LAURI JUHAN LIIVAMÄGI

Properties and spatial distribution of
galaxy superclusters

TARTU ÜLIKOOL · UNIVERSITAS TARTUENSIS · 1632

DISSERTATIONES ASTRONOMIAE UNIVERSITATIS TARTUENSIS

18

LAURI JUHAN LIIVAMÄGI

Properties and spatial distribution of galaxy superclusters

This study was carried out at the Tartu Observatory, Estonia.

The Dissertation was admitted on December 6, 2016, in partial fulfilment of the requirements for the degree of Doctor of Philosophy in physics (astrophysics), and allowed for defence by the Council of the Institute of Physics, University of Tartu.

Supervisor: Dr. Enn Saar,
                 Tartu Observatory,
                 Tõravere, Estonia

Opponents: Dr. Maciej Bilicki,
                 Leiden Observatory, Leiden University,
                 Leiden, Netherlands

                 Dr. Andi Hektor
                 National Institute of Chemical Physics and Biophysics,
                 Laboratory of High Energy and Computational Physics
                 Tallinn, Estonia

Defense: February 21, 2017, University of Tartu, Estonia

CONTENTS

6

LIST OF ORIGINAL PUBLICATIONS

This thesis is based on the following publications:

    I **Liivamägi, L. J.**, Tempel, E., Saar, E. 2012, *SDSS DR7 superclusters. The catalogues*, Astronomy & Astrophysics, 539, A80

   II Einasto, M., **Liivamägi, L. J.**, Tago, E., Saar, E., Tempel, E., Einasto, J., Martínez, V. J., Heinämäki, P. 2011, *SDSS DR7 superclusters. Morphology*, Astronomy & Astrophysics, 532, A5

  III Einasto, M., **Liivamägi, L. J.**, Saar, E., Einasto, J., Tempel, E., Tago, E., Martínez, V. J. 2011, *SDSS DR7 superclusters. Principal component analysis*, Astronomy & Astrophysics, 535, A36

Other related publications of the dissertant:

  IV Tempel, E., Tago, E., **Liivamägi, L. J.** 2012, *Groups and clusters of galaxies in the SDSS DR8. Value-added catalogues*, Astronomy & Astrophysics, 540, A106

   V Einasto, M., **Liivamägi, L. J.**, Tempel, E., Saar, E., Vennik, J., Nurmi, P., Gramann, M., Einasto, J., Tago, E., Heinämäki, P., Ahvensalmi, A., Martínez, V. J. 2012, *Multimodality of rich clusters from the SDSS DR8 within the supercluster-void network*, Astronomy & Astrophysics, 542, A36

  VI Tempel, E., Saar, E., **Liivamägi, L. J.**, Tamm, A., Einasto, M., Einasto, J., Müller, V. 2011, *Galaxy morphology, luminosity, and environment in the SDSS DR7*, Astronomy & Astrophysics, 529, A53

 VII Einasto, M., Tago, E., Saar, E., Nurmi, P., Enkvist, I., Einasto, P., Heinämäki, P., **Liivamägi, L. J.**, Tempel, E., Einasto, J., Martínez, V. J., Vennik, J., & Pihajoki, P. 2010, *The Sloan Great Wall. Rich clusters*, Astronomy & Astrophysics, 522, A92

VIII Einasto, M., Lietzen, H., Tempel, E., Gramann, M., **Liivamägi, L. J.**, Einasto, J. 2014, *SDSS superclusters: morphology and galaxy content*, Astronomy & Astrophysics, 562, A87

  IX Lietzen, H.; Tempel, E.; **Liivamägi, L. J.**; Montero-Dorta, A.; Einasto, M.; Streblyanska, A.; Maraston, C.; Rubiño-Martín, J. A.; Saar, E. 2016, *Discovery of a massive supercluster system at z ∼ 0.47*, Astronomy & Astrophysics, 588, L4

Author's contribution to the publications

Author has made considerable contributions to the following original publications. The following list gives details on the author's work in each of the papers. The Roman numerals correspond to those in the list of publications.

   Publication I. The author prepared the data and performed all the necessary calculations to create the catalogues. He implemented a method to delineate superclusters and carried out the analysis of the catalogues. The author outlined the paper structure, wrote the majority of the text and drew all the figures.

   Publication II and III. The author prepared and provided all the source data regarding superclusters. He took part of the analysis in both publications and wrote parts of the text.

INTRODUCTION

Astronomy is a science that can offer plenty of unforgettable imagery, and the large-scale distribution of galaxies is no exception. An example is shown in the figure on the next page. As one can see, it is a strikingly intricate foamlike structure with great variations in its overall density. However, aside from being a sight to lose yourself in, it also reflects on the fundamental properties of our universe. The picture we see, i.e, why are the galaxies positioned where they are and why do they look like they do, are a result of multiple phenomenona, the behaviour of the space-time, the matter-energy content of the universe, the conditions in the beginning, the many laws of physics and the effect they have on the ingredients of the universe. Bond et al. (1996) referred to this characteristic pattern of large-scale inhomogeneities as the "cosmic web", and that name has stuck since then.

Among the first features the viewer's eye is likely to be drawn to, are large concentrations of galaxies, contrasting to the seemingly empty regions beside them. These are commonly called galaxy superclusters. Superclusters can extend from tens to over hundred megaparsecs, they contain from hundreds to thousands of galaxies, and many galaxy groups and clusters. Unlike galaxy clusters, superclusters are clearly unrelaxed systems, not gravitationally bound as crossing times exceed the age of the universe, and show little to no radial symmetry (Oort 1983). Although they may be visually distinctive, galaxy superclusters unfortunately lack an universally applicable definition and their physical nature is also open for discussion.

Superclusters, as part of the large-scale structure, are sensitive to the initial power spectrum and the following evolution, and can thus be a tool to discriminate among cosmological models and galaxy formation scenarios (Park et al. 2012). They are massive enough to leave an imprint on the cosmic microwave background radiation by the integrated Sachs-Wolfe effect (Nadathur & Crittenden 2016). Superclusters can provide an unique environment for their constituent galaxies (Giovanelli et al. 1986) and galaxy clusters (Einasto et al. 2012).

This thesis is an empirical work, intended to be a retrospective of more recent studies of superclusters at the Tartu Observatory, a series that goes back several decades. The thesis is structured as following. The first chapter gives an overview of the cosmological background of the topic; we also cover some observational basics and shortly sum up the history of supercluster reseach. In the second chapter, we describe in detail the method for creating supercluster catalogues – preparation of the source data samples, calculation of density fields, using density fields to delineate the superclusters and obtaining their properties. A closer examination of supercluster catalogues is presented in the third chapter.
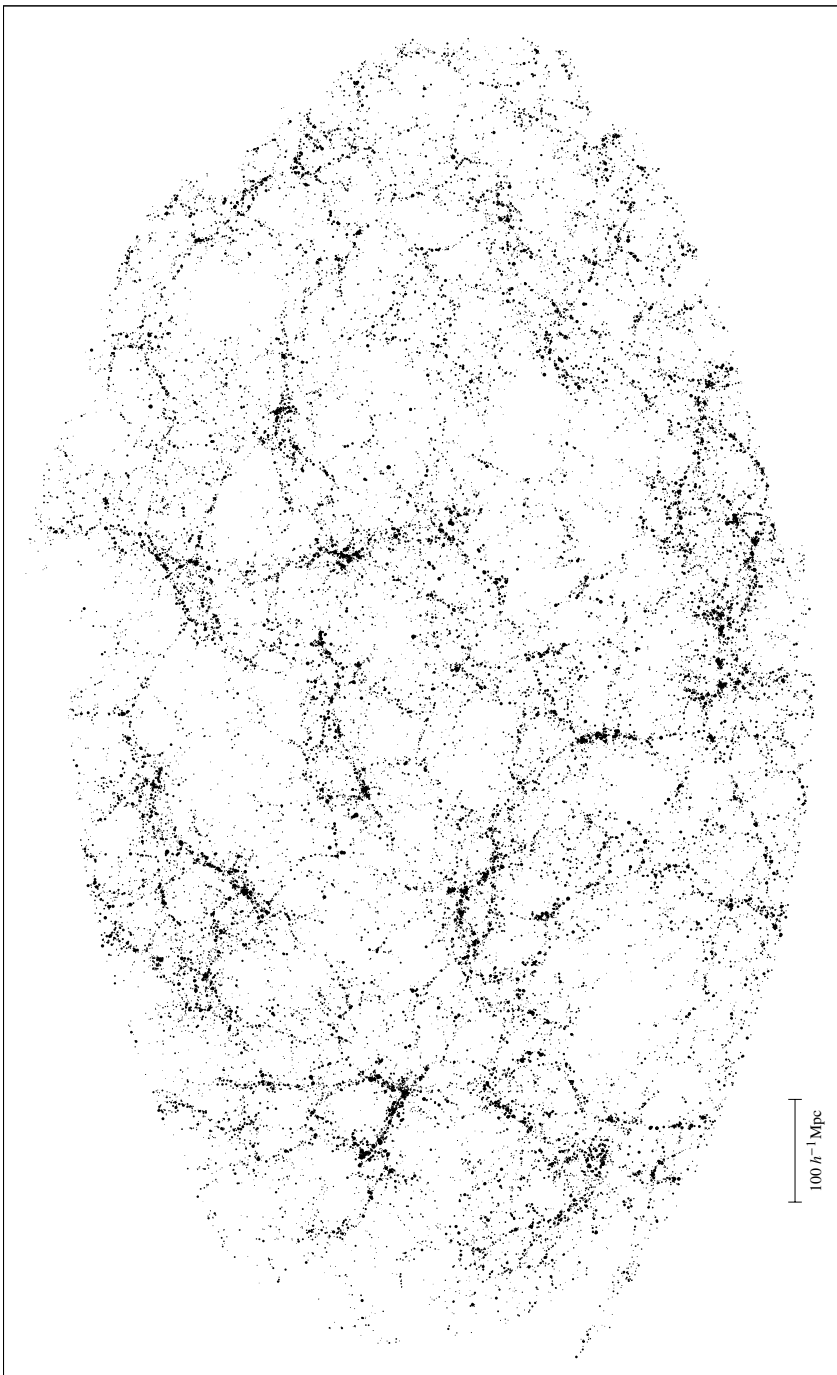
Figure 1.1: Large-scale galaxy distribution in the universe. A 20 h^{-1}Mpc thick slice from the SDSS main galaxy sample. *All figures in the thesis have been created with the* gnuplot *utility*.

1 BACKGROUND AND HISTORY

The large-scale structure of the galaxy distribution is characterised by large voids and a complex web of galaxy walls, filaments and clusters. In this chapter we outline the theory of structure formation and also present a cursory overview of observations of large-scale structures. We cover the primary observables and discuss their acquisition. Last, we describe a number of studies of superclusters, concentrating more on those which deal with defining superclusters and creating supercluster samples.

1.1 Formation of the large-scale structure

First we give a brief introduction to the basic cosmology and the formation of large-scale structures in the universe using Martínez & Saar (2002) and Mo et al. (2010) as reference. The basis of the current standard model of cosmology is a combination of the cosmological principle and the Einstein's theory of general relativity. Numerous observations imply that the universe is experiencing global expansion, it had a very hot and compact beginning (the Big Bang), and it has several components with very distinct properties – baryonic matter, dark matter, dark energy, etc. In this environment, the matter content of the universe has been shaped, mostly by gravity, into structures we can observe today.

1.1.1 Standard model of cosmology

General relativity is a theory of gravity, where gravity is the result of the local curvature of space. The geometry of the four-dimensional space-time is described by its metric. The cosmological principle states that the universe is spatially homogeneous and isotropic. Following the cosmological principle hypothesis we arrive at the Robertson-Walker metric

$$ds^2 = c^2 dt^2 - a^2(t) \left[\frac{dr^2}{1 - kr^2} + r^2(d\theta^2 + \sin^2 \theta \, d\phi^2) \right], \qquad (1.1)$$

where r, θ and ϕ are called comoving coordinates and the function $a(t)$ is the time-dependent scale factor. The curvature parameter k determines the global geometry of the universe. The curvature parameter can take three different values from $\{-1, 0, 1\}$, which correspond to models with open, flat or closed geometry. Using this form of the metric, the comoving coordinates correspond to a spherical coordinate system, where θ and ϕ are the angular position and r the radial coordinate (as is the standard astronomical practice). The scale factor describes the overall expansion of the

space, meaning that distances between objects that have constant comoving coordinates change in time as $a(t)$. By convention, the scale factor $a(t)$ is dimensionless and normalised to unity at the present epoch, $a(t_0) = 1$.

General relativity determines the relation between the space-time geometry of the universe and its physical content. The connection is given by the Einstein equation:

$$G_{\mu\nu} - \Lambda g_{\mu\nu} = \frac{8\pi G}{c^4} T_{\mu\nu}. \tag{1.2}$$

Here, the left side of the equation characterises the geometry of the universe. The Einstein tensor $G_{\mu\nu}$ describes the spacetime curvature, the metric is denoted by $g_{\mu\nu}$, and Λ is the cosmological constant. On the right hand side, $T_{\mu\nu}$ is the energy-momentum tensor, and G is the Newton's gravitational constant.

Using the Robertson-Walker metric, the Einstein equations can be reduced to the Friedmann equations:

$$\begin{aligned}
\frac{\ddot{a}}{a} &= -\frac{4\pi G}{3}(\rho + 3p/c^2) + \frac{\Lambda c^2}{3}, \\
\frac{\dot{a}^2}{a^2} &= \frac{8\pi G}{3}\rho + \frac{\Lambda c^2}{3} - \frac{kc^2}{a^2},
\end{aligned} \tag{1.3}$$

which describe the dynamical evolution of the scale factor. The variable ρ is the total energy density, p is the pressure and the two are linked by the equation of state $p = p(\rho)$. The main components of the energy density ρ are baryonic matter, non-baryonic (dark) matter and radiation. Similarly, the cosmological constant may represent the energy density of vacuum. Each of the components has a corresponding equation of state and they dominate different epochs in the evolution of the universe. The ratio $H(t) = \dot{a}(t)/a(t)$ is called the Hubble function and its value in the present epoch is the Hubble constant H_0. Even though the Friedmann equations were derived before the expansion of the universe was actually observed, they strongly inferred that the universe should not be static and has to be in an expanding (or contracting) state. For the present epoch, the Friedmann cosmological models are commonly parameterised by the three following variables:

$$\Omega_M = \frac{8\pi G\rho_0}{3H_0^2}, \quad \Omega_\Lambda = \frac{\Lambda c^2}{3H_0^2}, \quad \Omega_K = \frac{kc^2}{a_0^2 H_0^2}, \tag{1.4}$$

where Ω_M and Ω_Λ are the density parameters of matter and vacuum energy, respectively. The parameter Ω_K, which is not actually a *density* parameter, corresponds to the curvature of three-space. The second Friedmann equation in Eq. 1.3 then gives us $\Omega_K = 1 - \Omega_M - \Omega_\Lambda$ to describe the relation between the curvature and density parameters. The density parameters Ω_M and Ω_Λ, together with the Hubble constant

H_0, fully determine the evolution of the universe. Different parameter values can give very different estimates for, e.g., the age of the universe, and also its fate in the future – from collapsing to an eternally accelerating expansion. Current observations, however, agree rather consistently that the values for the density parameters are: $\Omega_M \approx 0.31$, $\Omega_\Lambda \approx 0.69$, the curvature $\Omega_K \approx 0.0$, i.e. the universe seems to be geometrically flat, and the Hubble constant $H_0 \approx 70$ km/(s Mpc) (Planck Collaboration et al. 2015). The matter component of the universe is dominated by cold dark matter $\Omega_{\text{CDM}} \approx 0.26$ and the contribution of baryonic matter is $\Omega_b \approx 0.05$.

1.1.2 Structure formation and evolution

The standard cosmological model we just presented describes the evolution of the universe as a whole. It arises from the cosmological principle of isotropy and homogeneity, which may well hold at sufficiently large scales, while at smaller scales the universe is obviously full of various prominent structures. Therefore, a departure from uniformity is needed at some stage of the evolution of the universe. Standard cosmological models as described above, being based on the theory of relativity, do not provide a mechanism for the emergence of initial perturbations from which the structures would grow. However, in very early times, general relativity is expected to be inadequate for wholly describing the universe and quantum effects have to be taken into account. It is assumed that the first primordial density deviations were generated by quantum fluctuations. At the very early stages, the universe also went through a phase of a rapid exponential expansion called inflation. While the exact physical nature of inflation remains unclear, several models can nonetheless successfully predict the resulting, now macroscopic, density perturbations consistent with the structures observable in today's universe. In addition to sctructure formation, inflation offers more or less satisfactory solutions to several cosmological problems like, e.g., the large-scale homogeneity of the cosmic microwave background, and the geometric flatness of the universe. After inflation, the universe settled to expanding in concordance with the Friedmann equations as described before.

The framework for the development of structures in this environment is provided by the gravitational instability theory. The majority of matter in the universe can be described as a non-relativistic fluid. This model is valid for both baryonic and dark matter. Fluctuations of the density $\rho(\mathbf{x})$ can be conveniently described with the so-called density contrast

$$\delta(\mathbf{x}) \equiv \frac{\rho(\mathbf{x}) - \rho_m}{\rho_m}, \tag{1.5}$$

where $\mathbf{x}$ are the comoving coordinates, and ρ_m is the average background density. In the expanding universe, the evolution of the gravitating fluid is described by the following set of equations. First, to maintain the conservation of mass, we need the

equation of continuity:

$$\frac{\partial \delta}{\partial t} + \frac{1}{a} \nabla [(1 + \delta)\mathbf{v}] = 0, \tag{1.6}$$

where $\mathbf{v} \equiv a\dot{\mathbf{x}}$ is the peculiar velocity. The conservation of momentum is described by the Euler equation

$$\frac{\partial}{\partial t}(a\mathbf{v}) + (\mathbf{v}\nabla)\mathbf{v} = -\nabla\Phi - \frac{\nabla P}{\rho_0(1 + \delta)}, \tag{1.7}$$

and the time-dependent gravitational potential Φ can be found by solving the Poisson equation

$$\nabla^2 \Phi = 4\pi G \rho_0 a^2 \delta, \tag{1.8}$$

In addition, the equation of state is needed to specify the fluid pressure P. Combining Eq. 1.6–1.8, one can obtain the equation for the evolution of the density contrast $\delta(\mathbf{x})$. Specific cosmological models are inserted into these equations via the scale factor $a(t)$.

Focussing on dark matter, which is dominant on large scales, allows making significant simplifications, e.g., for the collisionless dust model all pressure terms can be omitted. The actual evolution of structure is commonly outlined in three stages. We start with the *initial conditions*, i.e. the starting distribution of density perturbations. The initial density perturbations form a random field and thus, one can also regard the perturbations as a realisation of a random process and to describe it statistically. Observations strongly imply that the initial perturbations should be very well approximated by a homogeneous and isotropic random Gaussian field. A Gaussian field is also predicted by most common inflation models. In that case, the density perturbation field can be wholly described with the power spectrum $P(k) \equiv V_u \langle |\delta_{\mathbf{k}}|^2 \rangle$, where $\delta_{\mathbf{k}}$ refer to the Fourier modes of the density field and V_u is a normalisation constant. The simplest form for the power spectrum is a power law

$$P(k) \propto k^n, \tag{1.9}$$

where n denotes the spectral index. The special case with $n = 1$ is called the Harrison-Zeldovich or scale-invariant spectrum, where gravitational potential perturbations have the same amplitude at all scales. This initial distribution of perturbations was first suggested by Harrison (1970) and Zeldovich (1972) to consistently explain galaxy formation, and, it was later also found to be a natural result of inflationary scenarios.

In an universe dominated by non-relativistic matter, density perturbations are gravitationally unstable and will grow with time. Even slight overdensities are inevitably attracting matter from the less dense neighbouring regions, which in time

14

will become increasingly rarefied. The growth of instabilities would be exponential in a static universe. In an expanding universe, however, accretion is slowed and the exact growth rate depends on the cosmological model. While the density variations are still very low $\delta\rho/\rho_0 \ll 1$, the perturbations are said to grow in the *linear regime*. Leaving out the second-order and pressure terms, which are negligible, one can combine equations Eq. 1.6–1.8 to derive the equation for the evolution of density perturbations:

$$\frac{\partial^2 \delta}{\partial t} + 2\frac{\dot{a}}{a}\frac{\partial \delta}{\partial t} = 4\pi G\rho_0 \delta. \tag{1.10}$$

The general solution contains two independently evolving modes, one of which is growing and the other one decays quickly. The growing mode is commonly called the linear growth factor $D_1(t)$, since the density perturbation amplitude changes as the initial density $\delta(\mathbf{x})$ multiplied by the time-dependent coefficient:

$$\delta(\mathbf{x}, t) = D_1(t)\delta(\mathbf{x}). \tag{1.11}$$

The linear growth rate depends on the cosmological models and can be expressed as a function of redshift as

$$D_1(z) = E(z) \int_z^\infty \frac{(1 + z')dz'}{E^3(z')}, \tag{1.12}$$

where $E(z)$ is the normalised Hubble function (see Sect. 1.1.1), given here as a function of redshift:

$$E(z) = \left[\Omega_M(1 + z)^3 + (1 - \Omega_M - \Omega_\Lambda)(1 + z)^2 + \Omega_\Lambda\right]^{\frac{1}{2}}. \tag{1.13}$$

As the perturbation amplitudes grow and become comparable to unity, $\delta\rho/\rho_0 \sim 1$, overdense regions separate from the overall expansion of space and start to collapse. The governing equations cannot be linearised anymore, and the structure evolution moves on to what is called the *non-linear regime*. In general, the equations become too complex to solve analytically, when higher order effects must be taken into account. However, there still exist empirical models that are capable of describing the non-linear evolution.

On large scales, structure formation can be described by the so-called *Zeldovich approximation* introduced by Zeldovich (1970). It uses the Lagrangian approach, describing the displacements of mass elements (or particles) rather than density. Zeldovich argued that, having determined the initial coordinates and velocities of mass elements, the evolution of structures can be extrapolated using the linear theory (as formulated above) into the regime, where matter has already been transported considerably from the original location and the density perturbations are no longer small.

When matter is displaced from the initial position $\mathbf{x}^\star$ to the final location $\mathbf{x}$, the corresponding density can be found with the following formula:

$$1 + \delta = \left| \frac{\partial \mathbf{x}}{\partial \mathbf{x}^\star} \right|^{-1} = \frac{1}{(1 - \lambda_1 D_1)(1 - \lambda_2 D_1)(1 - \lambda_3 D_1)}. \qquad (1.14)$$

Here, the coefficients $\lambda_1 \geqslant \lambda_2 \geqslant \lambda_3$ denote the three eigenvalues of the deformation tensor $\partial_j \partial_k \left(\Phi^\star / 4\pi G \rho_0 a^3 \right)$. Crucially, we can see that according to the approximation, collapse amplifies any initial anisotropy of the matter distribution and contraction is locally dominated by one dimension at a time, which is determined by the eigenvalues. This leads to the formation of flattened structures, often named *Zeldovich pancakes*, i.e. walls, and later filaments and knots, easily distinguishable in observations and numerical simulations (see Figures 1.1 and 1.2). The Zeldovich formalism thus concludes that the morphology of the overall structure is embedded in the tidal field.

The formation of structure on scmaller scales can be described by spherical collapse. In models with dominating cold dark matter, structure formation follows a "bottom-up" scenario, meaning that smaller clumps of matter collapse and virialise earlier, providing progenitors for larger structures. The process is called hierarchical clustering (Peebles 1971) and it is happening concurrently with the large-scale structure formation described above (Bond et al. 1996).

On the one hand, lengths of galaxy superclusters (in the order of tens to a hundred of megaparsecs) place them firmly in the linear regime of the structure formations (e.g., Zeldovich et al. 1982). However, due to their asymmetrical shapes, other dimensions of supercluster can be considerably smaller and thus be influenced by non-linear processes.

Most of the structure formation can be approximately described using dark matter, as the baryonic (i.e. regular matter) has, for several reasons, a much weaker overall impact. First, baryons constitute a much smaller fraction of the total energy balance. Also, they remain coupled with radiation until recombination and the start of the gravitational evolution is delayed compared to dark matter. On large scales, baryon acoustic oscillations are a small but detectable feature in the galaxy distribution that are caused by oscillations in the pre-recombination matter-radiation medium. Baryonic matter becomes important on smaller scales (comparable to galaxy sizes) where the local density of baryons can become equal or greater than that of dark matter.

Numerical computer simulations are without doubt the most powerful tool for studying the large-scale structures and their formation. Numerical models develop the full view of cosmic web as a complex pattern of sheets, filaments, and roughly spherical small density peaks (knots), framing large under-dense void regions (Doroshkevich et al. 1980; Klypin & Shandarin 1983; Davis et al. 1985; Sheth & van
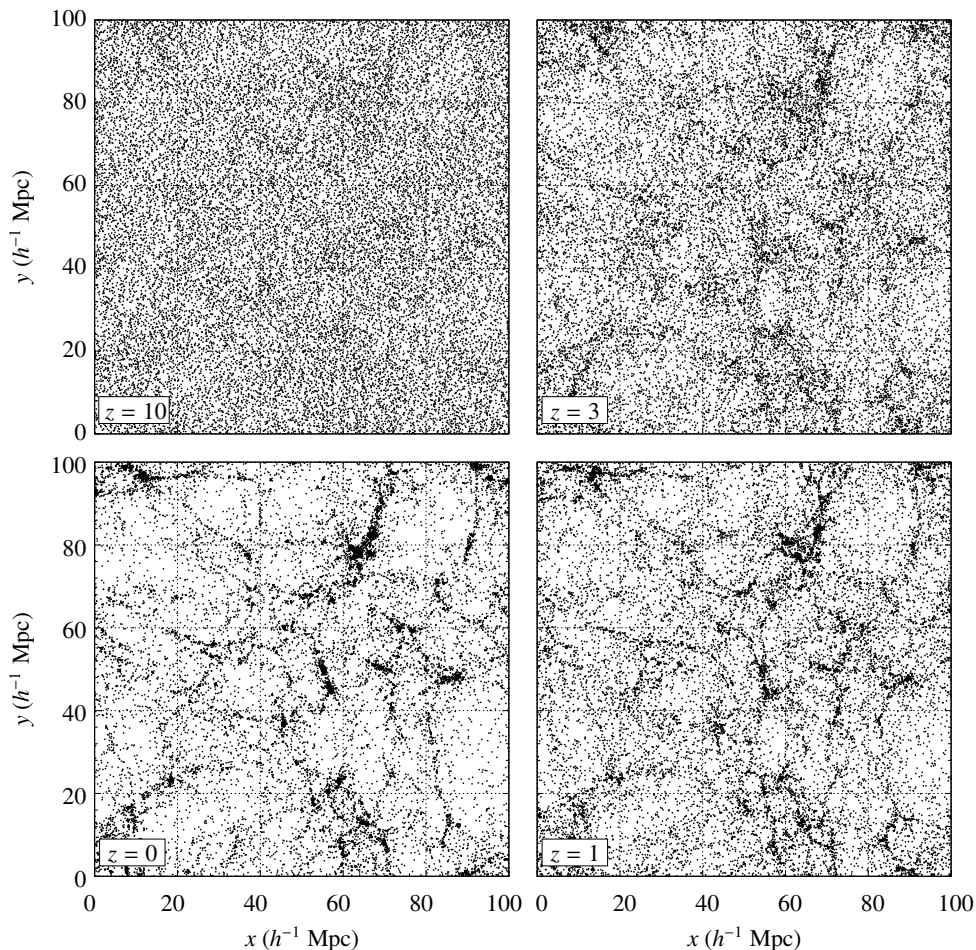
Figure 1.2: Example of the formation of large-scale structure in a dark matter only numerical simulation with the standard ΛCDM cosmological model (Suhhonenko et al. 2011). Images represent N-body simulation snapshots taken at redshifts $z = 10$, 3, 1 and 0 (moving clockwise from the *top-left panel*). Dots denote dark matter particles.

de Weygaert 2004). Within this network, matter flows along well-defined paths, with voids feeding sheets, sheets feeding filaments, and filaments finally channeling everything towards knots (see Cautun et al. 2014, for an overview). Computer simulations also offer great flexibility, as one can define a problem and attempt to combine all the relevant knowledge – different types of particles (dark matter, baryons), and the known physical processes (gravity, gas dynamics), to solve it (e.g., Vogelsberger et al. 2014).

Figure 1.2 shows four snapshots of a N-body simulation from Suhhonenko et al.

(2011). Even this simple and relatively small example displays the important characteristics of the large-scale structure of the universe. As described above, one can see the emergence of large flattened structures as well as the formation of haloes and their coalescence into larger clusters thereafter.

1.2   Observations of large-scale structures

In reality, observations of large-scale structures mean observing the galaxies. The structures themselves are then inferred from the acquired galaxy samples. Therefore, the analysing the large-scale structures is essentially based on making maps of the galaxy positions. Thus, the minimal data set we can start with is a sample of galaxy coordinates. The simplest estimator for the galaxy size is its luminosity. We will now cover the basics of obtaining both.

1.2.1   Galaxy coordinates and luminosities

It is fairly easy to measure the coordinates of celestial objects in the sky with reasonable accuracy. Various astrometric techniques allow automatic calculation of sky coordinates from the images taken with one's instrument (galaxies, being extended objects, may require more complex fitting than stars). Finding the third coordinate – distance, is much less straightforward. To determine distances using stellar parallaxes, i.e. fully geometrically, is possible only for the closer stars in our own galaxy, the Milky Way. To estimate the distances for more far-away objects, one has to utilise the so-called cosmological "distance ladder", where several methods are incrementally used. All these estimators are intercompared to reduce errors and scatter, and to build a consistent distance scale. The methods include motion of stellar clusters, main sequence fitting of stellar clusters, the period-luminosity relations for variable stars, e.g., the Cepheids, and other objects with known intrinsic brightnesses either from theory or observations, e.g., type Ia supernovae, which are commonly named "standard candles". Having an independent way to obtain absolute luminosities of certain objects is a powerful tool for finding distances, because one can then use the well-known distance modulus formula

$$m - M = 5\log_{10} d - 5, \tag{1.15}$$

where m is the apparent, i.e. observed magnitude, M the absolute magnitude, and d the distance to the source in parsecs. In addition, galaxies themselves have also other observable properties that are correlated to their absolute luminosities, e.g., the Tully-Fisher (Tully & Fisher 1977) relation for spiral, and the fundamental plane relation (Faber & Jackson 1976; Djorgovski & Davis 1987) for elliptical galaxies. Both of these tie the galaxy luminosities to the motion of matter in them.

In large galaxy surveys, which are the primary data for analysing the large-scale structure, the standard way to deduce distances is using redshifts. The redshift is the observed shift of the wavelength of the emitted radiation:

$$z_{obs} = \frac{\lambda_{obs} - \lambda_{em}}{\lambda_{em}}, \tag{1.16}$$

where λ_{em} is the true and λ_{obs} the observed wavelength. Galaxy redshifts are the net result of the expansion of the universe, galaxy motion and gravitation. The contribution from the latter is usually ignored, being much smaller than the first two. Then, following Harrison (1974), the observed redshift can be broken down as:

$$
\begin{aligned}
1 + z_l &= \frac{\lambda_{obs}}{\lambda_l} = \left(\frac{c - v_l}{c + v_l}\right)^{1/2}, \\
1 + z_c &= \frac{\lambda_l}{\lambda_c} = \frac{1}{a(t)}, \\
1 + z_p &= \frac{\lambda_c}{\lambda_{em}} = \left(\frac{c + v_p}{c - v_p}\right)^{1/2}.
\end{aligned}
\tag{1.17}
$$

The first contribution z_l is the wavelength shift resulting from our own peculiar motion with the velocity v_l. The second is the cosmological redshift z_c, where $a(t)$ corresponds to the scale factor at the time of emission. The last contribution is again due to the peculiar velocity v_p, when the emitting galaxy itself is moving with a velocity different from that of the general expansion. The first and third shifts occur because of the Doppler effect, while the second one is caused by the expansion of the universe, which continously stretches all wavelengths as the radiation passes across space.

Combining the equations in Eq. 1.16 and 1.17 gives us

$$1 + z_{obs} = (1 + z_l)(1 + z_c)(1 + z_p). \tag{1.18}$$

The contribution from the local Doppler shift is usually corrected using the cosmic microwave background dipole anisotropy, leaving us with

$$1 + z_{obs}^* = (1 + z_c)(1 + z_p), \tag{1.19}$$

where z_{obs}^* is now the CMB rest frame corrected redshift. After multiplication it becomes $1 + z_{obs}^* = 1 + z_c + z_p + z_p z_c$. If redshifts are small i.e. $z_c \ll 1$ and $z_p \ll 1$, the observed redshift can be approximated as

$$z_{obs}^* \approx z_c + z_p. \tag{1.20}$$

In our case (see Sect. 2.1 for the data description), these conditions are satisfied. The problem of separating the contributions of both components to the redshift can be handled in many ways.

The most primitive way to calculate distance using redshift, is to assume that the whole measured redshift is due to the Doppler effect, i.e. $z = v/c$ (approximately valid for $z < 0.1$). The Hubble law then states that the line-of-sight recession velocity v is proportional to the source distance (Hubble & Humason 1931), i.e. when using our notation the law becomes $v = H_0\, d$. At larger distances, proper calculation of distances is more complicated and depends on the cosmological model. Comoving distances are commonly used for spatial analysis because their values do not depend on the value of the scale factor at the specific time. Looking at the metric formula in Eq. 1.1 and the definition of the cosmological redshift in Eq. 1.17, we can see that it is possible to find a relation between the observed redshift and the comoving distance. We can derive the following formula for the comoving distance (in units of length)

$$r(z) = \frac{c}{H_0} \int_0^z \frac{dz'}{E(z')},$$
(1.21)

where the expansion of the universe is described by the Hubble function (1.13).

Correct distances are also necessary to calculate the absolute luminosities of galaxies. Knowing the galaxy distance, we can return to Eq. 1.15 and calculate its absolute magnitude M. From the absolute magnitude we can find the luminosity of galaxy

$$L = L_\odot \cdot 10^{\frac{M_\odot - M}{2.5}},$$
(1.22)

where $M_\odot$ and $L_\odot$ are the absolute magnitude and the luminosity of the sun. Galaxy luminosities are also usually presented in solar units $L_\odot$.

1.2.2 Redshift surveys

Since redshift is the most robust distance estimator for faraway galaxies, specific observation programs to create three-dimensional maps of the universe were designed. From the 1950s to 1970s the first galaxy catalogues complemented with redshifts were published, containing hundreds and later a few thousands of objects (Giovanelli & Haynes 1991). Before that, large galaxy surveys like the well-known catalogues by, e.g., Shapley & Ames (1932) or Shane & Wirtanen (1954), had hardly any distance information at all. One of the first large systematic surveys of galaxy redshifts was the The *Center for Astrophysics* (CfA) survey in the beginning of the 1980s. With its successor CfA2, they contained over $10\,000$ galaxy redshifts, and resulted in the first picture of the large-scale structure that clearly showed us all the major components of the cosmic web – voids, filaments, clusters and superclusters (de Lapparent et al. 1986). In addition to optical spectroscopy, 21 cm radio observations were
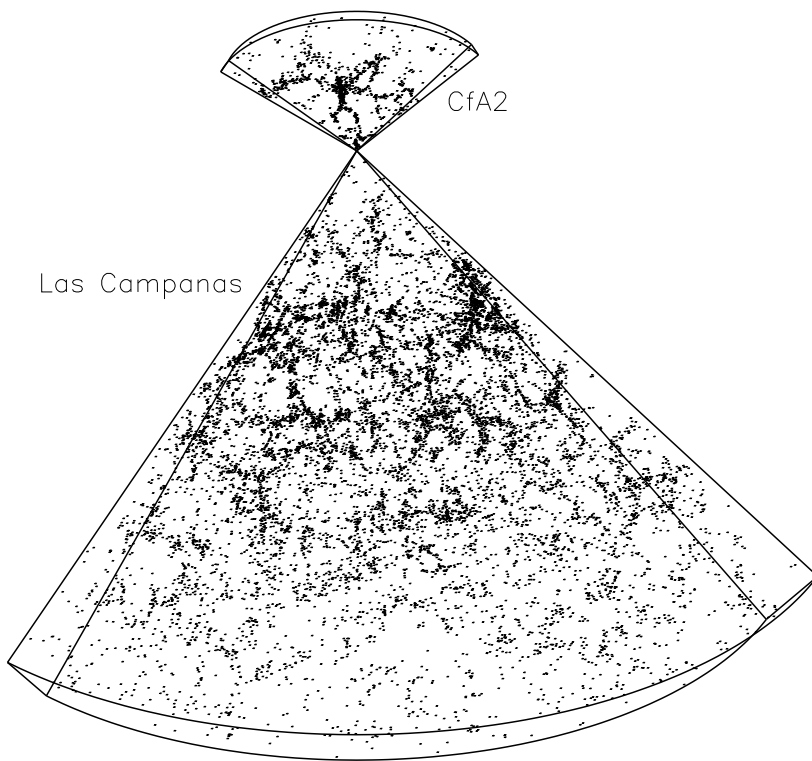
Figure 1.3: The galaxy distribution for the southern slices of the Las Campanas redshift survey and the CfA2 galaxy catalogue in the northen hemisphere. Reproduced from Martinez (1999).

used to measure redshifts, although in smaller quantities (e.g., Chincarini et al. 1983). The *Las Campanas Redshift Survey* (LCSR), carried out in the mid-1990s, consisted of over 23 000 galaxies in several thin slices. It was much deeper than any of the previous surveys and showed that distant large-scale structures are similar to those already discovered (Fig. 1.3). The quantitative leap forward came with the *2 degree Field Galaxy Redshift Survey* (2dFRS) that pushed the number of measured redshifts into hundreds of thousands (Colless et al. 2003). For the last decade, arguably, the most important redshift measuring program has been the *Sloan Digital Sky Survey* (SDSS). It has produced more than 3 million redshifts in several consecutive phases. Major galaxy samples include the so-called main galaxy sample and the Luminous Red Galaxy sample (both used in this work), the more recent *Baryon Oscillation Spectroscopic Survey* (BOSS) CMASS sample of luminous galaxies, and also the quasar sample. Each of these cover different redshift intervals moving from smaller

to larger distances. Redshift surveys like the *IRAS Point Source Catalog* (PSCz) or the *2 Micron All Sky Survey* (2MASS), which are unique for their wide sky coverage due to the lower galactic extinction in infrared wavebands (Saunders et al. 2000; Huchra et al. 2012) are also worth mentioning. While containing less than 50 000 galaxies and being relatively shallow, the 2MASS sample covers almost the whole sky except around the Galactic plane. A smaller, but still important galaxy catalogue is *Cosmicflows-2* (Tully et al. 2013), which is a compilation of over 8 000 independently acquired galaxy distances and redshifts, thus also giving peculiar velocities. There are naturally many more surveys in addition to those mentioned; however, most of them are rather specialised.

A redshift survey can be characterised by its sky coverage and depth. Theoretically, these are imposed by the survey instrumentation, but in practice however, they are set to obtain a fair sample of objects that would satisfy the scientific aims of the survey. The depth of the survey is fixed by apparent magnitude limits outside of which galaxies are not visible to the telescope or not selected for observation. Obviously, the resulting catalogue will not be uniform as intrinsically brighter galaxies are visible to greater distances. The survey properties depend on the direction in the sky, e.g., the high extinction of light near the plane of the Milky Way strongly hinders observations in a large proportion of the sky. Additionally, the completeness of the survey is also affected by interfering stars, technical issues with the instruments, or even the prevailing weather conditions during the observations.

Due to the nature of spectroscopy, obtaining galaxy spectra with a sufficient resolution and signal-to-noise ratio can require lots of observing time. Thus, conducting a redshift survey always means finding a compromise between the exposure time for individual objects and the total number of galaxies. In order to be useful for studies of the large-scale structure, surveys usually need to encompass a significant volume in space as well as to maintain a reasonable number density of galaxies.

The biggest bottleneck in redshift surveys is obtaining all of the galaxy spectra, although modern multifiber spectrographs can take up to thousands of measurements in one pointing. Alternatively, it is also possible to estimate the redshift from a series of photometric measurements in different filters. As for now, these require several assumptions about the source galaxy and are much more inaccurate. However, the future photometric galaxy surveys with a large number of filters like, e.g., the *Javalambre Physics of the Accelerating Universe Astrophysical Survey* (J-PAS), using over 50 narrow-band filters, are going to approach the precision of spectroscopic observations and will be suitable for detailed studies of the large-scale structure (Benitez et al. 2014). Simpler instrumentation, easier target selection and much faster operation form their added benefits.

1.3 Observational studies of superclusters

Only a handful of galaxies are visible to the naked eye, which made the discovery and observations of galaxies in greater numbers possible only after the proliferation of telescopes. The well-known *Catalogue of Nebulae and Star Clusters* published by C. Messier in the 18th century contained 40 objects now known to be galaxies. The majority of over thirteen thousand objects in the *New General Catalogue* (NGC) and *Index Catalogue* (IC) by J.L.E Dreyer (1888–1907) were actually galaxies. However, the extragalactic nature of these "nebulae" was not determined until the 1920s (e.g., Öpik 1922)[1].

That the galaxy distribution is not entirely homogeneous and exhibits clustering properties became also apparent with the first larger galaxy catalogues (Abell 1965; Oort 1983, and references therein). Notably, the nearby massive Virgo cluster is already evident in the earliest surveys of nebulae by W. and C. Herschel (which were the precursors for the NGC and IC). As stated before, the earliest larger systematic surveys of galaxies were compiled using photographic plates and were mostly two-dimensional. The distribution of galaxy clusters was also found not to be random as hints of even a more encompassing assembling of galaxies emerged in the 1930s (e.g., Shapley 1930; Zwicky 1937). Based on the existing observations at the time, Zwicky (1938) reached the conclusion that practically all galaxies probably belong to some aggregations. After studying his catalogue of about 2700 rich galaxy clusters, Abell (1958) suggested that the distribution of clusters indeed shows a real tendency toward "second-order" clustering, and later determined several such systems.

The first supercluster identified as such, however, was the Local supercluster (also called the Virgo supercluster) described by de Vaucouleurs (1953), who argued that the Virgo cluster togther with the Local Group, the Ursa Major cluster, and other smaller groups and clusters comprise a large flattened system with a length of tens of megaparsecs. About that time the term "supercluster" appeared, replacing "super-galaxy" which was in use before. Subsequently, the galaxy distribution in the Local supercluster and surrounding it was mapped in greater detail, benefitting from the accumulating redshift information (Fisher & Tully 1981). After that, several other superclusters were identified, most notably the Coma supercluster (Gregory & Thompson 1978), the Hercules supercluster (Tarenghi et al. 1978), with some evidence of the bridge between the two (Oort 1983). The Perseus-Pisces supercluster (Jõeveer et al. 1978) was assembled using galaxies and galaxy clusters from the cluster catalogue by Zwicky et al. (1961–1968). The massive Hydra-Centaurus supercluster (Chincarini & Rood 1979), also neighbouring the Local Supercluster but located opposite

---

[1]See also Trimble (1995) for a description of "the Great Debate" between H. Curtis and H. Shapley on the scale of the Universe in 1920.
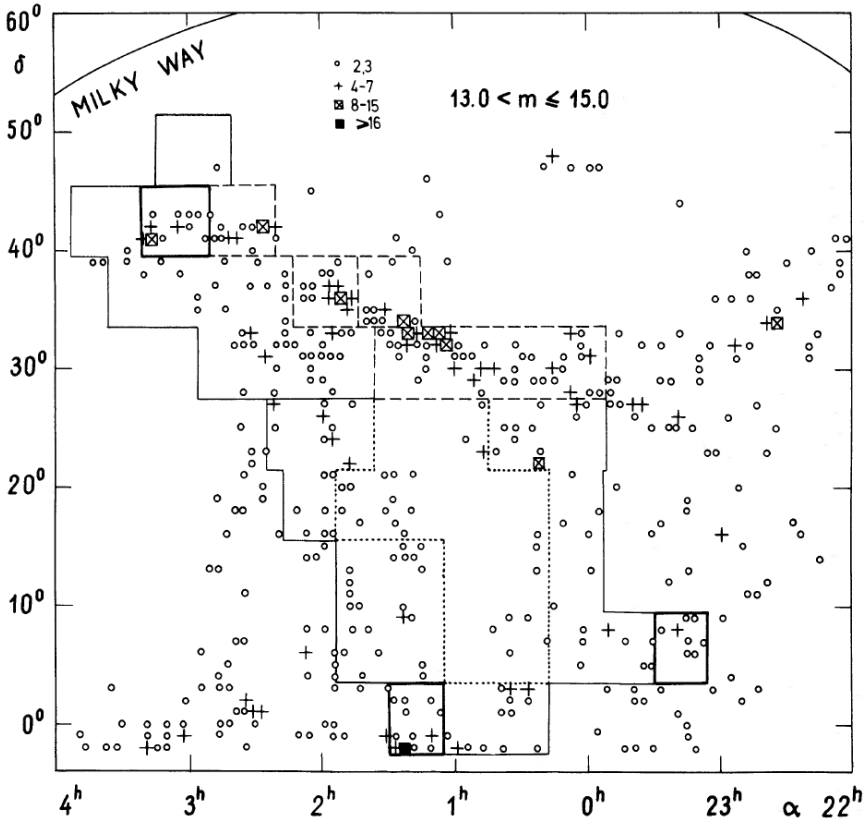
Figure 1.4: Galaxy distribution in the Perseus supercluster. Different symbols represent galaxy counts in 1x1 degree fields. Reproduced from Einasto et al. (1980).

to the Perseus-Pisces supercluster, was later marked by Lynden-Bell et al. (1988) as the approximate centre for the local galaxy bulk flows – the "Great Attractor". Also contributing to the local bulk motion is the very massive Shapley supercluster, which lies roughly in the same direction but is situated further away (Raychaudhury 1989; Branchini et al. 1999), and perhaps other structures in the difficult-to-observe regions behind the Galactic plane (Kraan-Korteweg et al. 2016).

Together with the growing number of superclusters, large zones of opposite nature were also noticed, overdense regions were accompanied by considerably more sparse "holes" in the galaxy distribution (Jõeveer et al. 1978; Kirshner et al. 1981). The studies of nearby superclusters revealed an increasingly complex internal structure, thus changing the view of superclusters from being somewhat vague "clusters of clusters" (Abell 1961) to large-scale filamentary and wall-like structures (Einasto

24

et al. 1980; Davis et al. 1982). Einasto et al. (1980) also concluded that neighbouring superclusters are truly in contact and form an interlocking network that surrounds the galaxy voids. The term "supercluster-void network" was later suggested and used in, e.g., Einasto et al. (1997b), to describe this "cellular" view of the large-scale galaxy distribution. Einasto et al. (1997a) reported a periodicity in the spatial distribution of superclusters; however, in later studies with more abundant data this perceived regularity dissolved.

Before the deeper galaxy redshift surveys were published in the late 1990s, the most convenient way to probe large-scale structures further away was to use galaxy clusters. Cluster samples, while being obviously much sparser than those of galaxies, can maintain better homogeneity to far greater distances. For example, the afore-mentioned Abell catalogue of rich clusters has been the basis for creation of many supercluster catalogues, from the first by Abell (1961) himself up to the 2000s (e.g., Einasto et al. 1994, 1997b, 2001), in the latter case together with X-ray clusters. Rich superclusters from the catalogue by Einasto et al. (1997b) are shown in Fig. 1.5. During the last decade, X-ray cluster samples have also grown large enough to be used to create supercluster catalogues (e.g., Chon et al. 2014). For example, six groupings of galaxy clusters were recently identified as superclusters in a deep XXL X-ray survey (Koulouridis et al. 2016). Another class of relatively rare but bright objects, visible to very large distances, are quasars. The use of quasars to characterise the large-scale structures had been suggested already by Oort (1983). However, the structures de-rived from the SDSS quasar sample by Clowes et al. (2013) and Einasto et al. (2014b) appear considerably larger than what is commonly regarded as superclusters.

The most common method to assemble structures in sparser samples has been the well-known friends-of-friends algorithm (Press & Davis 1982; Zeldovich et al. 1982). The friends-of-friends method has also been used in galaxy samples, e.g., by Basilakos (2003) to compile superclusters from the SDSS main galaxy sample, and by Park et al. (2012) to compare structures in the SDSS galaxy sample and N-body simulations. While it is rather straightforward to implement, results of this method can be very sensitive to the arbitrary choice of the linking parameter and, therefore, strongly affected by noise (the density is essentially estimated at the galaxy locations and determined by two data points).

With greater density of data points, one can turn to gridded smoothed density fields to describe and analyse the galaxy distribution. Galaxy structures like super-clusters can then be delineated by drawing isodensity surfaces. A number of filtering methods have been developed to map the galaxy density to a grid. Many early studies used counting the number of galaxies in equal area squares on the plane of the sky (most famously perhaps in the so-called *Lick counts* by Shane & Wirtanen 1954). An extremely widespread method is to calculate the density field on a regular grid using
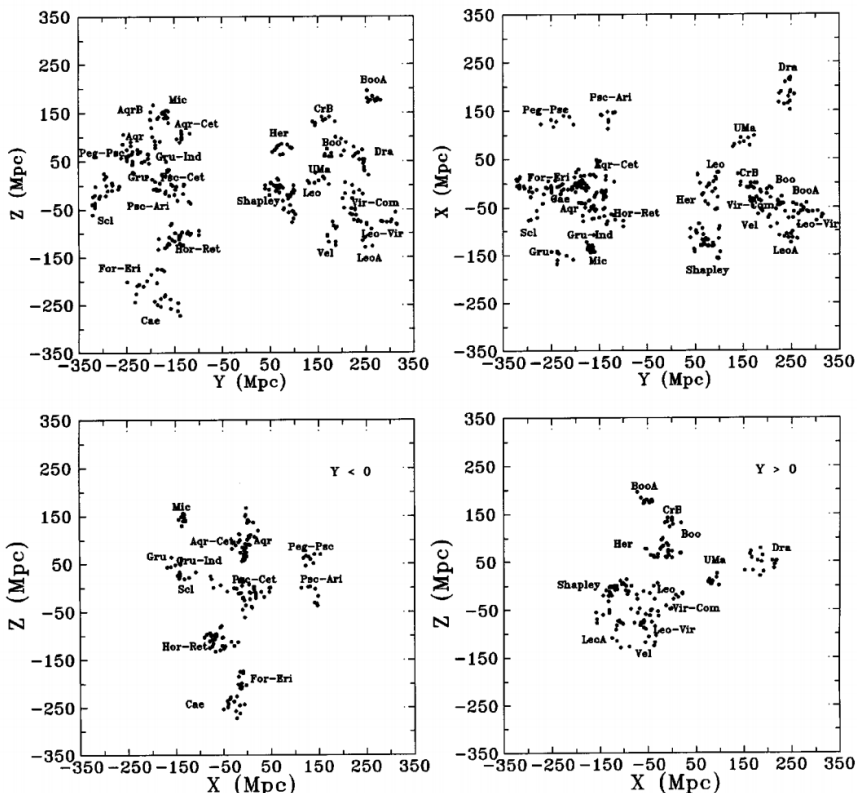
Figure 1.5: Rich superclusters as delineated by rich galaxy clusters. Reproduced from Einasto et al. (1997b).

some sort of smoothing kernel, e.g., the Gaussian being a popular choice. Basilakos et al. (2001) used the PSCz catalogue to calculate the galaxy density field and create superclusters. In Einasto et al. (2003a,b) the LCSR and early version of SDSS galaxy catalogues were used to calculate two-dimensional density fields from which superclusters were extracted. In case of the LCSR sample, galaxy luminosities were utilised as weights. Einasto et al. (2007b) assembled superclusters from the 2dFRS galaxy luminosity density field and in a similar way from the Millennium N-body simulation. Luminosity density fields were also used in the studies by Costa-Duarte et al. (2011) and Luparello et al. (2011). In all these studies the distributions of resulting objects on several density levels were studied before settling to what was deemed an optimal density threshold.

Besides kernel smoothing, Wiener filtering technique has been used to calculate

density fields and find superclusters in the PSCz survey (Schmoldt et al. 1999) and the 2dFRS catalogue (Erdoğdu et al. 2004). Wiener filtering requires a prior model corresponding to the expected properties of the data, it can take the selection and incompleteness effects into account by adjusting the resolution.

One of the disadvantages of kernel smoothing is that it imposes its own properties, like the shape and scale, onto the analysis and the resulting density field reflects them as well as those of the input data. In a different approach, Kitaura & Enßlin (2008) presented a method employing the Bayesian inference to estimate matter density and also peculiar velocity fields from the galaxy samples. It has been applied to a number of galaxy surveys to calculate density fields and extract the cosmic web elements, e.g., the SDSS main (Jasche et al. 2010) and CMASS samples (Ata et al. 2016), but also elsewhere (Kitaura et al. 2012; Granett et al. 2015). So-called natural, parameter-free density estimation methods also exist that use the Delauney or its dual, the Voronoi tesselation (see, e.g., Schaap 2007, for a detailed overview). Tesselation-based density estimators are widely used in describing the large-scale galaxy distribution, especially the cosmic web classification (e.g., Aragón-Calvo et al. 2010; Platen et al. 2011; Sousbie 2011; Cautun et al. 2014; Koulouridis et al. 2016), although they seem to be more popular in studies of numerical simulations. Nadathur & Hotchkiss (2014); Nadathur & Crittenden (2016) created supercluster catalogues of the SDSS main, LRG and later also CMASS samples using a modification of the ZOBOV algorithm. ZOBOV (Neyrinck 2008) was first deleveoped as a watershed-type void-finding algorithm using the Voronoi tesselation; however, it can also be used to detect "positive" structures like clusters and superclusters by inversing the density field.

Above-listed studies analyse supercluster multiplicity and luminosity functions with respect to both galaxies and clusters. Many of them are dedicated to describing the supercluster shape distribution. As this is also the method described in this work, some of the studies will be brought up again later in the thesis, where we will compare their results or methods to ours.

Most of the studies mentioned above do not have an entirely physical definition for supercluster. Some utilise the spherical collapse condition to derive the boundary density threshold, since simulations suggest that while supercluster-size objects are not gravitationally bound at the moment, they will become that in the future (Araya-Melo et al. 2009). However, this approach is not completely accurate for all cases as superclusters are highly asymmetrical at the current epoch. Proper investigation of the dynamical state of the large-scale structure is needed in order to determine the gravitational potential and galaxy velocities, i.e. what is actually bound to what. Currently, the observational precision needed to measure adequately the peculiar motions of galaxies is limited to our close neighbourhood. Tully et al. (2014) studied

the nearby large-scale structure in the Cosmicflows-2 catalogue using the Wiener algorithm, applying it on both the galaxy density and velocity fields. They were able to map the galaxy density and motions up to 150 h^{-1}Mpc distance from us and to delineate the so-called basins of attraction – volumes of space where all the galaxy flows inside it are *converging*. Natural boundaries of superclusters are therefore the surfaces where the flows *diverge* into separate potential depressions. This brought on a considerable reinterpretation of the local large-scale structure. As a result, the Local (Virgo) Supercluster was lumped together with the Hydra-Centaurus supercluster and several other smaller nearby superclusters forming an extensive structure named the Laniakea supercluster, with the "Great Attractor" region being the convergence point for the matter flow.

This short overview is certainly not complete and there are many other studies concerning galaxy superclusters. There are also several studies that are based on or related to the catalogue presented in this thesis (Liivamägi et al. 2012), which we will list and briefly describe in Sect. A.2.

2 BUILDING SUPERCLUSTERS USING LUMINOSITY DENSITY FIELDS

In this chapter we present our procedure of generating supercluster sets from the galaxy distribution. A simplified description of the supercluster creating procedure is as follows. We consider superclusters as large-scale enhancements in the density field that is dominated by dark matter. Supposing that the bias (the ratio of the dark matter density to the stellar density) is approximately constant on supercluster scales, the observational counterpart for the total density is the luminosity density. Density fields are calculated by interpolating the luminosity of galaxies into a predefined grid. Superclusters are extracted from the field by fixing a density threshold and grouping the grid cells with higher values into continuous structures. Several thresholds are used to obtain multiple sets of density field objects. We also discuss the question of choosing reasonable density thresholds and a method to find "natural" density limits for individual objects. Later we assign galaxies to density field objects according to their location in the field.

2.1 Galaxy and group data

In this section we describe the data samples used and the prepatory steps we take before constructing the density fields. We constructed our catalogues using both the Sloan Digital Sky Survey (SDSS) main and Luminous Red Galaxy (LRG) flux-limited samples. The main sample has a high spatial density and allows to follow the superclusters in detail, while the LRG sample, although sparse, is much deeper. The locations and luminosities of galaxies are affected by a number of observational effects, of which some can be taken into account and corrected. When analysing observations, it is also useful to apply the same methods in parallel to simulated data sets. For this purpose, we drew a comparable sample from one of the galaxy catalogues based on the Millennium N-body simulation. Throughout the whole work the following cosmological parameter values are assumed: we denote the Hubble constant as $H_0 = 100h$ km s^{-1} Mpc^{-1} and take the matter density $\Omega_m = 0.27$, and the dark energy density $\Omega_\lambda = 0.73$.

2.1.1 The SDSS main galaxy sample

Our main galaxy sample is the main sample from the 8th data release (DR8) of the (Aihara et al. 2011a). We used the data from the contiguous 7221 square degree area in the North Galactic Cap, the so-called Legacy Survey (see the sky coverage
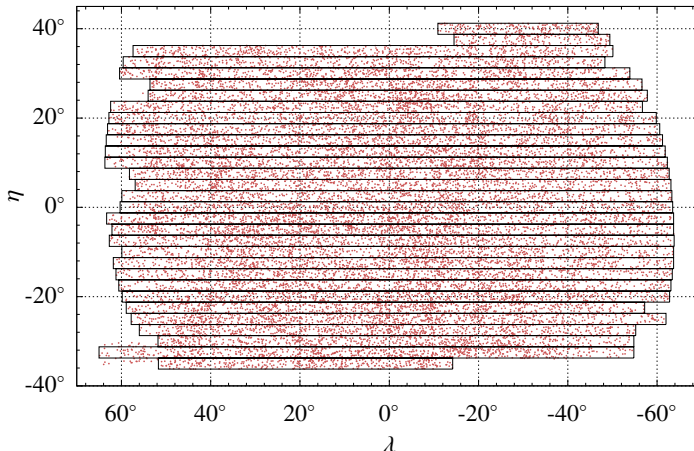
Figure 2.1: The sky projection of the DR8 galaxies and the simplified survey mask in the SDSS η and λ coordinates.

in Fig. 2.1). The sample selection is described in detail in the SDSS DR8 group catalogue paper by Tempel et al. (2012).

As the first step, we selected objects which had both photometric and spectroscopic data and were classified as galaxies in the Catalog Archive Server of the SDSS. The SDSS spectroscopic galaxy sample is complete for objects brighter than the apparent r-band magnitude $m_r = 17.77$, which we use as the lower limit. This magnitude limit was imposed after Galactic extinction correction, yielding an uniform sample. We also had to apply the upper limit $m_r = 12.5$ since bright nearby galaxies are unobservable due to the sensor saturation. The upper magnitude limit affects the sample up to distances $d = 60\ h^{-1}\mathrm{Mpc}$. The survey completeness is further affected by several technical limitations to the observations. We estimate that a total of 8% of the galaxy pairs are missing due to so-called fiber collisions (Tempel et al. 2012). This results from the fact that the minimum distance between the spectrograph fibers in the focal plane is 55″ (Strauss et al. 2002). The galaxy sample also needed to be cleaned of duplicate entries which where identified by comparing the redshifts and angular coordinates of closely located galaxies. In some cases visual inspection was needed. Also, a number of objects with unusual properties (extreme brightness or unphysical colours), were visually examined and excluded if necessary. Most of those were oversaturated and misclassified stars, but there were also other imaging artifacts. We corrected the redshifts of galaxies for the motion relative to the cosmic microwave background and computed comoving distances of galaxies (see, e.g., Martínez & Saar 2002). We put the lower redshift limit to $z = 0.009$ in order to exclude the local supercluster and the upper limit to $z = 0.2$ to avoid distant sparse regions. After all selections, the final sample contains 576493 galaxies.
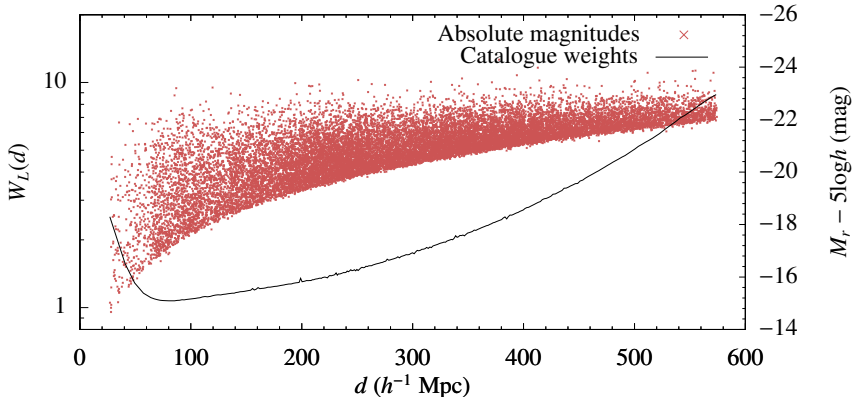
Figure 2.2: Absolute magnitudes M_r of the main sample galaxies (red dots) and the distance-dependent weights for the luminosity correction (solid line).

We calculated the absolute magnitudes of galaxies in the r-band as $M_r = m_r - 25 - \log_{10} d_L - K$, where m_r is the Galactic extinction corrected apparent magnitude, $d_L = d(1 + z)$ is the luminosity distance (d is the comoving distance) in h^{-1} Mpc and z the redshift, and $K = k + e$ is the sum of k and evolutionary corrections. The k-correction for the SDSS galaxies was calculated using the KCORRECT (v4_2) algorithm (Blanton & Roweis 2007). In addition, we corrected the magnitudes for evolution, using the luminosity evolution model of Blanton et al. (2003). The magnitudes correspond to the rest frame (at the redshift $z = 0$). The calculated galaxy magnitudes are shown in Fig. 2.2. As it is natural for a flux-limited sample, only brighter galaxies are included at larger distances.

Groups and clusters of galaxies were determined using a modified friends-of-friends (FoF) algorithm. With the FoF method, galaxies form a group if each of them has at least one neighbouring galaxy closer than a certain distance – the so-called linking length. Since the radial coordinate is in the redshift i.e. velocity space, two linking radius values are necessary. We chose 0.25 h^{-1}Mpc for the transversal and 250 km s^{-1} for the radial linking lengths, as this gives groups with reasonable properties. To take the selection effects into account when constructing groups from a flux-limited sample, the linking length has to increase slightly with distance. The scaling law for the linking distance was calibrated by shifting nearby groups to larger distances while requiring that the group would not break up (for details, see Tago et al. 2010). The group sizes and velocity dispersions of the resulting sample are similar at all distances. Our SDSS main sample contains 77858 galaxy groups and clusters.

The previous galaxy and group catalogue used in Liivamägi et al. (2012) was based on the SDSS DR7 main galaxy sample (Abazajian et al. 2009) and compiled

by Tago et al. (2010). The main discrepancies between those catalogues are caused by the different photometric processing of the SDSS releases and our criteria for sample cleaning. Comparing to DR7, there are 6869 galaxies and 942 galaxy groups less in our DR8 catalogue, most of which are galaxy pairs and small groups. Comparing galaxies from DR8 to those of DR7, we find that the luminosities of bright galaxies are increased (Tempel et al. 2012). We also note that, as a result, nearby galaxies are more luminous on average and distant galaxies are dimmer. This, of course, carries over to the group and cluster luminosities. The group and cluster properties are very similar between the two data releases. Individually, almost all of the larger groups were also identifiable, only pairs and smaller groups differ more noticeably. There is also a reported error in the astrometric measurements (Aihara et al. 2011b) but the inaccuracies in the positions of galaxies (less than 0.25") are too small to affect our analysis.

2.1.2 The SDSS LRG sample

The galaxies for the LRG sample were selected from the SDSS database by an SQL query requiring that the PrimTarget field should be either TARGET_GALAXY_RED or TARGET_GALAXY_RED_II. We demanded reliable redshifts (SpecClass = 2 and zConf > 0.95). We kept the galaxies within the same mask as the main galaxies (the compact continuous area in the Northern Galactic Cap). We calculated the absolute $M_g^\star(z = 0)$ magnitudes for the LRGs as in Eisenstein et al. (2001). We examined the photometric errors of the LRGs and deleted the galaxies brighter than $M_g^\star = -23.4$ from the sample to keep the magnitude errors small. In total, our sample includes 170423 LRGs up to the redshift $z = 0.6$ (the $k + e$-correction table in Eisenstein et al. (2001) stops at this redshift). It is worth mentioning that the LRG sample is approximately volume-limited (its number density is almost constant) between the distances from 400 h^{-1}Mpc to 1000 h^{-1}Mpc. The sky projection of the LRG sample is identical to the main sample (pictured in Fig. 2.1).

Galaxies closer than $d_0 = 435.6$ h^{-1}Mpc ($z_0 = 0.15$) are fainter and are "not officially" LRGs (Eisenstein et al. 2001). On the other hand, they are yet similar to LRGs by many properties and are sometimes also called Bright Red Galaxies (BRGs). Inclusion of these galaxies is necessary if we want to compare the main and LRG superclusters, because they increase the volume, where the two galaxy samples overlap.

The recently-published SDSS DR12 CMASS galaxy sample also has a large enough spatial density to allow finding supercluster-scale structures (Lietzen et al. 2016; Nadathur & Crittenden 2016). Unfortunately, the main bulk of its galaxies are located at distances greater than the LRG sample and well beyond the main sample making straight comparisons of superclusters between different catalogues difficult.

2.1.3 The Millennium galaxy sample

We chose a galaxy catalogue by Bower et al. (2006) that is an implementation of the Durham semi-analytic galaxy formation model on the Millennium Simulation by the Virgo Consortium (Springel et al. 2005). The catalogue is available from the Millennium database at the German Astrophysical Virtual Observatory[1]. A subsample of about one million galaxies was selected by the condition $M_r > -20.25$. This yielded a sample with almost the same number density of galaxies as that of the SDSS main sample (from 125 to 400 h^{-1}Mpc). We calculated the absolute luminosities for galaxies by taking $M_\odot = 4.49$ and using the SDSS r magnitudes (Vega) presented in the catalogue. This sample serves as a volume-limited test catalogue to study the performance of the supercluster finding algorithm.

2.1.4 Distance and luminosity corrections for the main sample

Spectroscopic galaxy samples such as SDSS are affected by the redshift distortions. These are caused by the peculiar velocities of the galaxies which add to the Hubble expansion (see Sect. 1.2.1). Consequently, the observed (and CMB corrected) redshifts of galaxies depend on both:

$$z_{\mathrm{obs}}^* \approx z_{\mathrm{c}} + \frac{v_{\mathrm{p}}}{c}, \tag{2.1}$$

where z_{c} originates from the true expansion of the universe and v_{p} is the velocity of the relative motion of a specific galaxy in its neighbourhood.

The most drastic of these effects are the cluster-finger redshift distortions, also known as the fingers-of-god effect. The cluster-finger effect is an apparent elongation of galaxy groups and clusters along the line of sight in redshift space. In groups and clusters, the velocity dispersion of galaxies can be very large and seriously obstruct our view of the real structures.

To suppress the cluster-finger redshift distortions, we use the rms (root mean square) sizes of galaxy groups and their radial velocity dispersions from the Tempel et al. (2012) group catalogue. We divide the radial distances between the group galaxies and group centres (d_{group}) by the ratio between the standard deviations σ_r and σ_v. This will remove the smudging of the density field by the cluster fingers. For groups with three or more members, the corrected galaxy distance d_{gal} is found as

$$d_{\mathrm{gal}} = d_{\mathrm{group}} + (d_{\mathrm{gal}}^\star - d_{\mathrm{group}}) \, \frac{\sigma_r}{\sigma_v / H_0}, \tag{2.2}$$

where $d_{\mathrm{gal}}^\star$ is the initial distance of the galaxy, σ_r the standard deviation of the projected distance in the sky from the group centre, σ_v the standard deviation of the

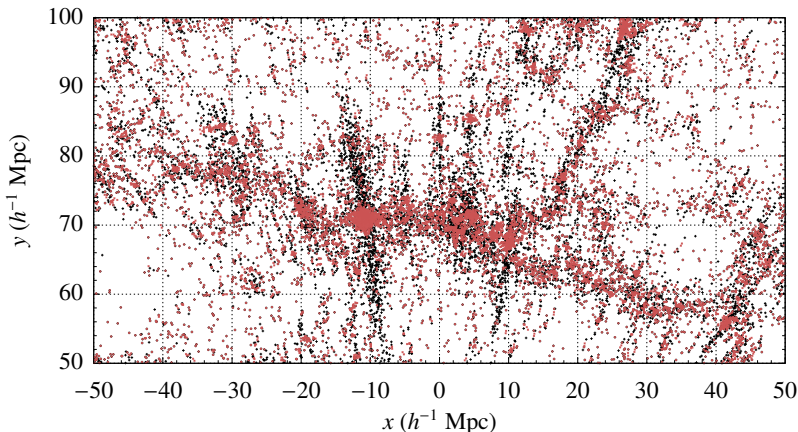---

[1]`http://www.g-vo.org/www/Products/MillenniumDatabases`

Figure 2.3: The result of the cluster-finger suppression on the galaxy distribution in a thin (40 h^{-1}Mpc) slice of the survey. Black dots show the original and red dots shifted galaxy positions.

radial velocity (both in physical coordinates at the group location). The radial velocities are given in units of km s^{-1} and formally converted to distance units by division with the Hubble constant H_0. In case of galaxy pairs, the extent of the system on the plane of the sky does not show its real size because of the projection effects. For pairs, we make corrections if comoving distance between galaxies along the line of sight is larger than the linking length $d_{\text{LL}(z)}$ at the redshift of the group. New distances are calculated as follows:

$$d_{\text{gal}} = d_{\text{group}} + (d^{\star}_{\text{gal}} - d_{\text{group}}) \frac{d_{\text{LL}(z)}}{|v_1 - v_2|/H_0}. \tag{2.3}$$

We do not alter the pairs, where it is smaller, i.e. $|v_1 - v_2|/H_0 < d_{\text{LL}}(z)$. Figure 2.3 illustrates the effects of the removal of the cluster-fingers. Kaiser effect, which distorts the redshifts on larger scales causing apparent flattening of structures along the line of sight (Kaiser 1987), is not taken into account. Shi et al. (2016) has shown that Kaiser effect can indeed have a significant effect on the shape of the superclusters, however, for the time being we don't have a method to correct this consistently.

We convert galaxy locations to Cartesian coordinates based on the so-called SDSS Survey coordinates η and λ (Stoughton et al. 2002). The Survey system is a spherical coordinate system defined in a way that each survey stripe forms a rectangle with a width of 2.5 degrees in η (see Fig. 2.1). Using the SDSS Survey coordinate system allows the most efficient placing of the galaxy sample cone inside a box, which is useful later when we construct the density field grid in the same coordinate

34

system. The galaxy coordinates are calculated as follows:

$$x = -d_{\mathrm{gal}} \sin \lambda,$$
$$y = d_{\mathrm{gal}} \cos \lambda \cos \eta, \qquad (2.4)$$
$$z = d_{\mathrm{gal}} \cos \lambda \sin \eta.$$

To compensate for selection effects and to ensure that the reconstructed density field does not depend on distance, we also have to account for the luminosity of galaxies that drop out of the survey magnitude window. One can find the total missing luminosity using the inverse of the selection function (Yahil et al. 1991). Considering every galaxy as a visible member of a density enhancement (a group or cluster) within the visibility range at the distance of the galaxy, we may add a proportional part of this lost luminosity to it. Observations and numerous group catalogues (e.g., Tago et al. 2006) have demonstrated that galaxies overwhelmingly reside in associations. We follow the procedure by Tempel et al. (2011) and estimate the amount of unobserved luminosity and weigh each galaxy as

$$L_{\mathrm{gal,w}} = W_L(d)\, L_{\mathrm{gal}}, \qquad (2.5)$$

where $L_{\mathrm{gal}} = L_\odot 10^{0.4(M_\odot - M)}$ is the observed luminosity of a galaxy with the absolute magnitude M (in the units of solar luminosity), and $M_\odot$ is the absolute magnitude of the Sun. The quantity $W_L(d)$ is the distance-dependent weight factor: the ratio of the expected total luminosity to the luminosity within the visibility window:

$$W_L(d) = \frac{\int_0^\infty L\, \phi(L)\mathrm{d}L}{\int_{L_1(d)}^{L_2(d)} L\, \phi(L)\mathrm{d}L}, \qquad (2.6)$$

where $L_{1,2}(d)$ are the luminosity limits corresponding to the survey magnitude limits $m_{1,2}$ at the distance d. For the main sample, we use the apparent magnitude limits $m_1 = 12.5$ and $m_2 = 17.77$ and take $M_\odot = 4.64$ mag in the r-band (Blanton & Roweis 2007) as the luminosity of the Sun.

We approximate the luminosity function by a double power law:

$$n(L)\mathrm{d}(L) \propto (L/L^*)^\alpha (1 + (L/L^*)^\gamma)^{(\delta-\alpha)/\gamma}\mathrm{d}(L/L^*). \qquad (2.7)$$

This form represents the bright-magnitude end of the luminosity function better than the usual Schechter function (Tempel et al. 2009). The meaning of the parameters in the luminosity function and their values are as follows: $\alpha = -1.305$ is the exponent at low luminosities $(L/L^*) \ll 1$, $\delta = -7.13$ is the exponent at high luminosities $(L/L^*) \gg 1$, $\gamma = 1.81$ is a parameter that determines the speed of the transition between the two power laws, and L^* (corresponds to $M^* = -21.75$) is the characteristic luminosity of the transition (Tempel et al. 2012). The luminosity weights are shown in Fig. 2.2.

2.1.5   Luminosity corrections for the LRG sample

Although the luminosity function of the SDSS LRGs has already been determined (Wake et al. 2006), it is difficult to calculate the luminosity weights for LRGs as we did above for the main sample. The reason is simple – the LRG sample does not have the two magnitude limits. Because of that, we find the observed comoving luminosity density $\ell(d)$ and defined the luminosity weight as its inverse:

$$W_L(d) = \ell(d_0)/\ell(d), \tag{2.8}$$

where d_0 is the fiducial comoving distance. We fix the fiducial distance at $d_0 = 435.6$ h^{-1}Mpc (see Sect. 2.1.2).

Again, we must draw attention to the fact, that both of these luminosity correction schemes (for the main and LRG samples) add luminosity to the observed galaxy locations, and cannot restore the real, unobserved galaxies. This increases the noise at larger distances, but that is unavoidable.

2.2   Estimation of the density field

As superclusters are searched for as regions with the luminosity density over a certain threshold in a compact region of space, we have to convert the spatial positions and luminosities of galaxies into a luminosity density field. A common approach is to assume a Cox model for the galaxy distribution, where the galaxies are distributed in space according to a inhomogeneous Poisson point process with the intensity $\rho(\mathbf{r})$ determined by an underlying random field (see, e.g., Martínez & Saar 2002). It means that, although expecting certain biases, we expect galaxies to trace the distribution of mass adequately. We will discuss it in greater detail in Sect. 2.2.3.

2.2.1   Kernel density estimates

A convenient way to estimate this intensity from the galaxy distribution is by finding the kernel sum (Davison & Hinkley 1997, Sect. 8.3.2):

$$\rho(\mathbf{r}) = \frac{1}{a^3} \sum_{i=1}^{N} K\left(\frac{\mathbf{r} - \mathbf{r}_i}{a}\right), \tag{2.9}$$

where the sum is over all N data points, $\mathbf{r}_i$ are the coordinates, $K(\cdot)$ is the kernel function, and a the smoothing scale. As we estimate luminosities, we multiply kernel amplitudes by weighted galaxy luminosities $L_{\text{gal,w}}$ and calculate the luminosity density field as

$$\ell(\mathbf{r}) = \frac{1}{a^3} \sum_{\text{gal}} K\left(\frac{\mathbf{r} - \mathbf{r}_{\text{gal}}}{a}\right) L_{\text{gal,w}}. \tag{2.10}$$
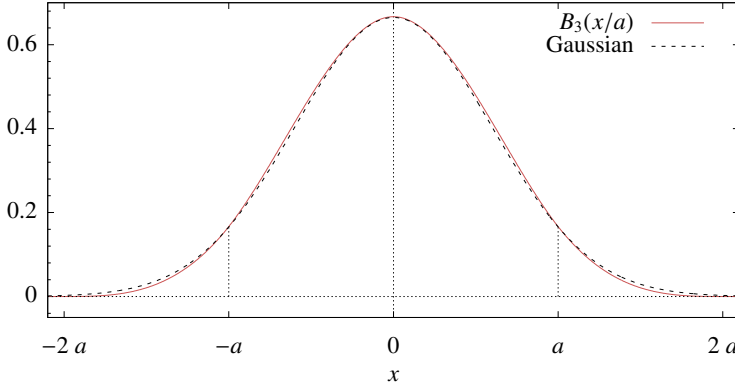
Figure 2.4: The shape of the $B_3(x)$ kernel function. Solid line – the $B_3(x)$ kernel; dashed line – a Gaussian with $\sigma \approx 0.6$.

The kernel functions $K(\cdot)$ are required to be distributions: positive everywhere and integrating to unity. In our case

$$\int K(\mathbf{y})d^3y = 1. \tag{2.11}$$

Good kernels for calculating densities on a spatial grid are the box splines B_k, where k denotes the degree of the spline. They are local and they are interpolating on a grid:

$$\sum_i B_k (x - i) = 1, \tag{2.12}$$

for any x and a small number of indices that give non-zero values for $B_k(x)$. To create our density fields we use the popular B_3 spline function, which can be given analytically as:

$$B_3(x) = \frac{|x - 2|^3 - 4|x - 1|^3 + 6|x|^3 - 4|x + 1|^3 + |x + 2|^3}{12}. \tag{2.13}$$

This function differs from zero only in the interval $x \in (-2, 2)$, meaning that the sum in (2.12) only includes values of $B_3(x)$ at four consecutive arguments $x \in (-2, 2)$ that differ by 1. In practice, we calculate the kernel sum (2.11) on a grid. Let the grid step be $\Delta \leqslant a$, and scale $a = k\Delta$, where $k \geqslant 1$ is an integer. Then the sum over the grid becomes

$$\sum_i B_3 \left(\frac{x - i\Delta}{a} \right) = k, \tag{2.14}$$

because it consists of k groups of four values of $B_3(\cdot)$ at consecutive arguments, differing by 1. Thus, the kernel

$$K_B^{(1)}(x/a; \Delta) = \frac{\Delta}{a} B_3(x/a) \tag{2.15}$$

37

differs from zero only in the interval $x \in (-2a, 2a)$ (Fig. 2.4) and preserves the interpolation property exactly for all values of a and Δ, where the ratio a/Δ is an integer (also, the error is very small even if this ratio is not an integer, but a is at least several times larger than Δ). We also refer to a as effective radius, because while the kernel extends to $\pm 2a$, it differs significantly from zero inside $\pm a$ (more than 90% of the area under the curve). The three-dimensional kernel $K_B^{(3)}$ is built as a direct product of three one-dimensional kernels:

$$K_B^{(3)}(\mathbf{r}/a; \Delta) \quad \equiv \quad K_B^{(1)}(x/a; \Delta)K_B^{(1)}(y/a; \Delta)K_B^{(1)}(z/a; \Delta) \tag{2.16}$$

$$= \quad \left(\frac{\Delta}{a}\right)^3 B_3(x/a)B_3(y/a)B_3(z/a), \tag{2.17}$$

where $\mathbf{r} \equiv \{x, y, z\}$. Although this is a direct product, it is practically isotropic (Saar 2009). This can be seen already from the fact that it is very close to a Gaussian with a mean zero and $\sigma \approx 0.6$ (Fig. 2.4), and the direct product of one-dimensional Gaussians is exactly isotropic. However, the difference in densities calculated by both kernels can be much greater due to the extended wings of the Gaussian (Tempel et al. 2014).

2.2.2 Construction of the luminosity density field

Construction of the luminosity density field in practise is carried out as follows. We denote the luminosity density field on a grid with $\ell_{\mathbf{i}}$, where $\mathbf{i} \equiv \{i_1, i_2, i_3\}$ are the indices of the vertices. Using Eq. 2.10, we calculate the luminosity densities by a kernel sum:

$$\ell_{\mathbf{i}} = \sum_{\text{gal}} K_B^{(3)}\left(\frac{\mathbf{r}_{\text{gal}} - \mathbf{r}_{\mathbf{i}}}{a}\right) L_{\text{gal,w}}. \tag{2.18}$$

where, as before, $L_{\text{gal,w}}$ is the weighted galaxy luminosity, and a the kernel scale.

We chose the smoothing width of $a = 8\ h^{-1}\text{Mpc}$ for the SDSS main sample. The choice of the kernel width is somewhat arbitrary, but an argument can be made that the scale has to correspond to the size of the structures we are searching for. For example, the kernel should be considerably wider than the diameters of galaxy clusters, which are a few megaparsecs. Also, we wish to be able to detect structures at large distances, where galaxies are sparser. We assume that the density field ties galaxies together if these are separated by $2a$. In Liivamägi et al. (2012), we show the nearest neighbour distributions for different distance intervals. We see that for the SDSS main sample the scale $a = 8\ h^{-1}\text{Mpc}$ is comfortably large enough to group galaxies together even at far distances (Fig. 2.5), and a slightly narrower kernel would also be sufficient. Historically, $a = 8\ h^{-1}\text{Mpc}$ has been used in previous supercluster
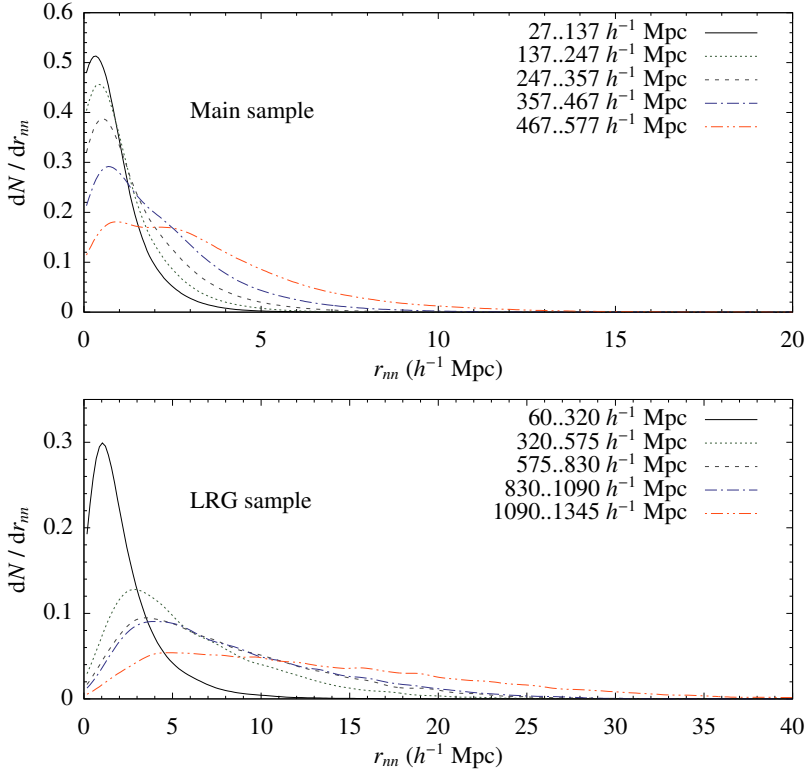
Figure 2.5: The nearest neighbour distributions for galaxies in the main (*upper panel*) and LRG samples (*lowe panel*).

catalogues (Einasto et al. 2007b) and in other supercluster studies (as in a more recent papers by Costa-Duarte et al. 2011; Luparello et al. 2011). As shown by Costa-Duarte et al. (2011), the density field method is actually not very sensitive to the choice of kernel width or type (e.g., Gaussian, Epanechnikov).

Following a similar analysis, we selected the kernel width for the SDSS LRG sample as $a = 16\ h^{-1}$Mpc , twice the scale of the kernel used for the main sample, since the LRG sample is sparser. However, most LRGs have at least one neighbour at distances up to $2a = 32\ h^{-1}$Mpc (Liivamägi et al. 2012). The density field of the Millennium sample is calculated with the same $a = 8\ h^{-1}$Mpc kernel width as for the main sample.

An image of the main sample density field is presented in the top panel of Fig. 2.6. The field seems to have a reasonably uniform look as the structures and density values are evenly distributed at large scales. We can now also estimate the effectiveness of our missing luminosity recovery. Figure 2.7 shows the dependence of the galaxy

Figure 2.6: A spatial slice of: the luminosity density field ℓ for the main sample (*top panel*), the standard deviation field of the luminosity density σ_ℓ (*middle panel*), and the signal-to-noise ratio field G (*bottom panel*). The slices have a thickness of $1\ h^{-1}$Mpc and are located at $z = 33\ h^{-1}$Mpc (in Cartesian coordinates, see Eq. 2.4). Colours are illustrative.

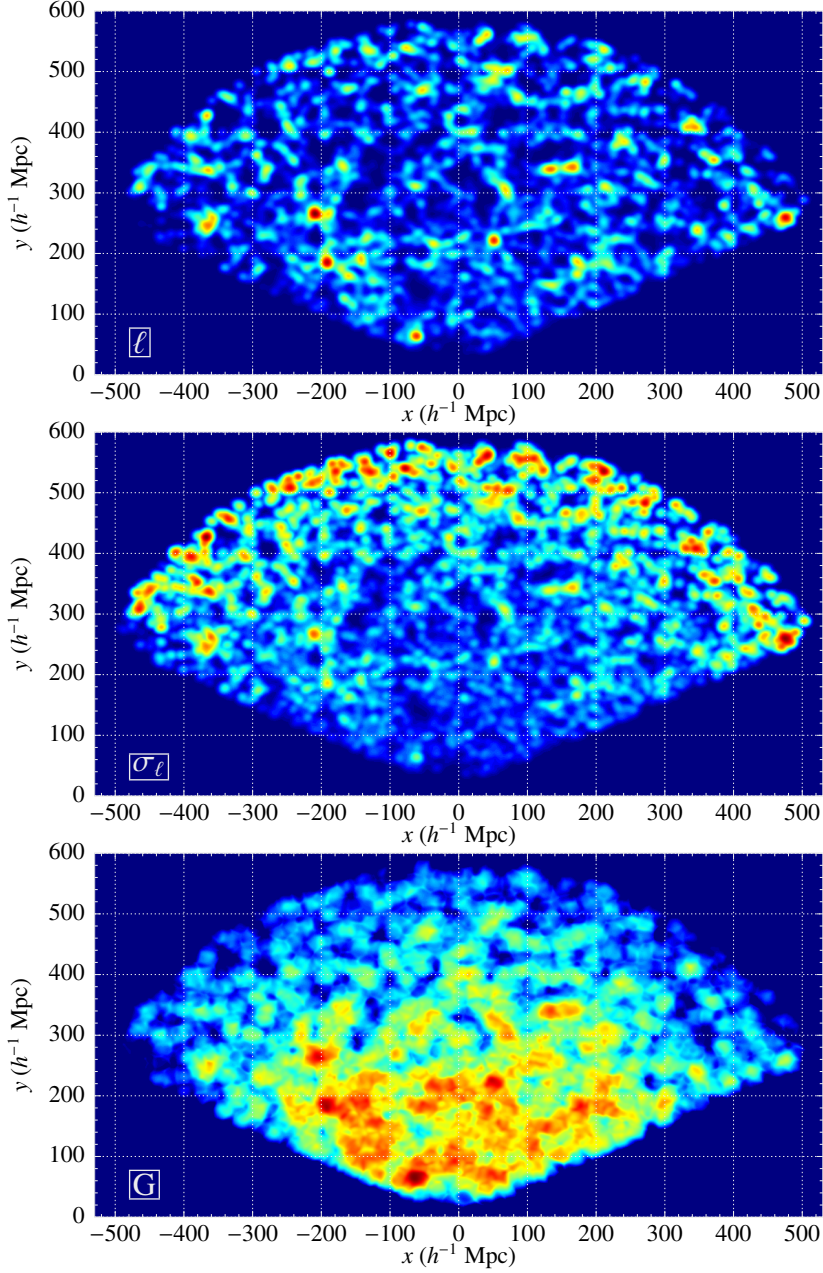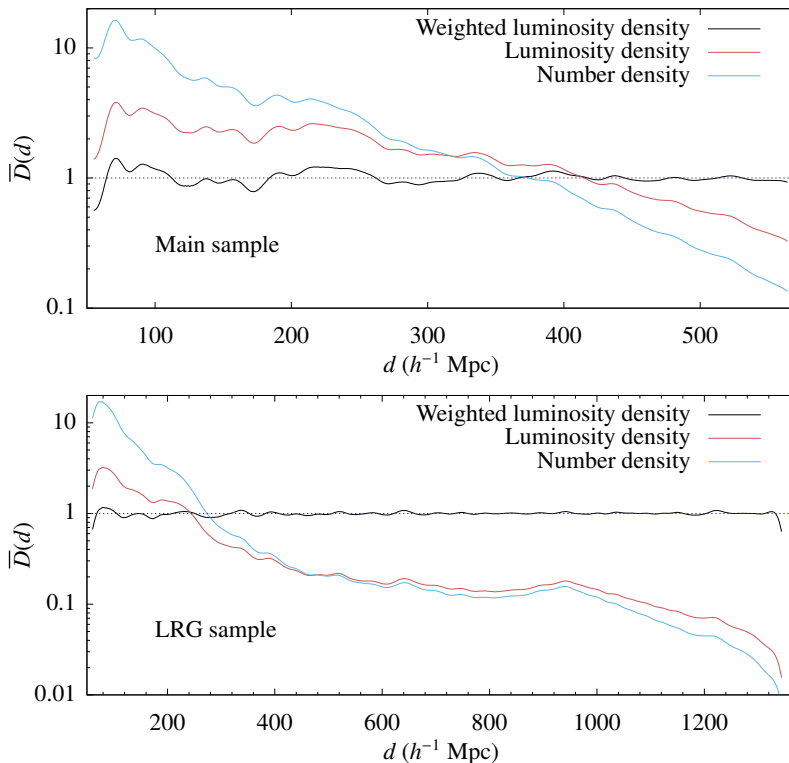**Figure 2.7**: Average normalised densities vs distance for the main (*upper panel*) and LRG samples (*lower panel*). The densities are averaged over thin (a few h^{-1}Mpc) concentric shells of the distance d. Solid line – the weighted luminosity density; dashed line – the observed luminosity density; dotted line – the galaxy number density.

number density, the observed luminosity density, and the weighted luminosity density of galaxies on distance. The weighting procedure seems to have restored total luminosities adequately, the luminosity density does not have a strong systematic dependence on distance.

As the last step before extracting superclusters, we convert densities into the units of mean density. The main purpose of this is to facilitate comparison between different density fields. For that, we construct a pixel mask that follows the sample edges. We first employ the sky projection mask (displayed in Fig. 2.1) used in Martínez et al. (2009) and then set the lower and higher limits for the distance. We do not need to use the more precise mask (e.g., the "mangle" mask provided by the NYU VAGC) because we are searching for structures of much larger dimensions. The angular diameter of the kernel at the far end of the sample is much larger (1.6 degrees for the

41

Table 2.1: Properties of the galaxy samples and of the density fields.

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Sample	N_{gal}	N_{groups}	$z_{min} \ldots z_{max}$	ℓ_{mean} $\frac{10^{10}h^{-2}L_\odot}{(h^{-1}\mathrm{Mpc})^3}$	V_{mask} $(h^{-1}\mathrm{Gpc})^3$	Δ $h^{-1}\mathrm{Mpc}$	a $h^{-1}\mathrm{Mpc}$	$d_{min} \ldots d_{max}$ $h^{-1}\mathrm{Mpc}$
Main	576493	77858	$0.009 \ldots 0.2$	$1.526 \cdot 10^{-2}$	0.132	1	8	$55 \ldots 565$
LRG	170423	-	$0.02 \ldots 0.5$	$8.148 \cdot 10^{-4}$	1.789	2	16	$60.5 \ldots 1346.4$
Millennium	1039919	-	-	$1.304 \cdot 10^{-2}$	0.125	1	8	-

Columns in the Table: (1) sample name; (2) number of galaxies used to calculate the density field; (3) number of groups in the sample; (4) redshift limits for the sample; (5) mean luminosity density value; (6) volume of the mask; (7) grid cell size; (8) smoothing scale; (9) distance limits of the mask.

main sample and 1.3 degrees for the LRGs) than these of the multitude of small holes inside the SDSS survey mask (with diameters less than an arcminute). The main sample density field mask is limited with the distances 55 to 565 h^{-1}Mpc. The distance limits here and also in case of the LRGs are chosen to avoid the nearest and the most distant incomplete regions. The mask for the Millennium sample is a cube with the side length of 500 h^{-1}Mpc.

The mean density of the field is determined as an average over all vertices inside the mask,

$$\ell_{mean} = \frac{1}{N_{mask}} \sum_{\mathbf{i} \in mask} \ell_{\mathbf{i}}, \qquad (2.19)$$

where N_{mask} is the number of grid vertices inside the mask. We finally normalise the whole density field as

$$D_{\mathbf{i}} = \frac{\ell_{\mathbf{i}}}{\ell_{mean}}, \qquad (2.20)$$

for all vertices with coordinates $\mathbf{i}$ inside the mask. The vertices outside the mask are marked and not used in further analysis. Main properties of luminosity density fields for all three samples are given in Table 2.1.

2.2.3 Bootstrap error analysis of the density field

To characterise the errors of our density field estimates we have to choose the statistical model for the galaxy distribution. The most popular model used for the statistics of the spatial distribution of galaxies in the universe is the "Poisson model" (Peebles 1980), an inhomogeneous Poisson point process where the local intensity of the process is defined by the amplitude of the underlying realisation of a random field. In statistics it is called the Cox random process, see an introduction and examples in Martínez & Saar (2002) and Illian et al. (2008). In cosmology, the random fields used are usually the Gaussian or log-Gaussian fields.

As any statistical model, it has been postulated to describe the galaxy distribution, and its success in applications describing the statistical properties of that distribution tends to support it. For example, this model was used to develop methods for estimating the two-point correlation function (Hamilton 1993) and the power spectrum of the galaxy distribution (Tegmark et al. 1998). These methods have been extensively used to study the galaxy distribution. The same model serves as the basis for a maximum-likelihood approach to recover the large-scale cosmological density field by Kitaura et al. (2010).

When using the kernel method to estimate the intensity of our Cox process, a popular procedure to estimate the uncertainties of kernel-based intensities for inhomogeneous Poisson processes is bootstrap (see, e.g., Davison & Hinkley 1997, Sect. 8.3.2). Because the kernel used in estimating the intensity in Eq. 2.9 is compact, there is only a finite and, in practice, a relatively small number of members in this sum. Bootstrap is used to estimate the sample errors (discreteness errors) caused by the discrete sampling.

As stressed by Silverman & Young (1987), bootstrap consists of two separate elements. Let our sample be $(X_1, \ldots, X_n)$. First, to estimate the discreteness error caused by the finite sample size n of the sample parameter $\theta(X)$ that we are estimating, we use the sampling method, drawing a large number of samples of size n from the (integral) population distribution function $F(X)$. Technically, it is the simplest method for generating random numbers with a given distribution – select n uniform random numbers U_i in the interval (0,1) and select the sample value X given by $F(X) = U$ for each U. The other element of bootstrap is to assume that the population distribution F can be approximated by the empirical distribution function F_n defined by all the n observed values $(X_1, \ldots, X_n)$ that form the sample. If all X_i-s are i.i.d. (independent and identically distributed), this function can be defined as a step function with increments $1/n$ at every $X_j^\star$, where $(X_1^\star, \ldots, X_j^\star)$ is an ordered growing sequence of the original sample values X_i. If we select from F_n, any bootstrap sample consists of the values of the original sample, selected from the original sample randomly with replacement (Efron & Tibshirani 1993). Using the values of θ for all bootstrap samples, we can find the sampling errors (usually the bias and the variance) of the parameter θ. These are the bootstrap error estimates we search for.

Theorems that prove the effectiveness of the bootstrap error estimates are usually proved for the case where both the sample size and the number of the bootstrap samples approach infinity (Shao & Tu 1995), and for a finite sample size simulations are used. In our case, the sample that defines the intensity estimate for a given point in space consists of these galaxies, the positions of which are within the kernel volume around that point. The usual rule of thumb says that bootstrap error estimates may be considered reliable when the sample size is more than 30 (even sizes as small

as 14 have been used in simulation examples, as in Efron & Tibshirani (1993) and Silverman & Young (1987)). Our kernel volumes include, on average, 150 galaxies in the case of the SDSS main sample and 25 galaxies for the LRG sample. At the density levels where we choose to define our superclusters, (from 2.5 to 5.5 times the mean density ℓ_{mean}), the corresponding numbers are 750 and 125, so our error estimates should be reliable enough.

For the inhomogeneous Poisson process, where the points $\mathbf{X}_i$ are identically and independently distributed with the locally defined intensity λ, bootstrap can be used to estimate the errors of the kernel estimate (Eq. 2.9). In practice, for that a bootstrap version that is called smoothed bootstrap is used. This is a version of parametric bootstrap, using, instead of the empirical distribution function, its smoothed version. Silverman & Young (1987) demonstrated that it is an effective way of estimating the variance of intensity. To use that, in practice, we generate bootstrap samples of the same size as the original sample, selecting the galaxies from our sample randomly with replacement, as usual in bootstrap, but giving the selected galaxies random displacements. As explained in Davison & Hinkley (1997) and Silverman & Young (1987), the random spatial displacements are required to have the probability density of the same form as the kernel function, but it is useful to undersmooth, using for the displacements a kernel that is narrower than the kernel used for calculating the intensity estimates. We undersmooth by a factor of two, meaning that the maximum shifting length was 8 h^{-1}Mpc for the main and 16 h^{-1}Mpc for the LRG sample (smoothing scale a divided by 2). We use 100 bootstrap samples for each grid vertex. This number has been found to be large enough to estimate the sample variance, based on simulation studies (Efron & Tibshirani 1993).

Another point that has to be taken care of when estimating global (population) statistics as correlation functions or power spectra (spectral densities) of Cox processes is to account for the difference in the measured statistic for the specific realisation of the random field and for the random field as a whole (the so-called cosmic noise problem, see, e.g., Szapudi & Colombi (1996), Peacock (1999, p. 522)). The discreteness errors, estimated by bootstrap, and the realisation variance combine in a subtle way (Cohn 2006). In our case, fortunately, the spatial density – the intensity of the Cox process that we are estimating (the geography of the large-scale structure) is exactly the underlying random realisation itself – we are measuring the cosmic noise, so we are not interested in the mean density of the universe. The only errors found in our intensity estimates are discreteness errors, and these can be estimated by bootstrap.

We select the galaxies for the bootstrap samples, together with their measured luminosities, and we consider galaxy distribution as a marked Cox process, with luminosities as marks. If we could statistically model the luminosity distribution
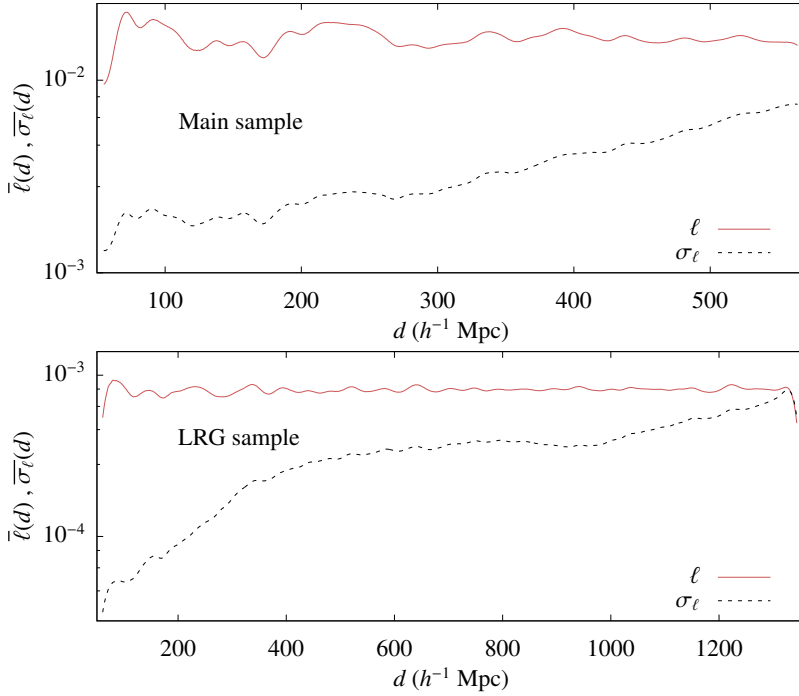
Figure 2.8: The distance dependence of the average density and of the standard deviations for the main (*upper panel*) and LRG samples (*lower panel*).

among galaxies as random (a random marks model, see, e.g., Illian et al. 2008), we could build bootstrap samples by randomly relabelling galaxies, and choosing their luminosities as in the usual bootstrap from the luminosities of the sample galaxies inside the kernel volume. This, however, would not be right, as galaxies are well known to be segregated by luminosity – more luminous galaxies populate the regions of higher number density of galaxies (Hamilton 1988; Girardi et al. 2003). We chose another way and tried modelling the luminosity errors. These consist of a small error in the luminosity weights, generated by the errors in the luminosity function, and an error in modelling the evolution correction (Blanton et al. 2003). We tested the effect of these errors by selecting them randomly from the observed distributions, compared the intensity estimates with modified luminosities and with fixed luminosities, and found no significant differences. As the luminosity errors were much smaller than the deviations of the intensity estimates generated by bootstrap, the discreteness errors, and we did not find a good statistical model to describe them, we ignored these errors.

After finding the positions for the galaxies of a bootstrap sample, we find a new intensity estimate, i.e., calculate a corresponding luminosity density field. We re-

peated the procedure a number of times (as said before, for this work, we generated 100 bootstrap samples for every grid point where we estimated the intensity) and found the standard deviation for the intensity σ for each grid vertex $\mathbf{i}$ as

$$\sigma_{\mathbf{i}} = \sqrt{\frac{1}{N} \sum_{m=1}^{N} \left(\ell_{m\mathbf{i}}^* - \overline{\ell_{\mathbf{i}}^*} \right)^2}, \tag{2.21}$$

where N is the number of bootstrap realisations, ℓ_m^* the intensity for a bootstrapped sample, and $\overline{\ell^*}$ its mean over all realisations. We also found the "signal-to-noise ratio" for each grid point:

$$G_{\mathbf{i}} = \frac{\ell_{\mathbf{i}}}{\sigma_{\mathbf{i}}}, \tag{2.22}$$

which is used later to estimate confidence levels of superclusters.

Figure 2.8 shows the dependence of σ_ℓ on distance. We can see the expected rise of σ_ℓ with a distance that is mainly caused by the decrease in the galaxy number density. We also see that the absolute values of the standard deviation are very low when compared to the density. This can be attributed to both the stability of the large-scale structures and the large smoothing scale for the density fields – several tens of galaxies contribute to the density at any point. Example maps of the standard deviation, and signal-to-noise fields for spatial slices of the main sample are shown in lower panels of Fig. 2.6. Looking at the features on the images of the standard deviation and the signal-to-noise fields, we can relate them to the observed large-scale structures in the luminosity density field on the upper panel. Nearby peaks in the density field stand out also in the signal-to-noise map, but the distant peaks are drowned in the noise.

2.3 Supercluster assembly

The conventional way to extract structures from the density field is to choose a threshold and define your objects as contiguous volumes above that level (see, e.g., Einasto et al. 2007b; Luparello et al. 2011). Depending on your further plans, these levels can be chosen differently. Because we do not know in advance, which density thresholds are suitable for our purposes, we create many sets of contour surfaces for different density thresholds and analyse their properties afterwards.

2.3.1 Exctraction of density field objects

We take density threshold values D_n from the range between $D_{\min}$ and $D_{\max}$, sampling it uniformly with a constant step δD. In this study we have chosen $\delta D = 0.1$ for all galaxy samples. Given the threshold D_n, density field objects are created in
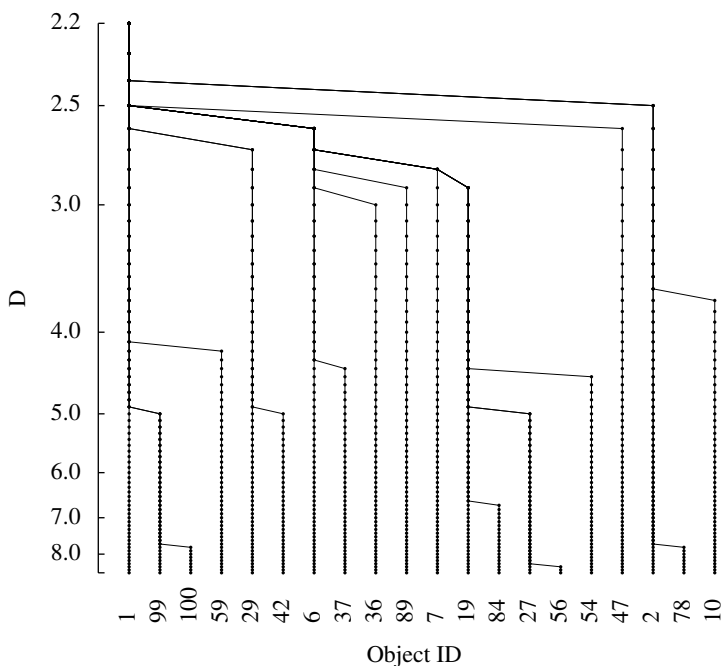
Figure 2.9: Excerpt of the splitting tree of the density field objects in the main sample.

the following way. First we locate the luminosity peaks in the density field. These density peaks are used to identify superclusters (density field objects). Contiguous supercluster regions are then grown cell by cell around the peaks in the density field resulting in what we call a "marker field":

$$M_{n,\mathbf{i}} = \mathrm{ID}_{\mathrm{peak}}, \ \mathbf{i} \in \{\mathbf{i}|D_{\mathbf{i}} > D_n\}, \tag{2.23}$$

where $\mathrm{ID}_{\mathrm{peak}}$ is the density peak number. All the vertices belonging to one object are assigned the same mark value.

We start scanning the field at high densities and move towards lower density levels. Each time an object first appears, it is assigned a unique identification number that will be used for this supercluster throughout the catalogue. We keep track when an object emerges from the field and when or if it is eventually incorporated into another density field object. If such a merger occurs, the identifier of the object with the higher peak value will be used to designate that object later on. To record the merging history of the density field objects, we order them into a tree-like structure covering all the density thresholds (Fig. 2.9).

Superclusters are finally assembled by distributing galaxies among density field objects. We do this for each density threshold by correlating galaxy positions with the

47

corresponding marker field. For the SDSS main sample we also assign galaxy groups to superclusters. If a group or a cluster is found to be in a supercluster (its centre coordinates lie inside the supercluster contours), all its member galaxies automatically also belong to the same supercluster. In order to avoid small spurious density field objects which include no galaxies, we also enforce a minimum volume limit of $(a/2)^3$ for the size of a supercluster (a being the smoothing scale for the particular density field).

2.3.2 Selection of density thresholds

With the multitude of available thresholds comes the question – which is the "correct" one? Just as there is no clear-cut definition for superclusters, there is also no single answer for this. In this work, we cover two possibilities for tackling this problem. The first one is the conventional way of choosing a fixed density level as a free parameter. This gives a set of objects that are comparable within the whole sample volume, where the density level D_n can be selected according to the properties of superclusters one wishes to study. As an example, for identifying structures, low density levels are better, but for studying the details of the structure, higher levels are useful; and sometimes it is necessary to use a set of luminosity levels.

Density level $D = 4.9$ was employed by Einasto et al. (2011c) to define structures in a detailed study of the Sloan Great Wall as this gave a good match with the visual impression of the system. In an earlier study by Einasto et al. (2007b), a wide range of density thresholds were studied while compiling a catalogue of superclusters in the 2dF redshift survey galaxy sample. In order to get a sensible distribution of supercluster dimensions, they finally settled on $D = 4.6$. Luparello et al. (2011) derived the density threshold for their sample using a spherical collapse model by Dünner et al. (2006) to delineate the so-called future virialised systems (see also Araya-Melo et al. 2009). They calibrated the model with numerical simulations and arrived at the density threshold $D = 5.5$. They also compared their supercluster sample to our catalogue on the same density level and found a very good correspondance (which is not surprising since the method and the data are very similar). Other examples include Costa-Duarte et al. (2011), who used $D = 3.0$ and $D = 6.0$ to obtain different samples with different properties, and Lietzen et al. (2009), also using a set of several density levels. However, this approach is susceptible to Poisson noise, especially in sparser environments. It also does not take into account the richness differences of superclusters. We show both effects in Sect. 3.1.

Another prospect is to use the so-called persistence filtering similar to that described by Sousbie (2011) or Nadathur & Hotchkiss (2014). In this case the difference between superclusters peak and adaptively found threshold density values can be used to derive the significance of its existance.

Because of that, we offer an alternative procedure that assigns an individual threshold to each supercluster, adapting to the local density level. The idea is to follow the growth of individual superclusters from a compact volume around its centre by lowering the density level and observing the supercluster mergers. By defining a supercluster as the volume within the density contour until the first major merger, we can break the large-scale structure into a collection of compact components. Every component (supercluster) then has its own limiting density level D_{scl}. This is an usual practise for other astronomical objects – we do not define galaxies by a common limiting stellar density level. As a result, we get a set of superclusters that forms the connected large-scale cosmic web.

To identify such superclusters in practise, it is easier to begin from lower densities and to proceed upwards. The mergers can now be seen as breakups of structures. We trace the splitting events in the density field objects tree. When a split occurs, the lower density filament ceases to be a "bridge" between two higher density regions. After the split, we pick the density value just above of the bridge, as the defining density level for these two objects. After that we continue increasing the threshold, but from now on both objects are handled separately. If one is broken up again at some higher threshold, it will not affect the other one.

As a downside, this technique still requires manual setting of several limits. First, the minimum size of a supercluster must be selected, for obviously some of the breaks involve objects that are too small to be of interest. However, this is not unique to us. In previous studies, a 100 $(h^{-1}\text{Mpc})^3$ lower volume limit was used by Einasto et al. (2007b), Costa-Duarte et al. (2011) used ten galaxies as a minimum for their superclusters (in combination with the volume limit of 64 $(h^{-1}\text{Mpc})^3$), Luparello et al. (2011) used the object luminosity of 10^{12} $L_\odot$ as the lower limit. In this study we use a minimum diameter of the supercluster. As demonstrated in the next chapter, we have a reason to believe that the diameters of our objects are relatively selection-free with respect to the distance and therefore a good property for constraining.

We must also choose the maximum threshold D_{lim}, because while we observe that most of the superclusters are defined at similar density levels, some objects in very rich environments, even taking into account the minimum size condition, can be fragmented at much higher thresholds and the algorithm may break up well-established structures. Because of that, we proceed in several steps. First, we calculate the distribution of structure breaks for all density levels. Then we select only the events that can be used to define a supercluster, i.e., involve objects larger than our minimum size limit. The threshold value where 95% of the "size-allowed" splits occur is chosen as the maximum density limit D_{lim}. To obtain the final set of superclusters, we repeat the procedure again, but this time we stop further splitting of structures above the limiting density value.
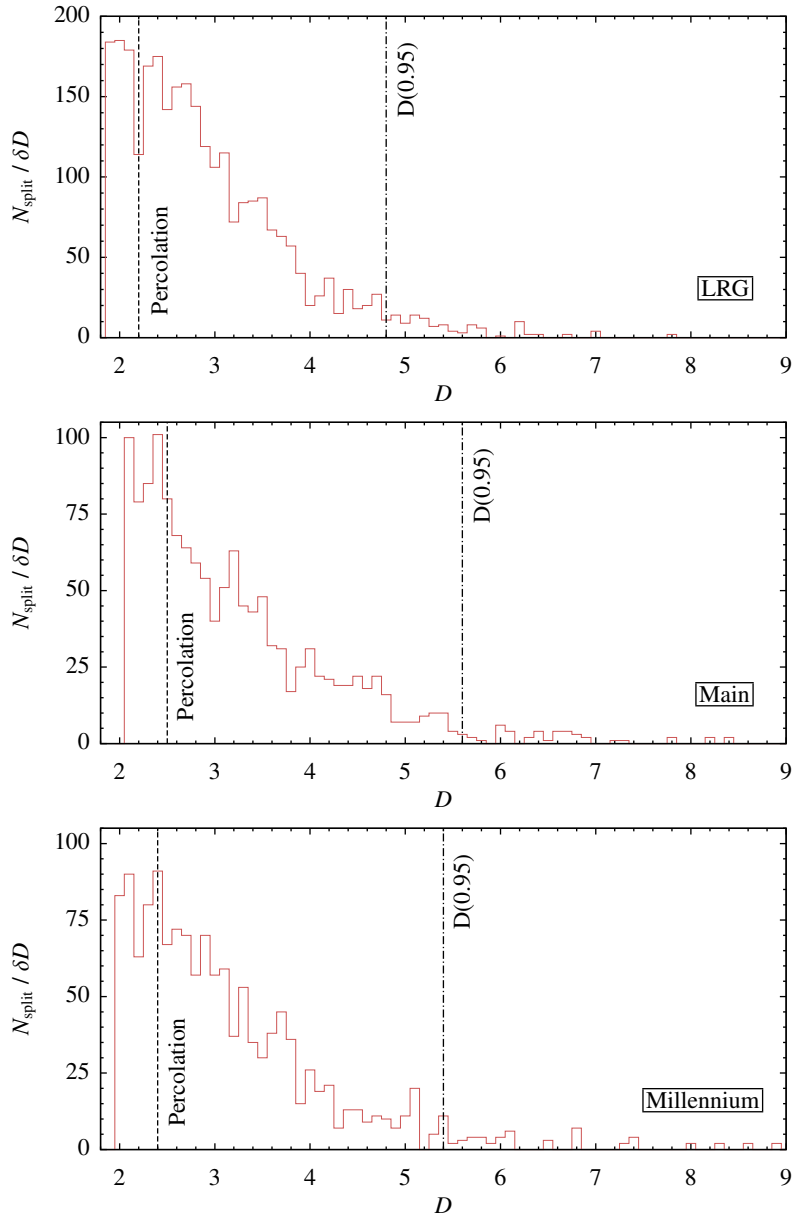
Figure 2.10: The number of supercluster defining structure breaks (anti-mergers) per density threshold. Vertical lines denote the percolation and the upper limit densities. From *top* to *bottom*: LRG, main, Millennium.

Designating a lower density limit is not strictly necessary, in principle we could track all the structures down to the point of merging. However, it is useful from the practical point of view. Slicing the field at very low densities yields huge structures which clearly include lots of "conventional" superclusters and require extra computational effort to take care of. One natural lower density limit is the percolation density level. Percolation happens when the largest structure begins to fill the whole sample volume. In practise we define this level as when the richness of the second richest structure starts to decrease when lowering the density level (Martínez & Saar 2002). The distribution of structure splits/mergers together with the lower and upper density constraints is shown in the Fig. 2.10.

Shifting the maximum density threshold upwards will fragment structures further and therefore slightly increase the number of objects. Reducing the minimum size of a supercluster will introduce a large number of new smaller objects and also break larger structures. Changing the lower density limit affects only the size of small and very isolated objects in low density areas, which may or may not merge into larger structures or reach the minimum size limit. A simplified 2D diagram in Fig. 2.11 outlines the differences between using the fixed or adaptive density threshold contours. The upper panel depicts the case with a single fixed density level D_n: only two objects are found, density field features I, II, III and IV are lumped together into one structure, and V is separate. However, with individually derived thresholds (lower panel), we break the bridge between the object I and the group of II, III, IV at the density level D_{n+1}. Then, II, III and IV are further split at the level D_{n+2}. II and III remain together because either III is, e.g., too small to be an independent object, or the threshold D_{n+3} is found to be too high. On the other hand, the link between I and V forms at smaller densities – thus, V is assigned the lower density threshold of D_{n-2}.

Structures in fixed and adaptive threshold catalogues can be very different (for visual impression see Fig. 2.12) Therefore, we present the SDSS main and LRG supercluster catalogues in two versions, one of them with adaptive density thresholds. Differences between these catalogues are discussed in more detail in Sect. 3.1. Another important issue is that in order to maintain the constant average luminosity density globally, we alter it significantly on a more local level (with galaxy luminosity weighting). Consequently, it also affects the threshold-related matters and we return to this problem while studying the spatial distribution of superclusters in Sect. 3.4.

2.4 Catalogue overview

The catalogue consists of several tables for superclusters, galaxies and galaxy groups (if available). For each density level D there are tables with properties of all super-
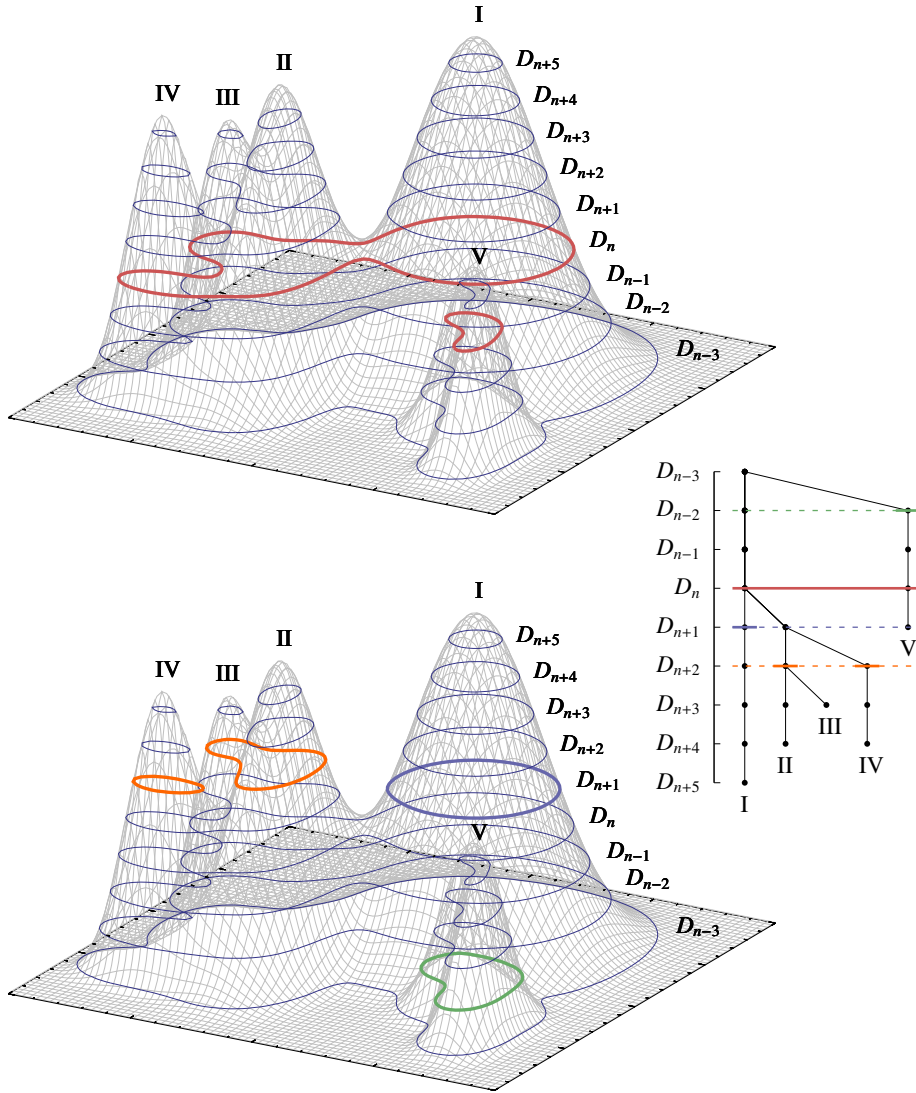
Figure 2.11: Extracted density field objects (thick coloured contours) found with a single fixed density level D_n (*upper panel*) or with various individually derived thresholds (*lower panel*). Thin contours denote all available isodensity cuts of the field. The corresponding tree-structure of the density field object is drawn in the *middle*.
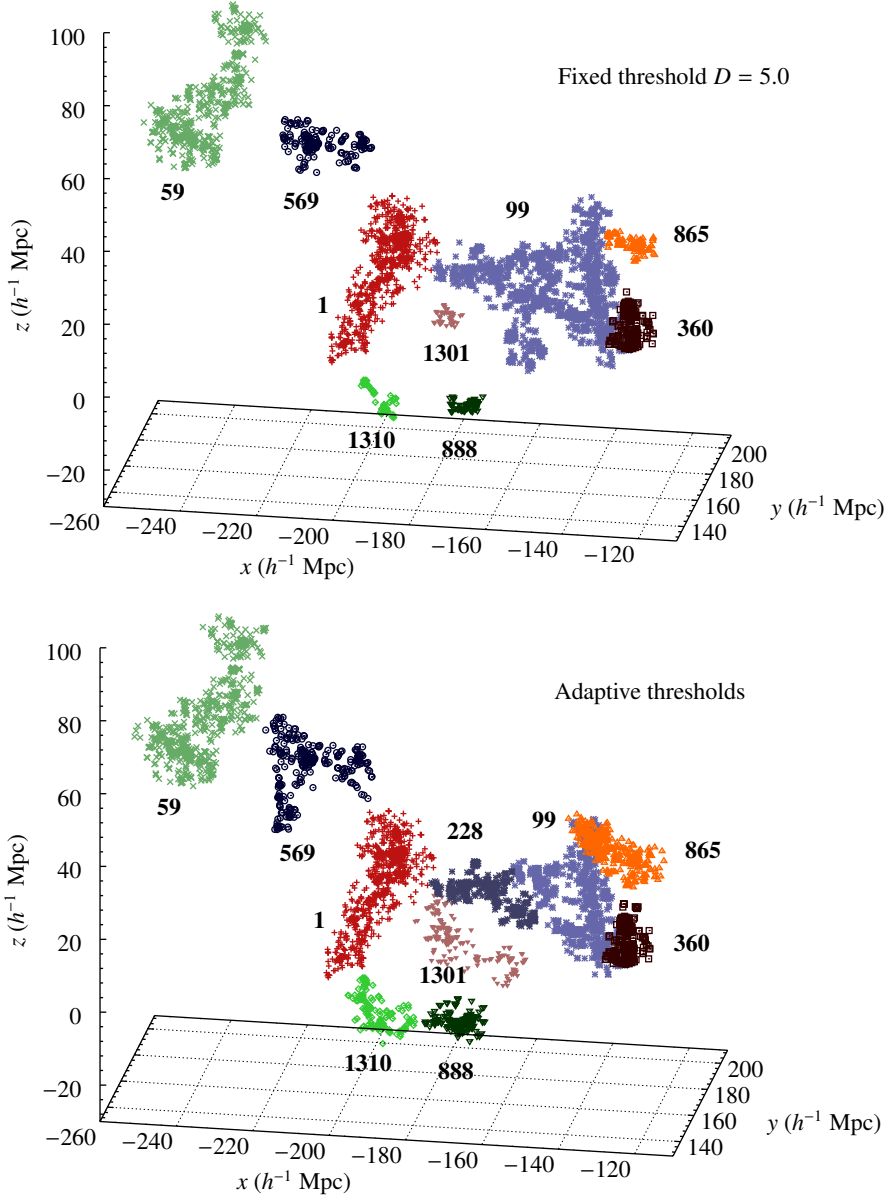
Figure 2.12: The supercluster SCl 1 and the surrounding superclusters found with different methods for assigning the limiting threshold D. Different colour symbols designate different superclusters. In the *upper panel* all superclusters have the same threshold $D_{\text{fix}} = 5.0$. The adaptive density levels D_a in *lower panel* are: SCl 1 ($D_a = 5.0$); SCl 59 (4.2); SCl 99 and SCl 228 (5.6); SCl 360 (4.8); SCl 569 (4.4); SCl 865 (3.9); SCl 888 (3.8); SCl 1301 (3.3); SCl 1310 (3.8).

53

Table 2.2: Supercluster catalogue properties.

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Sample	N_{scl}	n_{scl} $(h^{-1}\ \mathrm{Gpc})^{-3}$	f_{edge}	$f_{\mathrm{gal}\in\mathrm{scl}}$	$f_{\mathrm{vol}\in\mathrm{scl}}$	D_{fix} $^{\dagger}\ell_{\mathrm{mean}}$	$D(N_{\mathrm{max}})$ ℓ_{mean}	$n_{\mathrm{scl}}(D(N_{\mathrm{max}}))$ $(h^{-1}\ \mathrm{Gpc})^{-3}$
Fixed threshold								
Main	980	7417	0.187	0.145	0.013	5.0	2.8	11928
LRG	3761	2102	0.151	0.184	0.024	4.4	3.0	2671
Millennium	844	6752	0.147	0.153	0.010	5.0	3.3	10528
(1)	(2)	(3)	(4)	(5)	(6)	(10)	(11)	(12)
Sample	N_{scl}	n_{scl} $(h^{-1}\ \mathrm{Gpc})^{-3}$	f_{edge}	$f_{\mathrm{gal}\in\mathrm{scl}}$	f_{vol}	D_{lim} ℓ_{mean}	D_{perc} ℓ_{mean}	$\oslash_{\mathrm{lim}}$ $h^{-1}\mathrm{Mpc}$
Adaptive threshold								
Main	1188	8991	0.221	0.272	0.037	5.6	2.5	16
LRG	2701	1509	0.153	0.339	0.059	4.8	2.2	32
Millennium	1214	9712	0.194	0.282	0.035	5.4	2.4	16

Columns in the Table: (1) sample name and threshold assigning method; (2) number of superclusters; (3) number density of superclusters; (4) fraction of superclusters close to the sample edge; (5) fraction of galaxies in superclusters; (6) volume fraction of superclusters; (7) fixed threshold value; (8) density threshold with most objects; (9) number density of objects for the threshold $D(N_{\mathrm{max}})$; (10) maximum allowed value for adaptive thresholds D_a; (11) percolation threshold; (12) minimum allowed supercluster diameter.
Notes: †mean density value for the sample (see Table 2.2).

clusters found at that threshold. A similarly structured supercluster catalogue with adaptively assigned density thresholds is also presented. A short list of more important information given for each object is as follows:

- number of galaxies and groups (latter for the SDSS main sample alone);

- supercluster luminosity estimates, both from the density field and the galaxy data;

- different definitions for the supercluster location;

- supercluster size (volume, diameter);

- confidence estimate for the supercluster found from the signal-to-noise field;

- morphological parameters describing the supercluster shape.

Additionally, we provide lists of galaxies and groups, together with the supercluster identifiers they are attributed to, for all density levels. We cover more important attributes of superclusters in the next chapter: describe how they are calculated and also analyse their properties. We also present the supercluster splitting tree in the form of a table, where each supercluster is given the identifier of the object it belongs to at all given thresholds. A more detailed description of the catalogues can be found in Appendix A.1. Table 2.2 contains the summary information about the catalogues.

3 Properties and the spatial distribution of superclusters

In the previous chapter we outlined a method to produce a catalogue (actually many catalogues) consisting of specific galaxy structures. To give meaning to the nature of these objects we now try to analyse their properties. There are many directions for possible research here; however, in order to keep the volume of this work within reason, we limit ourselves to more general ones. First, we cover the most basic attributes as the size and luminosity of superclusters and then move on to supercluster shape characteristics and their spatial distribution. To uncover selection effects in the observational samples, we check the distance dependence of all the properties. Using principal component analysis, we also examine the relations between different properties and attempt to find correlations.

3.1 Basic properties of superclusters

After delineating superclusters, we can find a number of supercluster properties for all density levels using both the density field and the galaxy data. In the following we list the most basic attributes of superclusters in the catalogues and describe how they were acquired. Along the way we present and compare the supercluster catalogues for all our three galaxy samples, and investigate also the effects of both the fixed and adaptive density level assignment schemes.

3.1.1 Location

Defining the location of a supercluster is not straightforward. As we will show in the next paragraphs, they are extremely extended objects with dimensions ranging from 10 to almost $100\,h^{-1}$Mpc. More than one definition can be used to derive supercluster coordinates. First, we can use the initial density peak, from which the supercluster "grew" and which usually indicates the presence of a large galaxy cluster, to define the supercluster position. Another way is to calculate the superclusters' centre of mass (or rather, luminosity) using the coordinates and luminosities of its member galaxies

$$\mathbf{r}_{\mathrm{scl}} = \frac{1}{L_{\mathrm{scl,wgal}}} \sum_{\mathrm{gal}\in\mathrm{scl}} \mathbf{r}_{\mathrm{gal}} \cdot W_L(d_{\mathrm{gal}})L_{\mathrm{gal}}. \qquad (3.1)$$

Both methods have their merits – for certain studies like correlating CMB maps with supercluster positions (e.g., Cai et al. 2013), using the geometrical centre is more justified. However, this point is not related to any physical object and may even be
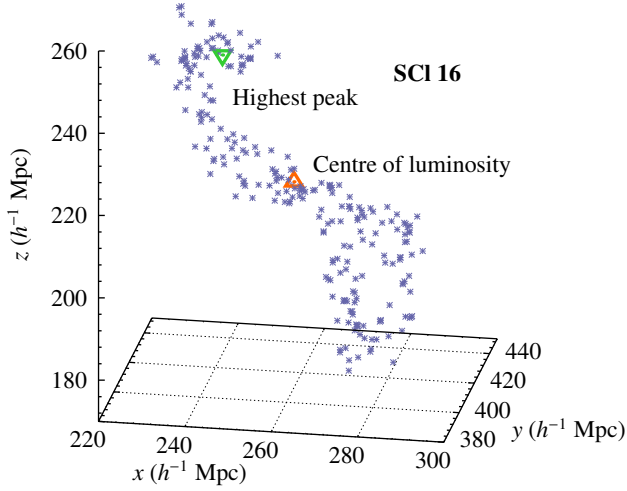
Figure 3.1: An example of the discrepancy between two supercluster location designators. The distance from the highest density peak (*green triangle*) to the centre of luminosity (*orange triangle*) is about 35 h^{-1}Mpc.
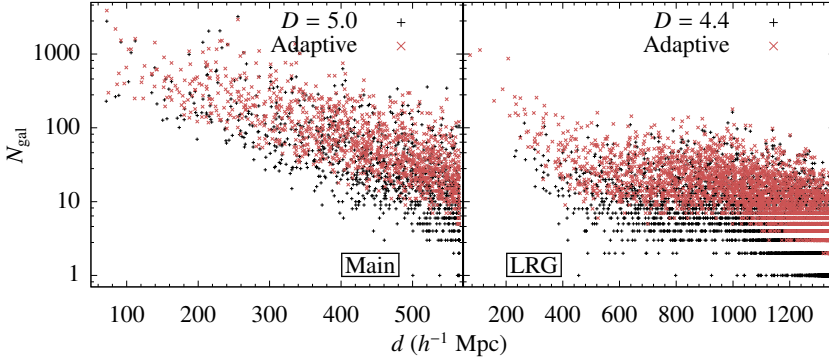
Figure 3.2: Number of galaxies in superclusters vs their distance for the fixed threshold (black symbols) and adaptively derived density levels (red symbols). The fixed threshold is $D = 5.0$ for the main (*left*) and $D = 4.4$ for the LRG sample (*right*).

located outside supercluster boundaries, depending on the shape of the supercluster. Figure 3.1 gives an example of the difference between these estimates.

3.1.2 Richness

Supercluster richness N_{gal} is the count of all galaxies inside the contour of a specific object of a given density threshold. Figure 3.2 shows the number of galaxies in
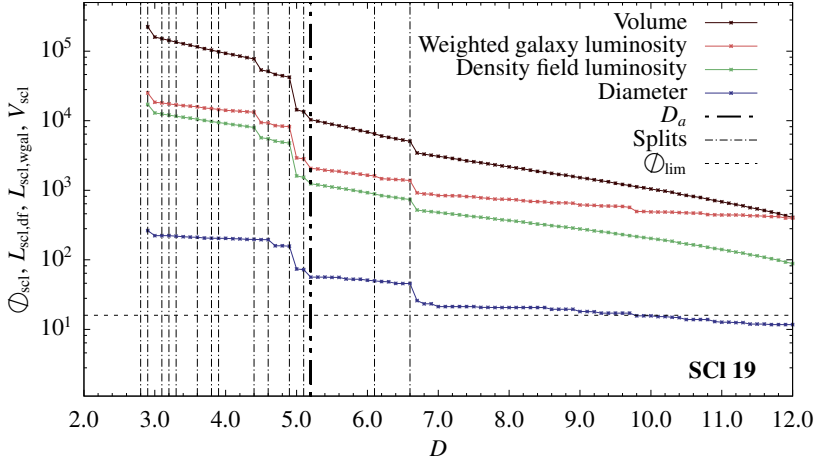
Figure 3.3: An example (supercluster SCl 19 in the catalogue) of the dependence of the supercluster diameter, total luminosity defined by the density field and galaxies, and supercluster volume on the density level D. Vertical thin dashed lines show splitting/merger events, and the thick dashed line shows the adaptive density threshold. The minimum diameter limit ($16\ h^{-1}$Mpc for the main sample) is also shown. Lines begin at the density level where the object separates from the larger structure.

superclusters for the main and LRG samples. As we use flux-limited galaxy samples, the richness depends heavily on the distance. The distributions of both samples differ as they have different selection functions. We can also see that the adaptive threshold catalogue has in general richer superclusters, and also, there are noticeably less very poor objects (especially in the LRG sample). For the SDSS main sample, where the galaxy groups are available, we can also find the number of clusters and groups N_{gr} in a supercluster.

3.1.3 Size and volume

The supercluster volume is derived from the density field as the number of grid vertices inside the supercluster boundary, multiplied by the cell volume:

$$V_{scl} = N_{cell \in scl} \Delta^3, \tag{3.2}$$

where Δ is the grid cell size. We define the supercluster diameter $\mathcal{D}_{scl}$ as the maximum distance between its galaxies. Figure 3.3 gives an example of how supercluster volume and diameter change with the density level. If defined as above, the volume of an object depends strongly on the density threshold, because a change in the limiting threshold can significantly increase or decrease the number of grid vertices in the

Figure 3.4: Supercluster diameters vs distance for fixed and adaptive thresholds. The main sample is shown in the *upper row* and LRG in the *lower*. Red symbols mark the diameters of individual superclusters; solid line shows the average within 10 h^{-1}Mpc bins for the main and 25 h^{-1}Mpc bins for the LRG sample. The standard deviations in bins are shown with grey contours.

supercluster. At the same time, the galaxy content inside an object boundary is more invariant to such changes. We see that major shifts in the diameter are usually caused by splits/mergers with similarly sized objects. When looking at the relationship between the supercluster diameter and distance (Fig. 3.4), we see that on average, there does not seem to be any serious systematic effects – while we may lose dimmer galaxies with the increasing distance, the brighter ones still mark the supercluster region sufficiently (Tempel 2011). The average and dispersion of the diameter distribution are both reasonably constant for the large part of all samples. Due to the cone-like shape of the sample volume, the spread of diameters is somewhat smaller at closer distances (a smaller volume makes it less likely to find large objects). One can also notice a significant decline in the average size only at the far end of the LRG sample (and perhaps to a lesser extent in the main sample), where the sample apparently

becomes too sparse to form superclusters consistently at this smoothing scale. These spurious objects are not present in the adaptive threshold catalogue.

The distribution of supercluster diameters for a set of fixed density thresholds is shown in the left panels of Fig. 3.6. The distribution maxima are at the same scale of about $10\,h^{-1}$Mpc. The number of smaller objects decreases, because we are rejecting ones with no galaxies inside. At larger diameters we can see series of maxima in the distribution. These indicate structures, which are distinctively larger than most of the objects, and they are present even at high density thresholds ($D = 8.0$). Moving to lower thresholds ($D = 6.0$), the total number of objects rises while the diameters also increase. The maxima caused by very large structures also become more pronounced, as they start to absorb a growing number of smaller objects. At one point, they become separated from the main part of the distribution and start moving towards larger sizes ($D = 4.0$). At even lower densities, below percolation, there is one enormous structure that extends through the whole volume (at $D = 2.0$). The lines which correspond to the adaptive catalogue start at the minimum diameter limit and show no very large structures. They are generally higher than for the fixed threshold supercluster distributions because they include objects from several density levels. The shape of the curves is very similar for all three samples, even if the absolute number of objects differs. The total amount of objects is the smallest in the LRG sample, which is due to the significantly lower number density of galaxies. When comparing the SDSS main and Millennium catalogues, we can see that the curves have comparable heights on all density levels. Also, the shape of the distribution at the highest density level ($D = 8.0$) signals the existence of large, Sloan Great Wall-like systems. The picture differs for the lowest threshold ($D = 2.0$), where in addition to the largest, percolated structure, there are also a couple of other sizeable structures present. However, we would like to stress that the properties of low-density regions in observational samples are probably strongly affected by various selection effects. On closer inspection, we can also see that the objects in the Millennium sample are systematically smaller at higher densities and in the adaptive catalogue than in the SDSS main sample, which does not contradict earlier similar findings (Einasto et al. 2006).

3.1.4 Luminosity

Similarly to the volume, we can also estimate the total luminosity of a supercluster as the sum of the luminosity densities at the grid vertices within the supercluster

$$L_{\text{scl,df}} = \sum_{\mathbf{i} \in scl} \ell_{\mathbf{i}} \cdot \Delta^3, \tag{3.3}$$

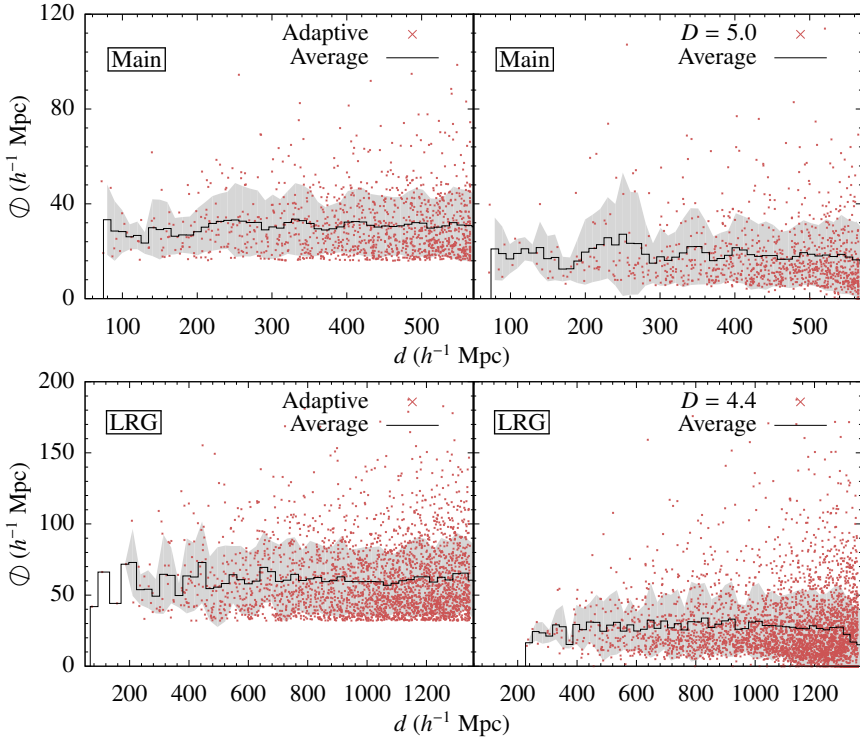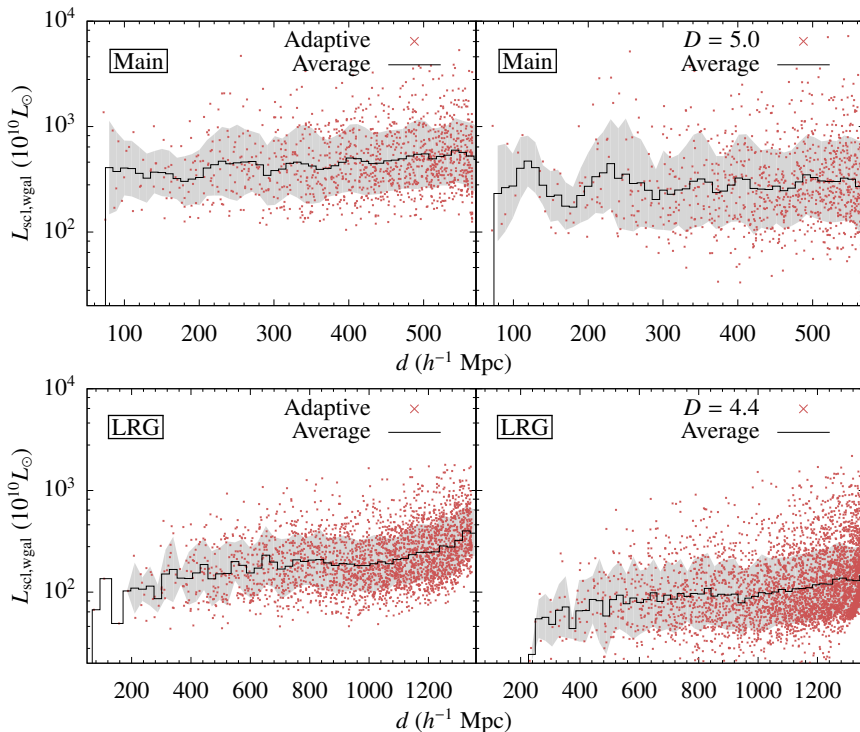**Figure 3.5**: Supercluster total weighted luminosities vs distance for fixed and adaptive thresholds. The main sample is shown in the *upper row* and LRG in the *lower*. Red symbols mark the luminosities of individual superclusters; solid line shows the average within 10 h^{-1}Mpc bins for the main and 25 h^{-1}Mpc bins for the LRG sample. The standard deviations in bins are shown with grey contours.

and, also similarly to the supercluster volume, it is very sensitive to the density level (Fig. 3.3). More robust estimates for the total luminosity of the supercluster are obtained by using its member galaxy luminosities – the sum of the observed galaxy luminosities, and the sum of the weighted galaxy luminosities:

$$L_{\text{scl,gal}} = \sum_{\text{gal}\in\text{scl}} L_{\text{gal}}, \tag{3.4}$$

$$L_{\text{scl,wgal}} = \sum_{\text{gal}\in\text{scl}} L_{\text{gal,w}}. \tag{3.5}$$

The latter is distance-corrected and therefore probably the most reasonable estimate of the supercluster brightness. We assume that we have taken into account most of the distance dependent selection effects and have restored the overall luminosity. And, as

Figure 3.6: *On the left*: supercluster diameter distributions in the LRG, main, and Millennium samples (*from top to bottom*). *On the right*: supercluster weighted luminosity distributions in the LRG, main, and Millennium samples (*from top to bottom*). Different lines correspond to different density levels $D = 2.0, 4, 0, 6.0, 8.0$, and also to adaptive thresholds for superclusters.

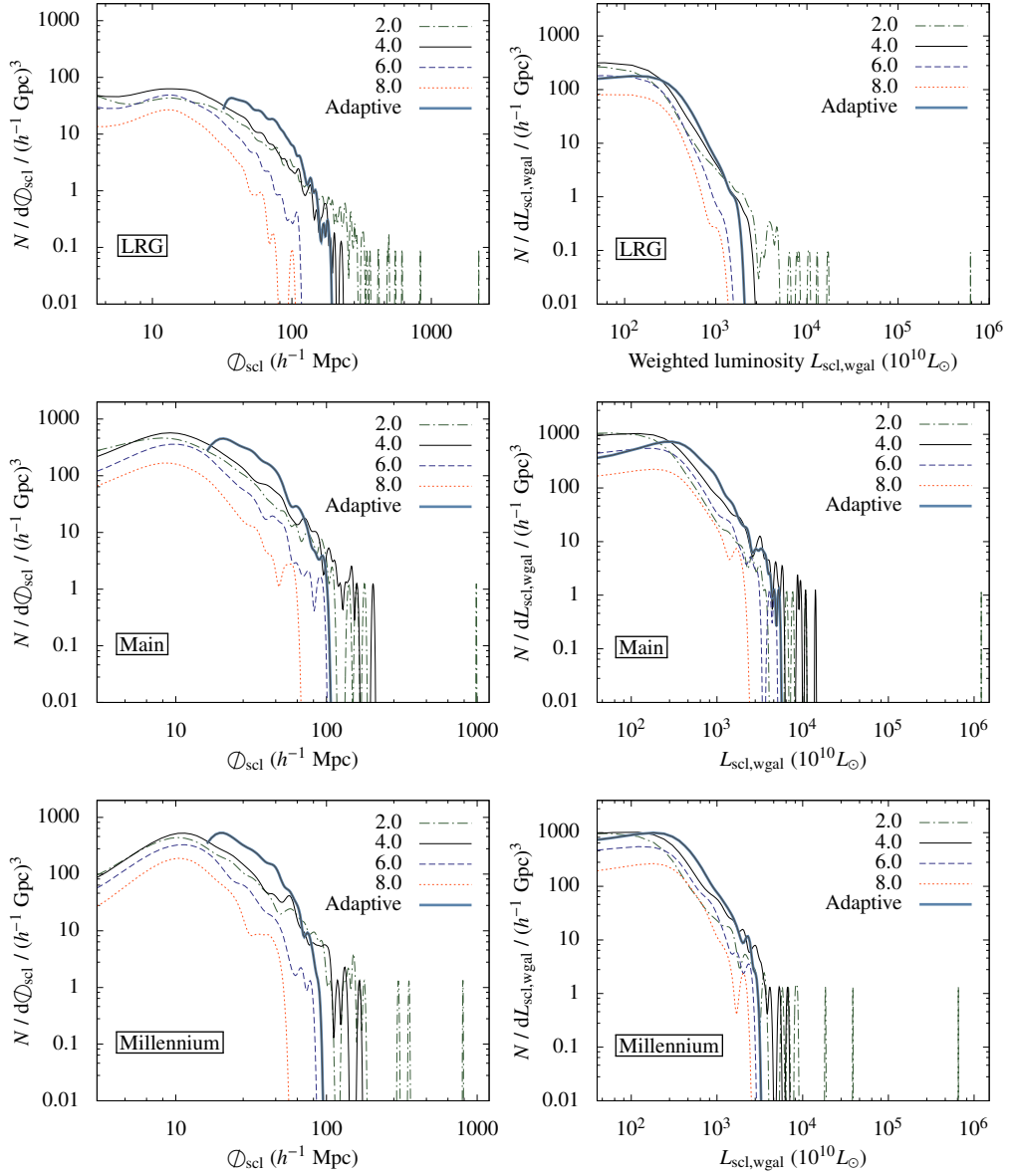the diameter, it is less dependent on the exact shape of the supercluster boundary. As it can be seen in Fig. 3.3, the galaxy-based luminosity estimate is always higher than the field-based one, and the difference roughly shows how much of the smoothed luminosity, that is located in the wings of the density profiles, is cut by the supercluster border. Compared to the physical dimensions, supercluster luminosities are slightly more dependent on the distance (Fig. 3.5). When comparing diameters and luminosities, we must conclude that superclusters of comparable size are more luminous at larger distances, which implies some overweighting. This is especially apparent in the SDSS LRG sample. However, we remind that because of the different weighting of the LRG luminosities, which is not based on the luminosity function, the weighted luminosity cannot there be used as an approximation to the total luminosity in the same manner as for the main sample. As was the case with diameters, the dispersion of the supercluster luminosities remains unchanged with the distance. The luminosity distributions of superclusters for different density thresholds (Fig. 3.6, right panels) show characteristics similar to those of the diameter, although they appear noisier due to the larger range of values.

When looking at the properties of the superclusters in the adaptive catalogue, they are clearly influenced by the hard cut of the diameter. In future analysis, perhaps a catalogue created by purely persistence filtering is able to perform better.

3.1.5 Confidence estimates

Using the signal-to-noise field G (see Section 2.2.3), we can calculate a confidence estimate for each supercluster

$$C_{\text{scl}} = \frac{1}{N_{\text{gal}}} \sum_{\text{gal}\in\text{scl}} G(\mathbf{r}_{\text{gal}}).$$

(3.6)

We interpolate the signal-to-noise ratio values of the density estimate to the galaxy positions and find the average over all galaxies belonging to the supercluster. Figure 3.7 displays the dependence of the supercluster confidence estimates on the supercluster richness (the number of its member galaxies) and distance. Both plots exhibit expected behaviour – the confidence estimates diminish with distance, and richer superclusters also have generally higher signal-to-noise ratios. Predictably, the confidence estimates for superclusters in the LRG sample are significantly lower. The confidence estimates depend on the density threshold, but at lower density levels, more galaxies from the density field regions with higher variance are included. Because of that, fixed threshold superclusters have slightly higher confidence estimates in Fig. 3.7. We also check whether a supercluster touches the edge of the pixel mask of the density field. A location next to the sample boundary implies incompleteness
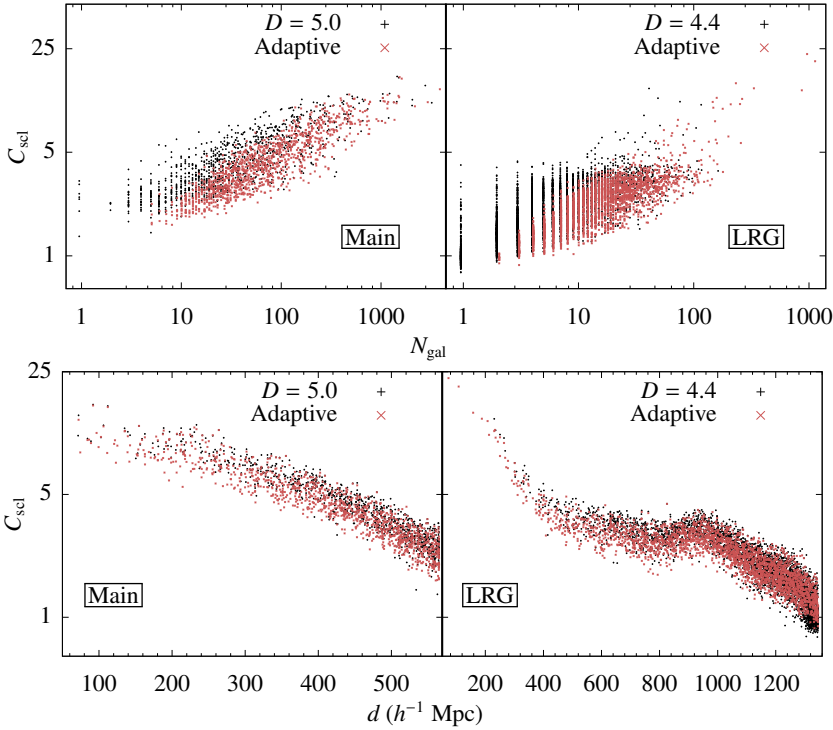
Figure 3.7: Supercluster confidence estimates vs their richness (the number of member galaxies) (*top panels*) and vs their distance (*bottom panels*). The superclusters defined by a fixed threshold are marked with black, and those found using adaptive thresholds with red points.

of a supercluster, and its parameters may not be reliable. These properties can be used to select samples of structures for further studies.

3.1.6 Density thresholds

We studied the distribution of structure splits in the previous chapter (see Sect. 2.3.2). We now look at the distance dependence of the adaptively assigned supercluster thresholds in Fig. 3.8. The results are very encouraging for the main sample, where the individually derived thresholds seem to be alike regardless of the location in the sample volume, while in the LRG sample we can observe a slight upward trend with the distance. Also, we can clearly discern the effects of specific structures – the Sloan Great Wall and the SCl 1 region superclusters cause the two peaks between $200\ h^{-1}$Mpc and $300\ h^{-1}$Mpc. The choice of the threshold selection method (adaptive or fixed) has profound effects on the spatial distribution and even the nature of the objects and we look at it in greater detail in Sect. 3.4.
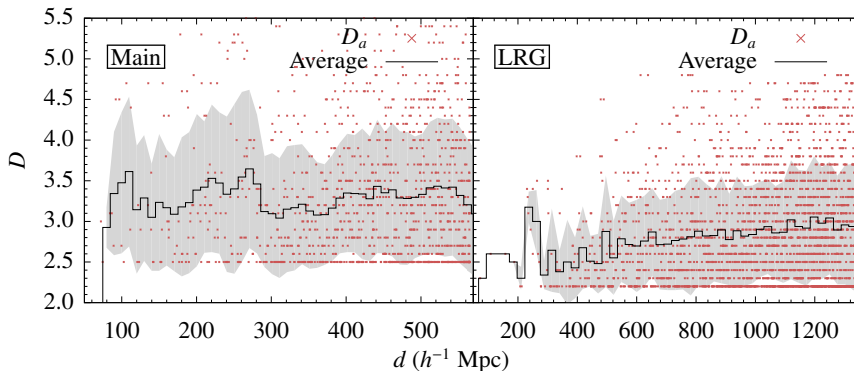
Figure 3.8: Adaptively assigned thresholds for individual superclusters (red symbols) in the main (*left*) and LRG samples (*right*). The average and the standard deviation vs distance are shown with the black line and grey contours, correspondingly.

3.1.7   Naming

By default, the objects in our catalogue are numbered by the order of the peak density value. However, a somewhat random number from a list may not be the clearest way to identify objects. Therefore, we associate superclusters with a "marker galaxy" that is located near the highest density peak in the supercluster volume. The aim of this is to connect a supercluster to an actual observational object and to construct an identifier that is not specific to the current catalogue. The long identification number is given in the following format: $AAA \pm BBB + CCCC$, where AAA and BBB are the integer parts of the equatorial coordinates α and δ of the marker galaxy and $CCCC$, its redshift multiplied by 1000.

The extended list of supercluster attributes in the catalogue is given in Appendix A.1.

3.2   Morphology of superclusters

Superclusters can exhibit very different shapes and sizes (see, e.g., Fig. 2.12). That inspired us to try another another angle for research – morphology (literally the "study of shapes"). In principle, morphology of superclusters can serve as a means to compare observations with simulations and perhaps to distinguish between different cosmological models (Kolokotronis et al. 2002; Einasto et al. 2007a). Superclusters also possess diverse internal structure and a wide range of environments: dense cores of clusters, filamentary parts in-between and also lower-density outskirts, where the boundary between the supercluster and the field is already fading. All these may have distinct effects on the evolution of galaxies and therefore the distribution and

64

properties of galaxies (Einasto et al. 2014a). The same goes for groups and clusters of galaxies. Here, morphological studies can offer a more sophisticated insight into the environments' impact on their properties (Aragón-Calvo et al. 2010; Einasto et al. 2012). First, we present an overview of the morphological properties of superclusters and later, we conduct a more throrough analysis of a few specific objects.

3.2.1 Minkowski functionals and shapefinders

In our approach, a supercluster is in essence an excursion set of the luminosity density field. The supercluster geometry is defined by its boundary isodensity surface and enclosed volume, and is thus completely described by the four Minkowski functionals $V_0 \ldots V_3$ (Einasto et al. 2011b). If we take an excursion set F_{ϕ_0} of a field $\phi(\mathbf{r})$ (i.e. the set of points where the density is above the limit $\phi(\mathbf{r}) \geqslant \phi_0$), then the first Minkowski functional is the volume of the region:

$$V_0(\phi_0) = \int_{F_{\phi_0}} dV. \qquad (3.7)$$

The second functional is proportional to the surface area of the set boundary δF_ϕ:

$$V_1(\phi_0) = \frac{1}{6} \int_{\delta F_{\phi_0}} dS(\mathbf{r}). \qquad (3.8)$$

The third one is proportional to the integrated mean curvature C of the boundary:

$$V_2(\phi_0) = \frac{1}{6\pi} \int_{\delta F_{\phi_0}} \left(\frac{1}{R_1(\mathbf{r})} + \frac{1}{R_2(\mathbf{r})} \right) dS(\mathbf{r}), \qquad (3.9)$$

where R_1 and R_2 are the boundary's principal radii of curvature. The first three functionals are sometimes denoted in the literature also by V, S and C. The fourth Minkowski functional is proportional to the integrated Gaussian curvature of the boundary surface:

$$V_3(\phi_0) = \frac{1}{4\pi} \int_{\delta F_{\phi_0}} \frac{1}{R_1(\mathbf{r})R_2(\mathbf{r})} dS(\mathbf{r}). \qquad (3.10)$$

The fourth functional is also known as the Euler characteristic χ and is closely related to the topological genus g: $V_3 = \chi = (1 - g)/2$. This functional gives us the sum of isolated clumps and void bubbles in the sample, subtracted by the number of tunnels (Martínez & Saar 2002; Saar et al. 2007):

$$V_3 = N_{\text{clumps}} + N_{\text{bubbles}} - N_{\text{tunnels}}. \qquad (3.11)$$

Apart from the fourth one, we do not use the functionals directly, but rather, the combinations of them called shapefinders (Sahni et al. 1998; Shandarin et al. 2004).

There are three shapefinders with dimension of length: the thickness $H_1 = 3V_0/V_1$, width $H_2 = V_1/V_2$ and length $H_3 = V_2/4\pi$. Sahni et al. (1998) also defined two dimensionless characteristics: the planarity $K_1 = (H_2 - H_1)/(H_2 + H_1)$ and filamentarity $K_2 = (H_3 - H_2)/(H_3 + H_2)$. As the names suggest, the ratio between the values of K_1 and K_2 should tell us if an object has a planar pancake-like overall shape or is it more elongated. In our short overview of the supercluster morphology we use the fourth Minkowski functional V_3 and the dimensionless shapefinders K_1, K_2.

In order to obtain more information about the inner structure of a supercluster, we implement the following scheme. We calculate the fourth Minkowski functional and shapefinders for the series of isodensity surfaces in an individual supercluster in the increasing order. The surfaces are chosen with constant intervals of the excluded mass fraction m_f, i.e. the ratio between the mass at the lower density and the total supercluster mass. The argument mass fraction value m_f for a supercluster is zero when the entire supercluster is enclosed by the isodensity surface, and one, when we have reached the highest peak density. We find and draw the relation between V_3 and m_f, which characterises the richness of the internal structure, e.g., the number of high-density cores. We also plot pairs of (K_1, K_2) depending on the m_f, the so-called morphological signature, which is unique to each object (Einasto et al. 2007c). In the catalogue, for each supercluster, the maximum value of $V_{3,\max}$ is given, as well as K_1 and K_2 for the base threshold (i.e. where $m_f = 0$). Values of the Minkowski functionals were calculated using the grid-based algorithm described in Saar et al. (2007).

The values of shapefinders K_1, K_2 and their ratio are somewhat difficult to interpret. Sahni et al. (1998) studied numerically the dependence of shapefinders on various shapes of ellipsoids and toroids. We use these results as an approximate guide to classify our objects. For example, ellipsoids, where two semi-axes are about the same length and significantly larger than the third semi-axis, have a large K_1 and a small K_2 value, while ellipsoids with one large and two much smaller semi-axes have low K_1 and high K_2 values. Thus, in general, $K_1/K_2 > 1$ indicates a planar and $K_1/K_2 < 1$ an elongated or filamentary form. For a spherical body, both K_1 and K_2 are zero, and in the case of a triaxial ellipsoid, shapefinders have also similar, but non-zero values. Negative values of shapefinders can imply the presence of concave features like dents or dimples in the shape of an object; however, they can also be a result of numerical noise. The shapefinders are to some degree proportional to object sizes. For all these reasons, the shapefinder ratios for individual structures can sometimes be the divisions of two very small numbers, and therefore noisy and largely meaningless. In order to find models for supercluster morphology, Einasto et al. (2007c) studied the morphological parameters of various more complex test objects, such as "kitchen table", etc.
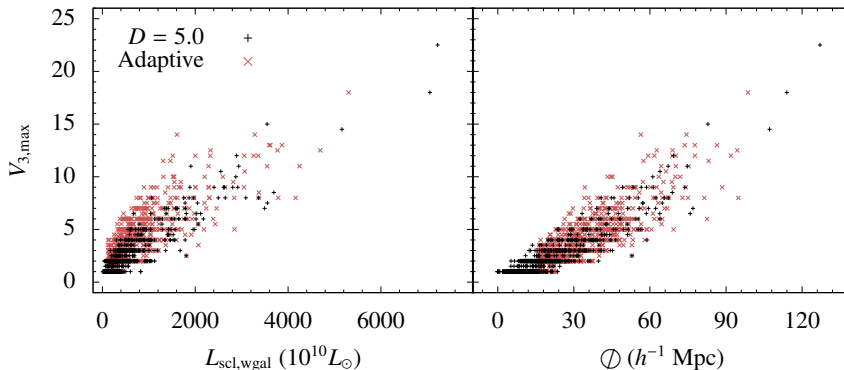
Figure 3.9: Dependence between the supercluster total weighted luminosity (*left panel*), its diameter (*right panel*) and the maximum value of the third Minkowski functional $V_{3,\mathrm{max}}$, i.e. the clumpiness in fixed (black symbols) and adaptive threshold samples (red symbols).

Supercluster shapefinders have been shown to be sensitive to cosmological models by Kolokotronis et al. (2002), who found that cosmological simulations with a significant neutrino contribution produce superclusters with demonstrably different (flatter) shapes when compared to the standard ΛCDM model. In observations, the shapes of large-scale structures are affected by the Kaiser effect, which causes the apparent flattening of objects along the line-of-sight.

3.2.2 Supercluster morphology overview

In Einasto et al. (2011b), we studied superclusters from the SDSS DR7 catalogue. To minimise selection effects, the supercluster sample was limited to distances from 90 h^{-1}Mpc to 320 h^{-1}Mpc, and to the objects with the number of member galaxies at least 300. Also, only one constant density level ($D = 5.0$) catalogue was used. Here, we try to study the whole volume of the main galaxy sample and all superclusters, regardless of their richness, using both adaptive and fixed threshold catalogues.

From Fig. 3.9 we can establish that, rather reasonably, the relation between $V_{3,\mathrm{max}}$ and the total weighted luminosity of supercluster, as well as the diameter, is monotonously increasing and even approximately linear. Figure 3.10 shows the distribution of the shapefinders K_1 and K_2 of all objects in the adaptive and fixed density level ($D = 5.0$) catalogues. Since we intend to expand on the analysis of Einasto et al. (2011b), we break the supercluster sample into three parts. First we create a subset that is comparable to the one used in Einasto et al. (2011b) – distances are limited to the range from 90 h^{-1}Mpc to 320 h^{-1}Mpc, but instead of setting a limit of 300 to the minimum number of member galaxies, we recquire the total weighted luminosities to be greater than $L_{\mathrm{scl,wgal}} > 400 \times 10^{10} L_\odot$. This is necessary if we wish to find
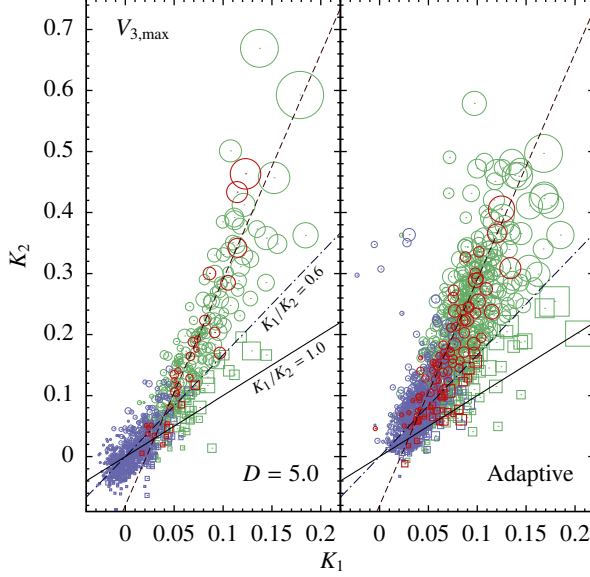
Figure 3.10: Distribution of the shapefinders K_1 and K_2 in the fixed (*left panel*) and adaptive threshold (*right panel*) supercluster samples. Objects with the shapefinder ratio $K_1/K_2 < 0.6$ are marked with circles and objects with $K_1/K_2 \geqslant 0.6$ with squares. Red symbols are used to denote superclusters with a total weighted luminosity greater than $L_{\mathrm{scl,wgal}} > 400 \times 10^{10} L_\odot$, that are within the distance interval of $d = 90..320\ h^{-1}$Mpc, green symbols present superclusters with the same luminosities but outside the distance limits, and blue symbols stand for objects at all distances and less luminous than $400 \times 10^{10} L_\odot$. Symbol sizes are proportional to the maximum value of the fourth Minkowski functional $V_{3,\mathrm{max}}$. The $K_1/K_2 = 0.6$ and $K_1/K_2 = 1.0$ guides are shown with the dash-dot and solid lines.

similarly sized superclusters outside the distance range where the selection effects begin to affect the number of observed galaxies more severely. This subset is shown with red symbols in the plot. Green symbols denote superclusters that are also more luminous than $400 \times 10^{10} L_\odot$, but are located outside that distance interval (closer than $90\ h^{-1}$Mpc or further away than $320\ h^{-1}$Mpc). Objects with total luminosities less than $400 \times 10^{10} L_\odot$ (at all distances) are drawn with blue symbols. Symbol sizes are proportional to $V_{3,\mathrm{max}}$, i.e. the clumpiness. Figure 3.10 thus reflects both the inner structure and the outer shape of superclusters.

Using multidimensional normal mixture modelling (`mclust` package in *R* by Fraley & Raftery (2006)), Einasto et al. (2011b) concluded that superclusters could be grouped into two subsets depending on the ratio of the shapefinders, with the division going along the line where $K_1/K_2 = 0.6$. We are not repeating these procedures here but rather make a cursory comparison of the results. We must also stress that it would

not be correct to call the second group of objects more *planar*, since for both groups the shapefinder ratio is generally $K_1/K_2 < 1$, and both are therefore generally elongated. Looking at the common sample (red symbols in the left-hand plot) we see a very similar trend between the values of K_1 and K_2 as in Einasto et al. (2011b), which is expected as the sample should be nearly the same. In general, it is immediately obvious that most of the objects in the catalogue are of elongated, filamentary type with only a small number of them having $K_1/K_2 > 1$. With only very few exceptions, the majority of larger objects are strongly elongated and the shapefinder ratio tends to increase with the supercluster size. The largest two superclusters have the same ratio of $K_1/K_2 \approx 0.3$, and moreover, if we linearly fit a line through the red sample we find that they follow rather well the relation $K_1 = 0.3\,K_2 - \epsilon$ (dashed line), where $\epsilon \approx 0.1$. Adding the previously discarded large superclusters to the mix (green symbols), we can see that these are placed in a very similar fashion with only a somewhat increased scatter. The largest of the added superclusters also possess $K_1/K_2 < 0.3$ ratios of the shapefinders. Still, one can see that new objects tend to have larger shapefinder and clumpiness values, which may be the result of selection effects. There are now also some superclusters present with triaxial shapes ($K_1/K_2 \approx 1, K_{1,2} > 0$), and a couple of superclusters which have both a substantial size and planar morphology ($K_1/K_2 > 1, K_{1,2} > 0$). The last subset of low-luminosity objects (blue symbols) has distinctly different properties – both shapefinders and clumpiness have very small values, and shapefinders can be zero or even negative. This is characteristic to small roughly spherical bodies and smoothed density isosurfaces of a small group of galaxies are very likely to have a more or less shpere-like shape. Looking at the clumpiness of the superclusters (shown by the size of the symbols), we see again that the objects gathering near the origin have only one or two higher density cores. We can also see that the relation between K_1 and K_2 has a noticeably different slope than that for larger objects, being closer to the $K_1/K_2 = 0.6$ line. In the first two samples 70% and 67% of the superclusters have the shapefinder ratio below 0.6; in the third sample, the fraction drops to 42%. Combining the observations from Figures 3.10 and 3.9, it can be fairly safely concluded that shape-wise, superclusters are basically massive filaments of galaxies. Bigger structures are made by concatenating richer clusters (together with their surroundings) like links to a chain.

These results are mostly in concordance with other studies dedicated to the shapes of large-scale galaxy structures (Zeldovich et al. 1982; Basilakos 2003; Einasto et al. 2007c; Luparello et al. 2011). Costa-Duarte et al. (2011), who also use shapefinders to describe supercluster morphology, conclude that while larger structures are preferably more elongated, they also find that about half of the superclusters have planar shapes. However, this has not been supported by other studies. Einasto et al. (2011b, 2012) go into even more detail in the morphology analysis, dividing superclusters fur-
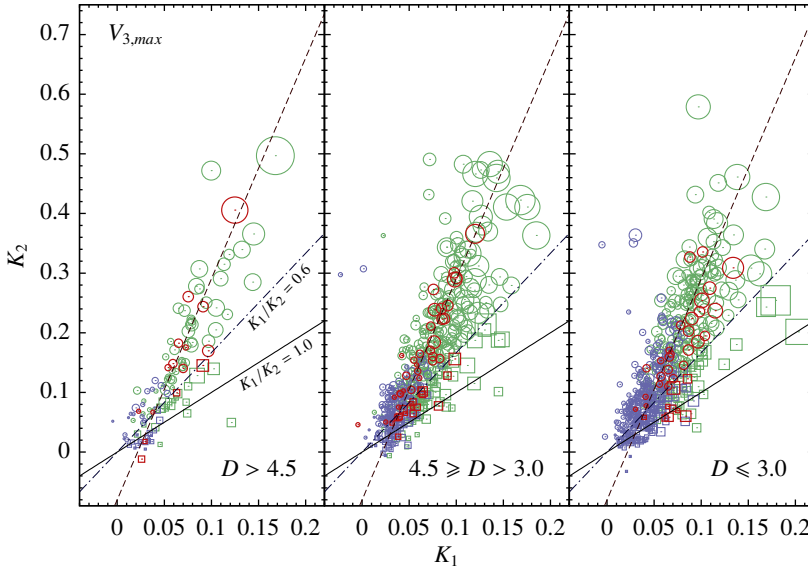
Figure 3.11: Distribution of the shapefinders K_1 and K_2 in the adaptive supercluster catalogue in different density threshold intervals. From *left* to *right*: $D > 4.5$, $4.5 \geqslant D > 3.0$ and $D \leqslant 3.0$.

ther into various subclasses like "filaments" and "spiders" depending on the number of high-density cores and visual impression.

We used the same conditions as before to create similar subsets for the adaptive level catalogue (the right panel in Fig. 3.10). Here, all the subsamples are more similar and the scatter is much bigger. Correspondingly, 67%, 77% and 61% of superclusters belong to the strongly elongated objects group. Due to the minimum diameter limit, there are no very small objects present and the number of objects with shapefinder values close to zero is smaller. They also do not seem to form a separate population on the plot. On the other hand, the largest superclusters are also slightly smaller, and they have more diverse combinations of shapefinders. While similarily to the previous cases, the majority of the objects have very elongated forms, there are now noticeably more large superclusters which have the shapefinder ratio implying triaxial or planar shapes. Breaking the adaptive catalogue into three density intervals (Fig. 3.11) reveals that high (boundary) density superclusters form a relatively tight group and most of the scatter, as well as most extreme cases of different shapes originate from the lower density objects. There can be multiple reasons for that. First, the structures in low-density environments are obviously less evolved. On the other hand, these regions of the density field are simply more noisy. In all three density intervals, about 70% of objects have the shapefinder ratio $K_1/K_2 < 0.6$. Taking the

edge effects into account, i.e. selecting only objects fully inside the sample volume, does not change the previous findings by more than a few percent.

3.2.3 Description of individual superclusters

In this section we take a closer look at some of the superclusters, trying to describe different types of objects. We have chosen a few of the largest superclusters, a couple of smaller and poorer objects, and also some well-known superclusters from the literature. We include objects from both near and far distances, which were previously usually omitted from the analysis due to selection effects. All superclusters are taken from the adaptive threshold catalogue and have varying boundary densities, meaning they come from high and low density regions which were described in Sect. 3.4.2.

Selected superclusters are shown in Figures 3.13 – 3.16. Left column displays the appearance of the supercluster: its galaxy and group distribution in Cartesian coordinates. If there are any Abell clusters present, they are also indicated. In the middle column we show the fourth Minkowski functional depending on the excluded mass fraction m_f. The curves start with $V_3 = 1$ at $m_f = 0$, when all of the supercluster is enclosed in the isodensity surface. When raising the density level we begin leaving parts of the supercluster outside, and as a result, the value of V_3 increases when the isodensity surface is broken into separate pieces, and decreases if either an isolated clump disappears or a tunnel is formed through an object. Thus, the plot illustrates the amount and diversity of the internal structure of a supercluster. The morphological signature is shown in the right column. The curves show the objects' tracks in the (K_1, K_2) plane as m_f increases. The value of the mass fraction m_f is indicated with the line colour. Morphological signatures do not have a clear quantitive meaning, mostly because they break down when the object is divided into multiple parts and cease to describe it as a whole. Nonetheless, they are unique to an individual supercluster and there are systematic correspondences between the (K_1, K_2) curves and the object's properties. The starting point of all signature curves designates the shapefinder combination describing the overall shape of the object. Most signatures start at the right hand side of the plot and then draw a counter-clockwise arc towards the left. Sometimes the K_1 value increases slightly, however when the object starts to fragment, the curve turns towards higher K_2 and lower K_1 values. The maximum value of K_2 (the height of the curve) depends strongly on the size of the supercluster. Large jumps in the morphological signature occur when fragmentation results in objects of very different shapes, and when the signature is mostly smooth, all pieces have largely similar shapes. The fourth Minkowski functional curve shows that tunnels are first to appear (if an object is large enough), and later the object is split as V_3 starts to rise. Later, at higher m_f values, it decreases again as individual clumps fall below the threshold. In the morphological signature plot, this corresponds to the

71
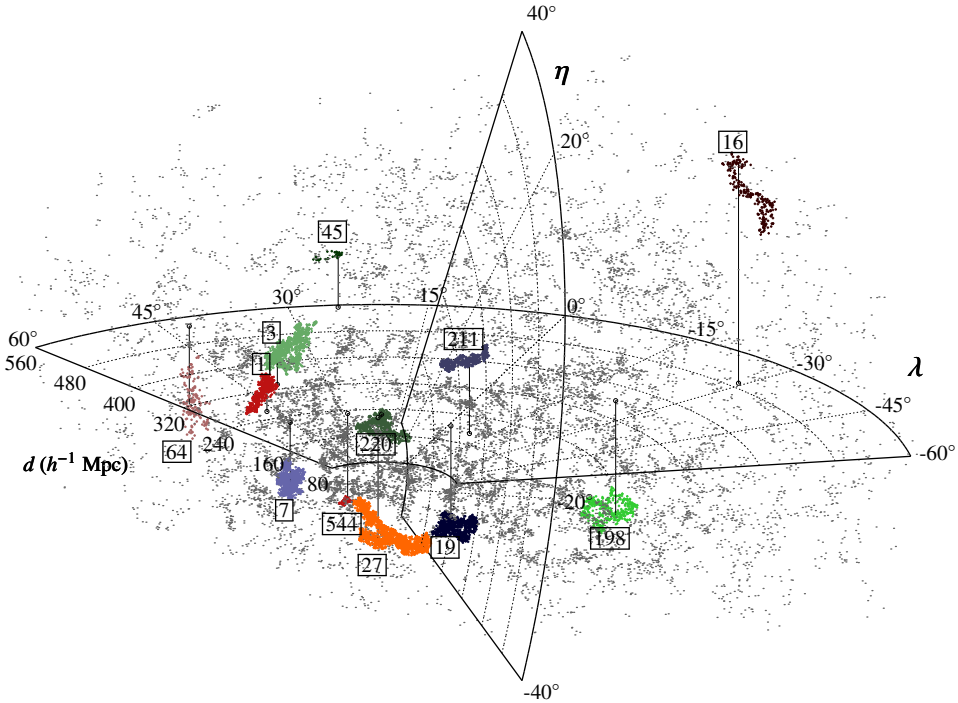
Figure 3.12: Spatial placement of the example objects. Supercluster short identification numbers are given with framed labels.

part where the (K_1, K_2) curve turns downwards. The locations of all described superclusters in the survey volume are shown in Fig. 3.12 and the basic properties of the superclusters are listed in Table 3.1.

Supercluster SCl 1 is the first object in the catalogue, its designator number 1 indicates that it contains the highest luminosity density peak in the survey volume. This is mostly due to the very rich system Abell A2142, which is an ongoing merger between two galaxy clusters (Markevitch et al. 2000), and the system is also a bright X-ray source. Supercluster SCl 1 is a part of a larger system of high boundary density superclusters, somewhat reminiscent of the Sloan Great Wall (Fig. 2.12). It is one of the most unusual-looking superclusters, being quite strongly elongated and very straight, and pointed at about 45° from the line-of-sight towards us. Other than that, the supercluster SCl 1 has a relatively simple morphology with only two higher density cores. The region with the highest density may be undergoing collapse (Einasto et al. 2015).

Supercluster SCl 3 is a relatively large supercluster in a relatively high density region. It is located not far from the SCl 1 supercluster complex, although it is at a 40 h^{-1}Mpc
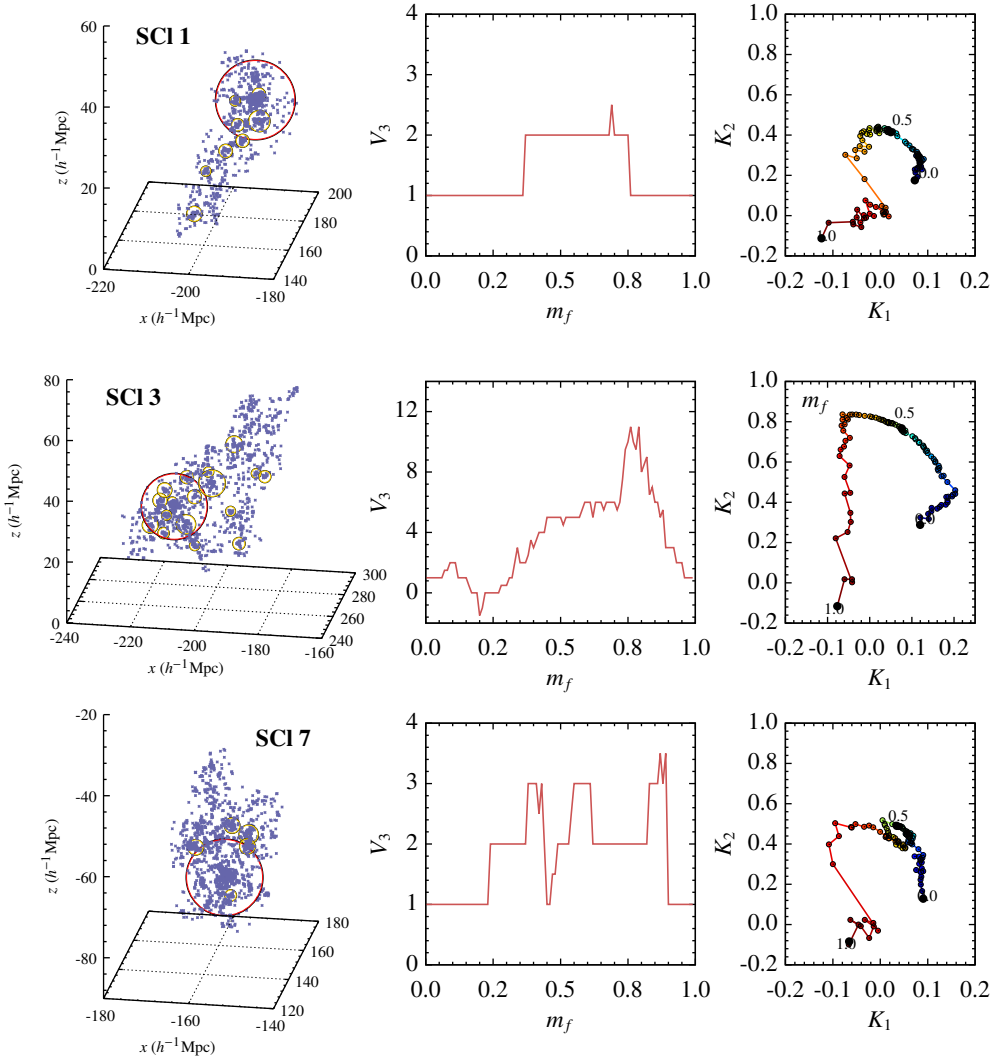
Figure 3.13: Superclusters Scl 1, SCl 3 and SCl 7: the distribution of galaxies (blue dots) and large galaxy clusters with luminosities over $L_{gr} > 30 \cdot 10^{10} h^{-2} L_\odot$ (black circles) in superclusters (*left panel*). Group symbol sizes are proportional to the sum of weighted galaxy luminosities in the group. Red circles mark the Abell clusters in the supercluster, if present. We show the dependence of the fourth Minkowski functional V_3 for the isosurfaces ordered by the mass fraction m_f (*middle panel*), and the dependence of the morphological signature, i.e. the shapefinder pairs' (K_1, K_2) dependence on the m_f. Line colour changes according to the m_f value (blue – 1, red – 0) and the locations for m_f values 0, 0.5 and 1.0 are also marked with black circles and labels.

larger distance and separated by a lower density void area. Supercluster SCl 3 has a considerable number of clusters including one large Abell X-ray cluster A2069. Its morphology is more similar to that of very large superclusters, it has multiple tunnels and cores. However, it is dominated by one single cluster, which is the cause for the morphological signature to turn back towards zero.

Supercluster SCl 7 is a medium-sized supercluster in an intermediate density region. It is located at about the same distance as the supercluster SCl 1 and the Sloan Great Wall, and lies almost between them. It is relatively similar to the supercluster SCl 1, being of about the same size and having comparable morphological properties. It possesses a few high density cores, including a very large concentration of galaxies, the Abell cluster A2040.

Supercluster SCl 16 is a very large supercluster in a high density environment. It is formally the most luminous object in the catalogue. However, it is also at a large distance, almost at the far end of the galaxy sample, and is therefore probably strongly influenced by selection effects and galaxy luminosity weighting. Morphologically, it has a very elongated overall shape with some branching and there is a large number of tunnels and dense cores. The morphological signature is smooth and is confined to the upper part of the (K_1, K_2) plot.

Supercluster SCl 19 is a medium-sized supercluster in a high density region. It is part of the Sloan Great Wall extensive galaxy supercluster system. It is the second richest object in the system and actually contains the highest density region. The morphology of SCl 19 is somewhat comparable to that of the superclusters SCl 1 and SCl 7, however, it does not have an obvious single dominant core. The whole Sloan Great Wall region was analysed in detail by Einasto et al. (2007c, 2011c).

Supercluster SCl 27 is the largest object in the Sloan Great Wall (Einasto et al. 2011c) and the richest (having most galaxies) supercluster in the catalogue. It contains nine Abell clusters, including five X-ray clusters (Böhringer et al. 2004). The richest of them is the cluster A1750, which is a merging X-ray cluster. SCl 27 is located along the bottom survey boundary and is not completely inside the SDSS main sample volume. The morphology of SCl 27 looks typical for very large superclusters, with many high density cores and a strong elongation. Visually, the supercluster has a distinct shape with multiple thick branches.

Supercluster SCl 45 is a distant smaller sized supercluster with a high adaptively derived density threshold. It has the most planar shape (the highest shapefinder ratio) of the superclusters. However, due to the great distance a large number of galaxies is unobserved and the exact shape of the structure is hard to estimate. Its morphology can nonetheless be considered similar enough to that of closer and more complete superclusters of comparable sizes.
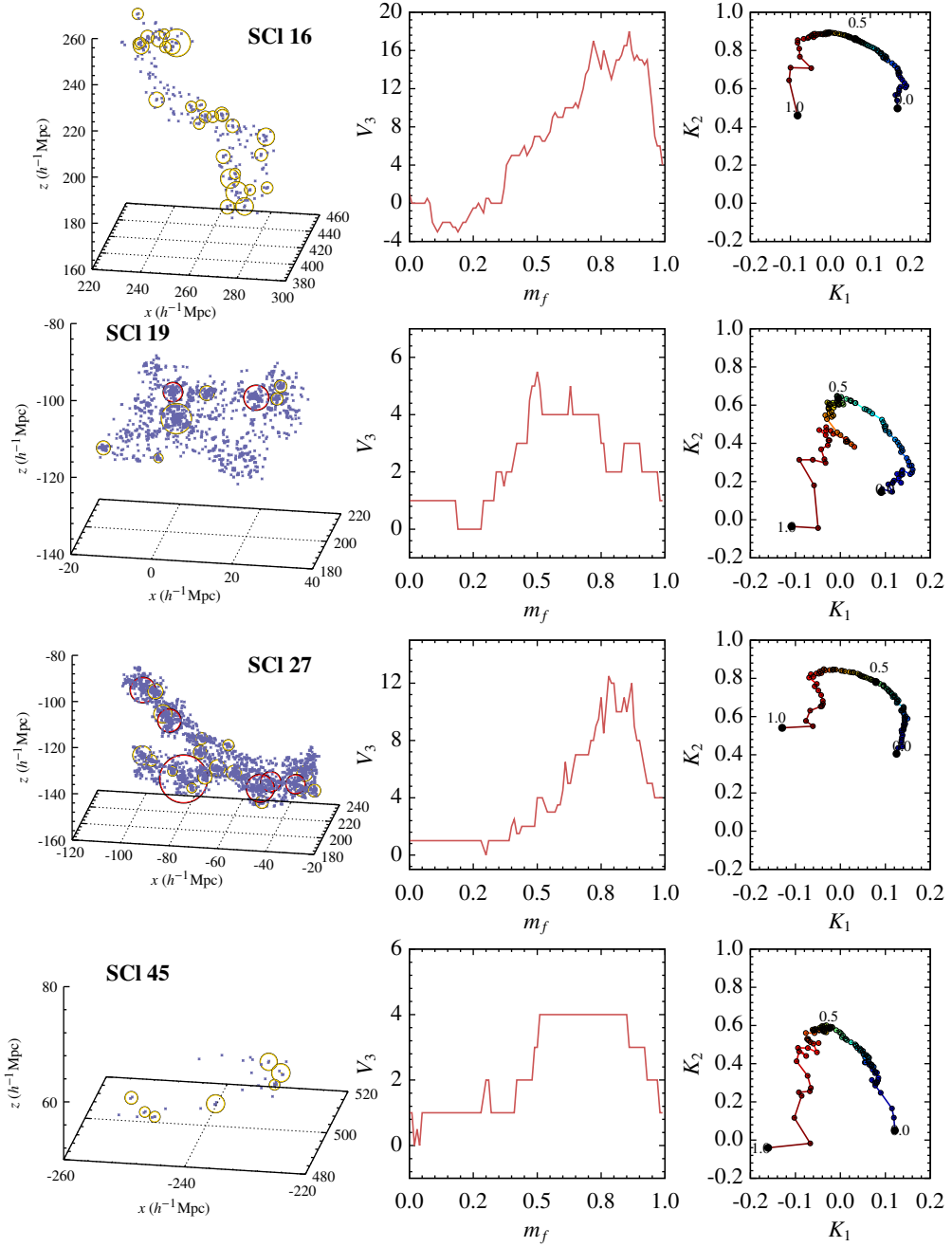
Figure 3.14: Superclusters SCl 16, SCl 19, SCl 27 and SCl 45. Panels and symbols are the same as in Fig. 3.13.

Supercluster SCl 64 is a very large supercluster in a low density region. It has the largest volume in the catalogue. Despite being at the far end of the survey and having a low density threshold, its morphology is typical to other extended superclusters.

Supercluster SCl 198 is a medium-large supercluster with a very low density threshold. The shapefinder values of SCl 198 indicate it has a rare planar overall shape. It is located in a poor environment and has an unusually high number of cores for its size.

Supercluster SCl 211 forms a large part of the Ursa Majoris supercluster. Historical superclusters described in the literature are usually larger than the objects in our catalogue. It contains two Abell clusters A1279 and A1436. It is a medium-sized supercluster with a high boundary density.

Supercluster SCl 220 belongs to the Boötes supercluster. It is of medium size and is located in an intermediate density environment. SCl 220 has multiple high-density cores, including the Abell clusters A1775 and A1831.

Supercluster SCl 544 is an example of a very small supercluster, located in a high density region. It belongs to the Sloan Great Wall and could also be considered as an extension of the SCl 27. Being so small, this object has a very simple morphology, it consists of only one clump. Based on its morphological signature, which mostly jumps around the origin, we can roughly approximate it by a slightly elongated ellipsoid.

It is difficult to draw any deep conclusions from these examples as our approach was purely empirical and the sample is very small. Especially when considering the obvious wealth of different shapes and sizes of selected objects. However, we can still make a few observations. First, there seems to be a fairly robust correlation between the supercluster morphology and its physical size. This is most apparent for the smallest and largest superclusters. While being fairly obvious for the small superclusters, which have very simple shapes, we can see that very large superclusters are also very much alike. For midsized superclusters, the relation between size and morphology appears to hold too, although there seems to be a slight difference whether the supercluster contains a dominant cluster or not. When taking into account the whole catalogue, there are no distinct classes of objects and the supercluster distribution is actually continuous. The second observation is that there is a visible rapport between objects in low and high density regions, meaning that similarily sized superclusters in rich and poor environments have also similar morphological properties. Concerning the largest objects, the Sloan Great Wall has long been touted as a very unique galaxy structure and a possible challenge to the models of structure formation (e.g., Einasto et al. 2006; Sheth & Diaferio 2011), while others say that the modelled galaxy superstructures are similar to the observed ones in both properties and
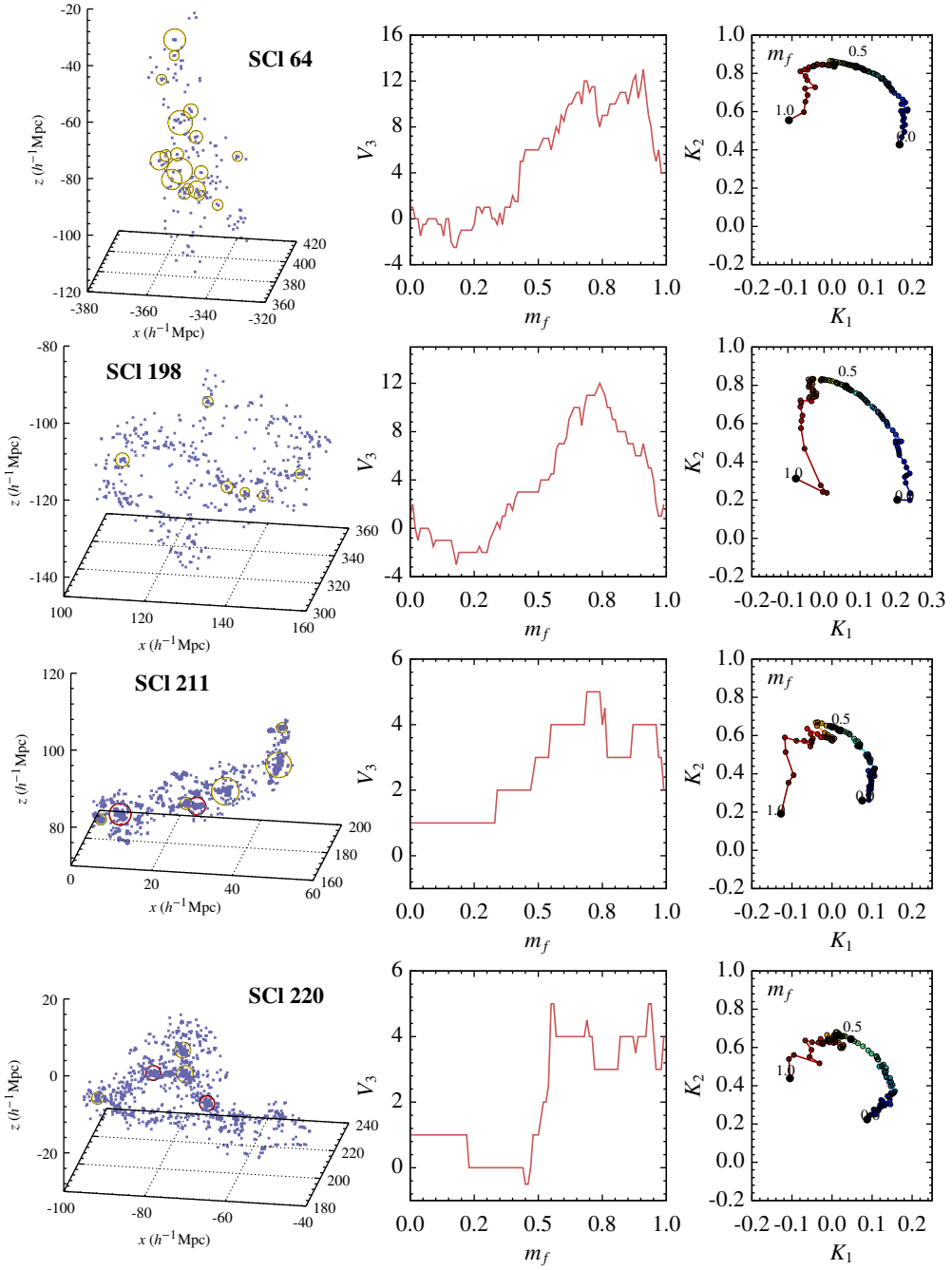
Figure 3.15: Superclusters SCl 64, SCl 198, SCl 211 and SCl 220. Panels and symbols are the same as in Fig. 3.13.
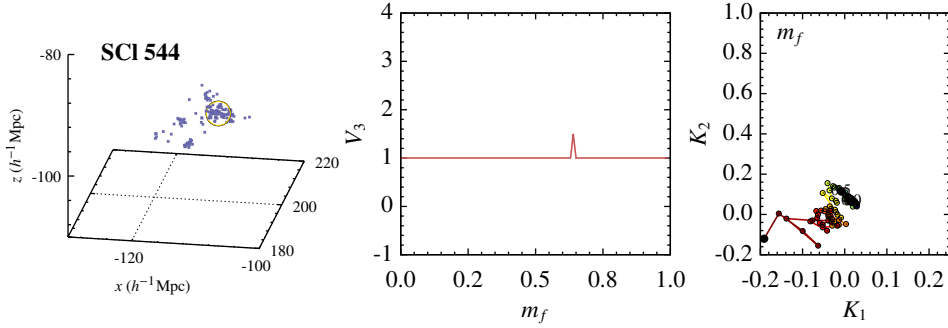
Figure 3.16: Supercluster SCl 544. Panels and symbols are the same as in Fig. 3.13

abundance (Park et al. 2012). Comparing the superclusters SCl 27, which forms the main body of the Sloan Great Wall, and SCl 16, using their physical properties and morphology, we cannot but notice strong similarities. Thus, adding morphological analysis allows us to better delineate certain types of objects, whose incidence the models should be able to predict. In general, we feel that objects with more unusual morphology or the most extreme combinations of shapefinders tend to reside at large distances or in low-density environments where the supercluster properties can be more seriously affected by selection effects and shot noise.

Table 3.1: Individual supercluster properties.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
	Name	ID	d	n_{gal}	n_{cl}	$L_{scl,wgal}$	$\oslash$	V	D	$V_{3,max}$	K_1, K_2
			h^{-1}Mpc			$10^{10}h^{-2}L_\odot$	h^{-1}Mpc	$(h^{-1}$Mpc$)^3$	$^\dagger \ell_{mean}$		
239+027+0091	1	263.9	1041	9	1809.6	52.9	8930	5.0	2.5	0.07,0.18
231+030+0117	3	335.8	1364	18	4249.7	75.3	26076	4.2	11.0	0.12,0.29
227+006+0078	7	233.4	1521	6	2069.1	52.9	14238	3.6	3.5	0.09,0.13
139+052+0190	16	547.8	244	29	5306.5	98.6	23689	5.4	18.0	0.17,0.50
*184+003+0077	19	230.5	1498	8	2064.3	56.5	10338	5.2	5.5	0.09,0.15
*202‑001+0084	27	255.7	2932	22	4694.7	94.5	22201	5.3	12.5	0.12,0.41
215+034+0193	45	556.7	36	7	959.5	32.1	3540	5.2	4.0	0.12,0.05
227+014+0186	64	540.2	188	18	3869.7	86.4	31231	2.9	13.0	0.17,0.43
163+011+0131	198	383.2	509	6	2327.4	59.3	20483	2.8	12.0	0.20,0.20
172+054+0071	211	206.5	1060	7	1235.9	53.6	4877	5.3	5.0	0.08,0.26
207+028+0077	220	225.7	1159	5	1454.0	58.6	12134	3.3	5.0	0.09,0.22
*210+006+0083	544	250.1	138	1	231.8	17.3	741	5.3	1.5	0.01,0.06

Columns in the Table: (1) name; (2) short ID; (3) distance to the highest density peak; (4) number of galaxies; (5) number of large clusters ($L_{gr} > 30 \cdot 10^{10}h^{-2}L_\odot$); (6) sum of weighted galaxy luminosities; (7) diameter; (8) volume; (9) adaptively assigned density threshold; (10) maximum clumpiness; (11) shapefinders.

Notes: †mean density value for the main sample is $\ell_{mean} = 1.526 \cdot 10^{-2} \frac{10^{10}h^{-2}}{(h^{-1}\text{Mpc})^3}L_\odot$ (Table 2.2); *superclusters belonging to the Sloan Great Wall.

78

3.3 Relations between supercluster properties

In this section we investigate relations between supercluster properties, both physical and morphological. The selection effects are a special issue to study, we pay close attention to any correlations between the supercluster properties and its distance.

First we will find the non-parametric Spearman's rank correlation coefficients between pairs of various supercluster properties. After that, we select a smaller number of attributes and study them simultaneously. For this, we have chosen to use principal component analysis, which is a widely used tool for multivariate data analysis of a large number of parameters. We follow the analysis presented in Einasto et al. (2011a), and as before, instead of a smaller sample we cover now the full fixed and adaptive threshold supercluster catalogues. We explore the underlying dimensional structure of supercluster properties. However, we are also interested in finding subsets and outliers in the supercluster catalogue, and finding possible scaling relations between supercluster properties.

3.3.1 Correlation analysis

Spearman's rank-order correlation is a non-parametric measure of the statistical relationship between two variables (Sheskin 2007). It evaluates the hypothesis whether the relationship between two variables can be described by a monotonic function. The correlation coefficient r_s is in the range between -1 and 1. The absolute value of the coefficient r_s shows the strength of the relation: $|r_s|$ being close to zero means weak or no correlation, and $|r_s| = \pm 1$ indicate cases of perfectly monotone correlation and anticorrelation. Spearman's rank correlation does not make any assumptions neither about the distribution of the data nor the relation between variables. Thus, it is well suited for our analysis as we are comparing supercluster characteristics of very different nature. We are looking for correlations between the following attributes: distance (d), total weighted galaxy luminosity (L_{wgal}), diameter ($\oslash$), number of galaxies (n_{gal}), volume (V), confidence estimate (C_{scl}), highest peak density (D_{peak}), clumpiness, i.e. the maximum value of the fourth Minkowski functional (V_3) and the shapefinders ($K_{1,2}$). In Tables 3.2 and 3.3 we present the coefficients r_s and the corresponding significance estimates p (the probability of the two sets being uncorrelated) between supercluster properties for the fixed and adaptive threshold catalogues. Traditionally, the correlation is considered statistically significant, if the p-value is less than 0.05. Due to the large sample sizes ($\sim$1000), the statistical significance of the coefficients is generally very high. We calculated the coefficients using the `scipy.stats` statistical package from `SciPy` library.

We will start by looking at the correlations between the supercluster distance and other properties. In the fixed threshold catalogue, the correlation between the dis-

Table 3.2: Spearman's rank correlation coefficients between supercluster attributes in the fixed threshold catalogue ($D = 5.0$).

	d	$L_{scl,wgal}$	$⌀$	n_{gal}	V	C_{scl}	D_{peak}	V_3	K_1	K_2
d		0.032 0.326	-0.129	-0.690	0.000 0.997	-0.945	0.093 0.004	0.092 0.004	0.046 0.156	0.052 0.102
$L_{scl,wgal}$	0.032 0.326		0.933	0.645	0.976	0.146	0.896	0.784	0.896	0.841
$⌀$	-0.129	0.933		0.733	0.926	0.297	0.772	0.832	0.935	0.887
n_{gal}	-0.690	0.645	0.733		0.658	0.804	0.546	0.474	0.570	0.525
V	0.000 0.997	0.976	0.926	0.658		0.196	0.927	0.757	0.887	0.823
C_{scl}	-0.945	0.146	0.297	0.804	0.196		0.112	0.046 0.154	0.120	0.094 0.003
D_{peak}	0.093 0.004	0.896	0.772	0.546	0.927	0.112		0.581	0.744	0.678
V_3	0.092 0.004	0.784	0.832	0.474	0.757	0.046 0.154	0.581		0.850	0.809
K_1	0.046 0.156	0.896	0.935	0.570	0.887	0.120	0.744	0.850		0.841
K_2	0.052 0.102	0.841	0.887	0.525	0.823	0.094 0.003	0.678	0.809	0.841	

Supercluster properties in the Table: d - distance to the highest density peak; $L_{scl,wgal}$ - total weighted galaxy luminosity; $⌀$ - diameter; n_{gal} - number of galaxies; V - volume; C_{scl} - confidence estimate; D_{peak} - highest luminosity density peak value; V_3 - maximum clumpiness; $K_{1,2}$ - shapefinders.
Notes: Each cell contains the Spearman's coefficient r_s and the p-value below it; p-values smaller than 0.001 are not shown.

tance and total weighted luminosity is weak and the scatter is large. The supercluster volume and the morphological parameters also show little correlation and have larger dispersions. Remaining attributes, however, are affected by selection effects. The supercluster diameter is slightly declining with distance, the number of galaxies in superclusters is decreasing strongly, as is the confidence estimate. Lastly, the peak density of superclusters is not correlated with distance, which again confirms the consistency of the luminosity weighting procedure. Now, the relations between the properties in the catalogue with adaptive thresholds are somewhat different. A weak but significant correlation between the supercluster distance and total luminosity has appeared, the same applies to the peak density and all morphological characteristics. At the same time, the correlation between distance and diameter has vanished. As in the fixed threshold catalogue, the object volumes seem to be independent of distance, and the numbers of galaxies and confidence estimates are strongly anti-

Table 3.3: Spearman's rank correlation coefficients between supercluster attributes in the adaptive threshold catalogue.

	d	$L_{scl,wgal}$	$\oslash$	D	n_{gal}	V	C_{scl}	D_{peak}	V_3	K_1	K_2
d		0.147	-0.015 0.609	0.001 0.960	-0.768	0.092 0.001	-0.939	0.170	0.266	0.188	0.290
$L_{scl,wgal}$	0.147		0.844	0.332	0.468	0.929	0.068 0.019	0.796	0.715	0.823	0.708
$\oslash$	-0.015 0.609	0.844		0.092 0.002	0.537	0.840	0.137	0.485	0.772	0.873	0.851
D	0.001 0.960	0.332	0.092 0.002		0.189	0.032 0.268	0.207	0.537	0.022 0.446	0.037 0.197	0.005 0.854
n_{gal}	-0.768	0.468	0.537	0.189		0.485	0.869	0.340	0.218	0.337	0.190
V	0.092 0.001	0.929	0.840	0.032 0.268	0.485		0.075 0.009	0.686	0.702	0.823	0.698
C_{scl}	-0.939	0.068 0.019	0.137	0.207	0.869	0.075 0.009		0.103	-0.184	-0.078 0.007	-0.204
D_{peak}	0.170	0.796	0.485	0.537	0.340	0.686	0.103		0.320	0.448	0.367
V_3	0.266	0.715	0.772	0.022 0.446	0.218	0.702	-0.184	0.320		0.831	0.789
K_1	0.188	0.823	0.873	0.037 0.197	0.337	0.823	-0.078 0.007	0.448	0.831		0.799
K_2	0.290	0.708	0.851	0.005 0.854	0.190	0.698	-0.204	0.367	0.789	0.799	

Supercluster properties in the Table: d - distance to the highest density peak; $L_{scl,wgal}$ - total weighted galaxy luminosity; $\oslash$ - diameter; D - adaptive density threshold; n_{gal} - number of galaxies; V - volume; C_{scl} - confidence estimate; D_{peak} - highest luminosity density peak value; V_3 - maximum clumpiness; $K_{1,2}$ - shapefinders.
Notes: Each cell contains the Spearman's coefficient r_s and the p-value below it; p-values smaller than 0.001 are not shown.

correlated. Also, adaptively assigned thresholds themselves are not correlated with distance. Many of these trends were also visible in plots in Sect. 3.1.

We can conclude that there are certain selection effects present in both types of catalogue. In the fixed threshold catalogue, for increasing distance, the supercluster luminosity is preserved (together with its morphology) while the diameter decreases. In the adaptive catalogue the diameter is independent of distance, however, the total luminosity now begins to increase. Only object volumes are largely independent of distance in both cases. Most of the reasons for these trends have already been discussed in Sect. 3.1. In addition to these, we also have the sample volume increasing with distance, which increases the likelihood of finding larger objects.

Regarding other supercluster properties, there are no important differences between catalogues. All the properties that characterise the size of the superclusters –

luminosity, diameter, volume and richness – are tightly correlated with each other and also with the morphological characteristics. It is worth noting that the filamentarity K_1 has slightly higher correlation coefficient values than the planarity K_2. The confidence estimate has the strongest correlation with the number of galaxies, which is natural as it is directly connected to the number density of galaxies in the sample. The adaptive density threshold is correlated with properties that naturally depend on the density of the large-scale environment – total luminosity, richness and peak density, but is independent of the size and morphology.

3.3.2 Principal component analysis

Principal component analysis (together with closely related factor analysis) is a statistical procedure mainly employed to eliminate redundancies in a large body of data (Sheskin 2007). PCA finds the linear combinations of correlated parameters that describe most of the variation in the dataset and replaces them with the same number of new, independent parameters. The data is converted into a new coordinate system, where the greatest variation by any projection of the data lies along the first coordinate (the first principal component), the second largest variance – along the second coordinate, and so on. Principal components PC_k ($k \in \mathbb{N}$) are linear combinations of the original parameters

$$\text{PC}_k = \sum_{i=1}^{N} a_{ik} V_i, \tag{3.12}$$

where $|a_{ik}| < 1$ are the coefficients of the linear transformation, V_i are the original parameters and N is the number of the parameters. The number of principal components is equal to the number of initial parameters, however, typically only the first few are needed to describe most of the total variation. For each component, the absolute values of the coefficients for different parameters characterise their relative importance (Sheskin 2007).

Linear combinations of principal components can be used to reproduce the characteristics of the objects, i.e. they can be used to derive scaling relations. For example, if data points lie on a plane defined by the first two principal components, then the scaling relations on this plane are defined using the third principal component (Efstathiou & Fall 1984). We use standardised parameters, centred on their means $(V_i - \overline{V_i})$ and normalised (divided by their standard deviations $\sigma(V_i)$). The scaling relations are then obtained as follows:

$$\sum_{i=1}^{N} a(i)_k \frac{(V_i - \overline{V_i})}{\sigma(V_i)} = 0|_{k=3}. \tag{3.13}$$
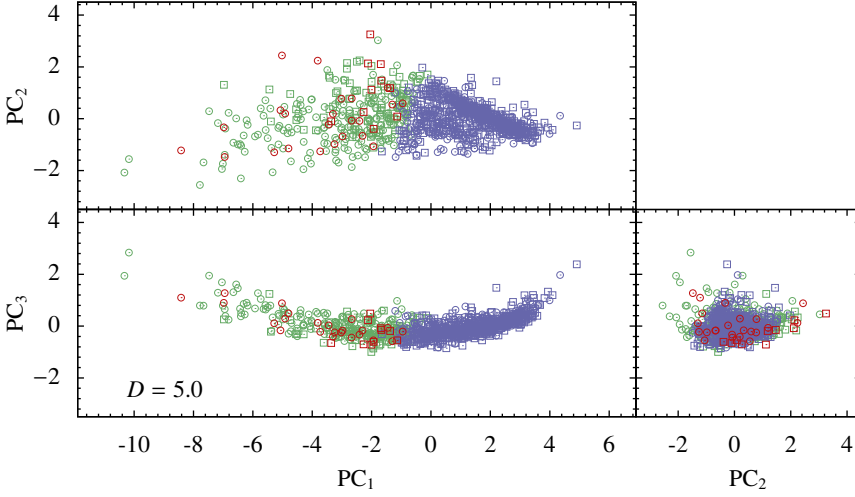
Figure 3.17: Principal components for the superclusters in the fixed threshold catalogue. Objects with the shapefinder ratio $K_1/K_2 < 0.6$ are marked with circles and objects with $K_1/K_2 \geqslant 0.6$ with squares. Red symbols are used to denote superclusters with a total weighted luminosity greater than $L_{\rm scl,wgal} > 400 \times 10^{10} L_{\odot}$, that are within the distance interval of $d = 90..320\ h^{-1}$Mpc, green symbols present superclusters with the same luminosities as previous but outside the limits, and blue symbols stand for objects at all distances and less luminous than $400 \times 10^{10} L_{\odot}$.

The parameters should be normally distributed to apply the PCA, therefore, for most cases we use logarithms of parameters as it makes the distributions more Gaussian, and the range of parameter values more reasonable (like in case of luminosity and size). Deviations from normal distributions are usually caused by, e.g., very large superclusters with extreme luminosities and sizes, or in case of morphology, the shapefinder ratio K_1/K_2 is noisy for small objects. The use of supercluster diameters in the adaptive catalogue may also require some allowances, because it has a sharp lower limit. However, Skibba & Macciò (2011) argue that cut distributions do not alter results significantly. It is still not possible to use adaptive density thresholds or the density peak values because their distribution is highly non-Gaussian (see Fig. 2.10).

Principal component analysis is a popular tool in astronomy, it has been used for, e.g., studying the properties of stars (Tiit & Einasto 1964), the spectral classification of galaxies (Sánchez Almeida et al. 2010), and multiple studies of galaxies, galaxy groups, dark matter haloes and relations between them (e.g., Efstathiou & Fall 1984; Lanzoni et al. 2004; Ferreras et al. 2006; Skibba & Macciò 2011; Toribio et al. 2011) and for various other uses.
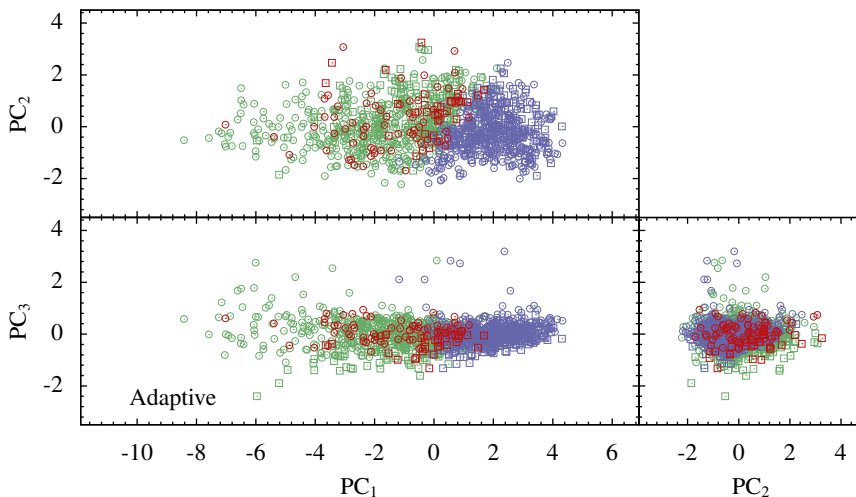
Figure 3.18: Principal components for the objects in the catalogue with adaptive density thresholds. Symbols are as described in Fig. 3.17.

We used `mlab` from the `matplotlib` package (Hunter 2007) for calculations, however, there are many other software libraries available for the principal component analysis.

3.3.3 PCA results and scaling relations

We used a number of physical and morphological parameters to conduct the principal component analysis. Results of the PCA of both catalogues are shown in Table 3.4. For better clarity we only present coefficients for the first three principal components. We see that the first component accounts for the large majority of variation in the data: 86 % for the fixed and 78 % for the adaptive level catalogue. Furthermore, the first three components account for about 96% and 93% of the total variation in fixed and adaptive catalogues, correspondingly. Looking at the first component, we see that formally, supercluster luminosity $L_{scl,wgal}$ has the highest and the peak density D_{peask} the lowest coefficient. However, the difference between the coefficients is actually not that significant. For the second component, the peak density is the most important property and for the third, the second shapefinder K_2, which characterises the filamentarity. There are some discrepancies between the fixed and adaptive threshold catalogues, but the coefficients of the first two principal components are largely similar, while the third principal component differs more substantially.

In Figures 3.17 and 3.18 we show superclusters from both catalogues in the coordinates of the first three principal components. As before in Sect 3.2, we divide
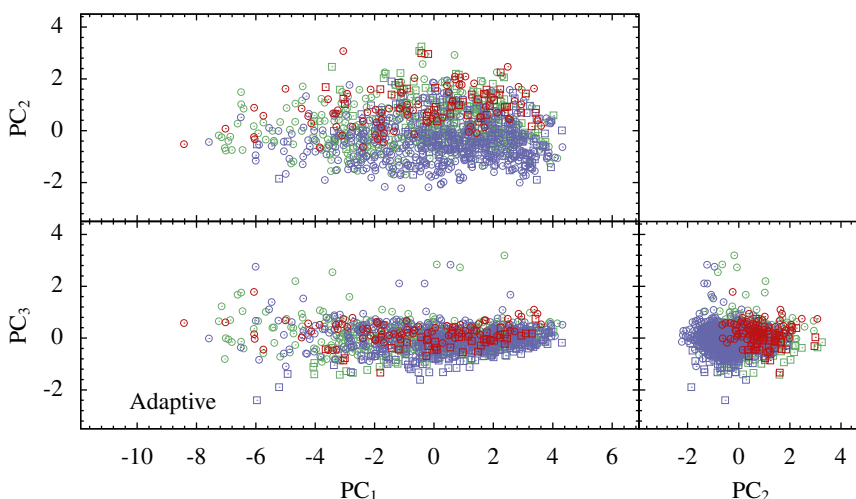
Figure 3.19: Principal components for the objects in different adaptive density level ranges. Red symbols denote superclusters with boundary densities $D > 4.5$, green symbols $4.5 \geqslant D > 3.0$, and blue symbols objects with adaptive densities $D \leqslant 3.0$. Squares designate less and circles more elongated structures.

both supercluster catalogues into three subsamples according to their sizes and distances. We can see that superclusters form an elongated cloud that is most stretched along the direction of PC_1. The extent of the distribution along the axes of PC_2 and PC_3 is much smaller and the overall shape of the distribution is almost plane-like. The plane is somewhat thinner for the fixed level sample, however, it is also bent in the PC_1–PC_2 plane. With red and green symbols we show luminous superclusters at different distances. We can see that they are similarily distributed. The dense bar-like structure among the low-luminosity objects in the top panel of Fig. 3.17 has no apparent direct connection to any specific supercluster property. In the case of the adaptive threshold catalogue, the distribution is more or less linear on all three principal axes. At the same time, the scatter is increased in the PC_3 direction when compared to the fixed density level superclusters. Excluding the outliers, the overall impression is still rather plane-like. The relation between the morphological parameters and PC_3 is most apparent in the PC_2–PC_3 panel in Fig. 3.18, where the more elongated objects have clearly higher and less elongated have lower values of PC_3. Figure 3.19 shows superclusters from the adaptive level catalogue in different boundary density intervals. We can see that there is some segregation of objects in the PC_1–PC_2 plane along the directions of PC_2 – superclusters in denser environments (red and green symbols) have also preferably higher values of the second principal component, while lower density objects (blue symbols) fill the plane more uniformly.

Einasto et al. (2011a) interpret the results of the PCA as to suggest that the first two principal components seem to define a kind of fundamental plane for superclusters and there is a possibility to derive the so-called scaling relations between supercluster properties. Famous examples of scaling relations for galaxies are the Tully-Fisher (Tully & Fisher 1977) and Faber-Jackson (Faber & Jackson 1976) relations, which tie galaxy luminosities to their rotational velocities (for spirals) or stellar velocity dispersions (for ellipticals), correspondingly. The latter is understood to be a projection of the fundamental plane for elliptical galaxies, which is a set of bivariate relations between several properties (also including size, mass, metallicity, colour, etc).

Following the example from Einasto et al. (2011a), we try to derive a relation between various supercluster parameters and the logarithm of its total weighted luminosity. As in Einasto et al. (2011a), we first carry out the principal component analysis using the following parameters – the logarithm of the supercluster luminosity: $\log L_{\mathrm{scl,wgal}}$, and combinations of the supercluster diameter and both shapefinders: $(1 - K_{1,2}) \log \oslash$ (where $(1 - K_{1,2})$ rather than $K_{1,2}$ were chosen to minimise numerical noise). This allows us to investigate the possibility of recovering supercluster luminosity utilising information on its size and shape. Again, we included full supercluster catalogues, both with fixed and adaptive thresholds, in the analysis. Results of the PCA are shown in Table 3.5.

Combining Eq. 3.13 with principal components from Table 3.5 we derive scaling relation formulas for supercluster luminosities. For the fixed density threshold

Table 3.4: Principal component analysis of selected supercluster properties.

	$D = 5.0$			Adaptive		
Property	PC_1	PC_2	PC_3	PC_1	PC_2	PC_3
$\log L_{\mathrm{scl,wgal}}$	-0.399	0.175	-0.091	-0.411	0.261	-0.098
$\log \oslash$	-0.387	-0.054	-0.597	-0.402	-0.152	0.187
$\log V$	-0.390	0.308	-0.268	-0.398	0.191	-0.250
$\log D_{\mathrm{peask}}$	-0.343	0.676	0.441	-0.292	0.757	0.196
$\log V_3$	-0.367	-0.454	-0.035	-0.367	-0.395	-0.276
K_1	-0.391	-0.224	0.028	-0.393	-0.203	-0.439
K_2	-0.367	-0.399	0.606	-0.369	-0.320	0.765
Standard deviations	2.446	0.737	0.423	2.332	0.927	0.516
Proportional variation	0.855	0.078	0.025	0.777	0.123	0.038
Cumulative variation	0.855	0.932	0.958	0.777	0.899	0.937

Supercluster properties in the Table: $L_{\mathrm{scl,wgal}}$ - total weighted galaxy luminosity; $\oslash$ - diameter; V - volume; D_{peak} - highest luminosity density peak value; V_3 - maximum clumpiness; $K_{1,2}$ - shapefinders.
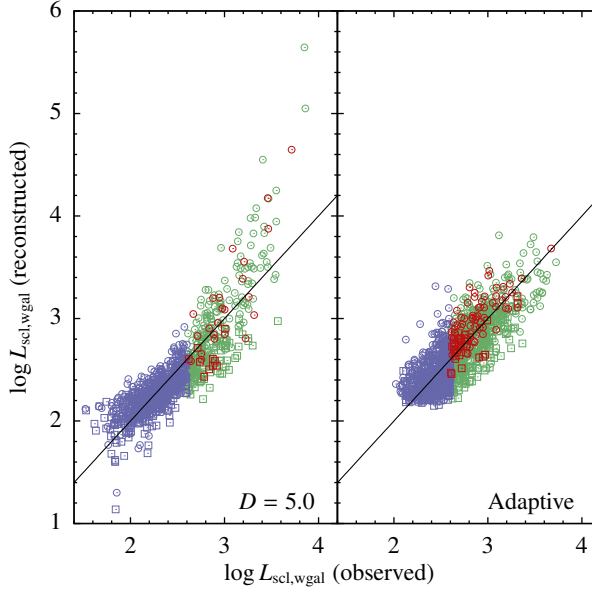
Figure 3.20: The relation between the observed and reconstructed total weighted galaxy luminosities for the superclusters in the fixed (*left panel*) and adaptive threshold (*right panel*) catalogues. Symbols are as explained in Fig. 3.17. The solid line denotes equality.

catalogue ($D = 5.0$) we obtain

$$\log L_{scl,wgal} = (2.486K_2 - 3.353K_1 + 0.867) \log \oslash + 1.386. \qquad (3.14)$$

Observed and recovered luminosities for the fixed level superclusters are shown on the left panel of Fig. 3.20. Comparing the results to those in Einasto et al. (2011a), which were obtained for a smaller sample, one can see that both the formula itself and the distribution of luminosities are similar. However, it is obvious that while the relation is relatively tight, it is also clearly not linear. A cursory glance at the figure seems to suggest that smaller (blue symbols) and larger (red and green symbols) objects might follow slightly different laws, as was also the case with morphological parameters. We can also observe that the luminosity of more elongated superclusters (circles) is generally overestimated and that of less elongated objects (squares) is underestimated.

The relation formula for the supercluster luminosities in the adaptive threshold catalogue is the following:

$$\log L_{scl,wgal} = (0.313K_2 - 2.695K_1 - 2.383) \log \oslash - 0.609, \qquad (3.15)$$

and the results are plotted in the right panel of Fig. 3.20. Here we can see increased scatter while the relation itself now appears linear. There are no discernible differ-

ences between rich and poor superclusters; however, we can see that more elongated superclusters are systematically given higher luminosities. In conclusion, the attempt to derive scaling relations between supercluster properties using regular PCA was not entirely succesful, and one might achieve better results with non-linear methods (e.g., machine learning).

In conclusion, the results of the PCA suggest that superclusters are relatively simple things. The first component is dominant and others have, at best, some marginal importance. (Skibba & Macciò 2011) used PCA to study properties of dark matter haloes in simulations and found them, in contrast, to be rather more complex objects, with 4 principal components having comparable significance. The PCA results obtained here are mostly in agreement with those in (Einasto et al. 2011a) – the biggest difference is that when using the full catalogue we do not see separate subsamples of superclusters according to their morphology.

Table 3.5: Principal component analysis for deriving a scaling relation.

	$D = 5.0$			Adaptive		
Property	PC_1	PC_2	PC_3	PC_1	PC_2	PC_3
$\log L_{scl,wgal}$	-0.561	0.744	-0.361	-0.672	0.245	-0.699
$(1 - K_1) \log \oslash$	-0.600	-0.066	0.797	-0.683	0.160	0.713
$(1 - K_2) \log \oslash$	-0.570	-0.665	-0.484	-0.288	-0.956	-0.060
Standard deviations	1.644	0.508	0.199	1.370	0.961	0.446
Proportional variation	0.901	0.086	0.013	0.626	0.308	0.066
Cumulative variation	0.901	0.987	1.000	0.626	0.934	1.000

Supercluster properties in the Table: $L_{scl,wgal}$ - total weighted galaxy luminosity; $\oslash$ - diameter; $K_{1,2}$ - shapefinders.

3.4 Spatial distribution of superclusters

For the last part of the supercluster catalogue overview we look at the spatial placement of superclusters in the universe.

3.4.1 Spatial density

Again, we are interested in looking for possible distance-dependent selection effects in the distribution of catalogue objects. Figure 3.21 shows the dependence of the spatial number density of superclusters on distance. Here we see a significant and interesting discrepancy between fixed and adaptive threshold catalogues. The number density of objects in the fixed level catalogue increases systematically with the
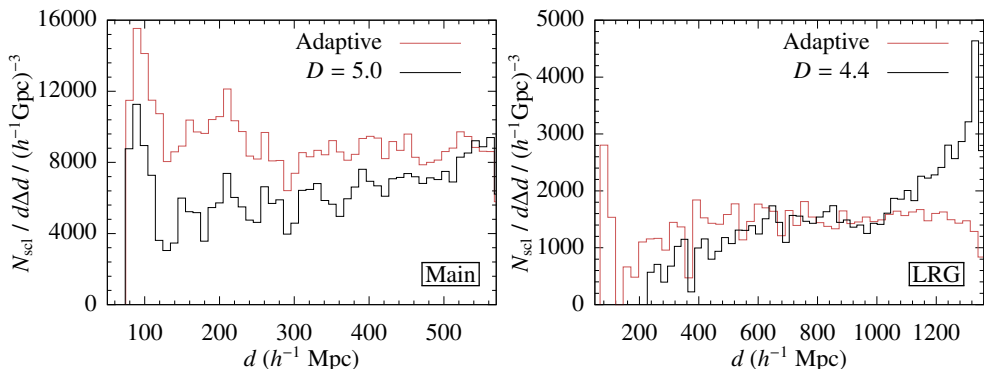
Figure 3.21: The dependence of supercluster number densities on distance for adaptive and fixed threshold objects in the main (*left*) and the LRG samples (*right*).

distance, while it stays more or less the same in the adaptive catalogue. Although the general behaviour of the distributions is common to both galaxy samples, a closer look also reveals some differences. In the case of the main sample, the adaptive catalogue contains more objects at all distances and the fixed threshold catalogue "catches up" only at the very far end. On the other hand, in the LRG sample the number density of fixed density level objects actually becomes larger at distances greater than $1000 \, h^{-1}\mathrm{Mpc}$.

The growth of the number of fixed level superclusters with the distance is caused by two processes. First, lower density objects are elevated above the threshold by luminosity weighting and, at the same time, the decreasing number density of galaxies leads to greater fragmentation of density field structures. The latter is especially apparent in the far regions of the LRG sample.

In Fig. 3.22 we take a look at the distance distribution of the adaptive density level objects in several density threshold intervals to see if there is a specific class of objects, which makes up the difference. We can see that this is not the case and that the spatial densities of superclusters behave similarly in all subsamples – at low, medium and high density levels. They are relatively constant with the distance (when compared to the fixed level case) and differ mostly in the number of objects. This is also true for the LRG sample.

Comparing both types of catalogues it seems that a fixed level catalogue is dominated by the distribution of density peak values while the adaptive catalogue depends more on the overall topology of the cosmic web. Both catalogues are clearly affected by the luminosity weighting procedure. However, regarding the distance distribution of the objects, the adaptive catalogue appears to have a more consistent spatial distribution.
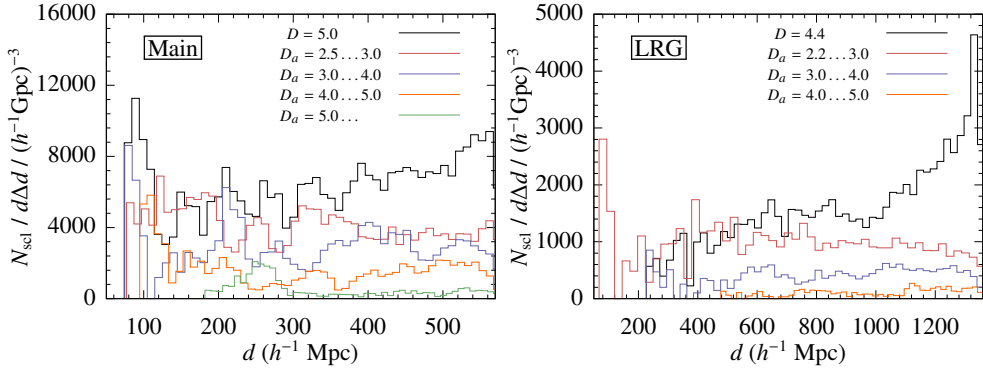
89

Figure 3.22: Comparison of distance distributions of the number density between adaptive and fixed level D_f supercluster catalogues when the adaptive catalogues are divided into subsamples according to threshold values D_a. The main sample is in the *left* and the LRG sample in the *right panel*.

3.4.2 Large-scale supercluster distribution

As with many three-dimensional structures, it is somewhat difficult to give a good overall visual impression of the spatial distribution of superclusters. Also, in case of observational samples, the whole volume has an irregular shape, which adds to the confusion. We give two views of the supercluster catalogue trying to convey some characteristic properties. To keep it shorter, we confine ourselves only to the adaptive density level sample.

First, we plot all the positions of superclusters in two projections (Fig. 3.23). The upper panel shows the "top down" and bottom panel the "face" projection of the SDSS sample volume. For each object, we show only the location of the highest density peak. As before, symbol colours show the adaptive density threshold values in three intervals, and symbol sizes are proportional to the total luminosities of superclusters. Even when using projections, we can see that the spatial distribution is markedly different for superclusters in different density intervals.

If we look at the positions and adaptive density levels of superclusters on the map of the SDSS main sample, we see that there are strong variations in the supercluster thresholds depending on the region where they are located. The threshold level needed to define a supercluster is tightly correlated with the large-scale mean density. The spatial scale of these variations is very large – more than $200 \ h^{-1}$Mpc. Substantial differences in the levels can be seen, e.g., in the *upper panel*, in the regions around $x \approx -60 \ h^{-1}$Mpc, $y \approx 300 \ h^{-1}$Mpc and at the Sloan Great Wall region at $x \approx 0$ and $y \approx 220 \ h^{-1}$Mpc. One can also discern the dominant supercluster plane (Einasto et al. 1997b) and a system of large voids behind it (described in detail by Platen 2009). Su-

90

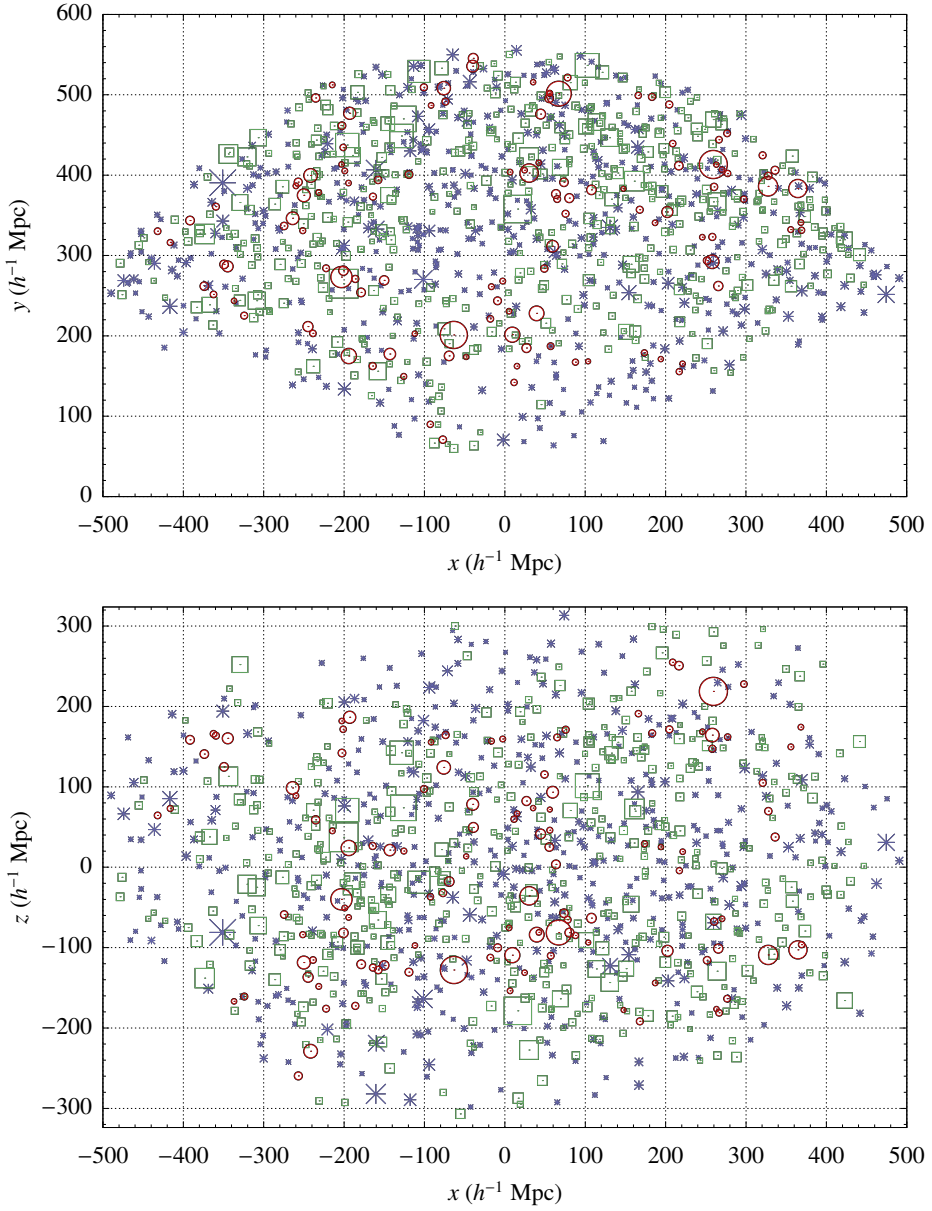**Figure 3.23**: The SDSS main sample supercluster map. Projection to the *xy*-plane in the *upper panel* and to the *xz*-plane in the *lower panel* Different symbol types show the density threshold levels used to delineate the superclusters; blue points: $D_a \leqslant 3.0$, green squares: $3.0 < D_a \leqslant 4.5$, and red rings: $D_a > 4.5$. Symbol sizes are proportional to the total luminosities of superclusters.

perclusters with highest adaptive densities (red circles) are usually in smaller more or less separated groups. Medium range superclusters (green squares) can be seen forming somewhat chainlike formations. Objects with the lowest limiting densities (blue symbols) have the most uniform distribution and are present in places like the aforementioned central large void region.

There are also some more subtle effects like the impact by the proximity of the survey borders. Objects with the highest boundary densities (red circles) tend to be further inside the sample volume and objects with the lowest densities (blue symbols) are closer to the borders. This is natural considering that the densities are somewhat underestimated at the borders. Also, the size of these objects is generally smaller being constrained by the sample borders. There are obvious exceptions to these tendencies, e.g., the Sloan Great Wall region, which contains the catalogues largest supercluster, is located at the bottom face of the survey volume (and is also partially outside). The reason for these variations is presently unclear, as are the influences of the selection and border effects, so we leave their quantification and study for the future.

Next we also look at the placement of large superclusters. Figure 3.24 shows the placement of 30 superclusters with the highest total weighted galaxy luminosities. Superclusters are depicted using their member galaxies while the inset shows the sky projection of supercluster positions. The colours again indicate the adaptive density thresholds. We restrict ourselves to only a limited number of objects to avoid cluttering the figure. Of course, there are not many conclusions one can draw by just looking at one figure, but we can still make a few observations.

We can see that, not unexpectedly, large superclusters keep some distance from each other. However, some of them still appear to be part of pairs or even small "groups". The general emptiness of the middle of the sample is also well visible here. Moreover, the entire distribution of large objects seems to be somewhat lop-sided. If we divide the whole sample in Fig. 3.24 into four sectors along the coordinate planes (see the inset) and count the number of objects in each sector, we can find rather large differences. For example, the top right part of the volume contains only 3 superclusters, but bottom right part – 11 superclusters. Similarily for the left side of the sample, there are 5 objects in the top sector, and 11 in the bottom one. Looking more closely, the distribution on the left actually looks slightly top-heavy instead, as 5 of the superclusters are located rather close to the coordinate plane. One might therefore tentatively suggest, that the volume of the SDSS main sample may not be fully homogeneous.

The homogeneity of the galaxy distribution in the SDSS main sample was studied by Pandey & Sarkar (2015) using Shannon entropy. They find that the galaxy distribution is homogeneous on scales larger than $140\,h^{-1}$Mpc. Also in the SDSS, the scale
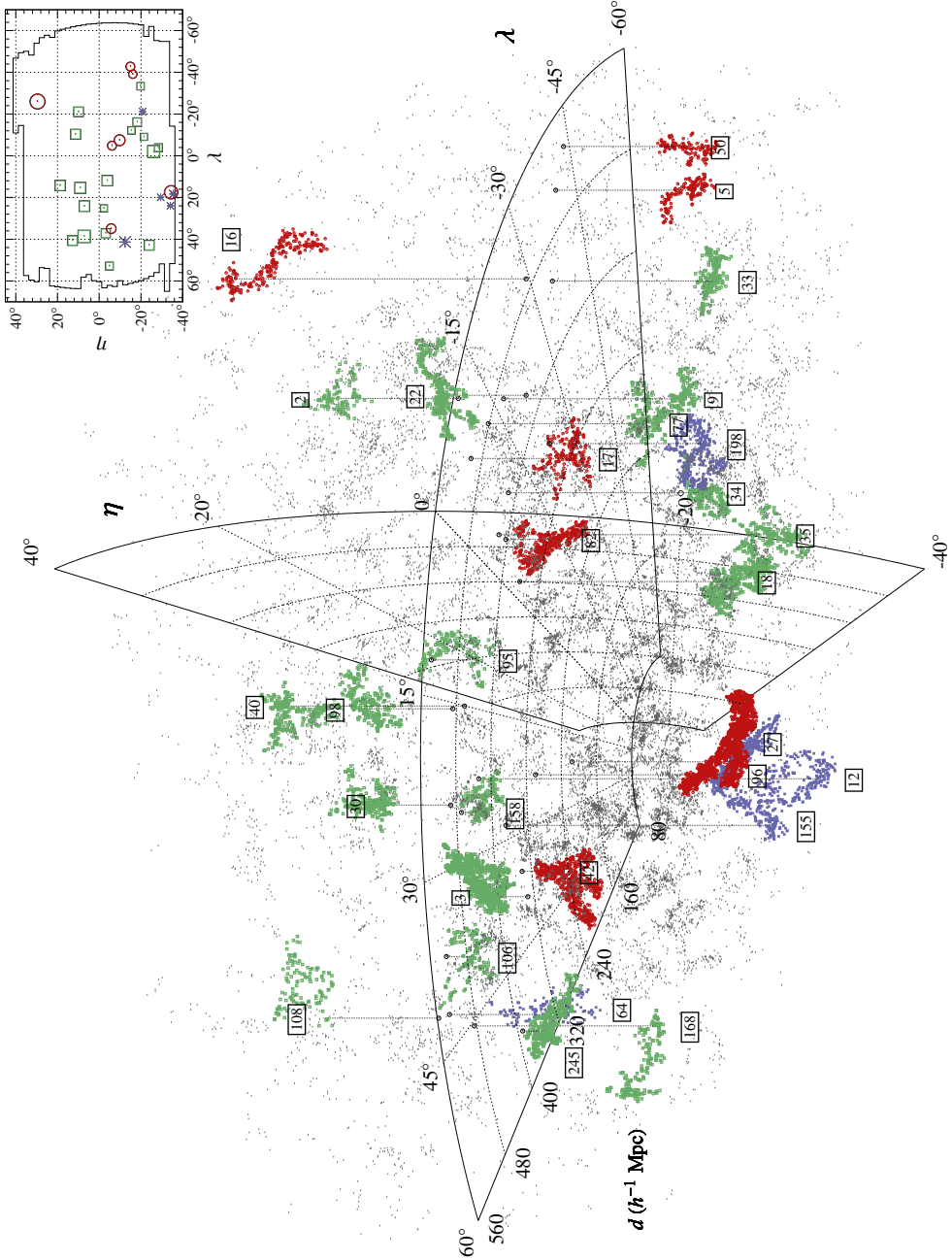
Figure 3.24: Placement of the 30 most luminous superclusters from the catalogue with adaptive thresholds. Blue points show superclusters with density leveles $D_a \leq 3.0$, green points: $3.0 < D_a \leq 4.5$, and red points: $D_a > 4.5$. Galaxies in other superclusters are shown with grey dots. Boxed labels show supercluster short ID numbers. Inset shows the sky projection of the supercluster positions in the SDSS η and λ coordinates.

where the galaxy distribution becomes isotropic was determined to be 150 h^{-1}Mpc by Marinoni et al. (2012). However, superclusters are structures built by us on top the galaxy distribution, using also the galaxy luminosities (and with all the additional selection effects), so their spatial distribution can be different. On the other hand, and more similar to our situation, there are multiple studies claiming to have found various very large-scale structures (several hundred megaparsecs) of sparse objects like quasars (Clowes et al. 2013) or events such as gamma ray bursts (Balázs et al. 2015) which, according to the respective authors, challenge the notion of the cosmological homogeneity. However, some of these results have been refuted by later analysis as not statistically significant (e.g., Nadathur 2013; Pilipenko & Malinovsky 2013). Without going into detailed analysis at present, we feel that this is also the case with the perceived inhomogeneity here, and the distribution of large superclusters is Poissonian. We will still look at the supercluster's spatial distribution in the future and try to find the most suitable statistical tools.

3.5   Concluding remarks and discussion

Developing catalogues of various objects is a standard procedure in astronomy, whether it is done to discover systematics, check the validity of models, or correlate with the properties of other objects. In this study we have created several catalogues of structures that, we think, correspond to what are generally considered galaxy superclusters. We also feel that, from the procedural point of view, this exercise has been a success.

Superclusters in our catalogues are fairly robust and well-established systems. Supercluster properties change radically with structure splits or mergers when changing the density level, but are relatively stable in between, not gaining nor losing significant amount of galaxies. We found that with the appropriate weighting of galaxy luminosities, it is possible to use flux-limited galaxy catalogues to create samples of superclusters which are relatively free of selection effects. Examining the supercluster properties revealed little dependence on the distance. To counter the inevitable arbitrariness when selecting the density thresholds we utilised a scheme to derive the boundary densities individually for each object. Adaptive thresholds for supeclusters permit to take the global density contrast into account and include more galaxies in superclusters, arguably thus giving a more complete perspective of the cosmic web. It also suppresses the selection effects, allowing us to generate practically volume-limited supercluster samples.

In addition, we also looked at the shape characteristics of superclusters and confirmed the visual imperssion that superclusters are primarily filamentary structures. The elongation grows with the size and richness, meaning that increasing the super-

cluster size, in essence, means lengthening it. Principal component analysis further supported this observation by showing that the relations between supercluster properties are not complicated, with the most important parameters being again the length and total luminosity.

We used two galaxy catalogues from the SDSS project with significantly different qualities. While the LRG sample is very sparse and the number density of superclusters in its volume is much lower than for the main sample, one can still construct a supercluster sample with relatively consistent properties. However, objects near the far edge of the LRG sample clearly suffer from the selection effects, due to the low density of galaxies. In the LRG sample, the vast majority of galaxies are enclosed in superclusters. This is natural since LRGs are bright galaxies presumably residing in the cores of large galaxy groups, which in turn are very likely to be situated in superclusters (Einasto et al. 2003b). We also compared the SDSS superclusters with the superclusters from the N-body simulation, which were built using the same principles, and the objects looked very similar.

We note that in addition to the usual variations in the galaxy density – from voids to clusters, etc; galaxy superclusters also exhibit secondary fluctuations at even larger scales. We can see clearly that large superclusters in the SDSS main sample volume are not distributed uniformly, and adaptively assigned density thresholds appear to have a semi-regular spatial distribution.

There are still several issues with the catalogue that need discussing. For example, the question of boundary effects. Using a fixed distance from the sample edge to limit the supercluster sample, as is sometimes done, is not entirely justified. First, it removes a large fraction of galaxies from the present samples. Second, many of the large superclusters (e.g., SCl 126 of the Sloan Great Wall) are touching the SDSS survey mask edges, but are nonetheless among the largest in the sample. In fact, most of the nearby superclusters are incomplete because of the cone-like shape of the survey. In the catalogue, objects in contact with the survey boundary are included but accordingly marked.

Then there are the redshift distortions in the galaxy sample. We take care of the cluster-finger effect by making the radial dispersion of galaxy velocities in clusters equal to the spatial dispersion on the sky plane. The assumption may not be necessarily accurate, however, the result is likely to be close enough to the actual galaxy distribution for our purposes. Kaiser effect, which occurs on the larger scale is more difficult to account for. Shi et al. (2016) showed that in the SDSS main sample the supercluster scale redshift distortions can boost clustering almost by half. This may not necessarily affect the results of delineating structures, if the effect is homogeneous enough at the scale of supercluster sizes, the shape of superclusters, however, is considerably altered. While we had to ignore the Kaiser effect in our current study,

techniques exist to estimate peculiar velocities of galaxies and subsequently correct for this (e.g., Kitaura et al. 2016, and the references within).

The the biggest ambiguity concerning superclusters lies probably in the definition of the objects themselves. For example, there is a principal difference between the two approaches to delineating superclusters in this study. First, using the fixed threshold corresponds to the most basic method for finding things – the existence of a phenomenon is simply determined by detecting an excess of something. As we described in the introduction (see Sect. 1.3), the first telltale signs of superclusters were the overabundance of galaxies in certain regions and second-order clustering. The situation has remained similar today, e.g., in cases where the data is sparse or incomplete for more detailed analysis, like the distant superclusters of X-ray clusters (Pompei et al. 2016) or the Vela supercluster behind the Zone of Avoidance (Kraan-Korteweg et al. 2016). The exact threshold could be justified by some theoretical considerations or, as well as be more or less intuitive. The alternative method, using individually assigned thresholds assumes that the galaxy distribution can be considered representative in the sample volume and can thus be *partitioned* into objects in a consistent manner. From this point of view, superclusters are like the constructor set pieces of the Universe's large-scale structure. We would contend that studies which use the spherical collapse condition, watersheds, or dynamical modelling of structures (see Sect. 1.3 for references) are essentially based on this notion.

Arguably, the most comprehensive way of partitioning the observed large-scale structure was shown in Tully et al. (2014) where the Laniakea supercluster was introduced. With the galaxy peculiar velocity field, one can divide the whole space into basins of attraction, which are defined by the divergence surfaces of the bulk flows. This is probably the most straightforwardly physical definition for superclusters. Using the same method (and the same data set), Pomarède et al. (2015) also delineated the small Arrowhead supercluster. We consider this important, since, while very large superclusters might be the biggest headline-grabbers, they still appear to form only a fraction of all the objects (e.g., Einasto et al. 2007a; Chon et al. 2014, and many other studies). As we have already mentioned, methods exist to estimate peculiar velocities of galaxies directly from the redshift samples (Kitaura & Enßlin 2008; Lavaux 2016). Galaxy velocities can also be obtained using the so-called constrained N-body simulations, which are tuned to give a result that resembles the observed large-scale structure. In principle, it should be therefore possible to define superclusters dynamically also in any of the more extended galaxy catalogues.

While being physically motivated, this dynamical definition for superclusters is also extremely general. For example, such superclusters would include the full range of environments – from the very low density voids to the massive galaxy clusters. From a sample of the dynamically delineated superclusters, one can probably put

constraints to some cosmological models or draw conclusions about the initial conditions, however, it's objects can be too fuzzy to, e.g., determine the relations between galaxy properties or galaxy group dynamics and the large-scale structure. Even the density-based division into superclusters and voids might be too vague. So, for many kinds of problems using a different, more detailed classification scheme can be more suitable.

A traditional complementary/alternative way for describing the large-scale structure is in terms of voids, walls, filaments and clusters. In this approach, different sets of web elements differ mainly in their dimensionality. Studies of numerical simulations, where the structures are much more resolved than those seen in the astronomical observations, mostly employ this classification. Indeed, whereas in observations the galaxy superclusters can be considered as visually dominating, they are somewhat lost in N-body simulations, where the whole matter distribution in its intricate glory can be seen. This is also a dynamical framework, and mathematically stringent prescriptions exist too, to classify space elements as belonging to one of the cosmic web components. Using either the gravitational tidal tensor (e.g., Aragón-Calvo et al. 2007; Hahn et al. 2007) or the velocity shear tensor (Hoffman et al. 2012), they are best suited to be applied to simulations, although still usable with the observational data (Sousbie et al. 2008; Pomarède et al. 2015). Most importantly, these methods have been formalised taking into account the structure formation theory, and are therefore also physically justified.

In Oort (1983), superclusters were described as "ill-defined", but "undoubtedly real". As it sometimes tends to happen in science, the accumulation of data over time can muddy the waters rather than clearing them. The most interesting question for us now is, what is the actual status of superclusters in the hierarchy of large-scale galaxy structures? Are the observational superclusters simply a superposition of walls, filaments and clusters in the underlying dark matter distribution that we are unable to see, or do they indicate some qualitative difference? For now that will remain to be investigated in future studies.

APPENDIX

A.1 Description of the catalogue

Here we give a detailed description of the supercluster catalogue. The catalogue is available in two forms. The first one consists of several tables with some redundancies between them. For each density level D there exists a table with all superclusters found at that threshold. These tables contain the following information (some less important properties are omitted here, but can be found in the documentation):

- unique identification number in the long and short forms;

- number of galaxies and groups (the latter for the main sample alone);

- supercluster volume as the number of the constituent grid cells times the cell volume (Eq. 3.2);

- supercluster luminosity as the sum of densities at grid vertices (Eq. 3.3);

- supercluster luminosity as the sum of the observed galaxy luminosities (Eq. 3.4);

- supercluster luminosity as the sum of the weighted galaxy luminosities (Eq. 3.5). For the main sample supercluster catalogue, we consider this as the best estimate of the total luminosity of the supercluster;

- maximum density in the supercluster;

- equatorial coordinates (J2000 here and hereafter) and the comoving distance of the highest density peak;

- equatorial coordinates and the comoving distance of the centre of mass (Eq. 3.1);

- cartesian coordinates (Eq. 2.4) of the highest peak and of the centre of mass;

- supercluster diameter as the maximum distance between the galaxies in the supercluster;

- identifier of the "marker" galaxy in the corresponding galaxy catalogue;

- equatorial coordinates and the redshift of the "marker" galaxy;

- confidence estimate for the supercluster found from the signal-to-noise field G (Eq. 2.22);

- indicator showing whether a supercluster is in contact with the mask boundary (1 – yes, 0 – no);

- number supercluster scale substructures (the number of objects that will split from the supercluster when increasing density threshold);

- maximum value of the 4th Minkowski functional V_3, i.e. the "clumpiness" of the supecluster;

- shapefinders K_1 – planarity and K_2 – filamentarity.

A similarly structured supercluster catalogue with adaptively assigned density thresholds has been compiled by combining the supercluster data in the tables described above. For each supercluster we take the data from the fixed level catalogue that corresponds to its defining density level and add the threshold value.

Additionally, we provide lists of galaxies and groups, together with the supercluster identifiers they are attributed to, for all density levels. We also present the supercluster splitting tree in the form of a table, where each supercluster is given the identifier of the object it belongs to at all given thresholds. Catalogue files can be downloaded from: `http://atmos.physic.ut.ee/~juhan/super/`, where the older SDSS DR7 main sample supercluster catalogue is also available.

The supercluster catalogue is also accessible via database interface at: `http://cosmodb.to.ee`, where it is fully integrated with the galaxy, cluster, and filament catalogues. Complete description of data is given in the documentation. Parts of the DR7 main and LRG catalogues have been uploaded to the CDS (Strasbourg astronomical Data Center).

A.2 Studies using the supercluster catalogue

In addition to Einasto et al. (2011b) and Einasto et al. (2011a) which form part of the basis of this work, several other studies have used this catalogue in one way or another. Following is a list of publications where versions of this catalogue were used.

Einasto et al. (2011c), Einasto et al. (2015), Gramann et al. (2015) and Einasto et al. (2016) are detailed studies of individual superclusters from the SDSS main galaxy sample.

Einasto et al. (2012) studied the relationship between the dynamical properties of galaxy clusters and their host superclusters.

Einasto et al. (2014a) analysed the galaxy content in superclusters from the main galaxy sample.

Planck Collaboration et al. (2011) used the LRG supercluster catalogue to confirm that a galaxy cluster discovered by the Planck satellite via the Sunyaev–Zeldovich effect is a member of a supercluster – the first time for this kind of observations.

Lietzen et al. (2011) examined the large-scale environment of active galaxies (AGN) in both the main and LRG samples.

Lietzen et al. (2012) studied the distribution of several types of galaxies in groups depending on their environment.

Roukema et al. (2015) used the LRG supercluster catalogue to investigate the hypothesis of variable spacetime metric as opposed to the standard spatially rigid FLRW metric. They find a weak signal for a shift of the characteristic BAO (baryon acoustic oscillations) peak scale in the correlation function depending on the large-scale environment.

Lietzen et al. (2016) applied the algorithm to the BOSS sample from the SDSS-III catalogue (Alam et al. 2015) to create a supercluster sample and found a massive galaxy structure similar to the Sloan Great Wall.

REFERENCES

Abazajian, K. N., Adelman-McCarthy, J. K., Agüeros, M. A., et al. 2009, *The Seventh Data Release of the Sloan Digital Sky Survey*, ApJS, 182, 543

Abell, G. O. 1958, *The Distribution of Rich Clusters of Galaxies.*, ApJS, 3, 211

Abell, G. O. 1961, *Evidence regarding second-order clustering of galaxies and interactions between clusters of galaxies*, AJ, 66, 607

Abell, G. O. 1965, *Clustering of Galaxies*, ARA&A, 3, 1

Aihara, H., Allende Prieto, C., An, D., et al. 2011a, *The Eighth Data Release of the Sloan Digital Sky Survey: First Data from SDSS-III*, ApJS, 193, 29

Aihara, H., Allende Prieto, C., An, D., et al. 2011b, *Erratum: "The Eighth Data Release of the Sloan Digital Sky Survey: First Data from SDSS-III" (2011, ApJS, 193, 29)*, ApJS, 195, 26

Alam, S., Albareti, F. D., Allende Prieto, C., et al. 2015, *The Eleventh and Twelfth Data Releases of the Sloan Digital Sky Survey: Final Data from SDSS-III*, ApJS, 219, 12

Aragón-Calvo, M. A., Jones, B. J. T., van de Weygaert, R., & van der Hulst, J. M. 2007, *The multiscale morphology filter: identifying and extracting spatial patterns in the galaxy distribution*, A&A, 474, 315

Aragón-Calvo, M. A., van de Weygaert, R., & Jones, B. J. T. 2010, *Multiscale phenomenology of the cosmic web*, MNRAS, 408, 2163

Araya-Melo, P. A., Reisenegger, A., Meza, A., et al. 2009, *Future evolution of bound superclusters in an accelerating Universe*, MNRAS, 399, 97

Ata, M., Kitaura, F.-S., Chuang, C.-H., et al. 2016, *The Clustering of Galaxies in the Completed SDSS-III Baryon Oscillation Spectroscopic Survey: Cosmic Flows and Cosmic Web from Luminous Red Galaxies*, ArXiv e-prints

Balázs, L. G., Bagoly, Z., Hakkila, J. E., et al. 2015, *A giant ring-like structure at $0.78 < z < 0.86$ displayed by GRBs*, MNRAS, 452, 2236

Basilakos, S. 2003, *Shape statistics of Sloan Digital Sky Survey superclusters*, MNRAS, 344, 602

Basilakos, S., Plionis, M., & Rowan-Robinson, M. 2001, *PSCz superclusters: detection, shapes and cosmological implications*, MNRAS, 323, 47

Benitez, N., Dupke, R., Moles, M., et al. 2014, *J-PAS: The Javalambre-Physics of the Accelerated Universe Astrophysical Survey*, ArXiv e-prints

Blanton, M. R., Hogg, D. W., Bahcall, N. A., et al. 2003, *The Galaxy Luminosity Function and Luminosity Density at Redshift z = 0.1*, ApJ, 592, 819

Blanton, M. R. & Roweis, S. 2007, *K-Corrections and Filter Transformations in the Ultraviolet, Optical, and Near-Infrared*, AJ, 133, 734

Böhringer, H., Schuecker, P., Guzzo, L., et al. 2004, *The ROSAT-ESO Flux Limited X-ray (REFLEX) Galaxy cluster survey. V. The cluster catalogue*, A&A, 425, 367

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

Bower, R. G., Benson, A. J., Malbon, R., et al. 2006, *Breaking the hierarchy of galaxy formation*, MNRAS, 370, 645

Branchini, E., Teodoro, L., Frenk, C. S., et al. 1999, *A non-parametric model for the cosmic velocity field*, MNRAS, 308, 1

Cai, Y.-C., Li, B., Cole, S., Frenk, C. S., & Neyrinck, M. 2013, *The Integrated Sachs-Wolfe effect in $f(R)$ gravity*, ArXiv e-prints

Cautun, M., van de Weygaert, R., Jones, B. J. T., & Frenk, C. S. 2014, *Evolution of the cosmic web*, MNRAS, 441, 2923

Chincarini, G. & Rood, H. J. 1979, *Space distribution of SC galaxies - Clues to the large-scale structure*, ApJ, 230, 648

Chincarini, G. L., Giovanelli, R., & Haynes, M. P. 1983, *21 centimeter observations of supercluster galaxies - The bridge between Coma and A1367*, ApJ, 269, 13

Chon, G., Böhringer, H., Collins, C. A., & Krause, M. 2014, *Characterising superclusters with the galaxy cluster distribution*, A&A, 567, A144

Clowes, R. G., Harris, K. A., Raghunathan, S., et al. 2013, *A structure in the early Universe at z   1.3 that exceeds the homogeneity scale of the R-W concordance cosmology*, MNRAS, 429, 2910

Cohn, J. D. 2006, *Power spectrum and correlation function errors: Poisson vs. Gaussian shot noise*, New Astronomy, 11, 226

Colless, M., Peterson, B. A., Jackson, C., et al. 2003, *The 2dF Galaxy Redshift Survey: Final Data Release*, ArXiv Astrophysics e-prints

Costa-Duarte, M. V., Sodré, Jr., L., & Durret, F. 2011, *Morphological properties of superclusters of galaxies*, MNRAS, 411, 1716

Davis, M., Efstathiou, G., Frenk, C. S., & White, S. D. M. 1985, *The evolution of large-scale structure in a universe dominated by cold dark matter*, ApJ, 292, 371

Davis, M., Huchra, J., Latham, D. W., & Tonry, J. 1982, *A survey of galaxy redshifts. II - The large scale space distribution*, ApJ, 253, 423

Davison, A. C. & Hinkley, D. V. 1997, Bootstrap Methods and Their Application (Cambridge University Press, Cambridge, UK)

de Lapparent, V., Geller, M. J., & Huchra, J. P. 1986, *A slice of the universe*, ApJ, 302, L1

de Vaucouleurs, G. 1953, *Evidence for a local super,galaxy*, AJ, 58, 30

Djorgovski, S. & Davis, M. 1987, *Fundamental properties of elliptical galaxies*, ApJ, 313, 59

Doroshkevich, A. G., Kotok, E. V., Poliudov, A. N., et al. 1980, *Two-dimensional simulation of the gravitational system dynamics and formation of the large-scale structure of the universe*, MNRAS, 192, 321

Dünner, R., Araya, P. A., Meza, A., & Reisenegger, A. 2006, *The limits of bound structures in the accelerating Universe*, MNRAS, 366, 803

Efron, B. & Tibshirani, R. J. 1993, An introduction to the bootstrap (Chapman and Hall, London, UK)

Efstathiou, G. & Fall, S. M. 1984, *Multivariate analysis of elliptical galaxies*, MNRAS, 206, 453

Einasto, J., Einasto, M., Gottlöber, S., et al. 1997a, *A 120-Mpc periodicity in the three-dimensional distribution of galaxy superclusters*, Nature, 385, 139

Einasto, J., Einasto, M., Hütsi, G., et al. 2003a, *Clusters and superclusters in the Las Campanas redshift survey*, A&A, 410, 425

Einasto, J., Einasto, M., Saar, E., et al. 2006, *Luminous superclusters: remnants from inflation?*, A&A, 459, L1

Einasto, J., Einasto, M., Saar, E., et al. 2007a, *Superclusters of galaxies from the 2dF redshift survey. II. Comparison with simulations*, A&A, 462, 397

Einasto, J., Einasto, M., Tago, E., et al. 2007b, *Superclusters of galaxies from the 2dF redshift survey. I. The catalogue*, A&A, 462, 811

Einasto, J., Hütsi, G., Einasto, M., et al. 2003b, *Clusters and superclusters in the Sloan Digital Sky Survey*, A&A, 405, 425

Einasto, J., Joeveer, M., & Saar, E. 1980, *Structure of superclusters and supercluster formation*, MNRAS, 193, 353

Einasto, M., Einasto, J., Tago, E., Dalton, G. B., & Andernach, H. 1994, *The Structure of the Universe Traced by Rich Clusters of Galaxies*, MNRAS, 269, 301

Einasto, M., Einasto, J., Tago, E., Müller, V., & Andernach, H. 2001, *Optical and X-Ray Clusters as Tracers of the Supercluster-Void Network. I. Superclusters of Abell and X-Ray Clusters*, AJ, 122, 2222

Einasto, M., Gramann, M., Saar, E., et al. 2015, *Unusual A2142 supercluster with a collapsing core: distribution of light and mass*, ArXiv e-prints

Einasto, M., Lietzen, H., Gramann, M., et al. 2016, *Sloan Great Wall as a complex of superclusters with collapsing cores*, A&A, 595, A70

Einasto, M., Lietzen, H., Tempel, E., et al. 2014a, *SDSS superclusters: morphology and galaxy content*, A&A, 562, A87

Einasto, M., Liivamägi, L. J., Saar, E., et al. 2011a, *SDSS DR7 superclusters. Principal component analysis*, A&A, 535, A36

Einasto, M., Liivamägi, L. J., Tago, E., et al. 2011b, *SDSS DR7 superclusters. Morphology*, A&A, 532, A5

Einasto, M., Liivamägi, L. J., Tempel, E., et al. 2011c, *The Sloan Great Wall. Morphology and Galaxy Content*, ApJ, 736, 51

Einasto, M., Liivamägi, L. J., Tempel, E., et al. 2012, *Multimodality of rich clusters from the SDSS DR8 within the supercluster-void network*, A&A, 542, A36

Einasto, M., Saar, E., Liivamägi, L. J., et al. 2007c, *The richest superclusters. I. Morphology*, A&A, 476, 697

Einasto, M., Tago, E., Jaaniste, J., Einasto, J., & Andernach, H. 1997b, *The supercluster-void network I. The supercluster catalogue and large-scale distribution*, A&AS, 123, 119

Einasto, M., Tago, E., Lietzen, H., et al. 2014b, *Tracing a high redshift cosmic web with quasar systems*, A&A, 568, A46

Eisenstein, D. J., Annis, J., Gunn, J. E., et al. 2001, *Spectroscopic Target Selection for the Sloan Digital Sky Survey: The Luminous Red Galaxy Sample*, AJ, 122, 2267

Erdoğdu, P., Lahav, O., Zaroubi, S., et al. 2004, *The 2dF Galaxy Redshift Survey: Wiener reconstruction of the cosmic web*, MNRAS, 352, 939

Faber, S. M. & Jackson, R. E. 1976, *Velocity dispersions and mass-to-light ratios for elliptical galaxies*, ApJ, 204, 668

Ferreras, I., Pasquali, A., de Carvalho, R. R., de la Rosa, I. G., & Lahav, O. 2006, *A principal component analysis approach to the star formation history of elliptical galaxies in compact groups*, MNRAS, 370, 828

Fisher, J. R. & Tully, R. B. 1981, *Neutral hydrogen observations of a large sample of galaxies*, ApJS, 47, 139

Fraley, C. & Raftery, A. E. 2006, *MCLUST version 3: an R package for normal mixture modeling and model-based clustering*, Tech. rep., WASHINGTON UNIV SEATTLE DEPT OF STATISTICS

Giovanelli, R. & Haynes, M. P. 1991, *Redshift surveys of galaxies*, ARA&A, 29, 499

Giovanelli, R., Haynes, M. P., & Chincarini, G. L. 1986, *Morphological segregation in the Pisces-Perseus supercluster*, ApJ, 300, 77

Girardi, M., Rigoni, E., Mardirossian, F., & Mezzetti, M. 2003, *Morphology and luminosity segregation of galaxies in nearby loose groups*, A&A, 406, 403

Gramann, M., Einasto, M., Heinämäki, P., et al. 2015, *Characteristic density contrasts in the evolution of superclusters. The case of A2142 supercluster*, A&A, 581, A135

Granett, B. R., Branchini, E., Guzzo, L., et al. 2015, *The VIMOS Public Extragalactic Redshift Survey. Reconstruction of the redshift-space galaxy density field*, A&A, 583, A61

Gregory, S. A. & Thompson, L. A. 1978, *The Coma/A1367 supercluster and its environs*, ApJ, 222, 784

Hahn, O., Carollo, C. M., Porciani, C., & Dekel, A. 2007, *The evolution of dark matter halo properties in clusters, filaments, sheets and voids*, MNRAS, 381, 41

Hamilton, A. J. S. 1988, *Evidence for biasing in the CfA survey*, ApJ, 331, L59

Hamilton, A. J. S. 1993, *Toward Better Ways to Measure the Galaxy Correlation Function*, ApJ, 417, 19

Harrison, E. R. 1970, *Fluctuations at the Threshold of Classical Cosmology*, Phys. Rev. D, 1, 2726

Harrison, E. R. 1974, *Interpretation of Redshifts of Galaxies in Clusters*, ApJ, 191, L51

Hoffman, Y., Metuki, O., Yepes, G., et al. 2012, *A kinematic classification of the cosmic web*, MNRAS, 425, 2049

Hubble, E. & Humason, M. L. 1931, *The Velocity-Distance Relation among Extra-Galactic Nebulae*, ApJ, 74, 43

Huchra, J. P., Macri, L. M., Masters, K. L., et al. 2012, *The 2MASS Redshift Survey – Description and Data Release*, ApJS, 199, 26

Hunter, J. D. 2007, *Matplotlib: A 2D graphics environment*, Computing In Science & Engineering, 9, 90

Illian, J., Penttinen, A., Stoyan, H., & Stoyan, D. 2008, Statistical Analysis and Modelling of Spatial Point Patterns (John Wiley & Sons, Chischester, UK)

Jõeveer, M., Einasto, J., & Tago, E. 1978, *Spatial distribution of galaxies and of clusters of galaxies in the southern galactic hemisphere*, MNRAS, 185, 357

Jasche, J., Kitaura, F. S., Li, C., & Enßlin, T. A. 2010, *Bayesian non-linear large-scale structure inference of the Sloan Digital Sky Survey Data Release 7*, MNRAS, 409, 355

Kaiser, N. 1987, *Clustering in real space and in redshift space*, MNRAS, 227, 1

Kirshner, R. P., Oemler, Jr., A., Schechter, P. L., & Shectman, S. A. 1981, *A million cubic megaparsec void in Bootes*, ApJ, 248, L57

Kitaura, F.-S., Ata, M., Angulo, R. E., et al. 2016, *Bayesian redshift-space distortions correction from galaxy redshift surveys*, MNRAS, 457, L113

Kitaura, F. S. & Enßlin, T. A. 2008, *Bayesian reconstruction of the cosmological large-scale structure: methodology, inverse algorithms and numerical optimization*, MNRAS, 389, 497

Kitaura, F.-S., Erdoğdu, P., Nuza, S. E., et al. 2012, *Cosmic structure and dynamics of the local Universe*, MNRAS, 427, L35

Kitaura, F.-S., Jasche, J., & Metcalf, R. B. 2010, *Recovering the non-linear density field from the galaxy distribution with a Poisson-lognormal filter*, MNRAS, 403, 589

Klypin, A. A. & Shandarin, S. F. 1983, *Three-dimensional numerical model of the formation of large-scale structure in the Universe*, MNRAS, 204, 891

Kolokotronis, V., Basilakos, S., & Plionis, M. 2002, *Supercluster properties as a cosmological probe*, MNRAS, 331, 1020

Koulouridis, E., Poggianti, B., Altieri, B., et al. 2016, *The XXL Survey. XII. Optical spectroscopy of X-ray-selected clusters and the frequency of AGN in superclusters*, A&A, 592, A11

Kraan-Korteweg, R. C., Cluver, M. E., Bilicki, M., et al. 2016, *Discovery of a supercluster in the ZOA in Vela*, ArXiv e-prints

Lanzoni, B., Ciotti, L., Cappi, A., Tormen, G., & Zamorani, G. 2004, *The Scaling Relations of Galaxy Clusters and Their Dark Matter Halos*, ApJ, 600, 640

Lavaux, G. 2016, *Bayesian 3D velocity field reconstruction with VIRBIUS*, MNRAS, 457, 172

Lietzen, H., Heinämäki, P., Nurmi, P., et al. 2011, *Large-scale environments of z < 0.4 active galaxies*, A&A, 535, A21

Lietzen, H., Heinämäki, P., Nurmi, P., et al. 2009, *Environments of nearby quasars in Sloan Digital Sky Survey*, A&A, 501, 145

Lietzen, H., Tempel, E., Heinämäki, P., et al. 2012, *Environments of galaxies in groups within the supercluster-void network*, A&A, 545, A104

Lietzen, H., Tempel, E., Liivamägi, L. J., et al. 2016, *Discovery of a massive supercluster system at z ˜ 0.47*, A&A, 588, L4

Liivamägi, L. J., Tempel, E., & Saar, E. 2012, *SDSS DR7 superclusters. The catalogues*, A&A, 539, A80

Luparello, H., Lares, M., Lambas, D. G., & Padilla, N. 2011, *Future virialized structures: an analysis of superstructures in the SDSS-DR7*, MNRAS, 862

Lynden-Bell, D., Faber, S. M., Burstein, D., et al. 1988, *Spectroscopy and photometry of elliptical galaxies. V - Galaxy streaming toward the new supergalactic center*, ApJ, 326, 19

Marinoni, C., Bel, J., & Buzzi, A. 2012, *The scale of cosmic isotropy*, JCAP, 10, 036

Markevitch, M., Ponman, T. J., Nulsen, P. E. J., et al. 2000, *Chandra Observation of Abell 2142: Survival of Dense Subcluster Cores in a Merger*, ApJ, 541, 542

Martinez, V. J. 1999, *COSMOLOGY:Enhanced: Is the Universe Fractal?*, Science, 284, 445

Martínez, V. J., Arnalte-Mur, P., Saar, E., et al. 2009, *Reliability of the Detection of the Baryon Acoustic Peak*, ApJ, 696, L93

Martínez, V. J. & Saar, E. 2002, Statistics of the Galaxy Distribution (Chapman & Hall/CRC, Boca Raton, USA)

Mo, H., van den Bosch, F. C., & White, S. 2010, Galaxy Formation and Evolution (Cambridge University Press, Cambridge, UK)

Nadathur, S. 2013, *Seeing patterns in noise: gigaparsec-scale 'structures' that do not violate homogeneity*, MNRAS, 434, 398

Nadathur, S. & Crittenden, R. 2016, *A detection of the integrated Sachs-Wolfe imprint of cosmic superstructures using a matched-filter approach*, ArXiv e-prints

Nadathur, S. & Hotchkiss, S. 2014, *A robust public catalogue of voids and superclusters in the SDSS Data Release 7 galaxy surveys*, MNRAS, 440, 1248

Neyrinck, M. C. 2008, *ZOBOV: a parameter-free void-finding algorithm*, MNRAS, 386, 2101

Oort, J. H. 1983, *Superclusters*, ARA&A, 21, 373

Öpik, E. 1922, *An estimate of the distance of the Andromeda Nebula.*, ApJ, 55

Pandey, B. & Sarkar, S. 2015, *Testing homogeneity in the Sloan Digital Sky Survey Data Release Twelve with Shannon entropy*, MNRAS, 454, 2647

Park, C., Choi, Y.-Y., Kim, J., et al. 2012, *The Challenge of the Largest Structures in the Universe to Cosmology*, ApJ, 759, L7

Peacock, J. A. 1999, Cosmological Physics (Cambridge University Press, UK)

Peebles, P. J. E. 1971, Physical cosmology (Princeton University Press, Princeton, U.S.)

Peebles, P. J. E. 1980, The large-scale structure of the universe (Princeton University Press, Princeton, U.S.)

Pilipenko, S. & Malinovsky, A. 2013, *Testing statistical significance of large quasar groups with sheets model of large scale structure*, ArXiv e-prints

Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2015, *Planck 2015 results. XIII. Cosmological parameters*, ArXiv e-prints

Planck Collaboration, Aghanim, N., Arnaud, M., et al. 2011, *Planck early results. IX. XMM-Newton follow-up for validation of Planck cluster candidates*, A&A, 536, A9

Platen, E. 2009, *A Void Perspective of the Cosmic Web*, PhD thesis, University of Groningen

Platen, E., van de Weygaert, R., Jones, B. J. T., Vegter, G., & Calvo, M. A. A. 2011, *Structural analysis of the SDSS Cosmic Web - I. Non-linear density field reconstructions*, MNRAS, 416, 2494

Pomarède, D., Tully, R. B., Hoffman, Y., & Courtois, H. M. 2015, *The Arrowhead Mini-supercluster of Galaxies*, ApJ, 812, 17

Pompei, E., Adami, C., Eckert, D., et al. 2016, *The XXL Survey. VII. A supercluster of galaxies at z = 0.43*, A&A, 592, A6

Press, W. H. & Davis, M. 1982, *How to identify and weigh virialized clusters of galaxies in a complete redshift catalog*, ApJ, 259, 449

Raychaudhury, S. 1989, *The distribution of galaxies in the direction of the 'Great Attractor'*, Nature, 342, 251

Roukema, B. F., Buchert, T., Ostrowski, J. J., & France, M. J. 2015, *Evidence for an environment-dependent shift in the baryon acoustic oscillation peak*, MNRAS, 448, 1660

Saar, E. 2009, *Multiscale Methods*, in Lecture Notes in Physics, Berlin Springer Verlag, Vol. 665, Data Analysis in Cosmology, ed. V. J. Martínez, E. Saar, E. Martínez-González, & M.-J. Pons-Bordería, 523

Saar, E., Martínez, V. J., Starck, J.-L., & Donoho, D. L. 2007, *Multiscale morphology of the galaxy distribution*, MNRAS, 374, 1030

Sahni, V., Sathyaprakash, B. S., & Shandarin, S. F. 1998, *Shapefinders: A New Shape Diagnostic for Large-Scale Structure*, ApJ, 495, L5

Sánchez Almeida, J., Aguerri, J. A. L., Muñoz-Tuñón, C., & de Vicente, A. 2010, *Automatic Unsupervised Classification of All Sloan Digital Sky Survey Data Release 7 Galaxy Spectra*, ApJ, 714, 487

Saunders, W., Sutherland, W. J., Maddox, S. J., et al. 2000, *The PSCz catalogue*, MNRAS, 317, 55

Schaap, W. E. 2007, *DTFE: the Delaunay Tessellation Field Estimator*, PhD thesis, Kapteyn Astronomical Institute

Schmoldt, I. M., Saar, V., Saha, P., et al. 1999, *On Density and Velocity Fields and beta from the IRAS PSCZ Survey*, AJ, 118, 1146

Shandarin, S. F., Sheth, J. V., & Sahni, V. 2004, *Morphology of the supercluster-void network in ΛCDM cosmology*, MNRAS, 353, 162

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

Shao, J. & Tu, D. 1995, The Jackknife and Bootstrap (Springer-Verlag GmbH, Heidelberg, Germany)

Shapley, H. 1930, *Note on a Remote Cloud of Galaxies in Centaurus*, Harvard College Observatory Bulletin, 874, 9

Shapley, H. & Ames, A. 1932, *A survey of the external galxies brighter than the thirteenth magnitude*, Annals of Harvard College Observatory, 88, 41

Sheskin, D. J. 2007, Handbook of Parametric and Nonparametric Statistical Procedures, 4th edn. (Chapman & Hall/CRC, Boca Raton)

Sheth, R. K. & Diaferio, A. 2011, *How unusual are the Shapley supercluster and the Sloan Great Wall?*, MNRAS, 417, 2938

Sheth, R. K. & van de Weygaert, R. 2004, *A hierarchy of voids: much ado about nothing*, MNRAS, 350, 517

Shi, F., Yang, X., Wang, H., et al. 2016, *The real space clustering of galaxies in SDSS DR7: I. Two point correlation functions*, ArXiv e-prints

Silverman, B. W. & Young, G. A. 1987, *The bootstrap: To smooth or not to smooth?*, Biometrika, 74, 469

Skibba, R. A. & Macciò, A. V. 2011, *Properties of dark matter haloes and their correlations: the lesson from principal component analysis*, MNRAS, 416, 2388

Sousbie, T. 2011, *The persistent cosmic web and its filamentary structure - I. Theory and implementation*, MNRAS, 414, 350

Sousbie, T., Pichon, C., Courtois, H., Colombi, S., & Novikov, D. 2008, *The Three-dimensional Skeleton of the SDSS*, ApJ, 672, L1

Springel, V., White, S. D. M., Jenkins, A., et al. 2005, *Simulations of the formation, evolution and clustering of galaxies and quasars*, Nature, 435, 629

Stoughton, C., Lupton, R. H., Bernardi, M., et al. 2002, *Sloan Digital Sky Survey: Early Data Release*, AJ, 123, 485

Strauss, M. A., Weinberg, D. H., Lupton, R. H., et al. 2002, *Spectroscopic Target Selection in the Sloan Digital Sky Survey: The Main Galaxy Sample*, AJ, 124, 1810

Suhhonenko, I., Einasto, J., Liivamägi, L. J., et al. 2011, *The cosmic web for density perturbations of various scales*, A&A, 531, A149

Szapudi, I. & Colombi, S. 1996, *Cosmic Error and Statistics of Large-Scale Structure*, ApJ, 470, 131

Tago, E., Einasto, J., Saar, E., et al. 2006, *Clusters and groups of galaxies in the 2dF galaxy redshift survey: A new catalogue*, Astronomische Nachrichten, 327, 365

Tago, E., Saar, E., Tempel, E., et al. 2010, *Groups of galaxies in the SDSS Data Release 7 . Flux- and volume-limited samples*, A&A, 514, A102

Tarenghi, M., Tifft, W. G., Chincarini, G., Rood, H. J., & Thompson, L. A. 1978, *The Structure of the Hercules Supercluster*, in IAU Symposium, Vol. 79, Large Scale Structures in the Universe, ed. M. S. Longair & J. Einasto, 263

Tegmark, M., Hamilton, A. J. S., Strauss, M. A., Vogeley, M. S., & Szalay, A. S. 1998, *Measuring the Galaxy Power Spectrum with Future Redshift Surveys*, ApJ, 499, 555

Tempel, E. 2011, *Tracing galaxy evolution by their present-day luminosity function*, PhD thesis, University of Tartu

Tempel, E., Einasto, J., Einasto, M., Saar, E., & Tago, E. 2009, *Anatomy of luminosity functions: the 2dFGRS example*, A&A, 495, 37

Tempel, E., Saar, E., Liivamägi, L. J., et al. 2011, *Galaxy morphology, luminosity, and environment in the SDSS DR7*, A&A, 529, A53

Tempel, E., Tago, E., & Liivamägi, L. J. 2012, *Groups and clusters of galaxies in the SDSS DR8. Value-added catalogues*, A&A, 540, A106

Tempel, E., Tamm, A., Gramann, M., et al. 2014, *Flux- and volume-limited groups/clusters for the SDSS galaxies: catalogues and mass estimation*, ArXiv e-prints

Tiit, E. & Einasto, J. 1964, *Faktornyj analiz krasnyh karlikov. Factor Analysis of Red Dwarfs.*, Publications of the Tartu Astrofizica Observatory, 34, 156

Toribio, M. C., Solanes, J. M., Giovanelli, R., Haynes, M. P., & Martin, A. M. 2011, *H I Content and Optical Properties of Field Galaxies from the ALFALFA Survey. II. Multivariate Analysis of a Galaxy Sample in Low-density Environments*, ApJ, 732, 93

Trimble, V. 1995, *The 1920 Shapley-Curtis Discussion: Background, Issues, and Aftermath*, PASP, 107, 1133

Tully, R. B., Courtois, H., Hoffman, Y., & Pomarède, D. 2014, *The Laniakea supercluster of galaxies*, Nature, 513, 71

Tully, R. B., Courtois, H. M., Dolphin, A. E., et al. 2013, *Cosmicflows-2: The Data*, AJ, 146, 86

Tully, R. B. & Fisher, J. R. 1977, *A new method of determining distances to galaxies*, A&A, 54, 661

Vogelsberger, M., Genel, S., Springel, V., et al. 2014, *Introducing the Illustris Project: simulating the coevolution of dark and visible matter in the Universe*, MNRAS, 444, 1518

Wake, D. A., Nichol, R. C., Eisenstein, D. J., et al. 2006, *The 2df SDSS LRG and QSO survey: evolution of the luminosity function of luminous red galaxies to z = 0.6*, MNRAS, 372, 537

Yahil, A., Strauss, M. A., Davis, M., & Huchra, J. P. 1991, *A redshift survey of IRAS galaxies. II - Methods for determining self-consistent velocity and density fields*, ApJ, 372, 380

Zeldovich, I. B., Einasto, J., & Shandarin, S. F. 1982, *Giant voids in the universe*, Nature, 300, 407

Zeldovich, Y. B. 1970, *Gravitational instability: An approximate theory for large density perturbations.*, A&A, 5, 84

Zeldovich, Y. B. 1972, *A hypothesis, unifying the structure and the entropy of the Universe*, MNRAS, 160, 1P

Zwicky, F. 1937, *On a New Cluster of Nebulae in Pisces*, Proceedings of the National Academy of Science, 23, 251

Zwicky, F. 1938, *On the Clustering of Nebulae*, PASP, 50, 218

Zwicky, F., Herzog, E., Wild, P., Karpowicz, M., & Kowal, C. T. 1961–1968, Catalogue of galaxies and of clusters of galaxies (California Institute of Technology, Pasadena, USA)

SUMMARY IN ESTONIAN

Galaktikate superparvede omadused ja ruumjaotus

I sissejuhatus

Käesoleva töö eesmärgiks on kirjeldada suuremastaabilist galaktikate paiknemist universumis, uurides selleks suuri galaktikate ja galaktikaparvede kogumeid, mida nimetatakse superparvedeks.

Superparved on esimeses lähenduses ulatuslikud, valdavalt ebasümmeetrilised, märgatava galaktikate ületihedusega alad, mis võivad sisaldada kümneid kuni sadu galaktiparvi ning tuhandeid üksikuid või väikestes gruppides olevaid galaktikaid. Koos neid ümbritsevate hõredate piirkondade – tühikutega – moodustavad nad iseloomuliku suuremastaabilise struktuuri, mida selle välimuse tõttu „kosmiliseks võrguks" kutsutakse. Superparvede suurus, kuju ja ruumiline jaotus võimaldavad anda hinnangut erinevatele kosmoloogilise struktuuri tekke, arengu ja algtingimuste mudelitele ning seada tingimusi neis sisalduvatele parameetritele. Lisaks peaksid nad massiivsete aine kontsentratsioonidena jätma jälgi kosmilisse mikrolainelisse taustkiirgusse ja mõjutama neis asuvat galaktikate omadusi.

Universumi arengut tervikuna kirjeldavad kosmoloogiline printsiip, mis ütleb, et universum on homogeenne ja isotroopne, ning Einsteini üldrelatiivsusteooria. Fikseerides kosmoloogilist printsiipi hõlmava Fridman-Robertson-Walkeri meetrika, saab Einsteini võrrandist leida Fridmani võrrandid, mis kirjeldavad universumi skaalateguri muutust ajas. Fridmani võrrandid ennustavad muuhulgas, et universum ei saa olla staatilises seisundis. Seda kinnitavad arvukad vaatlused, mis näitavad, et universum paisub kiirenevalt. Tänapäevased vaatlustulemused on vastavuses nn. standardmudeliga, mille järgi koosneb universum 69% ulatuses tumedast energiast, 26% tumedast ainest ja 5% tavalisest, barüonainest. Geomeetriliselt on universum seejuures tasane. Paisumise kiirust iseloomustava Hubble'i konstandi väärtus on umbes 70 km/(s Mpc). Standardmudelis mõjutab tumeenergia universumi üldise paisumise kiirust; tume ja tavaline aine erinevad üksteisest selle tõttu, et tumeaine on nii enda kui kõige muuga ainult gravitatsioonilises vastastikmõjus[1]. Kuna tumeainet on barüonainest oluliselt rohkem, siis võib – kui väga universumi väga varajane staadium välja arvata – pea kogu struktuuri tekkimist kirjeldada gravitatsioonilise ebastabiilsuse teooria abil. Barüonidega seotud protsessid muutuvad struktuuri arengu seisukohalt oluliseks alles üksikute galaktikate mastaabis. Hälvete arengut kirjeldab sel juhul aine pidevuse, liikumis- (e. Euleri) ja gravitatsiooni (e. Poissoni) võrrandeid koondav võrrandisüsteem. Esialgsete häirituste allikaks arvatakse olevat väga varaja-

---

[1]Vähemalt struktuuri tekke seisukohalt.

se universumi kvantfluktuatsioonid, mis inflatsiooni ehk universumi eksponentsiaalse paisumise faasi järel makroskoopiliste mastaapideni laienesid. Alghäirituste jaotus ei ole väga oluline, ka lihtne skaalainvariantne lähendus, nn. Harrison-Zeldovitši spekter, on kooskõlas nii inflatsiooni mudelite kui praegu vaadeldava suuremastaabilise struktuuriga. Üldjuhul ei ole häirituste arengut kirjeldava võrrandisüsteemi lahendamine analüütiliselt võimalik ja praktikas kasutatakse numbrilisi mudeleid. Kasutades nn. N-keha arvutusi, on võimalik väga paindlikult simuleerida struktuuri teket ja arengut – katsetada erinevaid mudeleid ning algtingimusi, keskenduda huvipakkuvale suuruskaalale ning võtta arvesse ka keerulisemaid füüsikalisi protsesse. Laias laastus võib struktuuri tekkimist kirjeldada kahe samaaegselt toimuva protsessi abil. Esiteks anisotroopne kollaps, mis leiab aset kõige suurematel pikkusskaaladel ja mille mehhanismi pakkus välja J.Željdovitš 1970.-l aastatel, ning J. Peeblesi poolt samal ajal kirjeldatud hierarhiline kuhjumine. Esimene neist seletab suurte, karakteristike seina- ja filamendilaadsete struktuuride, so. ka superparvede, teket, teine aga kirjeldab galaktikate ja galaktikaparvede evolutsiooni.

Vaatluslikult on suuremastaabilise struktuuri uurimiseks esmajärjekorras vaja kaardistada galaktikate asukohad. Taevakehade koordinaatidest on kõige keerulisem määrata nende kaugust, otsene mõõtmine aastaparallaksi abil on võimalik vaid lähimate tähtede korral. Kaugemate objektide kauguste hindamiseks kasutatakse mitmesuguseid efekte, mis tuleb omavahel võimalikult täpselt kalibreerida. Tulemuseks on nn. kosmiline kauguste redel, mille viimaseks astmeks on seos galaktika kauguse ja mõõdetud spektrijoonte punanihke vahel. Galaktika punanihe moodustub peamiselt kahe protsessi koosmõjul: esiteks Doppleri nihe, mis tekib galaktikate omaliikumise tõttu ja teiseks nn. kosmoloogiline punanihe universumi üldisest paisumisest. Kosmoloogilise punanihke suurus on üheselt seotud skaalateguri, millest sõltuvad kõik pikkusmõõdud universumis, muutusega kiirguse eraldumise ja vastuvõtu vahel. Galaktika kauguse saab seejärel leida kosmoloogilise mudeli, so. Fridmani võrrandite ja meetrika avaldise kaudu. Galaktikate kaugust on samuti vaja näivatest heledustest absoluutsete heleduste arvutamiseks. Punanihete määramiseks vajaliku täpsusega spektrite mõõtmine on aeganõudev protsess, mille tõttu oli heade kaugushinnangutega galaktikate hulk pikka aega äärmiselt piiratud. Alles 1990.-l ja 2000.-l aastatel valmisid esimesed arvestatava suurusega punanihete kataloogid. Need sisaldasid algul kümnete (CfA, LCRS kataloogid), kuid hiljem juba sadade tuhandete (2dF, SDSS kataloogid) galaktikate koordinaate. Praeguseks hetkeks on erinevates taevaülevaadetes avaldatud mitmed miljonid galaktikate punanihked.

Arusaam galaktikatest kui valdavalt isoleeritud, suurtest tähtede, gaasi ja tolmu süsteemidest tekkis alles 20. sajandi alguses – väga hilja, kui arvestada astronoomiateaduse üldist vanust. Selle põhjuseks on peamiselt asjaolu, et ilma abistavate vaatlusinstrumentideta on vaadeldavad vaid üksikud galaktikad, ning veel tükk aega

pärast teleskoopide leviku algust 16. sajandil arvati nad meie enda galaktikas leiduvate hajusate objektide hulka. Esimesed õiges suurusjärgus hinnangud galaktikate kaugustele, mis paigutasid nad kindlalt kaugele väljapoole Linnuteed, leiti 1920.-l aastatel (sh. eesti astronoomi E. Öpiku poolt). Galaktikate ebaühtlane paiknemine taevasfääril oli selgelt nähtav juba 19. sajandil koostatud suuremates „udukogude" kataloogides. Kasutades esialgu vaid taevakoordinaate sisaldavaid katalooge, hindas G. Abell eelmise sajandi keskpaigal, et lisaks galaktikate koondumisele erineva suurusega parvedesse on selgelt täheldatav ka galaktikaparvede endi, nö. teist järku grupeerumine. Terminit superparv kasutas esimesena G. de Vaucouleurs, kes nimetas Linnuteed hõlmavat galaktikate ja galaktikaparvede kogumit Kohalikuks superparveks. Tartu Observatoorumis alustati superparvede uurimist 1970.-l aastatel ja sellest ajast peale on avaldatud kümneid neile pühendatud töid. Viimaste aastate olulisemad arengud on seotud galaktikate pekuliaarsete kiiruste kasutuselevõtuga gravitatsioonipotentsiaali täpseks määramiseks ning superparvedele range füüsikalise definitsiooni andmisega. Esialgu on piisava täpsusega mõõtetulemused piiratud vaid meie lähema ümbrusega, kuid sellest hoolimata on need nõudnud juba Kohaliku superparve olemuse ümberhindamist, ning viinud Laniakea superparve avastamiseni B. Tully poolt.

II lähteandmed ja meetod

Käesolevas töös on lähteandmetena kasutatud Sloani Digitaalse Taevaülevaate (SDSS) 8. väljaande nn. põhivalimit ja 7. väljaande heledate punaste galaktikate (LRG) valimit. Lisaks nendele on võrdluseks võetud Millenniumi N-keha simulatsioonil põhinev galaktikate kataloog, mille ruumala ja galaktikate tihedus on valitud sarnastena põhivalimi omadele. Põhivalimi ruumala on 0,132 $(h^{-1}$ Gpc$)^3$ ja see sisaldab 576 493 galaktikat, Millenniumi valimi vastavad näitajad on 0,125 $(h^{-1}$ Gpc$)^3$ ja 1 039 919 galaktikat. LRG valim on põhivalimist üle kahe korra suurema sügavusega (ruumala 1,789 $(h^{-1}$ Gpc$)^3$), kuid selle eest keskmiselt peaaegu 50 korda hõredam, sisaldades 170 423 tuhat galaktikat. Suurus h tähistab siin dimensioonitut Hubble'i konstanti. Põhivalimi jaoks on Tartu Observatoorumis koostatud ka 77 858 objekti sisaldav galaktikagruppide ja -parvede kataloog. Tuleb rõhutada, et vaatluslike valimite täielikkus on piiratud objektide näiva heledusega. Suurematel kaugustel on vaadeldavad ainult heledamad galaktikad, mis tähendab, et valimi tihedus langeb kauguse suurenedes.

Superparvede leidmiseks arvutatakse galaktikate põhjal heledustiheduse väljad, kasutades silumistuumana B_3 splainfunktsiooni, ja kaaludes iga galaktikat tema heledusega. Enne silumist korrigeeritakse põhivalimis galaktikate kaugusi parvesõrmedeks nimetatud vaatekiiresuunaliste punanihete moonutuste eemaldamiseks. Nende põhjuseks on galaktikate suured omakiirused galaktikaparvedes, mis venitavad

parved vaatekiire sihis näiliselt mitmekordselt välja. Seda efekti saab statistiliselt kompenseerida, võrdsustades galaktikate kiiruste dispersiooni parvedes koordinaatide dispersiooniga taeva tasandis. Efektiivse silumisraadiusena kasutatakse põhivalimis ja Millenniumi valimis 8 h^{-1}Mpc , mis on umbkaudseks piiriks üleminekul tugevalt mittelineaarsesse režiimi struktuuri tekkel. LRG valimis on väiksema numbrilise tiheduse tõttu kasutatud 16 h^{-1}Mpc silumisraadiust. Nagu mainisime, väheneb valimite tihedus koos kaugusega. Seda on võimalik kompenseerida, hinnates, kui palju galaktikate heledusest jääb nn. vaatlusaknast välja, ning kaaludes kõikide galaktikate heledused antud kaugusel läbi vastava kordajaga. Ehkki konkreetsete vaatlemata galaktikate asukohad ja heledused ei ole teada, saab niiviisi arvutada suuremas mastaabis kaugusest sõltumatu tihedusega välja. Kordajad leitakse võrreldes kogu galaktikate heledusfunktsiooni vastaval kaugusel leituga. Põhivalimil on kindlalt fikseeritud näiva heleduse piirid, mis ei sõltu kaugusest ega asukohtast taevasfääril. See asjaolu võimaldab arvutada kaalud korrektselt ning pidada ka korrigeeritud heledusi füüsikalisteks. LRG valimi heleduspiirid sõltuvad paraku vaadeldavate galaktikate kaugusest ja kaalutud heledused on seetõttu tinglikud. Viimase sammuna normeeritakse tihedusväljad valimi keskmise heledustihedusega.

Superparved leitakse tihedusväljast piirkondadena, mille sees on tihedus kõrgem teatud läviväärtusest. Kokku leitakse tihedusvälja objektid paljudel nivoodel, mille väärtused on võetud iga 0.1 ühiku tagant vahemikus umbes 2 kuni 10 keskmist tihedust. Nagu mitmetes varasemates töödes, esitatakse üks superparvede kataloog tihedusnivool väärtusega 5 keskmise tiheduse ühikut. Kuna see on sisuliselt vaba parameeter, koostatakse lisaks teine kataloog kasutades algoritmi, mis leiab igale superparvele individuaalse piirtiheduse vastavalt teda ümbritsevatele teistele struktuuridele.

III tulemused ja kokkuvõte
Töö peamiseks tulemuseks on kõigi kolme galaktikavalimi kohta jaoks koostatud superparvede kataloogid. Iga valimi kohta on leitud kaks kataloogi, neist üks fikseeritud ja teine adaptiivselt määratud tihedusnivoodel. Kataloogides esitatakse järgmised superparvede omadused: asukoht, galaktikate ja galaktikaparvede arv, ruumala, diameeter, koguheledus ning morfoloogilised kujuparameetrid. Mitmetele parameetritele on antud eri meetoditel arvutatud hinnangud. Kumbagi tüüpi kataloogides on igas valimis umbes sama arv superparvi, ent nende omadused on pisut erinevad. Superparvede omaduste jaotused osutavad, et valitud meetod ja selle parameetrid on olnud põhjendatud. Omaduste sõltuvused kaugusest ehk selektsiooniefektid on mõlemat tüüpi kataloogis üldjuhul väikesed, mis näitab, et vaatuslikke piiranguid on õnnestunud edukalt kompenseerida. Superparvede morfoloogia näitab, et superparved on valdavalt filamentaarse kujuga, kusjuures nende väljavenitatus suureneb koos objektide kogu-

heledusega. Superparvede omaduste vaheliste seoste leidmiseks on kasutatud peakomponentide analüüsi. See kinnitab samuti, et superparvede olulisimad parameetrid on diameeter ja koguheledus, mis omavahel on tihedalt seotud, ning ülejäänud omadused on sekundaarsed. Neist kahest tulemusest võib järeldada, et superparvi võib käsitleda kui massiivseid galaktikate ja galaktikaparvede ahelaid. Superparvede ruumilises jaotuses võib omakorda märgata suure tihedusega ja rikaste objektide teatava regulaarsusega grupeerumist. Publitseeritud katalooge ja superparvede leidmise meetodit on kasutatud mitmetes uurimustöödes nii Tartu Observatoorumis kui ka mujal.

Kataloogide puuduseks on suureskaalaliste punanihkemoonutuste, ehk nn. Kaiseri efekti mitte arvesse võtmine. Kaiseri efekt kirjeldab, kuidas suurtes mastaapides korreleeritud galaktikate omakiirused moonutavad kosmilise võrgustiku vaadeldavat kuju. Samuti on superparvede definitsioon nõrgal füüsikalisel alusel. Ehkki superparved galaktikate jaotuses visuaalselt domineerivad, on neile täpsema definitsiooni andmine osutunud keerukaks. Superparved on kaugelt liiga suured, et, võttes arvesse universumi eluiga, olla galaktikaparvede kombel gravitatsiooniliselt seotud. Samuti on nad pideva võrgustiku osadena omavahel ühendatud, mis teeb raskeks üldiste üheselt põhjendatud piirtingimuste seadmise. Mõlemat probleemi võib aidata edaspidi lahendada viimaste aastate jooksul tehtud edusammud galaktikate pekuliaarsete kiiruste leidmisel, mille kaudu on võimalik leida tegelik gravitatsioonipotentsiaal ja kogu aine jaotus.

ACKNOWLEDGEMENTS

First and foremost I would like to thank my supervisor, dr. Enn Saar, who patiently saw me through this long and arduous process. He usually let me do things as I saw fit, but always came to offer his unwavering support and direction, when I most needed it.

I am sincerely thankful to my many colleagues who have provided me with supportive and inspiring atmosphere at the Tartu Observatory. My greatest appreciation goes to Maret Einasto and Elmo Tempel, who have been my closest collaborators, but also to Jaan Einasto, Jaan Pelt, Antti Tamm, Taavi Tuvikene and Ilmar Ansko, from whom I have often received invaluable advice. I am indebted to my friends Sander Mirme and Kristjan Kannike, who have been my pillars of knowledge in science and many other areas. I thank Enn Saar and Marika Liivamägi for the English language editing of the script.

There would have been no chance for me to reach this point without the limitless enthusiasm and warmth supplied by all of my family and friends.

I am also deeply grateful to Vicent Martínez from the Valencia University, Volker Müller from the Potsdam Astrophysical Institute, and Pekka Heinämäki from the Tuorla Observatory, whose hospitality I have been able to enjoy on several occasions while working on the thesis.

PUBLICATIONS

Curriculum vitae

Personal data

Name	Lauri Juhan Liivamägi
Date and place of birth	6 May 1981, Käina, Estonia
Citizenship	Estonian
Current employment	Tartu Observatory (research associate)
Address	Tartu Observatory
	61602 Tõravere
	Tartumaa, Estonia
Phone	(+372) 5561 3831
E-mail	juhan.liivamagi@ut.ee

Education

1996 – 1999	Hugo Treffner Gymnasium
1999 – 2003	University of Tartu, undergraduate student, BSc 2003 (astrophysics)
2003 – 2005	University of Tartu, graduate student, MSc 2005 (astrophysics)
2005 – 20xx	University of Tartu, PhD student

Employment

2005 – 2010	Tartu Observatory, extraordinary research associate
2010 – . . .	Tartu Observatory, research associate

Professional training

31.05 – 04.06 2004	Summer school "Cosmology and Planck", Tuorla, Finland
06.09 – 10.09 2004	Summer school "Data analysis in Cosmology", Valencia, Spain
17.07 – 11.08 2006	Summer school "Helmholtz Summer School 2006: Supercomputational Cosmology", Potsdam, Germany
10.09 – 21.09 2007	Summer school "Novicosmo 2007: fiat lux – formation and evolution of cosmic structures", Novigrad, Croatia
18.05 – 20.05 2009	Summer school "Scientific writing for young astronomers", Blankenberge, Belgium

08.06 – 18.06 2009	Summer school "Nordic-Baltic Optical/NIR and Radio Astronomy Summer School: Star Formation in the Milky Way and Nearby Galaxies", Tuorla, Finland
01.07 – 03.07 2009	Summer school "International summer school: future cosmic sky surveys and huge databases", Tartu, Estonia
12.06 – 20.06 2010	Summer school "CSC summer school in scientific and high-performance computing", Espoo, Finland
25.07 – 03.08 2012	Summer school "4th International Summer School on Astroparticle Physics: NIJMEGEN12", Nijmegen, Netherlands and Kleve, Germany

Conference presentations

20.06 – 21.06 2005	Workshop "Tartu – Tuorla annual meeting 2005: From galaxies to superclusters", Tõravere, Estonia *Oral presentation*: "Reduction of cluster-fingers by means of wavelet analysis"
01.10 – 04.10 2008	Workshop "Tartu – Tuorla annual meeting 2008: Cosmology: from observations to simulations and beyond", Elva, Estonia *Oral presentation*: "Density fields and superclusters"
02.09 – 03.09 2010	Workshop "Tuorla – Tartu annual meeting 2010: Observational cosmology", Tuorla, Finland *Oral presentation*: "Superclusters as observational objects"
12.08 – 17.08 2011	Workshop "Cosmic Web Morphology and Topology", Warsaw, Poland *Oral presentation*: "Superclusters and their properties in the SDSS main and LRG galaxy samples"
21.09 – 22.09 2011	Workshop "Tartu – Tuorla annual meeting 2011: Observational cosmology", Tartu, Estonia *Oral presentation*: "Superclusters in Planck mission observations"
20.09 – 21.09 2012	Workshop "Tuorla – Tartu annual meeting 2012: Dark and visible Universe", Tuorla, Finland *Oral presentation*: "BAO and LSS"
25.05 – 29.05 2014	IAU Symposium 306: Statistical Challenges in 21st Century Cosmology", Lisbon, Portugal *Poster presentation*: "Density field projection analysis in search for WHIM"

| 1.10 – 3.10 2014 | Workshop "Tuorla – Tartu annual meeting 2014: Small and Large scale Universe", Tuorla, Soome *Oral presentation*: "Density field projection analysis" |

Language skills

| Estonian | the first language |
| English | good |

Honours and Awards

| 2007 | E. Öpik stipend (Tartu Observatory) |
| 2009 | E. Öpik stipend (Tartu Observatory) |

Fields of research

Galaxy clusters and superclusters, statistics of the large-scale structure.

Publications

1. Einasto, J., Einasto, M., Saar, E., Tago, E., **Liivamägi, L. J.**, Jõeveer, M., Suhhonenko, I., Hütsi, G., Jaaniste, J., Heinämäki, P., Müller, V., Knebe, A., Tucker, D. 2006, *Luminous superclusters: remnants from inflation?*, Astronomy & Astrophysics, 459, 1, pp. L1–L4

2. Hirv, A., Eenmäe, T., **Liivamägi, L. J.**, Pelt, J. 2007, *Estimation of Time Delays from Two Blended Light Curves of Gravitational Lenses*, Baltic Astronomy, 16, pp. 241–250

3. Einasto, J., Einasto, M., Saar, E., Tago, E., **Liivamägi, L. J.**, Jõeveer, M., Suhhonenko, I., Hütsi, G., Jaaniste, J., Heinämäki, P., Müller, V., Knebe, A., Tucker, D. 2007, *Superclusters of galaxies from the 2dF redshift survey. II. Comparison with simulations*, Astronomy & Astrophysics, 462, 2, pp. 397–410

4. Einasto, J., Einasto, M., Tago, E., Saar, E., Hütsi, G., Jõeveer, M., **Liivamägi, L. J.**, Suhhonenko, I., Jaaniste, J., Heinämäki, P., Müller, V., Knebe, A., Tucker, D. 2007, *Superclusters of galaxies from the 2dF redshift survey. I. The catalogue*, Astronomy & Astrophysics, 462, 2, pp. 811–825

5. Hirv, A., Eenmäe, T., Liimets, T., **Liivamägi, L. J.**, Pelt, J. 2007, *Estimation of time delays from unresolved photometry*, Astronomy & Astrophysics, 464, 2, pp. 471–478

6. Einasto, M., Einasto, J., Tago, E., Saar, E., **Liivamägi, L. J.**, Jõeveer, M., Hütsi, G., Heinämäki, P., Müller, V., Tucker, D. 2007, *Superclusters of galaxies in the 2dF redshift survey. III. The properties of galaxies in superclusters*, Astronomy & Astrophysics, 464, 3, pp. 815–826

7. Einasto, M., Saar, E., **Liivamägi, L. J.**, Einasto, J., Tago, E., Martínez, V. J., Starck, J.-L., Müller, V., Heinämäki, P., Nurmi, P., Gramann, M., Hütsi, G. 2007, *The richest superclusters. I. Morphology*, Astronomy & Astrophysics, 476, 2, pp. 697–711

8. Einasto, M., Saar, E., Martínez, V. J., Einasto, J., **Liivamägi, L. J.**, Tago, E., Starck, J.-L., Müller, V., Heinämäki, P., Nurmi, P., Paredes, S., Gramann, M., Hütsi, G. 2008, *Toward Understanding Rich Superclusters*, The Astrophysical Journal, 685, 1, pp. 83–104

9. Heinämäki, P., Nurmi, P., Tago, E., Saar, E., **Liivamägi, L. J.**, Tempel, E., Einasto, M., Einasto, J., Lietzen, H., Takalo, L. 2008, *Nearby quasars in SDSS*, Proceedings of the international conference held at Russian Geographical Society, Problems of practical cosmology, ed. by Y. V. Baryshev, I. N. Taganov, and P. Teerikorpi, Vol. 1, pp. 123–128

10. Lietzen, H., Heinämäki, P., Nurmi, P., Tago, E., Saar, E., **Liivamägi, L. J.**, Tempel, E., Einasto, M., Einasto, J., Gramann, M., Takalo, L. O. 2009, *Environments of nearby quasars in Sloan Digital Sky Survey*, Astronomy & Astrophysics, 501, 1, pp. 145–155

11. Lietzen, H., Heinämäki, P., Nurmi, P., **Liivamägi, L. J.**, Saar, E., Tago, E., Tempel, E., Einasto, M., Einasto, J., Gramann, M., Takalo, L. O. 2010, *Large Scale Environments of Nearby Quasars*, AIP Conference Proceedings, Hunting for the dark: the hidden side of galaxy formation, ed. V. P. Debattista and C. C. Popescu, Vol. 1240, pp. 249–250

12. Einasto, M., Tago, E., Saar, E., Nurmi, P., Enkvist, I., Einasto, P., Heinämäki, P., **Liivamägi, L. J.**, Tempel, E., Einasto, J., Martínez, V. J., Vennik, J., Pihajoki, P. 2010, *The Sloan great wall. Rich clusters*, Astronomy & Astrophysics, 522, A92

13. Tempel, E., Saar, E., **Liivamägi, L. J.**, Tamm, A., Einasto, J., Einasto, M., Müller, V. 2011, *Galaxy morphology, luminosity, and environment in the SDSS DR7*, Astronomy & Astrophysics, 529, A53

14. Einasto, J., Hütsi, G., Saar, E., Suhhonenko, I., **Liivamägi, L. J.**, Einasto, M., Müller, V., Tago, E., Starobinsky A. A., Tempel, E. 2011, *Wavelet analysis of the formation of the cosmic web*, Astronomy & Astrophysics, 531, A75

15. Suhhonenko, I., Einasto, J., **Liivamägi, L. J.**, Saar, E., Einasto, M., Hütsi, G., Müller, V., Starobinsky, A. A., Tago, E., Tempel, E. 2011, *The cosmic web for density perturbations of various scales*, Astronomy & Astrophysics, 531, A149

16. Einasto, M., **Liivamägi, L. J.**, Tempel, E., Saar, E., Tago, E., Einasto, P., Enkvist, I., Einasto, J., Martínez, V. J., Heinämäki, P., Nurmi, P. 2011, *The Sloan Great Wall. Morphology and Galaxy Content*, The Astrophysical Journal, 736, 1, 51

17. Einasto, M., **Liivamägi, L. J.**, Tago, E., Saar, E., Tempel, E., Einasto, J., Martínez, V. J., Heinämäki, P. 2011, Astronomy & Astrophysics, 532, A5

18. Einasto, J., Suhhonenko, I., Hütsi, G., Saar, E., Einasto, M., **Liivamägi, L. J.**, Müller, V., Starobinsky, A. A., Tago, E., Tempel, E. 2011, *Towards understanding the structure of voids in the cosmic web*, Astronomy & Astrophysics, 534, A128

19. Lietzen, H., Heinämäki, P., Nurmi, P., **Liivamägi, L. J.**, Saar, E., Tago, E., Takalo, L. O., Einasto, M. 2011, *Large-scale environments of z < 0.4 active galaxies*, Astronomy & Astrophysics, 535, A21

20. Einasto, M., **Liivamägi, L. J.**, Saar, E., Einasto, J., Tempel, E., Tago, E., Martínez, V. J. 2011, *SDSS DR7 superclusters. Principal component analysis*, Astronomy & Astrophysics, 535, A36

21. **Liivamägi, L. J.**, Tempel, E., Saar, E. 2012, *SDSS DR7 superclusters. The catalogues*, Astronomy & Astrophysics, 539, A80

22. Tempel, E., Tago, E., **Liivamägi, L. J.** 2012, *Groups and clusters of galaxies in the SDSS DR8. Value-added catalogues*, Astronomy & Astrophysics, 540, A106

23. Einasto, M., Vennik, J., Nurmi, P., Tempel, E., Ahvensalmi, A., Tago, E., **Liivamägi, L. J.**, Saar, E., Heinämäki, P., Einasto, J., Martínez, V. J. 2012, *Multimodality in galaxy clusters from SDSS DR8: substructure and velocity distribution*, Astronomy & Astrophysics, 540, A123

24. Einasto, M., **Liivamägi, L. J.**, Tempel, E., Saar, E., Vennik, J., Nurmi, P., Gramann, M., Einasto, J., Tago, E., Heinämäki, P., Ahvensalmi, A., Martínez, V. J. 2012, *Multimodality of rich clusters from the SDSS DR8 within the supercluster-void network*, Astronomy & Astrophysics, 542, A36

25. Stoica, R. S., Tempel, E., **Liivamägi, L. J.**, Castellan, G., Saar, E. 2014, *Spatial Patterns Analysis in Cosmology based on Marked Point Processes*, EAS Publications Series, ed. by D. Fraix-Burnet, Vol. 66, pp. 197–226

26. Tempel, E., Libeskind, N. I., Hoffman, Y., **Liivamägi, L. J.**, Tamm, A. 2014, *Orientation of cosmic web filaments with respect to the underlying velocity field*, Monthly Notices of the Royal Astronomical Society: Letters, 437, 1, pp. L11–L15

27. Einasto, M., Lietzen, H., Tempel, E., Gramann, M., **Liivamägi, L. J.**, Einasto, J. 2014, *SDSS superclusters: morphology and galaxy content*, Astronomy & Astrophysics, 562, A87

28. Tempel, E., Stoica, R. S., Martínez, V. J., **Liivamägi, L. J.**, Castellan, G., Saar, E. 2014, *Detecting filamentary pattern in the cosmic web: a catalogue of filaments for the SDSS*, Monthly Notices of the Royal Astronomical Society, 438, 4, pp. 3465–3482

29. **Liivamägi, L. J.**, 2014, *Density field projection analysis in search for WHIM*, Proceedings IAU Symposium No. 306: Statistical Challenges in 21st Century Cosmology, ed. by A. Heavens, J.-L. Starck, and A. Krone-Martins, pp. 310–312

30. Tempel, E., Tamm, A., Gramann, M., Tuvikene, T., **Liivamägi, L. J.**, Suhhonenko, I., Kipper, R., Einasto, M., Saar, E. 2014 *Flux- and volume-limited groups/clusters for the SDSS galaxies: catalogues and mass estimation*, Astronomy & Astrophysics, 566, A1

31. Einasto, M., Tago, E., Lietzen, H., Park, C., Heinämäki, P., Saar, E., Song, H., **Liivamägi, L. J.**, Einasto, J. 2014, *Tracing a high redshift cosmic web with quasar systems*, Astronomy & Astrophysics, 568, A46

32. Einasto, M., Gramann, M., Saar, E., **Liivamägi, L. J.**, Tempel, E., Nevalainen, J., Heinämäki, P., Park, C., Einasto, J. 2015, *Unusual A2142 supercluster with a collapsing core: distribution of light and mass*, Astronomy & Astrophysics, 580, A69

33. Stoica, R. S.; Liu, S.; **Liivamägi, L. J.**; Saar, E.; Tempel, E.; Deleflie, F.; Fouchard, M.; Hestroffer, D.; Kovalenko, I.; Vienne, A. 2016, *An integrative approach based on probabilistic modelling and statistical inference for morpho-statistical characterization of astronomical data*, Romanian Astronomical Journal, 26, 11

34. Nevalainen, J., Tempel, E., **Liivamägi, L. J.**, Branchini, E., Roncarelli, M., Giocoli, C., Heinämäki, P., Saar, E., Tamm, A., Finoguenov, A., Nurmi, P., Bonamente, M. 2015, *Missing baryons traced by the galaxy luminosity density in the large-scale WHIM filaments*, Astronomy & Astrophysics, 583, A142

35. Einasto, M.; Heinämäki, P.; **Liivamägi, L. J.**; Martínez, V. J.; Hurtado-Gil, L.; Arnalte-Mur, P.; Nurmi, P.; Einasto, J.; Saar, E. 2016, *Shell-like structures in our cosmic neighbourhood*, Astronomy & Astrophysics, 587, A116

36. Lietzen, H.; Tempel, E.; **Liivamägi, L. J.**; Montero-Dorta, A.; Einasto, M.; Streblyanska, A.; Maraston, C.; Rubiño-Martín, J. A.; Saar, E. 2016, *Discovery of a massive supercluster system at z ~ 0.47*, Astronomy & Astrophysics, 588, L4

37. Bonamente, M.; Nevalainen, J.; Tilton, E.; **Liivamägi, J.**; Tempel, E.; Heinämäki, P.; Fang, T. 2016, *A possible Chandra and Hubble Space Telescope detection of extragalactic WHIM towards PG 1116+215*, Monthly Notices of the Royal Astronomical Society, 457, 4, pp. 4236–4247

38. Einasto, M.; Lietzen, H.; Gramann, M.; Tempel, E.; Saar, E.; **Liivamägi, L. J.**; Heinämäki, P.; Nurmi, P.; Einasto, J., *Sloan Great Wall as a complex of superclusters with collapsing cores*, Astronomy & Astrophysics, 595, A70

ELULOOKIRJELDUS

Isikuandmed

Nimi	Lauri Juhan Liivamägi
Sünniaeg ja -koht	6. mai 1981, Käina, Eesti
Kodakondsus	eesti
Praegune töökoht	Tartu Observatoorium (teadur)
Aadress	Tartu Observatoorium
	61602 Tõravere
	Tartumaa, Eesti
Telefon	(+372) 5561 3831
E-post	juhan.liivamagi@ut.ee

Haridus

1996 – 1999	Hugo Treffneri Gümnaasium
1999 – 2003	Tartu Ülikool, üliõpilane, BSc 2003 (astrofüüsika)
2003 – 2005	Tartu Ülikool, magistrant, MSc 2005 (astrofüüsika)
2005 – 20xx	Tartu Ülikool, doktorant

Teenistuskäik

2005 – 2010	Tartu Observatoorium, erakorraline teadur
2010 – . . .	Tartu Observatoorium, teadur

Täiendkoolitus

31.05 – 04.06 2004	Suvekool „Cosmology and Planck", Tuorla, Soome
06.09 – 10.09 2004	Suvekool „Data analysis in Cosmology", Valencia, Hispaania
17.07 – 11.08 2006	Suvekool „Helmholtz Summer School 2006: Supercomputational Cosmology", Potsdam, Saksamaa
10.09 – 21.09 2007	Suvekool „Novicosmo 2007: fiat lux – formation and evolution of cosmic structures", Novigrad, Horvaatia
18.05 – 20.05 2009	Suvekool „Scientific writing for young astronomers", Blankenberge, Belgia
08.06 – 18.06 2009	Suvekool „Nordic-Baltic Optical/NIR and Radio Astronomy Summer School: Star Formation in the Milky Way and Nearby Galaxies", Tuorla, Soome

01.07 – 03.07 2009	Suvekool „International summer school: future cosmic sky surveys and huge databases“, Tartu
12.06 – 20.06 2010	Suvekool „CSC summer school in scientific and high-performance computing“, Espoo, Soome
25.07 – 03.08 2012	Suvekool „4th International Summer School on Astroparticle Physics: NIJMEGEN12“, Nijmegen, Holland ja Kleve, Saksamaa

Konverentside ettekanded

20.06 – 21.06 2005	Konverents „Tartu – Tuorla annual meeting 2005: From galaxies to superclusters“, Tõravere *Suuline ettekanne*: „Reduction of cluster-fingers by means of wavelet analysis“
01.10 – 04.10 2008	Konverents „Tartu – Tuorla annual meeting 2008: Cosmology: from observations to simulations and beyond“, Elva *Suuline ettekanne*: „Density fields and superclusters“
02.09 – 03.09 2010	Konverents „Tuorla – Tartu annual meeting 2010: Observational cosmology“, Tuorla, Soome *Suuline ettekanne*: „Superclusters as observational objects”
12.07 – 17.07 2011	Konverents „Cosmic Web Morphology and Topology“, Varssavi, Poola *Suuline ettekanne*: „Superclusters and their properties in the SDSS main and LRG galaxy samples“
21.09 – 22.09 2011	Konverents „Tartu – Tuorla annual meeting 2011: Observational cosmology“, Tartu *Suuline ettekanne*: „Superclusters in Planck mission observations“
20.09 – 21.09 2012	Konverents „Tuorla – Tartu annual meeting 2012: Dark and visible Universe“, Tuorla, Soome *Suuline ettekanne*: „BAO and LSS“
25.05 – 29.05 2014	IAU sümpoosium 306: „Statistical Challenges in 21st Century Cosmology“, Lissabon, Portugal *Plakat*: „Density field projection analysis in search for WHIM“
1.10 – 3.10 2014	Konverents „Tuorla – Tartu annual meeting 2014: Small and Large scale Universe“, Tuorla, Soome *Suuline ettekanne*: „Density field projection analysis“

Keelteoskus

eesti keel	emakeel
inglise keel	hea

Uurimistoetused ja stipendiumid

2007	E. Öpiku stipendium (Tartu Observatoorium)
2009	E. Öpiku stipendium (Tartu Observatoorium)

Peamised uurimissuunad

Galaktikate parved ja superparved, universumi suuremastaabiline struktuur.

DISSERTATIONES ASTRONOMIAE
UNIVERSITATIS TARTUENSIS

1. **Tõnu Viik.** Numerical realizations of analytical methods in theory of radiative transfer. Tartu, 1991.
2. **Enn Saar.** Geometry of the large scale structure of the Universe. Tartu, 1991.
3. **Maret Einasto.** Morphological and luminosity segregation of galaxies. Tartu, 1991.
4. **Urmas Haud.** Dark Matter in galaxies. Tartu, 1991.
5. **Eugene A. Ustinov.** Inverse problems of radiative transfer in sounding of planetary atmospheres. Tartu, 1992.
6. **Peeter Tenjes.** Models of regular galaxies. Tartu, 1993.
7. **Ivar Suisalu.** Simulation of the evolution of large scale structure elements with adaptive multigrid method. Tartu, 1995.
8. **Teimuraz Shvelidze.** Automated quantitative spectral classification of stars by means of objective prism spectra: the method and applications. Tartu, 1999.
9. **Jelena Gerškevitš.** Formation and evolution of binary systems with compact objects. Tartu, 2002.
10. **Ivan Suhhonenko**. Large-scale motions in the universe. Tartu, 2003.
11. **Antti Tamm.** Structure of distant disk galaxies. Tartu, 2006.
12. **Vladislav-Veniamin Pustynski.** Modeling the reflection effect in pre-cataclysmic binary systems. Tartu, 2007.
13. **Anna Aret.** Evolutionary separation of mercury isotopes in atmospheres of chemically peculiar stars. Tartu, 2009.
14. **Mari Burmeister.** Characteristics of the hot components of symbiotic stars. Tartu, 2010.
15. **Elmo Tempel.** Tracing galaxy evolution by their present-day luminosity function. Tartu, 2011.
16. **Anti Hirv.** Estimation of time delays from light curves of gravitationally lensed quasars. Tartu, 2011.
17. **Rain Kipper**. Galaxy modelling: dynamical methods and applications. Tartu, 2016, 134 p.