Correlations and clustering in the universe

© **Francesco Sylos Labini**[1,2]
[1] Institute ``Enrico Fermi Center", Via Panisperna 89 A, Compendio del Viminale, 00184 Rome, Italy
[2] Istituto dei Sistemi Complessi" CNR, Via dei Taurini 19 00185 Rome, Italy
Email: sylos@roma1.infn.it

Abstract: One of the main assumptions to explain structure formation in an expanding universe is that a dominant mass component in the form of non-baryonic dark matter plays a central role. However, although dark matter is supposed to provide with more than 0.9 of the total fraction of the mass-energy in universe, its amount and properties can only be defined a posteriori. In this context, a crucial point concerns the identification of a possible clear feature of dark matter fields which is expected to hold on very general theoretical conditions. This property, in standard cosmological models, is represented by super-homogeneity, i.e. a very fine tuned balance between negative and positive (weak) correlations of density fluctuations, which must be imprinted both in the anisotropies of the Cosmic Microwave Background Radiation and in the large scale distribution of galaxies. We discuss the fact that the complex clumpiness characterizing galaxy structures corresponds to power-law correlations, i.e. a regime where the main feature is represented by scale-invariance of galaxy clustering. The transition between such a regime and the super-homogeneous one has not been detected yet. This would represent the main observational verification of standard models of galaxy formation, linking the early universe physics, through the statistical properties of the initial density field, to the non-linear structures observed today in galaxy structures. We discuss the measurements of galaxy correlations on large scales, focusing on the main finite size effects which may bias the estimations, and we point out several open issues for the next generation of galaxy surveys.

1. Introduction

In recent times there has been an impressive growth of observational data on cosmological relevant scales, and the development of cosmological models has became a sophisticated activity at the boundaries between different fields in physics and astronomy. The standard scenario of the Big-Bang cosmology is made of different pieces, each of them elaborated to explain a number of observations. The foundation are provided with General Relativity, which allows one to formulate an explanation for the Hubble's law that galaxies distances are linearly proportional to their redshifts; in particular the Friedmann-Robertson-Walker (FRW) solutions of Einstein's field equations, which assume the matter distribution to be a perfectly spatially uniform fluid. In order to explain the highly isotropy observed in the Cosmic Microwave Background Radiation (CMBR) and the apparent flatness of space it has been introduced in the eighties by the theory of inflation (see [1] for a recent critical discussion of these issues).

The highly isotropy of the CMBR together with the observation of galaxy structures, has given rise to the need of cosmological dark matter in the form of a non-baryonic type. While there are several astrophysical observations that in galaxies and galaxy clusters, cosmological arguments related to the growth of structures in an expanding universe pose strong constraints on the amount and type of dark matter. In fact, in contemporary cosmological models the structures observed today at large scales in the distribution of galaxies in the universe are explained by the dynamical evolution of purely self-gravitating matter (dark matter) from an initial state with low amplitude density fluctuations, the latter strongly constrained by satellite observations of the fluctuations in the temperature of the cosmic microwave background radiation (e.g. the satellites COBE [2] and WMAP [3]). Despite the apparent simplicity of the scheme, fundamental theoretical problems remain open and the overall picture is based on the assumption that the main mass component is not only dark but also with very specific properties.

In this theoretical framework one crucial element is represented by the initial conditions (IC) of the matter density field. Models of the early universe [4] predict certain primordial fluctuations in the matter density field, defining their correlation properties and their relation to the present day matter distribution. When gravity start to dominate the dynamical evolution of density fluctuations, which can generally be described by the Vlasov or ``collision-less Boltzmann" equations coupled with the Poisson equation, perturbations are still of very low amplitude. One of the most basic results (see e.g., [5]) about self-gravitating systems, treated using perturbative approaches to the problem (i.e. the fluid limit), is that the amplitude of small fluctuations grows monotonically in time, in a way which is independent of the scale. This linearized treatment breaks down at any given scale when the relative fluctuation at the same scale becomes of order unity, signaling the onset of the ``non-linear" phase of gravitational collapse of the mass in regions of the corresponding size. If the initial velocity dispersion of particles is small, non-linear structures start to develop at small scales first and then structures build up at successively larger scales. Given the finite time from the IC to the present day, the development of non-linear structures is limited in space, i.e., they can not be more extended than the scale at which the linear approach predicts that the density contrast becomes of order unity at the present time. This scale, estimated in current models to be of the order of several Mpc [4], is fixed by the amplitude of initial fluctuations, constrained by the CMBR, by the hypothesized nature of the dominating dark matter component and its correlation properties. Thus in order to explain the existence of contemporary visible structures, such as galaxies and clusters, which otherwise would never condense within the expanding universe, one must make the hypothesis that the "seeds" from which such structures have grown would be made of non-baryonic dark matter, i.e. a type of dark matter that has a very weak coupling with radiation (the so-called cold dark matter - CDM). This is so because dark matter could collapse into the structures we see today, thereafter dragging the much lesser amounts of ordinary matter in afterwards, without leaving large traces into the CMBR radiation, compatible with CMBR anisotropies of order of [2,3]

In this scenario there are therefore several very speculative and ad-hoc hypotheses to explain the growth of structures and observations of large scale galaxy distributions provide important tests for these models. On the one hand the first question concerns the extension of the regime of non-linear clustering and the intrinsic properties of galaxy structures. On the other hand according to the standard scenario, at some

$$\frac{\delta T}{T} \approx 10^{-5}$$

large scales where fluctuations are still today of small amplitude, the imprints of primordial correlations should be preserved and their detection represents a key observation for the validation of the model. Both the linear and non-linear phase are characterized by typical length scales which are fixed by the model (including cosmological parameters) and which represent the first target from the observational point of view.

The power spectrum of density fluctuations at a generic time $t>t^*$ can be written in the form

$$P(k,t) = A \times P(k,t^*) \times g(k,t)$$

where A is the amplitude fixed by the amplitude of CMBR anisotropies and the nature of dark matter, $P(k,t^*)$ is the power spectrum at the time of decoupling between matter and radiation (which we call t^*) and $g(k,t)$ is the time and wavelength dependent function which describes the modification of the primordial power spectrum due to gravitational clustering. In the linear regime, because different modes evolve independently on each other, we simply have [5]

$$g(k,t) = g(t) \propto t^{\frac{4}{3}}$$

Thus the growth factor in the linear regime and the initial amplitude of fluctuations determine the amplitude of perturbation at any time, and in particular they fix the legnth scale of non linear clustering λ_0 which is now estimated, in these theoretical models, to be of the order of 10 Mpc/h [5]. While at scales larger than λ_0 the evolution of perturbation is simply described by the linear amplification, at smaller scales there are non linear effects, usually studied through numerical N-body simulations, which distort the shape of the power spectrum from its initial one.

The power spectrum at the time of decoupling between matter and radiation can be written in the form

$$P(k,t^*) = P_{HZ}(k) \times T^2(k)$$

where the first term describes the power spectrum of the fluctuations generated in the early universe and the second term is the so-called transfer function and which takes into account the modification of the primordial spectrum due to the physical processes acting inside the casual horizon before the equality between matter and radiation [4]. This is chacterized by a typical length scale which is of the order of the size of the horizon at the time of equality between matter and radiation. For a CDM model the transfer function takes the form

$$T^2(k) \propto \frac{1}{[1+(k/k_c)^2]}$$

The length scale $r_c = 1/k_c$ is the first characteristic length scale of this class of models. The main difference between different types of dark matter concerns the functional behavior of this function at large wavenumbers [4]. As mentioned above the second length scale λ_0 is fixed by the amplitude of the initial fluctuations and by the gravitational evolution in the linear regime. Before discussing in more details the problems related to the measurements of these two length scales in galaxy redshift surveys let us now discuss the feature of the primordial power spectrum.

2. The signature of dark matter

The most prominent feature of the initial conditions in the early universe, in standard theoretical models, derived from inflationary mechanisms, is that matter density field presents on large scale super-homogeneous features [6]. This means the following. If one considers the paradigm of uniform distributions, the Poisson process where particles are placed completely randomly in space, the mass fluctuations in a sphere of radius *r* growths as

$$< \Delta M(r)^2 > \propto < M(r) >^{\beta}$$

i.e. like the volume of the sphere and $\beta=1$. A super-homogeneous distribution is a system where the average density is well defined (i.e. it is uniform) and where fluctuations in a sphere grow slower than in the Poisson case, e.g. like $\beta=2/3$. In this case there are the so-called surface fluctuations to differentiate them from Poisson-like volume fluctuations and they are characterized by a sort of long-range stochastic order.

Note that a uniform system with positive correlations present fluctuations which grow faster than Poisson. A simple example of a super-homogeneous distribution is represented by a perfect cubic lattice of particle while a well known system in statistical physics of this kind is the one component plasma [7] (OCP) which is characterized by a dynamics which at thermal equilibrium gives rise to such configurations. The OCP is simply a system of charged point particles interacting through a repulsive $1/r$ potential, in a uniform background which gives overall charge neutrality. Simple modifications of the OCP can produce equilibrium correlations of the kind assumed in the cosmological context [7]. In the cosmological context inflationary models predict a spectrum of fluctuations of this type.

In terms of the normalized mass variance in spheres of radius *R* the super-homogeneous property becomes

$$\sigma^2(R) = \frac{\langle \Delta M^2(R) \rangle}{\langle M(R) \rangle^2} \propto R^{-4}$$

which is the fastest possible decay for discrete or continuous distributions [6]. For comparison in a Poisson distribution, where there are not correlation between particles (or density fluctuations) at all, one simply has

$$\sigma^2(R) \propto R^{-3}$$

By comparing the last to equations one may notice that the long-range order of the distribution corresponds to the presence of mass fluctuations which are depressed with respect to the uncorrelated Poisson case. The reason for this peculiar behavior of primordial density fluctuations is the following. In a FRW cosmology there is a fundamental characteristic length scale, the horizon scale which is simply the distance light can travel from the Big Bang singularity until any given time in the evolution of the Universe, and it grows linearly with time. The Harrison-Zeldovich (H-Z) condition states that the mass variance at the horizon scale is constant. It is possible to show that this is equivalent to assume [6]

$$P(k) \propto k$$

In this situation, one may compute through the Poisson equation the gravitational potential fluctuations induced by matter fluctuations and the H-Z condition corresponds therefore case so that gravitational potential fluctuations become constant as a function of scale. This is a consistency constraint in the framework of FRW cosmology. In fact the FRW is a cosmological solution for a homogeneous Universe, about which fluctuations represent an inhomogeneous perturbation: if density fluctuations obey to a different condition than H-Z-criterion, then the FRW description will always break down in the past or future, as the amplitude of the perturbations become arbitrarily large or small. For this reason the super-homogeneous nature of primordial density field is a fundamental property independently on the nature of dark matter. This is a very strong condition to impose, and it excludes even Poisson processes [6].

Various models of primordial density fields differ for the behavior of the power spectrum at large wave-lengths, i.e. at relatively small scales. For example, for the case the CDM scenario, where elementary non-baryonic dark matter particles have a small velocity dispersion, the power spectrum decays as a power law so that an approximate expression is

For Hot Dark Matter (HDM) models, where the velocity dispersion is large, the power spectrum present an exponential decay at large wave-numbers while it still exhibits the H-Z tail which is in fact the common feature of all density fluctuations compatible with FRW models.

$$P(k,t^*) \propto \frac{k}{[1+(k/k_c)^2]}$$

3. Super-homogeneity in large scale galaxy structures

In terms of correlation function (the Fourier conjugate of the power-spectrum) CDM/HDM models present the following behavior (see Figure 1): it is positive at small scales, it crosses zero at a certain scale r_c and then it is negative approaching zero with a tail which goes as $-1/r^4$ (in the region corresponding to a power spectrum linear in k) [8]. The super-homogeneity (or H-Z) condition says that the volume integral over all space of the correlation function is zero

$$P(k) \propto k \Leftrightarrow \int \xi(r) d^3r = P(0) = 0$$

This means that there is a fine tuned balance between small-scale positive correlations and large-scale negative anti-correlations [6,8]. This is the behavior that one would like to detect in the data in order to confirm inflationary models. Up to now this search has been done through the analysis of the galaxy power spectrum which has to go correspondingly linearly in the wave-number at large scales. However one should consider an additional complication which is usually neglected in the current literature.

In standard models of structure formation galaxies result from a sampling of the underlying CDM density field: for instance one selects (observationally) only the highest fluctuations of the field which would represent the locations where galaxy will eventually form. It has been shown that sampling a super-homogeneous fluctuation field changes the nature of correlations [9]. The reason can be found in the property of super-homogeneity of such a distribution: the sampling necessarily destroys the surface nature of the fluctuations, as it introduces a volume (Poisson-like) term in the mass fluctuations, giving rise to a Poisson-like power spectrum (i.e. $P(k) = const$) on large scales. The ``primordial" form of the PS is thus not apparent in that which one would expect to measure from objects selected in this way (see Figure 2). This conclusion should hold for any generic model of bias and its quantitative importance has to established in any given model [9]. On the other hand one may show [9] that the negative $-1/r^4$ tail in the correlation function does not change under sampling: on large enough scales, where in these models (anti) correlations are small enough, the biased fluctuation field has a correlation function which is linearly amplified with respect to the underlying dark matter correlation function. For this reason the detection of such a negative tail would be the main confirmation of the super-homogeneous character of primordial density fields [8,9].

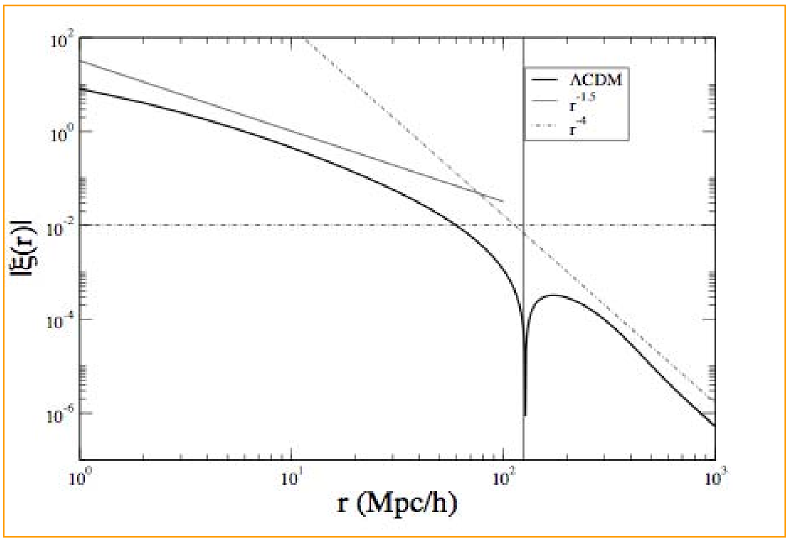

Figure 1: Correlation function predicted in the ΛCDM scenario. The main feature is represented by the negative power-law tail at large scales which corresponds to the a power spectrum which growths linearly as a function of the wave number. The length scale λ_0 which is predicted today to be of the order of 10 Mpc/h corresponds to the scale at which the correlation function is equal to unity. The scale r_c , which in this models is around 100 Mpc/h, corresponds to the scale at which the correlation function crosses zero (from [11]).

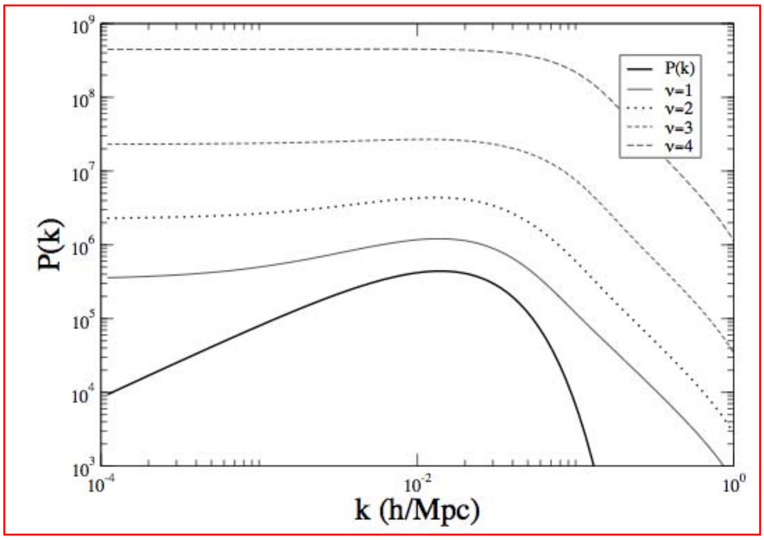

Figure 2: Power-spectrum of the ΛCDM model and of the ones corresponding to different selections of the highest fluctuations of the underlying density field, supposed to be Gaussian for simplicity. Because of the asymmetrical amplification of the correlation function at small and at large scales the condition of super-homogeneity is broken, i.e. the power-spectrum does not show anymore the H-Z linear tail [9,11] (from [11]).

Up to now, no quantitative evidences in this respect have been reported as measurements of galaxy data are not extended enough to reach the region where the correlation function shows the $-1/r^4$ tail. Future surveys, like the complete SDSS catalog [10] may sample this range of scales, but a precise study of the crossover to homogeneity, discretization effects, sampling effects and statistical noise is still needed [11]. At smaller scales instead there have been found evidences for the presence of highly clustered structures which exhibit power law correlations, i.e. a situation which is extremely different from a super-homogeneous distribution

4. Galaxy large-scale structures

In the past twenty years observations have provided several three dimensional maps of galaxy distribution, from which there is a growing evidence of large scale structures. This important discovery has been possible thanks to the advent of large redshift surveys: angular galaxy catalogs, considered in the past, are in fact essentially smooth and structure-less. In the CfA2 catalog[14] which was one of the first maps surveying the local universe, it has been surprisingly observed the giant ``Great Wall'' a filament linking several groups and clusters of galaxies of extension of about 200 Mpc/h and whose size is limited by the sample boundaries. Recently the SDSS project [10] has reveled the existence of structures larger than the Great Wall, as for example the so-called ``Sloan Great Wall'' which is almost double longer than the Great Wall. Nowadays this is the most extended structure ever observed, covering about 400 Mpc/h, and whose size is again limited by the boundaries of the sample [15].

The search for the ``maximum'' size of galaxy structures and voids, beyond which the distribution becomes essentially smooth, is still an open problem. (We refer the interested reader to the review by [23] for a detailed explanations of the different methods used to characterize galaxy structures and for a complete historical perspective on the field) Instead the fact that galaxy structures are strongly irregular and form complex patterns has become a well-established fact. From the theoretical point of view the understating of the statistical characterization of these structures represents the key element to be considered by a physical theory dealing with their formation. The primary questions that such a situation rises are therefore: (i) which is the nature of galaxy structures and (ii) which is the maximum size of structures ? Are the observed large scale structure compatible with the super-homogeneous features predicted by standard cosmological models ? A number of statistical concepts can be used to answer to these questions: in general one wants to characterize n-point correlation properties which are able to capture the main elements of points distributions [8]. The main difficulties when applying the different statistical methods to real galaxy samples concern the control of finite-size, systematic and selection effects. Finite size effects are mostly related to the fact that the samples have a limited volume and it becomes difficult to measure correlations on scales of the order of the size of the samples. Systematic effect are closely related to these ones, as they are due to un-averaged fluctuations on the scale of the samples. Finally selection effects are related to the observational techniques used to construct the catalogs as well as the various assumptions that have to be used in the data analysis (e.g. cosmological parameters, and in various cases assumptions on K-corrections). We discuss the problems related to these effects in the following section.

5. Samples and estimators

In order to briefly describe the present observational situation for what concerns the measurements of galaxy correlations we may recall the results obtained by [16] who have recently measured the conditional average density in a sample of Luminous Red Galaxies (LRG) from a data release of the SDSS (see Figure 3). Such a statistics is very useful to determine correlation properties in the regime of strong clustering and the spatial extension of strong fluctuations in a given sample. This was firstly introduced by [17] and then measured in many samples by [18]. The conditional density gives the average density of points in a spherical volume (or a spherical shell) centered around a galaxy (see e.g. [8]) The results obtained by [16] can be summarized following the discussion in [19]: (i) A simple power-law scaling corresponding to a correlation exponent $\gamma = 1$ gives a very good fit to the data up to at least 20 Mpc/h, over approximately a decade in scale. The power-law behavior of the correlation function is an evidence for the self-similarity of galaxy structures, i.e. of their fractal nature [8]. We note that these results are in good agreement with those obtained by [18] through the analyses of many smaller samples and more recently by [13] in another galaxy survey, the 2dFGRS. (ii) The second important result of [16] is that at larger scales (i.e. r >30 Mpc/h) the conditional density continues to decrease, but less rapidly, until about 70 Mpc/h, above which it seems to

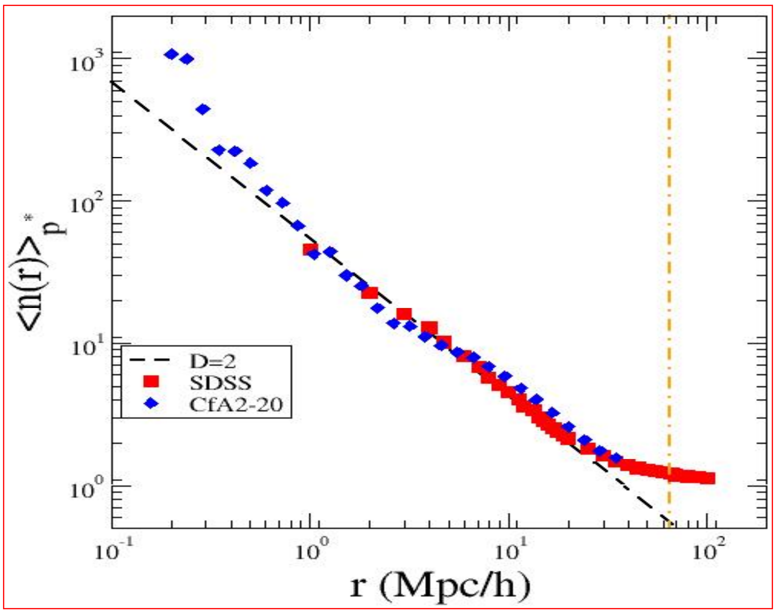

Figure 3: Behavior of the conditional density in the LRG sample of the SDSS [16] (squares) and in the CfA2 catalog [20] (diamonds). A reference line corresponding to a fractal dimension *D=2* is reported as references. The vertical line corresponds to the scale of 70 Mpc/h where [16] claim it occurs the crossover from the strong clustering regime (fractal) to an homogeneous one.

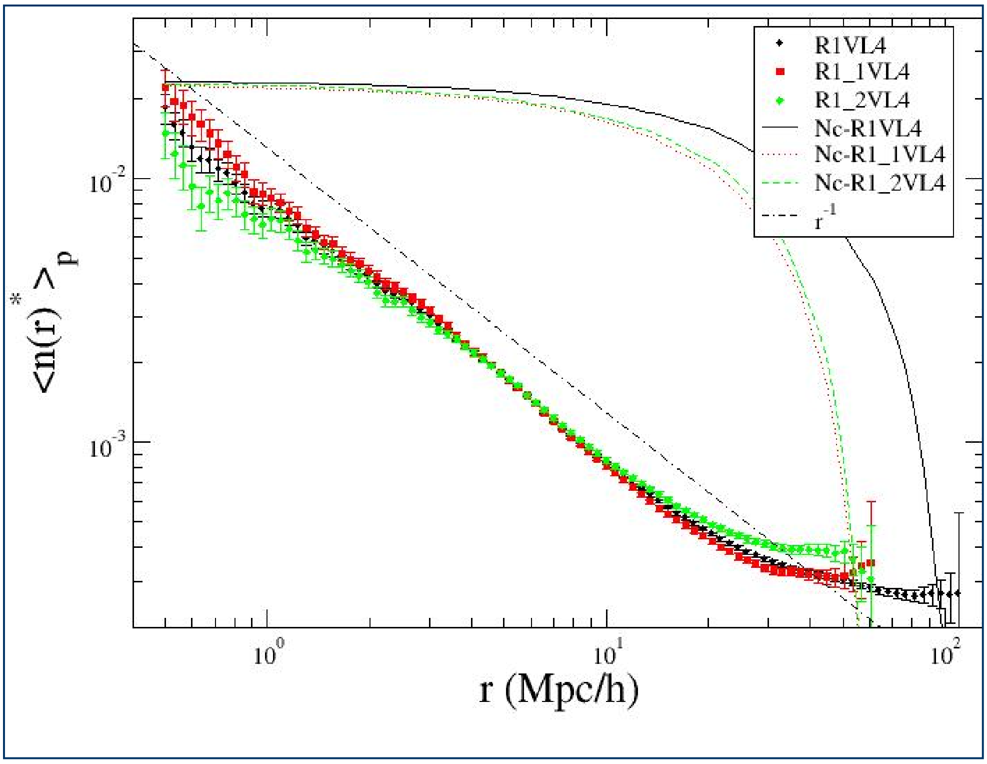

Figure 4: Behavior of the conditional density in the different samples of the SDSS DR4 (from [12]). One may see that the conditional density at large scales shows fluctuations which are larger than the statistical error bars. This is due to the fact that systematic effects on large scales perturb the estimations in different samples. In addition the solid lines show the behavior of the number of points which are considered to determine the average conditional density as a function of scales: one may see that when approaching the sample's boundaries the number of points shows a fast decrease. In order to reach an unambiguous conclusion on the large scale correlation properties in this sample it is needed to control these finite size effects and the systematic fluctuations which they induce.

flatten up to the largest scale probed by the sample, which are of the order of 100 Mpc/h. The transition between the two regimes is slow, in the sense that the conditional density at 20 Mpc/h is about twice the asymptotic mean density. Note that [18] found evidences for a continuation of the small scale power-law to distances of order hundreds of Mpc/h, although with a weaker statistics, which seems to be not confirmed by [16]. However one should consider that in the range of scales [30,100] Mpc/h there are evidences [12,13] for systematic un-averaged fluctuations corresponding to the presence of large scale structures extending up to the boundaries of the present survey, which require a detailed analysis of the problems induced by finite volume effects on the determination of the conditional density (see Figure 4). In addition there are evidences which suggest that in such range of scales the power-law index of the conditional density has a smaller value [12,13]. However future surveys will allow to distinguish between the two possibilities: that a crossover to homogeneity corresponding to a flattening of the conditional density occurs at scales smaller than 100 Mpc/h, or that correlations extend to scales of order 100 Mpc/h (with a smaller exponent $0 < \gamma < 1$). Thus the precise scale of homogeneity, if any, is not yet determined and a detailed analysis of the new coming data will clarify the situation in few years from now.

Finally we note that even if a transition toward a constant value of the conditional density will be finally detected this does not imply that the distribution becomes uncorrelated on larger scales. In fact, this means that structures, beyond the crossover scale, have small amplitude but they can be very well correlated on larger scales. It is then in this situation where the detection of anti-correlations, which as discussed above are predicted by all models of primordial density fields, become the relevant issue to be addressed.

In order to discuss the problems related to the detection of weak anti-correlations, as those predicted by standard cosmological models, let us discuss in some detail the estimators of the conditional density and of the correlation function (see [8,11] for more details). The average number of points in a sphere of radius r from an occupied point of the distribution, in general scales as

$$\langle N(r) \rangle_P = \sum_{i=1}^{M} N_i(r) = B r^D \quad D \le 3$$

where D is the fractal dimension and B is a constant. This is a conditional statistical quantity as it is defined with the condition that the center of the sphere lies on an occupied point. Thus the conditional density can be simply obtained by

$$\langle n(r) \rangle_P = \frac{\langle N(r) \rangle_p}{V(r)} = \frac{3B}{4\pi} r^{D-3}$$

where we find that it has a constant value when the distribution is uniform ($D=3$). In this situation the estimator of the correlation function in a finite spherical sample of radius R_s can be written

$$\xi(r, R_s) = \frac{\langle n(r) \rangle_p}{\langle n(R_s) \rangle_p} - 1 = \frac{D}{3}\left(\frac{r}{R_s}\right)^{D-3} - 1$$

The previous relation shows that the correlation function is strongly dependent on the sample size R_s when the distribution has power-law correlation of the fractal type. Thus the estimation of the conditional density represents a direct way to test the uniformity of a distribution inside a given sample. If this statistical quantity presents power-law correlations as those shown in Figure 3, it is then not surprising that the correlation function shows an amplitude which depends on the sample size (see Figure 5).

As shown in Figure 5 there is another sample-dependent feature of the estimation of the correlation function: the break of the power-law behavior. To understand this feature it is sufficient that one considers a situation in which correlations of weak amplitude extends beyond the size of the sample, as for a CDM-like model. In a spherical sample of radius R the estimator of the sample average is given by

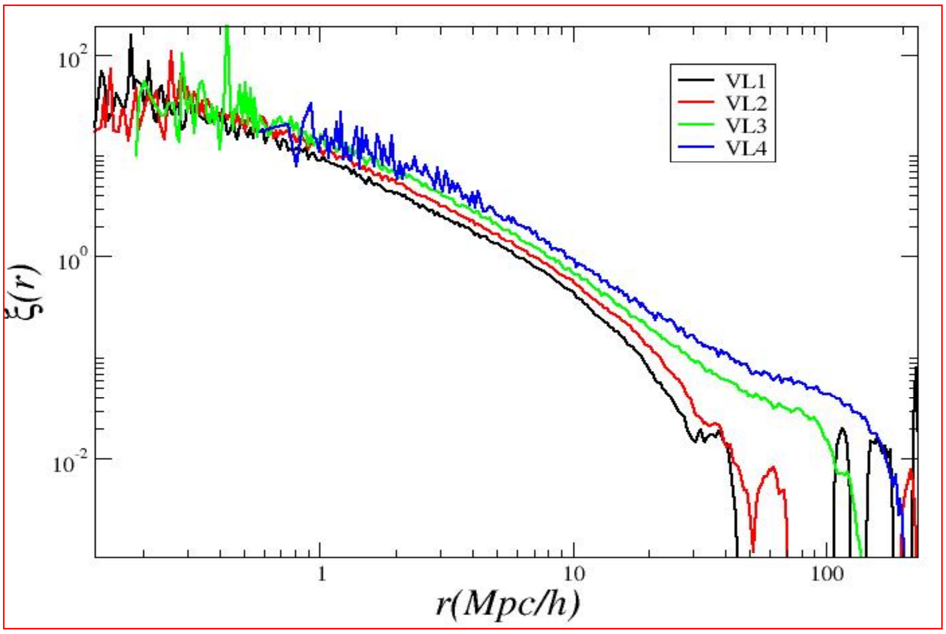

Figure 5: Estimation of the correlation function in different samples (volume limited) of the SDSS survey. One may show that the power law behavior has an amplitude which changes in the different samples and that the break occurs at a scale which also growth with the size of the sample (the size of the sample growths form VL1 to VL4). Both these behaviors maybe due to a finite size effects [21].

$$\overline{n} = \frac{N}{V} = \frac{3}{R^3}\int_0^R \xi(r) r^2 dr$$

where *ξ(r)* represents the ensemble average correlation function. This is clearly a biased estimator of the sample density as the average of an ensemble of realizations in a finite volume gives a different result from the infinite volume limit [8]. We can write the estimated correlation function in a finite sample in terms of the ensemble average correlation function as

$$\overline{\xi}(r) = \frac{1+\xi(r)}{1+\frac{3}{R^3}\int_0^R \xi(r) r^2 dr}$$

We may note that the following condition is always satisfied by the estimator, i.e. for any correlation function and for any volume of the sample, no matter which are the properties of the ensemble average correlation function

$$\int_V \overline{\xi}(r) d^3 r = 0 \quad \forall \overline{\xi}(r), V$$

This integral constraint is determined by the fact that the sample average has been estimated and it gives, in general, a different value from the ensemble average. For this reason the estimator presents a finite-size dependent distortion which affect in a non-trivial way the determinations of large scale correlations. The situation for a CDM-like model is shown in Figure 6 from which one may appreciate the difficulties in the measurements of the large scale correlations predicted by this model in a finite sample [11]. On the other hand the first observational target is to establish whether the break of the correlation function is scale-dependent or not, when the size of the sample considered is larger than r_c = 100 Mpc/h. Preliminary results in the DR6 of the SDSS show that the break of the correlation function is scale-dependent, although a detailed analysis of the statistical and systematic errors (finite size effects) is sill in progress [21].

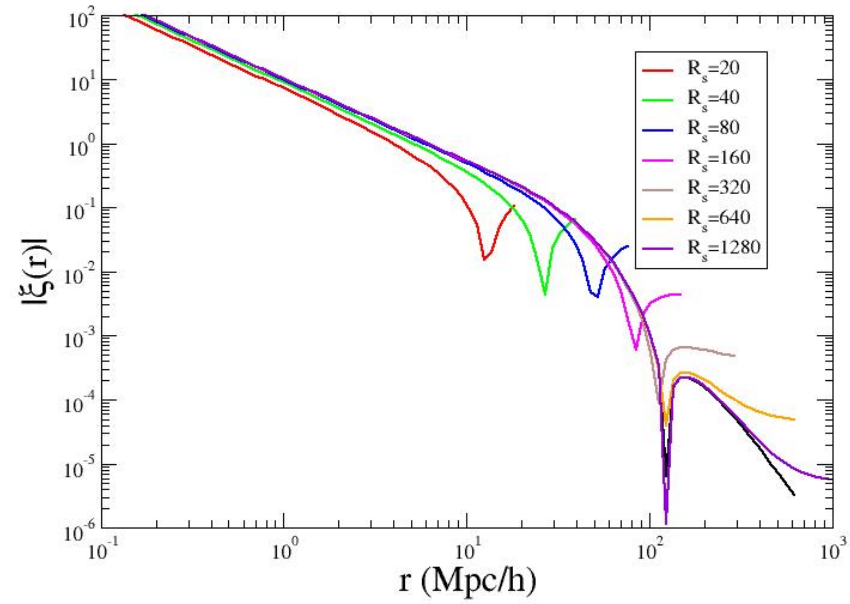

Figure 6: Expected estimation of the correlation function of the CDM model in samples of different sizes. One may note that the detection of the negative tail is very problematic even in a very large sample (from [11]).

6. Conclusions

Statistical properties of primordial density fields show interesting analogies with systems in statistical physics, like the one-component plasma, whose main characteristic is the ordered, or super-homogeneous, nature. In the FRW models the super-homogeneous (or Harrison-Zeldovich) condition arises as a kind of consistency constraint: other, more inhomogeneous, stochastic fluctuations, like the uncorrelated Poisson case, will always break down in the FRW models in the past or future as the amplitude of perturbations in the gravitational potential may become arbitrarily large. We discussed that the observational detection of the super-homogeneous character of the matter density field, through the observation of galaxy distribution or of the CMBR anisotropies, is still lacking. However the main feature of galaxy two-point correlation function is represented by its power-law character in the strongly non-linear region. We stressed that a clear crossover to homogeneity is also not well established in the data, and that the transition from the highly clustered phase to the highly uniform (super-homogeneous) one is the main observational tests for theories of the early universe. For galaxies this should be evidenced as a negative correlation function behaving as $-1/r^4$ at large scales. Future galaxy surveys will provide with samples large enough to clarify these issues. As a final remark it is interesting to note that [22] have studied the distribution of dark matter as it is probed through gravitational lensing. In their maps of the large scale distribution of dark matter they have found a network of filaments extending over several hundreds Mpc. While they claim that their results are consistent with the predictions of gravitationally induced structure formation, the analysis presented above on the correlation properties of theoretical density fields in standard cosmological model implies that there can be a problem in the observations of such structures. Indeed, the correlation function should cross zero at 100 Mpc/h and then it should remain negative up to the largest distance scale, corresponding the super-homogeneous character of the distribution. From this perspective the observation of structures extended over scales larger than 100 Mpc maybe problematic. Indeed the distribution should have negative correlations on such large scales, while the fact that one can identify a structure corresponds to the fact that this is characterized positive correlations. It is then clear that a very challenging problem for future observations will be to measure the two point correlation function of the dark matter density field as probed through gravitational lensing.

References

[1] M. Disney American Scientist, **95**, 35 (2007)
[2] C.L. Bennett, et al., Ap.J., **436**, 423, (1994)
[3] D.N. Spergel et al., Ap.J. Suppl., **148**, (2003) 175
[4] J.A. Peakock Cosmological Physics (Cambridge Astrophysics, Cambridge, 2001)
[5] P.J.E. Peebles, Large Scale Structure of the Universe, (Princeton University Press, Princeton, New Jersey, 1980)
[6] A. Gabrielli, M. Joyce and F. Sylos Labini Phys.Rev.D, 65, 083523 (2002)
[7] A. Gabrielli, B. Jancovici, M. Joyce, J.L. Lebowitz, L. Pietronero and F. Sylos Labini Phys.Rev.D **67**, 043406 (2003)
[8] Gabrielli A., Sylos Labini F., Joyce M., Pietronero L., Statistical physics for cosmic structures, (Springer Verlag, 2004)
[9] R. Durrer, A. Gabrielli, M. Joyce and F. Sylos Labini, Astrophys.J.Lett, **585** , L1-L4 (2003)
[10] D. York, et al., Astron.J., **120**, 1579 (2000)
[11] F. Sylos Labini and N. L. Vasilyev Astron.Astrophys. **477**, 381-395 (2008)
[12] F. Sylos Labini, Vasilyev N., Baryshev Y.V. , A&A **465**, 23 (2007)
[13] N. Vasilyev, Y.V. Baryshev, F. Sylos Labini A&A **447**, 431, (2006)
[14] R. Gott III et al. Astrophys.J. **624** (2005) 463
[15] V. De Lapparent, J. Huchra, M. Geller., et al., Ap.J.,**302**, (1986) L1
[16] D. Hogg, et al. (SDSS Collaboration). ApJ, **624**, 54 (2005)
[17] L. Pietronero, Physica A, **144**, 257 (1987)
[18] F. Sylos Labini, M. Montuori & L. Pietronero Phys Rep, **293**, 66 (1998)
[19] M. Joyce, F. Sylos Labini,M.Montuori and L. Pietronero Astron.Astrophys. **344** 387-392, (1999)
[20] M. Joyce, M. Montuori and F. Sylos Labini ApJL, **514**, L5 (1999)
[21] F. Sylos Labini, N. Vasilyev, Y.V. Baryshev, preprint (2008)
[22] R. Massey et al., Nature **445**, 286 (2007)
[23] Y.V. Barsyhev, P. Teerikorpi Bull. Spec. Astrophys. Obs. Russian Academy of Sciences, **59**, 92 (2006)