

**THE 2-POINT CORRELATION FUNCTION OF
DARK MATTER HALOES IN LIGHT-CONE
COSMOLOGICAL SIMULATION**

Undergraduate Thesis

By

AMVROSIADIS ARISTIDIS ¹

Thesis Advisor:

Prof. MANOLIS PLIONIS

¹*Physics Department, Aristotle University of Thessaloniki,
Thessaloniki Greece 54124*

27 May 2014

Contents

| | | |
|----------|---|-----------|
| 1 | The Morphology of the Large Scale Structure | 4 |
| 1.1 | The Origin of the primordial Density fluctuations | 5 |
| 1.2 | Gravitational instability of Density fluctuations | 5 |
| 1.3 | Linear Perturbation Theory | 7 |
| 2 | The Two-Point correlation Function | 10 |
| 2.1 | General description of what the correlation function represents | 10 |
| 2.2 | Method for calculating $\xi(r)$ or $w(\theta)$ | 12 |
| 3 | Evolution of Clustering | 14 |
| 4 | A brief review of the Cosmological simulations | 17 |
| 4.1 | An introduction to N-body simulations | 17 |
| 4.1.1 | Particle-Mesh Method (PM) | 19 |
| 4.2 | Light-cone Cosmological Simulations | 20 |
| 4.2.1 | Friends-of-Friends (FoF) method for halo identification | 21 |
| 5 | Simulation Halo Data Analysis | 22 |
| 5.1 | Introduction | 22 |
| 5.1.1 | Presenting the cosmological models | 22 |
| 5.1.2 | Comparing the model redshift distributions | 24 |
| 5.2 | Computing the angular Two-point Correlation function | 25 |
| 5.3 | Computing the spatial Two-point Correlation function | 30 |
| 5.4 | Evaluating the index of the growth factor $(n + 2)$ | 33 |
| 6 | Conclusions | 38 |
| 7 | REFERENCES | 39 |

Abstract

In this thesis project is presented an estimation of the angular and spatial two-point correlation function of dark matter halos for three different cosmological models, the Λ CDM, SUCDM and RPCDM. I use the corresponding dark matter halo catalogue from the DEUS simulation light-cones, which were constructed from N-body simulations with boxlength $2592 [Mpc/h]$, consisting of 1024^3 dark matter particles. Subsequently, I fit an evolution relation, between the correlation length $r_0(z)$ and the growth factor $D(z)$ taking into account also the bias evolution $b(z)$, for the corresponding data of the Λ CDM model. The aim of the latter procedure is to investigate which value of the modified index ($n+2$) of the growth factor corresponds better to the evolution of the halo correlation function at different redshifts.

1 The Morphology of the Large Scale Structure

A huge progress in the field of Large-Scale structure has been realized after the construction of redshift surveys of galaxies and cluster of galaxies; providing the spatial distribution of large regions of the sky and 3-dimensional maps containing thousand of galaxies (e.g The Sloan Digital Sky Survey (SDSS), or the Two-Degree Field Galaxy Redshift Survey (2dFGRS)) and Two Micron All-sky Survey (2MASS), see [24], [25] and [26] for the original papers).

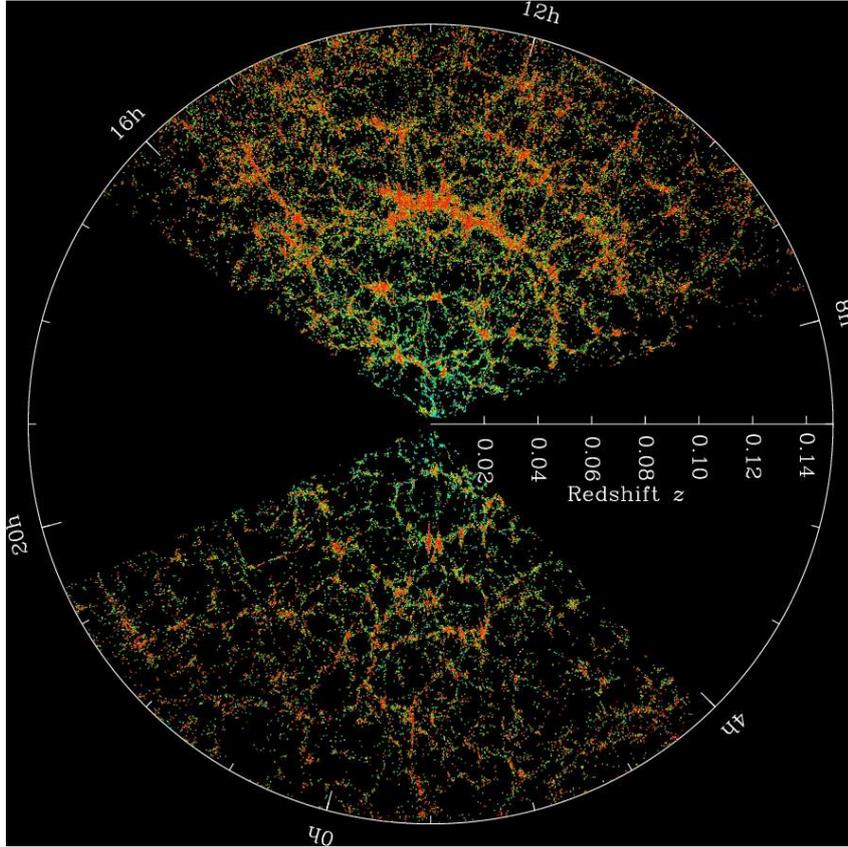


Figure 1: Slices through the three-dimensional map of the distribution of galaxies from the Sloan Digital Sky Survey (SDSS)

The distribution of galaxies on the sky is not uniform or random, but rather they form pairs, groups and clusters of galaxies with large (almost spherical) voids between them. Furthermore clusters of galaxies themselves are not distributed uniformly in the space, but their positions are correlated, grouped together to form even larger structures (superclusters) in the form of filaments, walls etc.

Even though these redshift surveys revealed that the universe is not homogeneous in scales of tens Megaparsecs, there is no evidence of structures with linear dimensions greater than a few hundred Megaparsecs. So, the universe can be considered homogeneous if averaged over scales of the latter scales. This property of the universe is fundamental and it is referred in the literature as the **Cosmological principle**, which

states that the universe is spatially homogeneous and isotropic on Large-Scales. Homogeneity implies that all *comoving* observers see identical properties, while isotropy implies that there is no preferred direction in the universe.

The most unquestionable proof that the universe is homogenous and isotropic was the discovery of the highly isotropic Cosmic Microwave Background radiation (**CMB**). Since this radiation dates from the epoch of decoupling ($z \sim 1100$) it constitutes direct evidence that the universe was homogeneous up to that epoch. However, even in the Cosmic Microwave Background ,with the emergence of high resolution data (especially from WMAP and later on from Planck satellites), there is a clear evidence of temperature fluctuations with a relative amplitude of $\Delta T/T \sim 10^{-5}$.

1.1 The Origin of the primordial Density fluctuations

The morphology of the Cosmic Web, as we know it today, represents the evolution of the primordial density fluctuation observed in the CMB (since fluctuations in density are proportional to temperature fluctuations when referred to the epoch of $z \sim 1100$). The origin of such fluctuations is not yet fully understood but it can be explained in the framework of inflationary models. In general inflation is a time period in the evolution of the universe, when a rapid expansion occurred (an exponential expansion within a time period measured in Planck time units). The Large-Scale Structure of the universe today should correspond to microscopic scales during the inflation period. In these tiny regions quantum zero point fluctuations occurred, which eventually generated the observable universe (see [2]).

In an abstract way the origin of the fluctuations can be explained as follows: During inflation, at each point in space the inflation field (**inflaton**) has a value. As inflation proceeds, quantum fluctuations of the inflaton are dragged from microscopic to macroscopic scales, during the exponential expansion of the universe. The result of this process, is the generation of the observable fluctuations in the CMB.

1.2 Gravitational instability of Density fluctuations

As discussed above the anisotropy in the CMB suggests that the spatial density fluctuations (inhomogeneities) must have had very small amplitudes at the epoch of recombination. On the other hand, numerous observations show that the universe (in the present epoch) in small-scales is highly inhomogeneous.

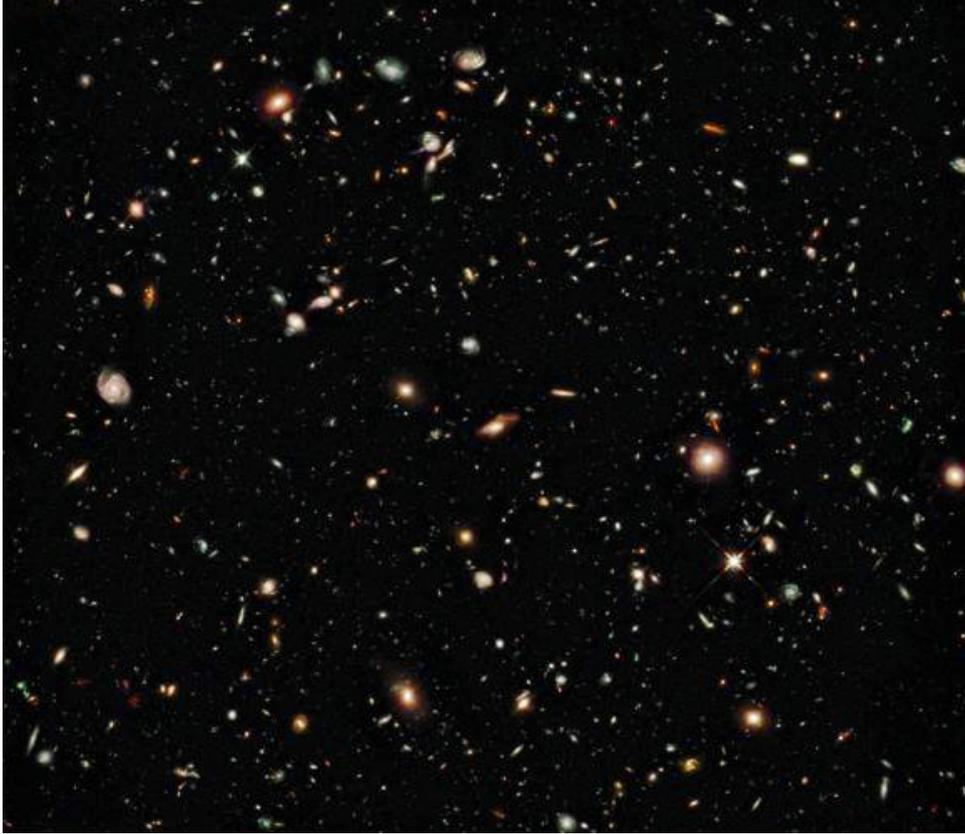


Figure 2: 'The Hubble Deep Field' image with the Wide Field camera 3 taken from the Hubble Telescope

The above picture of the Hubble Telescope clearly shows these density inhomogeneities. Obviously, during the evolution of the universe the latter became more and more inhomogeneous. The density fluctuation field at a region of the universe (with \mathbf{r} being the comoving position of that region) in relation to the mean cosmic matter density at time (t) is given by:

$$\delta(\mathbf{r}, t) = \frac{\rho(\mathbf{r}, t) - \tilde{\rho}(t)}{\tilde{\rho}(t)} \quad (1)$$

The dynamics of the expanding cosmos is controlled by the gravitational field of the average matter density, while the density fluctuations can be considered to generate an additional gravitational field. Considering a slightly overdense region of the universe (in a fixed time), where $\rho > \tilde{\rho}$ and proportionally $\delta > 0$, it is obvious that the gravitational field produced by that region is stronger than the average gravitational field. Therefore this region will expand slower than the average Hubble expansion and the difference in the expanding pace will cause the density to decrease slower than the average density of the universe. So, as time progresses the slightly overdense regions become ever more overdense due to their self-gravity. The opposite situation occurs in the underdense regions, where the density contrast decreases with time. Of course, since the minimum value of the density is zero, the condition $\delta \geq -1$ holds for δ . This scenario is called

gravitational instability and is considered to be the cause of the creation of the Large-Scale Structure from the primordial density fluctuations. The galaxies and cluster of galaxies represent the overdense regions of this scenario, while the underdense regions correspond to the large, nearly empty, voids in the universe (see also [1],[3] and [4]).

1.3 Liner Perturbation Theory

The evolution of structure in the universe can be understood in the framework of the above mechanism. A quantitative description of this model requires some physical approximations in order to be done. This description will be concentrated in length-scales, which are substantially smaller than the Hubble radius¹ in order for the Newtonian approximation to be valid (for fluctuations of length-scale $L \sim 2\pi/k \gtrsim R_H$, Newtonian perturbation theory is not valid). For simplicity, it is assumed that the universe consists of dust only, in order for the fluid approximation to be valid. The equations in the comoving coordinates² \mathbf{x} ($\mathbf{r} = \alpha(t) \mathbf{x}$) are given below:

The velocity field (\mathbf{v}) can be written in the form:

$$\mathbf{v}(\mathbf{r}, t) = \frac{\dot{\alpha}}{\alpha} \mathbf{r} + \mathbf{u}(\mathbf{x}, t) \quad (2)$$

where the first term represents the homogeneous Hubble expansion, whereas the last term describes the deviations from the latter and for this reason is called 'peculiar' velocity.

The perturbed continuity equation, which expresses that matter is conserved :

$$\frac{\partial \delta}{\partial t} + \frac{1}{\alpha} \nabla \cdot [(1 + \delta) \mathbf{u}] = 0 \quad (3)$$

with the gravitational potential Φ takes the following form:

$$\Phi(\mathbf{r}, t) = \frac{2\pi}{3} G \tilde{\rho}(t) |\mathbf{r}|^2 + \phi(\mathbf{x}, t) \quad (4)$$

The relation between the density field and the gravitational potential ϕ is established through the Poisson-Newton equation,

$$\nabla^2 \phi(\mathbf{x}, t) = 4\pi G \alpha^2(t) \tilde{\rho}(t) \delta(\mathbf{x}, t) = \frac{3}{2} H_0^2 \Omega_m \frac{\delta(\mathbf{x}, t)}{\alpha(t)} \quad (5)$$

While the Euler equation is given below,

¹Physical interactions can take place only on scales that are smaller than the Hubble radius R_H

²Considering a homogeneous universe which is radially expanding, we choose a point at time $t = t_0$ and introduce a coordinate system \mathbf{x} with its origin located in the center of the sphere. Now a particle inside the sphere with $\mathbf{r}(t_0) = \mathbf{x}$ will be located at some other time t at the position $\mathbf{r}(t)$, due to the expansion of the sphere.

$$\frac{\partial \mathbf{u}}{\partial t} + \frac{\mathbf{u} \cdot \nabla}{\alpha} \mathbf{u} + \frac{\dot{\alpha}}{\alpha} \mathbf{u} = -\frac{1}{\tilde{\rho}\alpha} \nabla P - \frac{1}{\alpha} \nabla \phi \quad (6)$$

In the linear regime (as long as the condition $|\delta| \ll 1$ holds), with the approximation that the universe contains only dust (thus $P = 0$), from the linearization of the above equations results the differential equation for the evolution of the density contrast δ , which allows the tracking of the growth of density perturbations, and takes the form:

$$\frac{\partial^2 \delta}{\partial t^2} + \frac{2\dot{\alpha}}{\alpha} \frac{\partial \delta}{\partial t} = 4\pi G \tilde{\rho} \delta \quad (7)$$

The solution to the above equation takes the following form,

$$\delta(\mathbf{x}, t) = D_+(t)\tilde{\delta}(\mathbf{x}) + D_-(t)\tilde{\delta}(\mathbf{x}) \quad (8)$$

which indicates that the spatial shape of the density fluctuations is frozen in the comoving coordinates, only their amplitude increases. The solution $D_-(t)$ decreases with time and supposing that both functional dependences were present, at some early time, the increasing solution will dominate at later times, thus the term $D_-(t)$ can be neglected and consider only the growing solution.³

$$\delta(\mathbf{x}, t) = D_+(t)\tilde{\delta}(\mathbf{x}) \quad (9)$$

The linear growth factor of density perturbations can be computed,

$$D(\alpha) = \frac{5\Omega_{m,0}H_0^2}{2} H(\alpha) \int_0^\alpha \frac{d\alpha'}{\alpha'^2 H^3(\alpha')} \quad (10)$$

However this approximation implies that the distribution of density fluctuations $\delta_0(\mathbf{x})$ today, would be homogeneous which is not valid. If the condition $|\delta| \ll 1$ does not hold the terms which were neglected in the linear approximation, must be included in the analysis (see [1] and [3]).

The non-linear structures, observed today, can not be interpreted by the growth of baryonic density fluctuations only. This contradiction can be resolved by taking dark matter, which is the dominant component of matter, into account. Perturbations of dark matter started in the Radiation Epoch (but mostly after equipartition) and by the time baryonic matter decoupled from radiation (Recombination Epoch) , dark matter had already formed potential wells in which baryonic matter 'fell'.

This observation is crucial to the understanding of the Large-Scale structure formation. In the current accepted cosmological model, the 'Concordance' or Λ CDM model, ordinary matter accounts for only 4.5 percent of the total matter in the universe. Thus the dynamics of the cosmic structure formation follows the dominant non-baryonic, cold dark matter (**CDM**) component. The nature of the cold dark matter particle is yet

³the growth factor $D(t)$ is usually normalized such as $D(t_0) = 1$

unknown, but there is a plethora of weakly-interacting massive particles candidates. 'Cold' means that these particles have rather small thermal velocities, which allows the formation of very small structures and is consistent with the hierarchy of structure formation. The latter means that dark matter halos form hierarchically by the aggregation, via gravitational interactions, of small collapsed structures that merge to form larger one, eventually forming clusters of galaxies. The existence of dark matter has been proven by many observational methods and it is also necessary in order to form cosmic structures without violating the observed amplitude of the **CMB** temperature fluctuations.

2 The Two-Point correlation Function

2.1 General description of what the correlation function represents

The previous sections develop the ideas of the transition from a homogeneous universe to a universe which consist of large inhomogeneities. The current section examines the question of how to describe an inhomogeneous universe quantitatively. Such a description can only be statistical in nature, since the primordial density field was a stochastic Gaussian random field (i.e it is impossible to predict that in a certain distance from our galaxy, another elliptical galaxy would form, in the framework of a cosmological model) That is why the analysis is focused on the statistical properties of the universe, which can be computed for every cosmological model and derive a conclusion about which model best describes the observed universe.

Such a statistical methodology for the study of the Large-scale distribution of objects in the universe, is provided by the ***Two-point Correlation Function***.

As mentioned, large redshift surveys have revealed a wealth of information regarding the clustering of matter. The conclusions for studying these surveys suggest that galaxies and cluster of galaxies are not randomly distributed in space, but rather they are strongly correlated. This means that the possibility of finding a galaxy in the vicinity of another is larger, than on average.

This phenomenon can be described as follows: Suppose that galaxies are distributed uniformly through space and \bar{n} is the average number density of the galaxies, then the probability of finding a galaxy in the volume element dV around a point \mathbf{x} is,

$$\delta P_1 = \bar{n} dV \tag{11}$$

which is independent of \mathbf{x} , with the assumption that the universe is statistically homogeneous.

Now assuming that \mathbf{x} and \mathbf{y} are the coordinate terms of two volume elements dV_1 and dV_2 respectively, then the possibility of finding a galaxy at the location \mathbf{x} and at the same time a galaxy exists in the location \mathbf{y} can be expressed as the following:

$$\delta P_2 = \bar{n}^2 dV_1 dV_2 [1 + \xi_g(\mathbf{x}, \mathbf{y})] , \tag{12}$$

with $\xi_g(\mathbf{x}, \mathbf{y})$ the two-point correlation function defined as the joint probability of finding an object (e.g. a galaxy) in each of the two volume elements dV_1 and dV_2 . The correlation function can also be defined as the conditional probability of finding an object in the volume element δV , considering than another object has been randomly selected from the ensemble

$$\delta P = \bar{n} dV [1 + \xi(r)]$$

If the distribution of objects is uncorrelated then the latter possibility would be $\delta P = \bar{n}^2 dV_1 dV_2 = \delta P_1 \cdot \delta P_2$

Since the universe is considered to be statistically homogeneous then the functional form of the correlation function $\xi_g(\mathbf{x}, \mathbf{y})$ can only depend on the modulus of the separation $r = |\mathbf{x} - \mathbf{y}|$ and not of the direction of $\mathbf{r} = \mathbf{x} - \mathbf{y}$, because the universe is also isotropic.

The definition of the process of finding the correlation function, in which a discrete ensemble of objects is assumed, is called a **Point Process**. It should be noted that if $\xi(r) > 0$ the objects in the distribution are considered to be correlated, conversely if $\xi(r) < 0$ then they are considered anticorrelated.

In angular space (i.e. the projection of the distribution of galaxies in the unit sphere) the clustering of objects is described by the so-called angular two-point correlation function which is defined as the conditional probability:

$$\delta P = N[1 + w(\theta)]d\Omega \quad (13)$$

, where N is the surface density and $w(\theta)$ the excess probability over random of finding an object in a solid angle $d\Omega$ under the condition that an object already been selected.

Another way to define the two-point correlation function is by the use of a continuous density field $\rho(\mathbf{x})$, **Poisson model**. Defining the average density as $\tilde{\rho} = \langle \rho(\mathbf{x}) \rangle$ and the density contrast $\delta(\mathbf{x})$ as,

$$\delta(\mathbf{x}) = \frac{\rho(\mathbf{x}) - \tilde{\rho}}{\tilde{\rho}} \quad (14)$$

Then the correlation function is defined as:

$$\begin{aligned} \xi(r) &= \langle \delta(\mathbf{x})\delta(\mathbf{x} + \mathbf{r}) \rangle = \left\langle \left(\frac{\rho(\mathbf{x}) - \tilde{\rho}}{\tilde{\rho}} \right) \cdot \left(\frac{\rho(\mathbf{x} + \mathbf{r}) - \tilde{\rho}}{\tilde{\rho}} \right) \right\rangle \\ &= \frac{\langle \rho(\mathbf{x})\rho(\mathbf{x} + \mathbf{r}) \rangle - \tilde{\rho}\langle \rho(\mathbf{x}) \rangle - \tilde{\rho}\langle \rho(\mathbf{x} + \mathbf{r}) \rangle + \tilde{\rho}^2}{\tilde{\rho}^2} \\ &= \frac{\langle \rho(\mathbf{x})\rho(\mathbf{x} + \mathbf{r}) \rangle - \tilde{\rho}^2}{\tilde{\rho}^2} \end{aligned}$$

2.2 Method for calculating $\xi(r)$ or $w(\theta)$

Estimating the correlation function is not always an easy task, since several survey boundary effects must be taken into account as well as other observable systematics, which are introduced by the very nature of the redshift surveys and human biases.

- The most important feature of sky catalogues is that they do not cover the whole region of the sky mainly due the limited capabilities of the research equipment or due to galactic absorption at low latitudes (i.e. the huge amount of dust of the galactic disk causes the absorption of the light originating from distant galaxies along the line-of-sight of the disk). For this reason objects near the boundaries of the surveys will not have the same number of neighbors at some angular or spatial separations (i.e. since the region of any survey does not correspond usually to the whole sky, the objects located near the survey boundaries all have less neighboring objects than they would have in a whole sky survey) , which causes biases to the average correlation function.
- Another feature which appears in three dimensional catalogues is the lack of objects as we move to higher distances (redshifts), due to the limiting magnitude of the survey.

One method of overcoming the boundary effects for a given separation is to use only a portion of the data sets, which includes objects that have a distance from the boundaries less than the specific separation. But, decreasing the amount of data used in the analysis may lead to low significance results. However using an adequately constructed random catalogue of points, based on *Monte-Carlo* simulations, one can overcome the above mentioned problems while using all the available data. This works as follows:

If in a volume V , the total number of objects is N , then

$$N = \bar{n} \cdot V$$

where \bar{n} in the average number of objects per unit volume while the expected number of objects in the element volume dV_i , when the distribution is clustered, is given by:

$$\begin{aligned} n_i &= \bar{n}[1 + \xi(r)]dV_i \rightarrow \\ \rightarrow \xi(r) + 1 &= \frac{n_i}{\bar{n}dV_i} = \frac{n_i V}{N dV_i} \end{aligned}$$

The calculation of $\xi(r)$ for a particular separation r (i.e. $r - \delta r/2 < |\mathbf{r}| < r + \delta r/2$) is performed as follows,

$$\xi(r) + 1 = \frac{\frac{V}{N} \sum_{i=1}^N n_i}{\sum_{i=1}^N dV_i}$$

and due to the fact that survey boundaries are complex, thus $\sum dV = N\langle dV \rangle$ the above equation takes the following form,

$$\xi(r) + 1 = \frac{2n_0V}{N^2\langle dV \rangle} \quad (15)$$

In the latter equation n_0 stands for the number of pairs for the particular separation range. Supposing N_r points are randomly placed in the same sample volume V , then by definition $\xi(r) = 0$ since the distribution of random points is uniform.

Taking into account these conditions, and supposing now that n_r is the number of random pairs for the same separation range, the above equation is reduced to,

$$\langle dV \rangle = \frac{2n_rV}{N_r^2}$$

and by substituting the latter into (12), we obtain:

$$\xi(r) + 1 = \frac{n_0}{N^2} \frac{N_r^2}{n_r} \quad (16)$$

The above formula can be used to derive the spatial correlation function from a data set by computing the number of data and random pairs for a certain common range of separations. N and N_r is the number of data and random points of the corresponding samples, respectively⁴

It should be noted that this is the crudest estimation of the 2-point correlation function. In the literature there are a variety of estimators such as (16) which give significantly better results and will be examined later (see [8] and [9]).

⁴Choosing the number of data and random points to be equal the formula (16) takes the form,

$$\xi(r) + 1 = \frac{n_0}{n_r}$$

3 Evolution of Clustering

Extragalactic objects (i.e. galaxies and cluster of galaxies) trace the underlying dark matter distribution in a way that is typically described with a multiplicative linear factor known as the bias $b(z)$. Such a biasing is considered to be statistical in nature, with galaxies and clusters being identified as high peaks of an underlying, initially Gaussian, random density field (see [12]). The linear and scale independent bias factor is defined as the ratio of the mass-tracer fluctuations δ_h (in this section dark matter haloes are used as a tracer), to those of the underlying mass δ_m .

$$\delta_h = b \delta_m \quad (17)$$

where the bias factor is considered to be a function of redshift $b = b(z)$.

Since the two-point correlation function in a continuous density field is defined as,

$$\xi_h(\mathbf{r}, z) = \langle \delta_h(\mathbf{x}, z) \delta_h(\mathbf{x} + \mathbf{r}, z) \rangle \quad (18)$$

Then by inserting equation (17) into the latter,

$$\xi_h(\mathbf{r}, z) = b^2 \langle \delta_m(\mathbf{x}, z) \delta_m(\mathbf{x} + \mathbf{r}, z) \rangle \quad (19)$$

As a result of the solution of the second-order differential equation for the density contrast (8), the latter can be written as a proportionality of the growth factor with the density perturbations.

$$\begin{aligned} \delta_m(\mathbf{x}, z) &\propto D(z) \\ \delta_m(\mathbf{x}, 0) &\propto D(0) \end{aligned}$$

which results to the equation,

$$\delta_m(\mathbf{x}, z) = \frac{D(z)}{D(0)} \delta_m(\mathbf{x}, 0) \quad (20)$$

From equation (17) the density contrast of the dark matter halos can be expressed in terms of the bias factor and the density contrast of the mass,

$$\begin{aligned} \delta_h(\mathbf{x}, z) &= b(z) \delta_m(\mathbf{x}, z) \\ \delta_h(\mathbf{x}, 0) &= b(0) \delta_m(\mathbf{x}, 0) \end{aligned}$$

So,

$$\begin{aligned}
\delta_h(\mathbf{x}, z) &= b(z) \delta_m(\mathbf{x}, z) \\
&= \frac{D(z)}{D(0)} b(z) \delta_m(\mathbf{x}, 0) \\
&= \frac{D(z)}{D(0)} \frac{b(z)}{b(0)} \delta_h(\mathbf{x}, 0)
\end{aligned}$$

Then from equation (19) the redshift evolution of the correlation function with respect to the correlation function at the present time, is presented by the formula below .

$$\xi_h(\mathbf{r}, z) = \left(\frac{D(z)}{D(0)} \right)^2 \left(\frac{b(z)}{b(0)} \right)^2 \xi_h(\mathbf{r}, 0) \quad (21)$$

Substituting the functional form of the correlation function $\xi(r, z) = r_0^\gamma(z) r^{-\gamma}(z)$, in the above equation:

$$r_0(z) = \left[\left(\frac{D(z)}{D(0)} \right)^2 \left(\frac{b(z)}{b(0)} \right)^2 \right]^{1/\gamma} r_0(0) \quad (22)$$

The growth factor of density perturbations $D(z)$ in the above equation has been computed in the linear regime, however the scales involved in the estimation of the two-point correlation function are at least in the mildly non-linear regime and thus we expect a modification of the above formula (see Peacock book).

The modification of the formula, in order to be valid in the mildly non-linear regime of the evolution of the growth factor, is the following,

$$r_0(z) = \left[\left(\frac{D(z)}{D(0)} \right)^{2+n} \left(\frac{b(z)}{b(0)} \right)^2 \right]^{1/\gamma} r_0(0) \quad (23)$$

In the above formula the growth factor of the density perturbations with respect to the redshift has the form,

$$D(z) = \frac{5}{2} \Omega_m E(z) \int_z^\infty \frac{1+z'}{E^3(z')} dz' \quad (24)$$

The function $E(z)$ depends on the Hubble constant in the present epoch H_0 , as well as on the Hubble function $H(z)$ (see [11]) and for a flat cosmology $\Omega_k = 0$ it takes the form:

$$E(z) = [\Omega_{m,0}(1+z)^3 + \Omega_\Lambda]^{1/2} \quad (25)$$

where $\Omega_i = \rho_i/\rho_c$ are the parameterized 'fluid' densities with respect to the total density of the universe (Ω_m for the total matter and Ω_Λ for the Cosmological constant).

For the evolution of the bias factor, the formula used is (see [9] for more information):

$$b(z) = 1 + \frac{b_0 + 1}{D(z)} + C_2 \frac{J(z)}{D(z)} \quad (26)$$

where,

$$J(z) = \int_0^z \frac{1 + z'}{E(z') dz'} \quad (27)$$

and with the constants being functions of the halo mass (see [10])

$$b_0(M_h) = C_b \left[1 + \left(\frac{M_h}{1 + 10^{14} h^{-1} M_\odot} \right)^B \right], \quad C_b = 0,857 \pm 0.021, \quad B = 0.55 \pm 0.06$$

and ,

$$C_2(M_h) = C_\mu \left[\frac{M_h}{10^{14} h^{-1} M_\odot} \right]^\mu \quad C_\mu = 1.105 \pm 0.018, \quad \mu = 0,255 \pm 0.005$$

4 A brief review of the Cosmological simulations

4.1 An introduction to N-body simulations

Observations reveal that the present universe is populated by large structures such as galaxies and clusters of galaxies. The current picture of structure formation suggests that these non-linear structures are formed from small fluctuations in the density field, via gravitational instability (i.e. gravitational amplification of the small perturbations, which existed in the early universe). In the non-linear regime the lack of analytic methods for tracking down the growth of fluctuations, has led to the development of N-body simulations which study the clustering of matter and the formation of galaxies.

Our understanding of the universe has substantially increased over the past years with the use of Cosmological N-body simulations, which model the growth of structure in the universe deeply within the non-linear regime of growth interactions. These cosmological simulations are an essential tool to study the gravitational instability in the expanding background as well as the formation of large-scale structures we observe today. A characteristic feature of such simulations, is driven from the fact that the universe is very large and can not be incorporated as a whole in the simulation. That is the reason why large chunks of it are simulated using the appropriate periodic boundary conditions to model the universe.

Since dark matter is dominant in terms of abundance over the baryonic matter, the dynamics of the universe is governed by it . It is known that fluctuations on small scales in the collisionless dark matter component grows via gravitational instability from the Radiation epoch, unlike fluctuations in the baryonic component, which start to grow only after the decoupling of radiation and baryonic matter (i.e. Recombination Epoch). The baryonic matter after the recombination Epoch “falls” in the potential wells of dark matter, which will govern in the evolution of fluctuations. This is the reason why dark matter particles are used in almost all cosmological simulations, which model the formation of the Large-Scale structure.

Another feature suggested by observations is that the universe is not dominated by relativistic matter. This property can allow the fluid approximation of the dark matter particles to take place in the simulation and therefore the motion of the dark matter particles can be described by the Newtonian equations. These equations can be written, for scales that are smaller than the Hubble radius ($\ll d_h = c/H_0$) and are valid for non-relativistic matter ($u \ll c$):

$$\begin{aligned}\delta(\mathbf{r}, t) &= \frac{\rho(\mathbf{r}, t) - \tilde{\rho}(t)}{\tilde{\rho}(t)} \\ \nabla^2 \phi &= 4\pi G \tilde{\rho}(t) \alpha^2(t) \delta_{DM} = \frac{3}{2} H_0^2 \Omega_0 \frac{\delta}{\alpha} \\ \ddot{\mathbf{x}} + 2\frac{\ddot{\alpha}}{\alpha} \dot{\mathbf{x}} &= -\frac{1}{\alpha^2} \nabla \phi\end{aligned}$$

The N-body codes consist of two basic modules: one part computes the force field for a given configuration of particles and the other one moves the particles according to this force field. These two are called in each step of the simulation in order to ensure that the force field and the particle trajectories evolve in a self consistent manner. The other basic feature of every simulation is the set of initial conditions.

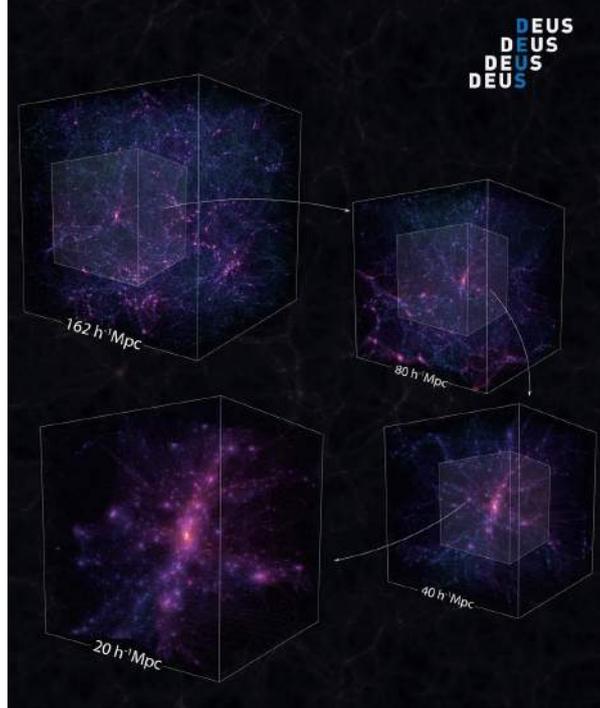


Figure 3 : Zooming into the Dark Energy Universe Simulation (DEUS consortium) of the Λ CDM model simulation with 1024^3 particles and $162h^{-1}Mpc$ cubic length (at $z=0$)

In every N-body simulation the particles populate a specific region of volume (V) and assuming that N_p particles are being used for the description of the density field then the following statements are the physical requirements that must be taken into account in the development of any N-body code (see [13],[14],[15],[16]):

- Any simulation cube of volume (V) is just a portion of the whole universe and cannot be treated as an isolated region. Hypothetically the space outside the simulation cube represents the rest of the universe (which is true of course), so the influence exerted in the cube must be incorporated in the simulation. The only viable solution are the periodic boundary condition (i.e. the walls of the periodic cub satisfy the boundary conditions of the simulation).²
- Another requirement is that the average density of the cube must be equal to the average density of the universe (i.e. averaged over the length of the cube the density must equal the average density of the universe). Regardless of the

²The periodicity of the boundary condition implies that the universe is homogeneous if averaged over large scales, so the gravitational force from all directions is the same. Without the boundary conditions the particles would collapse in the center of the cube, in that case the simulation studies a different procedure (thus i can not be considered as a Cosmological Simulation)

cosmological model, this requirement ensures that perturbations averaged over the boxlength will be negligible.

- The mass of each particle of the simulation depends on the mass of the structure of interest and must be significantly less than the mass of the latter. It results that the minimum number of particles, required for the study of a specific structure is:

$$N \gtrsim \frac{M_{structure}}{M_{particle}}$$

- For every cosmological simulation, each N-body particle represents a collection of a very large number of particles in the real universe as we would like to simulate the evolution of structures at scales that cosmology is interested. Therefore an essential property of the N-body particles is that the interaction between them is collisionless.

4.1.1 Particle-Mesh Method (PM)

In the analysis that will follow (in section 5.) the halo catalogues are from the DEUS light-cone simulation. In these simulations the evolution of the dark matter particles is computed using a PARTICLE-MESH (PM) code with an Adaptive Mesh Refinement (AMR) grid.

The basic concept of this method is to convert the system of particles into a grid of density values and then solve the potential for this density grid and apply forces to each dark matter particle based on which cell it is in and where this cell lies on the grid. Since the Poisson equation is a simple algebraic equation in Fourier space $\hat{\Phi} = -4\pi G\hat{\rho}/k^2$, where \vec{k} is the comoving wavenumber and the hats denote Fourier transform, the solution of this equation is trivial with the Fast Fourier Transform (FFT) method. This method is faster than computing all the forces extracted to each particle from the others in the simulating cube, basically because the grid cells are less than the number of particles and also the transformation to a Fourier space (in which the number of operations required for computing the Fourier transform scales with $N\log N$) is so fast that makes the computation of the forces in the Fourier space more efficient.

The advantages of this method are:

- It is the fastest compared to any other method used in cosmological simulations, but at the expense of small-scale accuracy.
- It allows the use of very large numbers of particles, which is essential in simulation like these.

The use of the (PM) method has the disadvantage of providing low resolution results. However, the addition of an Adaptive Mesh Refinement (AMR) grid mostly overcomes the limits in resolution inherent to the (PM) methods

4.2 Light-cone Cosmological Simulations

As it has already been mentioned above, modern galaxy redshift surveys (e.g. Sloan Digital Sky Survey, 2-degree Field Galaxy Redshift Survey, Two Micron All-sky Survey) have contributed vastly in the development of modern cosmology, by providing details of the LLS and constraining the cosmological model. To this end it was instrumental the construction of 'mock' galaxy catalogues, which also mimic the selection effects in real galaxy surveys. Mock catalogues can be obtained also from light-cone simulations, which are constructed during the run of N-body simulations.

A light-cone is the path that a light ray, (produced during a single event and then traveled in all directions) will follow in spacetime (it is also called the region of causal connection, as it describes all the possible locations light can ever reach from a single source, due to the constant value of the light speed). Events occurring outside the light-cone of an observer can not influence him because they do not have sufficient time to reach him.

Lightcone halo mock catalogues are constructed based on the above assumption, by using dark matter particles in a cubic N-body simulation. For an observer fixed in a position, a dark matter halo is placed in the lightcone to the epoch at which it first crossed the observer's past light cone. Because only then (supposing the halo emits light) the light emitted from it's location has just enough time to reach the observer and thus incorporate the evolution of structure with the cosmic time.

The ever growing volume size of galaxy surveys, have created a demand for simulations with boxes of sufficient size to match the volumes of these surveys. On the other hand, the lack of computing power means that a compromise must be made between the size of the simulating cube and the resolution at which the simulation is carried out. Then, in order to generate sufficient cosmological volumes the latter usually consists of replicants of the original simulating cube. The number of replicants per axis (n_{rep}), that need to be stacked in order to produce a large cosmological volume is,

$$n_{rep} = \left\lceil \frac{r_{max}}{L_{box}} \right\rceil + 1$$

where r_{max} is the maximum comoving radial distance that the final mock catalogue should have. The Cartesian coordinate system, $(\hat{X}, \hat{Y}, \hat{Z})$, of the combination of replicants is translated so that the observer is located at the origin.

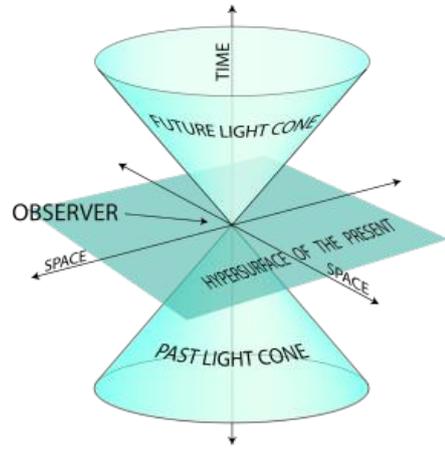


Figure 4: Light-cone in 2D plus a time dimension

The disadvantage of this method of producing a sufficient large cosmological cube, is that structures may appear repeatedly in the final volume. This method may cause contamination of the generated mock catalogues, which is then translated into falsely clustering statistics. Therefore this method of producing lightcone catalogues is not recommended.

The orientation of the observer is such, as to be looking down the $\hat{\mathbf{Z}}'$ axis. This axis defines the central axis of the conical volume of the lightcone and points to the center of the field of the lightcone. The angle θ'_r represents the angular extend of the field-of-view of the lightcone.

Therefore the orientation of the observer is translated into the dependence of the $\hat{\mathbf{Z}}'$ axis from the Cartesian coordinates $\hat{\mathbf{Z}}' = \hat{\mathbf{Z}}'(\hat{X}, \hat{Y}, \hat{Z})$ (for more information about lightcones see [17] and for the evolution of dark matter haloes [19],[20])

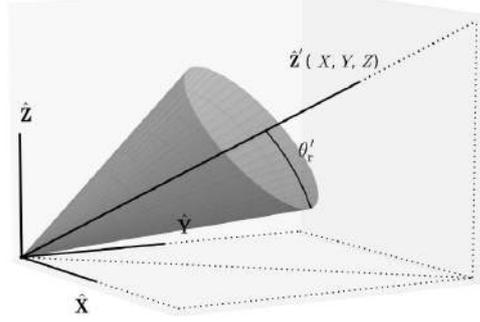


Figure 5: Haloes whose position vectors $\mathbf{r} = \mathbf{r}(\hat{X}, \hat{Y}, \hat{Z})$ is offset from the $\hat{\mathbf{Z}}'$ axis by an angle $\theta' > \theta_r$ are excluded from the lightcone

4.2.1 Friends-of-Friends (FoF) method for halo identification

Dark matter halos have been identified from the particles in the light-cone using the FRIENDS-OF-FRIENDS (FoF) method, which is the simplest method to identify clusters of objects in simulations.

The FoF algorithm is based on the notion of the 'friend', which means that two particles are friends if they are within a distance ϵ of each other. Moreover two particles are friend-of-friend if they are either direct friends or they can be reached through a series of intermediate friend relations. Imagining that each particle is surrounded by a sphere of radius $\epsilon = bd/2$, where \mathbf{d} represents the mean distance between particles and \mathbf{b} is a linking parameter (which has the value $b=0.2$ in the current DEUS simulation), then the algorithm computes the distances and all particles and those connected by a friend-of-friend relation belong to the same halo (or cluster, depending on the simulation).

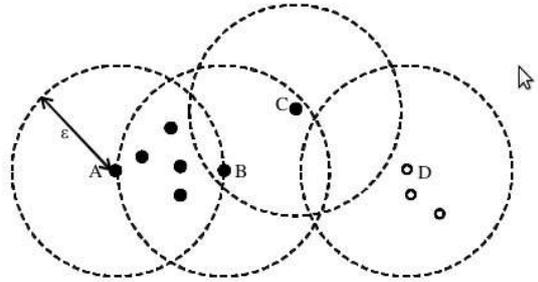


Figure 6: Two particles are considered to be friends if the distance between them is less than ϵ . In the figure A is friend with B but it is not friend with D. But since A can reach D through intermediate steps of the friend procedure then A and D are considered to be friend-of-friend

5 Simulation Halo Data Analysis

5.1 Introduction

The halo data are of three different DEUS light-cone cosmological simulations. The three cosmologies are a standard Λ CDM model calibrated to WMAP-5, a Ratra-Peebles quintessence model and SUGRA model both calibrated to WMAP-5 as well. The exact model parameter values (for a flat universe, $\Omega_\Lambda = 1 - \Omega_m$) are represented in the following table:

Table 1: SOME OF THE COSMOLOGICAL PARAMETERS FOR THE THREE COSMOLOGICAL MODELS: Λ CDM, RPCDM AND SUCDM

| Parameters | Λ CDM | RPCDM | SUCDM |
|----------------|---------------------|---------|------------------|
| Ω_m | 0.26 | 0.23 | 0.25 |
| $\Omega_b h^2$ | 0.02273 | 0.02273 | 0.2273 |
| α | 0 | 0.5 | 1 |
| $\lambda(eV)$ | $2.4 \cdot 10^{-3}$ | 4.9 | $2.1 \cdot 10^3$ |
| w_0 | -1 | -0.87 | -0.94 |

These are narrow light-cones covering an area of $\sim 160deg^2$ and a maximum redshift $z_{max} = 2.3$, which have been constructed during the run of N-body simulations with boxlength of 2592 Mpc/h and 1024^3 particles corresponding to a mass resolution of $10^{12} M_\odot$. From the particles in the light-cones, halos have been identified with at least 100 particles (which means that the minimum halo mass is $\sim 10^{14} M_\odot$) using FoF($b = 0.2$) (see [23]) for the parameters of the models presented in **Table 1**.

The halo catalogue files consist of:

- Halo-ID
- Halo-Mass [M_\odot/h]
- Halo's observed Redshift (which includes the peculiar velocity of the halo)
- Azimuthal-angle [degrees]
- Polar-angle [degrees]
- Peculiar velocity (along the line-of-sight) [km/s]
- Cosmic Distance [Mpc/h]

5.1.1 Presenting the cosmological models

The data-sets I used in the following analysis are generated with the Dark Energy Universe Simulation Series (**DEUSS**), the largest dark energy simulation conducted so far in terms of spatial dynamics. As mentioned above these sets correspond to three different realistic dark energy cosmological models, which hopefully probe the formation of the cosmic structure.

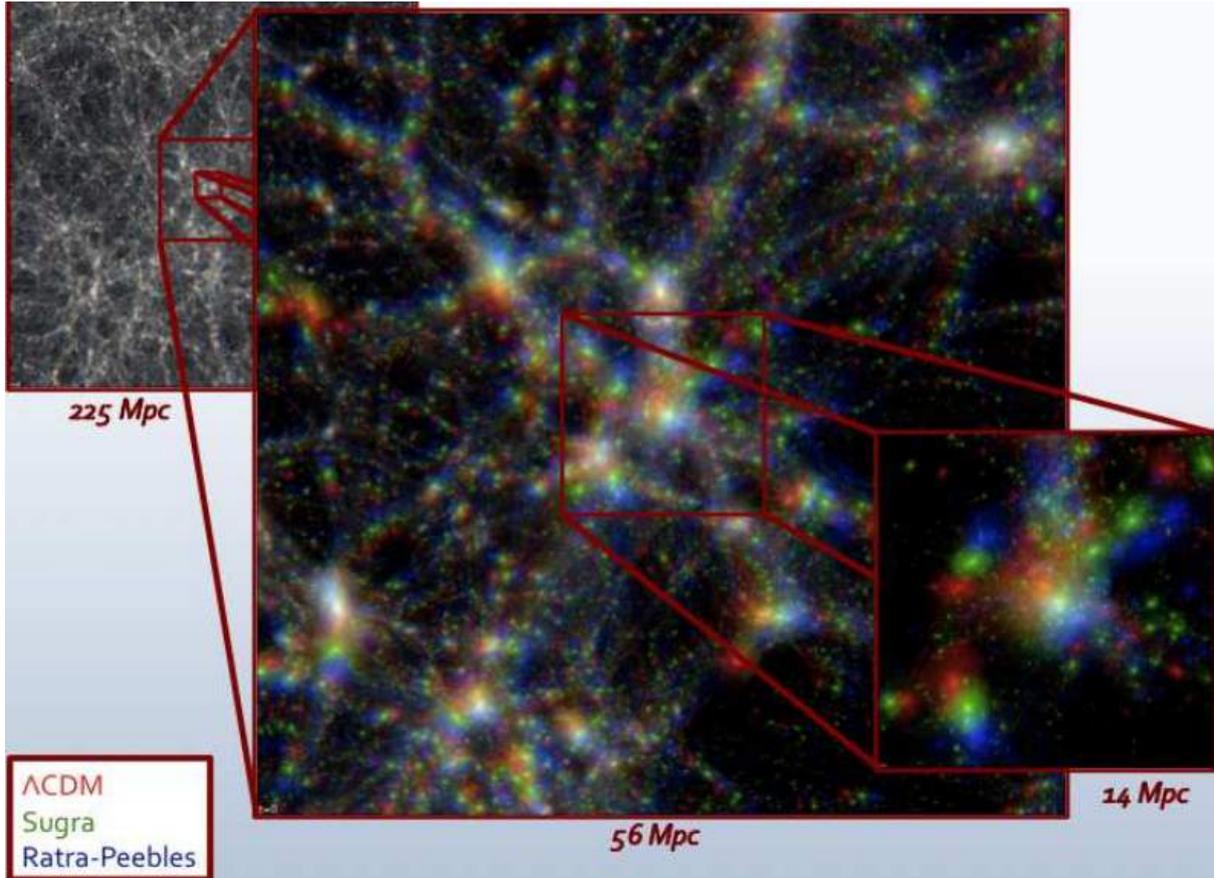


Figure 7 : The figure illustrates the projected density maps from the **DEUSS** simulation for the three cosmologies: Λ CDM (red), Sugra (green) and RPCDM (blue). The differences between the models are not clear on large scales but when zooming onto a halo (non-linear scales) the discrepancies between the models start to appear

Namely these cosmologies are the standard Λ CDM, the quintessence models of Ratra-Peebles (RPCDM) and Sugra (SUCDM) potentials, all calibrated on the latest SNIa and CMB (from WMAP-5) data. Quintessence is a hypothetical form of dark energy, introduced to describe the late acceleration of the universe. It is described by a slowly evolving, inhomogeneous and canonical scalar field ϕ , which can explain the acceleration of the universe by rolling down the potential $V(\phi)$, similar to the idea of inflation theory. Equation of state parameter, in the context of quintessence is given by,

$$w_0 = \frac{P_\phi}{\rho_\phi} = \frac{\dot{\phi}^2/2 - V(\phi)}{\dot{\phi}^2/2 + V(\phi)}$$

where P_ϕ and ρ_ϕ are the pressure and energy density of the quintessence, respectively (more information for quintessence models can be found [22] and for the SUGRA and Ratra-Peebles models in [23])

The quintessence field is assumed to be a neutral scalar field $\phi(\eta)$, where η is the conformal time, with a self-interaction potential $V(\phi)$ which couples to ordinary matter only through its gravitational influence. The evolution of the system is completely determined by specifying

the form of the quintessence potential, which are the RP inverse power law and the SUGRA model which is also an inverse power-law potential with supergravity corrections included:

$$V_{RP}(\phi) = \frac{\lambda^{4+\alpha}}{m_{Pl}^\alpha \phi^\alpha}$$

$$V_{SU}(\phi) = \frac{\lambda^{4+\alpha}}{m_{Pl}^\alpha \phi^\alpha} e^{4\pi\phi^2}$$

where α, λ are free parameters characterizing the slope and amplitude of the scalar self-interaction, respectively and m_{Pl} is the Planck mass. (These derived parameters are presented in **Table 1**).

5.1.2 Comparing the model redshift distributions

Here we make the first comparison of the different models. We investigate whether the halo redshift distribution, are distinct. The figure below represents the three different redshift distribution, which corresponds to each cosmology.

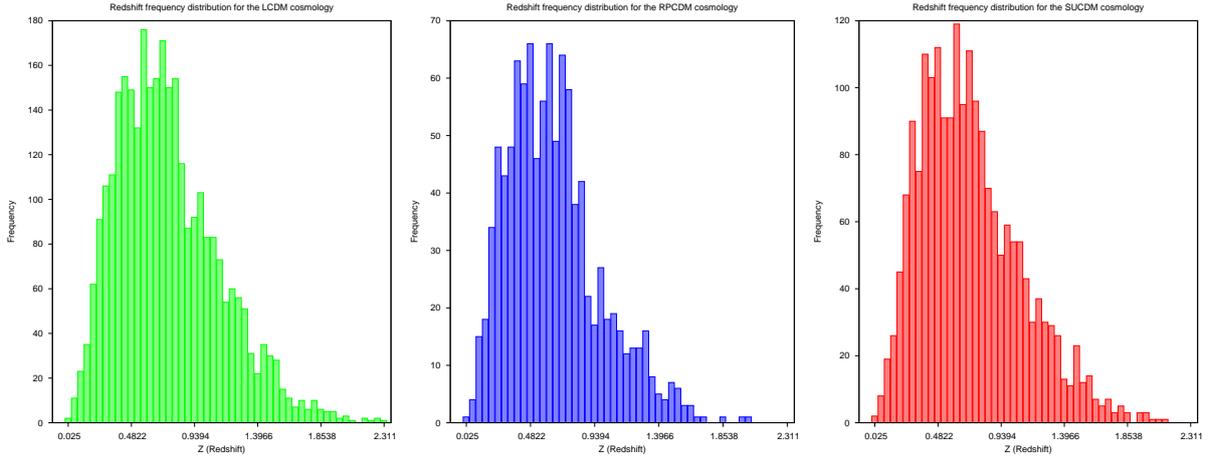


Figure 8: The graph shows the redshift frequency distributions for the Λ CDM-model (green), the RPCDM-model (blue) and the SUCDM-model (red) in the redshift interval $z = (0.024, 2.311)$

After deriving the redshift frequency distributions for the halo data for all the three cosmologies, the Kolmogorov-Smirnov (**K-S**) two-sample test is applied to them in order to compare them. This comparison aims to probe whether the $N(z)$ can be considered as being drawn from different parent distribution.

The K-S two-sample test can be applied to unbinned distributions that are functions of a single independent variable (i.e. redshift). In this case the data points can be converted to an unbiased estimator $S_N(x)$ of the cumulative distribution function of the probability distribution from which it was drawn. All cumulative distribution functions have a same property i.e. the smallest value that they can take is zero while the largest is one. The

difference between cumulative distribution functions, is the statistic used. These differences can be measured by the maximum value of the absolute difference between two cumulative distributions D , which is defined as:

$$D = \max_{-\infty < x < \infty} |S_{N_1}(x) - S_{N_2}(x)|$$

The K-S statistic enables the calculation of its distribution in the case of the null hypothesis (i.e. data sets are drawn from the same distribution). Then the significance level of the value D (i.e. the disproof that the distributions are the same) can be evaluated and is denoted as P_{prob} (see [21]).

Applying the K-S two sample test on the above redshift distributions,

Table 2: THE MAXIMUM VALUE OF THE ABSOLUTE DIFFERENCE BETWEEN TWO CUMULATIVE DISTRIBUTIONS D AND THE SIGNIFICANCE LEVEL OF IT.

| Parameters | Λ CDM-RPCDM | RPCDM-SUCDM | SUCDM- Λ CDM |
|------------|-----------------------------|----------------------------|----------------------------|
| D | 0.12400889 | 0.08986493 | 0.0417648554 |
| P_{prob} | $7.78404424 \cdot 10^{-11}$ | $2.99471285 \cdot 10^{-5}$ | $2.81944238 \cdot 10^{-2}$ |

From the resulting values of the K-S test statistics it is rigorously implied that there is a huge difference between the redshift distributions of the different dark energy models. We find that this difference is smaller when comparing the Λ CDM with the SUCDM model with respect to the other model pairs, but still not significant enough (of the order of $\sim 0.03\%$).

5.2 Computing the angular Two-point Correlation function

As a first test of the clustering we measure the two-point correlation function in angular space, which is derived from the halo catalogues using three different estimators. The choice of the estimators is based on the published paper (see [6]), in which the Landy & Szalay and Hamilton Estimators yield almost identical results and both of them significantly outperforms the rest of the estimators. For comparison the Natural Estimator is also included in the following procedure. The correlation functions in two-dimensions based on the three different estimators are given below for the Natural, Hamilton and Landy & Szalay respectively:

$$1 + w(\theta) = \frac{DD(\theta)}{RR(\theta)} \quad (28)$$

$$1 + w(\theta) = \frac{DD(\theta) \cdot RR(\theta)}{DR^2(\theta)} \quad (29)$$

$$1 + w(\theta) = \frac{DD(\theta) - 2DR(\theta) + RR(\theta)}{RR(\theta)} \quad (30)$$

The meaning of the symbols $DD(\theta)$, $RR(\theta)$ and $DR(\theta)$ will be further explained below.

In order to compute the correlation function, a random distribution of points must be generated. The random distribution must have the same spatial boundary conditions as the

real point distribution. Since the angular coordinates are between $5^\circ < \text{Azimuthial-angle} < 29^\circ$ and $-75^\circ < \text{Polar-angle} < -55^\circ$, the region that they cover is far from the equator, therefore is not a perfect square. This property can be visualized through the following figure:

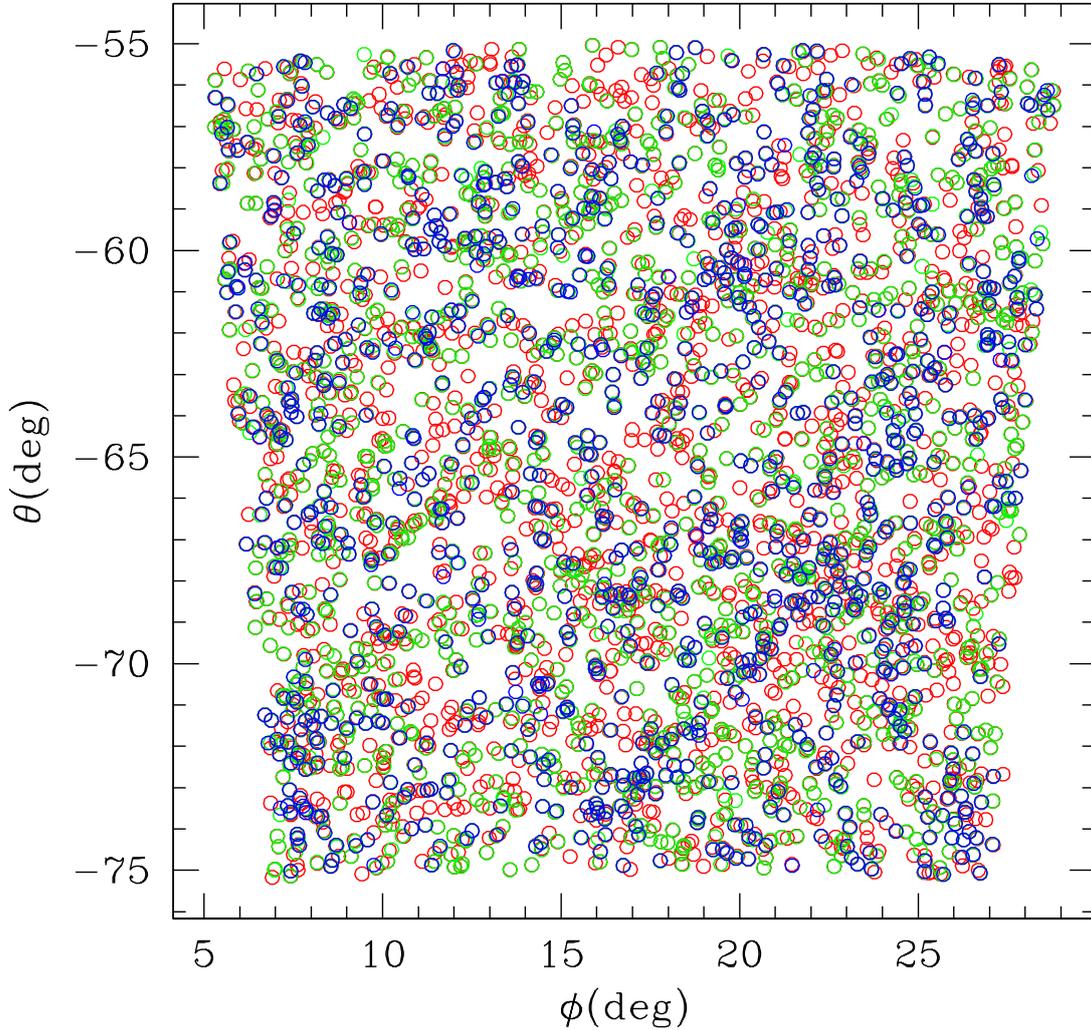


Figure 9: The graph shows the distribution of points in the Right-ascension Declination space for the three cosmologies (red- Λ CDM, blue-RPCDM and green-SUCDM).

The same property must be satisfied by the random point distribution, so during the selection of the random point, the curvature of the unit sphere must be taken into account. To encode this property a random distribution of points is generated in the Cartesian space $\{x,y,z\}$ and the value of each coordinate ranges in the interval $[-1,1]$. To derive the angular coordinates, a transformation to spherical coordinates is applied.

$$\begin{aligned}\phi &= \tan^{-1} \left(\frac{y}{x} \right) \\ \theta &= \sin^{-1} \left(\frac{z}{r} \right)\end{aligned}$$

and the points, which satisfies the conditions $5^\circ < \phi < 25^\circ$, $-75^\circ < \theta < -55^\circ$ and $r = \sqrt{(x^2 + y^2 + z^2)} \leq 1$ constitute the random distribution.

The next step is to count the data pairs within a certain range of separations as long as the random and the data-random pairs are within the same range. The mathematical interpretation of this process is being explained below.

Consider \mathbf{x} and \mathbf{y} as points in the angular space.

$$P_{DR}(\theta) = \sum_{\mathbf{x} \in D} \sum_{\mathbf{y} \in R} \Phi_\theta(\mathbf{x}, \mathbf{y}) \quad (31)$$

Here the summation runs over the angular coordinates of points in the data set D and points in the set R of randomly distributed points, respectively. The definition for Φ is that $\Phi_\theta(\mathbf{x}, \mathbf{y}) = [\theta \leq d(\mathbf{x}, \mathbf{y}) \leq \theta + \Delta\theta]$, where $d(\mathbf{x}, \mathbf{y})$ is the angular separation of two points of the respective set

$$d(\mathbf{x}, \mathbf{y}) = \arccos [\sin(\theta_1) \sin(\theta_2) + \cos(\theta_1) \cos(\theta_2) \cos(\phi_1 - \phi_2)] \quad (32)$$

and the condition in the brackets equals 1 when the condition holds and 0 otherwise.

Analogously, if both \mathbf{x} and \mathbf{y} are from the data set or the randomly distributed points (under the restriction that $\mathbf{x} \neq \mathbf{y}$) then P_{DD} and P_{RR} are defined, respectively. This method can be visualized as creating a histogram of angular separations (where the maximum and minimum value of the angular separation may vary), with a binwidth of $\Delta\theta$ and the value of $w(\theta)$ corresponding to the center of each bin. The value of $\Delta\theta$ depends on the number of θ -bins (N) and the maximum and minimum value of the angular separation:

$$\Delta\theta = \frac{\Theta_{max} - \Theta_{min}}{N}$$

For computing the two-point correlation function the normalized counts are being used, which are presented below:

$$\begin{aligned}DD(\theta) &= \frac{P_{DD}(\theta)}{N(N-1)} \\ RR(\theta) &= \frac{P_{RR}(\theta)}{N_R(N_R-1)} \\ DR(\theta) &= \frac{P_{DR}(\theta)}{N_R N}\end{aligned}$$

with N and N_R being the total number of data and random points in the survey volume, respectively.

The $w(\theta)$ uncertainty in each θ -bin is given by (see [26]):

$$\sigma_w = \frac{1 + w(\theta)}{\sqrt{DD(\theta)}}$$

After deriving the points of the two-point correlation function (in two-dimensions) $w(\theta)$, we apply the χ^2 -minimization method in order to fit the functional form of the correlation function in the data points. The functional form of the angular correlation function is (Peebles 1980):

$$w(\theta) = A \theta^{1-\gamma} \quad (33)$$

Therefore the χ^2 Minimization method (see [21]):

$$\chi^2 = \sum_N \frac{(w(\theta) - w_d(\theta))^2}{\sigma_w^2} \quad (34)$$

where the $w_d(\theta)$ are the data points of the measured correlation function and $w(\theta)$ is the two-parameter power law function, which is fitted to the data points. The summation runs over the number of bins N or a portion of them (in this case the fitting is done for points with $\theta \leq 1^\circ$) and in the range where a rough power-law is observed (usually we exclude the first few bins) and computes the χ^2 . The values of A and γ corresponding to the minimum value of χ^2 , gives the best fit functional form of the angular two-point correlation function.

The following figures shows the 2-D Angular correlation functions generated with the use of three Estimators in the different Cosmological models. The contour plots present the 1,2 and 3 σ range in the fitted (A, γ) parameter space:

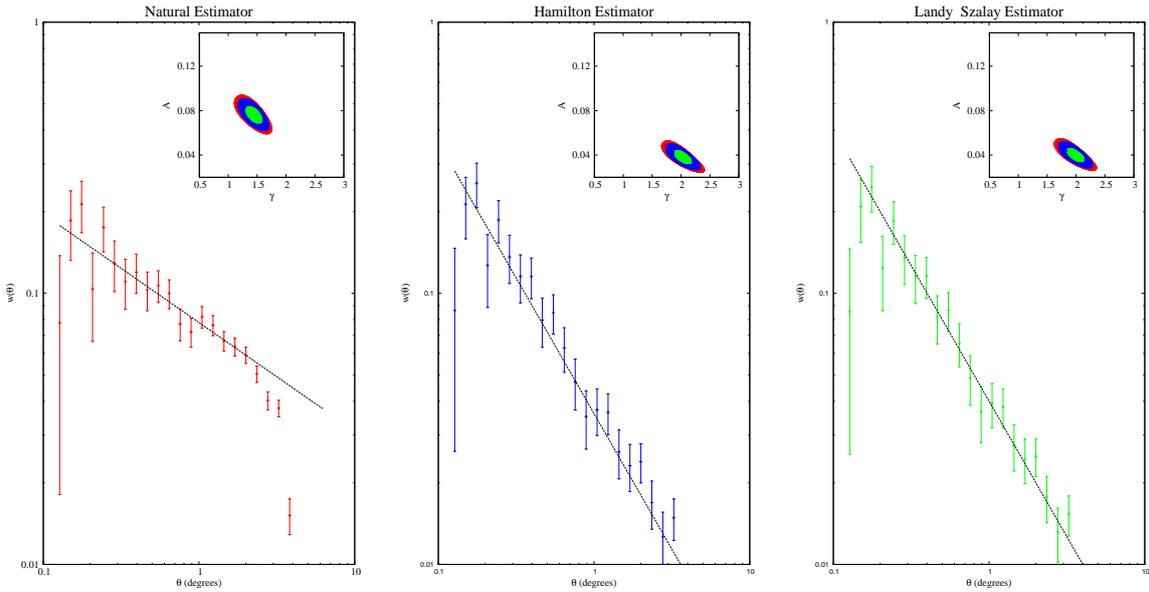


Figure 10: The graph shows the angular two-point correlation function for three different Estimators in the Λ CDM-cosmology (red-Natural, blue-Hamilton and green-L&Szalay).

Table 3: THE PARAMETERS OF THE CORRELATION FUNCTION FOR EACH ESTIMATOR IN Λ CDM-COSMOLOGY

| Parameters | Natural Estimator | Hamilton Estimator | Landy & Szalay Estimator |
|------------|---------------------------|---------------------------|---------------------------|
| A | $0.078^{+0.003}_{-0.003}$ | $0.039^{+0.002}_{-0.002}$ | $0.040^{+0.002}_{-0.002}$ |
| γ | $1.4^{+0.08}_{-0.05}$ | $2.0^{+0.07}_{-0.06}$ | $2.0^{+0.07}_{-0.08}$ |

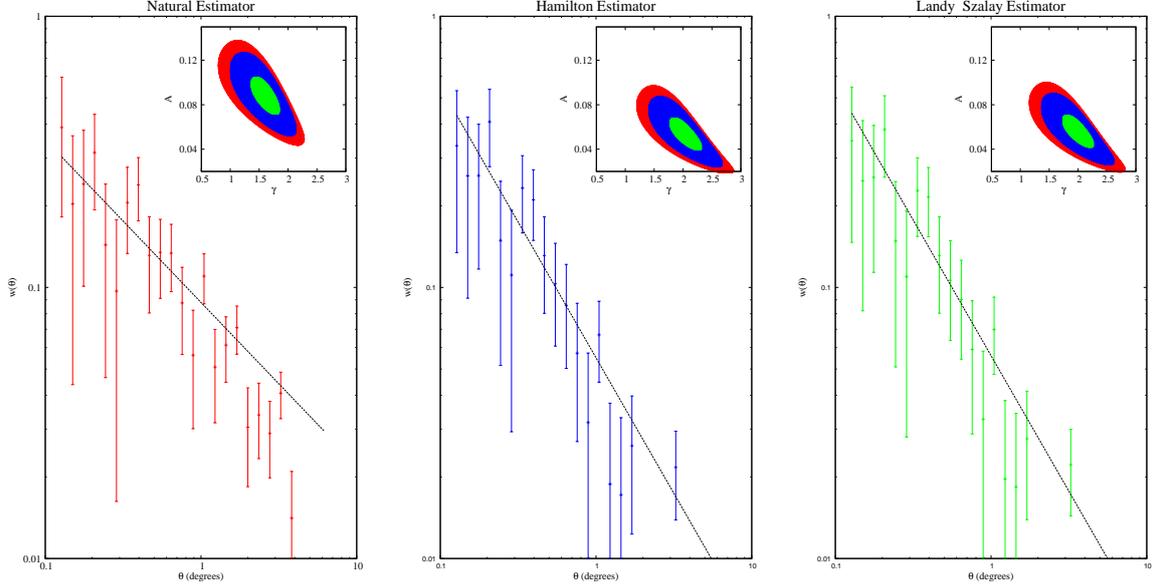


Figure 11: The graph shows the angular two-point correlation function for three different Estimators in the RCDM-cosmology (red-Natural, blue-Hamilton and green-L&Szalay).

Table 4: THE PARAMETERS OF THE CORRELATION FUNCTION FOR EACH ESTIMATOR IN RPCDM-COSMOLOGY

| Parameters | Natural Estimator | Hamilton Estimator | Landy & Szalay Estimator |
|------------|---------------------------|---------------------------|---------------------------|
| A | $0.088^{+0.010}_{-0.009}$ | $0.055^{+0.007}_{-0.007}$ | $0.056^{+0.007}_{-0.008}$ |
| γ | $1.6^{+0.16}_{-0.15}$ | $2.0^{+0.14}_{-0.14}$ | $2.0^{+0.12}_{-0.16}$ |

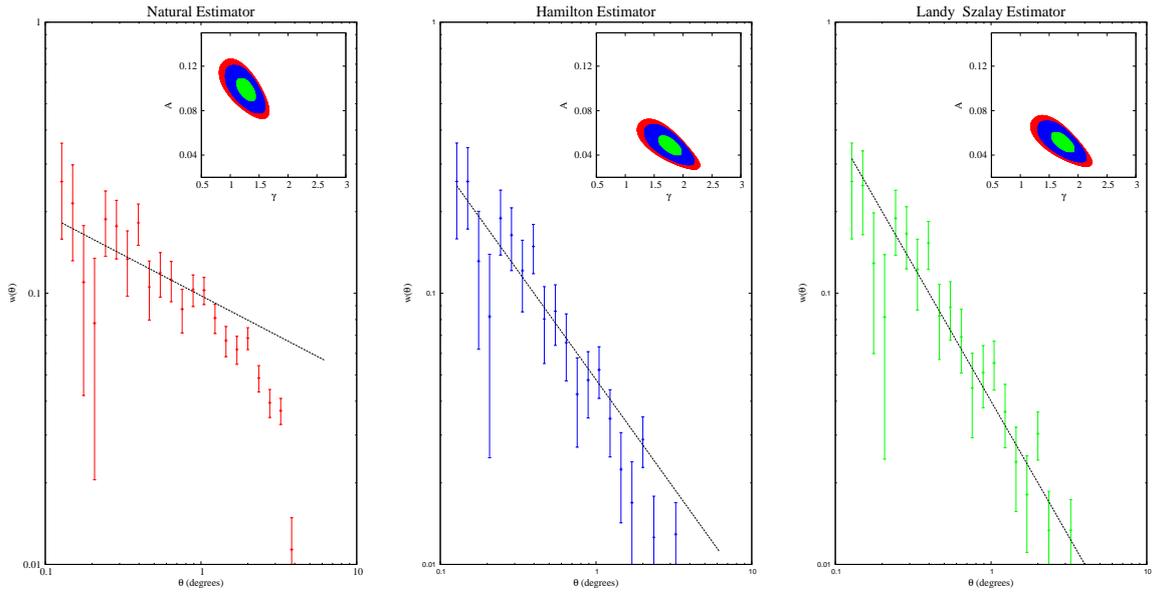


Figure 12: The graph shows the angular two-point correlation function for three different Estimators in the SUCDM-cosmology (red-Natural, blue-Hamilton and green-L&Szalay).

Table 5: THE PARAMETERS OF THE CORRELATION FUNCTION FOR EACH ESTIMATOR IN SUCDM-COSMOLOGY

| Parameters | Natural Estimator | Hamilton Estimator | Landy & Szalay Estimator |
|------------|---------------------------|---------------------------|---------------------------|
| A | $0.098^{+0.006}_{-0.006}$ | $0.048^{+0.004}_{-0.004}$ | $0.053^{+0.004}_{-0.005}$ |
| γ | $1.3^{+0.08}_{-0.12}$ | $1.8^{+0.08}_{-0.13}$ | $1.7^{+0.11}_{-0.10}$ |

We see that in all three models the Natural Estimator produce significantly different results with respect to the more robust Hamilton and Landy & Szalay estimators. Furthermore we see that in the Λ CDM and SUCDM models the $w(\theta)$'s are consistent to each other both in amplitude and slope. The RPCDM's $w(\theta)$ however has a significant difference from both the other two models.

5.3 Computing the spatial Two-point Correlation function

After generating the 2-D Angular two-point correlation function, a similar procedure is followed to derive the correlation function, now in Three Dimensions. The point process, described by the formula (4), is the same except that \mathbf{x} and \mathbf{y} are points in the 3-Dimensional Cartesian space. The distance between two points in the 3-D cartesian space is given by:

$$d(\mathbf{x}, \mathbf{y}) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2} \quad (35)$$

Another important feature must be taken into account, before computing the 3-D correlation function. When generating the random distribution in 2-D, the random points ought to have the same angular survey boundaries. In addition, the randomly distributed points in three dimensions, besides keeping the same angular survey boundaries, there should also have the same redshift distribution (or similarly the same Cosmic distance distribution in z-space). This is achieved by assigning randomly a halo distance to each random pointn in such a way so that no halo distance is selected more than once. In this way, both the data and random points have the same redshift distribution.

Furthermore, in the case of the 3-D correlation function, its functional form is:

$$\xi(r) = \left(\frac{r_0}{r}\right)^\gamma \quad (36)$$

which has to be incorporated in the formula (7), and apply the χ^2 -minimization method to derive the parameters r_0 and γ of the correlation function in three dimensions.

Analogously the $\xi(r)$ uncertainty in each r-bin is given by (Peebles 1980):

$$\sigma_\xi = \frac{1 + \xi(r)}{\sqrt{DD(r)}}$$

and the χ^2 correspondence method has the form:

$$\chi^2 = \sum_N \frac{(\xi(r) - \xi_d(r))^2}{\sigma_\xi^2} \quad (37)$$

Again in this case the fitting for the functional form of the correlation function $\xi(r)$ is done for points with $r \leq 90(Mpc/h)$, because for $r > 90Mpc h^{-1}$ a small peak is observed which corresponds to the baryonic acoustic oscillation (**BAO**) scale, which breaks the power-law nature of $\xi(r)$.

The following figures shows the 3-D correlation functions generated with the use of three Estimators for all three Cosmological models. The contour plots present the 1,2 and 3σ range in the fitted (r_0, γ) parameter space:

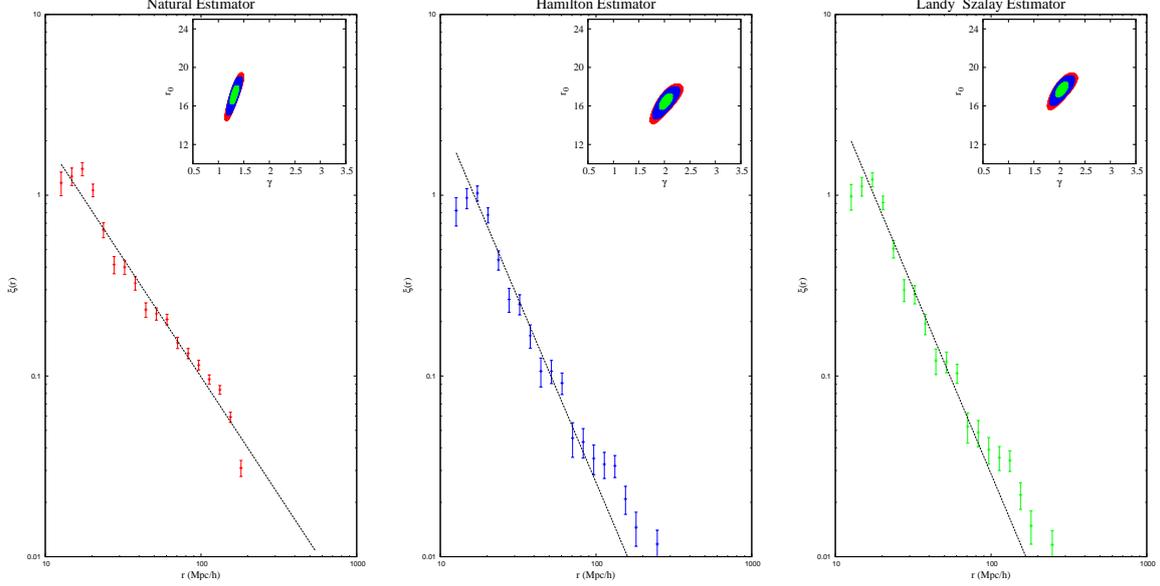


Figure 13: The graph shows the 3-D two-point correlation function for three different Estimators in the Λ CDM-cosmology (red-Natural, blue-Hamilton and green-L&Szalay). The **BAO** signal is evident at $r \sim 120h^{-1}Mpc$

Table 6: THE PARAMETERS OF THE CORRELATION FUNCTION FOR EACH ESTIMATOR IN Λ CDM-COSMOLOGY

| Parameters | Natural Estimator | Hamilton Estimator | Landy & Szalay Estimator |
|------------|--------------------------|--------------------------|--------------------------|
| r_0 | $17.1^{+0.33}_{-0.32}$ | $16.5^{+0.33}_{-0.38}$ | $17.7^{+0.38}_{-0.31}$ |
| γ | $1.31^{+0.022}_{-0.021}$ | $2.03^{+0.059}_{-0.050}$ | $2.05^{+0.049}_{-0.055}$ |

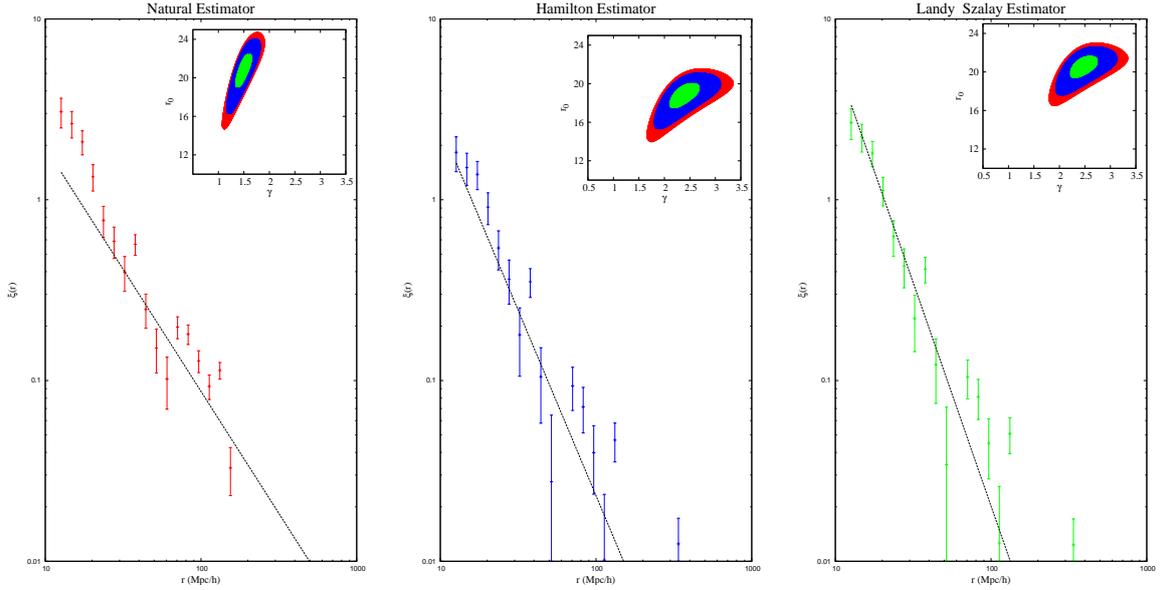


Figure 14: The graph shows the 3-D two-point correlation function for three different Estimators in the RPCDM-cosmology (red-Natural, blue-Hamilton and green-L&Szalay). The $\xi(r)$ is quite noisy and the **BAO** signal is not that clear

Table 7: THE PARAMETERS OF THE CORRELATION FUNCTION FOR EACH ESTIMATOR IN RPCDM-COSMOLOGY

| Parameters | Natural Estimator | Hamilton Estimator | Landy & Szalay Estimator |
|------------|--------------------------|----------------------------|----------------------------|
| r_0 | $20.9^{+0.78}_{-0.79}$ | $18.9^{+0.74}_{-0.78}$ | $20.6^{+0.77}_{-0.72}$ |
| γ | $1.50^{+0.067}_{-0.064}$ | $2.39^{+0.1903}_{-0.1653}$ | $2.47^{+0.1943}_{-0.1649}$ |

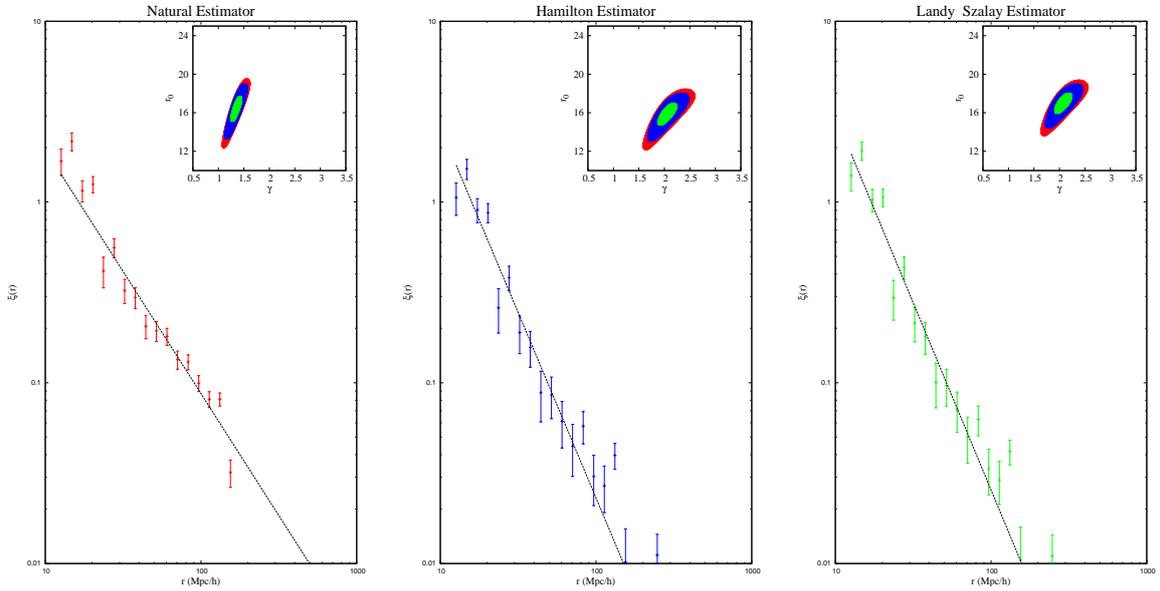


Figure 15: The graph shows the 3-D two-point correlation function for three different Estimators in the SUCDM-cosmology (red-Natural, blue-Hamilton and green-L&Szalay).

Table 8: THE PARAMETERS OF THE CORRELATION FUNCTION FOR EACH ESTIMATOR IN SUCDM-COSMOLOGY

| Parameters | Natural Estimator | Hamilton Estimator | Landy & Szalay Estimator |
|------------|--------------------------|--------------------------|--------------------------|
| r_0 | $16.4^{+0.48}_{-0.47}$ | $15.9^{+0.50}_{-0.55}$ | $17.0^{+0.51}_{-0.51}$ |
| γ | $1.35^{+0.034}_{-0.033}$ | $2.05^{+0.092}_{-0.075}$ | $2.07^{+0.083}_{-0.077}$ |

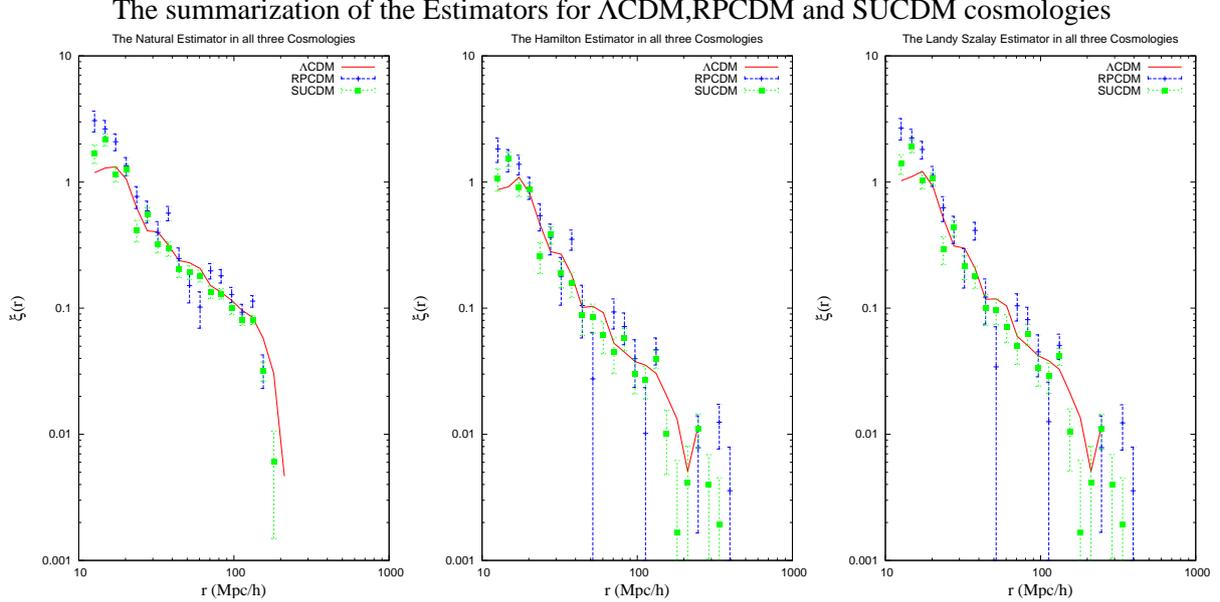


Figure 16: The graph shows the 3-D two-point correlation function for three different Estimators in the SUCDM-cosmology (red-Natural, blue-Hamilton and green-L&Szalay).

As general result we see that the Λ CDM, SUCDM models provide very similar $\xi(r)$, while the RPCDM significantly differ in the angular and spatial correlation function.

5.4 Evaluating the index of the growth factor ($n + 2$)

In the introductory chapters the growth factor of density perturbations was computed in the linear regime. Later on, with the use of the growth factor and the evolution of bias with redshift, the formula for the dependence of the correlation length with the redshift was introduced. Since this formula is only valid in the linear phases of structure formation, a modification of the formula for the correlation length was introduced, based on the expectation of the quasilinear regime of the growth factor (A. Peacock 2002).

The procedure, in order to evaluate the correlation length r_0 for different redshifts, is the following: The initial data set of the dark matter halos is separated into three new sets of data according to their redshift values as follows:

- $N_1 = 1023$, Data points in the redshift interval $z \in [0.024, 0.548]$
- $N_2 = 1023$, Data points in the redshift interval $z \in [0.548, 0.845]$
- $N_3 = 1025$, Data points in the redshift interval $z \in [0.845, 2.310]$

The limits of the redshifts are chosen by requiring that in each redshift bin to have the same number of haloes, in order to have similar uncertainties.

Now for each of these sets of data, the corresponding correlation function $\xi(r) = (r_0/r)^\gamma$ is evaluated (using only the Landy & Szalay Estimator, in order to compute the correlation length in each of the three redshift intervals.)

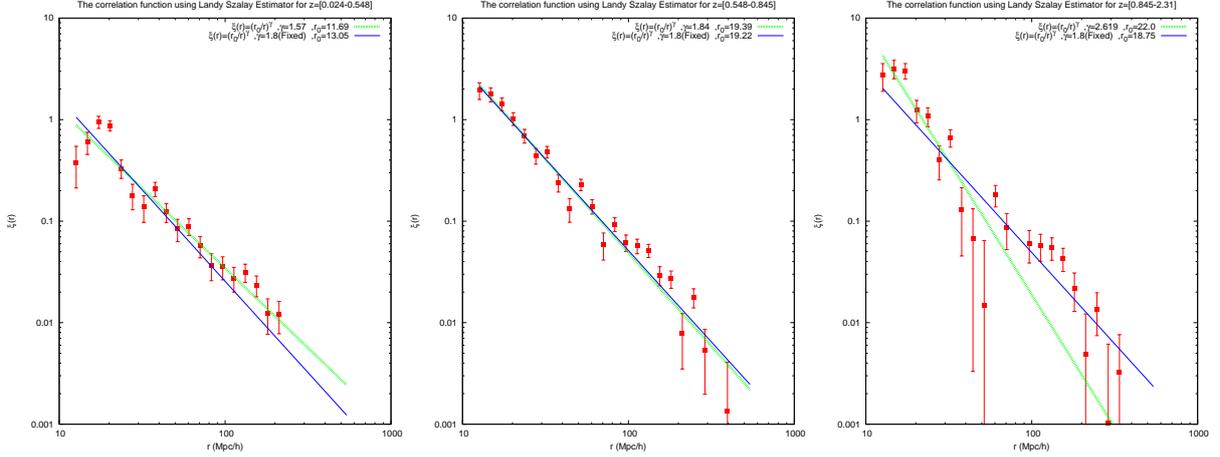


Figure 17: The graph shows the spatial correlation function computed in the three different redshift intervals using only the Landy & Szalay Estimator. The blue line corresponds to a fixed value for the index $\gamma = 1.8$, while in the green line the γ parameter was left free.

Table 9: THE PARAMETERS OF THE CORRELATION FUNCTION'S r_0, γ FOR EACH REDSHIFT, AS WELL AS THE CORRELATION LENGTH r_0 FOR $\gamma = 1.8$

| Parameters | $z_1 = 0.286$ | $z_2 = 0.696$ | $z_3 = 1.578$ |
|----------------------------|---------------|---------------|---------------|
| $r_0(Mpc/h)$ | 11.69 | 19.39 | 22.0 |
| γ | 1.57 | 1.84 | 2.619 |
| $r_0(Mpc/h)[\gamma = 1.8]$ | 13.05 | 19.22 | 18.75 |

However in the previous analysis we have at each different redshift a different fraction of the halo masses and therefore a mixing of the resulting correlation function's (which as it has already been discussed, depends on the mass of the halo). So the resulting correlation length can not be used for the evaluation of the slope of the growing mode of perturbations (n index), because there is a strong dependence in the halo masses of the resulting $\xi(r)$. In the last redshift interval the simulated universe has almost half of it's current age. Some of the halos are not yet virialized and the corresponding halo masses are smaller compared to the halo masses of the first redshift interval, where they had sufficient time to virialize.

This property can be visualized in the next figure, where the corresponding mass histograms reveals that in lower redshifts there is a smaller abundance in low halo masses compared to higher redshift where low halo masses dominate the histogram.

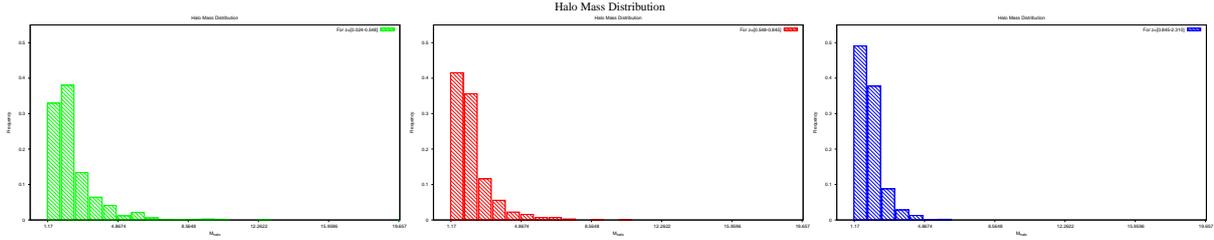


Figure 18: The distributions of the halo mass in each redshift interval

To eliminate the dependence on the variations of the halo masses in the different redshift intervals, we have chosen to reproduce the same mix of halo masses in each redshift bin. This is performed by the following routine:

- As the first step the, the number of halos in each M-bin and in every redshift interval is computed and the value which corresponds to the minimum number of halos in each M-bin for the three redshift intervals is determined
- The same number of halos (i.e. corresponding to the minimum value that the same M-bin will have in every redshift interval) is collected from every M-bin, in a random way, so that the resulting histograms from this process will be identical. The resulting halo data sets are independent of the halo mass
- This process is repeated five times and for every realization the spatial correlation function is generated.
- The value of the correlation function in every separation-bin is the average value of the five realization, which is then used to derive the corresponding correlation length for every redshift interval (which is independent of the mass distribution the halo masses)

This process can be visualized in the following figure, where the different colours denote the five different realizations of the random process.

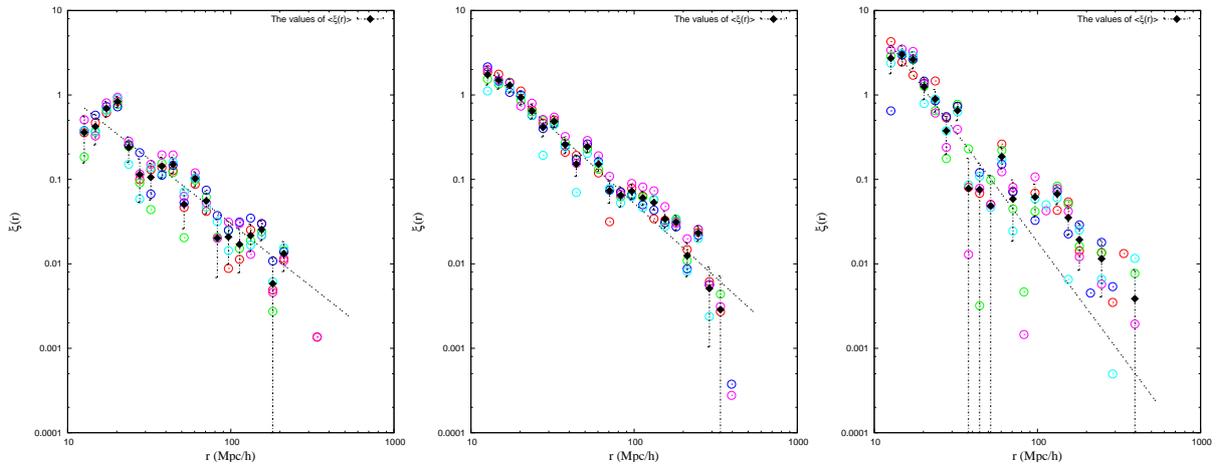


Figure 19: Black points denote the average correlation function $\langle \xi(r) \rangle$, while the other colours the different realizations of the random process

The table below summarizes the computed parameters for the average correlation function as well as the average correlation length and slope of the five realization of the random process.

Table 10: THE VALUES OF THE PARAMETERS

| Parameters | $z_1 = 0.286$ | $z_2 = 0.696$ | $z_3 = 1.578$ |
|--|-------------------|-------------------|-------------------|
| $\langle r_0 \rangle (Mpc/h)$ | 9.88 ± 0.93 | 18.54 ± 1.078 | 20.86 ± 1.256 |
| $\langle r_0 \rangle (Mpc/h) [\gamma = 1.8]$ | 11.64 ± 0.708 | 18.79 ± 0.992 | 17.86 ± 1.098 |
| $\langle \gamma \rangle$ | 1.52 ± 0.113 | 1.75 ± 0.089 | 2.59 ± 0.256 |
| For the mean value of $\langle \xi(r) \rangle$ | | | |
| $r_0 (Mpc/h)$ | 10.0 | 18.7 | 21.4 |
| $\langle r_0 \rangle (Mpc/h) [\gamma = 1.8]$ | 11.71 | 18.88 | 18.3 |
| γ | 1.52 | 1.76 | 2.61 |
| $\langle M_h \rangle [10^{-14} M_\odot h]$ | 1.2579302 | 1.2271066 | 1.2154511 |

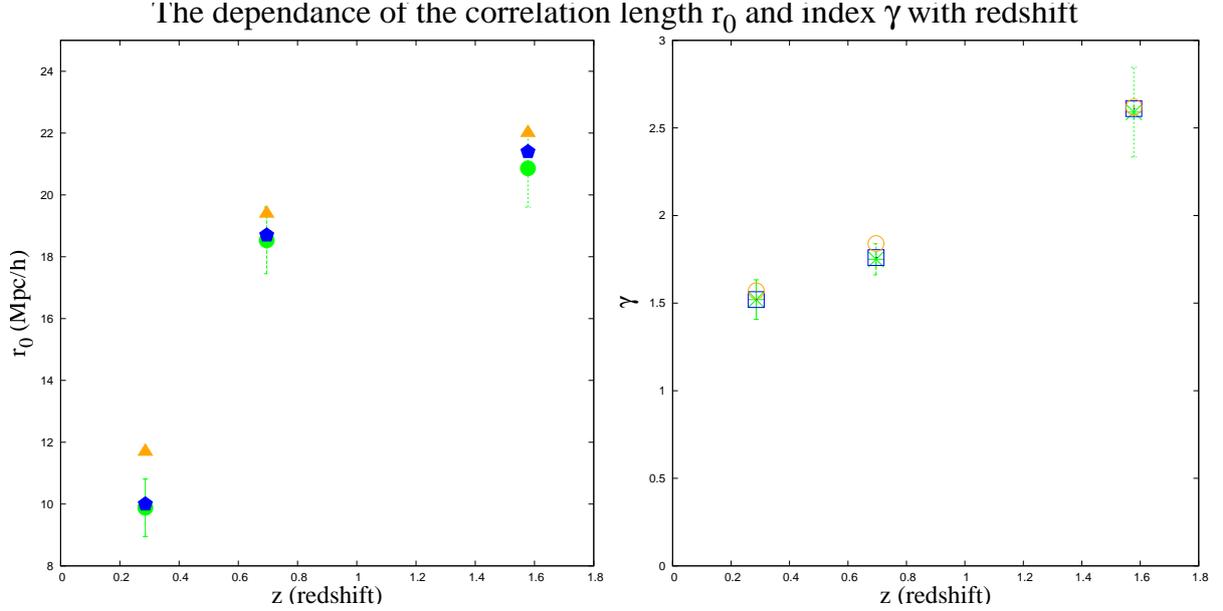


Figure 20: The orange points represents the original values of the correlation length r_0 and index γ in which the dependence on the different halo mass mix is still present. The green points are the mean values of the above parameters derived from the five realizations discussed in the text. Finally the blue points are the parameter values of the $\langle \xi(r) \rangle$.

As it has already been mentioned the evolution of the correlation length can not be interpreted by eq. (20), since the evolution of the density perturbation (22) is not valid in the mildly non-linear regime. That is why a modification of the above formula, by introducing a new index n , was necessary (A. Peacock 2002).

Since the value of the correlation length is not known in the present time $r_0(0)$, eq. (20) is normalized with respect to the value of the correlation length of redshift z_1 . This can be done

by dividing the value of the correlation length of z_2 with the value of it redshift z_1 . Using eq. (22) we can divide the value of r_0 at some redshift z_i , with respect to it's value at a redshift z_1 .

$$r_0(z_i) = \left[\left(\frac{D(z_i)}{D(z_1)} \right)^{2+n} \left(\frac{b(z_i)}{b(z_1)} \right)^2 \right]^{1/\gamma} r_0(z_1) \quad (38)$$

It should be noted that the index γ must be the same in all redshifts in order to use the simplified formula (38), thus the values of correlation lengths which are going to be used are the ones corresponding to a fixed slope $\gamma = 1.8$.

The last step of the process, for computing the index (n), is to apply the χ^2 -minimization method as follows:

$$\chi^2 = \sum_{i=1}^3 \frac{(r_{0,i} - r_0(z_i))^2}{\sigma_{r_{0,i}}^2}$$

The resulting value of n, which corresponds to the minimum of χ^2 , is $n = 0.9$ and it's corresponding evolution can be seen in the figure below as the red curve.

The whole process, that have been followed in this section, can be visualized in the next figure:

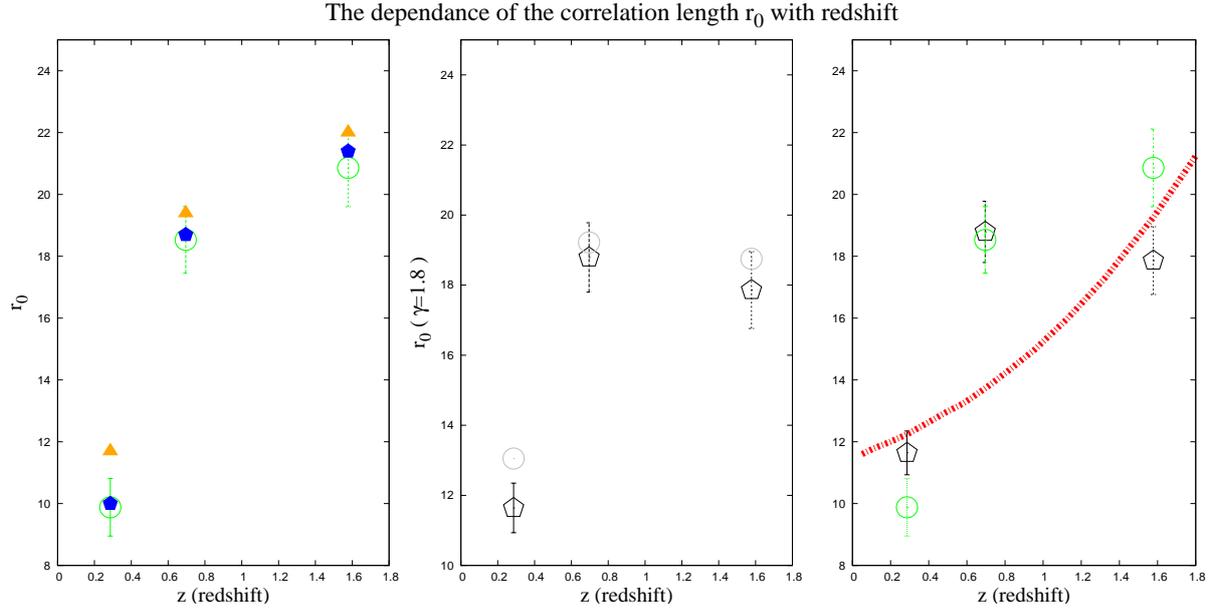


Figure 9: (The first figure is the same as in the previous and has been added for surveillant reasons. The second figure shows the dependence of the correlation length having the value of γ index fixed ($\gamma=1.8$). The black points corresponds the mean correlation legth computed from the five realization while the grey points are computed for the mean $\langle \xi(r) \rangle$. Finally in the last figure the green and black points are the mean values of the correlation length having the value of the γ index not-fixed and fixed respectively. The red line is the one which arise from equation (36) after substituting the value $n = 0.9$ in the modified index.)

As we see the red line in the previous figure, which represents the theoretical evolution of the correlation length with redshift, is an increasing function. However the evolution of the estimated halo correlation length does not satisfy well the theoretical relation. This may be attributed to an inadequacy of the functional form of the bias factor, $b(z)$, to trace the large redshifts or to the large redshift range covered by the last bin (implying an artificial mixing of different evolutionary phases in the development of the haloes). We plan to investigate this issue by taking smaller redshift intervals at lower redshift and repeating the process.

6 Conclusions

Several conclusions have been derived during the procedure of the analysis.

- 1) We have a clear difference between the correlation function estimated using the Landy & Szalay and or Hamilton (which exhibit almost identical results) and the natural Estimators. Previous research has shown that the first two estimators outperform the rest of the existing estimators (see [6]), among which is also the natural estimator. This result has also been confirmed in our analysis.
- 2) We find a relative good agreement between the halo redshift distribution, as well as the correlation length among the Λ CDM and *SUCDM* models. On the other hand, the *RPCDM* model have very different results from the other two models regarding the halo redshift distribution and the correlation length and slope (e.g computed using the Landy & Szalay estimator), which are clearly visible as follows: $r_0 = (17.7, 20.6, 17.0)$ and $\gamma = (2.05, 2.47, 2.07)$ for the *ACDM*, *RPCDM* and *SUCDM*, respectively.
- 3) From the analysis of the evolution of the correlation function for the Λ CDM model in the different redshift intervals we find that indeed the correlation length increases with increasing redshift. This analysis has taken into account the different halo mass distributions that arise in every different redshift interval.
- 4) The investigation of the correlation length with redshift, in the framework of Λ CDM cosmology, has yield for the mildly non-linear regime the value of $n = 0.9$ for the modified slope of the growth of density perturbations. As seen from (A. Peacock 2002) the value of this index should be around $n \sim 1.3$. This difference can be attributed to the lack of sufficient redshift intervals to trace down the evolution of the correlation length in our analysis. We plan to investigate the issue further in the future.

7 REFERENCES

- [1] M. Plionis et al. (Eds.), *A Pan-Chromatic View of Clusters of Galaxies and the Large-Scale Structure*, Lect. Notes Phys. 740 (Springer, Dordrecht 2008), DOI 10.1007/978-1-4020-6941-3
- [2] Richard Lieu, T.W.B. Kibble arxiv:astro-ph/2009
- [3] F. Bernardeau, S. Colombi, E. Gaztanaga, R. Scoccimarro, Phys. Rep. (2002)
- [4] Peebles, P.J.E, *The Large-Scale Structure of the Universe*, Princeton Univ. Press, New York (2008)
- [5] N. Espino-Briones, M. Plionis, C. Ragone-Figueroa, ApJ, 666: L5-L8 (2007)
- [6] M. Kerscher, I. Szapudi, A.S. Szalay, ApJ, 535:L13-L16 (2000)
- [7] M. Vargas-Magana , J. E. Bautista , J.-Ch. Hamilton , N.G. Busca , E. Aubourg, A. Labatie, J.-M. Le Goff , S. Escoffier , M. Manera , C. K. McBride , and D. P. Schneider Ch. N. A. Willmer, A&A (2013)
- [8] Alison L. Coil, *Large Scale Structure of the Universe*, (2012)
- [9] S. Basilakos, M. Plionis, A. Pouri, Phys. Rev. D 83, 123525 (2011)
- [10] S. Basilakos, J. B. Dent, S. Dutta, L. Perivolaropoulos, M. Plionis, Phys. Rev. D85, 123501 (2012)
- [11] Steven Weinberg, *Cosmology*, Oxford Univ. Press (2008)
- [12] Bardeen, J. M., Bond, J. R., Kaiser, N., & Szalay, A. S. 1986, ApJ, 304, 15
- [13] H.M.P. Couchman *Cosmological Simulations Using Particle-Mesh Methods*
- [14] A. Klypin arxiv:astro-ph/0005502 (2002)
- [15] J.S. Bagla, T. Padmanabhan arxiv:astro-ph/0411730v1 (2004)
- [16] J.S. Bagla arxiv:astro-ph/0411043v1 (2004)
- [17] Alexander I. Merson1 , Carlton M. Baugh1, John C. Helly , Violeta Gonzalez-Perez , Shaun Cole , Richard Bielby , Peder Norberg , Carlos S. Frenk , Andrew J. Benson , Richard G. Bower , Cedric G. Lacey , Claudia del P. Lagos arxiv:astro-ph1206.4049v2 (2013)
- [18] T. Hamana, S. Colombi, Y. Suto arxiv:astro-ph0010287v2 (2000)
- [19] T. Hamana, N. Yoshida, Y. Suto, A. E. Evrard arxiv:astro-ph/0110061v3 (2001)
- [20] P. Heinamaki, I. Suhhonenko, E. Saar, M. Einasto, J. Einasto, H. Virtanen A&A (2008)
- [21] W.H. Press, S.A. Teukolsky, W.T. Vetterling, B.P. Flannery, *Numerical Recipes in Fortran 77*, Cambridge Univ. Press (1986)
- [22] Adrian Vollmer, *Quintessence*, (2009)
- [23] J.-M. Alimi, A. Fuzfa, V. Boucher, Y. Rasera, J. Courtin P.-S. Corasaniti MNRAS **401**, 775-790 (2010)
- [24] *The Seventh Data Release of the Sloan Digital Sky Survey*, Abazajian et al. 2009, ApJS, 182, 543
- [25] Matthew Colless, arXiv:astro-ph/9804079v1 (1998)
- [26] THE TWO MICRON ALL SKY SURVEY, 2006, AJ, 131, 1163
- [26] A. Elyiv, M. Plionis, S. Basilakos A&A 537, A131 (2012)