

The supercluster-void network V.

The regularity periodogram

E. Saar¹, J. Einasto¹, O. Toomet¹, A. A. Starobinsky^{2,3}, H. Andernach⁴, M. Einasto¹, E. Kasak¹, and E. Tago¹

¹ Tartu Observatory, 61602 Tõravere, Estonia

² Research Center for the Early Universe, University of Tokyo, Tokyo 113-0033, Japan

³ Landau Institute for Theoretical Physics, Moscow 117 334, Russia

⁴ Depto. de Astronomía, Univ. Guanajuato, Apdo. Postal 144, C.P. 36000, Guanajuato, Mexico

Received 27 July 1999 / Accepted 28 May 2002

Abstract. We analyze the distribution of Abell clusters of galaxies to study the regularity of the supercluster-void network. We apply a new method, the regularity periodogram, that is sensitive to the geometry of the location of clusters, and measure the regularity of the network. We find that the supercluster-void network resembles a cubic lattice over the entire volume considered (out to the distance of $350 h^{-1}$ Mpc). The distribution of clusters in rich superclusters is not isotropic: it is periodic along a cubic lattice approximately aligned with the supergalactic coordinates SGX, SGY, SGZ . This large-scale inhomogeneity does not contradict recent CMB data, but its theoretical explanation remains a challenge.

Key words. cosmology: observations – cosmology: large-scale structure of the Universe

1. Introduction

The basic assumption in the standard Friedman-Robertson-Walker cosmology is that the Universe is homogeneous and isotropic on large scales (Peebles 1980). Clusters of galaxies and galaxy filaments form superclusters (Abell 1958); examples are the Local supercluster (de Vaucouleurs 1956; Einasto et al. 1984) and the Perseus-Pisces supercluster (Jõeveer et al. 1978; Giovanelli 1993). Superclusters and voids between them form a continuous network of high- and low-density regions which extends over the entire part of the Universe studied in sufficient detail up to a redshift of $z \approx 0.13$ (Einasto et al. 1994, 1997d, hereafter Paper I). This network has been called cellular (Jõeveer & Einasto 1978; Icke 1984), in which one cell is formed by a low-density region surrounded by superclusters and dense galaxy filaments. According to the conventional paradigm the location of superclusters and centres of voids in this cellular network is almost random, as density perturbations on all scales are believed to form a Gaussian random field (see, e.g., Peacock 1999).

A first hint that something may be wrong in our understanding of the formation of the large-scale structure of the Universe came when Broadhurst et al. (1990) measured redshifts of galaxies in a narrow beam along the direction of the northern and southern Galactic poles. They found that the distribution of galaxies along that beam is periodic: high-density regions alternate with low-density ones with a surprisingly constant interval of $128 h^{-1}$ Mpc (here h is the Hubble constant in units of $100 \text{ km s}^{-1} \text{ Mpc}^{-1}$). An attempt to explain this

distribution by a Voronoi foam model was made by van de Weygaert (1991). In the Voronoi model void centers are randomly distributed, clusters of galaxies occupy vertices of the network, and galaxies lie in sheets between voids. Quasi-periodicity is observed in approximately 10% of cases when a narrow beam is cut from the cellular network. Kaiser & Peacock (1991) showed that such a periodicity can occur even in a random distribution of galaxy clumps due to a specific geometry of the sample (very narrow beams). Thus the first reaction to the Broadhurst et al. result was that the observed quasi-periodicity can be explained in the framework of the conventional model of structure formation. However, more extensive simulations have shown recently that such regularity has a very small a priori probability, less than 10^{-3} (Yoshida et al. 2001).

There exists growing evidence that the supercluster-void network is more regular than expected from the Voronoi foam and the random clump model. The three-dimensional distribution of clusters shows clear signs of regularity, as shown by Einasto et al. (1994) and (1997d, Paper I of this series). One method to characterize this regularity is correlation analysis. Kopylov et al. (1984, 1988), Fetisova et al. (1993), Mo et al. (1992), and Einasto & Gramann (1993) have found evidence for the presence of a secondary peak of the correlation function of clusters of galaxies at $125 h^{-1}$ Mpc. This peak has been interpreted as the correlation of clusters in superclusters located at opposite sides of large voids. Such a secondary peak is expected also in the Voronoi foam model (see below). Recent studies show that the correlation function has not only one secondary maximum but a series of regularly spaced maxima and

Send offprint requests to: J. Einasto, e-mail: einasto@aai.ee

minima. In other words, it oscillates with a period equal to that of the periodicity of the supercluster-void network (Einasto et al. 1997b, hereafter Paper II). The existence of these oscillations has been confirmed using more recent data, although the statistical significance of the oscillations is not too high (Tago et al. 2000). An extremely low amplitude oscillation is seen also in the correlation function of LCRS galaxies (Tucker et al. 1997).

Regular oscillations of the correlation function correspond to a sharp peak in the power spectrum (Einasto et al. 1997c, hereafter Paper III). Baugh & Efstathiou (1993, 1994) and Gaztañaga & Baugh (1998) determined the three-dimensional power spectrum from the projected distribution of the APM galaxies; the Abell cluster power spectrum was derived by Einasto et al. (1997a) and Retzlaff et al. (1998), Tadros et al. (1998) calculated the power spectrum for the APM clusters. All these studies found the peak or turnover of the power spectrum near a wavelength of $\lambda_0 = 120\text{--}130 h^{-1} \text{ Mpc}$ or a wavenumber $k_0 = 2\pi/\lambda_0 = 0.05 h \text{ Mpc}^{-1}$.

This wavelength, however, corresponds to the case of the Einstein-de Sitter cosmological model ($\Omega_m = 1$). Recent observational data have changed our conception about the form of the matter power spectrum in the wavelength interval involved. First, observations of type Ia supernova explosions at high redshifts (Perlmutter et al. 1998, 1999; Garnavich et al. 1998; Riess et al. 1999) strongly support the existence of a positive cosmological constant in the Universe. When combined with results of the most recent CMB experiments BOOMERANG (de Bernardis et al. 2000) and MAXIMA-1 (Hanany et al. 2000), these data favor the flat Universe with $\Omega_m = 1 - \Omega_\Lambda = 0.3 \pm 0.1$. In the low density flat LCDM Universe with the flat ($n_s = 1$) initial spectrum of adiabatic perturbations, the maximum of the present matter power spectrum should lie at $k \approx 0.1\Gamma h \text{ Mpc}^{-1}$, where $\Gamma \approx \Omega_m h$ is the shape parameter, i.e. at $k \sim 0.02 h \text{ Mpc}^{-1}$. Thus, the additional peak at $k = k_0 = 0.05 h \text{ Mpc}^{-1}$ discussed above does not coincide with the main peak of the smooth power spectrum $P(k)$ (as would take place for $\Omega_m = 1$), but is located on its right slope. Moreover, if the empirical power spectrum $P(k)$ obtained in Einasto et al. (1999, hereafter E99) for $k \geq 0.03 h \text{ Mpc}^{-1}$ from galaxy and cluster catalogs is taken with the proper bias factor (see Einasto et al. 1999) and compared with the $n_s = 1$ power spectrum for the LCDM cosmological model, normalized to the COBE data on anisotropies of the cosmic microwave background (CMB) at very small k , the peak at $k = k_0 = 0.05 h \text{ Mpc}^{-1}$ appears to be only a local bump above a general depression in the empirical spectrum occurring for $k > 0.03 h \text{ Mpc}^{-1}$. So, there is no (or practically no) *excess* power in this region, as compared to the case of a featureless, scale-free initial matter power spectrum. In addition, a pronounced well at $k \sim 0.035 h \text{ Mpc}^{-1}$ arises. Therefore, now we suppose that a feature in the power spectrum corresponding to the large-scale quasi-regularity which we are considering probably has a more complicated structure than simply one bump (see Atrio-Barandela et al. 2001 for more discussion).

The most recent data on the power spectrum $P(k)$ from galaxy and cluster catalogs – the de-correlated power spectrum of the IRAS PSCz redshift survey (Hamilton et al. 2000)

that supersedes earlier results by Sutherland et al. (1999), the cluster spectrum by Miller & Batuski (2001), and the spectrum of the REFLEX X-ray cluster survey (Schuecker et al. 2001) all show the continuing increase of $P(k)$ to the direction of smaller k up to $k \sim 0.02 h \text{ Mpc}^{-1}$, in excellent agreement with the expected location of the main maximum in $P(k)$ given above. The situation with superimposed features that may reflect quasi-regularity is much more complicated. While the combined structure – the (relative) peak at $k = 0.05 h \text{ Mpc}^{-1}$ and the depression at $k \sim 0.035 h \text{ Mpc}^{-1}$ is clearly seen in the de-correlated PSCz spectrum and even in the Miller-Batuski data, still it is (almost) inside 1σ error bars for each point since errors are rather large in this region. Moreover, no sharp features are seen in the REFLEX spectrum. Once more, its error bars are sufficiently large to hide all expected deviations from the $n_s = 1$ initial spectrum. The recent determinations of the power spectrum of the 2dF survey also do not resolve the problem of the existence of local features. While Percival et al. (2001) find oscillations, in a later paper Tegmark et al. (2001) do not detect significant wiggles in the power spectrum.

A more detailed analysis of this question will be presented elsewhere. However, it is clear that presently available power spectra do not give a definite answer to the question if some local features in $P(k)$ over the range $k = 0.03\text{--}0.06 h \text{ Mpc}^{-1}$ related to the quasi-regularity exist or not. The same may be said about the present CMB data (see the discussion below, in Sect. 4).

The main reason why the recent precise results on $P(k)$ and the CMB multipoles C_l are still unable to resolve this problem (and still a significantly better accuracy is required) is because all these quantities are obtained from the module of the Fourier transform of inhomogeneities and averaging over rotations in 3-space. Valdarnini (2001) proposed to describe the density distribution by order parameters, which are sensitive both to radial and angular separations, but these are also found by averaging over rotations. Thus, these descriptors are insensitive to directional and phase information, and can describe regularity only indirectly. So far, the only positive results on regularity of the large-scale three-dimensional distribution of high-density regions have been found on the basis of oscillatory features of the correlation function of Abell clusters of galaxies, and confirmation of this result by independent methods is needed.

In this paper we shall use a new geometric method to investigate the distribution of clusters of galaxies, sensitive to the regularity of the distribution. Our main goal is to obtain independent evidence on the regularity problem. In the next section we describe the new method and check it using various test samples. Then we apply the method to the sample of Abell clusters of galaxies. Finally, we discuss and summarize our results.

2. The method

The method we use to study regularity of the distribution of clusters was proposed by Dr. J. Pelt and was first applied to the cluster distribution by O. Toomet (1997).

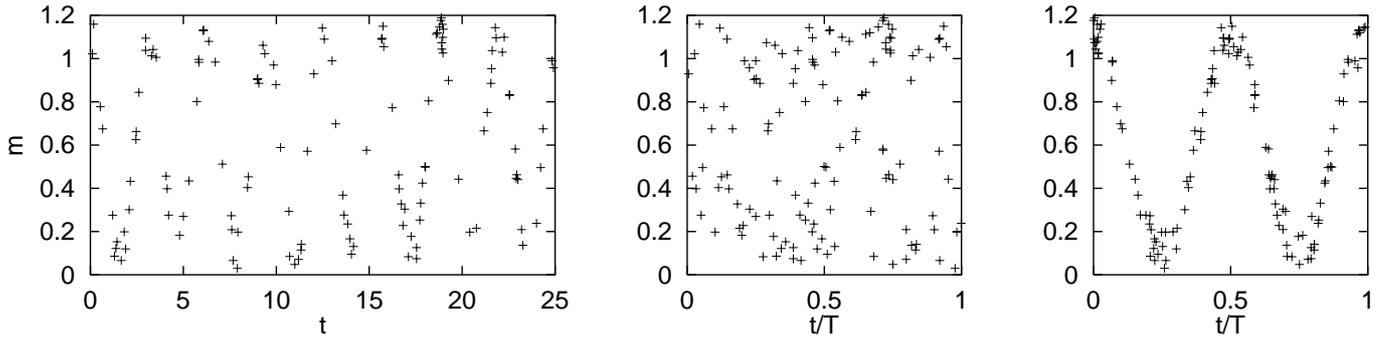


Fig. 1. Observed magnitudes of a variable star (left panel), folded using two trial periods, and the resulting phase diagrams – for a period of 4 units (middle panel) and for a period of 2π units (right panel).

2.1. The idea

As hints of regularity are seen in the correlation function and in the power spectrum, the first idea would be to use these statistics, but in their full three-dimensional form. For a general homogeneous density field the correlation function $\xi(\mathbf{d})$ is a function of a spatial displacement vector \mathbf{d} and the power spectrum $P(\mathbf{k})$ depends on the three-dimensional wave vector \mathbf{k} . The advantage of these statistics is that they describe directly the periodicity in the data. The disadvantages are larger, however. Firstly, these statistics are extremely noisy for real data samples, because we need much more individual amplitudes to populate a three-dimensional region than a line segment, as we have accustomed to do for isotropic $\xi(d)$ and $P(k)$. Secondly, even simple real space patterns give rise to a number of interconnected amplitudes in the \mathbf{d} and \mathbf{k} space; they practically transform one 3-D distribution to another that has to be analyzed again. So, it will be easier (and more transparent) to study the spatial distribution directly.

For this purpose we shall use similar methods as are employed to find periods in observed time series, say, the brightness of a variable star in different moments. The most intuitive method is to use a trial period to fold the series into a “phase diagram”, and to find the value of a statistic describing the reliability of that trial period. An illustration of that is shown in Fig. 1.

The data in the phase diagram (middle and right panels of Fig. 1) are usually binned, and mainly two methods are used to find the best period. The first one, called “phase dispersion minimization” (PDM), proposed by Stellingwerf (1978), does that by minimizing the sum of variances of data in individual phase bins (finding the narrowest light curve). Another method, called “epoch folding” (Leahy et al. 1983), maximizes the variance of the means of the phase bins (looking for a maximum amplitude of the light curve). Though surprising at the first glance, these two methods were shown to be equivalent by Schwarzenberg-Czerny (1989) and Davies (1990).

Phase diagrams already have been applied to the study of regularity in the galaxy distribution. Dekel et al. (1992) used phase diagrams to estimate the confidence levels of the regular signal, which was found in the pencil-beam surveys by Broadhurst et al. (1990). They selected as their statistics the maximum phase amplitudes and the asymmetry of the phase distribution. These statistics are more noisy and their sampling

distributions can be found only by simulations. In the contrary, the PDM and epoch folding, which are based on analysis of variance, allow to derive exact sampling distributions.

If we search for a cubic regularity, we have to fold our cluster distribution into a phase cube, as shown in Fig. 2. However, we cannot use directly the methods described above, as we do not observe densities at certain points in space, but discrete objects. However, we could, in principle, calculate a density for any point. As suggested by Dr. J. Pelt, the closest analog to the analysis of variance in our case will be a search for the maximum of the total variance of density in the phase cube. By the way, in the case of time series this variance (the total variance of data) is fixed and does not depend on the trial period.

So, the statistic we shall use to estimate the cubic regularity of the point distribution can be written as

$$\kappa(d) = \frac{1}{\bar{n}_p^2} \int_{C_p} n_p^2(\mathbf{x}_p) d\mathbf{x}_p, \quad (1)$$

where d is the trial period (cube size), C_p is the phase cube, \mathbf{x}_p are the phase coordinates, obtained from the data coordinates \mathbf{x} by the folding and scaling transformation

$$\mathbf{x}_p = \mathbf{x}/d - \lfloor \mathbf{x}/d \rfloor, \quad (2)$$

n_p is the number density in the phase cube and \bar{n}_p is the mean number density of the whole phase cube. The floor function $\lfloor \cdot \rfloor$ is defined to be the largest integer smaller or equal to x . As the phase coordinates (2) lie in the unit interval, the volume of the phase cube $V_p = 1$, and the regularity $\kappa = 1$ for a constant density distribution. Thus there is no regularity in the case $\kappa = 1$. If there is a regular cubic signal in the point distribution, it is magnified by folding and by squaring its total amplitude. This makes the statistic extremely sensitive; we shall estimate its sensitivity below.

This statistic is also invariant with respect to translations in data space – the phases of the regular signal will change, but not its amplitude distribution. This is strictly true, however, only if we can neglect edge effects, i.e. if we were able to cut the data cube from an infinite volume.

For simple distributions this statistic can be found analytically. If we have a constant density data cube of a fixed size, the regularity signal will come from edge effects, since the partial cells will create a nonuniform phase cube density distribution. It is a bit tedious to count the different contributions, but it

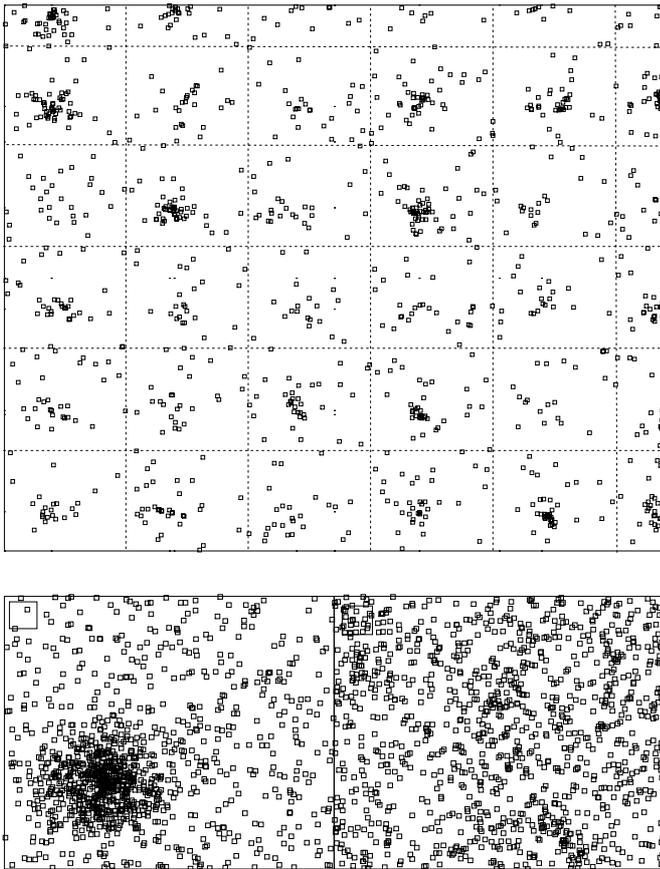


Fig. 2. 2-D cubic folding. The sample volume (upper panel) is divided into smaller trial squares (dotted lines) and all trial squares are stacked together (lower panels). If the size of the trial squares is close to the real period, there will be a clear density enhancement (lower left panel) in the stacked distribution. Otherwise the stacked distribution is almost random (lower right panel). A kernel window is shown in the upper left corner of the stacked cell.

can be done. We give the formula in the Appendix (it is useful for checking the program, should the reader want to write one) and illustrate the result in Fig. 3. Such a figure, the dependence of a statistic on a trial period, is called a periodogram; in our case the regularity periodogram. The regularity according to the definition above is shown by a solid line; the lower dashed line shows a slightly modified statistic where the points that are nearby in real space are ignored when calculating densities. We shall argue below that the latter version works better.

As we see, for a constant density cube the regularity κ is unity for the periods where there are no edge effects (the data cube size is an integer multiple of the period). As the period grows, the number of foldings decreases and the relative strength of edge effects increases. A regularity amplitude of 1.1 is already fairly large.

2.2. The estimator

In order to get an estimator for the regularity κ we have to be able to calculate the density of points in the phase cube. Noting

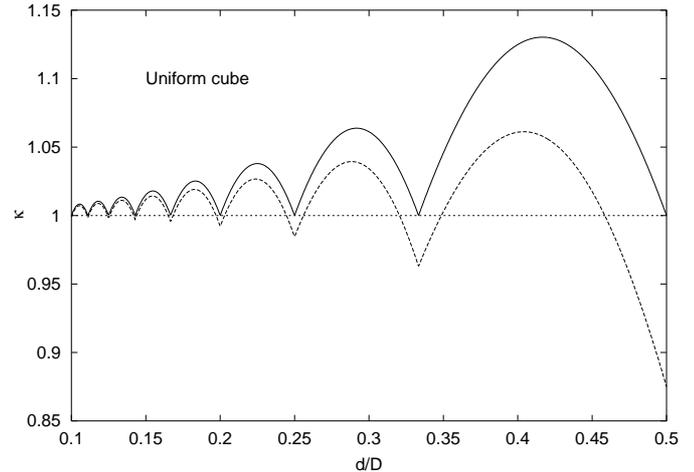


Fig. 3. The regularity periodogram for a constant density cube of size D (d is the test cube size, the period). The solid line is for the full density statistic, the dashed line – for the version without local correlations (see Sect. 2.4).

that the integral in the statistic can be written as

$$\int_{C_p} n_p^2 dx_p = \int_{C_p} n_p dN,$$

where N is the total number of points, and that

$$\bar{n}_p = \bar{n} \frac{V}{V_p} = N,$$

(\bar{n} is the mean density of the sample, V – its total volume, $V_p = 1$ the phase cube volume), a natural estimator can be written as

$$\hat{\kappa}(d) = \frac{1}{N^2} \sum_i \sum_j K(\mathbf{x}_i, \mathbf{x}_j), \quad (3)$$

where N is the total number of points in the sample and \mathbf{x}_i are their phase coordinates. The kernel function $K(\mathbf{x}_i, \mathbf{x}_j)$ is used to estimate the phase cube density near data points:

$$n_d(i) = \sum_j K(\mathbf{x}_i, \mathbf{x}_j). \quad (4)$$

Any kernel with compact support can be used here. We used the simplest one

$$K(\mathbf{x}_i, \mathbf{x}_j) = \frac{1}{V_\varepsilon} \begin{cases} 1, & \mathbf{x}_j \in C(\mathbf{x}_i, \varepsilon), \\ 0, & \text{otherwise.} \end{cases}$$

Here $C(\mathbf{x}_i, \varepsilon)$ is a cube of size ε , centered on \mathbf{x}_i and with volume $V_\varepsilon = \varepsilon^3$. This definition is symmetric in i, j , although at first glance it does not look as though.

As the phase distribution is periodic by definition, any kernel we use has also to take into account that periodicity, when calculating densities for points near the boundaries of the phase cube. In practice we do this by creating a padding for the phase cube that extends from $-\varepsilon/2$ to $d + \varepsilon/2$, and letting the index j in Eq. (3) to run over the padded phase cube. This speeds up calculations considerably, since we have to create this padding once for every trial period, without that

the periodicity conditions should have been checked for every point i in Eq. (4). The program to calculate the periodogram is written in C and it can be obtained by anonymous ftp from `ftp.aai.ee:/pub/saar/regularity.tar.gz`.

2.3. Error estimates

The probability distribution for the estimator (3) is easy to derive for Poisson-distributed data. As we use the regularity periodogram to search for small regular signals, this case will serve as a good point of reference.

Although we calculate the phase cube density by means of a kernel estimator, this is statistically equivalent to binning the phase cube into ν bins, where $\nu = 1/V_\varepsilon$. In this binning approximation we can write

$$\hat{k} = \frac{1}{N^2\nu} \sum_{i=1}^{\nu} (n_p)_i^2 = \frac{\nu}{N^2} \sum_{i=1}^{\nu} N_i^2,$$

where N_i are the occupation numbers of the phase bins and we have used the fact that $\Delta V_p = V_\varepsilon = 1/\nu$.

In practice, the occupation numbers N_i should be at least a few tens, to ensure proper estimates of the phase cube density. Thus, we can approximate the Poisson distribution of intensity $\lambda = N/\nu$ by the Gaussian one with the mean and variance λ .

Let us now define a sum

$$X_\nu = \sum_{i=1}^{\nu} \left(\frac{N_i - \lambda}{\sqrt{\lambda}} \right)^2. \quad (5)$$

In the Gaussian approximation made above the quantity X_ν obeys a χ_ν^2 distribution.

Expanding the sum in (5) and using the expression for λ , we get

$$\hat{k} = \frac{1}{N} X_\nu + 1. \quad (6)$$

This allows us to estimate the mean and variance of \hat{k} as

$$E[\hat{k}] = \frac{1}{N} E[X_\nu] + 1 = 1 + \frac{\nu}{N}$$

and

$$\text{Var}[\hat{k}] = \frac{1}{N^2} \text{Var}[X_\nu] = \frac{2\nu}{N^2}. \quad (7)$$

As we see, the estimate is effective, but biased. The bias may be removed easily, using $N_i(N_i - 1)$ in the sum instead of $(N_i)^2$. Then

$$E[\hat{k}] = \frac{1}{N} E[X_\nu] + 1 - \frac{\nu}{N} = 1.$$

The expression for the variance (7) shows that it depends only on the total number of data points and the number of bins used; the variance is uniform over all the period range. This is a very useful property. We tested the above formulae by calculating the regularity periodogram for a large number of simulations of Poisson cubes.

Now we are able to estimate the sensitivity of the estimator. Let us suppose that there is a small cubic signal, consisting of N_s additional points each in μ bins. Expanding the sum of

squares of bin population numbers for this case and taking into account the change in the normalization factor, we can write for the estimator

$$\hat{k} = 1 + \frac{\nu}{\mu} \left(\frac{S}{N} \right)^2, \quad (8)$$

where $S = \mu N_s$ is the total number of points in the signal, and $S \ll N$. Hence the amplitude of the signal is proportional to the square of the ‘‘signal-to-noise’’ ratio S/N and is inversely proportional to the ‘‘filling factor’’ of the structure μ/ν . Cubic arrangements that fill only grid vertices are easier to detect than those populating edges or faces of the grid.

Thus we are able to detect a signal, if

$$\frac{\nu}{\mu} \left(\frac{S}{N} \right)^2 \geq \gamma \sigma(\hat{k}) = \gamma \frac{\sqrt{2\nu}}{N},$$

where $\sigma(\hat{k}) = \sqrt{\text{Var}[\hat{k}]}$ is the rms error of \hat{k} , and γ is a coefficient that defines the confidence level required. We can write this as

$$S^2 \geq \gamma \mu N \sqrt{\frac{2}{\nu}}. \quad (9)$$

The regularity amplitude for another extreme case, when we are observing a purely cubic structure, is also easy to derive. In this case the amplitude depends only on the ‘‘filling factor’’,

$$\hat{k} = \frac{\nu}{\mu}.$$

This formula differs from the linear approximation above (8) only by the lack of the additive constant 1.

As an example, we take a typical application to galaxy clusters with $N \approx 1500$, $\nu = N/30 \approx 50$, and choose $\gamma = 3$. If we suppose that the number of signal bins is around 10, then a cubic signal with a total number of $S \geq 100$ clusters should be detectable already. If we know the amplitude of a periodogram and can estimate the filling factor of the structure, then we can use this formula also to estimate the number of points in the cubic arrangement.

An illustration of the sensitivity of the regularity statistic is shown in Fig. 4, where a tiny regular signal of a total of 80 points is mixed into a Poisson sample of 4000 points ($S/N = 0.02$). These points are distributed near a grid with a spacing of 100, while the total size of the cube is 700. We show also the range spanned by the regularity periodograms of 100 pure Poisson realizations of the data. We see that such a small signal is readily extracted from data, there is a peak with confidence higher than 99% at the period of 100. The filling factor of the cubic structure μ/ν is rather small for this arrangement, of course. A clear period will also generate both harmonics and sub-harmonics of smaller amplitude. The $d = 50$ harmonic is well seen in the Figure, while sub-harmonics are hidden by edge effects.

We also show a slice of the data cube in this figure. Note that the regular signal can not be traced there by eye.

There are two possible strategies to determine the size of the bin (the width of the density kernel). The first is to keep

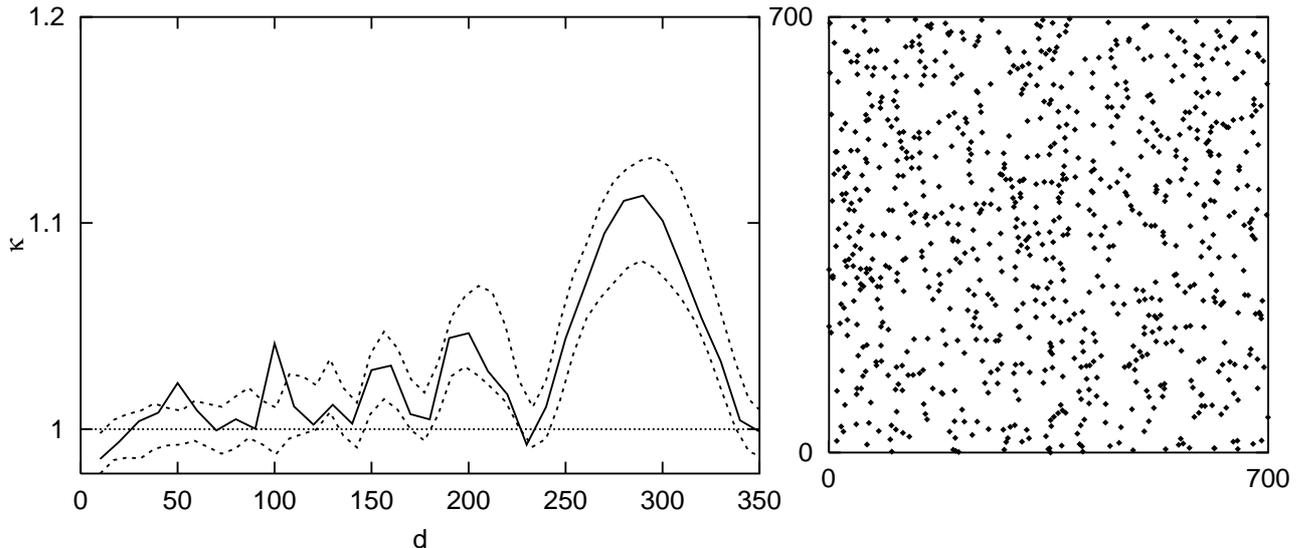


Fig. 4. A regularity periodogram (left panel) for a Poisson cube with a 2% mixture of points (80 versus 4000) forming a regular distribution near a grid with a spacing of 100 (the cube size is 700). This periodogram is shown by a solid line, the two dotted lines show the range spanned by 100 regularity histograms for pure Poisson data. The right panel shows a slice of a thickness of 150 of the data cube.

this width ε constant in data space. In that case the effective number of bins $\nu = (\varepsilon/d)^3$ (d is the period) and

$$\sigma(\hat{\kappa}) = \frac{\sqrt{2}}{N} \left(\frac{l}{d} \right)^{3/2}$$

will depend strongly on the period. Also, a particular choice of ε will restrict the usable period range. For periods only a few times larger than ε the density smoothing would be excessive, erasing most of the signal. For larger periods, when the number of points per bin $N_B = N(\varepsilon/d)^3$ becomes too small, the estimates may become very noisy.

Another strategy is to keep the phase resolution constant for all periods (the same number of bins, ν). In this case we get similar smoothing for all periods, i.e. the mean number of points per bin N_B remains the same. As $\nu = N/N_B$, the rms error can be written as

$$\sigma(\hat{\kappa}) = \frac{2}{NN_B};$$

the scatter of the estimate remains constant for all periods.

We have used both strategies, starting with that of the constant kernel width in data space, but we realized later that the second strategy is much better.

2.4. Effect of local correlations

The error analysis above was made for Poisson data samples. In real applications, our samples have substantial short-range correlations. This will increase both the amplitude and the variance of the estimate, and the clustered nature of the data could generate false regularity signals.

In order to demonstrate this we generated a segment Cox process with the same correlation properties as our cluster sample. The use of segment Cox processes in correlation analysis has been advocated by Pons-Bordería et al. (1999). A segment Cox process places random segments of length l in space and

then generates Poisson-distributed points along these segments. An important property of a segment Cox process is that while its long-range distribution is strictly homogeneous ($\xi(r) = 0$, if $r \geq l$), its short-range correlation function $\xi(r) \sim r^{-2}$ is similar to the correlations we find in observations.

We describe segment Cox processes and discuss how to specify the parameters of the process in the appendix. Here we list those for our sample: the size of the cube is 700, the segment length $l = 60$, the density of segments $\lambda_S = 1.65 \times 10^{-6}$ (the mean number of segments for the sample cube is 566), the line density of points $\lambda_L = 0.12$ (the mean number of points per segment is 7). The total number of points in the sample is 2057. As this is only an example, we use dimensionless units here, but if we would multiply these by $1 h^{-1} \text{ Mpc}$, all these sizes would be comparable to those of cluster samples.

We show the correlation function of our simulated segment Cox sample and plot a slice of the sample cube in Fig. 5. The spatial distribution of points looks slightly peculiar (like a segment Cox process should), but the correlation function that we show is close to that of rich clusters of galaxies. We plot the function $\xi(r) = (r/20)^{-2}$ there as an example.

Now that we have our simulated Cox sample we can calculate its regularity periodogram. In order to avoid edge effects we choose the periods this time by dividing the size of the data cube by successive integers, so that all cells fit exactly in the cube. This would not be a good recipe for observations, as we lose a lot of interesting periods this way, but here we know that there are no real periods in the data.

We show this periodogram in Fig. 6 with a dashed line. As expected, the estimate is biased, exceeding the expected value of unity for all periods, and the bias grows with period. The latter effect is also easy to understand, as for larger periods the number of stacked cells is smaller and the influence of clusters in individual cells is larger.

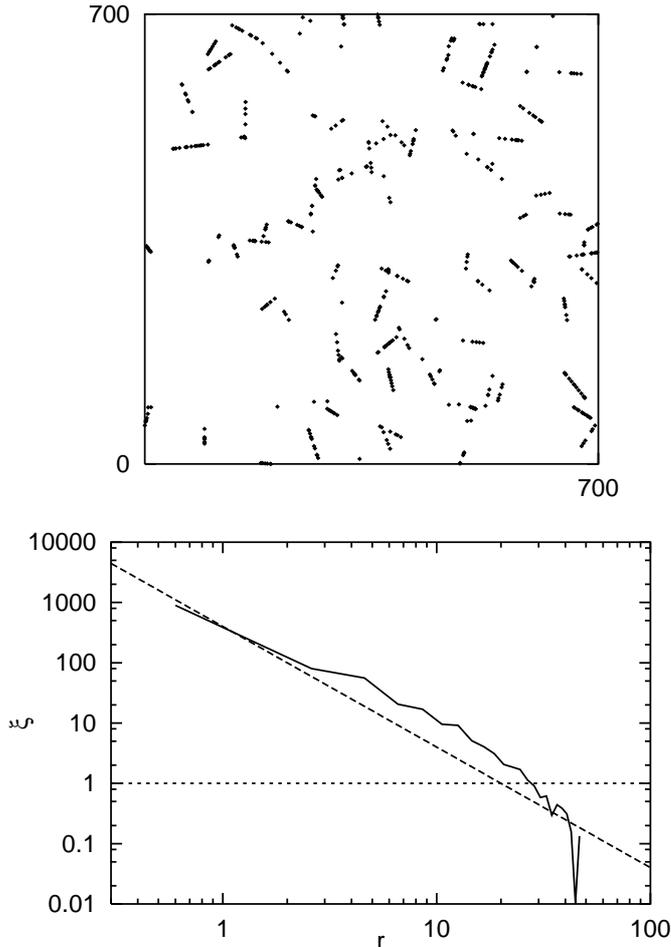


Fig. 5. The correlation function (lower panel) for a segment Cox process simulating the spatial distribution of rich clusters of galaxies. The straight line shows the approximation $\xi(r) = (r/20)^{-2}$. The upper panel shows a slice of the simulation cube with a thickness of $1/7$ of the cube size.

We could try to correct for this effect, determining the short-range correlation in advance, and comparing our periodograms with those built for Cox samples. This is not easy, however; while one can usually construct a segment Cox process that describes the two-point correlations of a sample, Cox samples often look rather different from observational ones. Hence one could never be sure if the comparison is a fair one.

Fortunately, there is an easier way to handle this, and again it is a trick known to the periodicity community. As Dr. J. Pelt suggested, we could modify our estimator, discarding any local contributions to the phase cube density. In other words, when calculating the sum in formula (4), we do not count these data points that lie in the same real space cell (of size d) as the point i itself.

This effectively eliminates local correlations and leaves only long-range correlations we are interested in. We show this periodogram as the solid line in Fig. 6. This line wiggles nicely around the expected value in the range predicted by the estimate of the variance, and it does not “feel” short-range correlations at all.

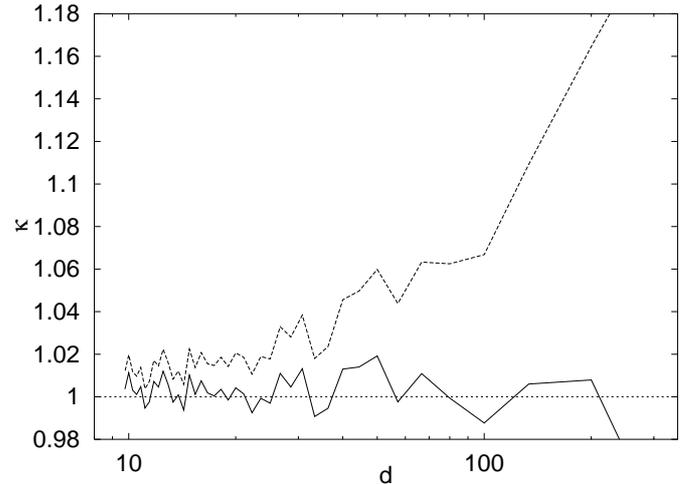


Fig. 6. Regularity histograms for a segment Cox process simulating short-range correlations. The dotted line shows the periodogram as defined above by formula (3), the solid line – for the estimator that discards local correlations.

We have used this estimator in all the rest of the paper. We have to note that in this case the normalization is strictly correct for very short periods only. The sum in the estimator (3) has to be changed to

$$\sum_v n_i^2 \rightarrow \sum_v n_i n'_i,$$

where

$$n'_i = n_i \left(1 - (d/L)^3\right)$$

(here d is the period and L is the size of the sample). Thus the expected value of $\hat{\kappa}$ will change to

$$E[\hat{\kappa}] = 1 - (d/L)^3 = f(d/L).$$

However, as we have to use the reduced estimator for observed samples, as explained below, we do not have to worry about normalization.

The regularity periodogram for this estimator for a constant density cube can also be found analytically; the formula is given in the appendix and demonstrated in Fig. 3.

2.5. Border and selection effects

It is good, of course, that our estimator is sensitive to extremely small cubic signals. However, the same sensitivity makes it also prone for edge effects as we see in Fig. 4; incomplete boundary cells may hide the signal we are seeking for.

Thus, first of all we have to account for edge effects very carefully. The case illustrated in Fig. 4 is a little extreme, as the incomplete cells form a clean cubic signal, but it shows clearly the possible amplitude of edge effects. One possibility would be to use a minus-estimator, discarding all those cells that intersect the boundary of the sample. This is a mathematically elegant possibility. However, it is of little use in practice, since boundaries of real samples are usually so complex that we would have to discard most of the data.

To show the effect of selection, we generated a spherical sample of the same linear dimensions, with a diameter of 700 and the same number of points, 2000. The density of this sample falls linearly with the radius, going to zero at the boundary of the sphere. This mimics observational samples rather well. Ten periodograms for this sample are shown in Fig. 7. The main effect is a radical change in the shape of the periodogram. Secondly, the boundary-generated oscillations are gone, as the boundaries are fuzzy. And thirdly, the scatter of the curves remains constant. This is expected, as both the total number of points (2000) and the number of points per density bin (30) are the same.

Selection effects could be accounted for by introducing weights for data points that are inversely proportional to the normalized density of the sample (the selection function) at that point \bar{n}_i , defining:

$$\hat{\kappa}'(d) = \frac{1}{N(N-1)} \sum_i^N \sum_j^N \frac{w_i w_j}{\bar{n}_i \bar{n}_j} K(\mathbf{x}_i, \mathbf{x}_j),$$

where the additional weights w_i can be chosen to minimize the sampling variance of $\hat{\kappa}'(d)$. As these weights have to be calculated only once, it also would speed up the calculation of periodograms. The trouble with this approach is that the choice of the weights is rather difficult. This approach also mimics a volume-limited constant mean density sample, leading to abrupt boundaries and thus amplifying edge effects.

An easier approach is to build reference samples using the selection functions determined for data samples, and to account for edge effects by comparing the real sample and a reference sample with the same geometry, using a reduced regularity periodogram k :

$$k(d) = \frac{\hat{\kappa}(d)_{\text{obs}}}{\hat{\kappa}(d)_{\text{ref}}}, \quad (10)$$

where the subscripts “obs” and “ref” denote the observed data and simulated reference samples, respectively. In order not to introduce discreteness noise by this operation, the number of points in the reference sample should be larger than in the data sample, but not too much (the variance of the estimate is proportional to N^{-2}). We have used in this paper reference samples with ten times more points than data samples have. We shall use this reduced periodogram throughout the rest of the paper, calling it simply the regularity periodogram. The reduced regularity $k(d)$ is an unbiased estimator of the regularity $\kappa(d)$, but its variance will grow as $\hat{\kappa}^{-2}(d)_{\text{ref}}$. This limits the range of the usable periods.

The reference periodogram will take account of both the edge and selection effects. Interestingly, the reduced estimator (10) might be the only one used in astronomy that actually gains from selection effects – the fact that the density of data points gets smaller at the sample boundaries helps to reduce significantly the amplitude of edge effects.

Apart from the spurious signal the boundaries create, they have another effect – they destroy the translational invariance of the statistic. This can be most easily seen when folding a one-dimensional analog of a data cube, a constant density line

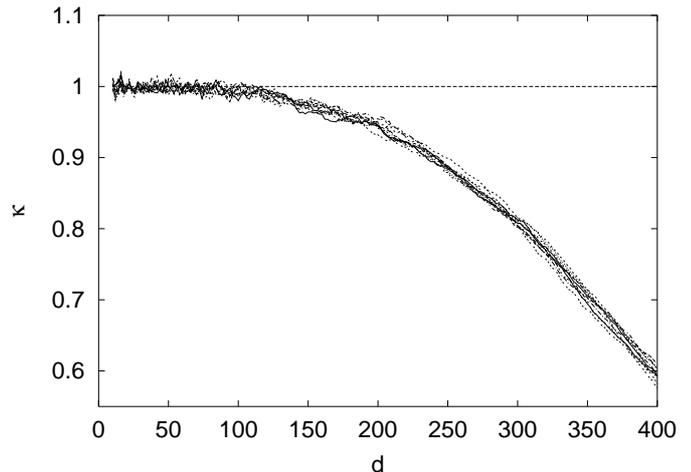


Fig. 7. Regularity periodograms for ten realizations of a sphere of diameter 700, with 2000 particles generated to represent a density falling linearly with radius; the limiting density at the surface of the sphere is zero. Notice the absence of edge effects.

segment. This means that we should create our reference samples very carefully, as any small shift in the shape between the reference sample and the observed sample would give a spurious signal. Fortunately, fuzzy boundaries of observed samples, due to selection effects, practically eliminate this problem.

We have tested this by shifting the reference samples generated to use with the observed galaxy cluster samples. The geometry of these samples is rather complicated and slight shifts between the observed and simulated shapes could easily happen. The test showed that shifts up to $20 h^{-1}$ Mpc in every coordinate direction (the samples consist of two cones with characteristic sizes of $350 h^{-1}$ Mpc) practically do not change regularity periodograms.

2.6. Voronoi samples

The regularity statistic is a typical universal frequentist statistic that can be applied to all point configurations. If we were using the Bayesian approach, we also should have specified a statistical model. However, this universality has also a drawback – without a statistical model we do not know if there exist non-cubic arrangements which also will produce a cubic signal. Thus, in order to understand the proposed statistic, we have to test it on various spatial arrangements.

The first class of arrangements we use are the Voronoi models of large-scale structure, introduced by van de Weygaert & Icke (1989) and van de Weygaert (1991).

These models are based on the Voronoi tessellation of space. In a Poisson–Voronoi point process the centers of voids serve as seeds of structure, and have a Poisson distribution, while the points (clusters of galaxies in our case) are located at the vertices of the structure formed by the expanding voids. Although the model starts from a Poisson distribution, it is well clustered and has a power-law correlation function $\xi \sim r^{-2}$, close to that of the galaxy clusters. This amazing fact was discovered by van de Weygaert & Icke (1989) and verified exactly by Heinrich et al. (1998). In this model clusters of galaxies

form superclusters, which together with voids form a cellular supercluster-void network, similar to that observed. The only free parameter of the model is the mean density of void centers that determines the mean diameter of voids.

We used a program by R. van de Weygaert to generate the cluster sample. The size of the simulation box was $700 h^{-1}$ Mpc; the number of seeds was 431, chosen to obtain a mean diameter of voids $115 h^{-1}$ Mpc in accordance with the distribution of Abell clusters (Paper I). We generated ten realizations of the model; the number of clusters in the samples varied between 2884 and 2948 (for details see Einasto et al. 1997c, hereafter Paper III).

To characterize the regularity of a model we calculated the reduced regularity periodograms $k(d)$ and the correlation functions $\xi(r)$ of clusters. The results for ten realizations are shown in Fig. 8. The correlation function has a deep minimum around $80 h^{-1}$ Mpc and a secondary maximum at $140 h^{-1}$ Mpc; on larger scales it flattens out to a mean value around zero. The geometric interpretation of this behaviour was discussed in Paper III: the minimum corresponds to the mean distance between superclusters and surrounding voids; the secondary maximum can be identified with the mean distance between superclusters across voids. In a Voronoi model the mean size of voids is well fixed, thus the presence of a secondary maximum is expected. On still larger scales the behavior of the correlation function depends on the regularity of the distribution of rich superclusters. In a Voronoi model void centers as well as rich superclusters are randomly distributed, thus we expect no correlation on very large scales. As we see from Fig. 8, on very large scales the correlation function is indeed close to zero which corresponds to a Poisson distribution. In contrast to the correlation function the features in the regularity periodograms $k(d)$ have a rather low amplitude (there are suspicious signals at $190 h^{-1}$ Mpc and $220 h^{-1}$ Mpc), but overall the periodograms are featureless, resembling the regularity periodograms of pure Poisson samples. Although the Voronoi model is cellular, the cells are not cubic or regular.

2.7. Quasi-regular models

Next we shall choose a quasi-regular model distribution. We use mixed models, i.e. samples with two populations of clusters. The clusters of one population are randomly distributed, while in the second population they are located in superclusters which form a regular rectangular network with a period of $130 h^{-1}$ Mpc. In our model superclusters are randomly located along rods which form a regular rectangular network with a period of $130 h^{-1}$ Mpc. The positions of the rods have been randomly shifted with the rms deviation of $\pm 20 h^{-1}$ Mpc (for details see Paper III).

The box size was taken as $690 h^{-1}$ Mpc, the period was set to $115 h^{-1}$ Mpc, the number of randomly located poor superclusters was chosen to be 3000, and the minimal number of rich superclusters on rods was 4. We generated ten realizations of the model; the total number of clusters varies between 9218 and 9519, and the total number of clusters in quasi-regular population in rods varied between 3328 and 4216.

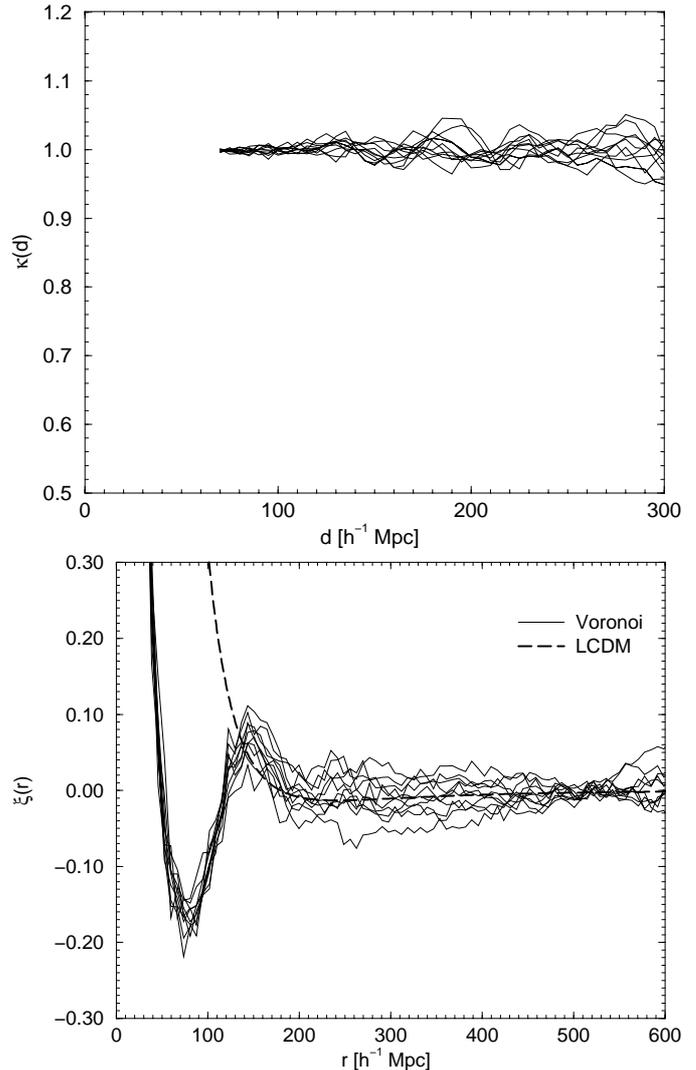


Fig. 8. Regularity periodograms (upper panel) and correlation functions (lower panel) for ten realizations of Voronoi models.

The regularity periodograms and the correlation functions are shown in Fig. 9. The correlation functions are oscillating with a period of $115 h^{-1}$ Mpc; the amplitude of oscillations decreases very slowly with distance r . The regularity periodograms have two well-pronounced maxima at $d = 115 h^{-1}$ Mpc and $d = 230 h^{-1}$ Mpc. The scatter of the regularity periodograms is very small.

2.8. Angular sensitivity

Obviously, the method is sensitive to the direction of the axes of the trial cubes. The orientation of a cube can be described by three Euler angles, so there is a considerable amount of freedom here. If the trial cubes are rotated 45° along one axis of the cubic alignment, there will be a mixture of cubic signals – one with the previous period from the direction perpendicular to that face, and another with the period of $\sqrt{2}$ times larger. If the trial cubes are oriented along the long diagonals, from a vertex to the opposing vertex of the original cube, the cubic network will be cubic again with a single period of $\sqrt{3}$ times

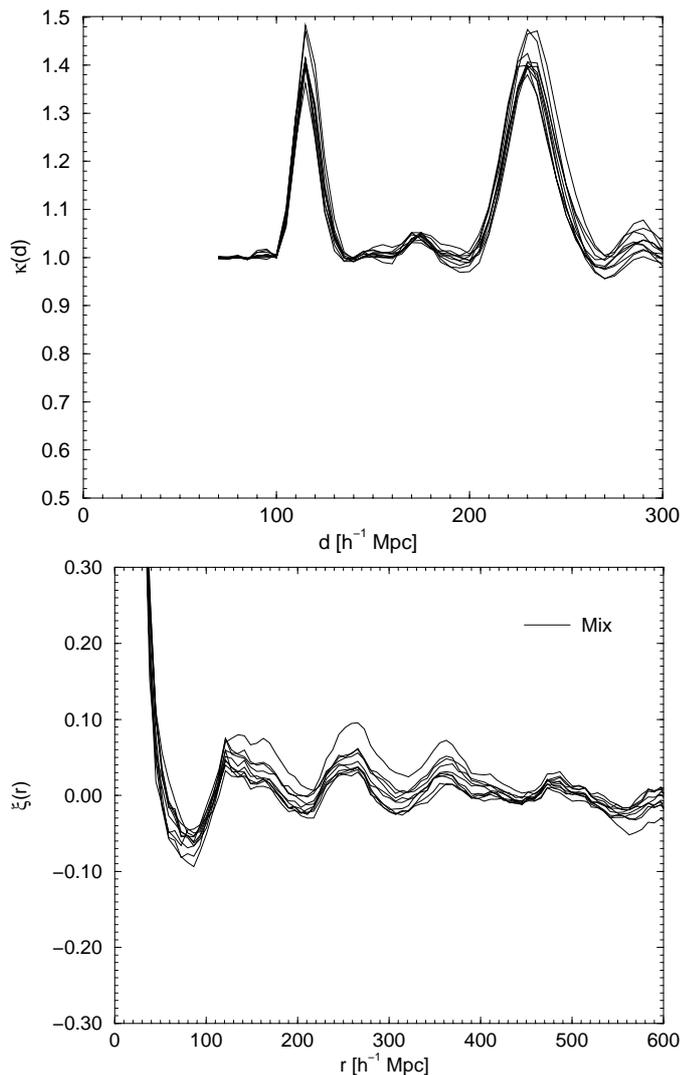


Fig. 9. Regularity periodograms (upper panel) and correlation functions (lower panel) for ten realizations of quasi-regular line models.

the original one. For other orientations the signal will be weak or absent. This property is illustrated in Fig. 10. The reason for this behaviour is clear: the value of $k(d)$ depends on the number of clusters which coincide in the stacked cell. If clusters located in superclusters of different original cells lie in the stacked cell in different locations (which happens when orientations differ considerably), then $k(d)$ cannot exceed by much the expected value of unity for random samples.

2.9. *N*-body cluster samples

We have compared the regularity periodograms and correlation functions also for several cluster samples found in numerical simulations of the large-scale structure of the Universe. We used four models analyzed by E99. These models were calculated using a PM code with 128^3 particles and a 256^3 mesh in a cube of size $720 h^{-1}$ Mpc. The models include the standard CDM model with the density of matter $\Omega_M = 1$ (model CDM61), and a LCDM model with the vacuum energy term, $\Omega_M = 0.2$ and $\Omega_\Lambda = 0.8$ (model CDM62). To obtain

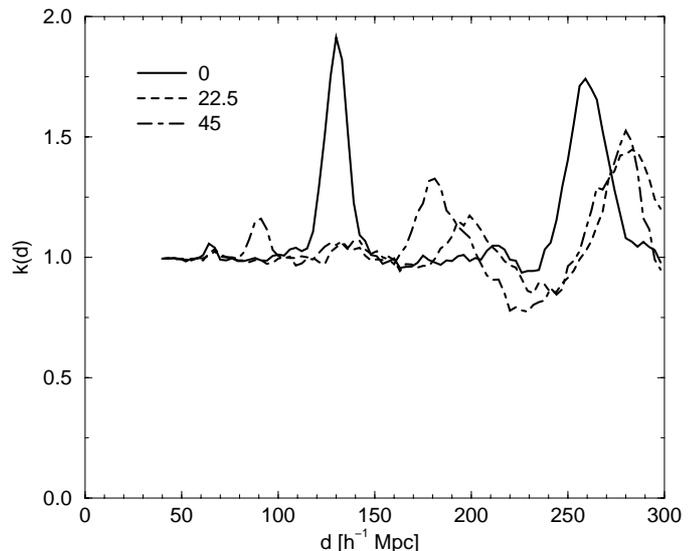


Fig. 10. Angular sensitivity of regularity periodograms (for a quasi-regular model). The solid, dashed and dot-dashed lines are for trial cubes oriented at $\theta = 0^\circ$, 22.5° and 45° with respect to the grid.

a model which better represents the observed power spectrum (see E99) we used two initial power-law spectra with the indices $n = 1$ on large scales, $n = -1.5$ on small scales, and a sharp transition on a scale of $k = 0.05 h \text{ Mpc}^{-1}$ (models DPS6 and DPS3, the latter with 256^3 particles). The last model we used (DPP6) has an additional high-amplitude peak in the power spectrum near the maximum. In most of the models clusters of galaxies were selected using the friends-of-friends algorithm. Only in the high-resolution model DPS3, computed by Gramann & Suhhonenko (1999), maxima of the smoothed density field were used for cluster identification.

Regularity periodograms and correlation functions for simulated cluster samples are shown in Fig. 11. These functions have been calculated for the whole cluster sample in the simulation box. The correlation functions of these model samples oscillate slightly, except for the model DPP6 with a high peak in the power spectrum that causes high-amplitude oscillations. The regularity periodograms of all models are practically featureless, even that for the model DPP6; small peaks and valleys are due to shot noise. This is understandable, as the initial density fluctuations are, by definition, isotropic and cannot develop into a cubic structure.

Thus, application of the regularity statistic to these three different, but typical classes of spatial arrangement of objects in cosmology, shows that it well detects the cubic signal, if it is present, and does not “feel” other regularities.

3. Regularity of the Abell cluster sample

3.1. Observational data

We shall use in this study the 1999 version of the compilation by Andernach & Tago (1998) of all published galaxy redshifts towards galaxy clusters in the catalogue of rich clusters of galaxies by Abell (1958) and Abell et al. (1989) (hereafter Abell clusters). The 1999 version of the compilation will be

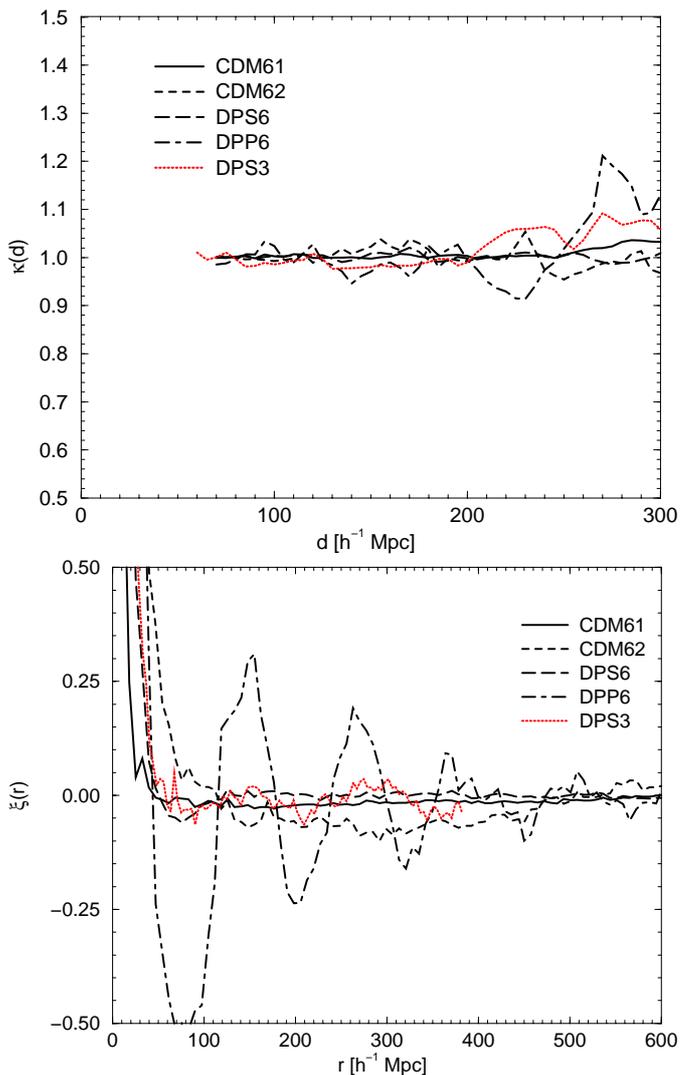


Fig. 11. Regularity periodograms (upper panel) and correlation functions (lower panel) for simulated CDM and double-power-law models.

described in detail elsewhere (Einasto et al. 2001). Individual galaxies were associated with a given Abell cluster if they lay within a projected distance of $\leq 1.5 h^{-1}$ Mpc (one Abell radius) and within a factor of two of the redshift estimated from the brightness of the clusters 10th brightest galaxy, using the photometric estimate of Peacock & West (1992). For the present analysis we used a sample of all rich clusters (richness class $R \geq 0$ and excluding clusters from ACO's supplementary list of S-clusters) in this compilation with redshifts up to $z = 0.13$. The sample contains 1665 clusters, 1071 of which have measured redshifts for at least two galaxies. Distances of clusters without measured redshifts and of clusters with only one galaxy measured have been estimated on the basis of the apparent magnitude of the 10th brightest galaxy of the cluster. The spatial distribution (in supergalactic coordinates) of the sample is shown in Fig. 12. The Galactic zone of avoidance breaks the distribution into two cones. The slight inclination of the cones with respect to the coordinate axes is due to the 6.3° angle between the supergalactic Y axis and the Galactic North Pole.

As representatives of high-density regions in the Universe we used *rich* superclusters from the list of superclusters presented in Paper I. Superclusters were identified using the friend-of-friends algorithm (used first in studies of large-scale structure by Zeldovich et al. 1982) with a neighbourhood radius of $24 h^{-1}$ Mpc. In this way all clusters of a supercluster have at least one neighbour at a distance not exceeding the neighbourhood radius. To illustrate the distribution of clusters in high-density regions we plot in Fig. 13 only clusters in rich superclusters, while in the quantitative analysis below we shall use both the sample shown in the figure and the full cluster sample. The sheets plotted are $300 h^{-1}$ Mpc thick, thus some superclusters overlap in projection. The sheet in the left panel of Fig. 13 crosses the majority of the cells of the supercluster-void network present in our cluster sample; the sheet in the right panel contains all clusters of the sample in the southern Galactic hemisphere. Figure 13 shows clearly the quasi-regular network of superclusters interspersed with voids. The three-dimensional distribution of all Abell and APM clusters in rich superclusters in the whole volume within a limiting radius $350 h^{-1}$ Mpc around us can be viewed at the home page of Tartu Observatory (<http://www.aai.ee>).

For comparison we plot in Fig. 13 also the distribution of the APM clusters of galaxies in rich superclusters. We see that the APM cluster sample covers a much smaller volume in space which makes it difficult to investigate the regularity of the distribution of high-density regions on large scales. The APM cluster sample is defined only in the southern Galactic hemisphere, and even here the APM sample containing clusters with measured redshifts covers only three high-density regions defined by very rich superclusters with at least 8 cluster members. These are the Sculptor (SC9), Pisces-Cetus (SC10), and Horologium-Reticulum (SC48) superclusters of the catalogue in Paper I. To investigate the regularity of the supercluster-void network, the sample volume must exceed the period of the network at least several times. For this reason we have used only Abell clusters in the following analysis.

Using the friend-of-friends algorithm, we generated from the full catalog (ACO1) catalogs of clusters belonging to superclusters with at least 2 members (ACO2), at least 4 members (ACO4) and at least 8 members (ACO8). As the strongest regularity was observed in the ACO8 sample, we concentrated our studies on that sample and compared it with the full cluster sample. We also studied separately the northern and southern (with respect to Galactic coordinates) half-cones ACO1N, ACO1S, ACO8N and ACO8S.

3.2. Reference samples

In order to eliminate the influence of the sample geometry and of selection effects we have to construct reference samples with the same geometry and selection function as those observed. As the regularity test is sensitive to small cubic signals, considerable care has to be taken when building reference samples.

We started with determining the selection functions for all samples we used in the analysis. As the main sources that determine the observability of a cluster are its distance and Galactic

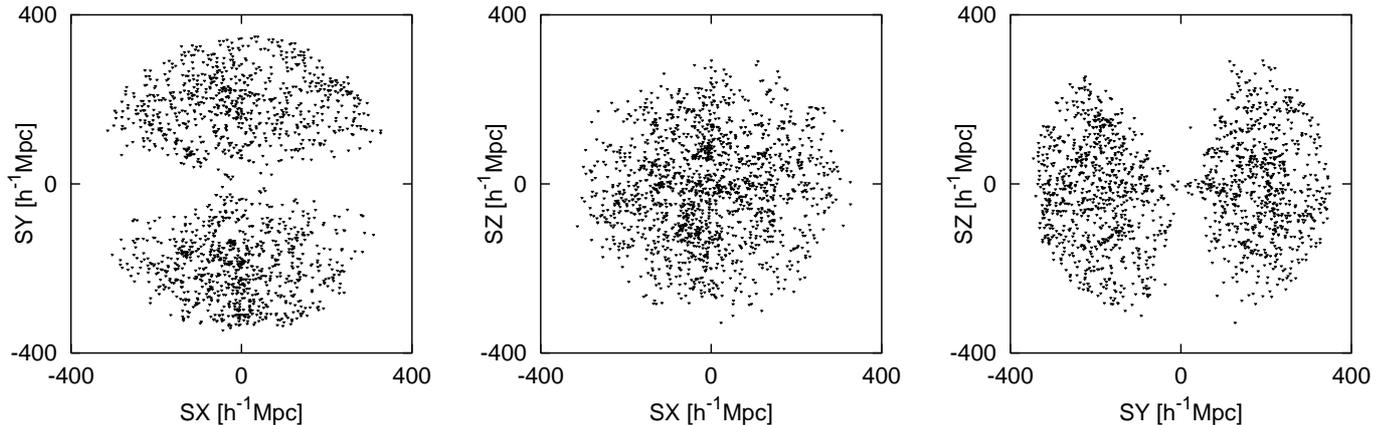


Fig. 12. Projections of the full Abell cluster sample in supergalactic coordinates.

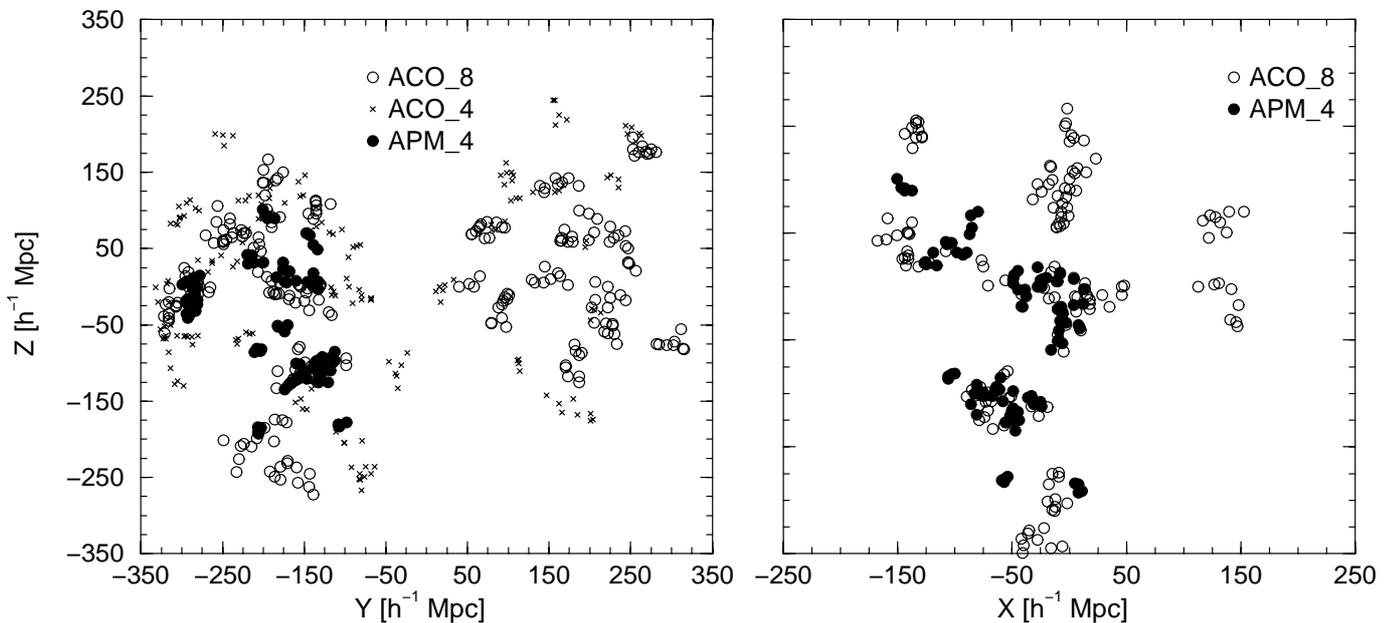


Fig. 13. Distribution of clusters in high-density regions in supergalactic coordinates. The left panel shows the clusters within a sheet, limited by $-100 \leq X \leq 200 h^{-1} \text{ Mpc}$ (in supergalactic coordinates). Abell and APM clusters in superclusters with at least 8 or 4 members are plotted with symbols as indicated. The supergalactic $Y = 0$ plane coincides approximately with the Galactic equatorial plane and marks the Galactic zone of avoidance. In the right panel only clusters in the southern Galactic hemisphere are plotted; here the depth is $-350 \leq Y \leq -50 h^{-1} \text{ Mpc}$.

obscurations that are independent of each other, the selection function $F(r, b)$ (r is the distance and b is the Galactic latitude) can be represented as a product of two selection functions, $F_r(r)$ and $F_b(b)$. We delineate our procedure on the example of the radial selection function $F_r(r)$.

We started by building an integral probability distribution $F_r(r)$ of cluster distances – ordering all r_i (the distance to the i th cluster) by their values and forming a step function with the increment $1/N$ (where N is the total number of points in the sample). We then found a least-squares polynomial fit to this distribution. Checking the form of the approximation and its derivative (which is equal to the spatial density multiplied by r^2), we chose the order of the approximation and the upper and lower boundaries r_{\min} and r_{\max} for the distribution.

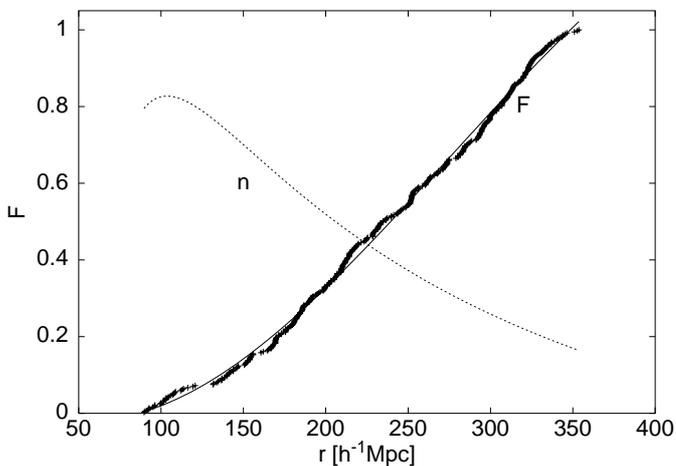
This procedure is illustrated in Fig. 14 for the radial selection function for the ACO8 sample. The number of clusters in that sample is fairly small (373) and binning this number would

not be useful. The radial distribution is rather lumpy, too, as seen from the wiggles in the observed probability distribution, shown with crosses in Fig. 14. The polynomial approximation smoothes these wiggles out nicely. The density distribution obtained from that function is also shown in the Figure.

Will smoothing of the selection functions introduce an artificial regularity signal? It is natural to assume that selection functions should be monotonic functions of their arguments, and the wiggles in the observed probability distribution are caused by real inhomogeneous large-scale structure. A regular cubic alignment will certainly introduce slight irregularities in the radial distribution of clusters. An attempt to exactly mimic these wiggles in the reference sample would introduce a specific undesirable regularity there. Thus we shall use reference samples with smooth radial and angular density distributions.

Table 1. Coefficients of the selection functions (columns a_0, a_1, a_2, a_3, a_4) and the coordinate ranges (d_1, d_2).

$F(x) = a_0 + a_1x + a_2x^2 + a_3x^3 + a_4x^4,$ $F_{\min} = F(d_1), F_{\max} = F(d_2).$								
Data	type	a_0	a_1	a_2	a_3	a_4	d_1	d_2
ACO1	sinb	0.181088	-1.53176	3.59138	-1.53302	0.292309	0.26	1
ACO1N	sinb	0.0760189	-0.860193	2.57105	-0.786879	0	0.19	1
ACO1S	sinb	0.107172	-0.718337	0.75158	1.88259	-1.023	0.28	1
ACO1	r	-0.00981878	-0.000147744	3.70636e-06	1.32363e-08	0	65.2	355.4
ACO1N	r	-0.0820001	0.00135584	-4.48581e-06	2.59908e-08	0	70.2	355.4
ACO1S	r	0.0615636	-0.00172044	1.19554e-05	8.86643e-10	0	72.0	355.4
ACO8	sinb	0.392894	-2.0418	2.6489	0	0	0.4	1
ACO8N	sinb	-0.12838	0.908411	-2.45742	2.67739	0	0.33	1
ACO8S	sinb	-0.5535	0.337642	1.21586	0	0	0.55	1
ACO8	r	0.076307	-0.00274636	2.65024e-05	-3.21016e-08	0	60.0	353.7
ACO8N	r	-0.717997	0.0111047	-3.97325e-05	6.23955e-08	0	89.1	353.7
ACO8S	r	-0.334387	-0.00102197	2.72916e-05	-3.87503e-08	0	151.0	352.7

**Fig. 14.** Determination of the radial selection function for the ACO8 sample with 373 clusters. Crosses show the observed radial distribution. The solid line (marked F) is the fitted selection function found, and the dashed line (marked n) is the spatial number density derived from the selection function, rescaled to fit in the same plot.

The last step is a linear transformation that ensures that the fitted function satisfies strictly the conditions $F_r(r_{\min}) = 0$ and $F_r(r_{\max}) = 1$:

$$F_r(r) \leftarrow \frac{F_r(r) - F_r(r_{\min})}{F_r(r_{\max}) - F_r(r_{\min})}.$$

This changes the function $F_r(r)$ very little, but facilitates its later use.

We proceeded in a similar way for the angular selection function. As found before (Einasto et al. 1997d), it can be best approximated as a distribution of $\sin(b)$.

When generating the reference samples, we used ten times more particles than in the observed samples to avoid the shot noise from normalization. As we know the integral distribution functions, generation of the random reference samples is straightforward. We generate a uniform random number $f \in [0, 1]$ and solve the equation $F_r(r) = f$ for r . Then we select another f and solve the equation $F_{\sin b}(\sin b) = f$ for $\sin(b)$,

select a uniformly distributed Galactic longitude $l \in [0, 2\pi]$ and place our point.

We fitted selection functions for each sample separately, as the spatial effects of the procedures used to select clusters for different samples are difficult to quantify.

The coefficients of the selection functions and the coordinate limits are given in Table 1.

3.3. Regularity analysis

Although we should search first for the best orientation of our cubic alignment, we rely on our visual impressions that tell us that the distribution of clusters in high-density regions looks fairly regular when seen in supergalactic coordinates (see Fig. 13). Thus we shall study the orientation problems later and suppose now that the best orientation is just along supergalactic SGX, SGY, SGZ . We shall learn that this is not a bad assumption at all.

Let us turn first to the analysis of the full sample. The (reduced) regularity periodogram for that sample is shown in Fig. 15 and the periodograms for the sub-cones in Fig. 16. In all these figures the periodogram is shown along with the confidence regions obtained from 100 Poisson models with the same number of clusters, same geometry and same selection functions as those of the observed sample. Dotted lines show the 99% and 80% confidence regions.

We see that the regularity periodogram shows a clear signal on scales of $130 h^{-1} \text{ Mpc}$ and $220\text{--}240 h^{-1} \text{ Mpc}$. If the distribution contained an ideal cubic signal, the second peak (the first sub-harmonic) should have been at $260 h^{-1} \text{ Mpc}$, but there is none. The first harmonic, if it exists, has a bit too large a period of $\sim 75 h^{-1} \text{ Mpc}$. The widening of the confidence regions at $\sim 350 h^{-1} \text{ Mpc}$ is due to the double-cone geometry of the sample, given that $350 h^{-1} \text{ Mpc}$ is the characteristic size of both cones.

Analysis of separate cones allows us to see where the regularity signal comes from. So, the northern cone (Fig. 16, upper panel) supports both the $130 h^{-1} \text{ Mpc}$ peak (rather weakly) and

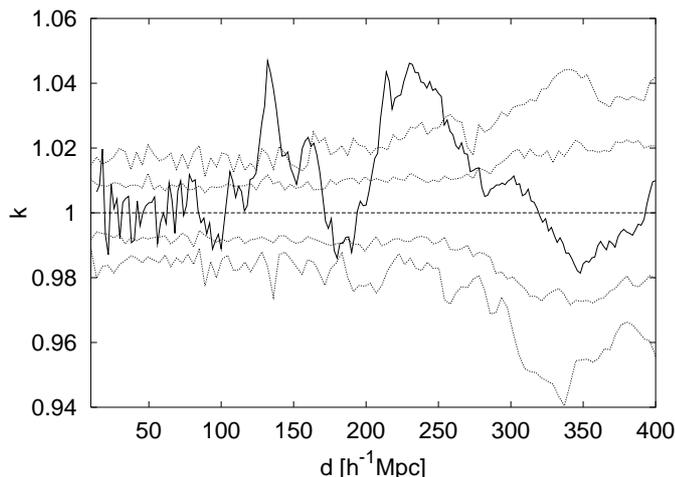


Fig. 15. Regularity periodogram for the full cluster sample (solid line). Dotted lines show the 80% and 90% confidence regions.

the $260 h^{-1}$ Mpc peak (note the exact match of the two periods) but the southern cone (Fig. 16, lower panel) supports only the $130 h^{-1}$ Mpc peak (and its harmonic, the $65 h^{-1}$ Mpc peak can be seen here, too). Surprisingly, the expected $260 h^{-1}$ Mpc signal is practically absent in the southern cone. Thus there seems to be a good regularity signal from both cones, but joining them together mixes the periods and amplitudes.

As suggested by correlation analysis (Tago et al. 2000), the subsample of clusters from dense regions (ACO8) should be especially regular. The regularity periodograms for that sample are shown in Figs. 17, 18 and 19.

The double-cone sample (ACO8) shows a very strong signal at $140 h^{-1}$ Mpc, the first harmonic at $60\text{--}70 h^{-1}$ Mpc, a signal at $225 h^{-1}$ Mpc and another at $340 h^{-1}$ Mpc. The latter is probably due to an extremely inhomogeneous nature of the distribution that is difficult to model correctly. The surprise comes from the northern cone (Fig. 18), where only a $60\text{--}70 h^{-1}$ Mpc period can be seen, moreover at a low confidence level. For most of the periods this sample is more uniform than the Poisson one (this is no surprise, in principle, as Poisson distributions are not too uniform). Such a uniformity could be the side effect of our recipe that discards nearby points in real space when estimating densities.

The southern cone of the ACO8 sample shows a pronounced regularity signal, although at a slightly larger period than expected, at $140 h^{-1}$ Mpc. The first harmonic at $70 h^{-1}$ Mpc is also clearly seen. The small size of the sample prevents the search for sub-harmonics.

3.4. Low-dimensional regularities

The first indications of large-scale periodicity in the distribution of galaxies and clusters were found by Broadhurst et al. (1990). The period they found was along a pencil beam and thus a one-dimensional regularity with a period in the range of those we are seeing here. A periodic signal along two axes, or a single axis of a cube will be detected by the regularity periodogram as a cubic signal, too, only of a lower amplitude.

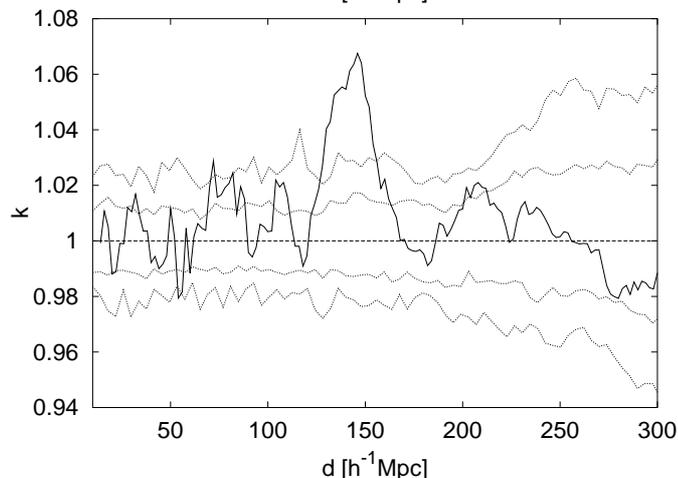
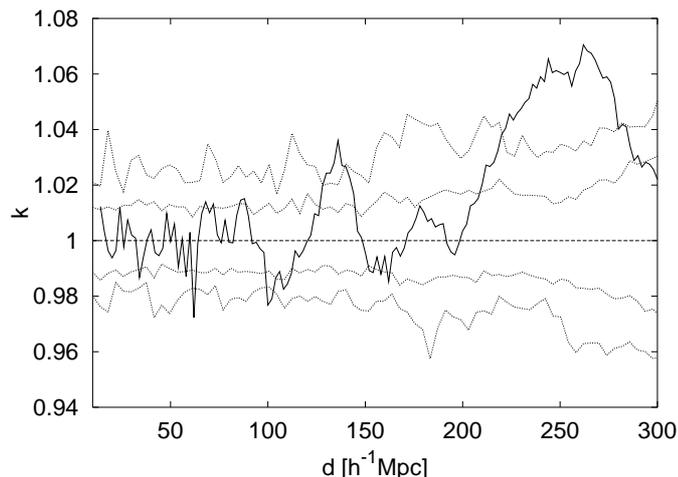


Fig. 16. Regularity periodograms for the northern sub-cone of the full cluster sample (upper panel), and for the southern sub-cone of the full cluster sample (lower panel). Periodograms are shown with solid lines, the 80% and 90% confidence regions with dotted lines.

Thence we have to check whether the periods seen in the periodograms above really imply a cubic network.

The approach we used can be easily modified to be applied to 2-D and 1-D “cubes”. For this purpose the points (clusters) have to be projected either on a coordinate plane or on a coordinate axis, and the kernel has to be modified accordingly. This needs only slight modifications of the programs. We performed this exercise for the ACO8 sample of clusters in dense regions and its southern sub-cone ACO8S. Given that the periodicity discovered by Broadhurst et al. (1990) was observed towards the Galactic poles and this direction is very close to the supergalactic Y-axis, we performed the 1-D analysis along this axis and the 2-D analysis for the supergalactic X-Z plane.

These 2-D and 1-D results are shown in Fig. 20. The periodogram for the full sample shows rather weak one-dimensional periodicity. The sequence of periods ($\sim 65 h^{-1}$ Mpc, $\sim 130 h^{-1}$ Mpc, $\sim 210 h^{-1}$ Mpc) indicates that $\sim 65 h^{-1}$ Mpc seems to be the main period in the supergalactic Y-direction.

The 2-D periodicity is stronger and clearer. However, the 1-D and 2-D periods are not well correlated. The rapid growth

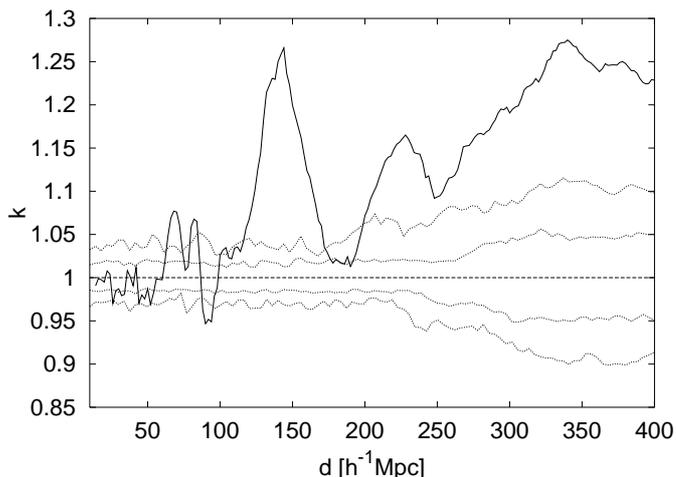


Fig. 17. Regularity periodogram for the subsample of clusters from dense regions (ACO8, solid line). Dotted lines show the 80% and 90% confidence regions.

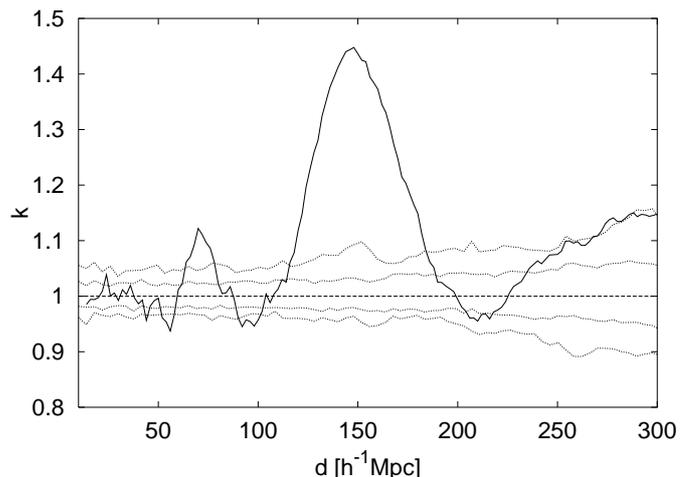


Fig. 19. Regularity periodogram for the southern sub-cone of the ACO8 cluster sample (solid line). Confidence regions are as above.

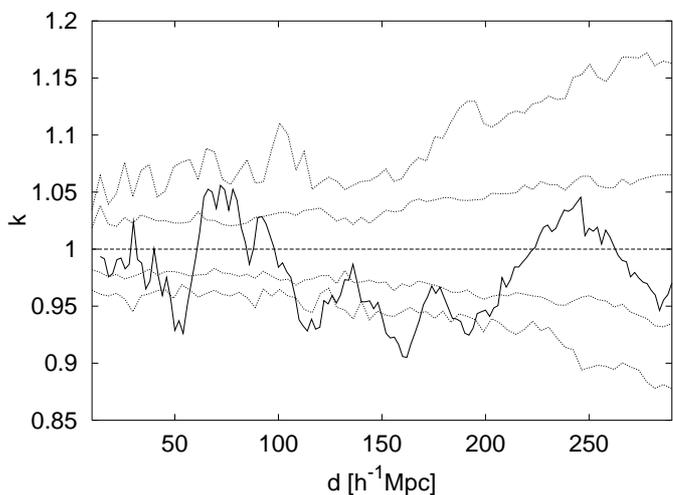


Fig. 18. Regularity periodogram for the northern sub-cone of the ACO8 cluster sample (solid line). Confidence regions are as above.

of the 2-D signal at $300 h^{-1} \text{Mpc}$ is probably a normalization effect.

The 2-D and 1-D periodograms for the ACO8S sample are closer to each other, although they are also shifted relative to each other. The small volume of this sample does not allow us to search for larger periods.

We conclude from this analysis that although the main cubic signal comes from the distribution of rich clusters in the supergalactic $X - Z$ plane, there is a substantial contribution from the perpendicular direction, too. This is, however, smaller than could be expected from the clear periodicity observed by Broadhurst et al. (1990) in that direction. And, as before, we see that the main source of this signal is the southern sample of clusters belonging to rich superclusters.

3.5. Search for the best alignment

The results presented above tell us that there is a clear regularity in the distribution of clusters of galaxies (especially those

from high-density regions) when the trial cubes are aligned along supergalactic coordinates. Is this the best orientation, or are there other preferred alignments, too?

A cube is a solid body and so its orientation in space (with respect to a fixed coordinate system) is described by three Euler angles. We have chosen the following convention for these angles: 1) α is the angle of rotation of the cube around the z -axis, 2) β is the angle of rotation around the resulting y -axis of the cube and 3) γ is the angle of rotation around the resulting z -axis of the cube. The first two angles determine the direction of the new z -axis and the last one rotates the cube around the latter. Because of the cubic symmetry all the possible alignments are described by the range of angles $\alpha, \beta, \gamma \in [0, 90^\circ]$. As we have to compute a full periodogram for every combination of the Euler angles, the full parameter space is 4-dimensional.

So far our analysis was one-dimensional, the only parameter we varied was the trial period d . This is similar to the conventional case of isotropic statistics, which depend only on the distance r . In contrast to that, the search for special spatial arrangements is much more difficult and time-consuming. In our case, we had to search the 4-dimensional parameter space, and this formed the main amount of work done in this study.

To start, we have also to quantify what we mean by a “better periodogram”. The maximum amplitude of a periodogram is probably not the best candidate for that – in the periodograms shown above we have seen examples when periodograms grow at large d because of insufficient compensation for edge effects. We would like to see well-separated maxima in periodograms, and we can quantify that by doing a discrete wavelet decomposition of a periodogram and choosing the maximum wavelet amplitude to describe the periodogram. Since the wavelet decomposition represents a function as a sum of localized “peaks” (see, e.g. Press et al. 1997), the result should be close to our expectations. Our experience has shown that in almost all cases we studied it corresponded to our notion of the best periodogram. An example of the procedure is shown in Fig. 21.

We used the fast discrete wavelet transform from Press et al. (1997) with the simplest Daubechies 4-coefficient wavelet filter. We tried also the “Mexican hat” wavelet, but the results

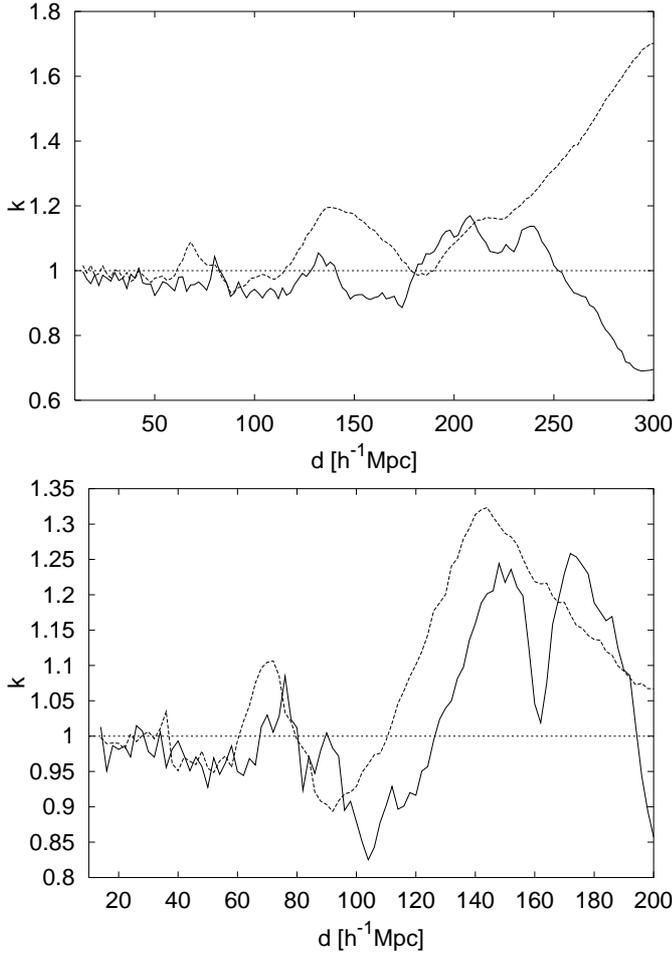


Fig. 20. Low-dimensional regularity periodograms for the ACO8 sample. The solid line show 1-D regularity along the supergalactic Y -axis (along the Galactic poles) and the dashed lines shows 2-D regularity in the $X-Z$ plane. The upper panel shows the periodogram for the full sample and the lower panel for the southern ACO8S subsample.

were similar. As the fast wavelet transform requires a data vector of a length equal to a power of 2 and supposes that this vector is periodic, we padded the remainder of our periodogram, interpolating linearly between its first and last points. This should diminish the amplitude of the first wavelet coefficient; to be on the safe side, we ignored this coefficient when finding the maximum amplitude, anyway. Given that the positive amplitude of basis wavelet functions is usually larger than the negative amplitude (this is certainly the case with the Daubechies wavelet we used), we chose the maximum wavelet transform amplitude as a result, not the maximum absolute magnitude.

As seen in Fig. 21, when comparing the two slightly different periodograms, the better periodogram gives a larger maximum wavelet amplitude, as expected. Since the padding procedure expanded the periodogram range to $512 h^{-1} \text{Mpc}$ (256 steps of $2 h^{-1} \text{Mpc}$ resolution), the scale of this (4th) amplitude is $128 h^{-1} \text{Mpc}$, and corresponds to the $\sim 130\text{--}140 h^{-1} \text{Mpc}$ peak in the upper panel.

Having replaced a periodogram by its amplitude (let us denote it by A), we have reduced our problem of period search to

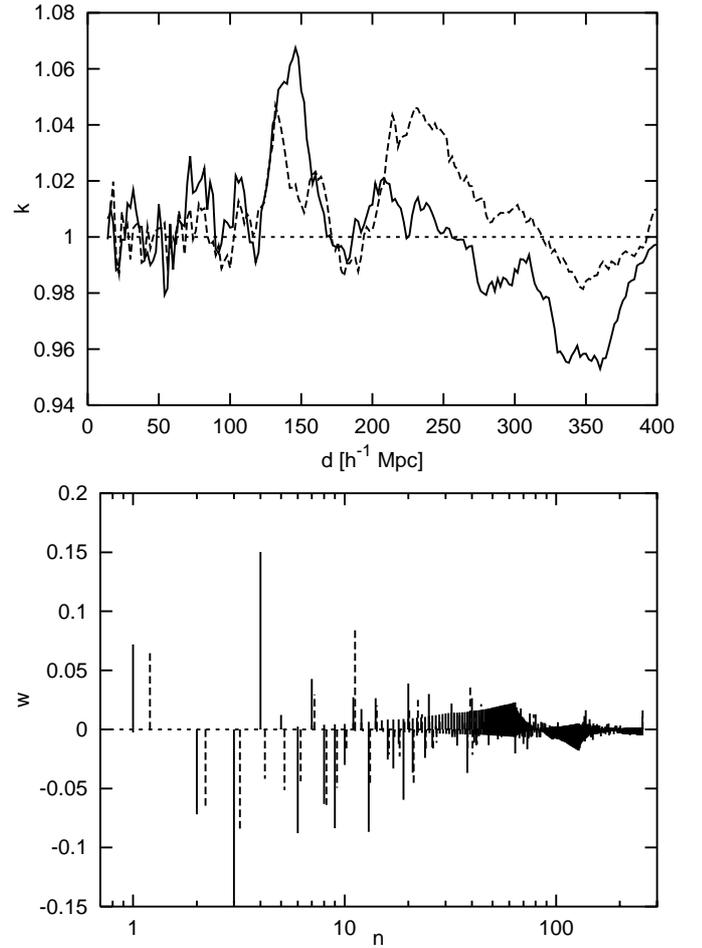


Fig. 21. Two regularity periodograms (upper panel) and their wavelet decompositions (lower panel). The “better” periodogram is shown by solid line in both panels. The wavelet amplitudes of the second (dashed line) periodogram in the lower panel are slightly shifted in the wavelet index n .

a 3-D cube of Euler angles that is easier to manage. We used for the period search the Powell algorithm described in Press et al. (1997) and the algorithm from that book. However, as we cannot suppose that the function $A(\alpha, \beta, \gamma)$ has only one (local) maximum, we have to explore first the full cube of Euler angles and to find the coordinates for the starting points for our search.

The computation of A for that cube is rather time-consuming. We computed a cube with a 5° resolution; it took about a full day on an SGI Origin 200 for the largest ACO1 cluster sample. We also computed only the unreduced periodogram, as otherwise the computational time would have grown too large. As the periodogram for the reference sample is fairly smooth, that probably did not change much the local properties of the amplitude field.

In order to better see local maxima, we then generated an inhomogeneous Poisson point field in the Euler cube with a density of

$$p(A) = \frac{1}{(1 + (A_{\max} - A) / \varepsilon)^P},$$

where the scale parameter ε and the exponent P were chosen to achieve the best visual impression. The Poisson points will

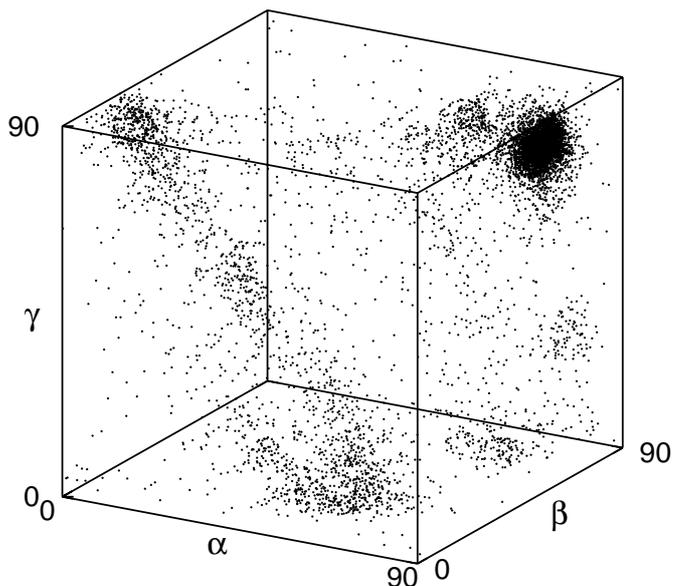


Fig. 22. An inhomogeneous Poisson point process for the cube of Euler angles. The high-density clumps show the local maxima of the wavelet amplitude of the regularity periodogram for the ACO8S sample.

fill preferentially the regions where the amplitudes A are close to the maximum one, in an interval of about ε . The higher the exponent P , the higher will be the density contrast. We found that good point sets were generated by choosing the number of points $N_E = 8000$, $\varepsilon = 0.3$ and $P = 8$ for the ACO8 and ACO8S samples and changed ε to 0.03 for the full ACO1 catalog. The value of A for a point was obtained by linear interpolation in the wavelet amplitude cube. An example of such a point distribution (let us call it the Euler cube) is shown in Fig. 22 for the ACO8S catalog. We see that there are only a few regions of Euler angles corresponding to periodograms with high maxima (where the Poisson process is clustered most). Because of a small size of the ACO8S sample, slight correlations of these regions also can be seen.

The coordinates of the densest clusters of points in the Euler cube are approximate solutions for the best orientation periodograms. In order to obtain the final solutions, we used the Powell algorithm. The periodogram interval for the search was limited to $d = 260 h^{-1}$ Mpc in order to avoid possible normalization effects. Usually the final location of the maximum remained close to the initial one; the change was in most cases less than 10° in every coordinate.

The most crowded cluster of final solutions for the Euler angles for the ACO8S sample was close to $(\alpha, \beta, \gamma) = (6^\circ, 3^\circ, 0^\circ)$ with a scatter of about 2° . The effect of that rotation is to bring the Y -axis closer to the direction to the (northern) Galactic pole (the angle between the supergalactic Y -axis and the Galactic pole is 6.3°). This gave also the maximum amplitude for the $140 h^{-1}$ Mpc peak of the periodogram. We illustrate this in Fig. 23. We do not show confidence limits on periodograms here and below to avoid crowding; they are similar to those shown in Figs. 15 to 19.

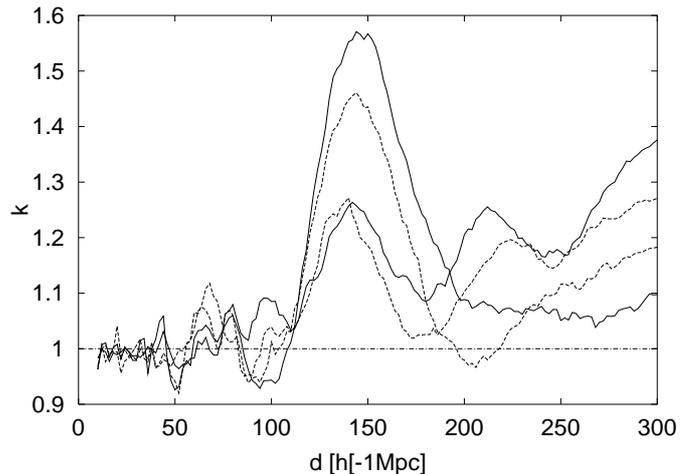


Fig. 23. Maximum amplitude periodograms for the ACO8 and ACO8S samples. The low amplitude periodograms are for ACO8, the high amplitude pair is for ACO8S. Dotted lines describe the cubic grid aligned with supergalactic coordinates, solid lines – the best solutions.

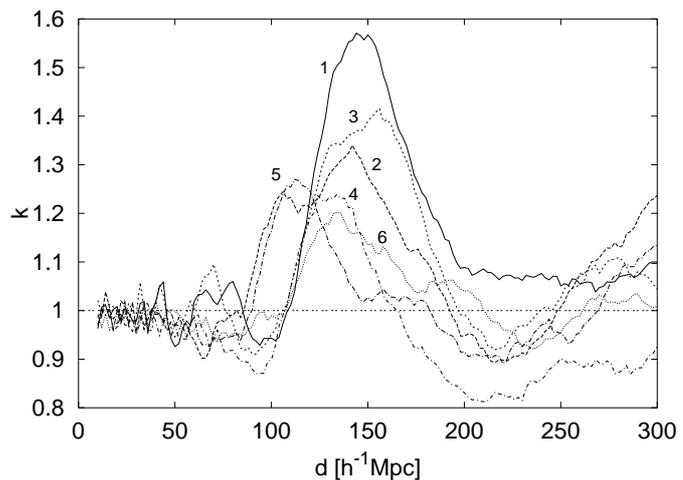


Fig. 24. The best periodograms (local maxima) for the ACO8S sample. The curve labeled “1” is the overall best solution. The Euler angles for all periodograms are given in Table 2.

As we see, the solution is better, but not much; for the full ACO8 sample the improvement is minimal, mainly in the amplitude of the second sub-harmonic.

We show the best solutions for the ACO8S sample in Fig. 24, with the very best solution labelled as “1”. The Euler coordinates for other solutions are given in Table 2. As we see, all other acceptable periodograms have considerably smaller amplitudes than the solution aligned approximately along the supergalactic coordinate axes. Some of the acceptable solutions in Fig. 24 also display shorter main periods, starting from $110 h^{-1}$ Mpc. The $130 h^{-1}$ Mpc periodicity favored in our previous papers can be seen here too at a rather high amplitude.

The best solution for the full ACO8 sample, also shown in Fig. 23, has a slightly different alignment than that for the ACO8S cone: $(10.1^\circ, 10.1^\circ, 81.9^\circ)$; the last rotation is equivalent to $\gamma = -8.9^\circ$. This reduces mainly to the rotation by $\sim 10^\circ$ around the supergalactic Y -axis. This solution has a period of

Table 2. The best alignments for cubic patterns for various samples. The numeration of the curves for the ACO8S sample is the same as in Fig. 24. All angles are in degrees.

No	α	β	γ	No	α	β	γ
ACO1				ACO8			
1	0.0	0.0	0.0	1	10.1	10.1	81.9
2	55.9	87.4	81.0	2	86.2	33.6	82.4
3	72.9	11.0	38.2	3	38.7	22.0	19.1
4	82.0	11.5	27.4	4	66.9	30.3	85.1
5	6.5	81.9	78.1	ACO8S			
6	77.0	41.3	85.9	1	6.7	3.9	0.0
7	22.2	11.0	50.1	2	72.0	81.8	69.6
8	68.3	12.1	41.4	3	10.0	14.0	82.3
				4	65.0	56.9	80.1
				5	55.7	25.1	0.7
				6	69.4	65.4	2.7

$130 h^{-1}$ Mpc, and the $260 h^{-1}$ Mpc sub-harmonic is also clearly seen.

All searches for the best solutions converged to that solution or to the other three solutions, which are listed in Table 2. In contrast with the ACO8S subsample, these solutions have amplitudes similar to the best solution. Rather than single well-defined periods there are whole ranges of acceptable periods.

The best solutions for the full sample are all of similar amplitude. They display a variety of periods, one of these always in the range 120 – $140 h^{-1}$ Mpc, as we have found for previous samples. The Euler angles for the solutions are given in Table 2. The total number of points in the signal can be estimated by formula (9), supposing that the filling factor of the structure is around 0.3. Using the amplitude ~ 0.06 of the best periodogram for the sample ACO1, we obtain ~ 0.13 for S/N , or ~ 200 points for the signal. The number of clusters in sample ACO8S, where most of the signal comes from, is 205. Does the periodogram see them all in the full sample? Probably it does, but it cannot define the structure well, due to the small amplitude of the signal. The large number of possible periods tells us that the cubic structure is rather vague so that the full sample can be tiled into cubes in different ways. In the contrary, the ACO8S sample displays a unique regular cubic tiling.

As the final best solutions are close to the supergalactic alignment that we started with, it may seem that the large amount of work done to analyze the full Euler cube was unnecessary. We do not think so; we can be sure that this is the best alignment only after having searched the full parameter space. We also stress that there are no shortcuts, and the full analysis can not be avoided.

3.6. Special orientations

Finally, we can check how perfect our cubic symmetry is by computing periodograms for special orientations, like e.g. those aligned along the diagonals of the original cube. In a perfect cubic world there would appear new periods $\sqrt{2}$ times larger, or $\sqrt{3}$ times larger than the basic period, depending on the orientation. In fact, this does not happen in the samples we

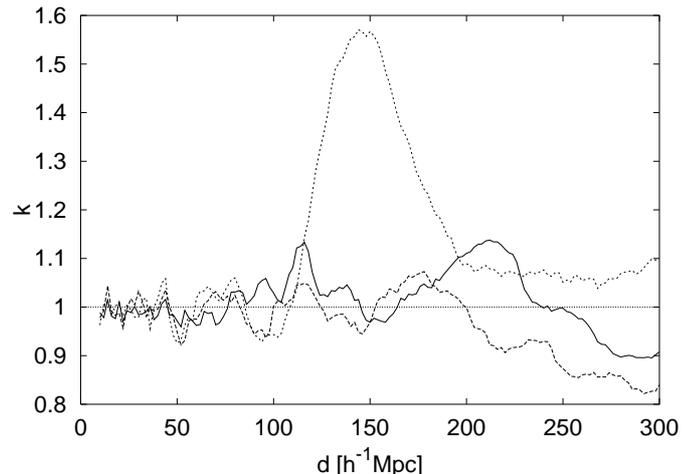


Fig. 25. The best periodograms for special orientations of the cubes (the cube axes form 45° angles with the supergalactic coordinate axes). Our overall best solution for ACO8S (short-dashed line) is shown to compare the amplitudes. The solid line shows the solution for the full ACO8 sample and the dashed line for the ACO8S sample.

have studied here. There is almost no trace of a signal for such alignments, and when we choose those as starting points for our period search, the search ends in the same (oblique) alignments summarized in Table 2, or returns to the almost perfect alignment. The two best special orientation periodograms are shown in Fig. 25. Their amplitude is a little larger than for other special orientations, but small compared with the best solution (for ACO8S), that is shown in that Figure for reference. The periodograms describe both the ACO8 and ACO8S samples and their orientation is the same, $(0^\circ, 45^\circ, 0^\circ)$, i.e. they describe cubes rotated around the supergalactic Y -axis by 45° . A look at Fig. 13 shows that one could really see such a pattern, knowing that it exists.

3.7. Resolution effects

We used the resolution of 30 points per density kernel throughout this analysis. From the point of view of density calculations, this seems the smallest one can safely choose. On the other hand, for smaller samples this will certainly give a very low spatial resolution (only 7 cubic bins for the 205 clusters of ACO8S). Let us see how the resolution influences the results, and how small one could take the kernel. We compared periodograms with different kernel widths for the ACO8 and ACO8S samples. The results are shown in Fig. 26.

We chose the best alignments for both samples, started with the standard 30 points per kernel, and then lowered this number down to 5 and 3. We see that the periodograms remain rather smooth even for extremely small kernels (individual densities might be noisy already, but summation smoothes them) and edge effects grow fast. Larger kernels smooth the signal and lower its amplitude, as expected. An interesting trend is that in both cases the periods move towards the magic $130 h^{-1}$ Mpc with smaller smoothing. The choice of a best kernel size is a matter of practice, but we think that our choice of a 30-point kernel was correct since otherwise edge effects could have

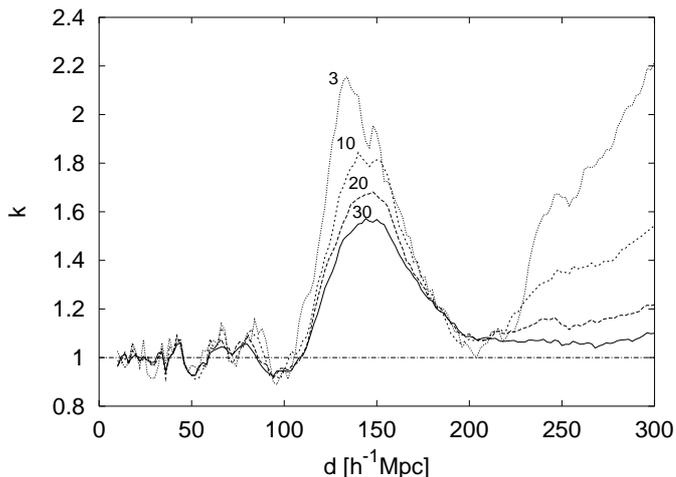


Fig. 26. Periodograms produced with different kernel sizes (expressed in points per kernel volume) for the ACO8S subsample. The periodograms are labeled with the kernel sizes used.

ruined our search for best alignments. Thus the strategy of using smoother periodograms first for alignment search and refining the periods by a subsequent use of smaller kernels, appears to be good.

3.8. Phase patterns

The regularity periodogram is an interesting curve to study, but it does not tell us exactly what the structure looks like. However, we can find it out easily by plotting the phase cubes. We show two of those in Fig. 27, for the ACO8 and ACO8S samples. We do not show the phase cube for the full sample; as expected, it does not look much different from a Poisson point distribution – the number of points is large and the amplitude of the cubic signal small.

The phase cubes shown correspond to the best alignments, counted as number one in the alignment Table 2. The best periods for these alignments are $120 h^{-1} \text{Mpc}$ for ACO8 and $140 h^{-1} \text{Mpc}$ for ACO8S. As indicated by high-resolution runs described above (see Fig. 26), we chose a period of $130 h^{-1} \text{Mpc}$ for both samples.

The phase cubes show, firstly, that the cubic regularity really exists; the density distribution in the phase cube is highly inhomogeneous. The signal is mainly due to a rich point cluster plus filamentary features, obtained easiest by folding a lattice with highest concentrations at its vertices. The cubes confirm the picture we get from the amplitude of the signal: for ACO8S with 205 clusters, the extra regularity signal is about one. Assuming that the filling factor of the structure is about 0.3, formula (9) tells us that about 60% of all points could belong to the regular structure. Apart from that, the phase cubes show also that the cubic structure is far from perfect and that one could try to fit other tilings.

4. Discussion and conclusions

We have confirmed that the Abell cluster sample contains a considerable fraction of clusters located in rich superclusters

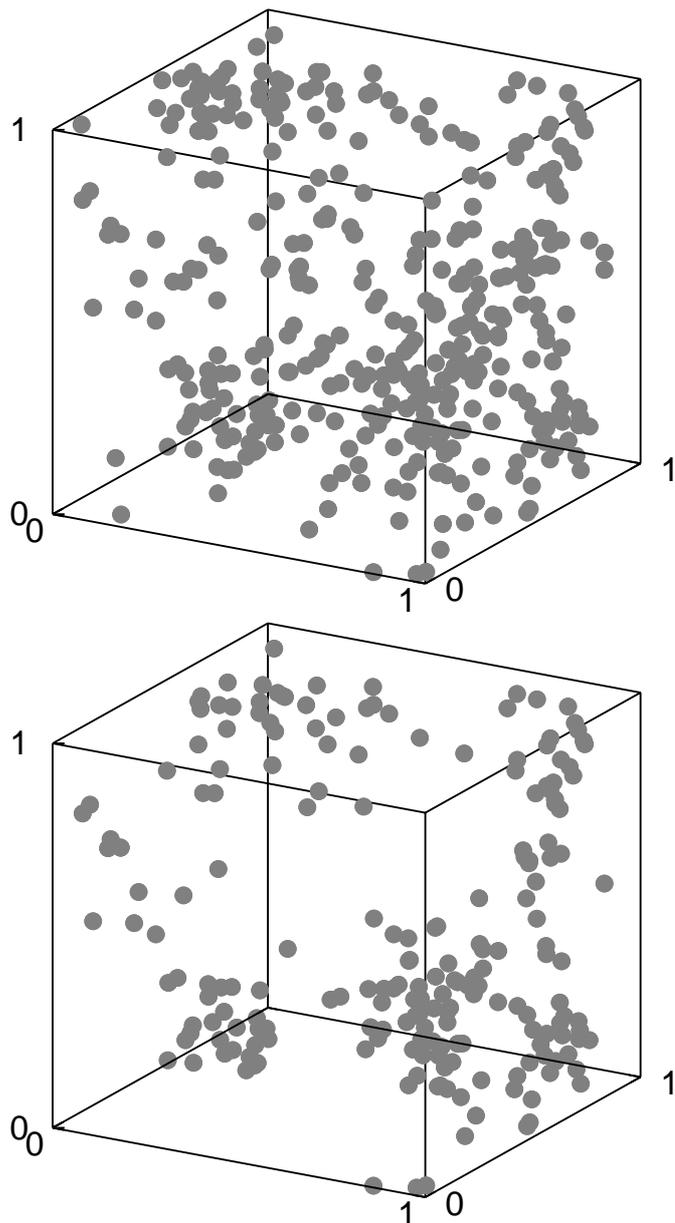


Fig. 27. Phase cubes for the best alignments and for the period $130 h^{-1} \text{Mpc}$ for the ACO8 sample (upper panel) and for the ACO8S subsample (lower panel).

which form a cubic lattice with a period about $130 h^{-1} \text{Mpc}$. This structure is best seen in the clusters populating high density regions, but it is present already in the whole Abell cluster sample. This supports the results obtained in Papers I and II by using other methods.

The cubic lattice we see in the distribution of clusters defines specific directions in space. A clear periodicity is observed only along a few alignments, while the best signatures of periodicity come from a cubic tiling that is nearly aligned with supergalactic coordinates.

Observationally, this result is not so surprising as it may seem at the first glance. It confirms earlier results on the presence of a high concentration of clusters and superclusters towards both the Supergalactic Plane (Tully et al. 1992), and

towards the Dominant Supercluster Plane, which are at right angles with respect to each other (Paper I). Tully et al. also noticed the rectangular character of the distribution of rich clusters. Thin deep slices, such as slices of the LCRS and the Century Survey, also show a weak periodicity signal (Landy et al. 1996; Geller et al. 1997). The scale length found in these studies is of the same order as obtained in the present paper.

The supergalactic Y axis is very close to the direction of Galactic poles, thus it is natural to expect a well-defined periodicity along Galactic poles as indeed observed by Broadhurst et al. (1990). As noted already by Bahcall (1991), the nearest peaks of the Broadhurst et al. survey coincide in position and redshift with nearby rich superclusters. We note, though, that the periodicity in the distribution of rich superclusters over the full sky is much less pronounced. This is in agreement with observational periodicity studies along beams oriented in other directions (Guzzo et al. 1992; Willmer et al. 1994; Etori et al. 1997).

Kerscher (1998) used a combination of the nearest neighbour distribution and the void probability function to measure the regularity of the structure in the Abell cluster sample. His method is more universal, and detects any regularity, but it is isotropic and cannot detect specific structures. He found that the rich supercluster sample shows clear signs of regularity, in agreement with Paper I. He also found that the regularity can be partly due to the friends-of-friends (FOF) procedure used to select the rich supercluster population. His analysis extends to scales about $60 h^{-1}$ Mpc and the neighbourhood distance $24 h^{-1}$ Mpc used for FOF can have an effect at these scales, but certainly not much at scales of $130 h^{-1}$ Mpc and larger.

The regularity of the large-scale structure of the Universe has been studied also using the distribution of centers of superclusters (Kalinkov et al. 1998). These authors looked for high-order clustering of superclusters using correlation analysis, and found that superclusters are not clustered. Our analysis in Papers I and II and that presented here suggests that the regularity is completely different. The supercluster-void regularity is seen in the distribution of clusters themselves, not in that of supercluster centers.

Theoretically, this result contradicts almost everything we know about the formation of structure in the Universe. It is not clear at all if it may be generated by an isotropic *Gaussian* random field of initial perturbations, even with bumps and wells in its power spectrum, since a rather large fluctuation would be necessary to generate a locally anisotropic realization required to explain the data. In particular, Eisenstein et al. (1998) speculate that it could be a domination of a few Fourier modes in our location, but the chance of that seems vanishingly small. If this were true, it would be easy to detect from future deeper surveys: the phase alignment observed in our neighbourhood will soon be lost at larger distances, and deeper surveys should not reveal any regularity.

Thus, if the regularity is real, then most probably it means a non-isotropic *non-Gaussian* admixture to a Gaussian density field in the region of $k = (0.03-0.06) h \text{Mpc}^{-1}$ where we tentatively see non-smooth features in $P(k)$. Of course, in this case $P(k)$ is not sufficient for description of stochastic perturbations, a whole hierarchy of higher moments or probability

distributions is required. Still, from a theoretical point of view, local non-Gaussian features superimposed on a Gaussian field of initial perturbations with an approximately flat power spectrum are fairly possible. The most natural mechanism of their generation is essentially the same as the mechanism for generation of local non-smooth features in the power spectrum – a fast phase transition in physical fields other than the inflaton field during an inflationary stage in the early Universe (occurring about 50 e-folds before its end), see, e.g., the discussion in Starobinsky 1998 – only a different kind of the phase transition is required. In particular (though not necessarily), cosmic strings or more complicated topological defects may be generated during it. However, we clearly know too little yet about the properties of the large-scale quasi-regularity to discuss specific scenarios.

A conclusive confirmation of large-scale regular patterns will demand deeper samples that we can presently use. In this respect, the future complete data release of the 2dF survey, and, especially, the BRG part of the SDSS survey will certainly help to clarify the situation.

The present paper is of an exploratory nature, so if traces of regularity will be seen in the future, there are many obvious ways to improve the methods used here. The notion of a “best periodogram” is not too clear yet; one should consider weighing of points again to speed up calculations etc. We have not touched the problem of false periods at all; for people who study time series this is one of the main problems. The most time-consuming part of the work at present is the calculation of the “Euler cube” to find starting values for the best period search. As the 3-D Fourier transform of density has directions as arguments, one could use that to find initial Euler values, looking for pairs of perpendicular wave vectors of maximum total Fourier amplitude (maximum change along the axes of the cube).

In conclusion, we have shown that in our neighbourhood up to a distance of $350 h^{-1}$ Mpc from us the Universe is not fully homogeneous and isotropic on scales of the order of hundreds of Mpc. High-density regions, delineated by galaxy clusters in rich superclusters, form a quasi-cubic lattice in this volume of space. This lattice is roughly aligned with the supergalactic coordinates, the regularity is well pronounced and has a period of $120-140 h^{-1}$ Mpc. This inhomogeneity does not contradict the isotropy and homogeneity of the Universe on larger scales, which follows from the data on the spatial distribution of galaxies and clusters, from the data on the angular anisotropies of the CMB, and from the observed degree of isotropy of the X-ray background in the Universe (emitted by baryons in the recent epoch).

Acknowledgements. We thank Jaan Pelt and Vicent Martínez for discussion and suggestions. Our special thanks are to our referee, Rien van de Weygaert, for stimulating suggestions and for the program to calculate the Voronoi model. The present study was supported by Estonian Science Foundation grants 2625 and 2882. A.S. was partially supported by the grant of the Russian Foundation for Basic Research No. 99-02-16224, and by the Russian Research Project “Cosmomicrophysics”. H.A. thanks CONACyT for financial support under grant 27602-E, and E.S. thanks the University of Valencia where this paper was finished.

Appendix A: Regularity periodograms for cubes

Regularity periodograms for constant density cubes can be found analytically, counting regions of different density generated by partially filled period cells at the boundaries. This is effectively a description of edge effects.

We found these expressions useful when debugging our code. Moreover, we discovered that our original estimator was biased, by comparing the results of our program with the analytic formula.

As this method can be used in a similar way also for 2-D and 1-D cubes (squares and line segments), we give formulae for these cases, too.

We shall use two auxiliary functions below:

$$m(x) = \lfloor 1/x \rfloor,$$

where the floor function $\lfloor \cdot \rfloor$ is defined as the largest integer smaller than or equal to its argument, and

$$v(x) = 1 - xm(x).$$

We use the normalized period $y = d/L$, where d is the test period and L is the size of the cube.

The formulae:

3-D, all points:

$$\begin{aligned} \kappa_{3D}(y) = & y^3 \{ [y - v(y)]^3 m^6(y) \\ & + 3v(y)[y - v(y)]^2 [m^3(y) + m^2(y)]^2 \\ & + 3v^2(y)[y - v(y)] [m^3(y) + 2m^2(y) + m(y)]^2 \\ & + v^3(y) [m^3(y) + 3m^2(y) + 3m(y) + 1]^2 \}; \end{aligned}$$

3-D, no local correlations:

$$\begin{aligned} \kappa'_{3D}(y) = & y^3 \{ [y - v(y)]^3 m^3(y) [m^3(y) - 1] \\ & + 3v(y)[y - v(y)]^2 [m^3(y) + m^2(y)] \\ & \times [m^3(y) + m^2(y) - 1] \\ & + 3v^2(y)[y - v(y)] [m^3(y) + 2m^2(y) + m(y)] \\ & \times [m^3(y) + 2m^2(y) + m(y) - 1] \\ & + v^3(y) [m^3(y) + 3m^2(y) + 3m(y) + 1] \\ & \times [m^3(y) + 3m^2(y) + 3m(y)] \}; \end{aligned}$$

2-D, all points:

$$\begin{aligned} \kappa_{2D}(y) = & y^2 \{ [y - v(y)]^2 m^4(y) \\ & + 2v(y)[y - v(y)] [m^2(y) + m(y)]^2 \\ & + 2v^2(y) [m^2(y) + 2m(y) + 1]^2 \}; \end{aligned}$$

2-D, no local correlations:

$$\begin{aligned} \kappa'_{2D}(y) = & y^2 \{ [y - v(y)]^2 m^2(y) [m^2(y) - 1] \\ & + 2v(y)[y - v(y)] [m^2(y) + m(y)] \\ & \times [m^2(y) + m(y) - 1] \\ & + 2v^2(y) [m^2(y) + 2m(y) + 1] \\ & \times [m^2(y) + 2m(y)] \}; \end{aligned}$$

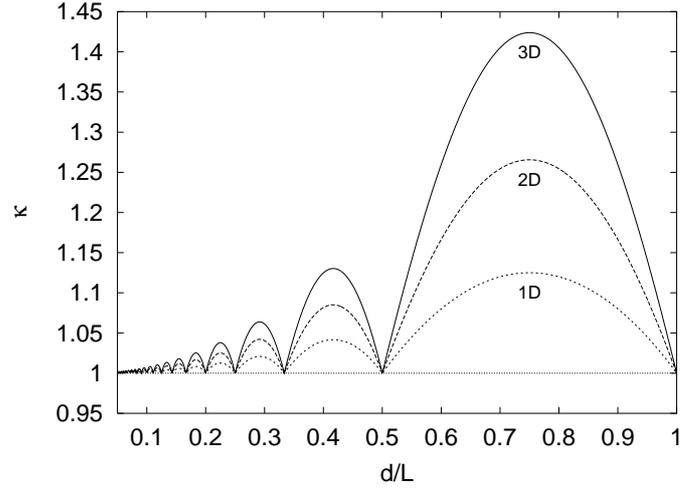


Fig. A.1. Periodograms for uniform density cubes (using the full estimator), for 1D, 2D and 2D cases. The full period range is shown.

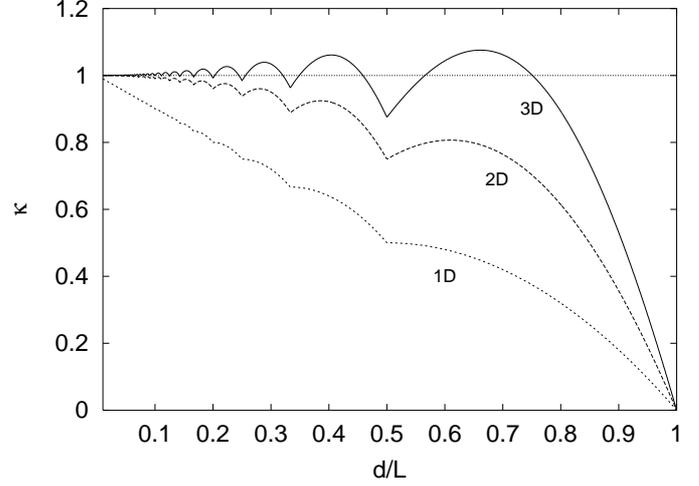


Fig. A.2. Periodograms for uniform density cubes (using the estimator that discards local correlations), for 1D, 2D and 2D cases.

1-D, all points:

$$\begin{aligned} \kappa_{1D}(y) = & y \{ [y - v(y)] m^2(y) \\ & + v(y) [m(y) + 1]^2 \}; \end{aligned}$$

1-D, no local correlations:

$$\begin{aligned} \kappa'_{1D}(y) = & y \{ [y - v(y)] m(y) [m(y) - 1] \\ & + v(y) [m(y) + 1] m(y) \}. \end{aligned}$$

We show all these periodograms in Figs. A.1 and A.2, which show the behaviour of these functions over the entire range of periods. In applications, of course, the periods $d > L/2$ do not make much sense.

The amplitudes of both estimators (edge effects) grow with dimension. Nevertheless the second estimator that discards local correlations, also reduces the edge effects considerably.

Appendix B: Segment Cox processes

We used the segment Cox point process (see Stoyan et al. 1995) to model the correlation properties of our observational samples. This process is constructed as follows: segments of length

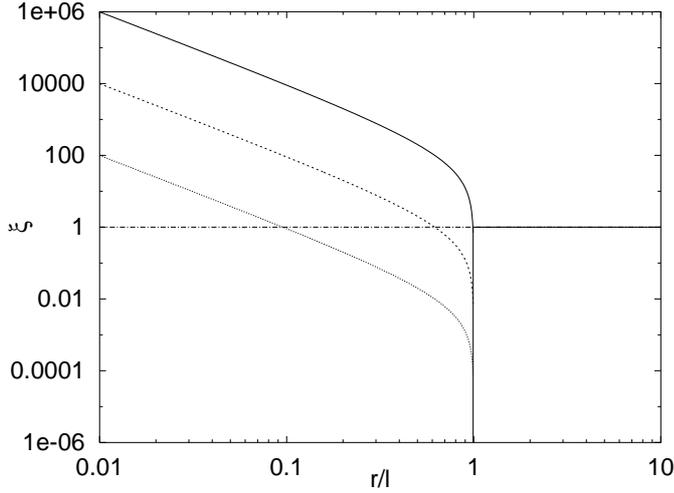


Fig. B.1. The correlation function of a segment Cox process for different correlation amplitudes.

l are randomly generated inside a cube of size L (a Poisson distribution of starting points and a uniform distribution of directions in space) and Poisson-distributed points are generated along these segments. The second parameter we have to choose after l is the intensity (mean density) of segments λ_S . Then the length density of all segments L_V is defined as

$$L_V = \lambda_S L.$$

The last parameter we have to choose in order to specify the process is the line intensity of points λ_L . The intensity of the point process λ is

$$\lambda = \lambda_L L_V = \lambda_L \lambda_S l.$$

Although this distribution of points might seem rather artificial, the good news is that the pair correlation function $\xi(r)$ for this process is known exactly (Stoyan et al. 1995):

$$\xi(r) = \begin{cases} \frac{1}{2\pi L_V} \left(\frac{1}{r^2} - \frac{1}{rl} \right), & x \leq r_0, \\ 0, & x > r_0. \end{cases} \quad (\text{B.1})$$

As we can see, the correlation function depends on r as $\xi(r) \sim (r/r_0)^{-2}$ for small r . This makes it suitable for modeling observed clustering of galaxies and galaxy clusters. It does not depend on the intensity λ_L at all. Also, the process is homogeneous on large scales by construction. We illustrate the correlation function in Fig. B.1.

If we want to simulate a cosmological sample, we are usually given the sample size L , the correlation length r_0 and the number of objects in the sample. How do we choose the parameters l , λ_S and λ_L that describe a segment Cox process?

First, we have to choose the segment length l . It certainly has to be larger than r_0 and can be chosen by the location of the first zero of the observed correlation function. Also, as seen in Fig. B.1, the smaller the ratio r_0/l , the better is the approximation $\xi(r) \sim r^{-2}$. This, however, frequently gives too large a value for d . A large d means that the intensity of segments will be small and their total number $N_S = \lambda_S L^3$ will be small, too; we get a rather unrealistic distribution of points along a few long segments.

On the other hand, if we do not worry too much about the exact slope of the correlation function, we could choose l rather close to r_0 (the upper curve in Fig. B.1). We have found that taking $l \approx 1.5r_0$ works well in practice.

Let us say that we have decided to choose $l = \alpha r_0$. Then, formula (B.1) for the correlation function tells us that

$$L_V = \frac{1}{2\pi r_0^2} \left(1 - \frac{1}{\alpha} \right).$$

We obtain for the intensity (mean density) of the segments:

$$\lambda_S = L_V/d,$$

and the last parameter λ_L :

$$\lambda_L = \frac{N}{L^3 L_V}.$$

An important point to remember when generating a segment Cox process is that having a mean number of segments $N_S = \lambda_S L^3$ does not mean that we have to generate N_S segments. The number of segments is a Poisson-distributed random number with intensity (mean) N_S , and it is useful to recall that its variance is also N_S . So every realization will give us a different number of segments. The same concerns the number of points on the segments. All this combines to give us a different total number of points with every realization.

References

- Abell, G. 1958, *ApJS*, 3, 211
 Abell, G., Corwin, H., & Olowin, R. 1989, *ApJS*, 70, 1
 Andernach, H., & Tago, E. 1998, in *Large Scale Structure: Tracks and Traces*, ed. V. Müller, S. Gottlöber, J. P. Mücke, & J. Wambsgans (World Scientific, Singapore), 147
 Atrio-Barandela, F., Einasto, J., Müller, V., Mücke, J. P., & Starobinsky, A. A. 2001, *ApJ*, 559, 1
 Bahcall, N. A. 1991, *ApJ*, 376, 43
 Baker, J. C., Grainge, K., Hobson, M. P., et al. 1999, *MNRAS*, 308, 1173
 Baugh, C. M., & Efstathiou, G. 1993, *MNRAS*, 265, 145
 Baugh, C. M., & Efstathiou, G. 1994, *MNRAS*, 267, 323
 de Bernardis, P., Ade, P. A. R., Bock, J. J., et al. 2000, *Nature*, 404, 955
 Broadhurst, T. J., Ellis, R. S., Koo, D. C., & Szalay, A. S. 1990, *Nature*, 343, 726
 Davies, S. R. 1990, *MNRAS*, 244, 93
 Dekel, A., Blumenthal, G. R., Primack, J. R., & Stanhill, D. 1992, *MNRAS*, 257, 715
 de Vaucouleurs, G. 1956, *Vistas Astron.*, 2, 1584
 Einasto, J., Klypin, A. A., Saar, E., & Shandarin, S. F. 1984, *MNRAS*, 206, 529
 Einasto, J., Einasto, M., Gottlöber, S., et al. 1997a, *Nature*, 385, 139
 Einasto, J., Einasto, M., Gottlöber, S., et al. 1997b, *MNRAS*, 289, 801 (Paper II)
 Einasto, J., Einasto, M., Frisch, P., et al. 1997c, *MNRAS*, 289, 813 (Paper III)
 Einasto, J., Einasto, M., Tago, E., et al. 1999a, *ApJ*, 519, 441
 Einasto, J., Einasto, M., Tago, E., et al. 1999b, *ApJ*, 519, 456
 Einasto, J., & Gramann, M. 1993, *ApJ*, 407, 443
 Einasto, M., Einasto, J., Tago, E., Dalton, G., & Andernach, H. 1994, *MNRAS*, 269, 301

- Einasto, M., Tago, E., Jaaniste, J., Einasto, J., & Andernach, H. 1997d, *A&AS*, 123, 119 (Paper I)
- Einasto, M., Einasto, J., Tago, E., Andernach, H., & Dalton, G. 2001 (in preparation)
- Eisenstein, D. J., Hu, W., Silk, J., & Szalay, A. S. 1998, *ApJ*, 494, L1
- Ettori, S., Guzzo, L., & Tarengi, M. 1997, *MNRAS*, 285, 218
- Fetisova, T. S., Kuznetsov, D. Y., Lipovetskij, V. A., Starobinsky, A. A., & Olowin, R. P. 1993, *Astron. Lett.*, 19, 198
- Garnavich, P. M., Kirshner, R. R., Challis, P., et al. 1998, *ApJ*, 493, L53
- Gaztañaga, E., & Baugh, C. M. 1998, *MNRAS*, 294, 229
- Geller, M., Kurtz, M. J., Wegner, G., et al. 1997, *AJ*, 114, 2205
- Giovanelli, R. 1993, *AJ*, 105, 1251
- Gramann, M., & Suhhonenko, I. 1999, *ApJ*, 519, 433
- Guzzo, L., Collins, C. A., Nichol, R. C., & Lumsden, S. L. 1992, *ApJ*, 393, L5
- Hamilton, A. J. S., Tegmark, M., & Padmanabhan, N. 2000, *MNRAS*, 317, L23
- Hanany, S., Ade, P., Balbi, A., et al. 2000, *ApJ*, 545, L5
- Heinrich, L., Körner, R., Mehlhorn, N., & Mücke, L. 1998, *Statistics*, 31, 325
- Icke, V. 1984, *MNRAS*, 206, 1
- Jöeveer, M., & Einasto, J. 1978, in *The Large Scale Structure of the Universe*, ed. M. S. Longair, & J. Einasto (Reidel, Dordrecht), 409
- Jöeveer, M., Einasto, J., & Tago, E. 1978, *MNRAS*, 185, 35
- Kaiser, N., & Peacock, J. A. 1991, *ApJ*, 379, 482
- Kalinkov, M., Valtchanov, I., & Kuneva, I. 1998, *A&A*, 331, 838
- Kerscher, M. 1998, *A&A*, 336, 29
- Kerscher, M., Szapudi, I., & Szalay, A. S. 2000, *ApJ*, 535, L13
- Kopylov, A. I., Kuznetsov, D. Y., Fetisova, T. S., & Shvarzman, V. F. 1984, *Astr. Tsirk.*, 1347, 1
- Kopylov, A. I., Kuznetsov, D. Y., Fetisova, T. S., & Shvarzman, V. F. 1988, in *Large Scale Structure of the Universe*, ed. J. Audouze, M.-C. Pelletan, & A. Szalay (Dordrecht: Kluwer), 129
- Leahy, D. A., Darbro, W., Elsner, R. F., et al. 1983, *ApJ*, 266, 160
- Landy, S. D., Shectman, S. A., Lin, H., et al. 1996, *ApJ*, 456, L1
- Landy, S. D., & Szalay, A. S. 1993, *ApJ*, 412, 64
- Miller, C. J., & Batuski, D. V. 2001, *ApJ*, 551, 635
- Mo, H. J., Deng, Z. G., Xia, X. Y., Schiller, P., & Börner, G. 1992, *A&A*, 257, 1
- Peacock, J. A. 1999, *Cosmological Physics* (Cambridge University Press, Cambridge)
- Peacock, J., & West, M. J. 1992, *MNRAS*, 259, 494
- Peebles, P. J. E. 1980, *The Large-scale Structure of the Universe* (Princeton Univ. Press)
- Percival, W. J., Baugh, C. M., Bland-Hawthorn, J., et al. 2001, *MNRAS*, 327, 1297
- Perlmutter, S., Aldering, G., Della Valle, M., et al. 1998, *Nature*, 391, 51
- Perlmutter, S., Aldering, G., Goldhaber, G., et al. 1999, *ApJ*, 517, 565
- Pons-Bordera, M., Martinez, V., Stoyan, D., Stoyan, H., & Saar, E. 1999, *ApJ*, 523, 480
- Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. 1997, *Numerical Recipes in C*, 2nd ed. (Cambridge University Press, Cambridge)
- Retzlaff, J., Borgani, S., Gottlöber, S., & Müller, V. 1998, *New Astron.*, 3, 631
- Riess, A. G., Filippenko, A. V., Challis, P., et al. 1999, *AJ*, 116, 1009
- Schuecker, P., Böhringer, H., Guzzo, L., et al. 2001, *A&A*, 368, 86
- Schwarzenberg-Czerny, A. 1989, *MNRAS*, 241, 153
- Scott, P. F., Saunders, R. Pooley, G., et al. 1996, *ApJ*, 461, L1
- Starobinsky, A. A. 1988, *Sov. Astron. Lett.*, 14, 166
- Starobinsky, A. A. 1998, in *Large Scale Structure: Tracks and Traces*, Proc. of the 12th Potsdam Cosmology Workshop (15–20 Sep. 1997), ed. V. Müller, S. Gottlöber, J. P. Mücke, & J. Wambsganss (World Scientific, Singapore), 375
- Stellingwerf, R. F. 1978, *ApJ*, 244, 953
- Stoyan, D., Kendall, W. S., & Mecke, J. 1995, *Stochastic Geometry and its Applications* (J. Wiley & Sons, Chichester)
- Sutherland, W., Tadros, H., Efstathiou, G., et al. 1999, *MNRAS*, 308, 289
- Szalay, A. S. 1998, in *Proceedings of 18th Texas Symposium on Relativistic Astrophysics*, ed. A. Olinto, J. Frieman, & D. Schramm (World Scientific, Singapore), 136
- Tadros, H., Efstathiou, G., & Dalton, G. 1998, *MNRAS*, 296, 995
- Tago, E., Einasto, J., Einasto, M., et al. 2000 [[astro-ph/0012537](#)]
- Tegmark, M., Hamilton, A. J. S., & Xu, Y. 2001 [[astro-ph/0111575](#)]
- Toomet, O. 1997, Master thesis (Tartu Univ.)
- Tucker, D. L., Oemler, A. Jr., Kirshner, R. P., et al. 1997, *MNRAS*, 285, L5
- Tully, R. B., Scaramella, R., Vettolani, G., & Zamorani, G. 1992, *ApJ*, 388, 9
- Valdarnini, R. 2001, *A&A*, 366, 376
- Willmer, C. N. A., Koo, D. C., Szalay, A. S., & Kurtz, M. J. 1994, *ApJ*, 437, 560
- van de Weygaert, R. 1991, *MNRAS*, 249, 159
- van de Weygaert, R., & Icke, V. 1989, *A&A*, 213, 1
- Yoshida, N., Colberg, J., White, S. D. M., et al. 2001, *MNRAS*, 325, 803
- Zeldovich, Ya. B., Einasto, J., & Shandarin, S. F. 1982, *Nature*, 300, 407