is to further restrict the number of pixels to the ones that lie "far away" from the cut. In order to find them, we consider the indicator function of the cut itself, smooth it with the Gaussian filter, and consider the values at pixels outside the cut as their level of "contamination". Now the sums from Equation (4.10) can be restricted to the pixels where the smoothed cut lies below a certain threshold. Figure 4.5 shows the results for an allowed level of 1%; in practice, even as much as

Figure 4.3 This plot shows the uncertainties due to a galactic cut. In order to separate them from fluctuations between different realizations in an ensemble (see Figure 4.2), we chose a single all–sky realization (central line) of the random field, and applied 1,000 straight cuts up to 30° latitude with randomly varying orientation of the equatorial plane. It turns out that the uncertainties caused by this deterioration of statistics are at least as strong as ensemble fluctuation. However, the correct mean value is still reproduced, so the estimator remains unbiased.
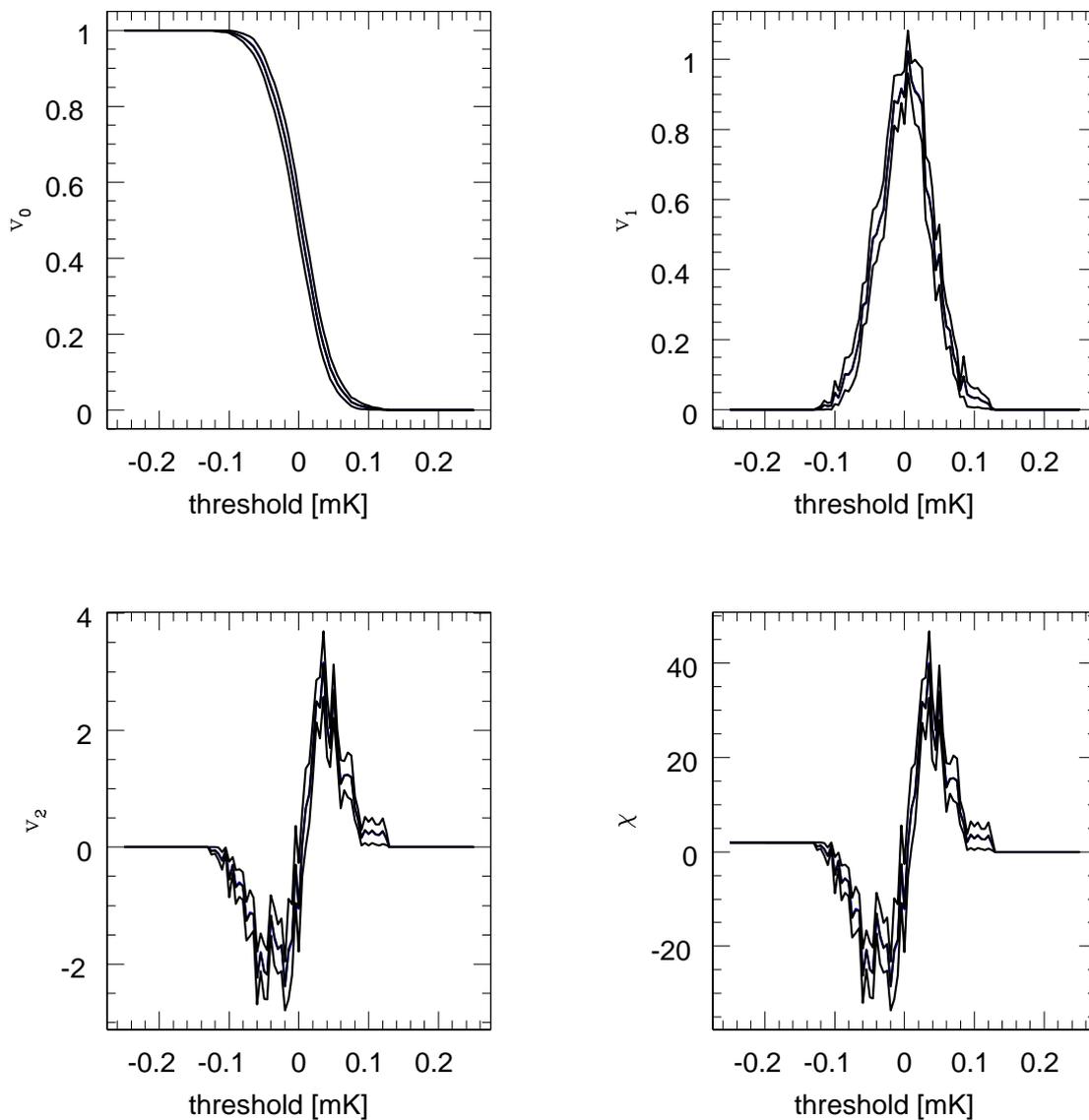
Figure 4.4 If the galactic cut is applied before smoothing the random field, as it should be done for real data, pixels in the vicinity of the galactic cut suffer from severe contamination. This leads to a visible bias in the estimated Minkowski functionals, particularly the integrated curvature v_2 and the Euler characteristic χ, although the area v_0 and the circumference v_1 are also affected. The shaded area shows average and fluctuations for the all–sky map already presented in Figure 4.2, while the empty area with thicker contour and central line indicates the same quantities for the maps biased through the cut.
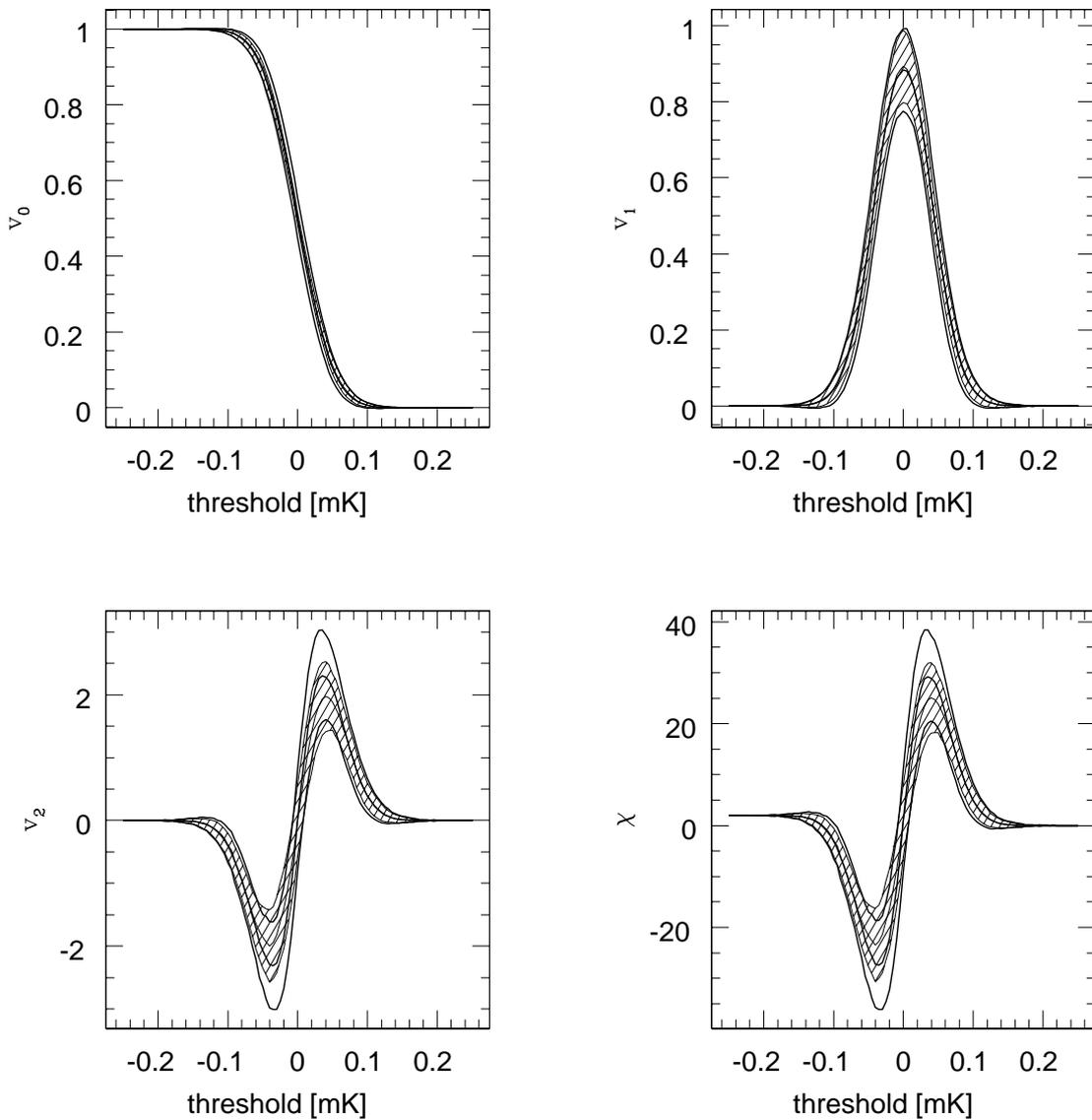
Figure 4.5 The biases demonstrated in Figure 4.4 can be removed by also smoothing the cut, and excluding points where the smoothed cut still reaches a certain level. In the example shown, this threshold was set to a very restrictive 1%; obviously the mean values agree completely. In practice, a level as high as 5% might still produce reliable results. Note that fluctuations have increased in comparison with the results obtained from all–sky maps, as already demonstrated in Figure 4.3.
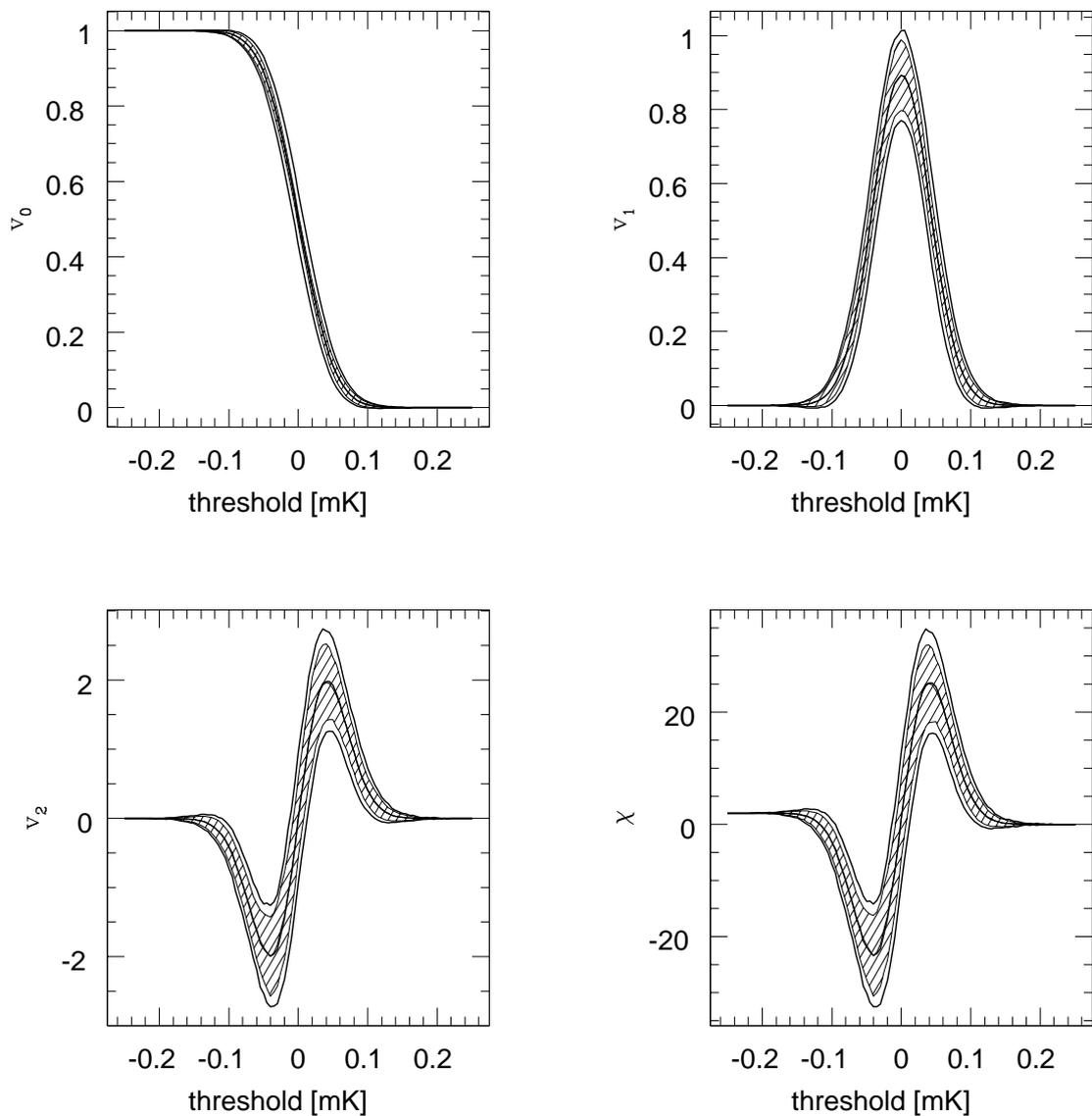
Figure 4.6 The same quantities as in Figure 4.4 are compared in this plot, but with 200 random pixels excluded instead of the galactic cut. Contamination affects more pixels than in the case of a galactic cut – if a residue of 1% is allowed, 2,842 out of the original 6,144 pixels remain, compared to 3,189 for a galactic cut. However, the estimated Minkowski functionals are more robust.
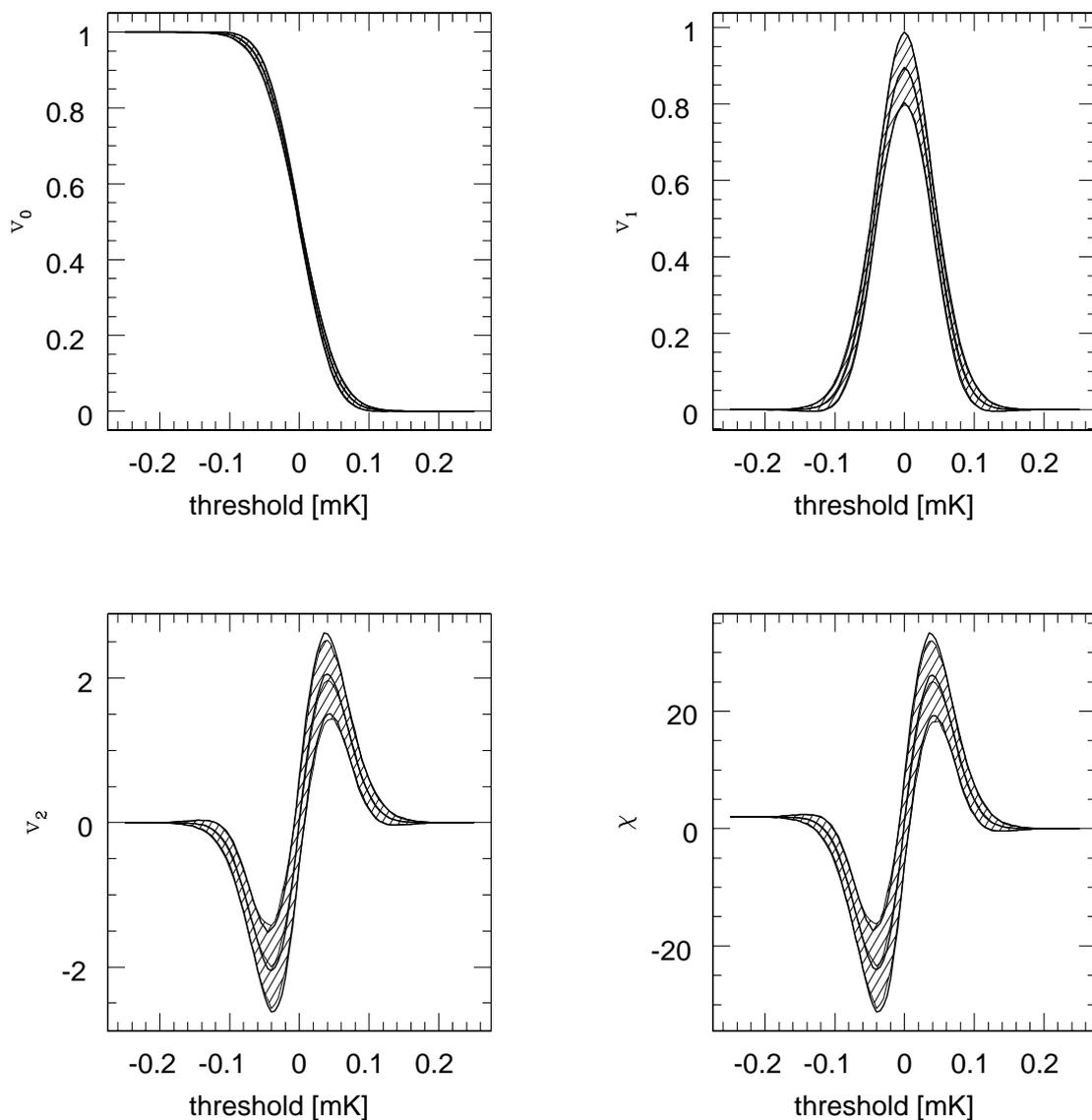
5% produces sufficiently unbiased estimates. Note that while the mean values agree completely after applying the correction, the variance of the estimators has increased, simply because fewer data points result in poorer statistics.

Apart from the galactic cut, point source contamination is another important source of incomplete sky coverage. Figure 4.6 shows the bias introduced by omitting 200 randomly scattered pixels. Obviously, the effect is much less pronounced compared to the realizations excluding the galactic cut; in fact the differences between the all–sky realizations and the restricted realizations are barely visible. Both the galactic cut and the random point cut affect roughly 3,000 of the 6,144 DMR pixels with a contribution of 1% or above, so at first sight our findings appear inconsistent. However, they can be explained with the prominent geometric features – namely, almost straight edges – in the galactic cut. These are missing in a random point distribution, so the errors remain smaller and average out.

4.2.4   How smoothing leads to noise reduction

In order to obtain a regular field, and to reduce the level of the additive noise present in the data, it is necessary to apply a smoothing kernel to the data before calculating the Minkowski functionals. Usually, the choice of a particular width is largely arbitrary. Here we show the example introduced in Section 4.3 with different degrees of smoothing applied to illustrate the behaviour of Minkowski functionals in the presence of noise.

The situation for $2°$ smoothing, where noise still makes an appreciable contribution, is shown in Figure 4.7. The surface area v_0 is much less affected than the other Minkowski functionals; this is due to the fact that noise is incoherent and forms comparatively small hot and cold spots. However, these spots are almost as intense as the signal contribution, as can be seen from the almost equal width of all curves, and far more numerous – the Euler characteristic for the noise field alone reaches a maximum of the order of 200. Even though the extrema in the pure noise maps are spread out over the whole range of thresholds when added to the signal, and hence their number at a specific threshold decreases, their contribution is still sufficiently high to make the signal appear completely different compared to the combination of signal and noise.

In Figure 4.8, where the results for $8°$ smoothing are displayed, noise is almost completely invisible in comparison to the signal. Only about two dozen extrema of either kind (compare the extrema of the Euler characteristic) remain, but since they have become extremely shallow, their contribution is not significant any more; the pure signal and the combination of signal and noise differ only marginally. Unfortunately, at a resolution of $8°$ the remaining signal does not carry too much cosmological information.

With this example, the behaviour of Minkowski functionals under filtering at different scales has only been hinted at. The two filter widths of $2°$ and $8°$ are chosen to show two extremal possibilities, namely total dominance of noise and total reduction of noise. In practice, the intermediate value of $3°$ turns out to give good enhancement of signal, while preserving small–

Figure 4.7 Signal (shallow curves, empty area), noise (sharply peaked curves, empty area) and combination of both (shaded area) at 2° smoothing. At this scale, the noise contribution contains far more features than the signal. Moreover, they are strong enough to persist when distributed over the whole range of the signal. Hence the combination of signal and noise displays a completely different morphology from the pure signal contribution.
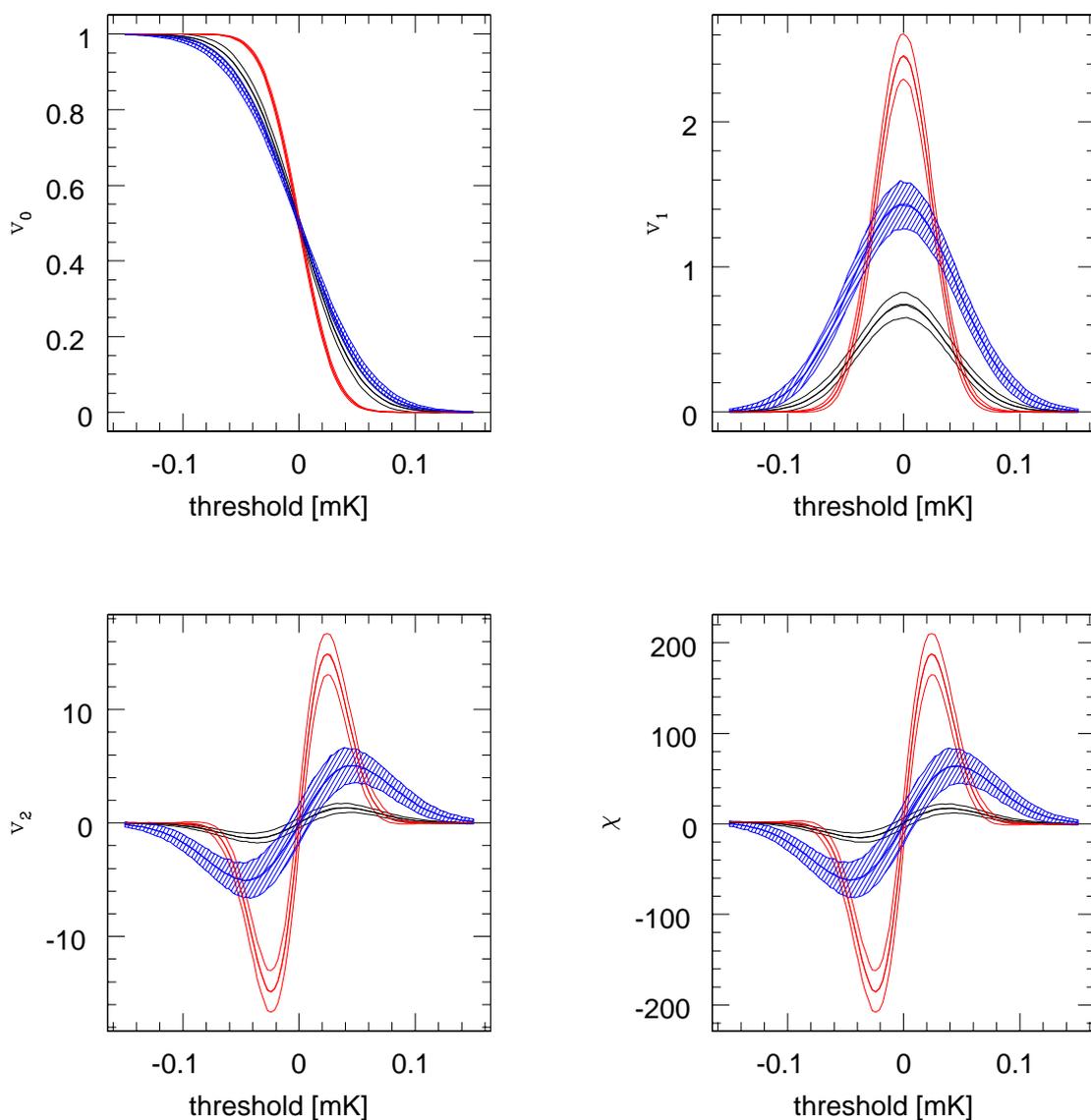
Figure 4.8 Signal (shallow curves, empty area), noise (sharply peaked curves, empty area) and combination of both (shaded area) of the same random field as in Figure 4.7, but at 8° smoothing. Although the pure noise map still features roughly two dozen extrema, their amplitude has been reduced to such extent that the pure signal is hardly affected when combined with the noise.
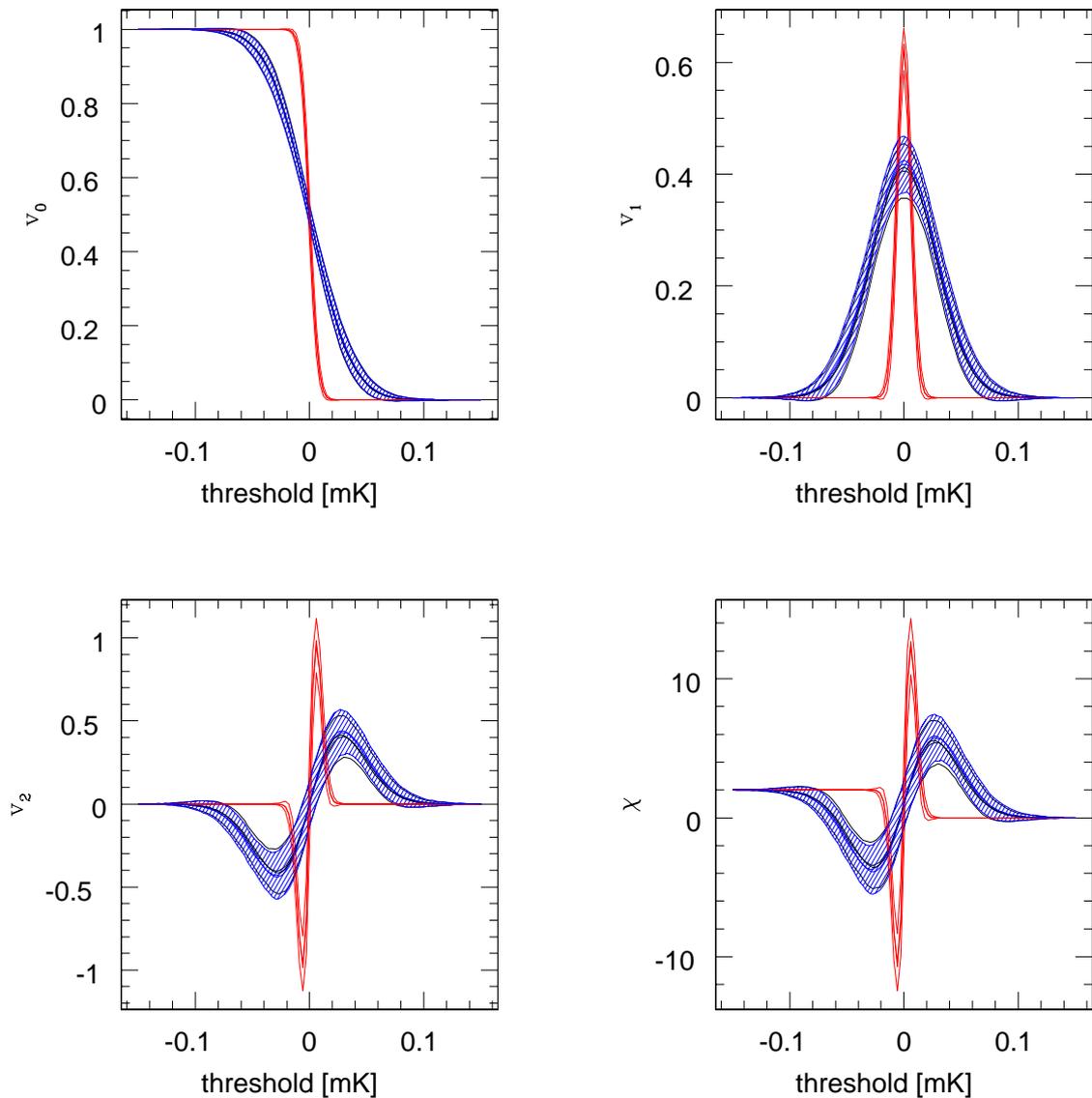
scale information as well.

Image analysis has developed the concept of scale space by introducing a continuously changing filter scale as an additional coordinate (ter Haar Romeny et al. 1991). It turns out that Minkowski functionals may be connected to this approach (Siegmund & Worsley 1995), being spatial averages of geometric invariants formed from derivatives. This connection is interesting as a theoretical problem itself, and with possible applications to astrophysics in mind, it certainly deserves deeper study.

So far, we have studied a simplified yet realistic model to test the theoretically derived estimators in their application to simulated CMB sky maps. Among other tests, we checked whether the estimators remain unbiased when incomplete data is smoothed over the edge of a galactic cut, and found a prescription to deal with this problem while preserving as much information as possible. To our knowledge, this is the only instance where both the curvature of the sky and the incompleteness of observations are addressed. However, a variety of applications have been performed on maps with full sky coverage (Novikov et al. 1999), and on nearly flat sections of the sky (Hobson et al. 1998).

4.3 Examples of non–Gaussian random fields

4.3.1 The χ^2 random field

The relevance of the χ^2 random field to the statistics of CMB temperature anisotropies was motivated from inflationary scenarios e.g. by Linde & Mukhanov (1997). Even before that, the impact of a variety of non–Gaussian initial fields on the process of structure formation was investigated (Moscardini et al. 1991; Coles et al. 1993b). In the standard model with Gaussian initial conditions, the CMB polarisation amplitude is a realisation of a χ^2 field with two degrees of freedom Naselsky & Novikov (1998).

In the context of pattern recognition, the statistics of the χ^2 random field have been investigated by Worsley (1994). Most notably, Siegmund & Worsley (1995) and Worsley (1998) present interesting applications to point source detection.

Consider N independent, stationary and sufficiently regular Gaussian random fields X_i with zero mean and correlation function $\xi_{\mathrm{G}}(r)$. Let us form a new random field U by

$$U = \sum_{i=1}^{N} X_i^2. \tag{4.14}$$

Th random field U is called a χ_N^2–random field, in analogy to a χ_N^2 random variable, which is the sum of the squares of N independent Gaussian random variables. By simply inserting

Figure 4.9 The Minkowski functionals of a χ^2 field with one degree of freedom.

Figure 4.10 The Minkowski functionals of a χ^2 field with nine degrees of freedom.

Equation (4.14) into the definition for the mean and correlation function, it is easily shown that

$$\mu = \langle U \rangle = N\xi_G(0),$$
$$\xi(x) = \langle U(\mathbf{x})U(0) \rangle - \langle U \rangle^2 = 2N\xi_G(x)^2. \tag{4.15}$$

Note that by Equation (4.15), the correlation function of a χ^2 random field is always non–negative. This leads into difficulties when trying to generate a χ^2 field with a given power spectrum.

The joint probability densities for the value and the first and second derivatives for the χ^2 field are calculated by Worsley (1994). Using these densities to calculate the average of the invariants given in Equation (4.8), we obtain the expected Minkowski functionals per unit area of the χ^2 random field with N degrees of freedom in two dimensions. In summary[3]:

$$v_0(u) = \frac{\Gamma\left(\frac{N}{2}, \frac{u}{2\sigma_G}\right)}{\Gamma\left(\frac{N}{2}\right)}$$

$$v_1(u) = \frac{1}{4}\sqrt{\frac{\pi\tau_G}{\sigma_G}}\left(\frac{u}{2\sigma_G}\right)^{\frac{N-1}{2}}\frac{\exp\left(-\frac{u}{2\sigma_G}\right)}{\Gamma\left(\frac{N}{2}\right)} \tag{4.16}$$

$$v_2(u) = \frac{\tau_G}{2\pi\sigma_G}\left(\frac{u}{2\sigma_G}\right)^{\frac{N-2}{2}}\frac{\exp\left(-\frac{u}{2\sigma_G}\right)}{\Gamma\left(\frac{N}{2}\right)}\left(\frac{u}{2\sigma_G} - (N-1)\right)$$

Here σ_G and τ_G are the variance of the underlying Gaussian fields X_i and their first derivatives $X_{i;j}$, respectively. They can be related to the corresponding quantities σ and τ of the chi^2 field U itself by

$$\sigma = \xi(0) = 2N\sigma_G \tag{4.17}$$

and

$$\tau = -\xi''(0) = -\left.\frac{\mathrm{d}^2}{\mathrm{d}r^2}\left[2N\xi_G(r)^2\right]\right|_{r=0} = 4N\sigma_G\tau_G, \tag{4.18}$$

using the fact that the correlation function's first derivative vanishes at $r = 0$ for a sufficiently regular field.

Figures 4.9 and 4.10 display the Minkowski functionals of χ^2 fields with one and nine degrees of freedom, respectively. The two central lines in each panel follow the analytical expectation value from Equation (4.16), and the average over 40 numerical realisations of the field. The areas indicate the standard deviation over all realisations.

Obviously, the morphology of the χ^2 random fields in our examples bears no resemblance to the Gaussian case. The most striking feature is probably the discontinuity of the circumference

---

[3]$\Gamma(\nu, x)$ is the incomplete Gamma function. For our purposes, the recurrence relation $\Gamma(\nu+1, x) = \nu\Gamma(\nu, x) + x^\nu \exp(-x)$ with starting values $\Gamma(\frac{1}{2}, x) = \sqrt{\pi}\left(1 - \Phi(\sqrt{x})\right)$ and $\Gamma(1, x) = \exp(-x)$ is sufficient.

Figure 4.11 A map of Earth's topography.
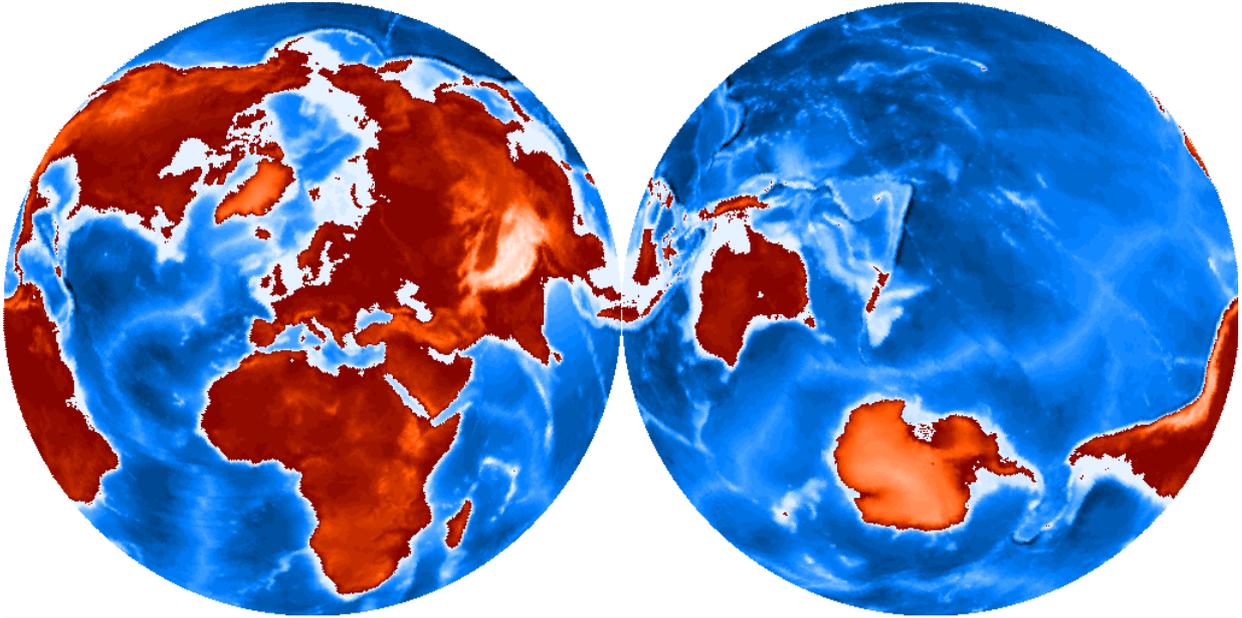
v_1 for the case of $N = 1$, i.e. a single degree of freedom. A similar jump takes place in the Euler characteristic v_2 for $N = 2$. However, even moderate amounts of Gaussian noise destroy this behaviour.

By the central limit theorem, the probability distribution of the χ^2 field approaches the one for the Gaussian field as the number of degrees of freedom increases. This makes the χ^2 field ideal for testing the sensitivity of morphological statistics towards non–Gaussian features. Kogut et al. (1996) use this in their analysis of non–Gaussian statistics of the four–year COBE DMR maps, and White (1999) proposes a similar approach to testing the statistics of the initial density field.

4.3.2   The Earth

Figure 4.13 shows the Minkowski functionals of the earth's topography. The map (see Figure 4.11) was constructed by binning the Etopo5 data[4] onto a HEALPix pixelisation (Gorski et al. 1999) of the sphere.

The curves reveal a number of characteristic features. All functionals experience a fairly sharp change at a depth between 6,000m and 5,000m, which is roughly the average depth of the seafloor. A peak of several 1,000m width and almost constant height of the boundary length v_1, and a corresponding minimum in the Euler characteristic χ indicate the rise of the oceanic ridges.

---

[4]The Etopo5 database gives elevations on a cylindrical grid of 5 arcminute spacing. The data files may be obtained from the net via `ftp://walrus.wr.usgs.gov/pub/data/`; see also *Data Announcement 88-MGG-02, Digital relief of the Surface of the Earth* by the NOAA, National Geophysical Data Center, Boulder, Colorado, 1988.

Figure 4.12 A Gaussian random field with identical first– and second–order characteristics.
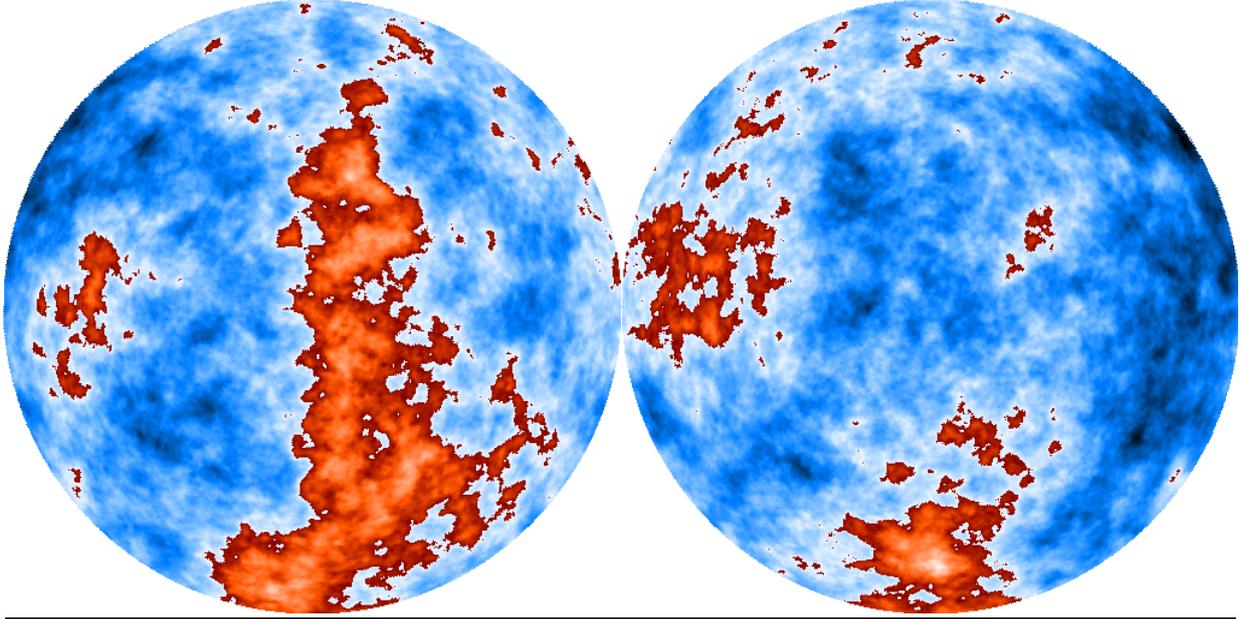
From 3,000m below sea level to slightly positive elevations, the boundary length remains largely constant, as the continental shelfs rise from the oceans; meanwhile, the Euler characteristic fluctuates with the disappearance of the oceanic ridges, and the opening of shallower, marginal parts of the oceans such as the Mediterranean, the Carribean sea or the Arctic sea. Most of the land mass does not rise beyond 1,000m, so all Minkowski functionals gradually decline after this height; a few small peaks in the Euler characteristic may be – cautiously – identified with Antarctica, the Rocky Mountains, the Andes, and the Himalaya.

Again, the Minkowski functionals in this example look strongly non–Gaussian. However, we may mix the original map of Earth's topography with a Gaussian random field of the same second–order characteristics (see Figure 4.12 for a realisation). In Figure 4.15, we compare a mixture of one part Earth and four parts Gaussian field to the purely Gaussian case. From looking at the simple Minkowski functional analysis, the mixed field appears perfectly consistent with Gaussianity.

However, after calculating the Minkowski functionals of a CMB map, we are left with three functionals as functions of a number of threshold values. These data will normally be strongly correlated. More information can be obtained by a principal component analysis. This is a standard procedure in statistics (Kendall 1980), and has already found its way into the statistical analysis of cosmological datasets (Springel et al. 1998).

We arrange the three Minkowski functionals at different thresholds as a vector–valued random variable $\mathbf{X}$, with N components X_i. The mean $\boldsymbol{\mu}$ and the covariance matrix $\boldsymbol{\sigma}$ of $\mathbf{X}$ are given by

$$\mu_i = \langle X_i \rangle, \qquad \sigma_{ij} = \langle (X_i - \mu_i)(X_j - \mu_j) \rangle. \tag{4.19}$$

Figure 4.13 Minkowski functionals of the earth's topography from the map in Figure 4.11 with a 12′ Gaussian filter applied. The resulting Minkowski functionals (jagged lines) look decidedly non–Gaussian; compare the analytical expectation values (smooth lines). Pronounced features in the Minkowski functional curves can be identified as corresponding to the main features of the earth's topography, such as oceanic ridges, continents and mountain ranges. See the main text for details.
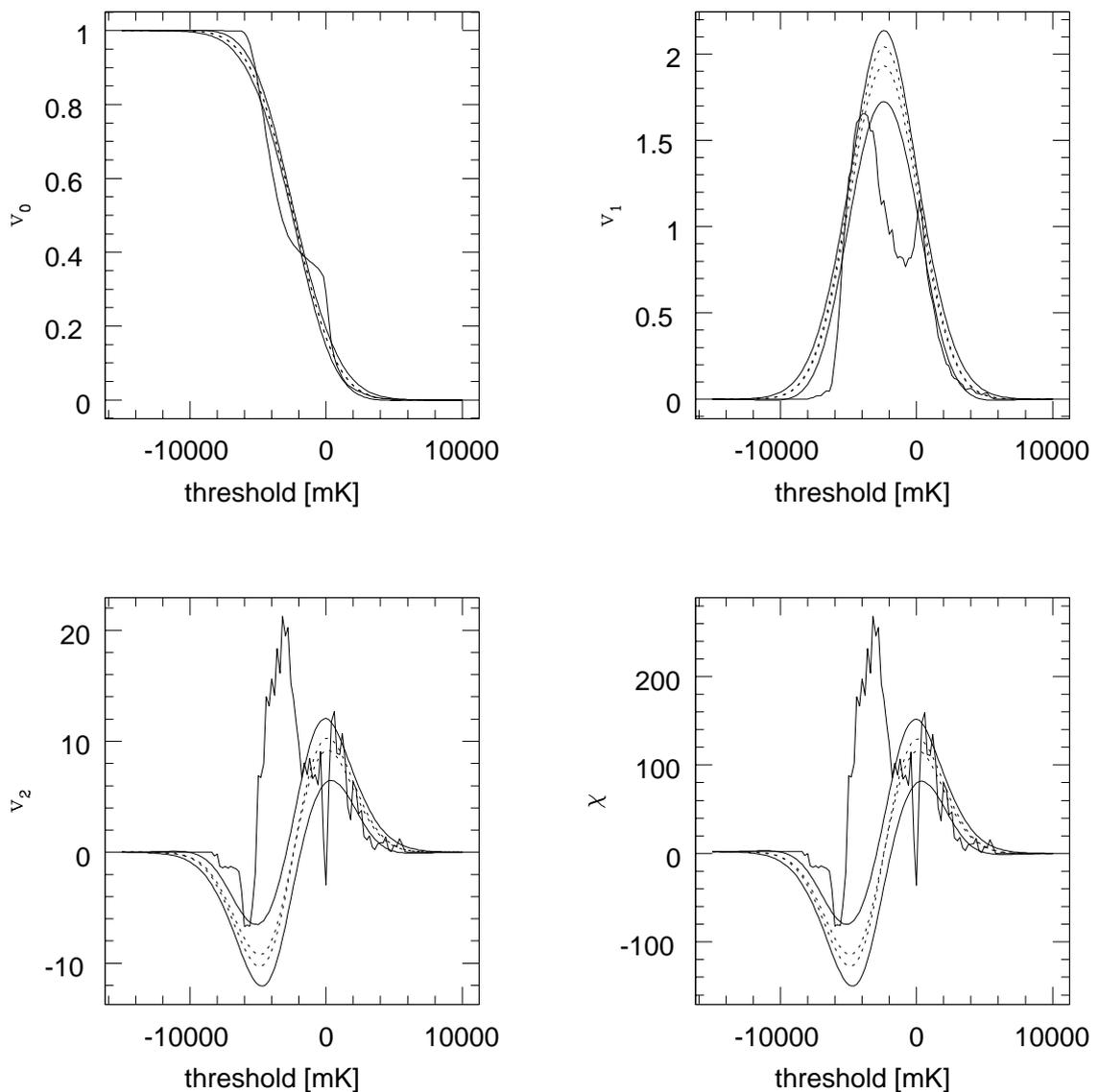
Figure 4.14 Minkowski functionals of a Gaussian random field with the same power spectrum as the Earth, shown in the panels in Figure 4.12.
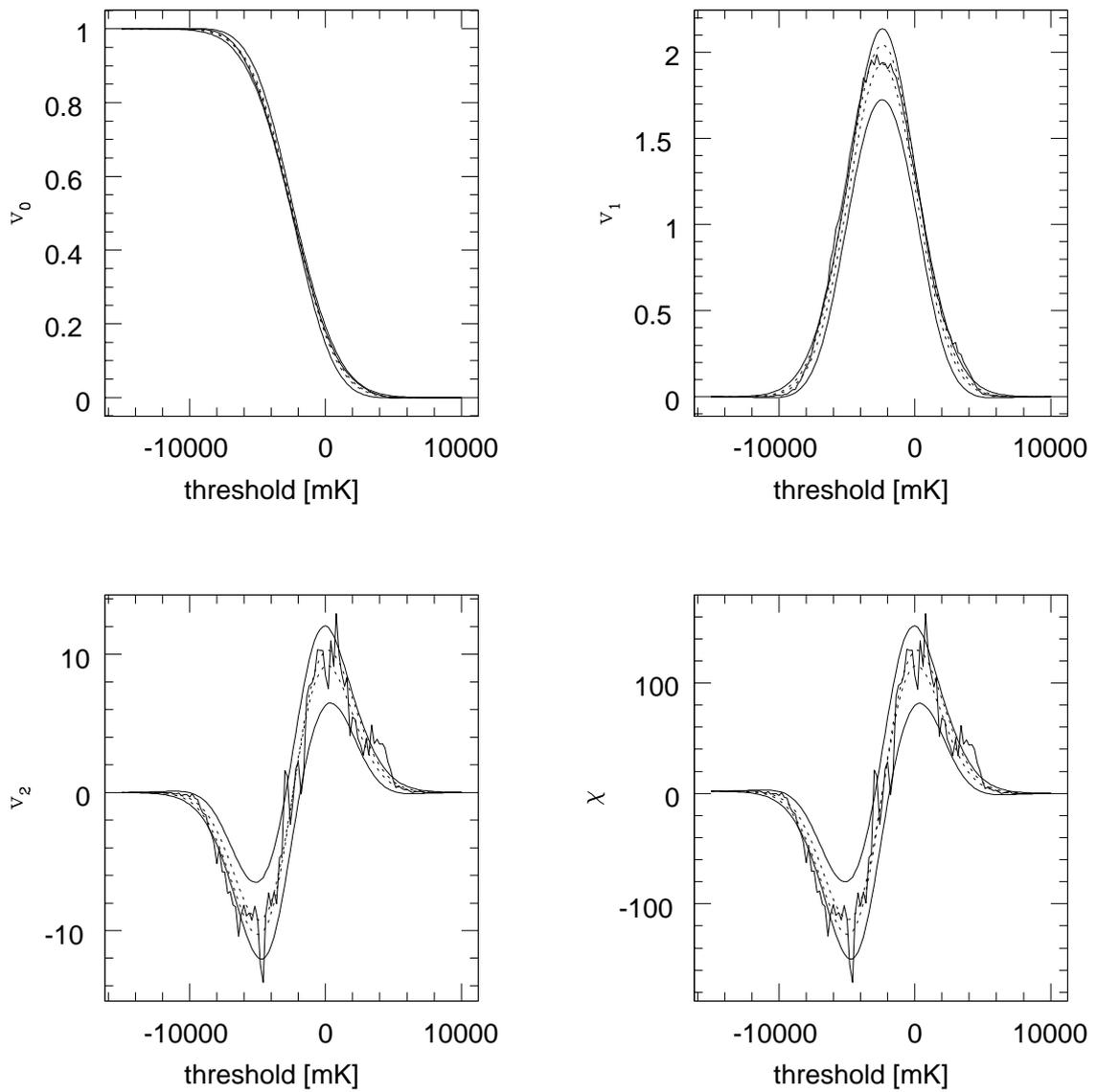
Figure 4.15 Minkowski functionals of a mixture of the Earth map from Figure 4.11 and a Gaussian contamination with the same power spectrum. The ratio of both was chosen to give a signal–to–noise ratio of 1:4.
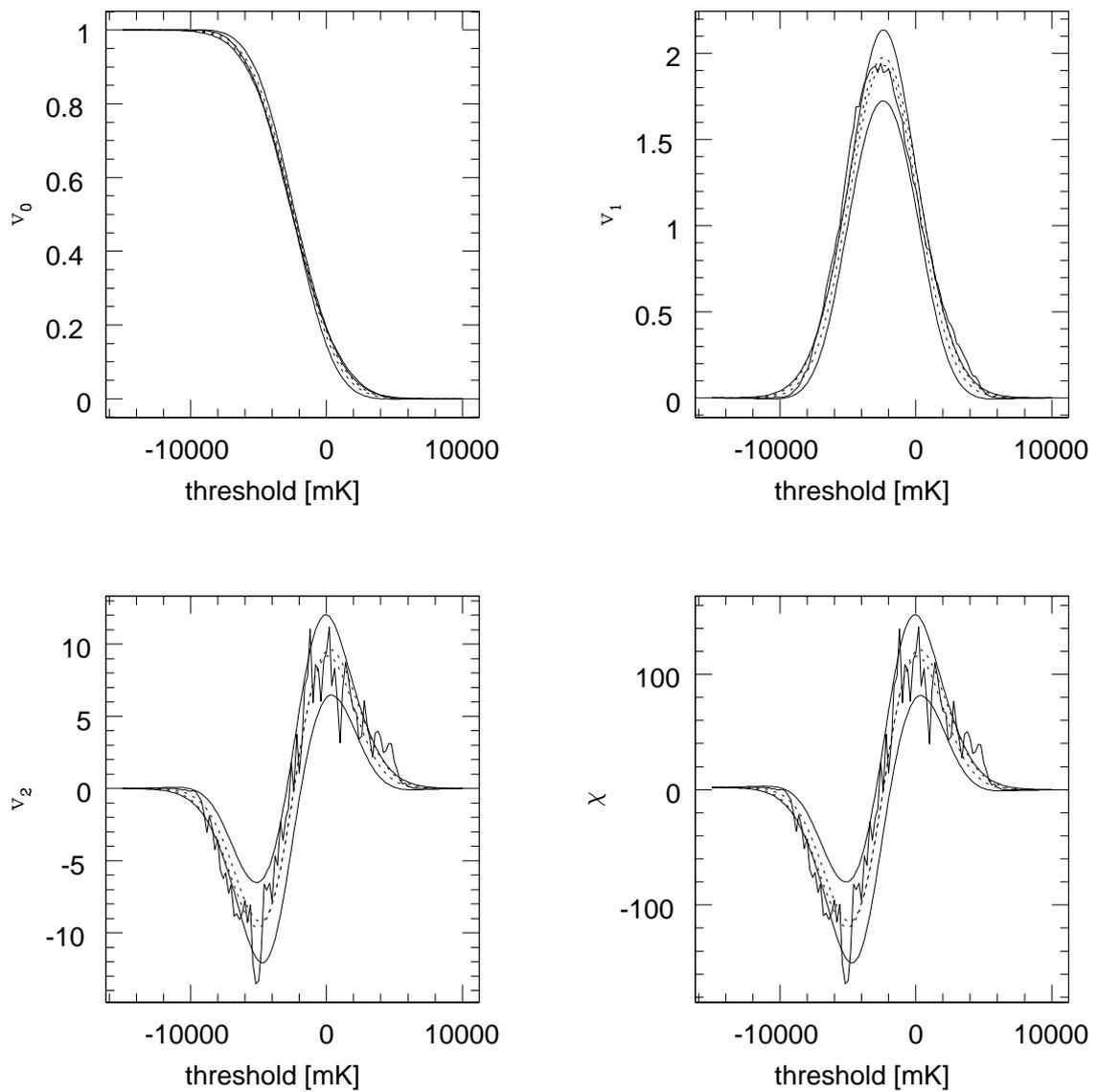
Figure 4.16 This panel shows the principal components of the Earth's topography mixed with a Gaussian random field in the ratio 1:4 (drude's feet) compared to a purely Gaussian random field (circles). On the x–axis the number of the principal component is displayed when the components are sorted by ascending eigenvalue. The y–axis gives the deviation of the data from the average of a Gaussian random field in units of the eigenvalue. Only 56 symbols fall into the dark shaded region, while for a purely Gaussian random field, 72.5 ± 6.8 hit that area. Also, the scatter is significantly larger.
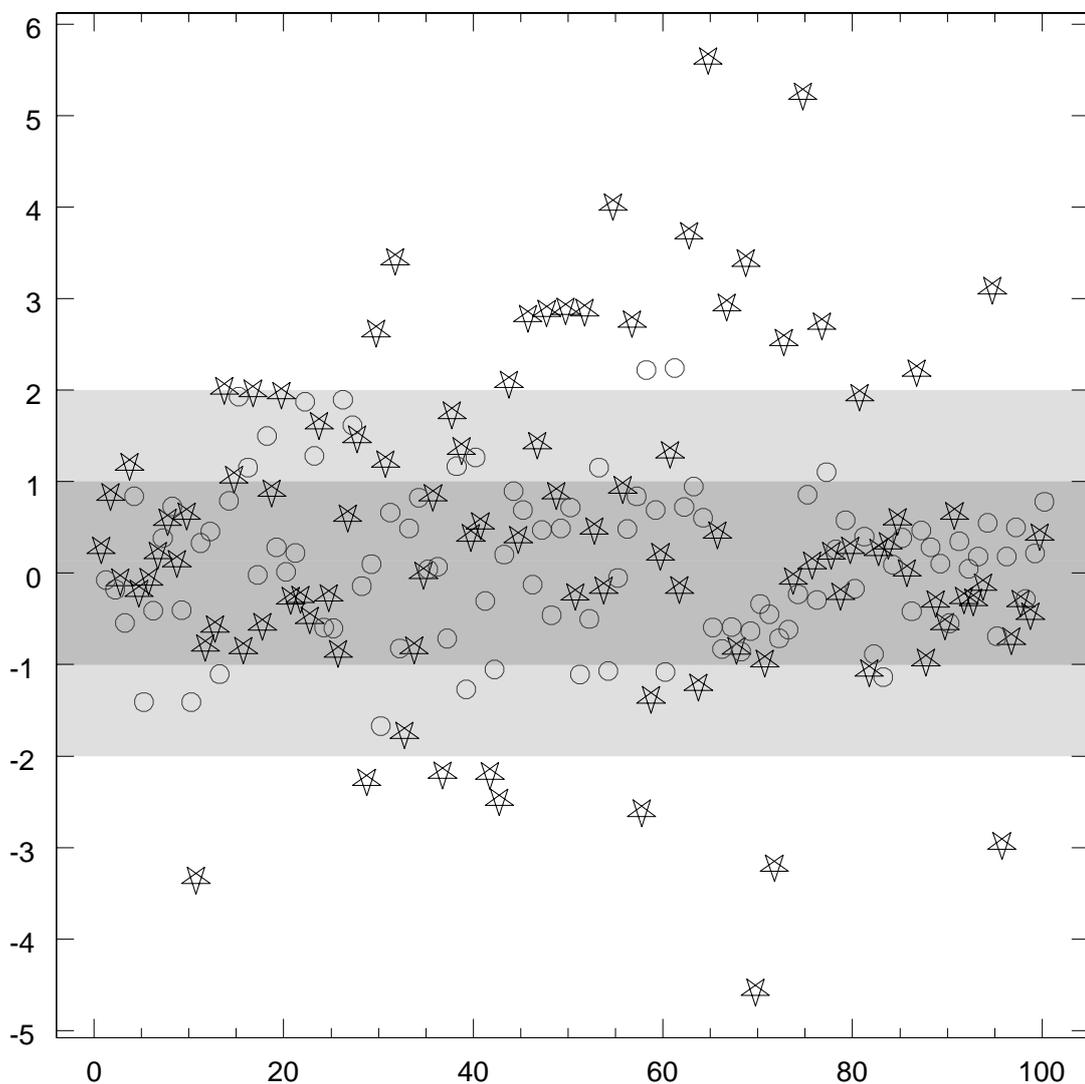
We estimate $\boldsymbol{\mu}$ and $\boldsymbol{\sigma}$ from a set of n realisations $\mathbf{X}^{(j)}$. In our example, we use $n = 10,000$ realisations of a Gaussian random field with angular power spectrum computed from the Earth's topography. For each realisation, the Minkowski functionals were computed in bins of 100 meter width ranging from -15,000 meter to 10,000 meter, so altogether, $N = 378$ functional values were obtained.

In order to identify the principal components, one writes $\boldsymbol{\sigma}$ as

$$\boldsymbol{\sigma} = \mathfrak{U}.\mathfrak{W}.\mathfrak{U}^{\mathrm{T}}, \tag{4.20}$$

where $\mathfrak{U}$ is an orthogonal $N \times N$ matrix, while $\mathfrak{W}$ is diagonal, containing the N eigenvalues of the covariance matrix. Since the strong correlations between data points lead to $\mathfrak{W}$ having many eigenvalues close to zero, the matrices $\mathfrak{U}$ and $\mathfrak{W}$ are best determined by singular value decomposition (see Press et al. 1987 for details). The vector $\mathfrak{U}.(\mathbf{X} - \boldsymbol{\mu})$ then consists of N uncorrelated random variables with zero mean.

The principal components of Earth's topography contaminated with Gaussian noise at a signal–to–noise ratio of 1:4 are shown in Figure 4.16. Although the Minkowski functionals themselves (Figure 4.15) show no significant deviation from the Gaussian expectation values, the principal component analysis reveals differences at a level of 5σ.

4.3.3 The COBE DMR data

As a last example, let us take a look at data that are both real and cosmologically relevant. Figure 4.17 shows a map of the microwave sky as seen at 53GHz by the COBE satellite after four years of observing (Bennett et al. 1996). The data are restricted to 3,189 pixels receiving less than 1% from a smoothed galactic cut, when a Gaussian filter of 3° width is applied.

As stated above, considerable uncertainties are introduced through the estimates of the parameters μ, σ and τ entering the analytical expectation values for the Minkowski functionals of a Gaussian random field. In order to make this statement more quantitative, Figure 4.18 summarizes the parameters determined from the 1,000 mock realizations used for the shaded area in Figure 4.19. Relative errors for the relevant parameters lie in the range of five to ten per cent, which is not too bad considering that little more than 3,000 data points enter our analysis.

Figure 4.19 displays the corresponding Minkowski functionals. Obviously, the analysis carried out on all 6,144 pixels is dominated by galactic emission, while the field with the galactic cut applied appears consistent with the assumption of a stationary Gaussian random field.

As explained above, we conduct a principal component analysis of the Minkowski functionals of the COBE DMR data, comparing the results calculated from the real data to 10,000 simulated realisations of a Gaussian random field. Figure 4.20 shows our results; no deviation from the Gaussian behaviour can be seen.

This finding does not come unexpected even in view of the recent detection of non–Gaussian signals from the DMR maps (see e.g. Ferreira et al. 1998; Novikov et al. 1999; Heavens 1999,

Figure 4.17 This panel shows the four–year data from the COBE DMR 53GHz channel, with a customized galactic cut and a smoothing filter of $3°$ width applied. Figure 4.19 displays the corresponding Minkowski functionals calculated both from this map, and from the all–sky map with galactic signal dominating.
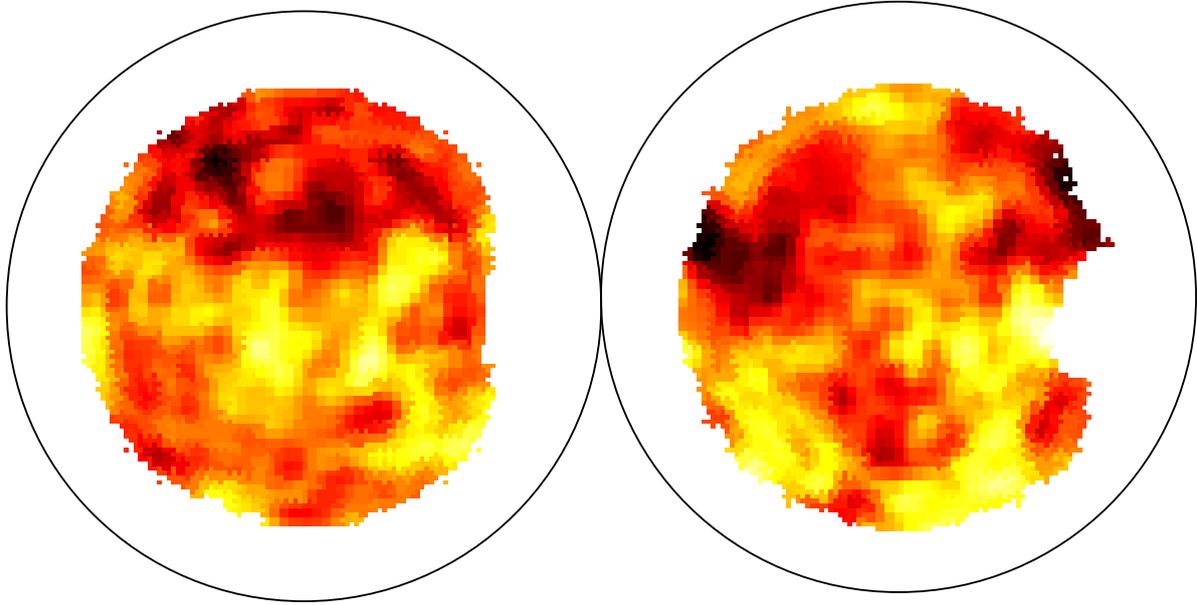
and especially the references in Bromley & Tegmark 1999). The first evidence for non–Gaussian contributions in the COBE DMR four–year data was reported by Ferreira et al. (1998) using collapsed bispectrum statistics. Their result sharply locates the non–Gaussian signal at the $\ell = 16$ multipole, while our analysis mainly probes the $4°$ scale in real space. Excessive tests by Banday et al. (1999), who eventually put forward a non–cosmological explanation, even show that the non–Gaussian signal may disappear completely through different data processing, for example if pixel locations are shifted. Apparently, we should either consider statistics which are tunable to the type of non–Gaussian field we wish to identify, or further refine the method of Minkowski functional analysis.

Figure 4.18 Uncertainties in the parameter estimates. Parameters were determined from each of the 1,000 simulated DMR maps, and binned into histograms. In order to give an idea of the amount of uncertainty, the range is extended to include zero in all plots. Obviously, the fluctuations are still fairly large in comparison to the mean value; to be precise, the relative r.m.s. fluctuation is roughly 8% for μ, almost 6% for τ, and more than 9% for $\sqrt{\sigma}$.
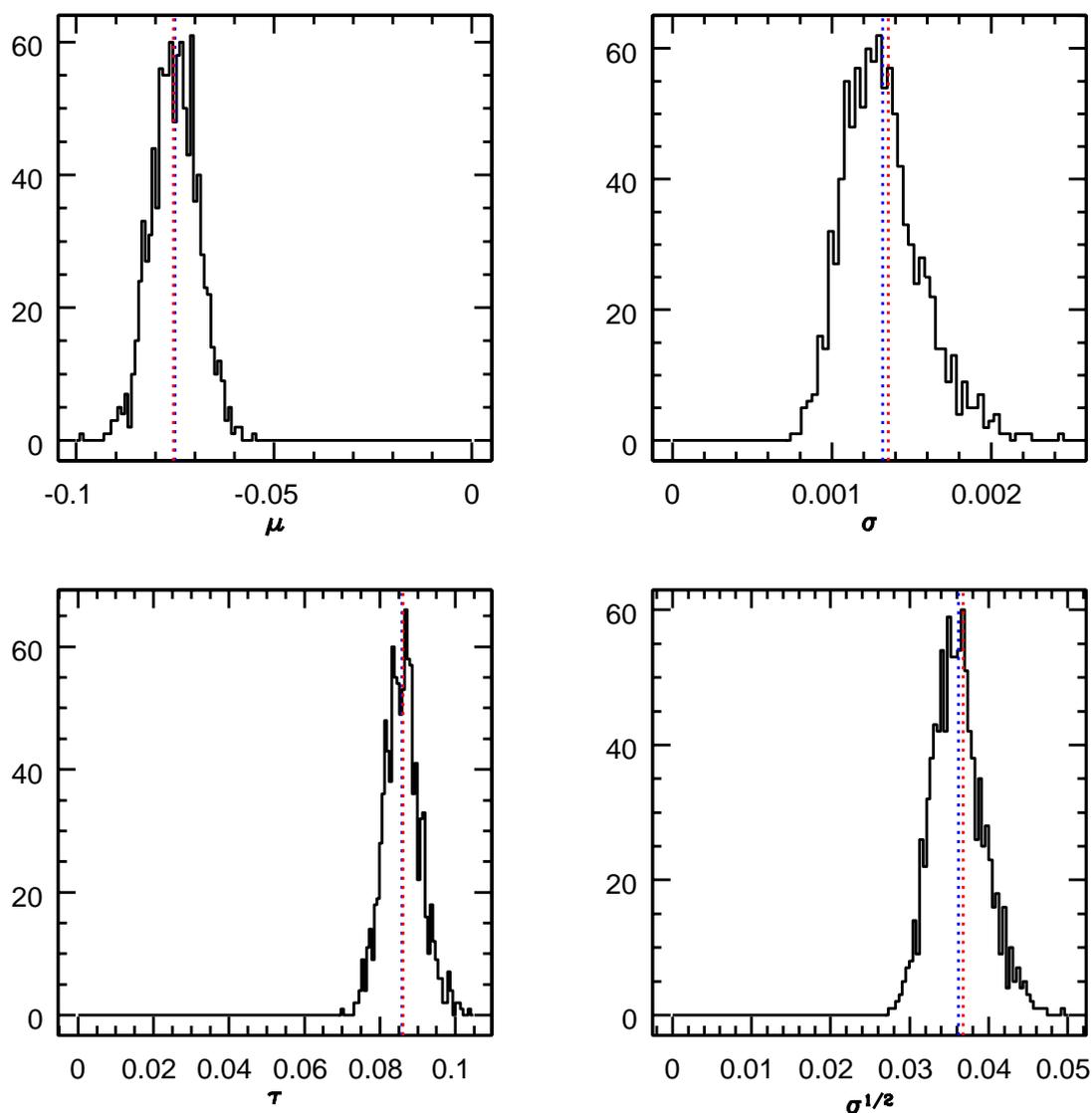
Figure 4.19 Minkowski functionals of the four–year COBE DMR map in the 53GHz channel. As expected, the full map of 6,144 pixels (dashed line) is severely affected by galactic emission. With a galactic cut, the functionals determined from the map (solid line) become consistent with a stationary Gaussian random field. The shaded area indicates the mean and variance of 1,000 realizations of a Gaussian random field with Harrison–Zel'dovich spectrum and pixel noise; normalisations were chosen to reproduce the parameters μ, σ and τ determined from the data via Equation (4.12).
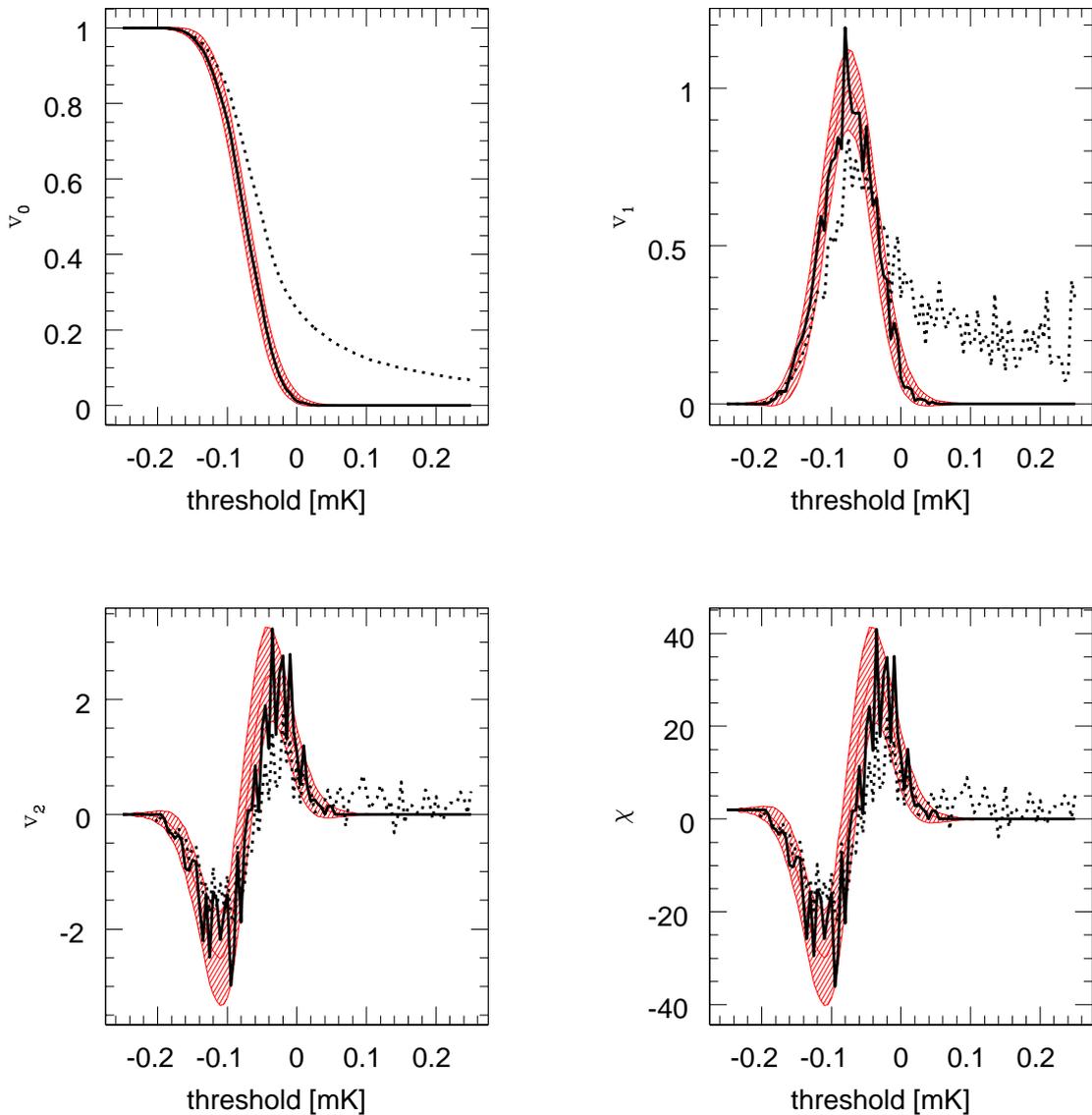
Figure 4.20 Principal components of the DMR compared to a Gaussian random field.
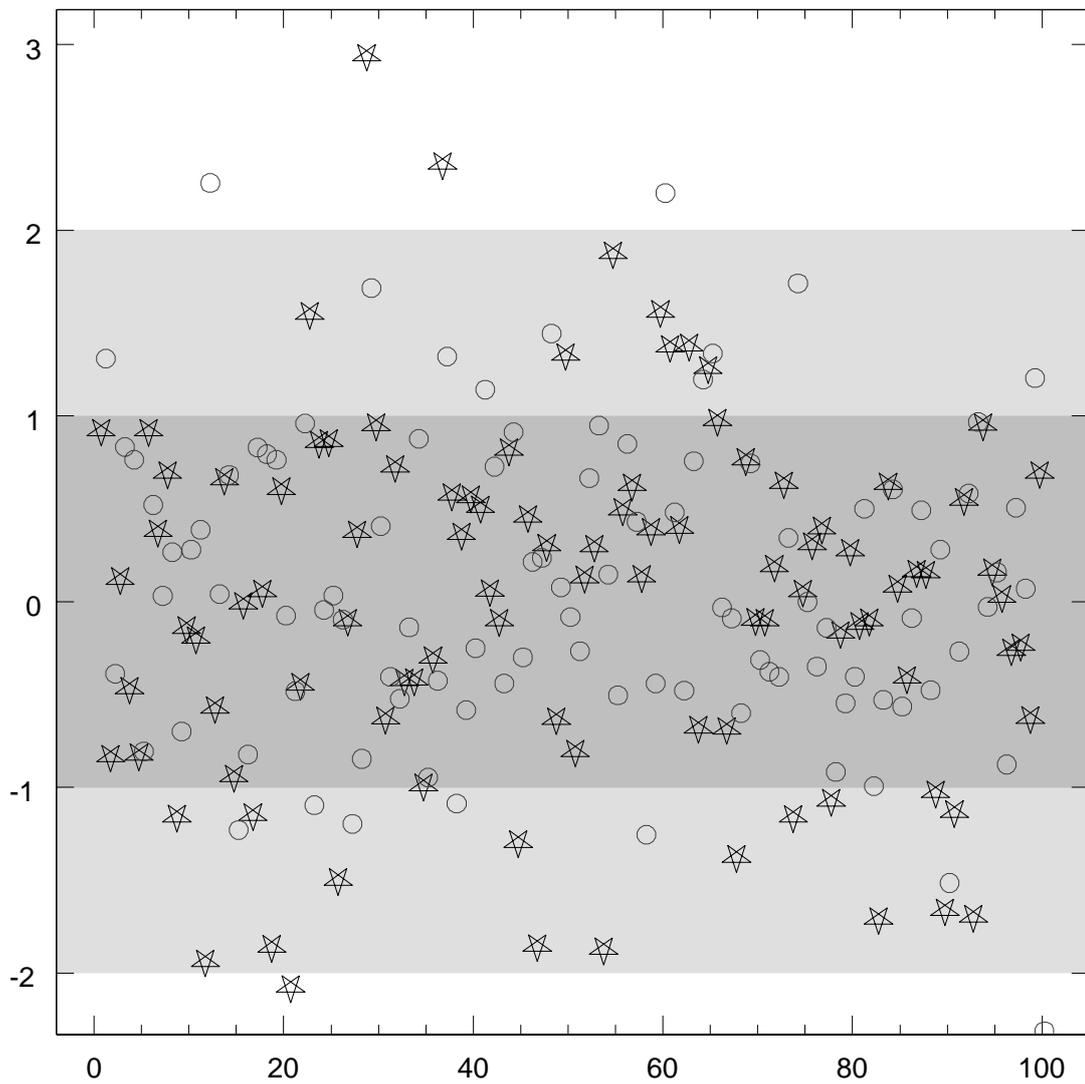
5 The distribution of dark matter

Wir weben emsig Tag und Nacht.
(Heine 1844)

The existence of the large–scale clustering of galaxies had already been well established by the early 1970's mainly due to the pioneering work of Totsuji & Kihara (1969) and Peebles (1974) who showed that the two–point correlation function for galaxies in the Lick and Zwicky catalogues was positive and had the power-law form $\xi(r) \propto r^{-1.8}$ on scales $\lesssim 10h^{-1}\mathrm{Mpc}$. Their result was later extended to three–dimensional galaxy catalogues as well. Although the clustering of galaxies is now a well–known fact, a complete description of clustering which includes its geometrical features has so far eluded researchers. This is perhaps due to the fact that the galaxy density field which we observe appears to be strongly non–Gaussian. A Gaussian random field is uniquely described by its power spectrum $P(k)$, or its two–point correlation function $\xi(r)$, since $\xi(r)$ and $P(k)$ form a Fourier transform pair. This is no longer true for a non–Gaussian field for which ξ must be complemented by other statistical descriptors which are sensitive to the structure of matter on large scales. In the so-called "standard model" of structure formation, an initially Gaussian density distribution becomes non–Gaussian due to mode-coupling and the resultant build up of phase correlations during the non–linear regime. These phase correlations give rise to the amazing diversity of form, which is characteristic of a highly evolved distribution of matter, and is often referred to as being cellular, filamentary, sheet–like, network–like, sponge–like, a cosmic web etc. Most of these descriptions are based either on a visual appearance of large–scale structure or on the presence of features that are absent in the reference Gaussian distribution which, by definition, is assumed to be featureless. Gravitational instability, for example, may cause CDM–like Gaussian initial perturbations to evolve towards a density field that percolates at a higher density threshold, i.e. at a lower filling factor, than a Gaussian field (Melott et al. 1983). Such distributions display greater connectivity and are sometimes referred to as being "network–like" (Yess & Shandarin 1996).

In order to come to grips with the rich textural possibilities inherent in large–scale structure, a number of geometrical indicators of clustering have been proposed in the past including minimal spanning trees (Barrow et al. 1985), the genus curve (Gott III et al. 1986), percolation theory (Zel'dovich & Shandarin 1982; Shandarin 1983) and shape analysis (Sathyaprakash et al. 1998a; Sahni & Coles 1995 and references therein). A major recent advance in our understanding of gravitational clustering has been associated with the application of Minkowski functionals (MFs) to cosmology (Mecke et al. 1994). The four MFs $V_0, \dots, V_3$ provide an

excellent description of the geometrical properties of a collection of point objects (galaxies) or, alternatively, of continuous distributions such as density fields in large–scale structure or brightness contours in the cosmic microwave background. The scope and descriptive power of the MFs is enhanced by the fact that both percolation analysis and the genus curve are members of the family. Additionally, as demonstrated by Sahni et al. (1998), ratios of MFs provide us with an excellent "shape statistic" with which one may attempt to quantify the morphology of large–scale structure, including the shapes of individual superclusters and voids. Spurred by the success of MFs in quantifying the geometrical properties of large–scale structure we apply the MFs to scale–invariant N–body simulations of gravitational clustering, with an attempt to probe both the global properties and the individual "bits and pieces" that might make up the "cosmic web" (Bond et al. 1996).

In this chapter we shall use the excursion set approach which is applicable to continuous fields. In Section 5.1 we attempt to reconstruct the density field from the discrete particle positions produced by the simulation run. We employ methods from statistical data analysis (Silverman 1986) to quantify the accuracy of the commonly used smoothing technique that simply convolves the particle set with a kernel of fixed width. A considerable refinement of the technique is possible if the kernel width is allowed to vary with the density. In Section 5.2 we illustrate the improvements of our methods using large N–body simulations of a Cold Dark Matter spectrum in various cosmologies. To explore a wider range of initial conditions, another analysis of density fields is performed in Section 5.3, this time on simulations of initial power–law spectra. Apart from the Minkowski functionals of the global field, we also identify individual coherent pieces of the isodensity contour and investigate the statistics of their Minkowski functionals using the shapefinders suggested by Sahni et al. (1998).

5.1   Reconstructing the density field

Throughout this Section, we consider a set $\mathfrak{X}$ of n data points $\mathbf{x}_i$ in d dimensions. The points are randomly sampled from a d–dimensional probability density, that is a non–negative and normalised field $u(\mathbf{x})$. In our context, this field is taken to be the normalised dark matter density,

$$u(\mathbf{x}) = \frac{\rho(\mathbf{x})}{\bar{\rho}} = \delta(\mathbf{x}) + 1. \tag{5.1}$$

The goal consists in reconstructing the underlying true field u given the set $\mathfrak{X}$.

A variety of density estimation procedures are described in the monograph by Silverman (1986), which is standard to statisticians but not widely known to cosmologists. Some applications have been introduced into cluster analysis (Pisani 1993, Pisani 1996, Fadda et al. 1998), but the methods presented there may not be naïvely applied to large–scale structure analysis.

We will consider a subclass of density estimation techniques using smoothing kernels for assigning a continuous density to a discrete particle set. In this context, the kernel $K(\mathbf{x})$ is a

non–negative, symmetric, and normalised function. Usually, the kernel is also chosen to have unit variance, that is

$$\int \mathrm{d}^d x \, x_i^2 K(\mathbf{x}) = 1 \text{ for } i = 1, \dots, d. \tag{5.2}$$

5.1.1 Fixed width smoothing

The simplest kernel density estimate treats each particle from the data set $\mathfrak{X}$ in the same manner, regardless of their environment. After stretching the kernel to a given bandwidth λ, it is convolved with the particle set, and gives the so–called fixed width kernel density estimate $\hat{u}_\lambda$.

$$\hat{u}_\lambda(\mathbf{x}) = \frac{1}{n\lambda^d} \sum_{i=1}^{n} K\left(\frac{\mathbf{x} - \mathbf{x}_i}{\lambda}\right). \tag{5.3}$$

A priori, the kernel width λ is a free parameter of the method. In order to choose an appropriate value, we attempt to quantify the deviation of the estimate $\hat{u}_\lambda$ from the true but unknown field u. A common measure for the error made when reconstructing a density field by smoothing is the Mean Integrated Squared Error, or MeISE. Using angular brackets to denote the expectation value over all possible samples $\mathfrak{X}$ from a fixed underlying field u, we have

$$\mathrm{MeISE} = \left\langle \int \mathrm{d}^d x \, [u(\mathbf{x}) - \hat{u}_\lambda(\mathbf{x})]^2 \right\rangle. \tag{5.4}$$

Obviously the MeISE decreases as one increases the sample length n and reduces the bandwidth λ. In many applications, it is sufficient to know the asymptotic behaviour of the MeISE as a function of n and λ for a fixed density field u. This important leading term is called the Asymptotic Mean Integrated Squared Error, or AMeISE, and we have

$$\mathrm{MeISE} = \mathrm{AMeISE} + \text{higher order terms in } \lambda \text{ or } n^{-1}. \tag{5.5}$$

Generalising the calculations by Silverman (1986), we can evaluate the AMeISE in arbitrary dimensions. To begin with, we observe the MeISE may be written as the sum of two integrals:

$$
\begin{aligned}
\mathrm{MeISE} &= \int \mathrm{d}^d x \, [\langle \hat{u}_\lambda(\mathbf{x}) \rangle - u(\mathbf{x})]^2 + \int \mathrm{d}^d x \, \left(\langle \hat{u}_\lambda(\mathbf{x})^2 \rangle - \langle \hat{u}_\lambda(\mathbf{x}) \rangle^2 \right) \\
&= \int \mathrm{d}^d x \, \mathsf{bias}(\mathbf{x})^2 + \int \mathrm{d}^d x \, \mathsf{var}(\mathbf{x}).
\end{aligned}
\tag{5.6}
$$

As indicated by the names in the last line, both integrals can be interpreted in an intuitive way. The first one contains the deviation of the kernel estimate from the true field and is therefore called the integrated square bias. The second integral is nothing but the integrated variance of

the kernel estimate. Using the explicit expression (5.3) for the kernel density estimate, the bias at a location $\mathbf{x}$ becomes

$$
\begin{aligned}
\text{bias}(\mathbf{x}) &= \langle \hat{u}_\lambda(\mathbf{x}) \rangle - u(\mathbf{x}) \\
&= \frac{1}{n\lambda^d} \sum_{i=1}^{n} \left\langle K\left(\frac{\mathbf{x} - \mathbf{x}_i}{\lambda}\right) \right\rangle - u(\mathbf{x}) \\
&= \lambda^{-d} \int \mathrm{d}^d y\, K\left(\frac{\mathbf{x} - \mathbf{y}}{\lambda}\right) u(\mathbf{y}) - u(\mathbf{x}) \\
&= \int \mathrm{d}^d t\, K(\mathbf{t})\left[u(\mathbf{x} - \lambda\mathbf{t}) - u(\mathbf{x})\right] \\
&\approx \frac{1}{2d}\lambda^2 \int \mathrm{d}^d t\, \mathbf{t}^2 K(\mathbf{t}) \Delta u(\mathbf{x}),
\end{aligned}
\tag{5.7}
$$

where we have used a Taylor series of the true field u around $\mathbf{x}$ to obtain the leading order of the integral. We proceed analogously with the variance.

$$
\begin{aligned}
\text{var}(\mathbf{x}) &= \langle \hat{u}_\lambda(\mathbf{x})^2 \rangle - \langle \hat{u}_\lambda(\mathbf{x}) \rangle^2 \\
&= \frac{1}{n^2\lambda^{2d}} \sum_{i,j=1}^{n} \left\{ \left\langle K\left(\frac{\mathbf{x}-\mathbf{x}_i}{\lambda}\right) K\left(\frac{\mathbf{x}-\mathbf{x}_j}{\lambda}\right) \right\rangle - \left\langle K\left(\frac{\mathbf{x}-\mathbf{x}_i}{\lambda}\right) \right\rangle \left\langle K\left(\frac{\mathbf{x}-\mathbf{x}_j}{\lambda}\right) \right\rangle \right\} \\
&= \frac{1}{n\lambda^{2d}} \left\{ \int \mathrm{d}^d y \left[K\left(\frac{\mathbf{x}-\mathbf{y}}{\lambda}\right)\right]^2 u(\mathbf{y}) - \left[\int \mathrm{d}^d y\, K\left(\frac{\mathbf{x}-\mathbf{y}}{\lambda}\right) u(\mathbf{y})\right]^2 \right\} \\
&\approx \frac{1}{n\lambda^d} \int \mathrm{d}^d t\, K(\mathbf{t})^2 u(\mathbf{x}).
\end{aligned}
\tag{5.8}
$$

Putting both results together we have

$$
\textsc{AMeISE} = \frac{\lambda^4 k_1^2}{4d^2} \int \mathrm{d}^d x\, \left[\Delta u(\mathbf{x})\right]^2 + \frac{k_2}{n\lambda^d},
\tag{5.9}
$$

where the constants k_1 and k_2 depend on the kernel $K(\mathbf{x})$ through

$$
k_1 = \int \mathrm{d}^d x\, \mathbf{x}^2 K(\mathbf{x}), \qquad k_2 = \int \mathrm{d}^d x\, K(\mathbf{x})^2.
\tag{5.10}
$$

Minimising this expression for the $\textsc{AMeISE}$ with respect to λ, one can in principle obtain an explicit expression for the optimal kernel width λ as given by (Parzen 1962). However, the − unknown − underlying true density u enters into this expression via the integrated squared Laplacian. Although this is somewhat disappointing, this problem can be remedied using the biased cross–validation technique suggested by Scott & Terrell (1987). The idea underlying this approach is to obtain an estimate for the integrated squared Laplacian of the true field from

the kernel estimate. We do this by calculating the expectation value of the integrated squared Laplacian for the kernel estimate. To first non–vanishing order in n^{-1} and λ the result is

$$\left\langle \int \mathrm{d}^d x \, [\Delta \hat{u}_\lambda(\mathbf{x})]^2 \right\rangle \approx \int \mathrm{d}^d x \, [\Delta u(\mathbf{x})]^2 + \frac{1}{n\lambda^{4+d}} \int \mathrm{d}^d x \, [\Delta K(\mathbf{x})]^2 . \qquad (5.11)$$

Thus, we obtain an expression for the AMeISE that can be calculated from the kernel estimate itself. We have

$$\mathrm{AMeISE} = \frac{\lambda^4 k_1^2}{4d^2} \int \mathrm{d}^d x \, [\Delta \hat{u}_\lambda(\mathbf{x})]^2 + \frac{1}{n\lambda^d} \left(k_2 - \frac{k_1^2 k_3}{4d^2} \right) , \qquad (5.12)$$

with another constant

$$k_3 = \int \mathrm{d}^d x \, [\Delta K(\mathbf{x})]^2 . \qquad (5.13)$$

Inserting a Gaussian kernel

$$K(\mathbf{x}) = (2\pi)^{-d/2} \exp\left(-\tfrac{1}{2}\mathbf{x}^2\right) , \qquad (5.14)$$

where the constants k_1, k_2, and k_3 are given by

$$k_1 = d, \quad k_2 = \left(2\sqrt{\pi}\right)^{-d}, \quad k_3 = \frac{d(d+2)}{4} k_2 . \qquad (5.15)$$

the AMeISE for this particular kernel becomes

$$\mathrm{AMeISE} = \frac{1}{4}\lambda^4 \int \mathrm{d}^d x \, [\Delta \hat{u}_\lambda(\mathbf{x})]^2 + \frac{17-(d+1)^2}{2^{d+4}\pi^{d/2}} \frac{1}{n\lambda^d} . \qquad (5.16)$$

With a good starting value, efficient minimisation techniques (Press et al. 1987), and Fast Fourier Transform techniques (Frigo & Johnson 1997) for evaluating integrated squared Laplacians, it becomes computationally feasible to push fixed width smoothing to its optimal value. Figure 5.1 shows an application to two simulations of the dark matter distribution in the Universe, namely the $240h^{-1}\mathrm{Mpc}$ boxes of the ΛCDM and the OCDM models calculated by the Virgo project. The 256^3 particles of a set were first binned onto a 256^3 grid using a cloud–in–cell kernel, and then smoothed with a kernel of fixed smoothing width. The resulting AMeISE is plotted as a function of the smoothing length. Given that the grid cells are roughly $1h^{-1}\mathrm{Mpc}$ in size, for both models the minimum AMeISE is not reached at a sensible smoothing length. The obvious reason for this is the high degree of singularity in the evolved density field. The AMeISE is dominated by peaks in high–density regions, and hence favours a small smoothing length, where areas of intermediate to low density suffer from overwhelming shot noise.

So far, we have shown that fixed width smoothing is a failure as far as the accurate reconstruction of the dark matter density field from simulations of gravitational clustering is concerned. This can be explained by the nature of dark matter density fields, and we expect the method to

Figure 5.1 The AMeISE as a function of the smoothing lengths for a variety of dark matter models. Note that the minimum at roughly $100h^{-1}\mathrm{kpc}$ corresponds to a smoothing length of one tenth of the average inter–particle distance. We conclude that fixed width smoothing provides at most a first guess of a cosmological density field.
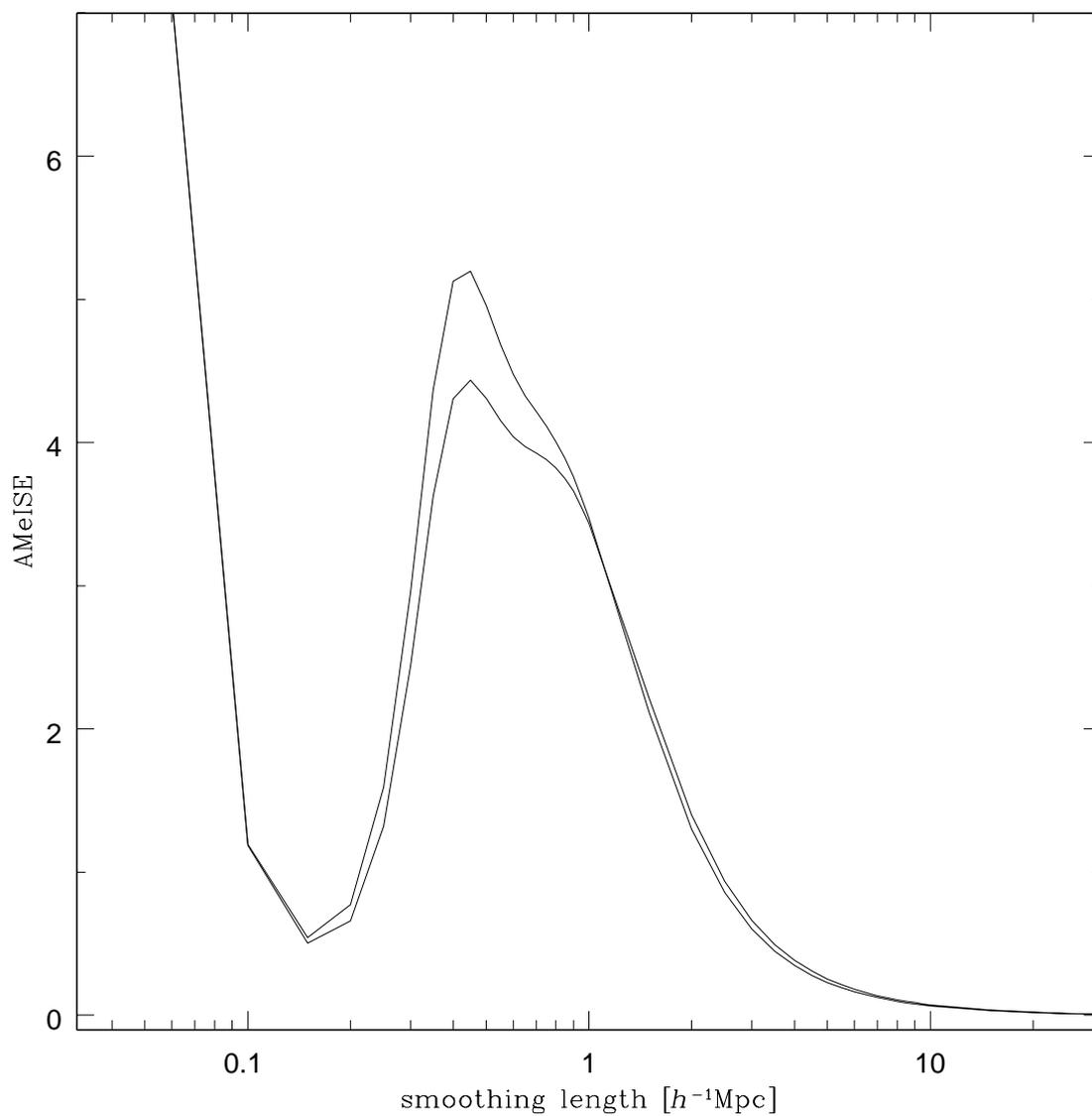
perform better with fields of less roughness. Examples are encountered in the reconstruction of individual clusters, and with sparser point samples, such as catalogues of galaxies and cluster of galaxies.

Still, fields constructed by fixed width smoothing provide a rough orientation of the density field. Hence they can serve as a starting point for the following considerable refinement of the kernel density estimate.

5.1.2   Adaptive smoothing

It appears natural to account for the strong density variations by varying the kernel width in Equation (5.3). Each particle from the data set is assigned an individual smoothing length, and the density estimate becomes

$$\hat{u}(\mathbf{x}) = \frac{1}{n} \sum_{i=1}^{n} \lambda(\mathbf{x}_i)^{-d} K\left(\frac{\mathbf{x} - \mathbf{x}_i}{\lambda(\mathbf{x}_i)}\right). \tag{5.17}$$

Regions of high density are well sampled and should be more accurately reconstructed with a smaller kernel, while low density regions need larger kernel widths to compensate the sparser sampling. Obviously, this approach requires a first estimate of the density, the so–called pilot field u_{pilot}, to guide the actual reconstruction. Since its accuracy is not too critical, fixed width smoothed field with some width $\overline{\lambda}$ is usually sufficient according to Silverman (1986). The local smoothing width $\lambda(\mathbf{x})$ can then be chosen as

$$\lambda(\mathbf{x}) = \overline{\lambda} \left(\frac{u_{\mathsf{pilot}}(\mathbf{x})}{\overline{u_{\mathsf{pilot}}}}\right)^{-\alpha}, \tag{5.18}$$

where $\overline{u_{\mathsf{pilot}}}$ denotes the geometric mean of the pilot field, that is

$$\overline{u_{\mathsf{pilot}}} = \exp\left(\int \mathrm{d}^d x \, \ln u_{\mathsf{pilot}}(\mathbf{x})\right). \tag{5.19}$$

The choice of the sensitivity parameter α is crucial for the outcome of the method and is discussed at length by Breiman et al. (1977) and Abramson (1982). For $\alpha = 0$, fixed width smoothing is recovered. In the following we shall use the value $\alpha = \frac{1}{2}$, which results in smoothing volumes that contain roughly the same number of neighbours regardless of the local density.

At first sight, it appears that adaptive smoothing can only be done in real space. This approach has been used on cosmological simulations by Colberg (1999), for example, and proved woefully inefficient, especially for the large smoothing lengths required to obtain an accurate estimate of the density in voids. We follow an approach that combines high accuracy and high computational efficiency.

Since the squared widths of Gaussian smoothing kernels are additive, m smoothing steps with widths $\Delta\lambda_j$, $j = 1 \ldots m$ are equivalent to a single smoothing with width $\lambda_m = \sqrt{\sum_{j=1}^{m} \Delta\lambda_j^2}$

Table 5.1 Parameters of the four Virgo simulations used in our analysis. The Hubble constant h, the particle mass m_p, and the comoving size L of the simulation box are in units of $H_0 = 100\mathrm{km\,s}^{-1}/\mathrm{Mpc}$, $10^{10}h^{-1}M_\odot$, and $h^{-1}\mathrm{Mpc}$, respectively.

model	Ω_0	Ω_Λ	h	σ_8	m_p	L
ΛCDM	0.3	0.7	0.7	0.90	6.8	240
OCDM	0.3	0.0	0.7	0.85	6.8	240
SCDM	1.0	0.0	0.5	0.51	22.1	240
τCDM	1.0	0.0	0.5	0.51	22.1	240

altogether. The $\Delta\lambda_j$ can be chosen freely, and we obtain $\lambda_m \propto m$ by setting $\Delta\lambda_j \propto \sqrt{2j-1}$. To begin with, we distribute the particles $\mathbf{x}_i$ onto a cubic grid using a cloud–in–cell kernel, which preserves accurate position information. Calculating a density field smoothed with a kernel of fixed width and applying Equation (5.18), we obtain the desired smoothing width for each grid cell. We then perform a series of smoothing steps, starting from an empty grid and adding particles from regions of higher density as more smoothing steps are performed. In the end, particles in high–density regions have undergone only few smoothing steps, equivalent to smoothing with a small kernel, while particles in voids have taken part from the very beginning and hence are effectively smoothed with a large kernel width. If the number of steps are chosen large enough, the discreteness error becomes negligible. Real space smoothing should be used for the last smoothing steps, where the width $\Delta\lambda_j$ falls below the lattice constant, and Fourier space smoothing would be adversely affected by aliasing. In practice, the reconstruction of a density field from 256^3 particles using 40 steps on a 256^3 cell grid consumes not more than fifteen minutes CPU time on an average workstation.

Note that the procedure described relies heavily on the additivity property of the Gaussian smoothing kernel. Although other kernels, most notably the Epanechnikov kernel (Epanechnikov 1969) are better in the sense that fewer points are necessary to obtain some desired kernel bandwidth (Kendall & Stuart 1973), the Gaussian kernel is merely marginally suboptimal. Furthermore, it is the only additive kernel at all (ter Haar Romeny et al. 1991), and is highly suitable both for Fourier space and real space smoothing.

This is enough theory for now, some applications are in order. We will reconstruct density fields from N–body simulations and calculate their Minkowski functionals.

5.2   Minkowski functionals of the Virgo simulations

5.2.1   The Virgo simulations

The first run of the so–called Virgo simulations follows the gravitational evolution of 256^3 particles in $240h^{-1}\mathrm{Mpc}$ boxes. Four different variants of the Cold Dark Matter model are used as initial conditions, with parameters summarized in Table 5.1. A detailed technical description

Figure 5.2 The two panels display the same 256^2 slice through density fields reconstructed from the Virgo project's simulation of a ΛCDM model. While the panel on the left is obtained by fixed width smoothing with a $1.5h^{-1}\mathrm{Mpc}$ kernel (the box itself is $240h^{-1}\mathrm{Mpc}$ long), the panel on the right was calculated by adaptive smoothing with an average width of $1.5h^{-1}\mathrm{Mpc}$.
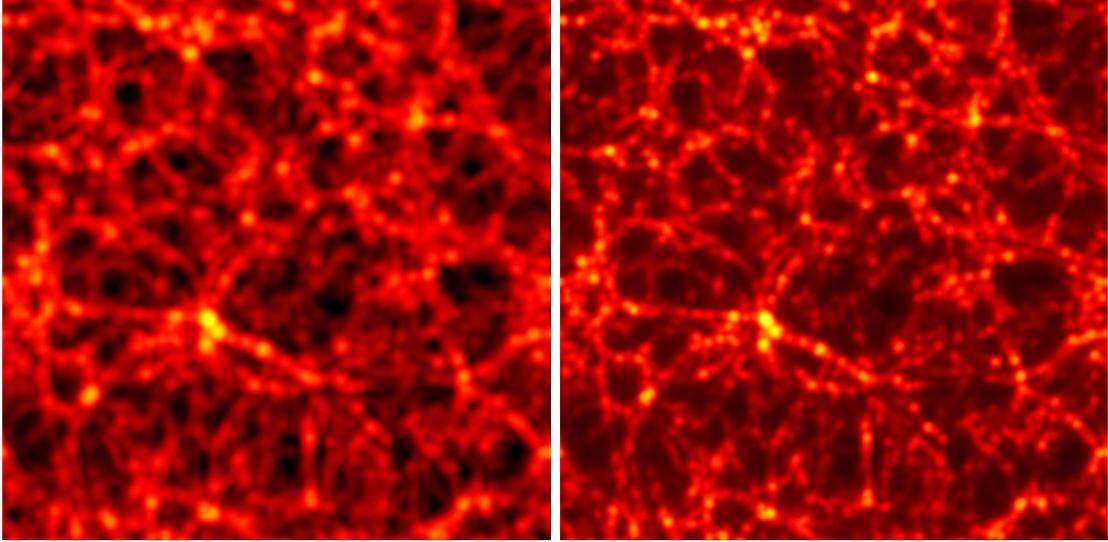
of the simulations, and a correlation and power spectrum analysis can be found in Jenkins et al. (1998). Further studies of large–scale structure use genus statistics (Springel et al. 1998) and cluster abundances (Thomas et al. 1998).

Pictures of the density field reconstructed from the ΛCDM model in Figure 5.2. The two panels were calculated from the same set of particles and with the same smoothing length of $1.5h^{-1}\mathrm{Mpc}$. However, the left panel simply uses fixed width smoothing, while the right panel was obtained through the adaptive smoothing technique described in Section 5.1.2. It can be clearly seen that the adaptive method suppresses shot noise in voids, and enhances high–density structures at the same time.

5.2.2   Minkowski functionals

The Minkowski functionals of these two density fields are compared in Figure 5.3. The functionals are plotted as functions of the rescaled threshold ν, which is obtained from the density contrast δ by numerically inverting the implicit connection

$$v_0(\delta) = \frac{1}{2} - \frac{1}{2}\Phi\left(\frac{\nu}{\sqrt{2}}\right), \tag{5.20}$$

where $v_0(\delta)$ is the measured volume functional at the density threshold δ. Obviously, using ν forces exact Gaussian behaviour of the volume functional v_0. This removes the deviations

Figure 5.3 Minkowski functionals of the density fields shown in Figure 5.2. The results for the adaptive and fixed smoothing techniques are drawn in solid and dashed lines, respectively. A number of drastic differences in the morphology are visible, even though both fields were constructed from the same point set with the same average filter width.
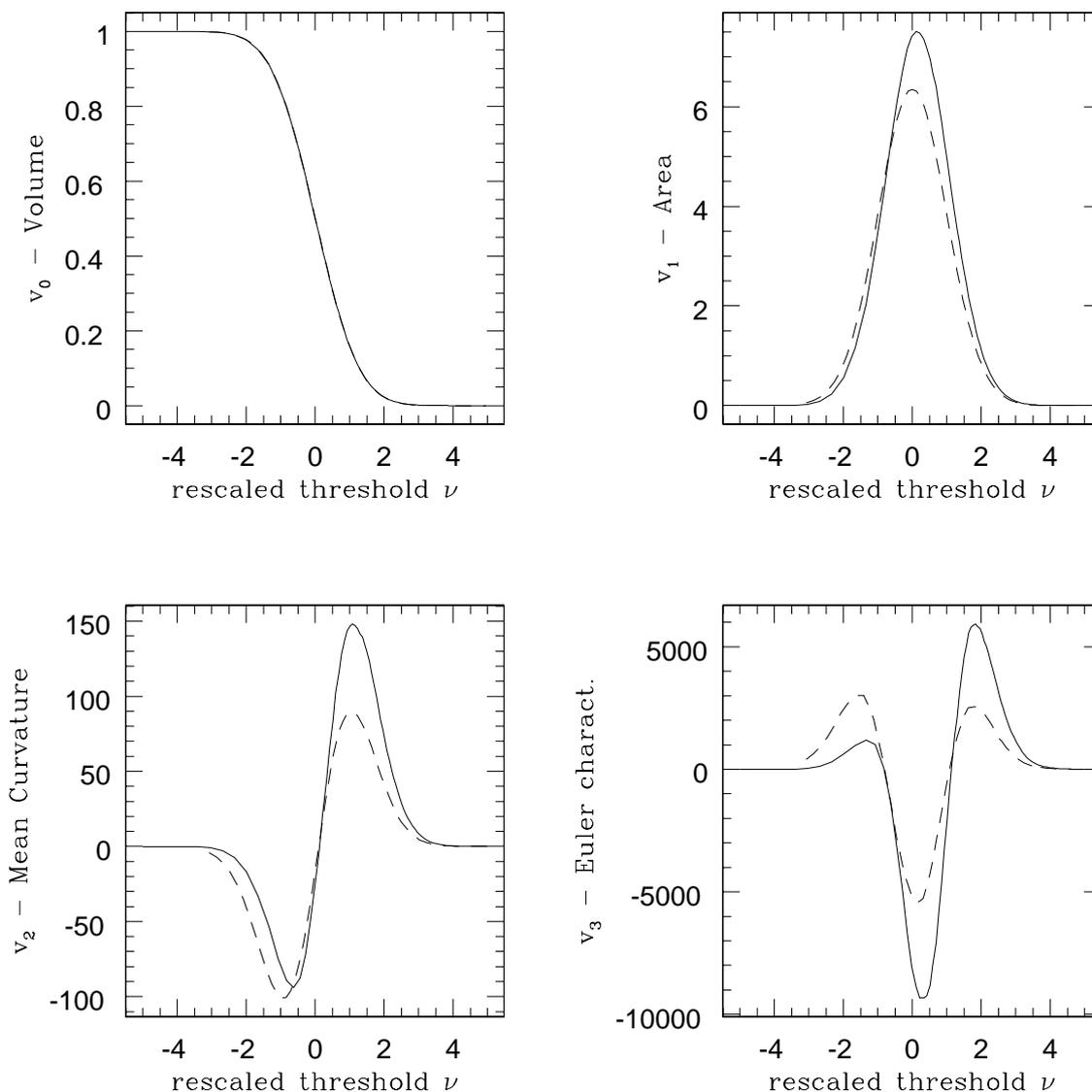
Figure 5.4 Just like Figure 5.3, this plot shows the Minkowski functionals of the density fields shown in Figure 5.2. This time the functionals are plotted against the filling factor v_0 of the global field. For each field, we display both the Minkowski functionals of the global field and the Minkowski functionals of the largest object.
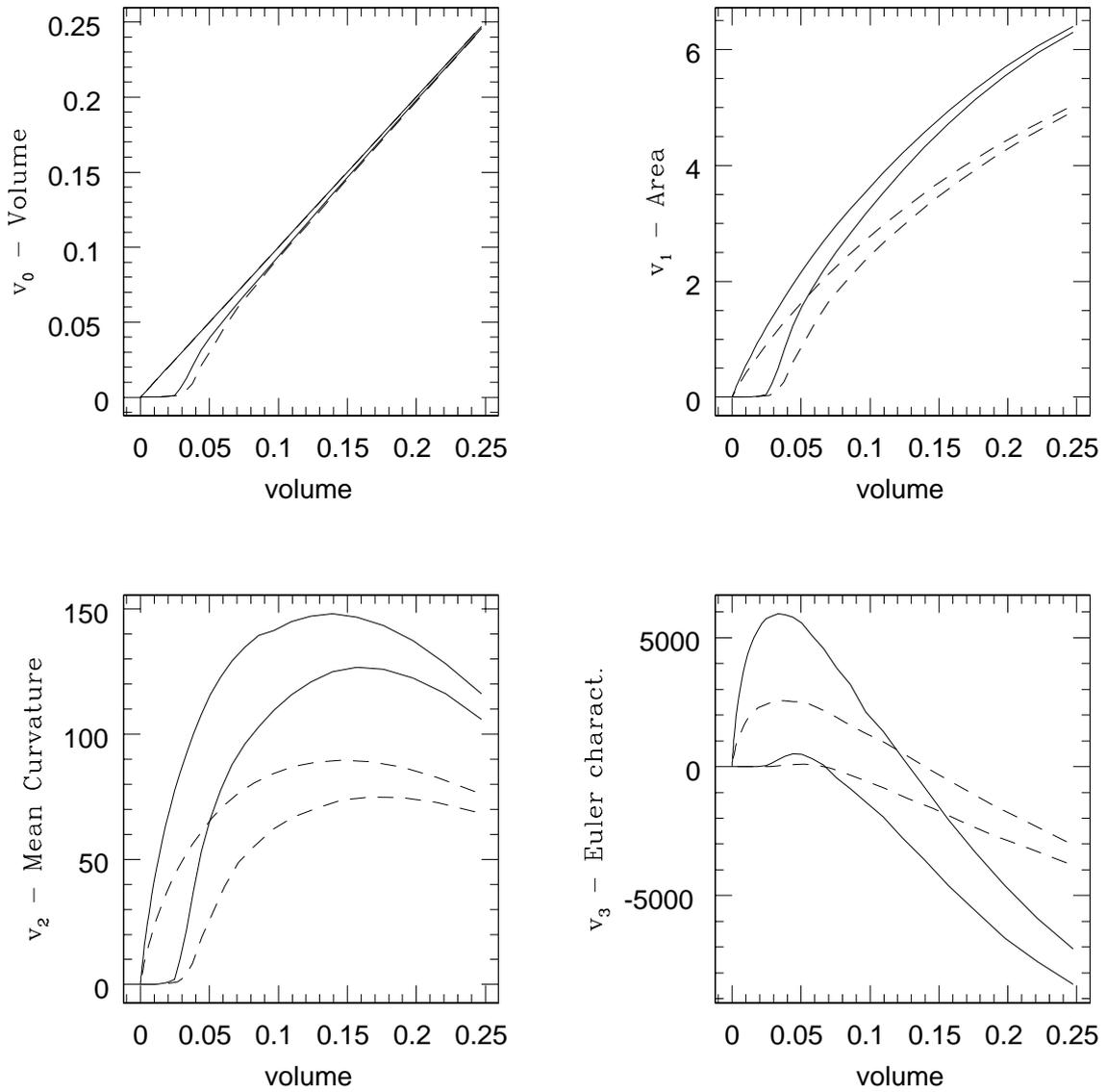
Figure 5.5 Here we compare the four Cold Dark Matter variants used in the calculations by the Virgo project. Shown are ΛCDM (solid), OCDM (dotted), SCDM (short–dashed) and τCDM (long–dashed).
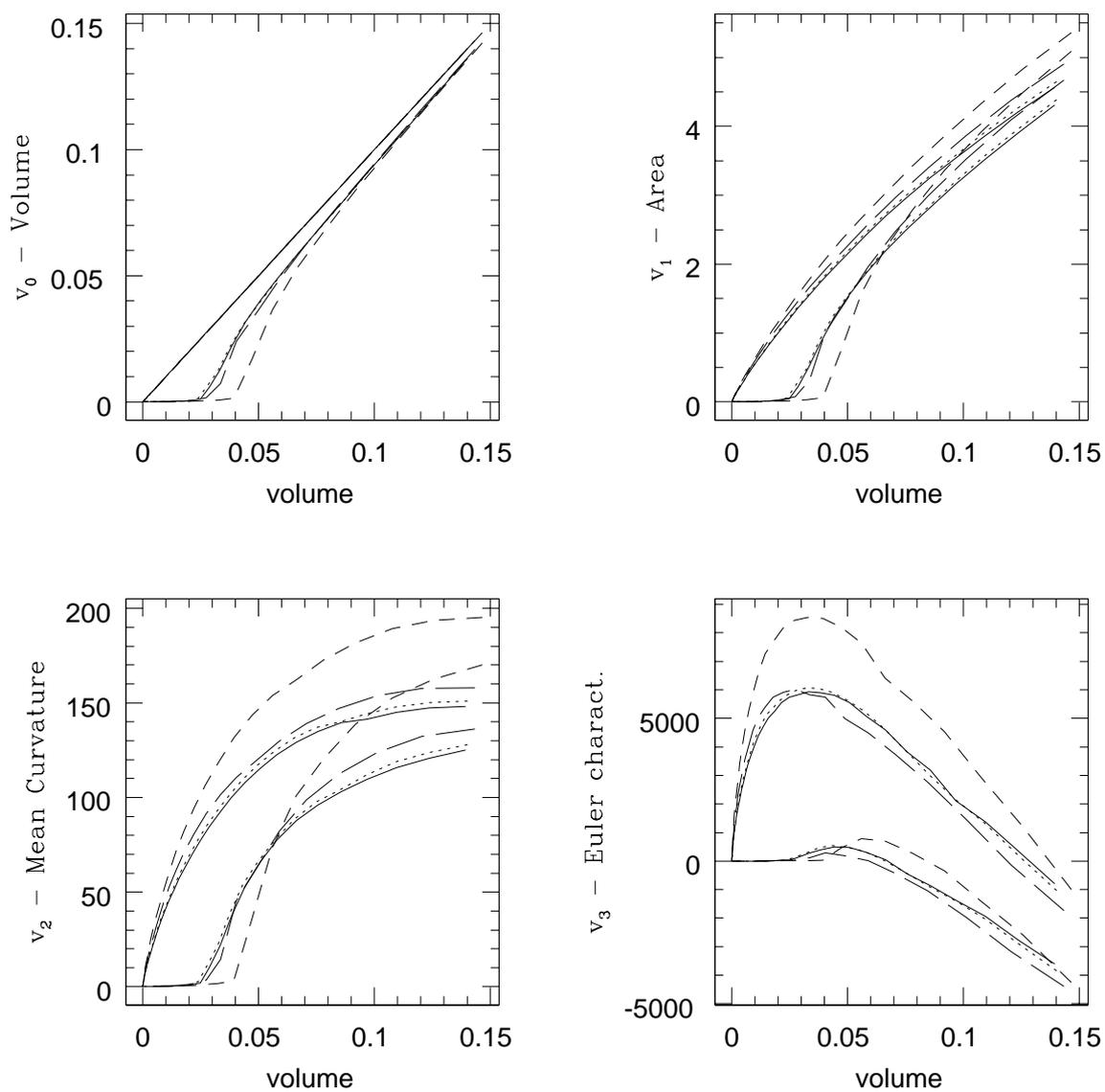
from Gaussianity that are due to changes in the one–point probability distribution function, and emphasises deviations due to higher–order correlations. Note that for the density field constructed with a kernel of fixed width (dashed lines) the difference of the Minkowski functional curves to the expectation value for a Gaussian random field is hardly visible to the naked eye – the curves are more or less symmetrical. The adaptively smoothed field, however, appears strongly non–Gaussian. All curves are slightly shifted towards higher density values. The area functional v_1 is just a little elevated, indicating that the adaptively smoothed field contains more features despite the fact that it underwent the same smoothing on average. The behaviour of the mean curvature v_2 is interesting: On the one hand, the minimum has hardly shifted, while on the other hand, the value at the maximum has almost doubled. The two maxima of the Euler characteristic v_3 follow the same trend. The maximum at negative rescaled threshold, measuring the highest possible number of cavities in the excursion set, is more than five times smaller than the maximum at positive rescaled threshold, which indicates the number of isolated lumps in the system. This provides a quantitative confirmation of the visual impression in Figure 5.2. On the one hand, the adaptive smoothing technique has a tendency to empty voids, thereby increasing their size and decreasing their number. On the other hand, the isolated high–density clusters of dark matter become more pronounced through adaptive smoothing. The much lower minimum of the Euler characteristic translates into the presence of many tunnels in the excursion set at intermediate thresholds, in accord with the observation that the filaments and sheets that line the voids and contain the clusters appear in a sharper and more coherent fashion in the adaptively smoothed field.

Minkowski functionals of the global field and the largest coherent object are shown in Figure 5.4. Only the range of low filling factor v_0, i.e. high threshold is shown; at higher global volume the largest object practically encompasses the whole excursion set. The most striking difference is again the much larger abundance of isolated objects at high densities in the adaptively smoothed field. Apart from that, we can also see that the largest coherent object breaks up at a lower value of the filling factor than in the field constructed by fixed width smoothing. This transition from a networked structure that is almost as large as the whole excursion set to a simply connected small lump also occurs more rapidly; the curves almost display a bend when the Minkowski functionals of the largest object reach a value close to zero. We can conclude that adaptive smoothing also enhances the connectivity of the network underlying the dark matter distribution.

Figure 5.5 compares all four models simulated by the Virgo project at the present epoch. Apart from the Standard Cold Dark Matter model, all variants produce a rather similar morphology. Most notably, both low–density models, the OCDM Universe and the ΛCDM Universe practically coincide. These findings are in agreement with the classical genus analysis of Springel et al. (1998).

5.3   Shapefinders

Apart from the largest coherent object, we can also identify the smaller isolated objects making up the excursion set at a given threshold and individually calculate the Minkowski functionals of each. Since the total excursion set is the disjoint union of all these parts, the global Minkowski functionals are given as sums of these Partial Minkowski functionals (PMFs) at the same threshold. This follows the spirit of Mecke et al. (1994), where partial Minkowski functionals are introduced to measure local morphology for the generalised Boolean grain model.

One task of morphological statistics consists in quantifying strongly non–Gaussian features such as filaments and sheets. Given the four Minkowski functionals, we aim at reducing their morphological information content to two measures of planarity and filamentarity, respectively, as has been done for example with various geometrical quantities (Mo & Buchert 1990), moments of inertia (Babul & Starkman 1992), and cumulants of counts–in–cells (Luo & Vishniac 1995).

5.3.1   Simple examples

Let us first explore the Blaschke diagram introduced in Section 2.4 by considering some simple families of convex bodies in three–dimensional space that can take both filamentary and planar shape. We use bodies with analytically known Minkowski functionals V_j, so Equation (2.45) yields their shapefinders $\mathcal{P}$ and $\mathcal{F}$.

A spheroid with two axes of length r and one axis of length λr has Minkowski functionals

$$V_0 = \frac{4\pi}{3}r^3\lambda, \qquad V_1 = \frac{\pi}{3}r^2\left(1 + f(1/\lambda)\right), \qquad V_2 = \frac{2}{3}r\left(\lambda + f(\lambda)\right), \qquad V_3 = 1, \quad (5.21)$$

where[1]

$$f(x) = \frac{\arccos x}{\sqrt{1 - x^2}}. \tag{5.22}$$

By varying the parameter λ from zero to infinity, we can change the morphology of the spheroid from a filament to a pancake via a spherical cluster.

A different way of deforming a filament into a pancake goes via generic triaxial ellipsoids; thereby one of the smaller axes of a strongly prolate spheroid is increased in size until it matches the larger axis and an oblate spheroid has been reached. An integral expression can be found in (Sahni et al. 1998).

Yet another transition from prolate to oblate shape is provided by cylinders of radius r and height λr. Here, the Minkowski functionals are given by

$$V_0 = \pi r^3\lambda, \qquad V_1 = \frac{\pi}{3}r^2(1 + \lambda), \qquad V_2 = \frac{1}{3}r(\pi + \lambda), \qquad V_3 = 1. \tag{5.23}$$

---

[1]Note that for arguments $x > 1$, one can use the relation $i\arccos x = \ln\left(x + \sqrt{x^2 - 1}\right)$ to recover an explicitly real–valued expression.

A Blaschke diagram, that is a plot of planarity $\mathcal{P}$ and filamentarity $\mathcal{F}$, summarising the morphological properties of these simple convex bodies is shown in Figure 5.6.

5.3.2 A set of power–law spectrum simulations

In the following, we shall analyse a family of initial power law spectra $P(k) \propto k^n$ with $n \in \{-2, -1, 0, +1\}$ set before a flat Friedmann background with vanishing cosmological constant (Ω=1, Λ=0). We conduct numerical experiments using a PM code (consult Melott & Shandarin 1993 for details). Four sets of phases were used for each model, making a total of 16 simulation runs. Each run consists of 128^3 particles sampled at an epoch well in the non–linear regime. This epoch is chosen such that the scale of non–linearity k_{nl}, defined in terms of the evolved spectrum

$$\sigma_{k_{\mathsf{nl}}}^2 = \int_0^{k_{\mathsf{nl}}} \mathrm{d}^3 k P(k) = 1, \tag{5.24}$$

is equal to eight in units of the fundamental mode of the simulation box. By using the stage $k_{\mathsf{nl}} = 8$, we make sure that structure is already sufficiently developed on scales much larger than the simulation's resolution, while it is not yet influenced by boundary effects.

Using a cloud–in–cell kernel, these particles were put onto a 256^3 grid, which is the maximum value an ordinary workstation can tackle with acceptable time and memory consumption. Subsequently, the density field was smoothed with a Gaussian kernel $\propto \exp(-x^2/2\lambda^2)$, where x is the distance in mesh units and the width λ is set to 3. Tests have shown that this value both leads to a reasonably smooth field, and preserves at least some detail on smaller scales. Finally, we re–scale the density to the density contrast δ, ranging from -1 to infinity with zero mean.

5.3.3 The global field

The global Minkowski functionals calculated from the density fields described in the previous section are shown in Figure 5.7. The four different line styles correspond to the different spectral indices. Figure 5.8 shows the Minkowski functionals for the same set of models, but instead of the density threshold δ, the rescaled threshold ν is used as the x–axis. ν is calculated from the volume Minkowski functional v_0, that is the filling factor f, as described by Gott III et al. (1987). Thus the deviations from Gaussianity that are due to changes in the one–point probability distribution function are removed, and deviations due to higher–order correlations are emphasised.

Obviously, the global Minkowski functionals clearly discriminate between the various models (Figures 5.7 and 5.8). Only the $n = -2$ model shows slight drawbacks as far as robustness is concerned, but that is due to the method's sensitivity to large–scale features of the smoothed density field. The total area, mean curvature, and Euler characteristic are sensitive to abundances of the structures and are easy to interpret: the more power on small scales (greater n) the more abundant structures and therefore the greater the amplitude of the curve.

5.3.4   The largest objects

At intermediate thresholds, the excursion sets consist of numerous isolated objects. We identify them by grouping adjacent occupied grid cells into one object, where adjacent means that the cells have a common face. Since the Minkowski functionals of the global field are calculated by integrating over quantities that can be assigned to individual grid cells, the partial Minkowski functionals of each object can be obtained at no extra cost once the cells belonging to each object have been identified.

Several plots in Figures 5.9, 5.10, 5.11, and 5.12 illustrate the behavior of the Minkowski functionals of these objects.

Obviously, the contribution of smaller objects to the volume is almost negligible compared to the largest one. Note that in all figures, the mean and standard deviation over all four realisations are shown instead of the individual curves. It is worth noting that the variance is largest in the $n = -2$ and -1 models, which are dominated by structures on large scales and hence show the strongest sample variance.

The models with various initial spectral indices n show qualitatively a similar behaviour. At small filling factors (high density thresholds) the two largest clusters give negligible contribution to each of the global characteristics. Then, at the percolation transition the largest cluster quickly becomes the dominant structure in terms of volume, area, and integrated mean curvature. The second largest cluster also grows at the percolation threshold but just a little and then quickly diminishes. The percolation transition is clearly marked in all three characteristics of the largest cluster by their sudden growth. However, this transition does not happen at a well–defined threshold. Instead, clusters gradually merge into the largest objects as the threshold is decreased (the filling factor grows). This continuous transition has also been observed using percolation analysis, i.e. the zeroth Minkowski functional alone (Shandarin & Zel'dovich 1983; Shandarin & Yess 1998; Klypin & Shandarin 1993; Sahni et al. 1997) and is explained by the finite size of the sample.

Nevertheless, the percolation transitions happen within fairly narrow ranges of the filling factor that are clearly distinct for different models in question: The filling factors are approximately $0.03 \pm 0.01, 0.07 \pm 0.015, 0.11 \pm 0.015, 0.14 \pm 0.015$ in the $n = -2, -1, 0, +1$ models, respectively. It is remarkable that all percolation transitions occur at smaller filling factors than in Gaussian fields (about 0.16) indicating that even in the most hierarchical model ($n = +1$) the structures tend to be more connected than in the "structureless" Gaussian field. Pauls & Melott (1995) showed positive correlation with networks based on the same phases all the way to $n = +3$. This provides further evidence to the conjecture of Yess & Shandarin (1996): The universality of the network structures results from the evolution of Gaussian initial conditions through gravitational instability. The Euler characteristic of the largest cluster also marks the percolation threshold but in a different manner: before percolation it is zero and after percolation it becomes negative in every model, however, in the $n = 0$ and $n = +1$ models it grows to a small positive peak before becoming negative.

While none of the global functionals have particular features at the percolation threshold, valu-

able information is obtained from looking at the Minkowski functionals of the largest coherent object at each threshold (Figures 5.9, 5.10, 5.11, and 5.12). All four Minkowski functionals of the largest cluster clearly consistently detect the percolation transition. Two points are worth stressing: Firstly, in all models the percolation transition happens at smaller filling factors than in the "structureless" Gaussian fields. Secondly, the more power on the large scales (i.e. the smaller n) the smaller the filling factor at percolation. Both conclusions confirm the results of Yess & Shandarin (1996) about the universality of the network structures in the power law models with $n \leq 1$. The results of Pauls & Melott (1995) present evidence that this should be expected all the way up to $n = +3$; at $n = +4$ mode coupling effects from smaller scales should begin to fully disrupt the network structure.

5.3.5   Small objects

As an example, the Blaschke diagram for the model $n = -2$ is shown in Figure 5.13. The distributions for the other models look qualitatively very similar and the average quantities for other models are presented in Figures 5.14 and 5.15. Figure 5.14 shows that most of the small objects are either spherical or slightly planar. There is also a considerable number of elongated clusters with filamentarities from 0.1 to 0.5. In some cases filamentarity reaches large values ~ 1. On the contrary planarity is much weaker, it hardly reaches the value of 0.2 (which is partly a consequence of the smoothing). There is a hint of a small correlation between filamentarity and planarity: the objects with the largest filamentarity also tend to have larger planarity.

Figures 5.14 and 5.15 display the shapefinders as functions of the cluster mass. The curves give averages over the realisations of each model. Apparently, the signal for filamentarity is much stronger than for planarity, regardless of the model which is in full agreement with Figure 5.13. The planarity and filamentarity distributions qualitatively look very similar except for the amplitude. Small objects ($5 \times 10^{-6} \lesssim m \lesssim 5 \times 10^{-4}$) display stronger planarity and filamentarity for models with more power on large scales. However, for greater masses ($m \gtrsim 5 \times 10^{-4}$) the situation is reversed: the less large-scale power the greater the filamentarity and planarity. If the former seems to be natural and was expected, the latter has been unexpected. Both the planarity and filamentarity monotonically grow and reach their maxima at largest clusters: $\mathcal{P}_m \approx 0.1$ and $\mathcal{P}_m \approx 0.5$ in all models. As expected the largest objects possess the largest planarities and filamentarities, but the independence of the maxima from the model again was unexpected.

Despite the successful application of Minkowski functional analysis to the global field and the largest coherent object, the shapefinder analysis of small objects has led to ambiguous results. Their abundance discriminates well, but is already determined by the difference in the Euler characteristic as well as by the total area and mean curvature of the whole contour (Figures 5.7 and 5.8) that are also sensitive to the abundance of structures. The morphology of small objects as measured by shapefinders shows little differences between models (Figures 5.14, 5.15). Both the maximum average planarity ($\mathcal{P} \approx 0.1$) and the maximum average filamentarity ($\mathcal{F} \approx 0.5$) are reached in the most massive non-percolating objects. None of the models showed ribbon–

like objects characterised by both large planarity and large filamentarity. We may speculate that the smaller objects are ones which formed earlier, are more nonlinear, and therefore more decoupled from initial conditions.

Blaschke diagrams are virtually insensitive to the wide variety of initial power spectra used in this section. Replacing fixed width kernel presented here by adaptive smoothing techniques hardly alters these results. However, all models use Gaussian initial conditions, and evolve under the influence of gravity. Hence the similar morphology of the clumps may point towards universal behaviour. Note that different initial conditions resulting e.g. from string wakes might produce different results.

The grouping and measurement techniques used in this study are less accurate for small objects than for large clusters. It is worth trying to study the morphology of small objects by applying more accurate methods of measuring partial Minkowski functionals such as the interpolation method of Novikov et al. (1999) generalised to three dimensions.

Since partial Minkowski functionals offer the possibility of probing the morphology of individual objects, or the object's environmental morphological properties, we expect that this concept will be more powerful when applied to continuous fields at high spatial resolution so that the details of structures are not smoothed out. Their application to point processes also delivers more direct information. In the next chapter, we will follow an approach that defines partial Minkowski functionals using the Boolean grain model.

Figure 5.6 The shapefinders of some convex bodies. The solid line shows the transition from a ball to a filament to a pancake and back to a ball via triaxial ellipsoids of various shapes. The dashed line corresponds to cylinders that also undergo the transition from a filament to a pancake by varying their height. The dots on the spheroid and cylinder curves indicate typical values of the λ parameter from Equations (5.21) and (5.23), respectively.
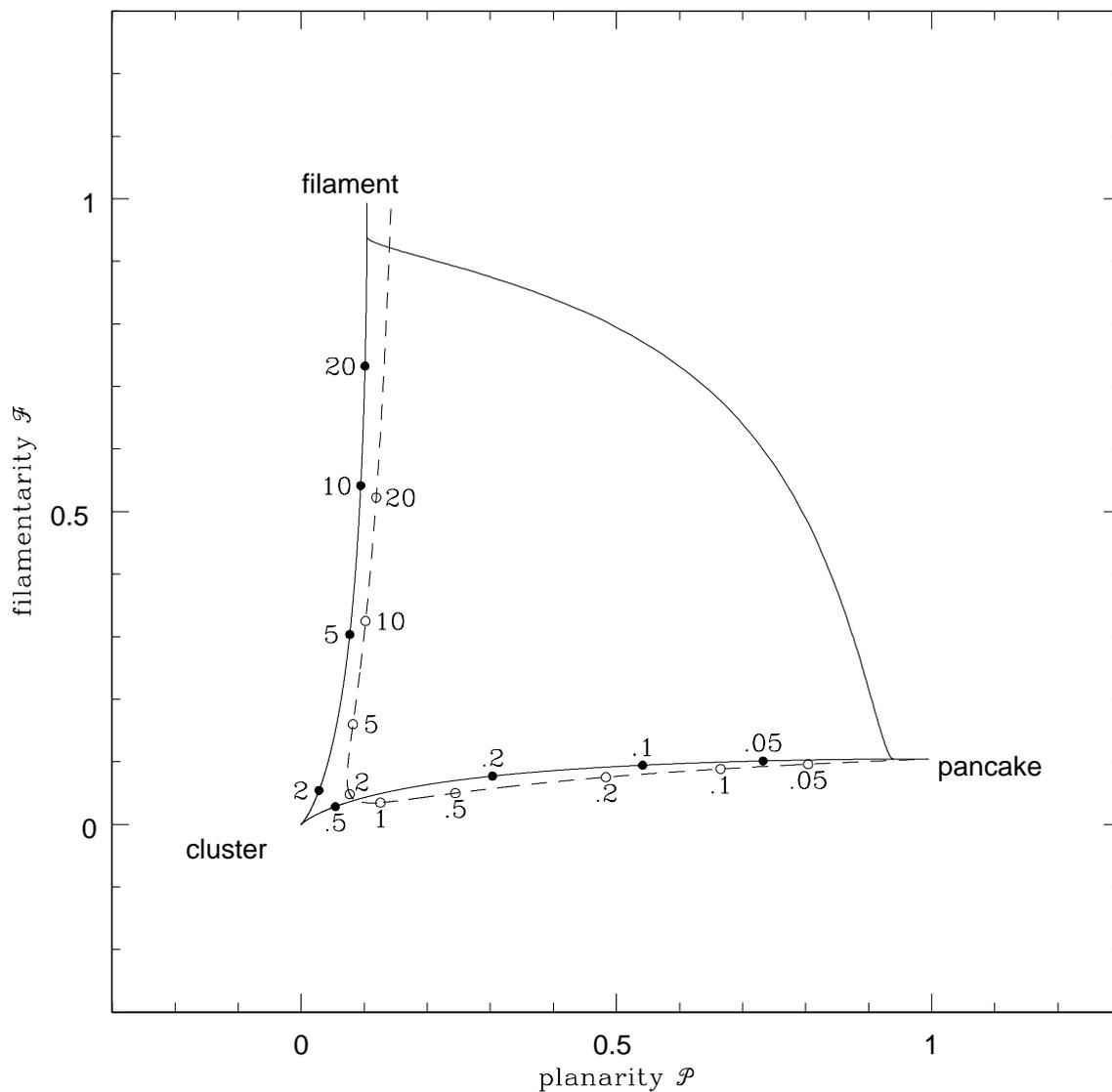
Figure 5.7 Minkowski functionals for the evolutionary stage $k_\mathrm{nl} = 8$ of four different models with initial power law spectra $P(k) \propto k^n$. The area between two lines gives the one–sigma deviations of four realisations from their mean. The solid lines correspond to $n = -2$, dotted is $n = -1$, dash–dotted lines give the results for $n = 0$, and $n = +1$ is shown with dashed lines.
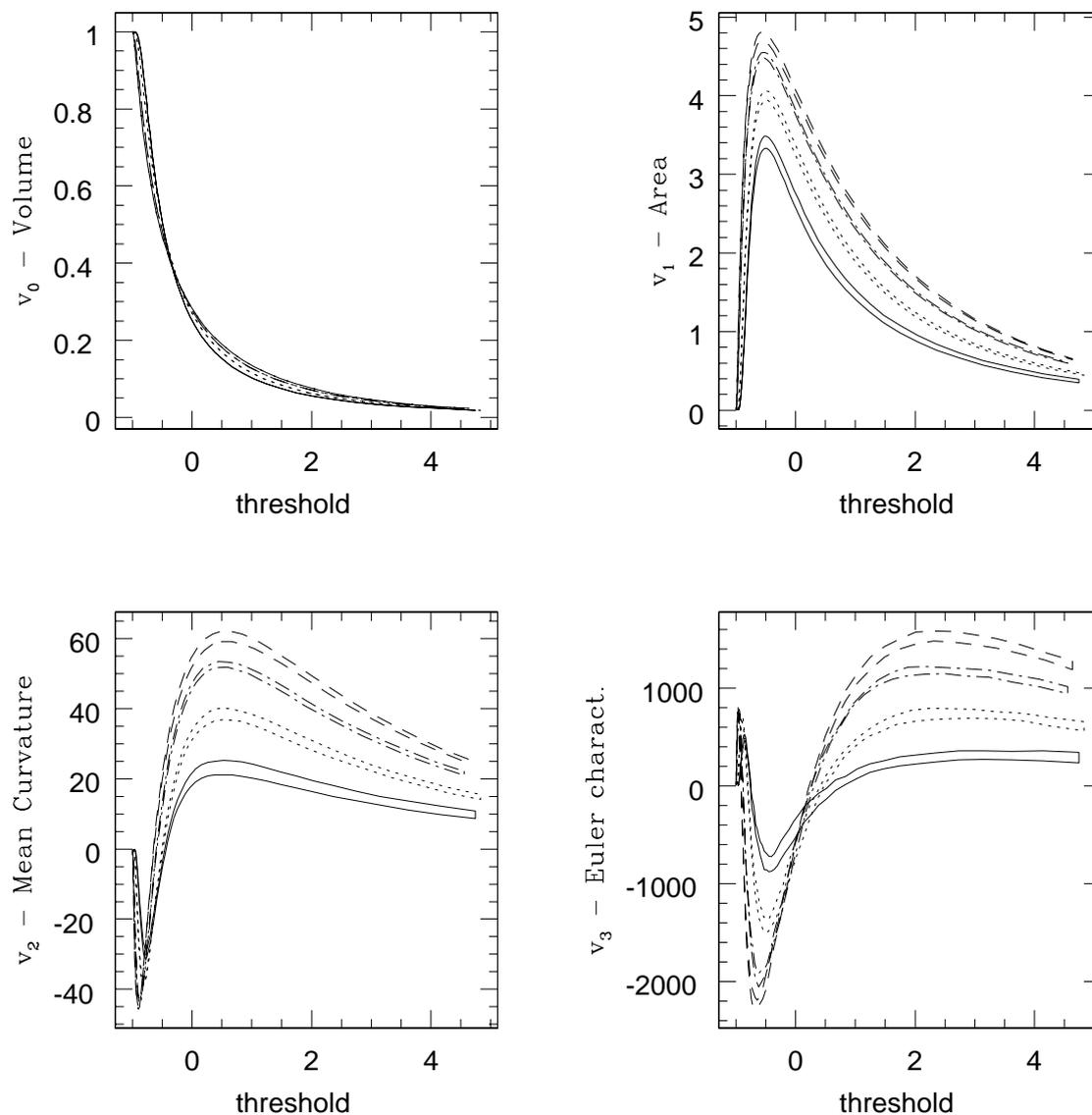
Figure 5.8 Minkowski functionals for the evolutionary stage $k_{nl} = 8$ of four different models with initial power law spectra. Line styles are the same as in the previous figure, but instead of the density threshold δ, the rescaled threshold ν introduced by Gott III et al. (1987) is used.
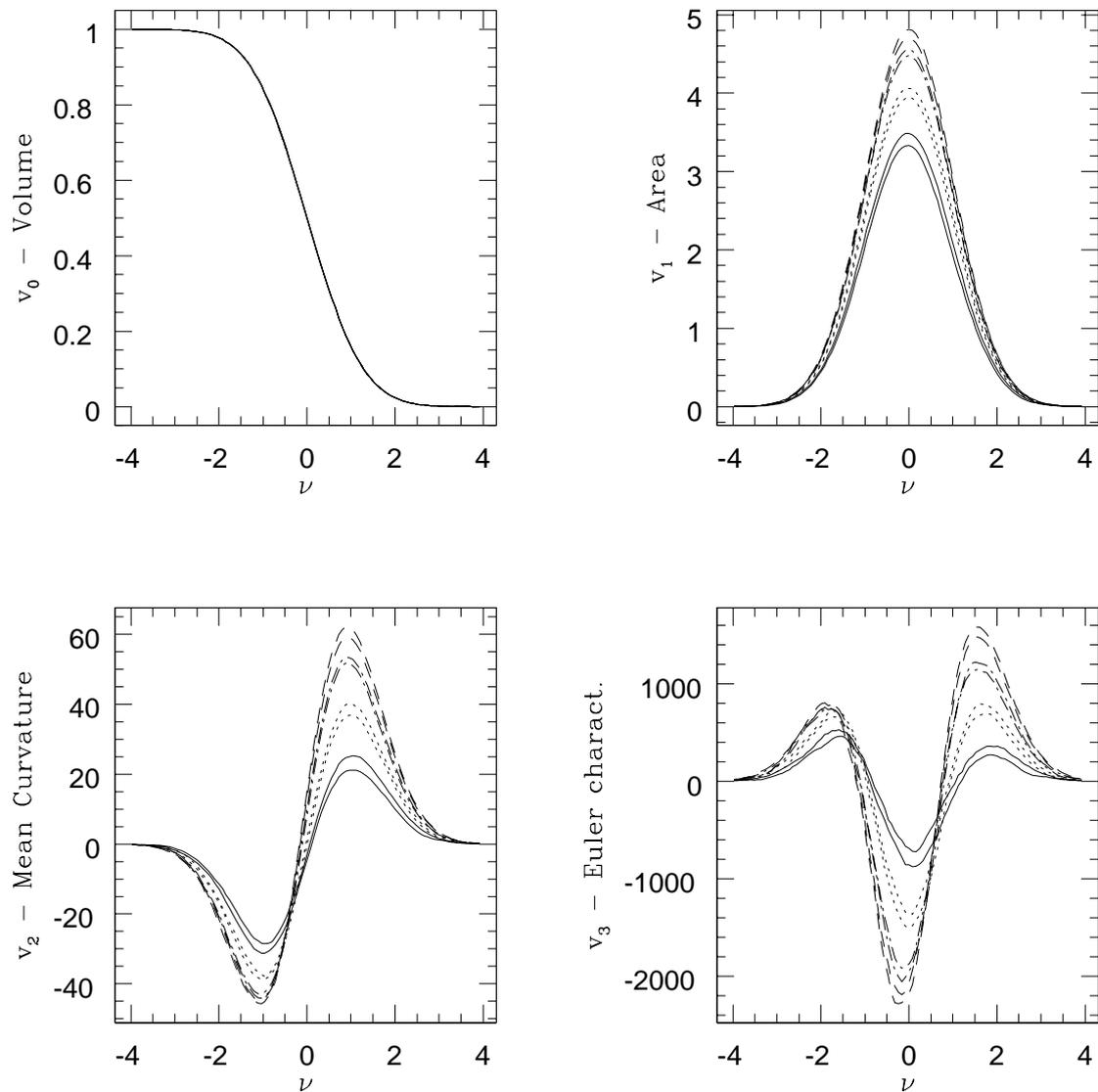
Figure 5.9 The following set of figures compares the global Minkowski functionals of the whole density field (solid line) to the Minkowski functionals of the largest (dashed) and second largest (dotted) objects. The areas indicate the mean and standard deviation over the four realisations of each model. In order to emphasise the region around the percolation transition, the Minkowski functionals are plotted against the filling factor, which is equal to the volume enclosed in all isocontours of the global field. In this particular plot, the values for an initially power–law spectrum with an index of $n = -2$ are shown.
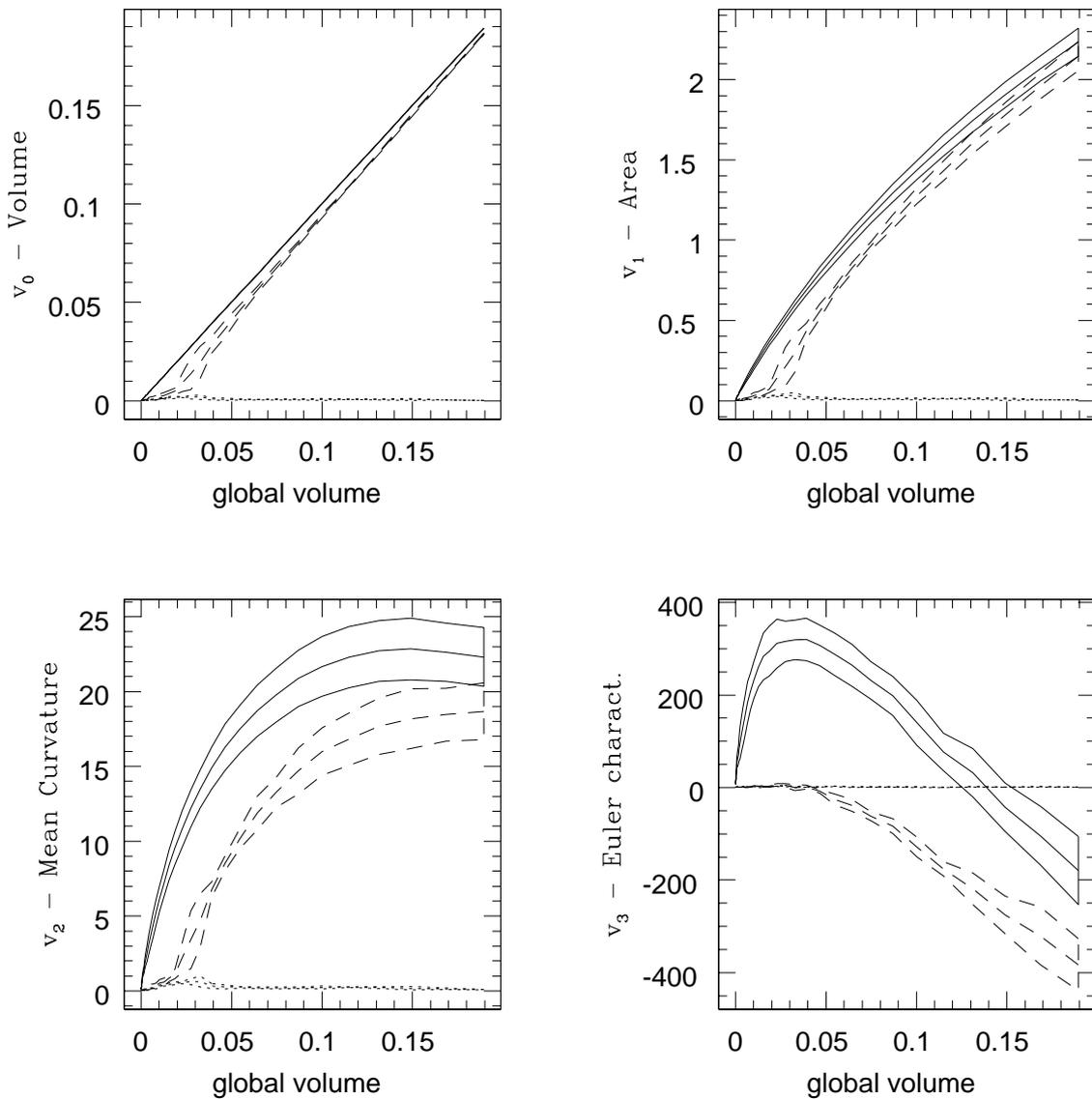
Figure 5.10 The same tendency as before is visible in this case where $n = -1$. As before, the dominance of large networked structures is clearly visible – for example, the Euler characteristic stays well below zero even for high thresholds, which indicates a spongy structure with many tunnels in it. However, the networked structures have fewer loops, since they tend to split up at lower thresholds.
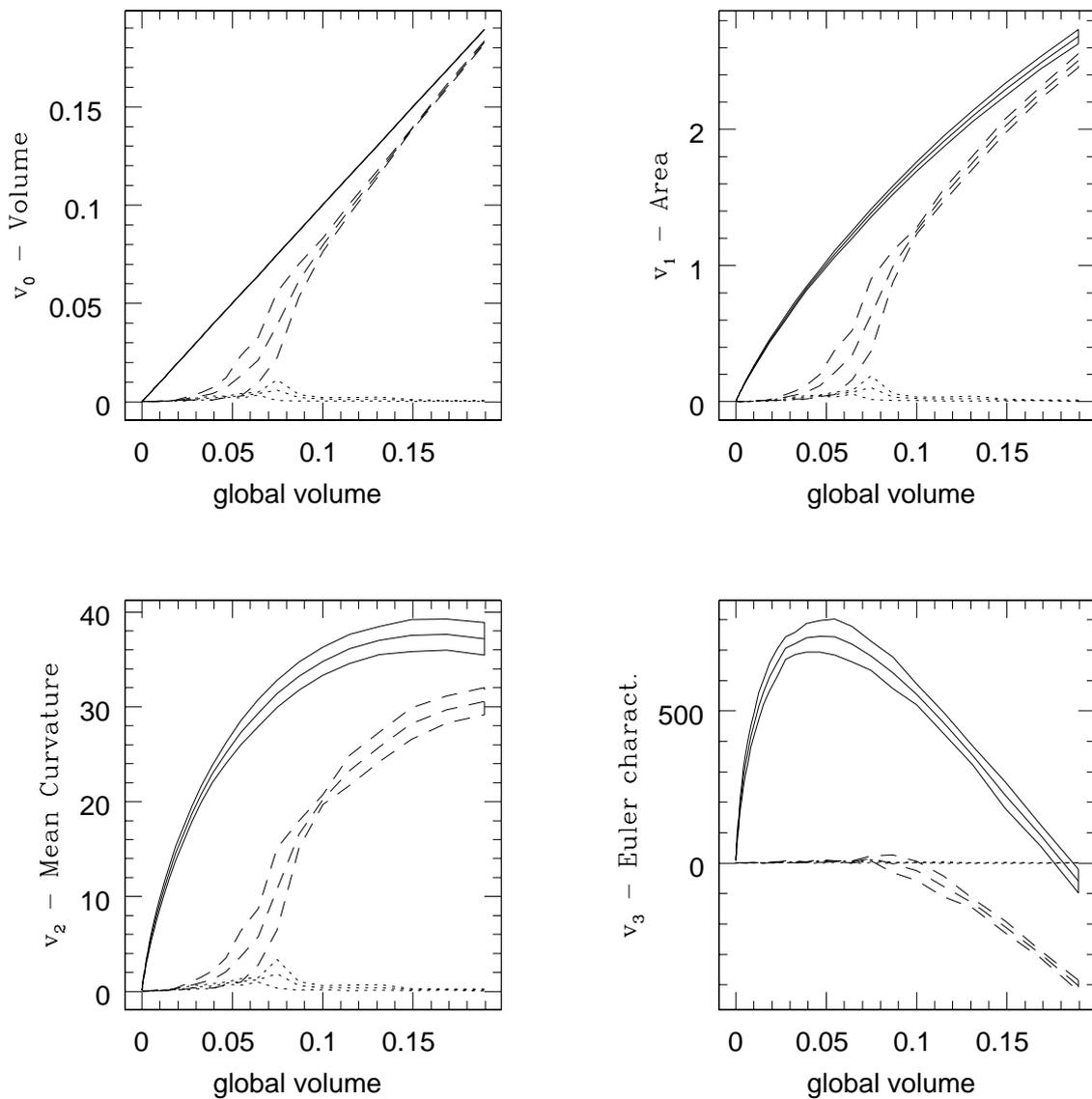
Figure 5.11 The situation is different in the case $n = 0$. The largest object splits up rapidly, and is already simply connected at a filling factor correspongding to the fairly low density contrast $\delta \approx 0.5$. Also, at thresholds around percolation, the second largest object becomes comparable to the largest one, so the field is not dominated by a single networked structure, but by several chunks.
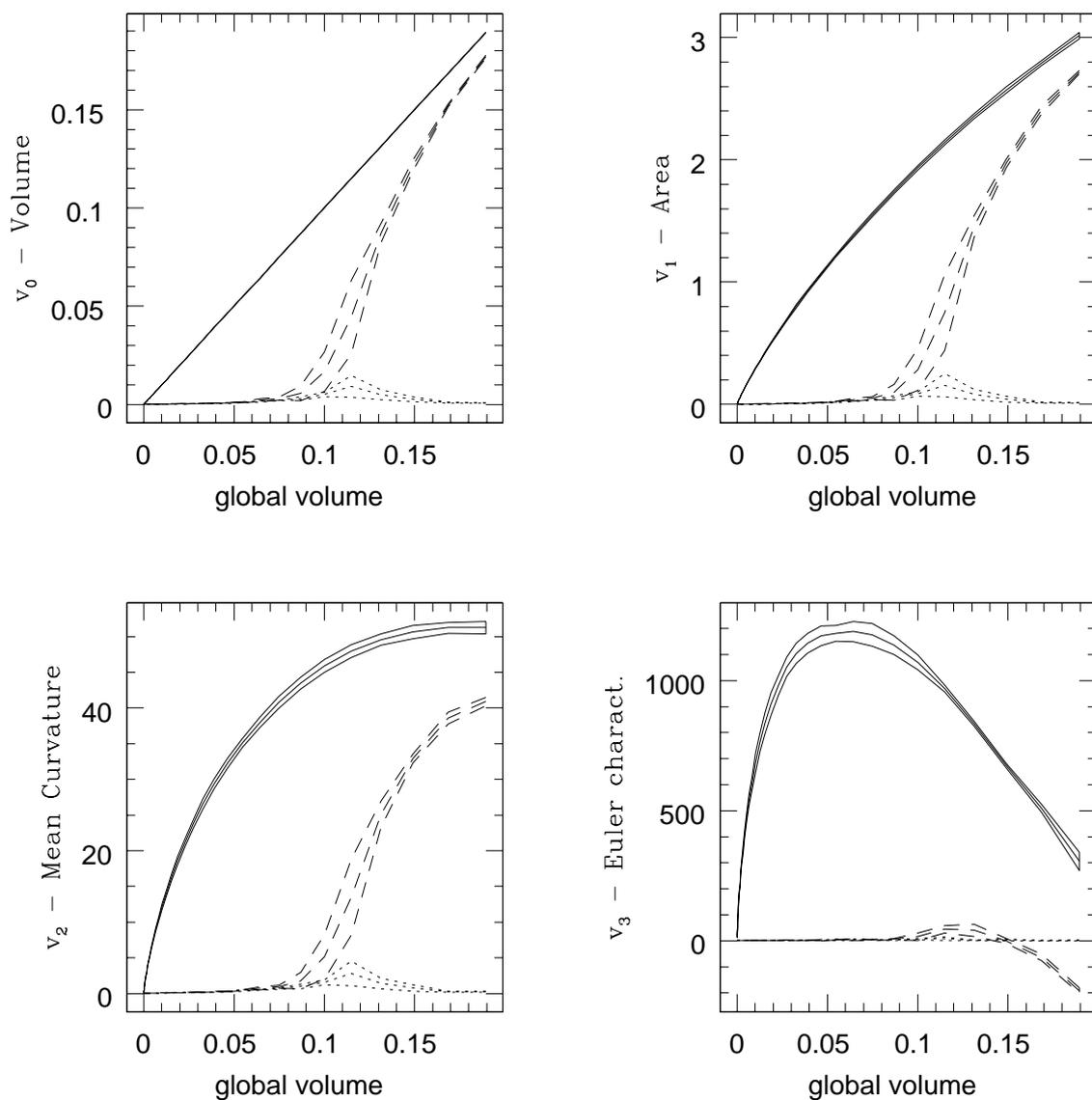
Figure 5.12 The effect described in the previous figure is even more pronounced for $n = +1$, since this model is dominated by small–scale structures. Contrary to the models with negative index n, where the Euler characteristic of the largest object approaches the value one for a simply connected structure from below, one can even see a small positive peak around the percolation threshold.
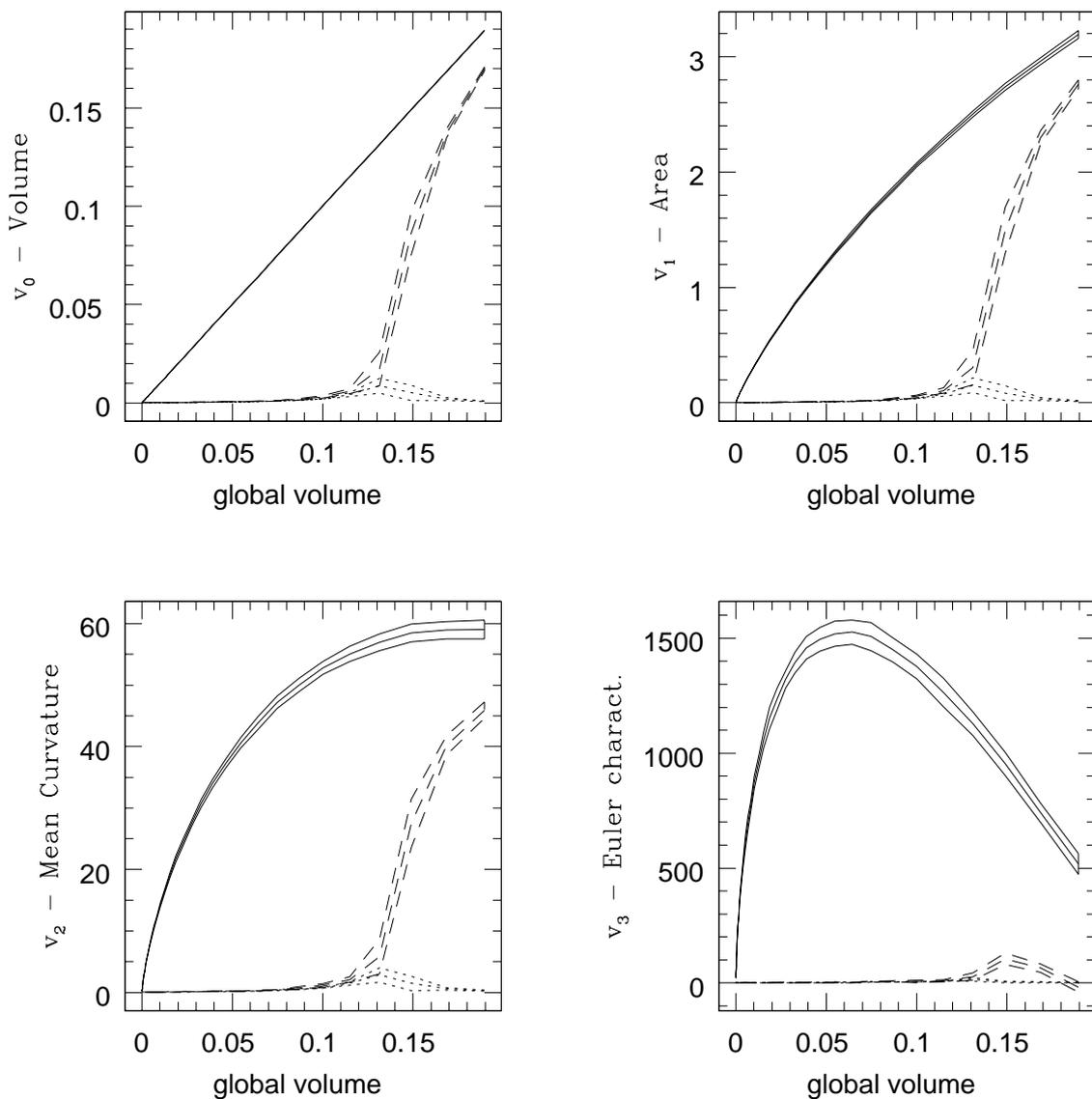
Figure 5.13 The shapefinder or Blaschke diagram of the model with $n = -2$. The area of each dot corresponds to the number of coherent objects whose shapefinders $(\mathcal{P}, \mathcal{F})$ lie around the centre of this dot. We only include objects larger than 10^{-4} times the simulation box, a value several times larger than the smoothing volume. Note that the shapefinders of all objects at all thresholds are used in the construction of the diagram. Nevertheless, most information comes from thresholds close to percolation, where small objects are abundant.
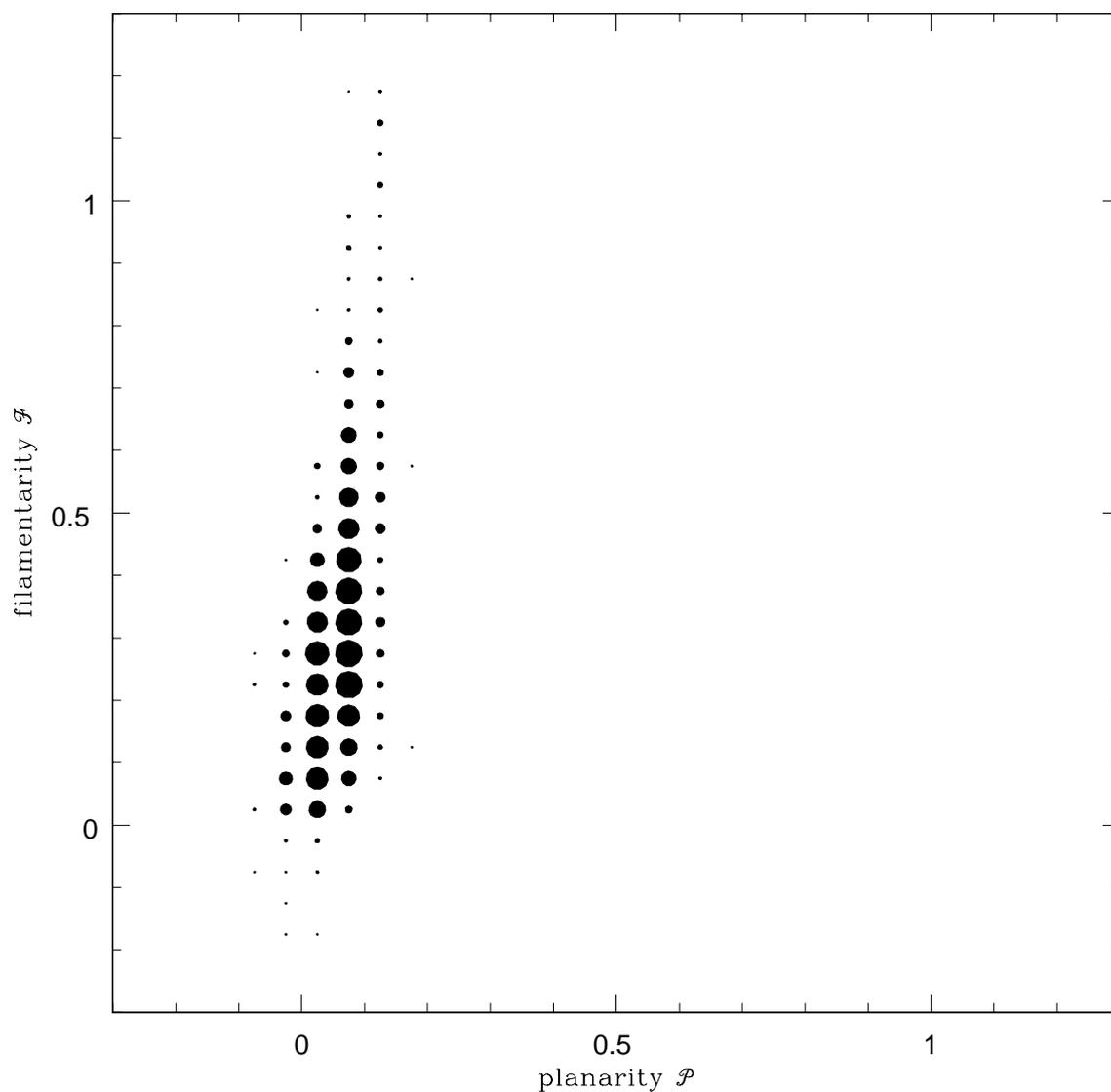
Figure 5.14 The average planarity shapefinder $\mathcal{P}$ as a function of the cluster mass. Line styles are explained in Figure 5.7. As in Figure 5.13, all objects at all thresholds are used to improve statistics.
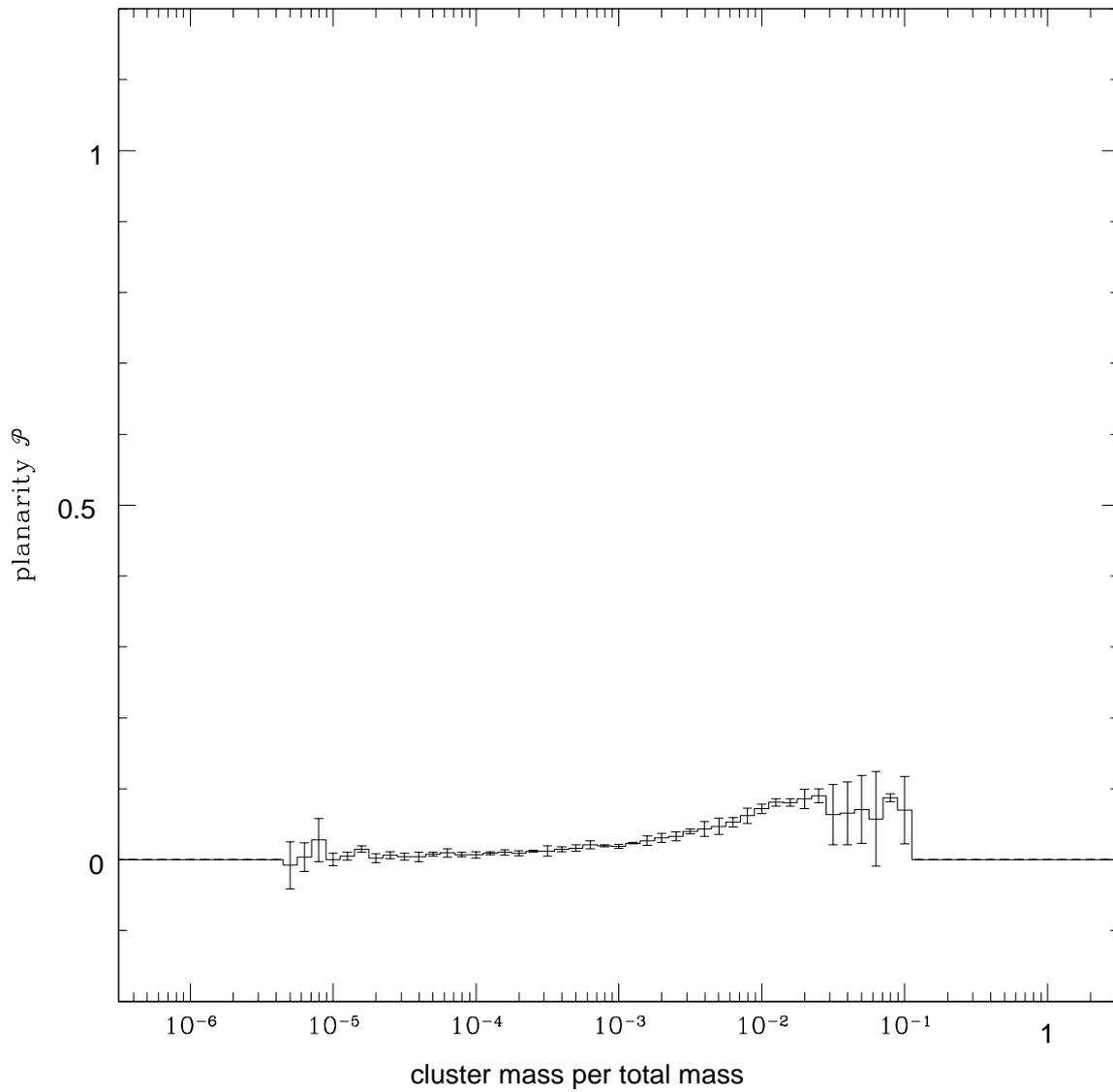
Figure 5.15 This plot shows the same type of distribution as the previous Figure 5.14, but this time the average filamentarity shapefinder $\mathcal{F}$ is plotted as a function of the cluster mass. Again, line styles correspond to models as explained in the caption of Figure 5.7.
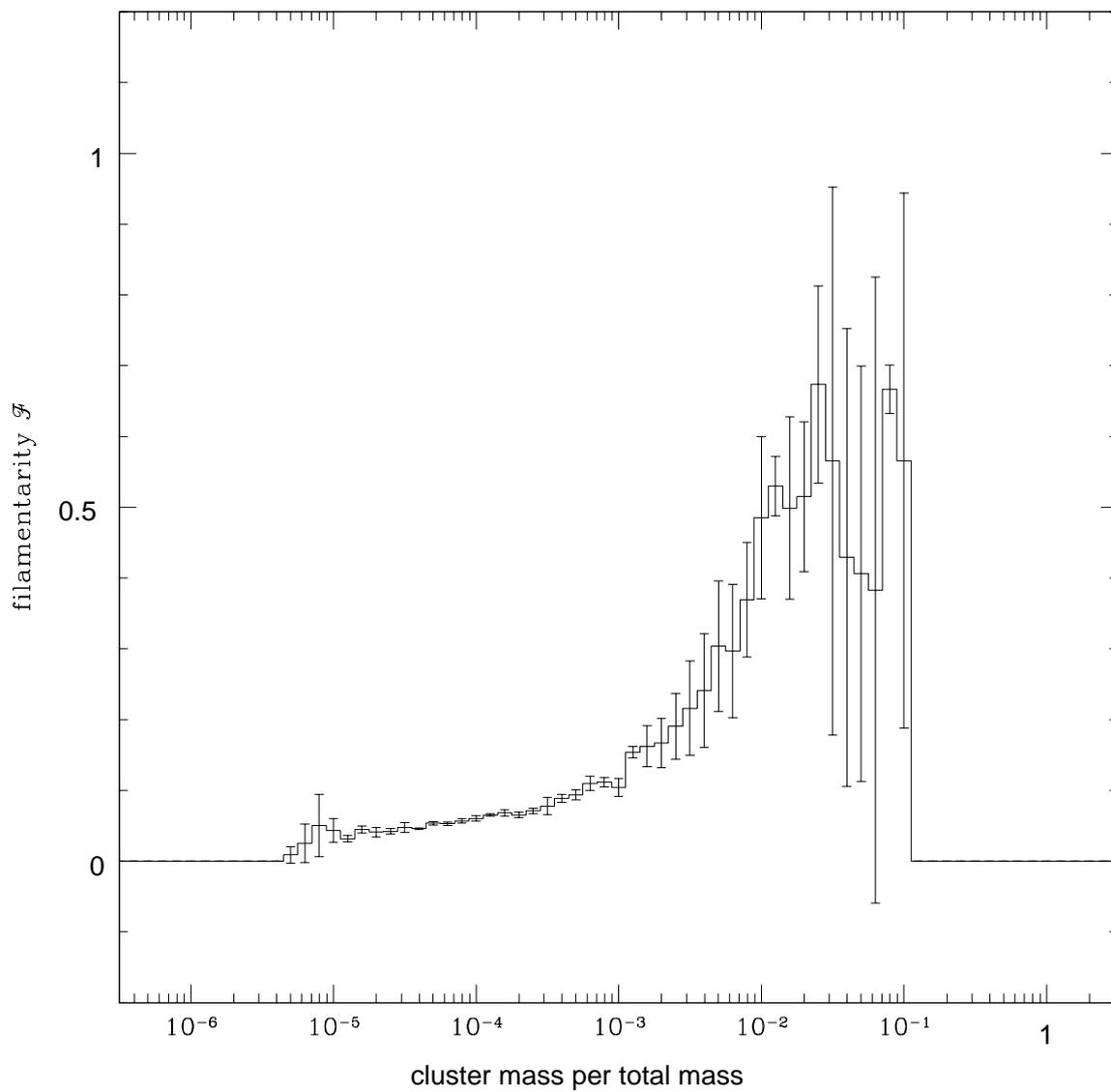
6 Structure in galaxy catalogues

Dis donc, tonton, demande Zazie,
quand tu déconnes comme ça,
tu le fais esprès ou c'est sans le vouloir?
(Queneau 1959)

In contrast to the previous chapter, where our analysis of the large–scale structure of the Universe focussed on a density field reconstructed from the distribution of dark matter particles, we will now turn our attention to the distribution of luminous matter. In this chapter, we analyse the redshift space topology and geometry of the nearby Universe by computing the Minkowski functionals of the Updated Zwicky Catalogue (UZC, Falco et al. 1999). The UZC contains the redshifts of almost 20,000 galaxies, and is 96% complete to the limiting magnitude $m_{\mathrm{Zw}} = 15.5$ in more than one fifth of the sky. The UZC includes the famous Center for Astrophysics (CfA) Redshift Survey (Huchra et al. 1999; Huchra et al. 1995; Huchra et al. 1990; Geller & Huchra 1989; Huchra et al. 1983; de Lapparent et al. 1986; Davis et al. 1982). From the UZC we can extract volume limited samples reaching a depth of $70h^{-1}\mathrm{Mpc}$ before sparse sampling dominates. We compute the Minkowski functionals of these samples, taking into account the boundary effects by using unbiased estimators (Section 6.1.1), and quantify the shape of the large–scale galaxy distribution by deriving measures of planarity and filamentarity from the Minkowski functionals (Section 6.1.2).

Recently, a similar analysis on the nearby Universe has been performed by using the LCRS (Bharadwaj et al. 1999). The main advantage of using the UZC over the LCRS is the geometry of the survey: the LCRS consists of six separated slices only $1°\!.5$ thick; thus, an analysis of the full three–dimensional structure cannot be performed. In fact, Bharadwaj et al. (1999) are unable to discriminate whether the "filaments" seen in the LCRS are actually sections of "sheets". Because the UZC consists of two coherent regions covering $\sim 14\%$ and $\sim 7\%$ of the sky respectively without discontinuities, we can address this point here. In fact, Geller & Huchra (1989) were the first to point out the two–dimensionality of the structures by identifying the "Great Wall" in the CfA2 redshift survey. On the other hand, with the UZC, we can only probe the topology of a volume $70h^{-1}\mathrm{Mpc}$ deep, which is too small to be representative of the Universe.

We also compare the results with redshift space mock catalogues constructed from high resolution N–body simulations of two Cold Dark Matter models with either a decaying massive

neutrino (τCDM) or a non–zero cosmological constant (ΛCDM). In these dark–matter simulations, semi–analytic modelling is employed to form and evolve galaxies (Kauffmann et al. 1999a). This means that the density bias, the distribution of galaxies relative to the dark matter, is taken self–consistently into account by including the physics of galaxy formation and evolution explicitly. We are able to compile redshift space mock catalogues which contain galaxies, along with their observable properties, rather than dark matter particles.

It turns out that in both models the large scale galaxy distribution is less coherent than the observed distribution, especially with regard to the large degree of planarity of the real survey. However, given the small volume of the region studied, this disagreement might still be a result of cosmic variance.

6.1 Method

As described in Chapter 2, we consider a volume–limited subsample from a galaxy catalogue as a set of points $\{\mathbf{x}_i\} \subset D$, where $i = 1 \dots N$ and D is the spatial region of the survey. Since the geometry of single points is trivial, we decorate each point with a ball B of radius r, and measure the Minkowski functionals of the resulting union set, the so–called Boolean grain model (Wicksell 1925). In short, we construct the union set

$$A_r = \bigcup_{i=1}^{N} B_{\mathbf{x}_i} \qquad (6.1)$$

and look upon the quantities $V_\mu(A_r)$ as functions of the ball radius r. By increasing the size of the balls, connections are established between points farther apart, until the survey volume is completely filled by the Boolean grain model. The radius can therefore be used as a diagnostic parameter.

6.1.1 Incomplete sky coverage and finite survey depth

Currently, estimation of geometrical quantities such as Minkowski functionals relies on homogeneous sampling of the underlying point distribution. Galaxy catalogues, however, are usually magnitude–limited, and therefore intrinsically inhomogeneous. The standard solution is to use a series of volume–limited subsamples. However, even with volume–limited subsamples, care needs to be taken to obtain meaningful results that are not affected by the sample geometry.

For comparing samples of different size, we may define volume densities v_μ of the Minkowski functionals of this objects by

$$v_\mu = \frac{1}{|D|} V_\mu(A_r), \qquad (6.2)$$

if D is a large box with periodic boundary conditions.

Figure 6.1 The volume fraction of the reduced window as a function of the shrinking distance. The dashed line corresponds to full sky coverage of 4π; the solid line is for the CfA2n window.
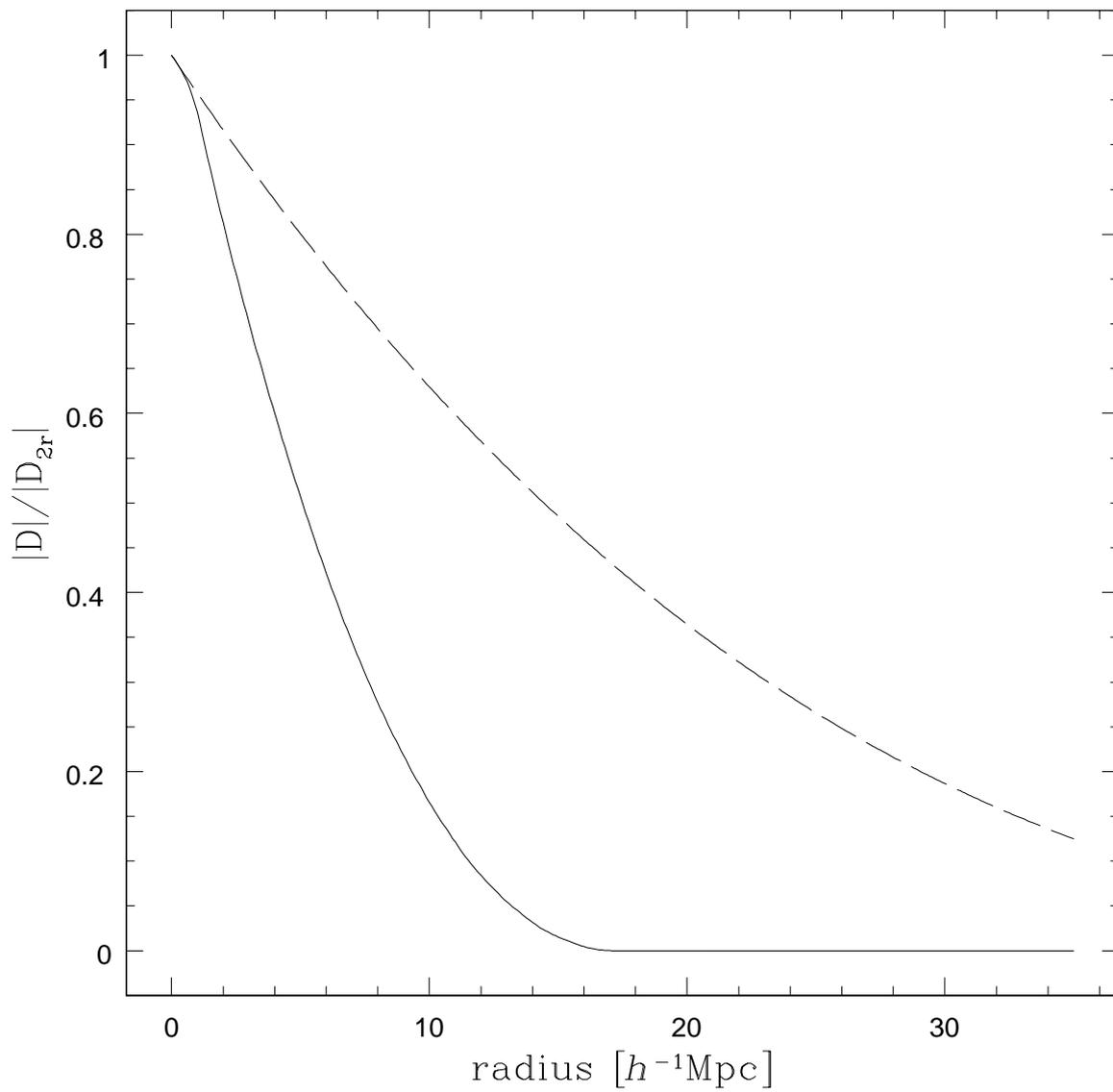
First of all, to take into account the boundary effects on the volume functional v_0, the sample window D must be reduced by the radius r to obtain an unbiased estimate. If D_r denotes the set of all points that are farther than r from the boundary of the sample window D, we have

$$v_0 = \frac{V_0(A_r \cap D_r)}{V_0(D_r)}. \tag{6.3}$$

This correction was introduced by Maurogordato & Lachièze–Rey (1987) for calculating the void probability function of the CfA1 catalogue, and extended by Schmalzing et al. (1996) for all Minkowski functionals of a point set in a convex window.

While for a smooth body, all Minkowski functionals v_μ, $\mu = 1, 2, 3$ can be written as surface integrals, this is no longer possible for the Boolean grain model, which has edges and corners at intersections of two or more balls. Nevertheless the contributions of these non–regular surface to the Minkowski functionals may still be evaluated using a more general concept of surface integration that includes edges and corners (Federer 1959; Schneider 1993). Following Mecke et al. (1994), we arrive at the so–called partition formula, which sorts contributions to the Minkowski functionals from the various types of boundary, i.e. surface points belonging to one, two or three balls[1]. The partition formula states that

$$V_\mu(A_r) = \sum_{i=1}^{N} V_\mu^{(i)}(A_r) + \frac{1}{2} \sum_{i,j=1}^{N} V_\mu^{(ij)}(A_r) + \frac{1}{6} \sum_{i,j,k=1}^{N} V_\mu^{(ijk)}(A_r). \tag{6.4}$$

$V_\mu^{(i)}(A_r)$ denotes the contribution from the uncovered surface of the ball around point $\mathbf{x}_i$, while $V_\mu^{(ij)}(A_r)$ comes from the intersection line of balls i and j, and $V_\mu^{(ijk)}(A_r)$ is located at the corner made by the intersection of balls i, j and k. Rearranging the summations, we have

$$V_\mu(A_r) = \sum_{i=1}^{N} V_\mu(A_r; \mathbf{x}_i), \tag{6.5}$$

where the partial Minkowski functional $V_\mu(A_r; \mathbf{x}_i)$ sums up contributions that include the point i located at $\mathbf{x}_i$. Since an intersection, and hence a non–zero contribution, is only possible if all balls are less than $2r$ apart, the partial Minkowski functionals measure the local morphology in a well–defined neighbourhood, namely a ball of radius $2r$, around each point.

Equation (6.5) is already useful for the practical evaluation of Minkowski functionals for a set of points in a rectangular box with periodic boundary conditions. Basically, one calculates all partial Minkowski functionals and sums them up to estimate the density through Equation (6.2), by

$$v_\mu = \frac{1}{|D|} \sum_{i=1}^{N} V_\mu(A_r; \mathbf{x}_i). \tag{6.6}$$

---

[1] Although intersections of four and more points may occur for very special point distributions (Platzöder 1995), they are irrelevant since they provide no contribution on average. In practice, slight displacements of the order of the numerical resolution can be used to resolve such situations without losing accuracy.

In passing, let us briefly note that for very large N, when calculating all partial Minkowski functionals becomes impractical, it is conceivable to select $M \leq N$ points at random, and estimate the functionals v_μ of the whole box by averaging over M points only.

Even for a complicated survey geometry, the global Minkowski functionals of the Boolean grain model can now be written as sums of the partial Minkowski functionals. Since only neighbours within $2r$ around a point contribute to its partial Minkowski functionals, we can simply restrict the summation to the part of the sample that is further than $2r$ from the boundary. Calling this shrunken window D_{2r}, we have

$$v_\mu = \frac{1}{|D_{2r}|} \sum_{i=1}^{N} \chi_{D_{2r}}(\mathbf{x}_i) V_\mu(A_r; \mathbf{x}_i) \qquad (\mu = 1, 2, 3),$$

(6.7)

where

$$\chi_{D_{2r}}(\mathbf{x}) = \begin{cases} 1 \text{ if } \mathbf{x} \in D_{2r} \\ 0 \text{ if } \mathbf{x} \notin D_{2r} \end{cases}$$

(6.8)

is its characteristic function. In the language of spatial statistics, this quantity is a minus estimator for the volume densities of the Minkowski functionals. Minus estimators especially for the two–point correlation function are already known in cosmology; they have been used for example by Coleman & Pietronero (1992), and thoroughly investigated by Kerscher (1999). One should keep in mind that minus estimators provide unbiased estimates only if applied to a sample from a stationary point process. This is why we always use volume–limited subsamples from catalogues when carrying out our analysis.

Our approach works for arbitrary sky coverage, including the non–convex boundaries of the CfA2 survey. Moreover, at any radius, we use all available points, since points outside the reduced sample window still contribute as neighbours of points remaining inside. The only drawback of our approach, when compared to the method of Schmalzing et al. (1996), is that the window now shrinks by twice the radius. This fact considerably reduces the maximum allowed size of the Boolean ball.

We may quantify the loss of accuracy by the ratio $|D_r|/|D|$ between the volumes of the reduced and the original windows. For complete sky coverage, we have

$$\frac{|D_r|}{|D|} = \left(1 - \frac{r}{R}\right)^3,$$

(6.9)

where R gives the depth of the volume–limited sample and r is the distance to be kept from the boundary. If the sky coverage is not complete, this fraction decreases rapidly. This is illustrated in Figure 6.1, which compares the reduced window size for full sky coverage, and for the 10% sky coverage of the CfA2n window (defined in Section 6.3.1). Apparently, analysing the CfA2 catalogue with Minkowski functionals becomes impractical at radii above $10h^{-1}\mathrm{Mpc}$. We set our maximum Boolean ball radius at $5h^{-1}\mathrm{Mpc}$, so our results are not affected by this problem.

Figure 6.2 These panels compare the best fit of our toy model (dotted) to the Minkowski functionals measured from the CfA2N subsample (solid).
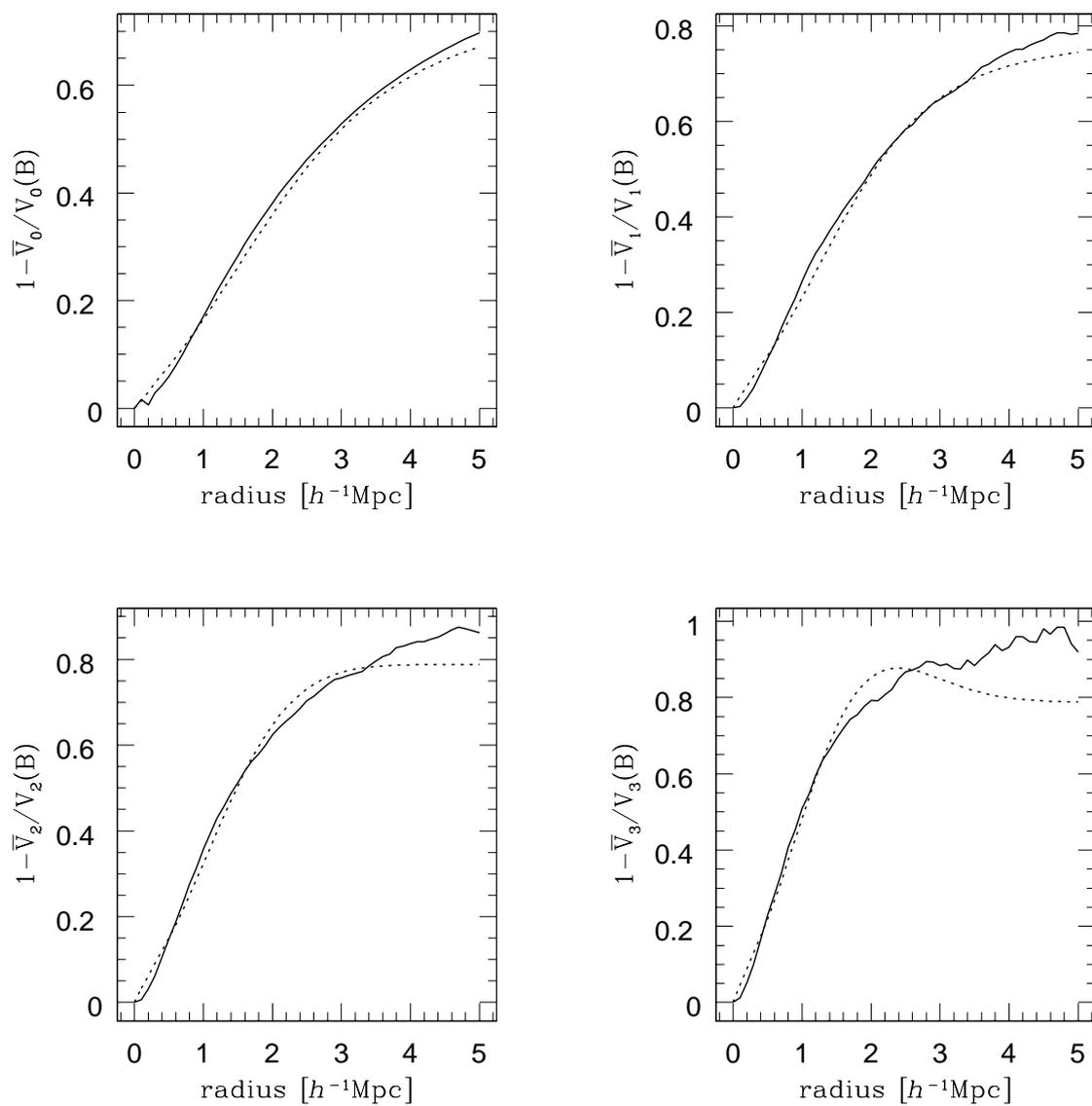
6.1.2 Definition of filamentarity and planarity through a toy model

Consider a random mixture of N bodies where, contrary to the Boolean grain model considered before, each body has different Minkowski functionals. The average Minkowski functionals of the resulting union are given by (Mecke 1994)

$$
\begin{aligned}
v_0 &= 1 - \mathrm{e}^{-\rho \overline{V}_0} \\
v_1 &= \mathrm{e}^{-\rho \overline{V}_0} \rho \overline{V}_1 \\
v_2 &= \mathrm{e}^{-\rho \overline{V}_0} \left(\rho \overline{V}_2 - \frac{3\pi}{8} \rho^2 \overline{V}_1^2 \right) \\
v_3 &= \mathrm{e}^{-\rho \overline{V}_0} \left(\rho \overline{V}_3 - \frac{9}{2} \rho^2 \overline{V}_1 \overline{V}_2 + \frac{9\pi}{16} \rho^3 \overline{V}_1^3 \right)
\end{aligned}
\tag{6.10}
$$

where $\overline{V}_\mu$ is the number weighted average over the isolated Minkowski functionals of all different bodies.

Although the Boolean grain model used in our analysis places identical balls around all points, we may use the formula above for calculating analytically the Minkowski functionals of a simple toy model. In fact, we can take into account the inhomogeneous distribution of identical balls by considering them as randomly distributed objects with Minkowski functionals characteristic of their respective environment.

Let us assume that points are located in one–dimensional filaments, in two–dimensional sheets, or in a homogeneous field. If each of these types of structure were isolated from the rest, we could calculate the Minkowski functionals per single point analytically (Schmalzing 1996). For the average point in the filament, sheet and field, respectively, we have[2]

$$
\begin{aligned}
\frac{\overline{V}_0^{(\mathrm{fil})}}{V_0(B)} &= \tfrac{3}{4} \frac{x + \mathrm{e}^{-2x}}{x^2} - \tfrac{3}{8} \frac{1 - \mathrm{e}^{-2x}}{x^3}, \\
\frac{\overline{V}_1^{(\mathrm{fil})}}{V_1(B)} &= \tfrac{1}{2} \frac{1 - \mathrm{e}^{-2x}}{x}, \\
\frac{\overline{V}_2^{(\mathrm{fil})}}{V_2(B)} &= \tfrac{1}{2} \frac{1 - \mathrm{e}^{-2x}}{x} - x \int_0^1 \mathrm{d}t\, \mathrm{e}^{-2xt} \sqrt{1 - t^2} \arcsin t, \\
\frac{\overline{V}_3^{(\mathrm{fil})}}{V_3(B)} &= e^{-2x},
\end{aligned}
\tag{6.11}
$$

---

[2]We have that $\Psi(x) = \frac{2}{\sqrt{\pi}} \int_0^x \mathrm{d}t \exp(t^2)$ and $I_0(x)$ is the modified Bessel function of order zero.

$$\frac{\overline{V}_0^{(\text{sheet})}}{V_0(B)} = \tfrac{3}{2\pi}\frac{1}{x^2} - \tfrac{3}{4\pi}\frac{e^{-\pi x^2}\Psi(\sqrt{\pi}x)}{x^3},$$

$$\frac{\overline{V}_1^{(\text{sheet})}}{V_1(B)} = \tfrac{1}{2}\frac{e^{-\pi x^2}\Psi(\sqrt{\pi}x)}{x},$$

$$\frac{\overline{V}_2^{(\text{sheet})}}{V_2(B)} = \tfrac{1}{2}\frac{e^{-\pi x^2}\Psi(\sqrt{\pi}x)}{x} - \tfrac{\pi}{2}x^2 e^{-\pi x^2} \tag{6.12}$$

$$\times \int_0^1 \mathrm{d}t \, \exp(\tfrac{\pi}{2}x^2 t^2) \, I_0(\tfrac{\pi}{2}x^2 t^2) \, t^2 \, \arcsin t,$$

$$\frac{\overline{V}_3^{(\text{sheet})}}{V_3(B)} = e^{-\pi x^2}(1 - \pi x^2),$$

$$\frac{\overline{V}_\mu^{(\text{field})}}{V_\mu(B)} = 1, \qquad (\mu = 0, \dots, 3) \tag{6.13}$$

where we have normalised the functionals by dividing through the values for an isolated ball of the same radius, and $x = r/d$ denotes the radius divided by the mean separation of the points on the filament or sheet.

If we further assume that those idealised structures are mixed randomly to form the point distribution we wish to study, the Minkowski functionals v_μ are given through Equation (6.10), where the $\overline{V}_\mu$ are a weighted average over the three types of idealised structure. We can therefore calculate the v_μ numerically, extract the $\overline{V}_\mu$, and perform a linear fit to obtain the percentage for each type of structure in the mixture. By using the quantities $1 - \overline{V}_\mu/V_\mu(B)$, which are exactly zero for field galaxies, we obtain

$$1 - \frac{\overline{V}_\mu}{V_\mu(B)} = \mathcal{F}\left(1 - \frac{\overline{V}_\mu^{(\text{fil})}}{V_\mu(B)}\right) + \mathcal{P}\left(1 - \frac{\overline{V}_\mu^{(\text{sheet})}}{V_\mu(B)}\right) \tag{6.14}$$

for the toy model. We fix the free parameters $\mathcal{F}$ and $\mathcal{P}$ by minimising the χ^2 of all four Minkowski functional profiles at the same time. This procedure yields two numbers for $\mathcal{F}$ and $\mathcal{P}$; we interpret them as measures of filamentarity and planarity in the point distribution.

The standard method of linear fitting solves the normal equations (Press et al. 1987), and also gives an error for our measurements, usually of the order of 0.01. As an example, Figure 6.2 compares the Minkowski functionals of the CfA2N sample and the best fit toy model. Refinements of the fitting procedure, such as a Principal Component Analysis (Kendall 1980) are conceivable, but appear inappropriate in view of the vast simplifications and resulting systematic problems of the toy model. The assumption that galaxies arrange in infinitely thin filaments and sheets is of course a strong simplification. Nevertheless, this simplified model already allows us to put a quantitative measurement on our intuitive impression of the local Universe.

6.2   The Updated Zwicky Catalogue

We now briefly describe the properties of the Updated Zwicky Catalogue (UZC, Section 6.2.1). In Section 6.2.2 we compute its Minkowski functionals with the method based on the Boolean grain model.

6.2.1   Description

Falco et al. (1999) provide new measurements of the positions and redshifts of most galaxies in the Zwicky Catalogue, which has been the basis for the CfA1 and CfA2 redshift surveys (see Falco et al. 1999 and references therein). The UZC represents a homogeneous and well calibrated redshift catalogue of almost 20,000 galaxies. Furthermore, the catalogue is now publicly available.

The UZC covers the northern celestial hemisphere which is not obscured by the Galactic disk; the redshift catalogue is $\sim 96\%$ complete to a limiting Zwicky magnitude $m_{\mathsf{Zw}} = 15.5$.

The CfA2 redshift surveys have been compiled over many years and this name has indicated progressively larger areas of the sky, as the surveys were reaching completion. The UZC contains all of the CfA2 survey analysed in the literature to April 1999. Following Falco et al. (1999), we refer to the area with right ascension ranges $8^{\mathrm{h}} \leq \alpha_{1950} \leq 17^{\mathrm{h}}$ and $20^{\mathrm{h}} \leq \alpha_{1950} \leq 4^{\mathrm{h}}$ and declination range $-2\overset{\circ}{.}5 \leq \delta_{1950} \leq 50°$ as the CfA2 region. The first right ascension range defines the CfA2N region containing 12,082 galaxies with measured redshift. The regions between 3^{h} and 4^{h} and between 20^{h} and 22^{h} are sparsely populated in the catalogue largely because of galactic obscuration. Therefore we restrict the CfA2S region to the 4,436 galaxy redshifts in the right ascension range $22^{\mathrm{h}} \leq \alpha_{1950} \leq 3^{\mathrm{h}}$. Both areas used in our analysis are more than 98% complete. Moreover the UZC contains some galaxies fainter than the magnitude limit. These galaxies are mainly in multiplets which were unresolved by Zwicky. They affect the properties of the galaxy distribution on very small scales but should leave our analysis unchanged. Note that the actual number of available redshifts in the UZC is currently increasing, as the catalogue is being updated.

Visual inspection shows well–defined structure in the CfA2 surveys, with large voids and sheet–like structures. Figure 6.3 shows a cone diagram of the volume–limited samples from CfA2N and CfA2S we analyse here. The dense feature close to the boundary of the northern part of the catalogue is the "Great Wall" (Geller & Huchra 1989, see also Figure 6 of Falco et al. 1999). In CfA2S, the apparent galaxy concentration is the Perseus–Pisces supercluster.

6.2.2   Topology and geometry

In our analysis, we use volume–limited samples $70h^{-1}\mathrm{Mpc}$ deep. This provides a good compromise between object number (of all volume–limited subsamples, the $70h^{-1}\mathrm{Mpc}$ one includes the highest number of objects) and depth (parts of the Great Wall are closer than $70h^{-1}\mathrm{Mpc}$).

Figure 6.3 Cone diagram of the volume–limited subsample of $70h^{-1}\mathrm{Mpc}$ depth we analyse here from the UZC. The larger and smaller cones correspond to the northern and southern region, respectively. The large dot marks the location of our galaxy and alludes to Adams (1986).
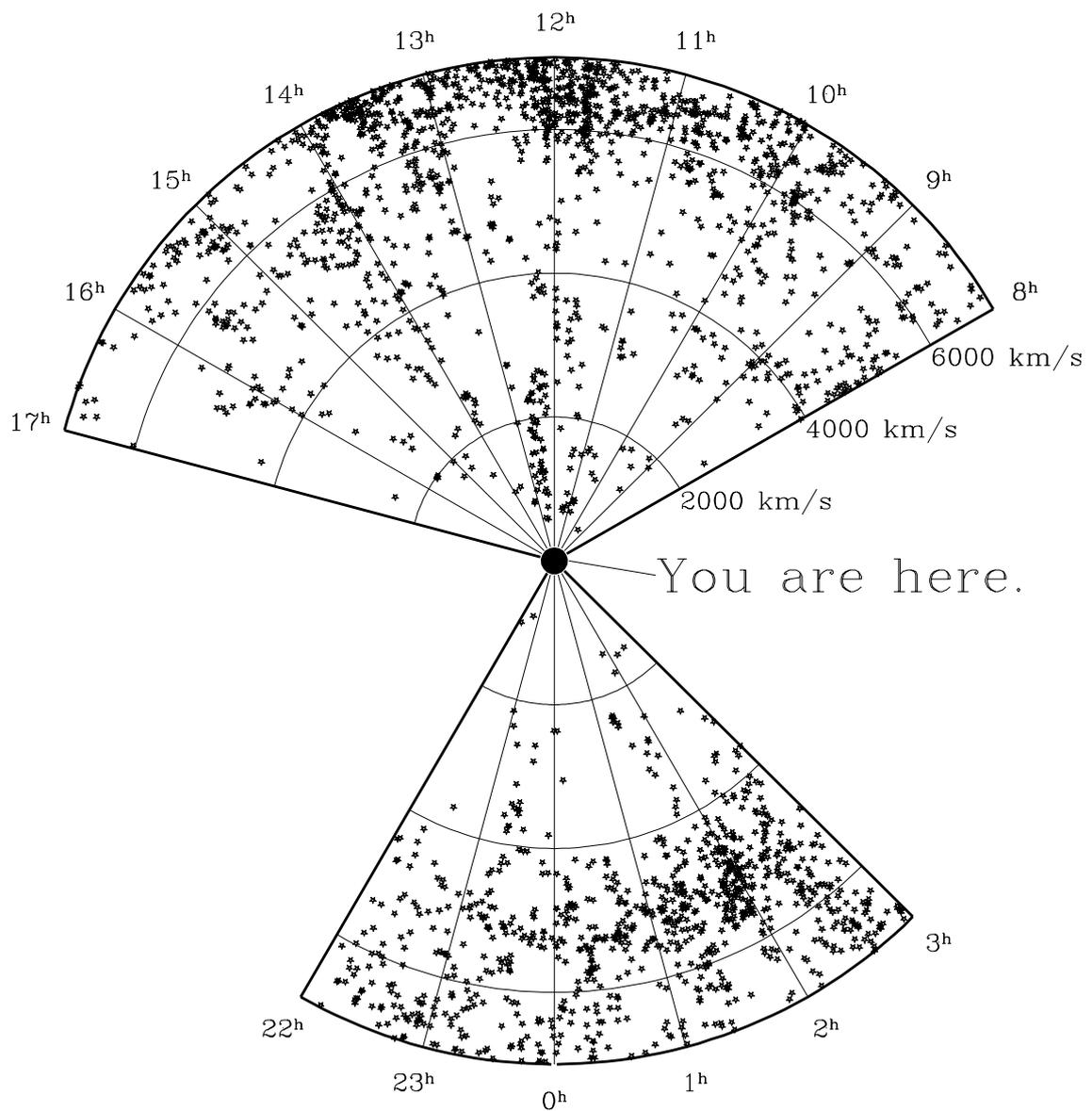
Figure 6.4 Minkowski functionals of CfA2N (solid line) and CfA2S (dashed line). All curves were calculated from volume–limited samples $70h^{-1}\mathrm{Mpc}$ deep. Note that the CfA2N Minkowski functionals are systematically larger than the CfA2S ones, indicating larger clustering (Marzke et al. 1995).
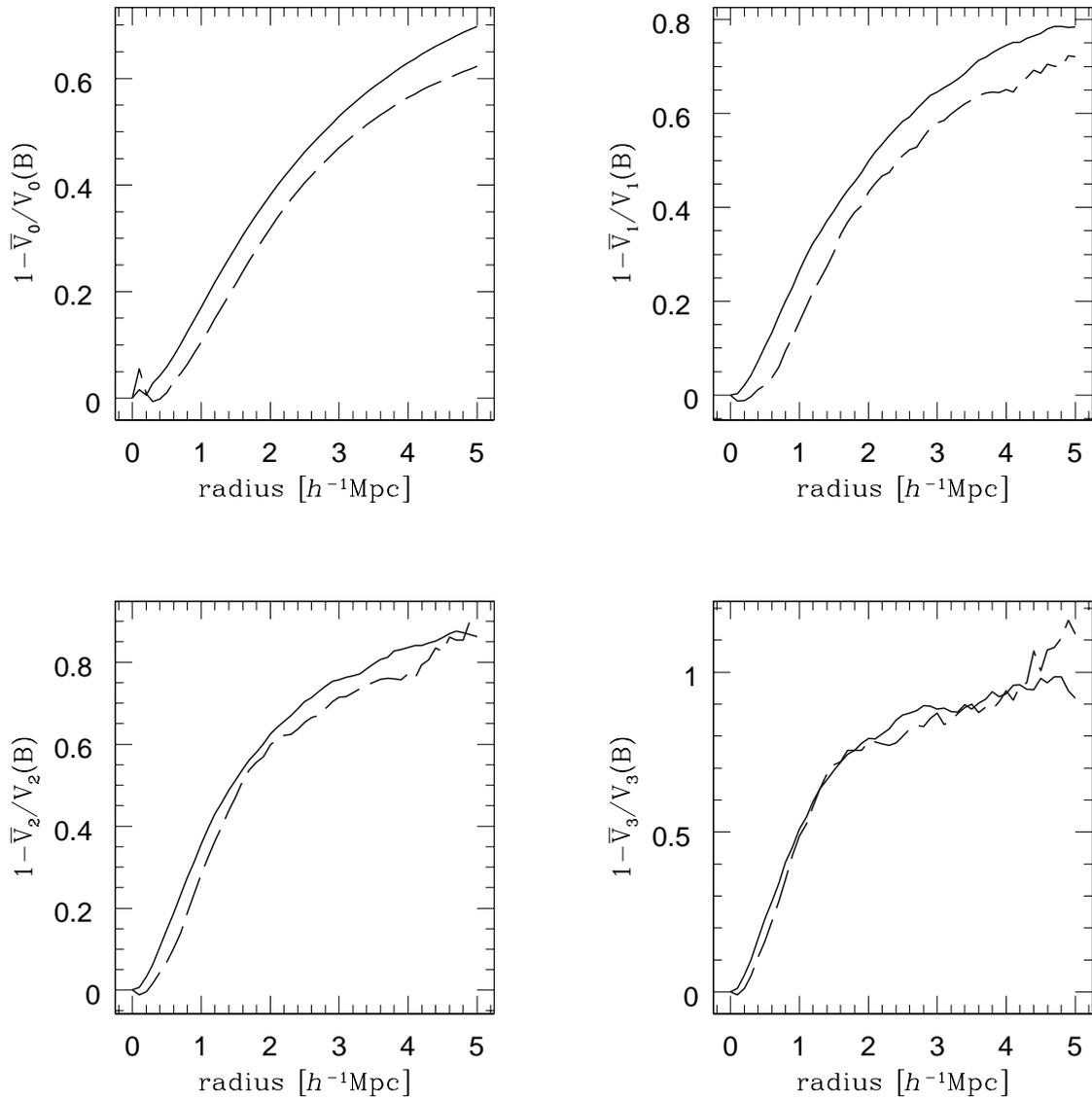
Since the scaling behaviour of Minkowski functionals with the number density is not known in general, we take care to always analyse samples with the same number density of galaxies. From all volume–limited samples, we choose a set of slightly smaller random samples such that the number density corresponds to exactly 1,000 galaxies in the CfA2n volume (see Section 6.3.1 for a definition).

The results of our analysis of the UZC are shown in Figure 6.4. Instead of the Minkowski functionals v_μ themselves, we plot the quantities $1 - \overline{V}_\mu / V_\mu(B)$. A thorough motivation of their use is given in Section 6.1.2; most notably they are exactly equal to zero for a stationary Poisson process, i.e. a random distribution of points with spatially constant density. Generally speaking, values above zero reveal a clustered distribution of galaxies. Note, however, that the $\overline{V}_\mu$ can be written as an alternating series containing the hierarchy of correlation functions (Schmalzing et al. 1999b), so clustering does not simply mean that the two–point function is positive. As an example, consider $1 - \overline{V}_0 / V_0(B)$. By Equation (6.10) this quantity increases monotonically with $1 - v_0$, the void probability function (White 1979). Hence larger values in our plots indicate larger voids in the point distribution, which in turn means that points have to cluster more tightly.

Some of the functionals, most notably the volume v_0, indicate stronger clustering for CfA2N than for CfA2S. In fact, CfA2N contains more clusters than CfA2S and the Great Wall is more apparent than the Perseus–Pisces supercluster in CfA2S. The two-point correlation function already showed that the CfA2N clustering is peculiarly large when compared to other surveys (Marzke et al. 1995; Diaferio et al. 1999a).

We use a toy model to characterise the geometry of CfA2: we assume that galaxies are distributed in one-dimensional filaments, two-dimensional sheets or in a homogeneous field. We can compute the Minkowski functionals for this model analytically. Two coefficients, $\mathcal{F}$ and $\mathcal{P}$, give the relative contribution of filaments and sheets in the galaxy distribution (see Section 6.1.2 for details). Note, however, that this approach has the main shortcoming of assigning structures which are loosely clustered to the planarity component, because alternative assignments would yield considerably worse fits. Therefore, $\mathcal{F}$ and $\mathcal{P}$ yield the relative importance of the two structural components, but their exact values should be considered cautiously.

By fitting our toy model to the CfA2N Minkowski functional curves, we obtain $\mathcal{F} = 0.12$ and $\mathcal{P} = 0.67$ for the filamentarity and planarity contribution, respectively. The high degree of planarity is expected from the visual appearance of the Great Wall in the catalogue. For the CfA2S we find $\mathcal{F} = 0.01$ and $\mathcal{P} = 0.75$. The large value of $\mathcal{P}$ in both regions reflects the tendency of galaxies to be distributed in two-dimensional structures or in loose structures rather than in filaments. Moreover, the smaller size of the CfA2S region also tends to increase slightly the signal for planar and loose structures.

Because of its simplicity, our toy model provides a straightforward way of measuring planarity and filamentarity, but it is inadequate to model more complex realistic situations. We are planning some refinements in order to obtain results in better agreement with our intuitive interpretations of the galaxy distribution.

Table 6.1 Parameters of the two GIF simulations (Kauffmann et al. 1999a) used in our analysis. The Hubble constant h, the particle mass m_p, and the comoving size L of the simulation box are in units of $H_0 = 100 \mathrm{km\,s^{-1}/Mpc}$, $10^{10} h^{-1} M_\odot$, and $h^{-1}\mathrm{Mpc}$, respectively.

model	Ω_0	Ω_Λ	h	σ_8	m_p	L
ΛCDM	0.3	0.7	0.7	0.90	1.4	141
τCDM	1.0	0.0	0.5	0.60	1.0	85

6.3 Mock galaxy redshift surveys

We compare the morphology of the CfA2n redshift survey with two variants of a CDM model. In Section 6.3.1 we briefly describe the simulations and how we model galaxy formation and evolution. Section 6.3.2 discusses how the luminosity limit of the galaxy sample, the shape of the sample volume and the peculiar velocity distortions in redshift space affect the Minkowski functionals. Section 6.3.3 contains the comparison of CfA2n with the mock galaxy redshift surveys.

6.3.1 Models

We consider two simulations from the GIF project (Kauffmann et al. 1999a): a CDM model with either $\Omega_0 = 1$ (τCDM) or $\Omega_0 = 0.3$ and $\Lambda = 0.7$ (ΛCDM). The simulations contain 256^3 dark matter particles each. The normalisation σ_8 of the power spectrum of the initial density perturbations was chosen to reproduce the present day abundance of rich galaxy clusters. Table 6.1 summarises the parameters of the models.

The simulations were run with Hydra (Pearce & Couchman 1997), the parallel version of the AP^{3}M code (Couchman 1991; Couchman et al. 1995), kindly provided by Jenkins et al. (1997).

Kauffmann et al. (1999a) combine these high resolution N-body simulations with semi-analytic modelling to form and evolve galaxies within dark matter halos: gas cooling, star formation, supernova feedback, stellar evolution and galaxy–galaxy merging are the relevant processes included. At any cosmological epoch, the observable properties of the galaxies, namely luminosity, colour, star formation rate, and stellar mass, are the result of the galaxy's merging history, its interaction with the environment and the passive evolution of its stellar content.

The models predict photometric and clustering properties of galaxies at redshift $z = 0$ (Kauffmann et al. 1999a; Diaferio et al. 1999a), and their evolution to high redshift (Kauffmann et al. 1999b; Diaferio et al. 1999b) which are in reasonable agreement with observations. Moreover, these analyses suggest further diagnostics for constraining the cosmological model and determining the relevance of the different galaxy formation processes.

These simulations represent the only models, available to date, where galaxies form and evolve in a self-consistent cosmological model and within a volume of the Universe sufficiently large for

Figure 6.5 Cone diagrams of two mock catalogues. The upper and lower cones are a volume–limited subsample constructed from the CfALF ΛCDM model and the CfALF τCDM model, respectively. The observer galaxy is indicated by the large dot.
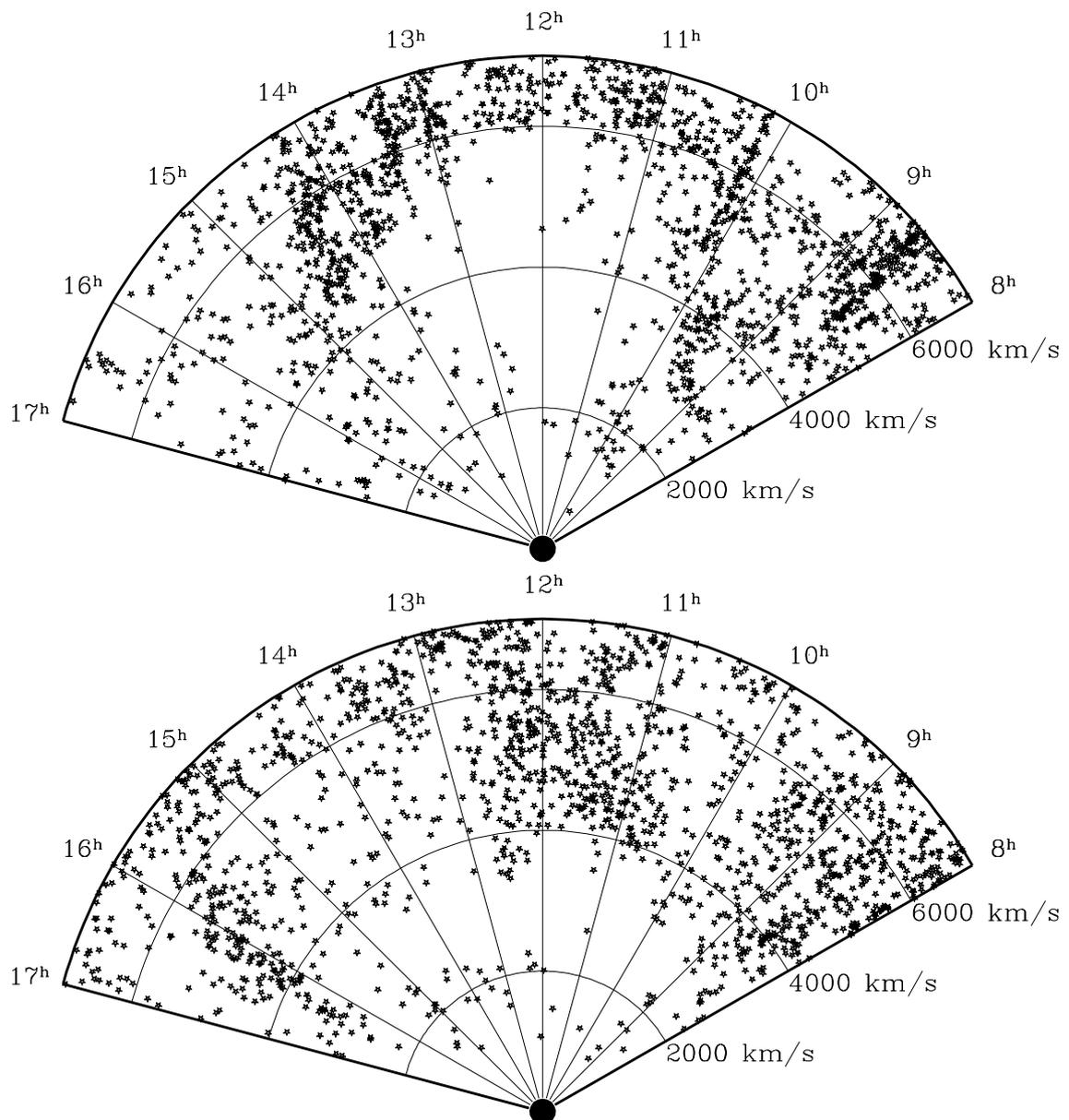
an analysis of the large scale galaxy distribution (see also Pearce et al. 1999). In fact, Diaferio et al. (1999a) extract mock galaxy redshift surveys from the simulations at $z = 0$ for a direct comparison with the CfA2N redshift survey, whose volume $\sim 7 \times 10^5 h^{-3} \mathrm{Mpc}^3$ is comparable to the volume of the simulation boxes. Specifically, they consider the area $8^\mathrm{h} \leq \alpha_{1950} \leq 17^\mathrm{h}$, $8°\!.5 \leq \delta_{1950} \leq 44°\!.5$ which has a declination range smaller than CfA2N. We refer to this region as the CfA2n region. In order to resemble the CfA2n appearance in the mock catalogues as closely as possible, the mock redshift surveys were compiled by locating the observer home galaxy within the simulation box on a galaxy similar to the Milky Way at $\sim 70 h^{-1} \mathrm{Mpc}$ away from a Coma–like massive cluster.

Diaferio et al. (1999a) analyse both the CfA2n and the mock redshift surveys with the same standard observational techniques. They show that the small scale clustering properties of galaxies, namely galaxy group properties, the redshift space correlation function and pairwise velocities, are reproduced by both models independently of Ω_0. Despite this success, a visual inspection of the mock redshift surveys shows remarkable differences in the large scale distribution of galaxies. Both models fail to produce structures as coherent and as sharply defined as in the CfA2n survey, although ΛCDM catalogues seem to be in better qualitative agreement with the data (see their Figures 5–8). Two examples of volume–limited subsamples constructed from these models are shown in Figure 6.5.

The B-band galaxy luminosity function predicted by the semi–analytical models is not a good fit to the CfA survey luminosity function (Marzke et al. 1994). To investigate the effect of the luminosity function on the small-scale clustering properties, Diaferio et al. (1999a) assign new luminosities to the galaxies in order to reproduce the CfA luminosity function exactly, while preserving the galaxies' luminosity rank. We thus have two sets of mock catalogues, where the galaxies have luminosities according to the semi–analytical model luminosity function (SALF) or the CfA luminosity function (CfALF). We have analysed both the SALF and CfALF mock catalogues. Similarly to other measures related to the underlying mass distribution, the Minkowski functionals are basically independent of the luminosity function adopted. Throughout this analysis, we show results from the CfALF catalogues alone. Moreover, we consider ten mock catalogues for each simulation to assess the robustness of our results.

6.3.2 Luminosity, geometry and redshift space effects

Kauffmann et al. (1999a) show that the two–point correlation function is similar for galaxy samples with different magnitude limits. The Minkowski functionals are able to extract more information from the galaxy distribution. In fact, bright galaxies cluster stronger than faint ones. Figure 6.6 shows the Minkowski functionals for the ΛCDM simulation; τCDM yields very similar plots. In order to keep constant the statistical significance of different galaxy samples, we compute the Minkowski functionals only of subsamples containing the same number density of galaxies, and average over ten subsamples. The solid lines show the Minkowski functionals for galaxies brighter than $M_B = -17.5 + 5 \log h$. The Minkowski functionals for galaxies brighter than $M_B = -18.7 + 5 \log h$ are indicated by dashed lines. This absolute magnitude corresponds

Figure 6.6 Effect of luminosity and geometry on the Minkowski functionals of the CfALF ΛCDM model. The dashed and solid lines are for all galaxies in the simulation box brighter than $M_{\mathrm{B}} = -18.7 + 5\log h$ and $M_{\mathrm{B}} = -17.5 + 5\log h$ respectively. The shaded area indicates the 1σ–spread over ten subsamples.
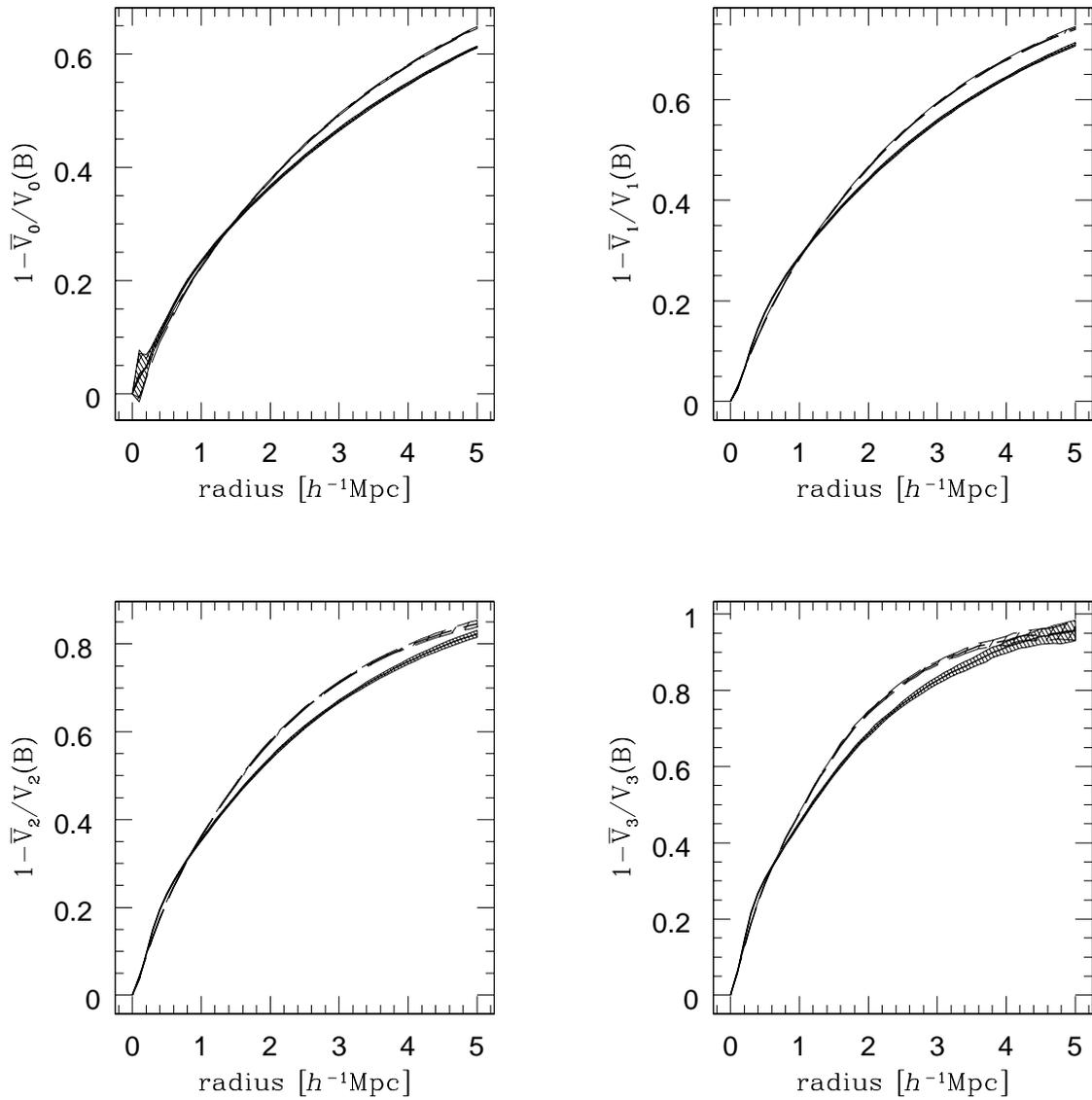
to an apparent magnitude $m_B = 15.5$ at $70h^{-1}\mathrm{Mpc}$. The difference between the two curves is significantly larger than the scatter between different subsamples from the same set of galaxies.

We have also checked that the Minkowski functionals computed with our procedure (Section 6.1.1) are insensitive to the shape of the sample volume. We do this by calculating the Minkowski functionals of the real space position of the galaxies averaged over the volume–limited subsamples extracted from the ten mock redshift surveys. The dashed lines in Figure 6.6 lie within the scatter of the resulting Minkowski functionals (shown as the dark shaded area in Figure 6.7). Since they come from volumes which have differing shape (periodic, rectangular box versus CfA2n volume) but contain galaxies with roughly the same luminosity, we can deduce that the difference due to the sample geometry is negligible in comparison to the scatter between individual mock catalogues. Note that for mock catalogues, the scatter is much larger than for subsamples from the whole simulation box, simply because a mock catalogue contains fewer galaxies.

Redshift space distortions tend to destroy structure on small scales, because groups and clusters in real space are diluted in redshift space by the finger–of–god effect. Figure 6.7 shows this difference in the ΛCDM model. The Minkowski functionals of the real and redshift space galaxy distributions almost coincide at larger values of the Boolean ball radius, because redshift space distortions become less relevant on large scale.

6.3.3 Comparison with the CfA2 catalogue

Figure 6.8 compares the morphological properties of CfA2n with the mock catalogues: neither of the two models is capable of reproducing the CfA2n properties. At Boolean ball radii $\sim 1 - 2h^{-1}\mathrm{Mpc}$, all the four Minkowski functionals of CfA2n are five times the r.m.s. above the average mock catalogue of the ΛCDM. Given the rather small number of ten realisations, we can conclude that the CfA2n clustering is not reproduced by the ΛCDM model at the 90% confidence level. The τCDM model shows an even smaller degree of clustering. At larger radii and for the mean curvature v_2 and Euler characteristic v_3, the disagreement between the models and observations is less dramatic, but also less significant, because the sample volume is filled by a considerably smaller number of Boolean balls. It is crucial to recall, however, that the scatter of quantities derived from the variance over all ten mock catalogues underestimates the sampling variance, because the mock surveys are constructed from the same small parent simulation.

The failure of the ΛCDM in reproducing the CfA2n Minkowski functionals is particularly interesting, because, on average, this model fits other observations reasonably well and in any case better than the τCDM model (Kauffmann et al. 1999a; Kauffmann et al. 1999b). The Minkowski functionals confirm that ΛCDM should be preferred to τCDM.

It is worth noting that, because of the smaller degree of clustering of CfA2S when compared to CfA2N (see Figure 6.4), the difference between the CfA2S and the models is less pronounced, although still significant. The result is an example of the power of the Minkowski functionals

Figure 6.7 Minkowski functionals of the ΛCDM mock catalogues in real space (dark shaded area) and in redshift space (light shaded area). The shaded area indicates the r.m.s. spread over the ten mock catalogues.
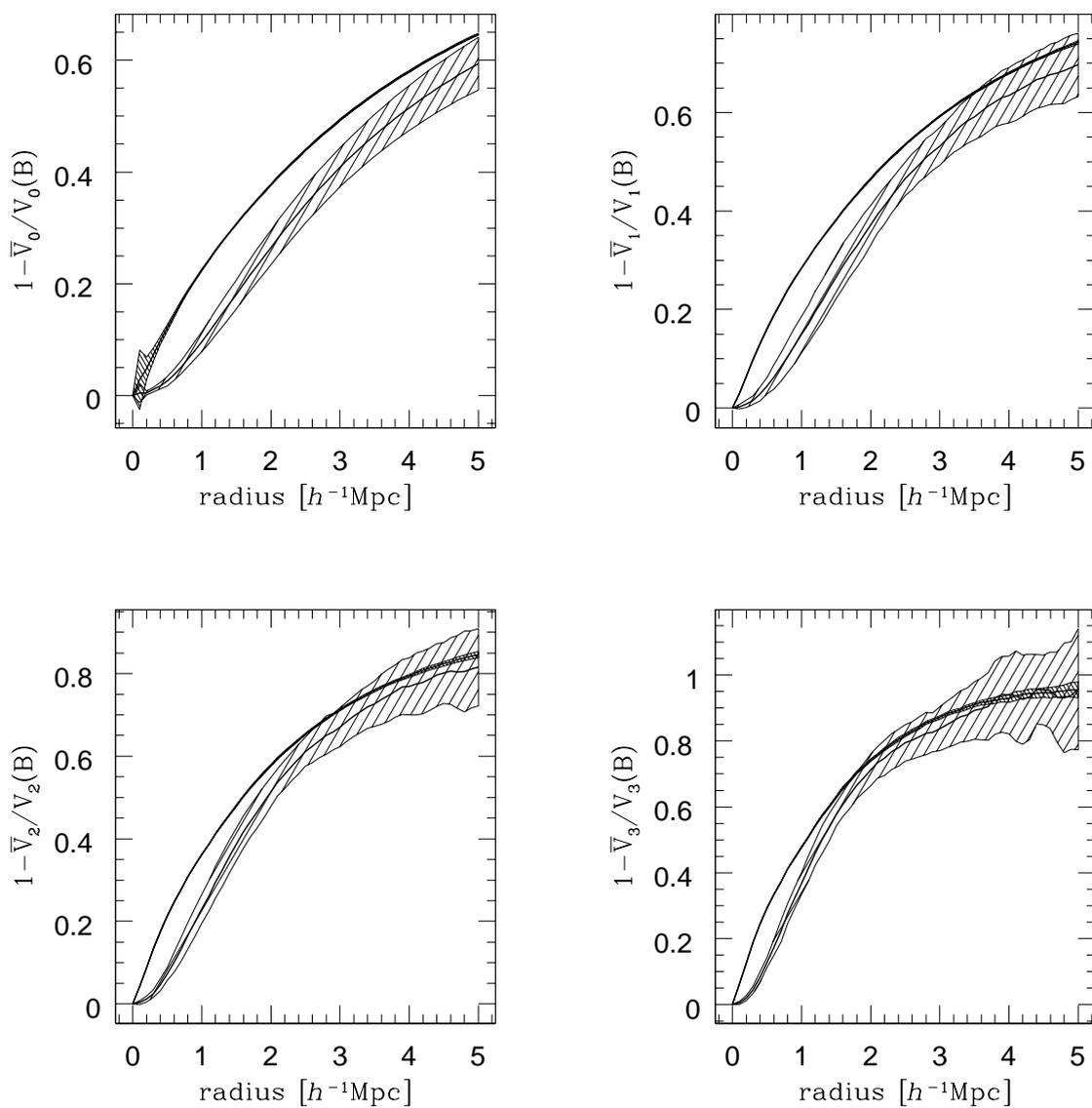
Figure 6.8 Structure in the CfA2n redshift survey, and in the mock catalogues extracted from the ΛCDM (dark shaded area) and the τCDM (light shaded area) models. The areas indicate the 1-σ scatter over the ten mock catalogues constructed for each model. All curves are for volume–limited samples $70h^{-1}\mathrm{Mpc}$ deep.
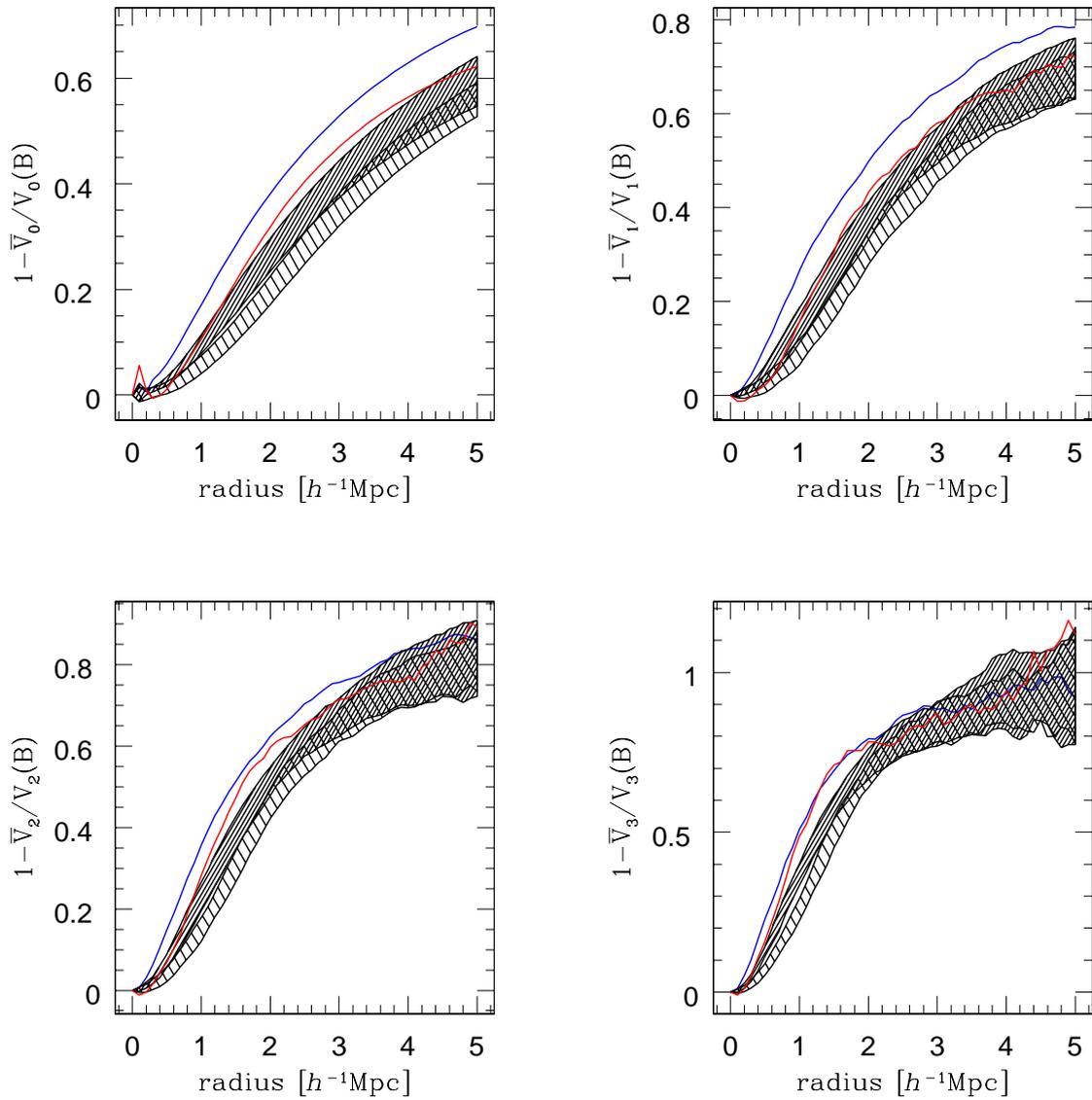
analysis. In fact, Diaferio et al. (1999a) show that the two point correlation function of the mock catalogues agrees quite well with the CfA2S function. As expected, the Minkowski functionals are able to extract more information on the clustering properties of galaxies.

We can also perform the fitting procedure described in Section 6.1.2 to estimate the planarity and filamentarity of the structure contained in the mock catalogues. ΛCDM yields $\mathcal{F} = 0.01$ and $\mathcal{P} = 0.64$, while τCDM results in $\mathcal{F} = 0.02$ and $\mathcal{P} = 0.60$ for the average mock catalogues. As pointed out in Section 6.2.2, the large values of $\mathcal{P}$ compared to $\mathcal{F}$ indicate that the galaxies tend to be distributed in two-dimensional or loose structures rather than filaments. Compared to the values determined from the real catalogue in Section 6.2.2 ($\mathcal{F} = 0.12$ and $\mathcal{P} = 0.67$), $\mathcal{P}$ is $\sim 10\%$ smaller in the models than in CfA2, again indicating a smaller degree of clustering.

We cannot exclude that the failure of the models in reproducing the Minkowski functionals of CfA2N is due to the small size of both the simulation box and the real survey. In fact, the initial power spectrum of the density perturbations does not contain Fourier components with wavelength larger than the box size. This prevents the development of large-scale non–Gaussian features such as the Great Wall. Moreover, as already indicated by an analysis of the IRAS 1.2Jy catalogue (Kerscher et al. 1998), a volume–limited subsample $70h^{-1}\mathrm{Mpc}$ deep is too small to be a fair sample of the Universe. Therefore, our analysis is likely to be dominated by local effects, which should average out when deeper catalogues are considered. This hypothesis could already be tested with simulations in larger volumes (see e.g. Doroshkevich et al. 1999), where one could estimate the cosmic variance, or with a constrained realisation (Mathis et al. 1999), which would not suffer from the fair sample problem.

6.4 Discussion

de Lapparent et al. (1986) and Geller & Huchra (1989) first pointed out that the large–scale galaxy distribution in the CfA2 redshift survey has a two–dimensional sheet–like structure rather than a one–dimensional filamentary appearance. Here we quantify this visual impression by analysing the redshift space topology and geometry of the UZC. By constructing a Boolean grain model of balls centred around the galaxies, we obtain unbiased estimates for the Minkowski functionals of the galaxy distribution. We then fit the Minkowski functionals to a toy model to extract quantitative information on the degree of filamentarity and planarity in CfA2. We find that filament– and sheet–like structures contribute 12% and 67% to the Minkowski functionals of CfA2N respectively, while the remainder comes from a homogeneous distribution of galaxies. Our results therefore strongly suggest that the filaments identified in the LCRS (Bharadwaj et al. 1999) are actually sections of sheets.

Apparently, our simple model is inadequate in other realistic situations, such as CfA2S (the southern part of the UZC), where looser structures are present. However, the encouraging result we obtain with CfA2N indicates that it is worth pursuing this approach with some refinements. When deeper surveys as the Sloan Digital Sky Survey will be available, our approach can represent a valuable tool to quantify the topology of the galaxy distribution. Most notably,

Figure 6.9 The average Minkowski functionals of mock CfA2 catalogues extracted from truncated Zel'dovich simulations (see Chapter 3 for details) of $60h^{-1}\mathrm{Mpc}$ (solid lines) and $240h^{-1}\mathrm{Mpc}$ (dashed lines) sidelength. The differences are not significant, especially on small scales.
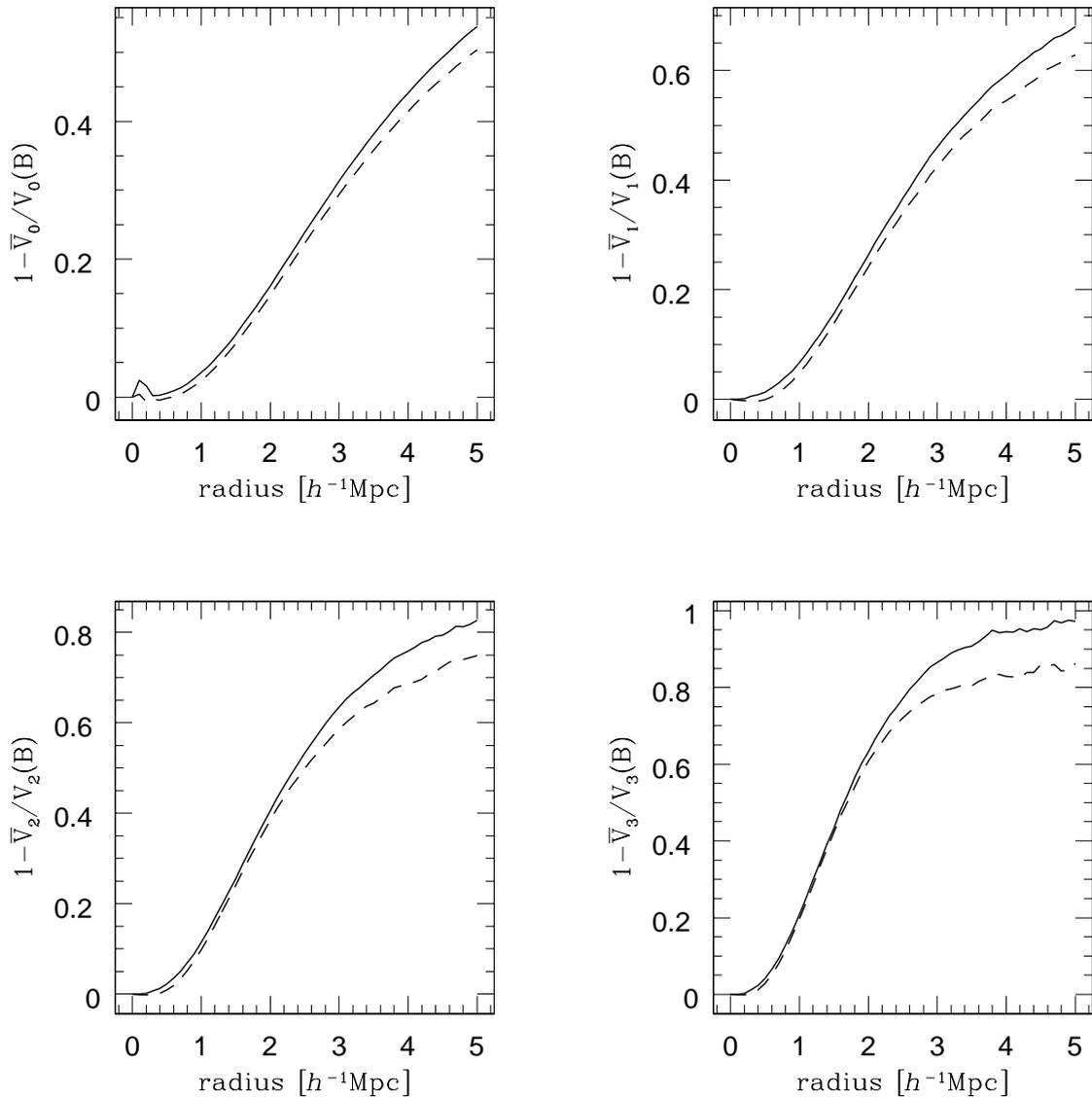
contrary to Sahni et al. (1998) or Schmalzing et al. (1999a), the method presented here does not rely on shape measurements of isolated isodensity contours. Hence we can also obtain meaningful results in low to intermediate density environments, where objects in a smoothed field already percolate and are inaccessible to measurements.

We also compare CfA2n, a smaller volume of CfA2N, with mock redshift surveys extracted from a ΛCDM and a τCDM model. The models do not show the high degree of clustering of CfA2n. However, CfA2n seems to be quite a peculiar region of the Universe and the disagreement between the models and, for example, CfA2S is less dramatic. Moreover, the Minkowski functionals of the ΛCDM are closer to observations than the Minkowski functionals of τCDM. In conclusion, although our topological analysis confirms earlier results which indicate that overall ΛCDM is a better representation of the real Universe than τCDM, we show, at the same time, that ΛCDM is not yet a satisfactory model of the Universe.

Vogeley et al. (1994b) compare the genus statistics of CfA2 with a ΛCDM model. They find that this model is consistent with CfA2 on scales $\geq 10h^{-1}\mathrm{Mpc}$, but it fails to reproduce CfA2 on smaller scales. They use a simple bias scheme to identify galaxies and extract mock catalogues from N-body simulations. So the physics acting on small scales, which is missing from their simulations, might be responsible for the disagreement. Here, we used semi–analytic modelling to form and evolve galaxies in N–body simulations, but still the agreement between ΛCDM and CfA2 does not improve.

Note that our Boolean grain model approach follows a direction complementary to the genus statistics analysis generally used. The standard method introduced by Gott III et al. (1986) employs a density field smoothed on varying scales. Increasing the smoothing kernel width successively erases structure on small scales (ter Haar Romeny et al. 1991). The analysis of Vogeley et al. (1994b), for example, employs widthes between $6h^{-1}\mathrm{Mpc}$ and $20h^{-1}\mathrm{Mpc}$ and addresses the issue of the Gaussianity of the distribution of the initial density perturbations. We adopt the Boolean grain model approach by Mecke et al. (1994) and decorate each galaxy with a sphere of radius r varying from zero to $5h^{-1}\mathrm{Mpc}$. This procedure captures the morphology of distinct structures on larger scales while preserving the small scale structure information.

Finally, the small volume of both the survey and our simulation box represents a severe limit to our analysis, because we are unable to quantify the effect of cosmic variance. Figure 6.9 shows the Minkowski functionals for mock CfA2 catalogues extracted from boxes of $60h^{-1}\mathrm{Mpc}$ and $240h^{-1}\mathrm{Mpc}$, respectively. Although some differences at larger radii are visible, they are far from significant, and the functionals at small radii practically coincide. Since the particle distributions were calculated using the second order Zel'dovich schemes presented in Chapter 3, and "galaxies" are not biased against the dark matter, these results should be taken with a grain of salt. However, they present a first indication that increasing the simulation box size would not remedy the observed discrepancies between the UZC data and the CDM models. We expect that in the end surveys covering larger volumes should enable our topological approach to discriminate between cosmological models successfully.

References

Abramowitz, M. X. & Stegun, I. A., *Handbook of mathematical functions*, 9th ed., Dover Publications, New York, 1970.

Abramson, I. S., *On bandwidth variation in kernel estimates*, Ann. of Stat. **10** (1982), 1217–1223.

Adams, D., *The hitch hiker's guide to the galaxy: A trilogy in four parts*, William Heinemann Ltd., London, 1986.

Adler, R. J., *The geometry of random fields*, John Wiley & Sons, Chichester, 1981.

Albrecht, A., Bayette, R. A., & Robinson, J., *The case against scaling defect models of cosmic structure formation*, Phys. Rev. Lett. **79** (1997), 4736–4739.

Alexandrov, A. D., *Neue Ungleichungen zwischen den gemischten Volumina und ihre Anwendungen*, Matem. Sb. SSSR **2** (1937), 1205–1238, in Russian, summary in German.

Allendoerfer, C. B. & Weil, A., *The Gauss–Bonnet theorem for Riemannian polyhedra*, Trans. Amer. Math. Soc. **53** (1943), 101–129.

Arnol'd, V. I., Shandarin, S. F., & Zel'dovich, Ya. B., *The large–scale structure of the universe. I – General properties of one– and two–dimensional models*, Geophysical and Astrophysical Fluid Dynamics **20** (1982), 111–130.

Babul, A. & Starkman, G. D., *A quantitative measure of structure in the three–dimensional galaxy distribution: Sheets and filaments*, Ap. J. **401** (1992), 28–39.

Balian, R. & Schaeffer, R., *Scale–invariant matter distribution in the Universe I. counts in cells*, Astron. Astrophys. **220** (1989), 1–29.

Balian, R. & Schaeffer, R., *Scale–invariant matter distribution in the Universe II. bifractal behaviour*, Astron. Astrophys. **226** (1989), 373–414.

Banday, A. J., Zaroubi, S., & Górski, K. M., *On the non–Gaussianity observed in the COBE–DMR sky maps*, Ap. J. (1999), submitted, astro-ph/9908070.

Bardeen, J. M., Bond, J. R., Kaiser, N., & Szalay, A. S., *The statistics of peaks of Gaussian random fields*, Ap. J. **304** (1986), 15–61.

Bardeen, J. M., Bond, J. R., & Efstathiou, G., *Cosmic fluctuation spectra with large–scale power*, Ap. J. **321** (1987), 28–35.

Bardeen, J. M., *Gauge–invariant cosmological perturbations*, Phys. Rev. D **22** (1980), 1882–1905.

Barrow, J. D., Sonoda, D. H., & Bhavsar, S. P., *Minimal spanning trees, filaments and galaxy clustering*, Mon. Not. Roy. Astron. Soc. **216** (1985), 17–35.

Bennett, C. L., Hinshaw, G., Jarosik, N. C., Mather, J., Meyer, S. S., Page, L., Skillman, D., Spergel, D. N., Wilkinson, D. T., & Wright, E. L., *The microwave anisotropy probe (MAP) mission concept*, Bull. American Astron. Soc. **187** (1995), no. 71.09, 1385.

Bennett, C. L., Banday, A., Górski, K. M., Jackson, P. D., Keegstra, P. B., Kogut, A., Smoot, G. F., Wilkinson, D. T., & Wright, E. L., *4-year COBE DMR cosmic microwave background*

observations: Maps and basic results, Ap. J. Lett. **464** (1996), L1–L4.

Bersanelli, M., Bouchet, F. R., Efstathiou, G., Griffin, M., Lamarre, J. M., Mandolesi, N., Norgaard-Nielsen, H. U., Pace, O., Polny, J., Puget, J. L., Tauber, J., Vittorio, N., & Volonté, S., *COBRAS/SAMBA. A mission dedicated to imaging the anisotropies of the cosmic microwave background. Report on the phase A study*, European Space Agency, February 1996.

Bertschinger, E., *Gravitational simulations of cosmic structure formation*, Bull. American Astron. Soc. **186** (1995), no. 26.02, 849.

Bharadwaj, S., Sahni, V., Sathyaprakash, B. S., Shandarin, S. F., & Yess, C., *Evidence for filamentarity in the Las Campanas Redshift Survey*, Ap. J. (1999), submitted, astro-ph/9904406.

Bildhauer, S. & Buchert, T., *The relation between peculiar velocity and density parameter for a class of solutions in Newtonian cosmology*, Progr. Theor. Phys. **86** (1991), 653–658.

Bildhauer, S., Buchert, T., & Kasai, M., *Solutions in Newtonian cosmology − The pancake theory with cosmological constant*, Astron. Astrophys. **263** (1992), 23–29.

Binney, J. J., Dowrick, , & Fisher, , *The thory of critical phenomena: An introduction to the renormalization group*, Clarendon Press, Oxford, 1992.

Blanchard, A., Buchert, T., & Klaffl, R., *Can the neutrino picture be revived? QSO constraints revisited*, Astron. Astrophys. **267** (1993), L1–L10.

Blaschke, W., *Integralgeometrie. Erstes Heft*, Bernd G. Teubner, Leipzig, Berlin, 1936.

Bode, P. W. & Bertschinger, E., *Accurate calculations of CMB anistropies in cold and mixed dark matter models*, Bull. American Astron. Soc. **186** (1995), no. 08.05, 823.

Bond, J. R. & Efstathiou, G., *The statistics of CBR fluctuations*, Mon. Not. Roy. Astron. Soc. **226** (1987), 655–687.

Bond, J. R., Kofman, L., & Pogosyan, D. Yu., *How filaments of galaxies are woven into the cosmic web*, Nature **380** (1996), 603–605.

Borgani, S., Coles, P., & Moscardini, L., *Cluster correlations in the Zel'dovich approximation*, Mon. Not. Roy. Astron. Soc. **271** (1994), 223–232.

Börner, G., *The early Universe, facts and fiction*, 3rd ed., Springer Verlag, Berlin, 1993.

Bouchet, F. R., Juszkiewicz, R., Colombi, S., & Pellat, R., *Weakly nonlinear gravitational instability for arbitrary Omega*, Ap. J. Lett. **394** (1992), L5–L8.

Bouchet, F. R., Colombi, S., Hivon, E., & Juszkiewicz, R., *Perturbative Lagrangian approach to gravitational instability*, Astron. Astrophys. **296** (1995), 575–608.

Breiman, L., Meisel, W., & Purcell, E., *Variable kernel estimates of multivariate densities*, Technometrics **19** (1977), 135–144.

Bromley, B. C. & Tegmark, M., *Is the cosmic microwave background really non–Gaussian?*, Ap. J. (1999), in press, astro-ph/9904254.

Brösel, R. F., *Werner, eiskalt!*, Semmel Verlach, Kiel, 1986.

Buchert, T. & Bartelmann, M., *High spatial resolution in three dimensions − A challenge for large–scale structure formation models*, Astron. Astrophys. **251** (1991), 389–392.

Buchert, T. & Ehlers, J., *Lagrangian theory of gravitational instability of Friedman–Lemaître cosmologies − second–order approach: an improved model for non–linear clustering*, Mon. Not. Roy. Astron. Soc. **264** (1993), 375–387.

Buchert, T., Karakatsanis, G., Klaffl, R., & Schiller, P., *The performance of Lagrangian perturbation schemes at high resolution*, Astron. Astrophys. **318** (1997), 1–10.

Buchert, T., *A class of solutions in Newtonian cosmology and the pancake theory*, Astron. Astrophys. **223** (1989), 9–24.

Buchert, T., *Galaxy formation in pancake–models*, in: *11th European Regional Astronomy Meeting* (La Laguna, Spain), Vol. 171, IAU, September 1990, pp. 141–145.

Buchert, T., *Lagrangian theory of gravitational instability of Friedmann–Lemaître cosmologies and the "Zel'dovich approximation"*, Mon. Not. Roy. Astron. Soc. **254** (1992), 729–737.

Buchert, T., *Lagrangian theory of gravitational instability of Friedmann–Lemaître cosmologies – a generic third–order model for non–linear clustering*, Mon. Not. Roy. Astron. Soc. **267** (1994), 811–820.

Buchert, T., *Lagrangian perturbation approach to the formation of large–scale structure*, in: *Proceedings of the international school of physics Enrico Fermi. Course CXXXII: Dark matter in the Universe* (Varenna sul Lago di Como) (Bonometto, S., Primack, J., & Provenzale, A., eds.), Società Italiana di Fisica, 1996.

Bunn, E. F. & White, M., *The four–year COBE normalization and large–scale structure*, Ap. J. **480** (1997), 6–21.

Bunn, E. F., Scott, D., & White, M., *The COBE normalization for standard cold dark matter*, Ap. J. Lett. **441** (1995), L9–L12.

Burago, Yu. D. & Zalgaller, V. A., *Geometric inequalities*, Springer Verlag, Berlin, 1988.

Canavezes, A., Springel, V., Oliver, S. J., Rowan-Robinson, M., Keeble, O., White, S. D. M., Saunders, W., Efstathiou, G., Frenk, C., McMahon, R. G., Maddox, S., Sutherland, W., & Tadros, H., *The topology of the IRAS point source catalogue redshift survey*, Mon. Not. Roy. Astron. Soc. **297** (1998), 777–793.

Catelan, P., *Lagrangian dynamics in non–flat universes and non–linear gravitational evolution*, Mon. Not. Roy. Astron. Soc. **276** (1995), 115–124.

Chern, S.-S. & Yien, C.-T., *Sulla formula principale cinematica dello spazio ad n dimensioni*, Bollettino della Unione Matematica Italiana **2** (1940), 434–437.

Chern, S.-S., *A simple intrinsic proof of the Gauss–Bonnet formula for closed Riemannian manifolds*, Ann. of Math. **45** (1944), 747–752.

Colberg, J. M., white, S. M. D., MacFarland, T. J., Jenkins, A., Frenk, C. S., Pearce, F. R., Evrard, A. E., Couchman, H. M. P., Efstathiou, G., Peacock, J. A., & Thomas, P. A., *Galaxy clusters in the hubble volume simulations*, in: *Wide Field Surveys in Cosmology, 14th IAP meeting held May 26-30, 1998, Paris.* (Château de Blois, France) (Colombi, S. & Mellier, Y., eds.), Editions Frontières, 1998, p. 247.

Colberg, J. M., *Parallel supercomputer simulations of cosmic evolution*, Dissertation, Ludwig–Maximilians–Universität München, 1999.

Coleman, P. H. & Pietronero, L., *The fractal structure of the Universe*, Physics Rep. **213** (1992), 311–389.

Coles, P. & Barrow, J. D., *Non–Gaussian statistics and the microwave background radiation*, Mon. Not. Roy. Astron. Soc. **228** (1987), 407–426.

Coles, P., Melott, A. L., & Shandarin, S. F., *Testing approximations for non–linear gravitational clustering*, Mon. Not. Roy. Astron. Soc. **260** (1993), 765–776.

Coles, P., Moscardini, L., Lucchin, F., Matarrese, S., & Messina, A., *Skewness as a test of non-Gaussian primordial density fluctuations*, Mon. Not. Roy. Astron. Soc. **264** (1993), 749–757.

Coles, P., Davies, A., & Pearson, R. C., *Quantifying the topology of large–scale structure*, Mon. Not. Roy. Astron. Soc. **281** (1996), 1375–1384.

Coles, P., Pearson, R. C., Borgani, S., Plionis, M., & Moscardini, L., *The cluster distribution as a test of dark matter models. IV - topology and geometry*, Mon. Not. Roy. Astron. Soc. **294**

(1998), 245–258.

Coles, P., *Statistical geometry and the microwave background*, Mon. Not. Roy. Astron. Soc. **234** (1988), 509–531.

Colley, W. N., Gott III, J. R., & Park, C., *Topology of COBE microwave background fluctuations*, Mon. Not. Roy. Astron. Soc. **281** (1996), L82–L84.

Colley, W. N., *Two-dimensional topology of large-scale structure in the las campanas redshift survey*, Ap. J. **489** (1997), 471–475.

Couchman, H. M. P., Thomas, P. A., & Pearce, F. R., *Hydra: An adaptive-mesh implementation of P3M-SPH*, Ap. J. **452** (1995), 797–813.

Couchman, H. M. P., *Mesh-refined P3M - a fast adaptive n-body algorithm*, Ap. J. **368** (1991), L23–L26.

Crofton, M. W., *On the theory of local probability, applied to straight lines drawn at random in a plane; the methods used being also extended to the proof of certain new theorems in the integral calculus*, Phil. Trans. Roy. Soc. London **158** (1868), 181–199.

Daley, D. J. & Vere-Jones, D., *An introduction to the theory of point processes*, Springer Verlag, Berlin, 1988.

Davé, R., Hellinger, D., Nolthenius, R., Primack, J., & Klypin, A., *Filament statistics: A quantitative comparison of cold + hot and cold dark matter cosmologies with CfA1 data*, Mon. Not. Roy. Astron. Soc. **284** (1996), 607–626.

Davis, M., Huchra, J. P., Latham, D. W., & Tonry, J., *A survey of galaxy redshifts. II – the large scale space distribution*, Ap. J. **253** (1982), 423–445.

de Lapparent, V., Geller, M. J., & Huchra, J. P., *A slice of the universe*, Ap. J. **302** (1986), L1–L5.

Dekel, A. & Rees, M. J., *Physical mechanisms for biased galaxy formation*, Nature **326** (1987), 455–462.

Dekel, A., *Scenarios of biased galaxy formation*, in: *Nearly normal galaxies: From the Planck time to the present; Proceedings of the Eighth Santa Cruz Summer Workshop in Astronomy and Astrophysics* (Faber, S. M., ed.), 1987, pp. 244–254.

Diaferio, A., Kauffmann, G., Colberg, J. M., & White, S. D. M., *Clustering of galaxies in a hierarchical universe: III. mock redshift surveys*, Mon. Not. Roy. Astron. Soc. **307** (1999), 537–552.

Diaferio, A., Kauffmann, G., & White, S. D. M., *Clustering of galaxies in a hierarchical universe: IV.*, Mon. Not. Roy. Astron. Soc. (1999), in preparation.

Diggle, P. J., *Statistical analysis of spatial point patterns*, Academic Press, New York and London, 1983.

Doroshkevich, A. G., Müller, V., Retzlaff, J., & Turchaninov, V., *Superlarge–scale structure in N–body simulations*, Mon. Not. Roy. Astron. Soc. **306** (1999), 575–591.

Doroshkevich, A. G., *Spatial structure of fluctuations and origin of galactic rotation in fluctuation theory*, Astrophysics **6** (1970), 320–330.

Ehlers, J. & Buchert, T., *Newtonian cosmology in Lagrangian formulation: Foundations and perturbation theory*, G. R. G. **29** (1997), 733–764.

Epanechnikov, V. A., *Non–parametric estimation of a multivariate probability density*, Theor. Prob. Appl. **14** (1969), 153–158.

Fadda, D., Slezak, E., & Bijaoui, A., *Density estimation with non–parametric methods*, Astron. Astrophys. Suppl. **127** (1998), 335–352.

Falco, E. E., Kurtz, M. J., Geller, M. J., Huchra, J. P., Peters, J., Berlind, P., Mink, D. J., Tokarz,

S. P., & Elwell, B., *The updated zwicky catalog (uzc)*, Publications of the Astronomical Society of the Pacific **111** (1999), 438–452.

Federer, H., *Curvature measures*, Trans. Amer. Math. Soc. **93** (1959), 418–491.

Fenchel, D., *Inégalités quadratiques entre les volumes mixtes des corps convexes*, C. R. Acad. Sci. Paris **203** (1936), 647–650, in French.

Ferreira, P. G., Magueijo, J. C. R., & Górski, K. M., *Evidence for non–gaussianity in the COBE DMR four year sky maps*, Ap. J. **503** (1998), L1–L4.

Frigo, M. & Johnson, S. G., *The Fastest Fourier Transform in the West*, Tech. report, Massachusetts Institute of Technology, 1997, MIT–LCS–TR–728, `http://theory.lcs.mit.edu/~fftw/`.

Geller, M. J. & Huchra, J. P., *Mapping the Universe*, Science **246** (1989), 897–903.

Gilliam, T. & Jones, T., *Monty Python's The Meaning of Life*, 1983.

Goenner, H., *Einführung in die kosmologie*, Spektrum Akademischer Verlag, Heidelberg, 1994.

Górski, K. M., Hinshaw, G., Banday, A. J., Bennett, C. L., Wright, E. L., Kogut, A., Smoot, G. F., & Lubin, P. M., *On determining the spectrum of primordial inhomogeneity from the COBE DMR sky maps: Results of two–year data analysis*, Ap. J. Lett. **430** (1994), L89–L92.

Górski, K. M., Banday, A. J., Bennett, C. L., Hinshaw, G., Kogut, A., Smoot, G. F., & Wright, E. L., *Power spectrum of primordial inhomogeneity determined from the four–year COBE DMR sky maps*, Ap. J. Lett. **464** (1996), L11–L16.

Gorski, K. M., Wandelt, B. D., Hansen, F. K., Hivon, E., & Banday, A. J., *The HEALPix primer*, Tech. report, Theoretical Astrophysics Center (TAC) Kopenhagen, 1999, `http://www.tac.dk/~healpix/`.

Gott III, J. R., Melott, A. L., & Dickinson, M., *The sponge–like topology of large–scale structure in the Universe*, Ap. J. **306** (1986), 341–357.

Gott III, J. R., Weinberg, D. H., & Melott, A. L., *A quantitative approach to the topology of large–scale structure*, Ap. J. **319** (1987), 1–8.

Gott III, J. R., Park, C., Juszkiewicz, R., Bies, W. E., Bennett, D. P., Bouchet, F. R., & Stebbins, A., *Topology of microwave background fluctuations – theory*, Ap. J. **352** (1990), 1–14.

Hadwiger, H., *Altes und Neues über konvexe Körper*, Birkhäuser, Basel, 1955.

Hadwiger, H., *Vorlesungen über Inhalt, Oberfläche und Isoperimetrie*, Springer Verlag, Berlin, 1957.

Hadwiger, H., *Das Wills'sche Funktional*, Mh. Math. **79** (1975), 213–221.

Hamana, T., *Performance of the optimized post-Zel'dovich approximation for cold dark matter models in arbitrary FLRW cosmologies*, Ap. J. **507** (1998), L1–L4.

Harrison, E. R., *Fluctuations at the threshold of classical cosmology*, Phys. Rev. D **1** (1970), 2726–2730.

Heath, D. J., *The growth of density perturbations in zero pressure Friedmann–Lemaître universes*, Mon. Not. Roy. Astron. Soc. **179** (1977), 351–358.

Heavens, A. F., *Estimating non–Gaussianity in the microwave background*, Mon. Not. Roy. Astron. Soc. **299** (1999), 805–808.

Heine, H., *Die schlesischen Weber*, 1844.

Hinshaw, G., Banday, A. J., Bennett, C. L., Górski, K. M., & Kogut, A., *Three–point correlations in the COBE DMR two–year anisotropy maps*, Ap. J. Lett. **446** (1995), L67–L70.

Hinshaw, G., Banday, A. J., Bennett, C. L., Górski, K. M., Kogut, A., Lineweaver, C. H., Smoot, G. F., & Wright, E. L., *Two–point correlations in the COBE DMR four–year anisotropy maps*, Ap. J. Lett. **464** (1996), L25–L28.

Hivon, E., Bouchet, F. R., Colombi, S., & Juszkiewicz, R., *Redshift distortions of clustering: a*

Lagrangian approach, Astron. Astrophys. **298** (1995), 643–660.

Hobson, M. P., Jones, A. W., & Lasenby, A. N., *Wavelet analysis and the detection of non–Gaussianity in the CMB*, Mon. Not. Roy. Astron. Soc. (1998), submitted, astro-ph/9810200.

Holtzman, J. A., *Microwave background anisotropies and large–scale structure in universes with cold dark matter, baryons, radiation, and massive and massless neutrinos*, Ap. J. Suppl. **71** (1989), 1–24.

Huchra, J. P., Davis, M., Latham, D., & Tonry, J., *A survey of galaxy redshifts. IV - The data*, Ap. J. Suppl. **52** (1983), 89–119.

Huchra, J. P., Geller, M. J., de Lapparent, V., & Corwin Jr., H. G., *The CfA redshift survey – data for the NGP + 30 zone*, Ap. J. Suppl. **72** (1990), 433–470.

Huchra, J. P., Geller, M. J., & Corwin Jr., H. G., *The CfA redshift survey: Data for the NGP + 36 zone*, Ap. J. Suppl. **99** (1995), 391.

Huchra, J. P., Vogeley, M., & Geller, M. J., *The CfA redshift survey: Data for the Sout Galactic Cap*, Ap. J. Suppl. **121** (1999), 287–368.

Jenkins, A., Frenk, C. S., Pearce, F. R., Thomas, P. A., Hutching, R., Colberg, J. M., White, S. D. M., Couchman, H. M. P., Peacock, J. A., Efstathiou, G. P., & Nelson, A. H., *The Virgo consortium: Simulations of dark matter and galaxy clustering*, in: *Dark and Visible Matter in Galaxies* (Sesto Pusteria) (Persic, M. & Salucci, P., eds.), Vol. 117, ASP Conference Series, 1997, p. 348.

Jenkins, A., Frenk, C. S., Pearce, F. R., Thomas, P. A., Colberg, J. M., White, S. D. M., Couchman, H. M. P., Peacock, J. A., Efstathiou, G., & Nelson, A. H., *Evolution of structure in cold dark matter universes*, Ap. J. **499** (1998), 20–40.

Kaiser, N. & Stebbins, A., *Microwave anisotropy due to cosmic strings*, Nature **310** (1984), 391–393.

Kanitscheider, B., *Kosmologie*, 2nd ed., Philipp Reclam jun., Stuttgart, 1991.

Karakatsanis, G., Buchert, T., & Melott, A., *Temporal optimization of Lagrangian perturbation schemes*, Astron. Astrophys. **326** (1997), 873–884.

Kates, R., Kotok, E. V., & Klypin, A. A., *High–resolution simulations of galaxy formation in a cold dark matter scenario*, Astron. Astrophys. **243** (1991), 295–308.

Kauffmann, G., Colberg, J. M., Diaferio, A., & White, S. D. M., *Clustering of galaxies in a hierarchical universe: I. methods and results at $z = 0$*, Mon. Not. Roy. Astron. Soc. **303** (1999), 188–206.

Kauffmann, G., Colberg, J. M., Diaferio, A., & White, S. D. M., *Clustering of galaxies in a hierarchical universe: II. evolution to high redshift*, Mon. Not. Roy. Astron. Soc. **307** (1999), 529–536.

Kellerer, H. G., *Minkowski functionals of Poisson process*, Z. Wahrscheinlichkeitstheorie verw. Gebiete **67** (1984), 63–84.

Kendall, M. G. & Stuart, A., *The advanced theory of statistics*, 3 ed., Vol. 2, Griffin, London, 1973.

Kendall, M. G., *Multivariate analysis*, 2 ed., Griffin, London, 1980.

Kerscher, M., Schmalzing, J., Retzlaff, J., Borgani, S., Buchert, T., Gottlöber, S., Müller, V., Plionis, M., & Wagner, H., *Minkowski functionals of Abell/ACO clusters*, Mon. Not. Roy. Astron. Soc. **284** (1997), 73–84.

Kerscher, M., Schmalzing, J., Buchert, T., & Wagner, H., *Fluctuations in the 1.2 Jy galaxy catalogue*, Astron. Astrophys. **333** (1998), 1–12.

Kerscher, M., *The geometry of second–order statistics – biases in common estimators*, Astron. Astrophys. **343** (1999), 333–347.

Klypin, A. & Shandarin, S. F., *Percolation technique for galaxy clustering*, Ap. J. **413** (1993), 48–58.

Koenderink, J. J., *The structure of images*, Biol. Cybern. **50** (1984), 363–370.

Kogut, A., Banday, A. J., Bennett, C. L., Hinshhaw, G., Lubin, P. M., & Smoot, G. F., *Gaussian statistics of the cosmic microwave background: Correlation of temperature extrema in the COBE DMR two–year sky maps*, Ap. J. Lett. **439** (1995), L29–L32.

Kogut, A., Banday, A. J., Bennett, C. L., Górski, K. M., Hinshaw, G., Smoot, G. F., & Wright, E. L., *Test for non–Gaussian statistics in the DMR four–year sky maps*, Ap. J. Lett. **464** (1996), L29–L34.

Lachièze-Rey, M., *Beyond the Zel'dovich approximation*, Ap. J. **408** (1993), 403–408.

Linde, A. & Mukhanov, V. F., *Nongaussian isocurvature perturbations from inflation*, Phys. Rev. D **56** (1997), 535–539.

Luo, X. & Schramm, D. N., *Kurtosis, skewness, and non–Gaussian cosmological density perturbations*, Ap. J. **408** (1993), 33–42.

Luo, S. & Vishniac, E., *Three–dimensional shape statistics: Methodology*, Ap. J. Suppl. **96** (1995), 001–005.

Marzke, R. O., Huchra, J. P., & Geller, M. J., *The luminosity function of the cfa redshift survey*, Ap. J. **428** (1994), 43.

Marzke, R. O., Geller, M. J., da Costa, L. N., & Huchra, J. P., *Pairwise velocities of galaxies in the CfA and SSRS2 redshift surveys*, A. J. **110** (1995), 477–501.

Mathis et al., H., *??*, in preparation (the German Israeli Foundation collaboration), 1999.

Maurogordato, S. & Lachièze–Rey, M., *Void probabilities in the galaxy distribution – scaling and luminosity segregation*, Ap. J. **320** (1987), 13–25.

Mecke, K. R., Buchert, T., & Wagner, H., *Robust morphological measures for large–scale structure in the Universe*, Astron. Astrophys. **288** (1994), 697–704.

Mecke, K. R., *Integralgeometrie in der Statistischen Physik: Perkolation, komplexe Flüssigkeiten und die Struktur des Universums*, Harri Deutsch, Thun, Frankfurt/Main, 1994.

Melott, A. L. & Shandarin, S. F., *Controlled experiments in cosmological gravitational clustering*, Ap. J. **410** (1993), 469–481.

Melott, A. L., Einasto, J., Saar, E., Suisalu, I., Klypin, A. A., & Shandarin, S. F., *Cluster analysis of the nonlinear evolution of large–scale structure in an axion/gravitino/photino–dominated Universe*, Phys. Rev. Lett. **51** (1983), 935–938.

Melott, A. L., Shandarin, S. F., & Weinberg, D. H., *A test of the adhesion approximation for gravitational clustering*, Ap. J. **428** (1994), 28–34.

Melott, A. L., Buchert, T., & Weiß, A. G., *Testing higher–order Lagrangian perturbation theory against numerical simulations. II: Hierarchical models*, Astron. Astrophys. **294** (1995), 345–365.

Minkowski, H., *Volumen und Oberfläche*, Mathematische Annalen **57** (1903), 447–495, in German.

Mo, H. & Buchert, T., *A statistical discriminator among galaxy samples of different large–scale topology and geometry*, Astron. Astrophys. **234** (1990), 5–19.

Moore, B., Frenk, C. S., Weinberg, D. H., Saunders, W., Lawrence, A., Ellis, R. S., Kaiser, N., Efstathiou, G., & Rowan-Robinson, M., *The topology of the qdot IRAS redshift survey*, Mon. Not. Roy. Astron. Soc. **256** (1992), 477–499.

Morse, M. & Cairns, S. S., *Critical point theory in global analysis and differential topology*, Academic Press, New York and London, 1969.

Moscardini, L., Matarrese, S., Lucchin, F., & Messina, A., *Non-Gaussian initial conditions in cosmological N-body simulations, II - Cold Dark Matter models*, Mon. Not. Roy. Astron. Soc.

248 (1991), 424–438.

Mukhanov, V. F., Feldman, H. A., & Brandenberger, R. H., *The theory of cosmological perturbations*, Physics Rep. **215** (1992), 203–333.

Naselsky, P. D. & Novikov, D. I., *General statistical properties of the cmb polarization field*, Ap. J. (1998), 31–39.

Negele, J. W. & Orland, H., *Quantum many–particle systems*, Addison–Wesley, Reading, MA, 1988.

Novikov, D., Feldman, H. A., & Shandarin, S. F., *Minkowski functionals and cluster analysis for CMB maps*, Int. J. Mod. Phys. **D8** (1999), 291–306.

Padmanabhan, T., *Structure formation in the Universe*, Cambridge University Press, Cambridge, 1993.

Parzen, E., *On estimation of a probability density function and mode*, Ann. Math. Stat. **33** (1962), 1065–1076.

Pauls, J. L. & Melott, A. L., *Hierarchical pancaking: why the Zel'dovich approximation describes coherent large–scale structure in N-body simulations of gravitational clustering*, Mon. Not. Roy. Astron. Soc. **274** (1995), 99–109.

Pearce, F. R. & Couchman, H. M. P., *Hydra: a parallel adaptive grid code*, New Astronomy **2** (1997), 411–427.

Pearce, F. R., Jenkins, A., Frenk, C. S., Colberg, J. M., White, S. D. M., Thomas, P. A., Couchman, H. M. P., Peacock, J. A., & Efstathiou, G., *A simulation of galaxy formation and clustering*, Ap. J. **521** (1999), 99L–102L.

Peebles, P. J. E., *The nature of the distribution of galaxies*, Astron. Astrophys. **32** (1974), 197–202.

Peebles, P. J. E., *The large scale structure of the Universe*, Princeton University Press, Princeton, New Jersey, 1980.

Peebles, P. J. E., *Principles of physical cosmology*, Princeton University Press, Princeton, New Jersey, 1993.

Penzias, A. A. & Wilson, R. W., *A measurement of excess antenna temperature at 4080 $\mathrm{mc/s}$*, Ap. J. **142** (1965), 419–421.

Pirsig, R. M., *Zen and the art of motorcycle maintenance*, The Bodley Head Ltd., United Kingdom, 1974.

Pisani, A., *A non–parametric and scale–independent method for cluster analysis – I. The univariate case*, Mon. Not. Roy. Astron. Soc. **265** (1993), 706–726.

Pisani, A., *A non–parametric and scale–independent method for cluster analysis – II. The multivariate case*, Mon. Not. Roy. Astron. Soc. **278** (1996), 697–726.

Platzöder, M., *Statistische Analyse der Galaxienverteilung mit Hilfe von Minkowski–Funktionalen*, Diplomarbeit, Ludwig–Maximilians–Universität München, 1995.

Plionis, M., Valdarnini, R., & Jing, Y.-P., *Clustering of galaxy cluster. II – Rare events in the cluster distribution*, Ap. J. **398** (1992), 12–32.

Pogosyan, D. Yu., Bond, J. R., Kofman, L., & Wadsley, J., *The cosmic web and filaments in cluster patches*, Bull. American Astron. Soc. **189** (1996), no. 13.03, 1289–1290.

Pompilio, M. P., Bouchet, F. R., Murante, G., & Provenzale, A., *Multifractal analysis of string–induced cosmic microwave background radiation anisotropies*, Ap. J. **449** (1995), 1–8.

Press, W. H., Flannery, B. P., Teukolsky, S. A., & Vetterling, W. T., *Numerical recipes in C*, Cambridge University Press, Cambridge, 1987.

Protogeros, Z. A. M. & Weinberg, D. H., *The topology of large–scale structure in the 1.2 Jy IRAS redshift survey*, Ap. J. **489** (1997), 457.

Queneau, R., *Zazie dans le métro*, Éditions Gallimard, 1959.

Rees, M. J., *Biasing and suppression of galaxy formation*, in: *Nearly normal galaxies: From the Planck time to the present; Proceedings of the Eighth Santa Cruz Summer Workshop in Astronomy and Astrophysics* (Faber, S. M., ed.), 1987, pp. 255–262.

Reiss, R.-D., *A course on point processes*, Springer Verlag, Berlin, 1993.

Sachs, R. K. & Wolfe, A. M., *Perturbations of a cosmological model and angular variations of the microwave background*, Ap. J. **147** (1967), 73–90.

Sahni, V. & Coles, P., *Approximation methods for nonlinear gravitational clustering*, Physics Rep. **262** (1995), 1–135.

Sahni, V., Sathyaprakash, B. S., & Shandarin, S. F., *Probing large-scale structure using percolation and genus curves*, Ap. J. **476** (1997), L1–L5.

Sahni, V., Sathyaprakash, B. S., & Shandarin, S. F., *Shapefinders: A new shape diagnostic for large–scale structure*, Ap. J. **495** (1998), L5–L8.

Santaló, L. A., *Integral geometry and geometric probability*, Addison–Wesley, Reading, MA, 1976.

Sathyaprakash, B. S., Sahni, V., & Shandarin, S. F., *Morphology of clusters and superclusters in N-body simulations of cosmological gravitational clustering*, Ap. J. **508** (1998), 551–569.

Sathyaprakash, B. S., Sahni, V., Shandarin, S. F., & Fisher, K. B., *Filaments and pancakes in the IRAS 1.2 Jy redshift survey*, Ap. J. **507** (1998), L109–L112.

Schmalzing, J. & Buchert, T., *Beyond genus statistics: A unifying approach to the morphology of cosmic structure*, Ap. J. Lett. **482** (1997), L1–L4.

Schmalzing, J. & Diaferio, A., *Topology and geometry of the CfA2 redshift survey*, Mon. Not. Roy. Astron. Soc. (1999), in press, astro-ph/9910228.

Schmalzing, J. & Górski, K. M., *Minkowski functionals used in the morphological analysis of cosmic microwave background anisotropy maps*, Mon. Not. Roy. Astron. Soc. **297** (1998), 355–365.

Schmalzing, J., Kerscher, M., & Buchert, T., *Minkowski functionals in cosmology*, in: *Proceedings of the international school of physics Enrico Fermi. Course CXXXII: Dark matter in the Universe* (Varenna sul Lago di Como) (Bonometto, S., Primack, J., & Provenzale, A., eds.), Società Italiana di Fisica, 1996, pp. 281–291.

Schmalzing, J., Buchert, T., Melott, A., Sahni, V., Sathyaprakash, B. S., & Shandarin, S., *Disentangling the cosmic web I: Morphology of isodensity contours.*, Ap. J. **526** (1999), 568–578.

Schmalzing, J., Gottlöber, S., Kravtsov, A., & Klypin, A., *Quantifying the evolution of higher order clustering*, Mon. Not. Roy. Astron. Soc. **309** (1999), 1007–1016.

Schmalzing, J., *Morphologische Maße für kosmische Strukturen*, Diplomarbeit, Ludwig–Maximilians–Universität München, 1996, in German, English excerpts available.

Schneider, P. & Bartelmann, M., *The power spectrum of density fluctuations in the Zel'dovich approximation*, Mon. Not. Roy. Astron. Soc. **273** (1995), 475–483.

Schneider, R., *Curvature measures of convex bodies*, Ann. Math. pura appl. **116** (1978), 101–134.

Schneider, R., *Convex bodies: the Brunn–Minkowski theory*, Cambridge University Press, Cambridge, 1993.

Scott, D. W. & Terrell, G. R., *Biased and unbiased cross–validation in density estimation*, J. Am. Statist. Assoc. **82** (1987), 1131–1146.

Shandarin, S. F. & Yess, C., *Detection of network structure in the Las Campanas Redshift Survey*, Ap. J. **505** (1998), 12.

Shandarin, S. F. & Zel'dovich, Ya. B., *Topology of the large–scale structure of the universe*, Comments Astrophys. **10** (1983), 33–45.

Shandarin, S. F., *Percolation theory and the cell–lattice structure of the Universe*, Sov. Astron. Lett. **9** (1983), 104–106.

Shane, C. D. & Wirtanen, C. A., *The distribution of extragalactic nebulae*, A. J. **59** (1954), 285.

Siegmund, D. O. & Worsley, K. J., *Testing for a signal with unknown location and scale in a stationary Gaussian random field*, Ann. of Stat. **23** (1995), 608–639.

Silverman, B. W., *Density estimation for statistics and data analysis*, Chapman and Hall, London, 1986.

Smoot, G. F., Bennett, C. L., Kogut, A., Wright, E. L., Aymon, J., Boggess, N. W., Cheng, E. S., De Amici, S., Hauser, M. G., Hinshaw, G., Jackson, P. D., Janssen, M., Kaita, E., Kelsall, T., Keegstra, P., Lineweaver, C., Loewenstein, K., Lubin, P., Mather, J., Meyer, S. S., Moseley, S. H., Murdock, T., Rokke, L., Silverberg, R. F., Tenorio, L., Weiss, R., & Wilkinson, D. T., *Structure in the COBE differential microwave radiometer first-year maps*, Ap. J. **396** (1992), L1–L5.

Smoot, G. F., Tenorio, L., J., B. A., Kogut, A., Wright, E. L., Hinshaw, G., & Bennett, C. L., *Statistics and topology of the COBE differential microwave radiometer first–year sky maps*, Ap. J. **437** (1994), 1–11.

Spergel, D. N. & Pen, U.-L., *Cosmology in a string–dominated universe*, Ap. J. **491** (1997), L67–L71.

Springel, V., White, S. D. M., Colberg, J. M., Couchman, H. M. P., Efstathiou, G. P., Frenk, C. S., Jenkins, A. R., Pearce, F. R., Nelson, A. H., Peacock, J. A., & Thomas, P. A., *Genus statistics of the Virgo N–body simulations and the 1.2Jy redshift survey*, Mon. Not. Roy. Astron. Soc. **298** (1998), 1169–1188.

Stebbins, A., *Cosmic strings and the microwave sky. I - Anisotropy from moving strings*, Ap. J. **327** (1988), 584–614.

Stoyan, D., Kendall, W. S., & Mecke, J., *Stochastic geometry and its applications*, 2nd ed., John Wiley & Sons, Chichester, 1995.

Susperregi, M. & Buchert, T., *Cosmic density and velocity fields in Lagrangian perturbation theory*, Astron. Astrophys. **323** (1997), 295–304.

ter Haar Romeny, B. M., Florack, L. M. J., Koenderink, J. J., & Viergever, M. A., *Scale space: Its natural operators and differential invariants*, in: *Lecture Notes in Computer Science*, Vol. 511, Springer Verlag, Berlin, 1991, pp. 239–255.

Thomas, P. A., Colberg, J. M., Couchman, H. M. P., Efstathiou, G. P., Frenk, C. S., Jenkins, A. R., Nelson, A. H., Hutchings, R. M., Peacock, J. A., Pearce, F. R., & White, S. D. M., *The structure of galaxy clusters in different cosmologies*, Mon. Not. Roy. Astron. Soc. **296** (1998), 1061–1071.

Tomita, H., *Statistics and geometry of random interface systems*, in: *Formation, dynamics and statistics of patterns* (Kawasaki, K., Suzuki, M., & Onuki, A., eds.), Vol. 1, World Scientific, 1990, pp. 113–157.

Torres, S., Cayón, L., Martínez–González, E., & Sanz, J. L., *Genus and spot density in the COBE DMR first year anisotropy maps*, Mon. Not. Roy. Astron. Soc. **274** (1995), 853–857.

Torres, S., *Topological analysis of COBE-DMR cosmic microwave background maps*, Ap. J. Lett. **423** (1994), L9–L12.

Totsuji, H. & Kihara, T., *The correlation function for the distribution of galaxies*, Publications of the Astronomical Society of Japan **21** (1969), 221.

Vanselow, M., Diplomarbeit, Ludwig–Maximilians–Universität München, 1995, in German.

Vogeley, M. S., Geller, M. J., Park, C., & Huchra, J. P., *Voids and constraints on the nonlinear*

clustering of galaxies, A. J. **108** (1994), 745–758.

Vogeley, M. S., Park, C., Geller, M. J., Huchra, J. P., & Gott III, J. R., *Topological analysis of the CfA redshift survey*, Ap. J. **420** (1994), 525.

Weil, W., *Stereology: A survey for geometers*, in: *Convexity and its applications* (Gruber, P. M. & Wills, J. M., eds.), Birkhäuser, Basel, 1983, pp. 360–412.

Weinberg, S., *Gravitation and cosmology*, John Wiley & Sons, Chichester, 1972.

Weiß, A. G. & Buchert, T., *High–resolution simulation of deep pencil beam surveys – analysis of quasi–periodicty*, Astron. Astrophys. **274** (1993), 1–11.

Weiß, A. G., Gottlöber, S., & Buchert, T., *Optimizing higher order Lagrangian perturbation theory for standard CDM and BSI models*, Mon. Not. Roy. Astron. Soc. **278** (1996), 953–964.

White, M. & Bunn, E. F., *The COBE normalization of CMB anisotropies*, Ap. J. **450** (1995), 477–485.

White, M. & Bunn, E. F., *The COBE normalization of CMB anisotropies: Erratum*, Ap. J. **460** (1996), 1071.

White, S. D. M., *The hierarchy of correlation functions and its relation to other measures of galaxy clustering*, Mon. Not. Roy. Astron. Soc. **186** (1979), 145–154.

White, M., *Higher order moments of the density field in a parametrized sequence of non–Gaussian theories*, Mon. Not. Roy. Astron. Soc. (1999), to appear, astro-ph/9811227.

Wicksell, S. D., *The corpuscle problem I.*, Biometrika **17** (1925), 84–99.

Winitzki, S. & Kosowsky, A., *Minkowski functional description of microwave background Gaussianity*, New Astronomy **3** (1997), 75–100.

Worsley, K., *Local maxima and the expected Euler characteristic of excursion sets of χ^2, f and t fields.*, Adv. Appl. Prob. **26** (1994), 13–42.

Worsley, K., *Testing for a signal with unknown location and scale in a χ^2 random field, with an application to fMRI*, Adv. Appl. Prob. (1998), accepted, `http://www.math.mcgill.ca/~keith/chisq/`.

Yamamoto, K. & Suto, Y., *Two-point correlation function of high-redshift objects: An explicit formulation on a light-cone hypersurface*, Ap. J. **517** (1999), 1–12.

Yess, C. & Shandarin, S. F., *Universality of the network and bubble topology in cosmological gravitational simulations*, Ap. J. **465** (1996), 2–13.

Yess, C., Shandarin, S. F., & Fisher, K. B., *Percolation analysis of a Wiener reconstruction of the IRAS 1.2 Jy redshift catalog*, Ap. J. **474** (1997), 553–560.

Zel'dovich, Ya. B. & Shandarin, S. F., *The maximum density in heavy–neutrino clouds*, Sov. Astron. Lett. **8 (3)** (1982), 139–140.

Zel'dovich, Ya. B., *Fragmentation of a homogeneous medium under the action of gravitation*, Astrophysics **6** (1970), 164–174.

Zel'dovich, Ya. B., *Gravitational instability: An approximate theory for large density perturbations*, Astron. Astrophys. **5** (1970), 84–89.

Zel'dovich, Ya. B., *A hypothesis, unifying the structure and the entropy of the Universe*, Mon. Not. Roy. Astron. Soc. **160** (1972), 1P–3P.

Zentsova, A. S. & Chernin, A. D., *Evolution of entropy perturbations in the post–recombination epoch. II – Nonlinear stage*, Astrophysics **16** (1980), 108–113.

Danksagung

Name	Jens Schmalzing
Geburtsdatum	11. Juni 1971
Geburtsort	Nürnberg
Familienstand	verheiratet mit Sabine Vahlensieck
	Sohn Jonathan geboren am 20. Februar 1994
	Sohn Johannes geboren am 9. Dezember 1998

Schulbildung

1977 – 1981	Gerhart-Hauptmann-Schule Nürnberg
1981 – 1990	Martin-Behaim-Gymnasium Nürnberg
29. Juni 1990	**Abitur** – sehr gut
	Leistungskurse Mathematik und Physik
	Grundkurse Englisch und Geologie

Studium der Physik

1990 – 1996	Ludwig-Maximilians-Universität München
1993 – 1994	University of Oxford – visiting research student
11. Mai 1992	**Vordiplom** – sehr gut
23. September 1996	**Diplom** – mit Auszeichnung
	Betreuer Prof. Herbert Wagner
	Thema „Morphologische Maße für kosmische Strukturen"

Promotion in Astrophysik

1996 – 1999	Ludwig-Maximilians-Universität München
	Betreuer Priv.-Doz. Thomas Buchert und Prof. Gerhard Börner
	Thema „On Statistics and Dynamics of Cosmic Structure"

Stipendien und Auszeichnungen

1990 – 1996	Hochbegabtenförderung des Freistaates Bayern
1990 – 1996	Stiftung Maximilianeum
1992 – 1996	Studienstiftung des deutschen Volkes
1996 – 1999	Promotionsstipendium des Max-Planck-Instituts für Astrophysik

Berufserfahrung

1992 – 1998	Übungsbetreuer an der Universität München
1998 – 1999	Systemverwalter an der Universität München